CONVERGENCE OF BS∆ES DRIVEN BY RANDOM WALKS TO BSDES: THE CASE OF (IN)FINITE ACTIVITY JUMPS WITH GENERAL DRIVER

DILIP MADAN, MARTIJN PISTORIUS, AND MITJA STADJE

Abstract. In this paper we present a weak approximation scheme for BSDEs driven by a Wiener process and an (in)finite activity Poisson random measure with drivers that are general Lipschitz functionals of the solution of the BSDE. The approximating backward stochastic difference equations (BS∆Es) are driven by random walks that weakly approximate the given Wiener process and Poisson random measure. We establish the weak convergence to the solution of the BSDE and the numerical stability of the sequence of solutions of the BS∆Es. By way of illustration we analyse explicitly a scheme with discrete step-size distributions.

1. Introduction

Backward stochastic differential equations (BSDEs) have turned up in a range of different setting, notably in many applications in mathematical finance such as portfolio optimization and utility indifference pricing, and also as non-linear expectations—see El Karoui *et al.* (1997) for an overview of applications of BSDEs in finance and Delong (2013) for a recent treatment of the case of BSDEs with jumps. Unlike in the case of BSDEs without jumps, exact sampling methods from the probability distribution of the increments of the driving Poisson random measures are in general not readily available, which is an issue in the practical implementation of approximation schemes. Motivated by this observation, we develop in this paper a weak approximation scheme for BSDEs driven by a Wiener process and independent Poisson random measure, allowing the approximating processes to be defined on filtrations that are different from the one the BSDE lives on. We also allow the drivers to take a general Lipschitz-continuous functional form (see (1.2) below), which is encountered in many applications.

Setting. Let $T > 0$ be a given horizon and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space endowed with a filtration $\mathbf{F} = (\mathcal{F}_t)_{t\in[0,T]}$ generated by a d_1-dimensional Wiener process W and an independent d_2-dimensional Lévy process X (*i.e.*, a càdlàg stochastic process with $X_0 = 0$ and stationary independent increments—refer to *e.g.* Sato (1999) for background on Lévy processes). We assume that X is a zero-mean square-integrable process without Gaussian component, in which case X is a pure-jump martingale given by

$$X_t = \int_{[0,t]\times\mathbb{R}^{d_2}\setminus\{0\}} x(N(\mathrm{d}s \times \mathrm{d}x) - \nu(\mathrm{d}x)\mathrm{d}s) = \int_{[0,t]\times\mathbb{R}^{d_2}\setminus\{0\}} x\tilde{N}(\mathrm{d}s \times \mathrm{d}x), \qquad t \in [0, T], \tag{1.1}$$

where ν denotes the Lévy measure of X, N is the Poisson random measure associated to the Poisson point process $(\Delta X_t, t \in [0, T])$ of jumps of X and $\tilde{N}(\mathrm{d}s\times\mathrm{d}x) = N(\mathrm{d}s\times\mathrm{d}x)-\nu(\mathrm{d}x)\mathrm{d}s$ is the corresponding compensated Poisson random measure. We consider in this paper BSDEs of the form

$$Y_t = F + \int_t^T f(s, Y_s, Z_s, \tilde{Z}_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s - \int_{(t,T]\times\mathbb{R}^{d_2}\setminus\{0\}} \tilde{Z}_s(x)\tilde{N}(\mathrm{d}s \times \mathrm{d}x), \quad t \in [0, T], \tag{1.2}$$

with driver function $f : [0, T] \times \mathbb{R} \times \mathbb{R}^{d_1} \times L^2(\nu(\mathrm{d}x), \mathcal{B}(\mathbb{R}^{d_2}\setminus\{0\})) \to \mathbb{R}$,

2000 *Mathematics Subject Classification.* 60H10; 60Fxx.

Key words and phrases. Backward stochastic differential equation (BSDE), Backward stochastic difference equation (BS∆E), convergence, Lévy process, infinite jump-activity.

Acknowledgements: MS acknowledges support by NWO VENI.

for $\mathcal{F}_T$-measurable terminal conditions $F \in L^2(\mathbb{P})$. A triplet $(Y, Z, \tilde{Z})$ is called a solution of this BSDE if (1.2) holds for all $t \in [0,T]$ and the triplet takes values in the product of the spaces $\mathcal{S}^2$, $\mathcal{H}^2$ and $\tilde{\mathcal{H}}^2$ of square-integrable $\mathbf{F}$-adapted semi-martingales Y, predictable processes Z and $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^{d_2} \setminus \{0\})$-measurable processes $\tilde{Z}$, respectively.* Under the standard setup, which we assume to be in force throughout, the driver f is assumed to be such that (i) f is continuous as function of $t \in [0,T]$ at any $(y, z, \tilde{z})$, and (ii) f is Lipschitz continuous in $(y, z, \tilde{z})$ uniformly for all $t \in [0,T]$, that is, there exists a positive K satisfying

$$|f(t, y_1, z_1, \tilde{z}_1) - f(t, y_0, z_0, \tilde{z}_0)| \leq K\left(|y_1 - y_0| + |z_1 - z_0| + \sqrt{\int_{\mathbb{R}^{d_2}\setminus\{0\}} |\tilde{z}_1(x) - \tilde{z}_0(x)|^2 \nu(\mathrm{d}x)}\right), \tag{1.3}$$

for any $y_0, y_1 \in \mathbb{R}$, $z_0, z_1 \in \mathbb{R}^{d_1}$ and $\tilde{z}_0, \tilde{z}_1 \in L^2(\nu(\mathrm{d}x), \mathcal{B}(\mathbb{R}^{d_2}\setminus\{0\}))$. Under these conditions it is well-known that the BSDE (1.2) has a unique solution (see Tang & Li (1994) and Royer (2006)).

Related literature. BSDEs with jumps of the form in (1.1) play an important role in many optimal control problems, see for instance Tang & Li (1994), Eyraud-Loisel (2005), Lim (2006), or Jeanblanc *et al.* (2010). Another main application of BSDEs arises in utility maximization, see for instance El Karoui & Rouge (2000), Hu, Imkeller & Müller (2005), Klöppel & Schweizer (2007), and Sircar & Sturm (2011) in a Brownian filtration. See Mania & Schweizer (2005) and Morlais (2009a) in a continuous filtration, and Becherer (2006) and Morlais (2009b) in a setting with finite jump activity, and Morlais (2009b) and Pelsser & Stadje (2014) in a setting with infinite jump acitivity. Royer (2006) studied BSDEs driven by Brownian motion and a Poisson random measure, and their application to g-expectations. In the references quoted above the optimal solutions were characterized in terms of solutions of BSDEs, but the problem of numerical approximation was not addressed in the case of BSDEs with jumps.

A common way to approximate a BSDE is by discretizing time, replacing the BSDE by an appropriate discrete time backward stochastic difference equation (BSΔE). We will consider the sequence of BSΔEs driven by d_1-dimensional and d_2-dimensional random walks $W^{(\pi)}$ and $X^{(\pi)}$ converging to W and X. In a setting without jumps, convergence results for general random walks have been obtained in Ma *et al.* (2002), Cheridito & Stadje (2013), and in Briand *et al.* (2001, 2002) using Picard iteration arguments as well as results on convergence of filtrations from Coquet *et al.* (2000). While many authors studied discrete schemes for the approximation of solutions of BSDEs in a purely Brownian setting, in a setting with jumps there is considerably less literature available. Lejay *et al.* (2007) is concerned with approximation schemes for BSDEs with one single degenerate jump for a specific approximating process. Contrary to the references mentioned earlier in this paragraph which took a random walk as the approximating process Bouchard & Elie (2007) considered numerical schemes in a pure finite activity jump setting (without a Brownian component) based on a direct discretization of the Lévy process. They showed convergence results for driver functions taking the form $f(t, y, \int_{\mathbb{R}^{d_2}\setminus\{0\}} \rho(x)\tilde{z}(x)x\nu(dx))$, for a bounded functional ρ and that for driver functions of this form it suffices to compute (recursively backwards in time) the integral $\int_{\mathbb{R}^{d_2}\setminus\{0\}} \rho(x)\tilde{z}(x)x\nu(\mathrm{d}x)$. Recently Aazizi (2013) has extended the convergence results of Bouchard & Elie (2007) to the setting of a forward-backward SDE driven by an infinite activity jump-process.

Contributions. In this paper we introduce a discrete-time scheme for the approximation of the solution of a BSDE driven by a Wiener process and an independent Poisson random measure allowing for a general Lipschitz-continuous driver function (where the driver may be a functional of $\tilde{z}$). We prove L^2-stability and convergence

*That is, these processes are square-integrable with respect to

$$|Y|_{\mathcal{S}^2} := \mathbb{E}\left[\sup_{t\in[0,T]} |Y_t|^2\right]^{1/2}, \quad |Z|_{\mathcal{H}^2} := \mathbb{E}\left[\int_0^T |Z_t|^2 \mathrm{d}t\right]^{1/2} \quad |\tilde{Z}|_{\tilde{\mathcal{H}}^2} := \mathbb{E}\left[\int_0^T \int_{\mathbb{R}^{d_2}\setminus\{0\}} |\tilde{Z}_t(x)|^2 \nu(\mathrm{d}x)\mathrm{d}t\right]^{1/2}$$

respectively, where $|\cdot|$ denotes the Euclidean norm. $\mathcal{P}$ denotes the predictable σ-algebra.

results for the solutions to BSΔEs generated by approximating random walks which may not be defined on the same filtration as the continuous-time processes. The prime examples are finite (bi-, tri- and multinomial) trees approximating the driving Brownian motion and Lévy process.

Unlike the schemes considered in Bouchard & Elie (2007) or Aazizi (2013) the weak approximation scheme considered in the current paper neither relies on the discrete process being a discretization of the continuous-time process, nor on the discrete-time process being defined on the same filtration as the continuous-time process, nor on a Markovian structure. In fact our results hold for any suitable random walk and any terminal condition and are more in the spirit of Briand *et al.* (2002). Furthermore, in all of the financial mathematics papers quoted above the BSDEs in question take the more general form given in (1.2) (with the need to compute the whole functional $\tilde{z}$) of which, up to this point, the numerical implementation has received little attention in the literature. One of the main contributions of the current paper is to analyze this case. In particular, our results include the case of a driving Lévy process with infinite jump activity. We show that when the probability of the random walks not moving is strictly positive our BSΔEs satisfy strong L^2-regularity conditions which lead to stable numerical schemes. Note that, in the infinite activity case, approximations schemes for Lévy processes often exclude or use a special technique to approximate the small jumps, see for instance Asmussen & Rosiński (2001) and the reference therein for a discussion.

The outline of the proof of convergence is as follows. We first prove convergence for terminal conditions and drivers satisfying regularity and differentiability conditions on the underlying Hilbert space. To overcome the difficulties arising from a non-continuous limit we apply results from Mémin (2003) concerning the extended convergence of filtrations and use that the solutions of the BSΔEs satisfy appropriate regularity properties. The latter is shown by an induction over the Picard sequences. General arguments on Hilbert spaces then conclude the proof for smooth terminal conditions and drivers. For the general case we deploy the L^2-regularity properties of the solutions of BSΔEs mentioned above.

Contents. The remainder of the paper is structured as follows. First, in Section 2, we present the random walk approximations and review the associated (extended) weak convergence results that form the basis of the approximation schemes under consideration. In Section 3 we show numerical stability of the sequence of approximating BSΔEs driven by these random walks, which forms an important step towards the main result, the convergence theorem, that we present together with its proof in Section 4. By way of illustration we present in Section 5 an example in our setting of an explicit approximating BSΔE scheme driven by a discrete random walk. Some proofs of auxiliary results are deferred to the appendix.

2. Preliminaries

As approximation to the BSDE (1.2) we consider a sequence of discrete-time BSDEs (also referred to as BSΔEs, backward stochastic difference equations) driven by processes with independent stationary increments $(W^{(\pi)}, X^{(\pi)})$ that are constant outside uniform time-grids π, with the collection of grids $\pi = \pi_N, N \in \mathbb{N}$ given by $\pi_N := \{t_0, t_1, \ldots, t_N\}$ with $t_i = iT/N$, $i = 0, \ldots, N$, with mesh denoted by $\Delta = \Delta_N = T/N$. In the sequel we often write $\pi = \pi_N$ when no confusion is possible and identify the process $(W^{(\pi)}, X^{(\pi)})$ with the random walk $(W^{(\pi)}_{t_i}, X^{(\pi)}_{t_i})_{t_i \in \pi}$. In this section we specify these approximating random walks and collect weak-convergence results that are deployed in the sequel.

2.1. **Random walk approximation.** We assume that $W^{(\pi)}$ and $X^{(\pi)}$ are independent, square-intergrable martingales defined on the probability space $(\Omega, \mathcal{F}^{(\pi)}, \mathbb{P})$ which are piecewise constant on $[t_i, t_{i+1})$. More specifically, let $W^{(\pi)} = (W^{(\pi),1}, \ldots, W^{(\pi),d_1})'$ (where $'$ denotes transpose) be a column-vector of zero-mean random walks that have independent stationary increments $\Delta W^{(\pi)}_{t_i} := W^{(\pi)}_{t_{i+1}} - W^{(\pi)}_{t_i}$ with second moment matching the corresponding second moment of a Wiener process subject to a uniform moment-condition, *i.e.*,

$$\mathbb{E}_{t_i}\left[\left(\Delta W^{(\pi)}_{t_i}\right)\left(\Delta W^{(\pi)}_{t_i}\right)'\right] = \Delta I_{d_1}, \qquad i = 0, \ldots, N-1, \tag{2.1}$$

$$\sup_{\pi} \mathbb{E}[|W^{(\pi)}_T|^{2+\epsilon}] < \infty, \quad \text{for some } \epsilon > 0, \tag{2.2}$$

where I_{d_1} the $d_1 \times d_1$ identity matrix and $\mathbb{E}_t[\cdot] = \mathbb{E}[\cdot|\mathcal{F}^{(\pi)}_t]$ for $t \in \pi$, with $\mathbf{F}^{(\pi)} = (\mathcal{F}^{(\pi)}_t, t \in \pi)$ denoting the standard filtration generated by $(W^{(\pi)}, X^{(\pi)})$. The increments of $W^{(\pi)}$ may be for instance be taken to follow suitably chosen multivariate Bernoulli or Gaussian distributions.

Moreover, let $X^{(\pi)} = (X^{(\pi),1}, \ldots, X^{(\pi),d_2})'$ be a (column-vector of) zero-mean random walk with independent stationary increments $\Delta X^{(\pi)}_{t_i} := X^{(\pi)}_{t_{i+1}} - X^{(\pi)}_{t_i}$ satisfying the moment conditions

$$\Delta^{-1/2}\mathbb{E}[|\Delta X^{(\pi)}_{t_i}|] \longrightarrow 0, \quad \Delta \to 0, \quad \text{and} \tag{2.3}$$

$$\Delta^{-1}\mathbb{E}_{t_i}\left[\left(\Delta X^{(\pi)}_{t_i}\right)\left(\Delta X^{(\pi)}_{t_i}\right)'\right] \longrightarrow (\nu_{k,l})^{d_2}_{k,l=1}, \qquad i = 0, \ldots, N-1, \quad \text{with} \tag{2.4}$$

$$\nu_{k,l} = \int h_k(x) h_l(x) \nu(\mathrm{d}x), \qquad h_k(x) = x_k, \;\; k = 1, \ldots, d_2, \text{ and}$$

$$\sup_{\pi} \mathbb{E}[|X^{(\pi)}_T|^{2+\epsilon}] < \infty, \quad \text{for some } \epsilon > 0. \tag{2.5}$$

Note that (2.3) is satisfied when we take $\Delta X^{(\pi)}_{t_i}$ equal to the increment $X_{t_{i+1}} - X_{t_i}$ of X over the interval $[t_i, t_{i+1}]$: since X is square-integrable by (2.5), the first absolute moment of X_t at small t satisfies $\mathbb{E}[|X_t|] = O(t)$ for $t \to 0$ (see Ludschgy & Pagès (2008), Theorem 1).

It is also assumed that the step-size distribution $G^{(\pi)}$ satisfies

$$\int_{\mathbb{R}^{d_2}\setminus\{0\}} g(x)\nu^{(\pi)}(\mathrm{d}x) \longrightarrow \int_{\mathbb{R}^{d_2}\setminus\{0\}} g(x)\nu(\mathrm{d}x), \tag{2.6}$$
$$\text{as } \Delta \to 0, \text{ with } \quad \nu^{(\pi)}(\mathrm{d}x) := \Delta^{-1}G^{(\pi)}(\mathrm{d}x),$$

for all continuous bounded functions $g : \mathbb{R}^{d_2}\setminus\{0\} \to \mathbb{R}$ that are 0 around $x = 0$ and have a limit as $|x| \to \infty$.

Finally, we assume that there is a positive probability that the random walk $X^{(\pi)}$ remains at the same location from one time-step to the next:

$$\liminf_{\Delta \to 0} \mathbb{P}\left(\Delta X^{(\pi)}_{t_i} = 0\right) \geq a, \quad \text{for some } a > 0. \tag{2.7}$$

Under condition (2.7) we establish that the corresponding sequence of BSΔEs is numerically stable (see Theorem 3.4). In the case that X has finite activity (2.7) is naturally satisfied by the strong scheme (X_{t_i}) taking $a = \mathrm{e}^{-\lambda}$ where $\lambda = \nu(\mathbb{R}^{d_2}\setminus\{0\})$ denotes the jump rate. Thus, all strong schemes based on direct discretizations of Lévy processes with finite jump-activity that are in $L^{2+\epsilon}$ satisfy all conditions specified above. While in the complementary case of infinite jump-activity (2.7) is generally not satisfied by a strong discretisation scheme, this condition can be incorporated in the construction of a weak scheme—see Section 5 for explicit examples of weak schemes satisfying (2.7) and all other conditions given above.

The conditions given in (2.1), (2.4) and (2.6) are sufficient to guarantee functional weak convergence of the processes $(W^{(\pi)}, X^{(\pi)})$ to the Lévy process (W, X) as the mesh size Δ tends to zero. More precisely, as $\Delta \to 0$ we have

$$(W^{(\pi)}, X^{(\pi)}) \xrightarrow{\mathcal{L}} (W, X), \tag{2.8}$$

where $\xrightarrow{\mathcal{L}}$ denotes weak-convergence in the Skorokhod J_1-topology. This assertion follows as a direct consequence of classical weak convergence theory (see Thm. VII.3.7 in Jacod & Shiryaev (2003)), given the conditions in (2.1), (2.4) and (2.6), the independent increments property of $(W^{(\pi)}, X^{(\pi)})$ and the independence of $W^{(\pi)}$ and $X^{(\pi)}$ on the one hand and that of X and W on the other hand.

In the sequel we assume that the random variables $(W^{(\pi)}, X^{(\pi)})_\pi$ have been defined such that the convergence in (2.8) holds in probability:

$$(W^{(\pi)}, X^{(\pi)}) \xrightarrow{\mathbf{P}} (W, X), \tag{2.9}$$

where $\xrightarrow{\mathbf{P}}$ denotes convergence in probability in the Skorokhod J_1-topology†. In the next results we collect for later reference a number of ramifications of the convergence in (2.9).

Lemma 2.1. (i) *Let $g : [0,T] \times \mathbb{R}^{d_2} \to \mathbb{R}$ be a continuous function that is 0 in a neighbourhood of 0. Then we have*

$$\sum_{t_i \in \pi \setminus \{T\} \cap [0,\,\cdot\,]} g(t_i, \Delta X^{(\pi)}_{t_i}) \xrightarrow{\mathbf{P}} \int_{[0,\,\cdot\,] \times \mathbb{R}^{d_2} \setminus \{0\}} g(s,x) N(\mathrm{d}s \times \mathrm{d}x),$$

$$\sum_{t_i \in \pi \setminus \{T\} \cap [0,\,\cdot\,]} \{g(t_i, \Delta X^{(\pi)}_{t_i}) - \mathbb{E}_{t_{i-1}}[g(t_i, \Delta X^{(\pi)}_{t_i})]\} \xrightarrow{\mathbf{P}} \int_{[0,\,\cdot\,] \times \mathbb{R}^{d_2} \setminus \{0\}} g(s,x) \tilde{N}(\mathrm{d}s \times \mathrm{d}x), \quad \text{as } \Delta \to 0.$$

The following limit holds in L^1:

$$\lim_{\epsilon \downarrow 0} \limsup_{\Delta \to 0} \sum_{j=1}^{T/\Delta} |\Delta X^{(\pi)}_{t_j}|^2 I_{\{|\Delta X^{(\pi)}_{t_j}| \le \epsilon\}} = 0. \tag{2.10}$$

(ii) *Let $\bar{Z} : [0,T] \times \mathbb{R}^{d_2} \setminus \{0\} \to \mathbb{R}$ be a bounded function that is jointly continuous, and zero in a neighbourhood of zero, and let the function $g^{(\pi)}_s(x)$ be piecewise constant (i.e. $g^{(\pi)}_s = g^{(\pi)}_{t_i}$ for $s \in [t_i, t_{i+1})$), $(\mathcal{F}^{(\pi)}_s \otimes \mathcal{B}(\mathbb{R}^{d_2} \setminus \{0\}))$-measurable for any $s \in [0,T]$, and uniformly Lipschitz continuous as function of x (*i.e.*, for some constant $\hat{K} > 0$ it holds $\sup_{N \in \mathbb{N}, s \in [0,T]} |g^{(\pi_N)}_s(x)| \le \hat{K}|x|$ a.s.). Then we have as $\Delta \to 0$*

$$\mathbb{E}\left[\sup_{i \in \{1,\dots,N\}} \left| \sum_{j=0}^{i-1} \int_{\mathbb{R}^{d_2} \setminus \{0\}} g^{(\pi)}_{t_j}(x)^2 \nu^{(\pi)}(\mathrm{d}x)\Delta - \int_{[0,t_i] \times \mathbb{R}^{d_2} \setminus \{0\}} g^{(\pi)}_s(x)^2 \nu(\mathrm{d}x)\mathrm{d}s \right| \right] \to 0, \tag{2.11}$$

$$\mathbb{E}\left[\sup_{i \in \{1,\dots,N\}} \left| \sum_{j=0}^{i-1} \int_{\mathbb{R}^{d_2} \setminus \{0\}} (\bar{Z}_{t_j} g^{(\pi)}_{t_j})(x) \nu^{(\pi)}(\mathrm{d}x)\Delta - \int_{[0,t_i] \times \mathbb{R}^{d_2} \setminus \{0\}} (\bar{Z}_s g^{(\pi)}_s)(x) \nu(\mathrm{d}x)\mathrm{d}s \right| \right] \to 0. \tag{2.12}$$

Proof. (i) The first relation is a direct consequence of the convergence in (2.8) and the fact that the map $\omega \mapsto (\omega, \sum_{s \le \cdot} g(s, \Delta\omega_s))$ (with $\Delta\omega_s = \omega_s - \omega_{s-}$) is continuous in the Skorokhod J_1-topology (see [18, Cor. VI.2.8]). The second relation follows from the first and the convergence in (2.6). Finally, we turn to (2.10). Note that by (2.4)

$$\limsup_{\Delta \to 0} \mathbb{E}\left[\sum_{j=1}^{T/\Delta} |\Delta X^{(\pi)}_{t_j}|^2\right] = \limsup_{\Delta \to 0} \mathbb{E}[|X^{(\pi)}_T|^2] = \mathbb{E}[|X_T|^2]. \tag{2.13}$$

Furthermore, for any collection of continuous functions $(h_\epsilon)_\epsilon$ satisfying $I_{\{|x| > 2\epsilon\}} \le |h_\epsilon(x)| < I_{\{|x| > \epsilon\}}$ the integrability conditions imply

$$\lim_{\epsilon \downarrow 0} \liminf_{\Delta \to 0} \mathbb{E}\Big[\sum_{j=1}^{T/\Delta} |\Delta X^{(\pi)}_{t_j}|^2 I_{\{|\Delta X^{(\pi)}_{t_j})| > \epsilon\}}\Big] \ge \lim_{\epsilon \downarrow 0} \mathbb{E}\Big[\sum_{t: \Delta X_t \ne 0} |\Delta X_t|^2 h_\epsilon(\Delta X_t)\Big] = \mathbb{E}[|X_T|^2]. \tag{2.14}$$

The combination of (2.13) and (2.14) yields (2.10).

†Such random variables can be constructed by the Skorokhod embedding theorem

(ii) For any $s \in [0,T]$ and $\epsilon > 0$, the triangle inequality implies

$$\left| \int_{\mathbb{R}^{d_2}\backslash\{0\}} g_s^{(\pi)}(x)^2 \nu^{(\pi)}(\mathrm{d}x) - \int_{\mathbb{R}^{d_2}\backslash\{0\}} g_s^{(\pi)}(x)^2 \nu(\mathrm{d}x) \right| \leq \left| I^{(\pi)}(g_s^{(\pi)}) \right| + J_\epsilon^{(\pi)}, \quad \text{with} \tag{2.15}$$

$$J_\epsilon^{(\pi)} = \int_{\{|x|\leq\epsilon\}} g_s^{(\pi)}(x)^2 \nu^{(\pi)}(\mathrm{d}x) + \int_{\{|x|\leq\epsilon\}} g_s^{(\pi)}(x)^2 \nu(\mathrm{d}x),$$

where, for any Borel-function $f \in L^2(\nu(\mathrm{d}x), \mathcal{B}(\mathbb{R}^{d_2}\backslash\{0\})) \cap L^2(\nu^{(\pi)}(\mathrm{d}x), \mathcal{B}(\mathbb{R}^{d_2}\backslash\{0\}))$, we denote

$$I^{(\pi)}(f) = \int_{\{|x|>\epsilon\}} f(x)^2 \nu^{(\pi)}(\mathrm{d}x) - \int_{\{|x|>\epsilon\}} f(x)^2 \nu(\mathrm{d}x). \tag{2.16}$$

Fix $\delta > 0$ arbitrary and choose an $\epsilon > 0$ from the set $\{a \in \mathbb{R}^{(d_2}\backslash\{0\} : \nu(\{x : |x| = |a|\}) = 0\}$ that satisfies

$$\hat{K} \left(\int_{\{|x|\leq\epsilon\}} |x|^2 \nu^{(\pi)}(\mathrm{d}x) + \int_{\{|x|\leq\epsilon\}} |x|^2 \nu(\mathrm{d}x) \right) < \delta, \tag{2.17}$$

uniformly over partitions π [which is possible in view of (2.10)]. Let us first show that $I^{(\pi)}(g_s^{(\pi)})$ converges to zero in L^1 for any $s \in [0,T]$. Let $X_\epsilon^{(\pi)}$ and X_ϵ be the pure-jump processes induced by $X^{(\pi)}$ and X by excluding all jumps smaller than ϵ. Then $X_\epsilon^{(\pi)}$ converges to X_ϵ in the Skorokhod J_1-topology in probability as $\delta \to 0$. Since the position at the epoch of firt exit from a ball is a continuous path-functional in the Skorokhod J_1-topology (see [18, Prop. VI.2.12], it follows in view of the integrability condtion (2.5) that $X_\epsilon^{(\pi)}(\tau_\epsilon^{(\pi)})$ converges to $X_\epsilon(\tau_\epsilon)$ in L^2, where $\tau_\epsilon^{(\pi)} = \inf\{t \geq 0 : |X_{\epsilon,t}^{(\pi)}| > \epsilon\}$ and $\tau_\epsilon = \inf\{t \geq 0 : |X_{\epsilon,t}| > \epsilon\}$ are equal to the first-passage times into the complement of the ball with radius ϵ. The observation that $\tau_\epsilon^{(\pi)}$ and τ_ϵ are equal to the first time that $X_\epsilon^{(\pi)}$ and X_ϵ jump in conjunction with the uniform Lipschitz-continuity of $g^{(\pi)}$, (2.5) and the fact $\nu^{(\pi)}(|x| > \epsilon) \to \nu(|x| > \epsilon)$ then imply

$$\begin{aligned} \lim_{\Delta\to 0} |I^{(\pi)}(g^{(\pi)})| &= \lim_{\Delta\to 0} \nu(|x| > \epsilon) \left| \int_{\{|x|>\epsilon\}} g^{(\pi)}(x)^2 \frac{\nu^{(\pi)}(\mathrm{d}x)}{\nu^{(\pi)}(|x| > \epsilon)} - \int_{\{|x|>\epsilon\}} g^{(\pi)}(x)^2 \frac{\nu(\mathrm{d}x)}{\nu(|x| > \epsilon)} \right| \\ &= \lim_{\Delta\to 0} \nu(|x| > \epsilon) \left| \mathbb{E}\left[\left| g^{(\pi)}(X_\epsilon^{(\pi)}(\tau_\epsilon^{(\pi)})) \right|^2 - \left| g^{(\pi)}(X_\epsilon(\tau_\epsilon)) \right|^2 \right] \right| = 0. \end{aligned}$$

Furthermore, by the uniform Lipschitz-continuity of the function $g_s^{(\pi)}$ we also have that (a) the sequence $I^{(\pi)}(g_s^{(\pi)})$ is uniformly bounded and (b) $J_\epsilon^{(\pi)}$ is bounded by the left-hand side of (2.17). As a consequence, the bounded convergence theorem and the bounds (2.15) and (2.17) imply that the limit as $\Delta \to 0$ of the left-hand side of (2.11) is smaller than $T\delta$. Since δ is arbitrary the convergence stated in (2.11) follows.

The proof of convergence in (2.12) is analogous, and is omitted. □

2.2. **Extended weak convergence.** In order to establish the convergence of BSDEs we also need to deploy the notions of extended weak convergence and weak convergence of filtrations, the definitions of which, we recall from Coquet *et al.* (2004) and Mémin (2003), are given as follows:

Definition 2.2. Given stochastic processes $Z = (Z_t)_{t\in[0,T]}$ and $(Z^n)_{n\in\mathbb{N}}$ with $Z^n = (Z_t^n)_{t\in[0,T]}$ defined on filtered probability spaces $(\Omega, \mathcal{G}, (\mathcal{G}_t), \mathbb{P})$ and $(\Omega, \mathcal{G}^n, (\mathcal{G}_t^n), \mathbb{P})$ respectively, we say (i) $\mathcal{G}^n$ weakly converges to $\mathcal{G}$ [denoted $\mathcal{G}^n \stackrel{w}{\to} \mathcal{G}$] if for every $A \in \mathcal{G}$ the sequence of processes $(\mathbb{E}[I_A|\mathcal{G}_t^n])_{t\in[0,T]}$ converges to the process $(\mathbb{E}[I_A|\mathcal{F}_t])_{t\in[0,T]}$ and (ii) $(Z^n, \mathcal{G}^n)$ weakly converges to $(Z, \mathcal{G})$ [denoted $(Z^n, \mathcal{G}^n) \stackrel{w}{\to} (Z, \mathcal{G})$] if for every $A \in \mathcal{G}$ the sequence of processes $(Z_t^n, \mathbb{E}[I_A|\mathcal{G}_t^n])_{t\in[0,T]}$ converges to the process $(Z_t, \mathbb{E}[I_A|\mathcal{F}_t])_{t\in[0,T]}$. In both cases the convergence is in probability under the Skorokhod J_1-topology (on the space D of càdlàg functions).

Remark 2.3. It is clear that the notion of extended weak convergence in general is stronger than the notion of weak convergence of filtration (see for instance Coquet *et al.* (2004) and Mémin (2003) for a discussion).

However, in the notation of the previous definition, if F_i^n converges to F_i in L^1 for $i = 1, \dots, m$ and $\mathcal{G}^n \xrightarrow{w} \mathcal{G}$, it may be shown by an application of Doob's maximal inequality (see Coquet *et al.* (2004), Remark 1) that we have the convergence $(\mathbb{E}[F_1^n|\mathcal{G}^n_\cdot], \dots, \mathbb{E}[F_m^n|\mathcal{G}^n_\cdot]) \to (\mathbb{E}[F_1|\mathcal{G}_\cdot], \dots, \mathbb{E}[F_m|\mathcal{G}_\cdot])$ in probability in the Skorokhod J_1-topology. In particularly, if $\mathcal{G}^n$ converges to $\mathcal{G}$ weakly, L^n is a $\mathcal{G}^n$-martingale and L is a $\mathcal{G}$-martingale then $L^n_T \to L_T$ in L^1 implies that $(L^n, \mathcal{G}^n) \xrightarrow{w} (L, \mathcal{G})$ in the extended sense, see also Proposition 7 in Coquet *et al.* (2004) or Proposition 1 in Mémin (2003).

We recall (from Proposition 2 in Mémin (2003)) that $(W^{(\pi)}, X^{(\pi)})$ converges to the Lévy process (W, X) in the sense of extended convergence, due to the independence of the increments of the two-coordinate processes $W^{(\pi)}$ and $X^{(\pi)}$, in conjunction with the fact that the filtration $\mathcal{F}^{(\pi)}$ is generated by the process $(W^{(\pi)}, X^{(\pi)})$:

Proposition 2.4 (Proposition 2, Mémin (2003))**.** *We have* $((W^{(\pi)}, X^{(\pi)}), \mathcal{F}^{(\pi)}) \xrightarrow{w} ((W, X), \mathcal{F})$ *as* $\Delta \to 0$. *In particular,* $\mathcal{F}^{(\pi)} \xrightarrow{w} \mathcal{F}$.

If a sequence of square-integrable martingales converges to a limit in the sense of extended convergence that is given above, the convergence of the corresponding quadratic variation and predictable compensator processes also holds true, which is a fact that is deployed in the proof of convergence of BSDEs.

Theorem 2.5 (Corollary 12, Mémin (2003))**.** *Let* $(L^{(\pi)})$ *be a sequence of square integrable* $\mathcal{G}^{(\pi)}$*-measurable martingales, and let* L *be a square integrable quasi-left continuous* $(\mathcal{G}_t)$*-martingale. If* $L_T^{(\pi)} \to L_T$ *in* L^2 *and* $(L^{(\pi)}, \mathcal{G}^{(\pi)}) \xrightarrow{w} (L, \mathcal{G})$*, then we have*

$$\big(L^{(\pi)}, [L^{(\pi)}, L^{(\pi)}], \langle L^{(\pi)}, L^{(\pi)}\rangle\big) \to \big(L, [L, L], \langle L, L\rangle\big)$$

in probability under the Skorokhod J_1*-topology, where, for any square integrable martingale* M*,* $[M, M]$ *and* $\langle M, M\rangle$ *denote the associated quadratic variation and predictable compensator, respectively.*

We record some results concerning the convergence of cross-variations which follow as implications of Theorem 2.5 and are deployed later in the paper.

Corollary 2.6. *Under the assumptions on the processes* $(L^{(\pi)})$ *and* L *in Theorem 2.5, the following hold true:*

(i) *As* $\Delta \to 0$*,* $\langle W^{(\pi)}, L^{(\pi)}\rangle \to \langle W, L\rangle$*, in probability in the Skorokhod* J_1*-topology.*

(ii) *Assume that* $\bar{Z} : [0, T] \times \mathbb{R}^{d_2} \to \mathbb{R}$ *is bounded, jointly continuous, and zero in an environment around zero, and consider the stochastic processes* $U^{(\pi)} = (U_t^{(\pi)})_{t\in[0,T]}$ *and* $U = (U_t)_{t\in[0,T]}$ *given by*

$$U_t^{(\pi)} := \sum_{t_i \in \pi\cap[0,t]} \{\bar{Z}_{t_i}(\Delta X_{t_i}^{(\pi)}) - \mathbb{E}_{t_{i-1}}[\bar{Z}_{t_i}(\Delta X_{t_i}^{(\pi)})]\}, \quad U_t := \int_{[0,t]\times\mathbb{R}^{d_2}\setminus\{0\}} \bar{Z}_s(x)\tilde{N}(\mathrm{d}s \times \mathrm{d}x).$$

As $\Delta \to 0$*,* $\langle U^{(\pi)}, L^{(\pi)}\rangle \to \langle U, L\rangle$*, in probability in the Skorokhod* J_1*-topology.*

Proof. (i) Since $W^{(\pi)}$ $(L^{(\pi)})$ converges to W (L, respectively) in probability in the Skorokhod J_1-topology and W is continuous, this entails that joint processes $(W^{(\pi)} + L^{(\pi)}, W^{(\pi)} - L^{(\pi)})$ converge in probability in J_1 to $(W + L, W - L)$. By Remark 2.3 and Proposition 2.4 this convergence holds true in the extended sense with the filtrations $\mathcal{F}^{(\pi)}$ and $\mathcal{F}$. Since $(W_T^{(\pi)} + L_T^{(\pi)}, W_T^{(\pi)} - L_T^{(\pi)})$ actually converges in L^2 to $(W + L, W - L)$ (by assumption for L and by conditions (2.1) and (2.2) for W), we deduce from Theorem 2.5 that $(\langle W^{(\pi)} + L^{(\pi)}\rangle, \langle W^{(\pi)} - L^{(\pi)}\rangle)$ converges to $(\langle W + L\rangle, \langle W - L\rangle)$ in probability in the Skorokhod J_1-topology. As a consequence, we have

$$\langle W^{(\pi)}, L^{(\pi)}\rangle = \frac{1}{4}\Big(\langle W^{(\pi)} + L^{(\pi)}\rangle - \langle W^{(\pi)} - L^{(\pi)}\rangle\Big) \to \frac{1}{4}\Big(\langle W + L\rangle - \langle W - L\rangle\Big) = \langle W, L\rangle \tag{2.18}$$

in probability in the Skorokhod J_1-topology, as stated.

(ii) We start by noting (from Lemma 2.1) that as $\Delta \to 0$ $U^{(\pi)}$ converges to U in probability in the Skorokhod J_1-topology. As $\bar{Z}$ is bounded and zero in a neighbourhood of zero, it follows from (2.5) that the collection $(U^{(\pi)})_\pi$ is bounded in $L^{2+\epsilon}$, so that in particular $U_T^{(\pi)} \to U_T$ in L^2. Since the filtration satisfy $\mathcal{F}^{(\pi)} \overset{w}{\to} \mathcal{F}$ we have (by Proposition 2.4 and Remark 2.3)

$$(U_\cdot^{(\pi)}, L_\cdot^{(\pi)}) = (\mathbb{E}[U_T^{(\pi)}|\mathcal{F}_\cdot^{(\pi)}], \mathbb{E}[L_T^{(\pi)}|\mathcal{F}_\cdot^{(\pi)}]) \overset{\Delta \to 0}{\to} (\mathbb{E}[U_T|\mathcal{F}_\cdot], \mathbb{E}[L_T|\mathcal{F}_\cdot]) = (U_\cdot, L_\cdot).$$

By similar arguments as in part (i) it then follows that we have the convergence of $\langle U^{(\pi)}, L^{(\pi)}\rangle$ to $\langle U, L\rangle$ in probability in the Skorokhod J_1-topology. □

3. BSΔEs

We turn next to the formulation of the approximating BSΔEs, the construction of their solutions and numerical stability.

3.1. **Formulation.** Since by switching from the Wiener process W to the process $W^{(\pi)}$ we lose the predictable representation property, it is well known that we need to include in the formulation of the BSΔE an additional orthogonal martingale term $(M^{(\pi)})$, which thus leads us to the following BSΔE on the grid π:

$$\begin{aligned} Y_{t_i}^{(\pi)} &= F^{(\pi)} + \sum_{j=i}^{N-1} f^{(\pi)}(t_j, Y_{t_j}^{(\pi)}, Z_{t_j}^{(\pi)}, \tilde{Z}_{t_j}^{(\pi)})\Delta - \sum_{j=i}^{N-1} Z_{t_j}^{(\pi)} \Delta W_{t_j}^{(\pi)} \\ &\quad - \sum_{j=i}^{N-1} \left\{ \tilde{Z}_{t_j}^{(\pi)}(\Delta X_{t_j}^{(\pi)}) I_{\{\Delta X_{t_j}^{(\pi)} \neq 0\}} - \mathbb{E}_{t_j}\left[\tilde{Z}_{t_j}^{(\pi)}(\Delta X_{t_j}^{(\pi)}) I_{\{\Delta X_{t_j}^{(\pi)} \neq 0\}} \right] \right\} - \left(M_T^{(\pi)} - M_{t_i}^{(\pi)} \right), \end{aligned} \tag{3.1}$$

where the random variable $F^{(\pi)} \in L^2(\mathcal{F}_T^{(\pi)})$ is the final condition, and the driver $f^{(\pi)} : [0,T] \times \mathbb{R} \times \mathbb{R}^{d_1} \times L^2(\nu^{(\pi)}, \mathcal{B}(\mathbb{R}^{d_2}\backslash\{0\})) \to \mathbb{R}$ is a function that is piecewise constant (*i.e.*, $f^{(\pi)}(s,\cdot) = f^{(\pi)}(t_i,\cdot)$ for $s \in [t_i, t_{i+1})$) and is uniformly Lipschitz-continuous in $(y, z, \tilde{z})$, *i.e.*, for some $K > 0$ we have for all $t \in [0,T]$

$$|f^{(\pi)}(t, y_1, z_1, \tilde{z}_1) - f^{(\pi)}(t, y_0, z_0, \tilde{z}_0)| \leq K \left(|y_1 - y_0| + |z_1 - z_0| + \sqrt{\mathbb{E}_{\nu^{(\pi)}}[(\tilde{z}_1(\xi) - \tilde{z}_0(\xi))^2]} \right), \tag{3.2}$$

where, for any Borel-function f, $\mathbb{E}_{\nu^{(\pi)}}[f(\xi)^2] := \int f(z)^2 \nu^{(\pi)}(\mathrm{d}z)$.

A quadruple $(Y^{(\pi)}, Z^{(\pi)}, \tilde{Z}^{(\pi)}, M^{(\pi)})$ is a solution of the BSΔE (3.1) if it satisfies (3.1) for all $t_i \in \pi$ where $Y_{t_i}^{(\pi)}$ and (the components of the row-vector) $Z_{t_i}^{(\pi)}$ are in $L^2(\mathrm{d}\mathbb{P}, \mathcal{F}_{t_i}^{(\pi)})$, $\tilde{Z}_{t_i}^{(\pi)}$ lies in $L^2(G^{(\pi)}(\mathrm{d}x) \times \mathrm{d}\mathbb{P}, \mathcal{B}(\mathbb{R}^{d_2}\backslash\{0\}) \otimes \mathcal{F}_{t_i}^{(\pi)})$ and $M^{(\pi)} = (M_{t_i}^{(\pi)})$ is a zero-mean square-integrable $\mathbf{F}^{(\pi)}$-martingale on π that is orthogonal to $(W_{t_i}^{(\pi)})$ and to the martingales $(M_{t_i}^k)$ with increments $\Delta M_{t_i}^k = k_{t_i}(\Delta X_{t_i}^{(\pi)}) - \mathbb{E}_{t_i}\left[k_{t_i}(\Delta X_{t_i}^{(\pi)})\right]$ for any function $(k_{t_i})_{t_i}$ with $k_{t_i} \in L^2(G^{(\pi)}(\mathrm{d}x) \times \mathrm{d}\mathbb{P}, \mathcal{B}(\mathbb{R}^{d_2}) \otimes \mathcal{F}_{t_i}^{(\pi)})$.

The BSΔE can be equivalently expressed in differential notation as

$$\begin{aligned} \Delta Y_{t_i}^{(\pi)} &= -f^{(\pi)}(t_i, Y_{t_i}^{(\pi)}, Z_{t_i}^{(\pi)}, \tilde{Z}_{t_i}^{(\pi)})\Delta + Z_{t_i}^{(\pi)} \Delta W_{t_i}^{(\pi)} \\ &\quad + \left\{ \tilde{Z}_{t_i}^{(\pi)}(\Delta X_{t_i}^{(\pi)}) I_{\{\Delta X_{t_i}^{(\pi)} \neq 0\}} - \mathbb{E}_{t_i}\left[\tilde{Z}_{t_i}^{(\pi)}(\Delta X_{t_i}^{(\pi)}) I_{\{\Delta X_{t_i}^{(\pi)} \neq 0\}} \right] \right\} + \Delta M_{t_i}^{(\pi)}, \end{aligned} \tag{3.3}$$

$$Y_T^{(\pi)} = F^{(\pi)}, \tag{3.4}$$

where $i = 0, \ldots, N-1$. We have the following result:

Proposition 3.1. *For $\Delta < 1/K$ the BSΔE* (3.1) *has a unique solution* $(Y^{(\pi)}, Z^{(\pi)}, \tilde{Z}^{(\pi)}, M^{(\pi)})$*, which satisfies the relations: for* $t_i \in \pi$*,*

$$Y_{t_i}^{(\pi)} = f^{(\pi)}(t_i, Y_{t_i}^{(\pi)}, Z_{t_i}^{(\pi)}, \tilde{Z}_{t_i}^{(\pi)})\Delta + \mathbb{E}_{t_i}[Y_{t_{i+1}}^{(\pi)}] \tag{3.5}$$

$$= \mathbb{E}_{t_i}\left[F^{(\pi)} + \sum_{j=i}^{N-1} f^{(\pi)}(t_j, Y_{t_j}^{(\pi)}, Z_{t_j}^{(\pi)}, \tilde{Z}_{t_j}^{(\pi)})\Delta\right], \tag{3.6}$$

$$Z_{t_i}^{(\pi)} = \Delta^{-1}\, \mathbb{E}_{t_i}\left[Y_{t_{i+1}}^{(\pi)} \Delta W_{t_i}^{(\pi)}\right], \tag{3.7}$$

$$\tilde{Z}_{t_i}^{(\pi)}(x) = \mathbb{E}_{t_i}\left[Y_{t_{i+1}}^{(\pi)} | \Delta X_{t_i}^{(\pi)} = x\right] - \mathbb{E}_{t_i}\left[Y_{t_{i+1}}^{(\pi)} | \Delta X_{t_i}^{(\pi)} = 0\right], \tag{3.8}$$

$$\begin{aligned} \Delta M_{t_i}^{(\pi)} &= Y_{t_{i+1}}^{(\pi)} - \mathbb{E}_{t_i}\left[Y_{t_{i+1}}^{(\pi)}\right] - Z_{t_i}^{(\pi)} \Delta W_{t_i}^{(\pi)} \\ &\quad - \left\{ \tilde{Z}_{t_i}^{(\pi)}(\Delta X_{t_i}^{(\pi)}) I_{\{\Delta X_{t_i}^{(\pi)} \neq 0\}} - \mathbb{E}_{t_i}\left[\tilde{Z}_{t_i}^{(\pi)}(\Delta X_{t_i}^{(\pi)}) I_{\{\Delta X_{t_i}^{(\pi)} \neq 0\}}\right]\right\}. \end{aligned} \tag{3.9}$$

Proof. First of all we verify that a given solution $(Y^{(\pi)}, Z^{(\pi)}, \tilde{Z}^{(\pi)}, M^{(\pi)})$ of the BSΔE (3.3) satisfies the stated relations. By taking conditional expectations with respect to $\mathcal{F}_{t_i}^{(\pi)}$ in (3.1) and (3.3) and using that the martingale increments $\Delta W_{t_i}^{(\pi)}$, $\tilde{Z}_{t_i}^{(\pi)}(\Delta X_{t_i}^{(\pi)}) - \mathbb{E}_{t_i}\left[\tilde{Z}_{t_i}^{(\pi)}(\Delta X_{t_i}^{(\pi)})\right]$ and $\Delta M_{t_i}^{(\pi)}$ are orthogonal and have zero mean we find (3.5) and (3.6). Similarly, multiplying the left- and right-hand sides of (3.3) with the coordinates of the vector $\Delta W_{t_i}^{(\pi)}$ and subsequently taking the $\mathcal{F}_{t_i}^{(\pi)}$-conditional expectations yields (3.7) in view of the moment condition in (2.1). Multiplying with an arbitrary function $g \in L^\infty(\mathcal{F}_{t_i}^{(\pi)} \otimes \mathcal{B}(\mathbb{R}^{d_2}))$ and taking conditional expectations and using (3.5) shows denoting $A = \{\Delta X_{t_i}^{(\pi)} \neq 0\}$

$$\mathbb{E}_{t_i}\left[\left\{Y_{t_{i+1}}^{(\pi)} - \mathbb{E}_{t_i}[Y_{t_{i+1}}^{(\pi)}]\right\} g(\Delta X_{t_i}^{(\pi)})\right] = \mathbb{E}_{t_i}\left[\left\{\tilde{Z}_{t_i}^{(\pi)}(\Delta X_{t_i}^{(\pi)}) I_A - \mathbb{E}_{t_i}[\tilde{Z}_{t_i}^{(\pi)}(\Delta X_{t_i}^{(\pi)}) I_A]\right\} g(\Delta X_{t_i}^{(\pi)})\right], \tag{3.10}$$

which implies $I_A \tilde{Z}_{t_i}^{(\pi)}(\Delta X_{t_i}^{(\pi)}) = C + \mathbb{E}_{t_i}[Y_{t_{i+1}} | \Delta X_{t_i}^{(\pi)}]$ for some $C \in L^2(\mathcal{F}_{t_i}^{(\pi)})$. By inserting this expression into (3.10) and taking $g(x) = I_{\{0\}}(x)$ we find with $A^c = \{\Delta X_{t_i}^{(\pi)} = 0\}$

$$-\left(C + \mathbb{E}_{t_i}[Y_{t_{i+1}}^{(\pi)}]\right) \mathbb{E}_{t_i}[I_{A^c}] = \mathbb{E}_{t_i}[Y_{t_{i+1}}^{(\pi)} I_{A^c}] - \mathbb{E}_{t_i}[Y_{t_{i+1}}^{(\pi)}]\mathbb{E}_{t_i}[I_{A^c}] \Rightarrow C = -\mathbb{E}_{t_i}\left[Y_{t_{i+1}}^{(\pi)} \middle| A^c\right],$$

which implies that we have (3.8). The relation (3.9) directly follows by combining (3.3) and (3.5).

Next we verify existence. Define the quadruple $(Y^{(\pi)}, Z^{(\pi)}, \tilde{Z}^{(\pi)}, M^{(\pi)})$ by the right-hand sides of (3.5), (3.7), (3.8) and (3.9). Note that $Y^{(\pi)}$ is determined uniquely by the implicit equation (3.5) (since the map $\Psi : L^2(\mathrm{d}\mathbb{P}, \mathcal{F}_{t_i}^{(\pi)}) \to L^2(\mathrm{d}\mathbb{P}, \mathcal{F}_{t_i}^{(\pi)})$ given by $\Psi(Y) = f^{(\pi)}(t_i, Y, Z_{t_i}^{(\pi)}, \tilde{Z}_{t_i}^{(\pi)})\Delta + \mathbb{E}_{t_i}[Y_{t_{i+1}}^{(\pi)}]$ is a contraction in case $K\Delta < 1$ as a consequence of the Lipschitz condition (3.2)). Furthermore, it is straightforward to verify that the measurability and integrability requirements are satisfied, as well as (3.3).

Finally, we verifty the orthogonality of the martingale $M^{(\pi)}$. To see that $M^{(\pi)}$ and $W^{(\pi)}$ are orthogonal, we note that since $\left\{\tilde{Z}_{\cdot}^{(\pi)}(\Delta X_{\cdot}^{(\pi)}) I_{\{\Delta X_{\cdot}^{(\pi)} \neq 0\}} - \mathbb{E}_{\cdot}\left[\tilde{Z}_{\cdot}^{(\pi)}(\Delta X_{\cdot}^{(\pi)}) I_{\{\Delta X_{\cdot}^{(\pi)} \neq 0\}}\right]\right\}$ and $\Delta W_{\cdot}^{(\pi)}$ are orthogonal, we have by definition of $Z_{t_i}^{(\pi)}$ and $\Delta W_{t_i}^{(\pi)}$

$$\mathbb{E}_{t_i}[\Delta M_{t_i}^{(\pi)} \Delta W_{t_i}^{(\pi)}] = \mathbb{E}_{t_i}[Y_{t_{i+1}}^{(\pi)} \Delta W_{t_i}^{(\pi)}] - \mathbb{E}_{t_i}[(Z_{t_i}^{(\pi)} \Delta W_{t_i}^{(\pi)}) \Delta W_{t_i}^{(\pi)}] = 0.$$

Furthermore, for any function $k_{t_i} \in L^\infty(\mathcal{F}_{t_i}^{(\pi)} \otimes \mathcal{B}(\mathbb{R}^{d_2}))$ it holds

$$\begin{aligned} &\mathbb{E}_{t_i}[\Delta M_{t_i}^{(\pi)} \{k_{t_i}(\Delta X_{t_i}^{(\pi)}) - \mathbb{E}_{t_i}[k_{t_i}(\Delta X_{t_i}^{(\pi)})]\}] \\ &= \mathbb{E}_{t_i}[Y_{t_{i+1}}^{(\pi)} k_{t_i}(\Delta X_{t_i}^{(\pi)})] - \mathbb{E}_{t_i}[Y_{t_{i+1}}^{(\pi)}]\mathbb{E}_{t_i}[k_{t_i}(\Delta X_{t_i}^{(\pi)})] \\ &\quad - \mathbb{E}_{t_i}[\tilde{Z}_{t_i}^{(\pi)}(\Delta X_{t_i}^{(\pi)}) k_{t_i}(\Delta X_{t_i}^{(\pi)})] + \mathbb{E}_{t_i}[\tilde{Z}_{t_i}^{(\pi)}(\Delta X_{t_i}^{(\pi)})]\mathbb{E}_{t_i}[k_{t_i}(\Delta X_{t_i}^{(\pi)})] = 0, \end{aligned}$$

where we used that $\tilde{Z}^{(\pi)}_{t_i}(0) = 0$, inserted the form (3.8) and used the tower-property of conditional expectation. Hence, $M^{(\pi)}$ is orthogonal to the martingales with increments $k_{t_i}(\Delta X^{(\pi)}_{t_i}) - \mathbb{E}_{t_i}[k_{t_i}(\Delta X^{(\pi)}_{t_i})]$, and the proof is complete. □

In the case that the final value $F^{(\pi)}$ is independent of $W^{(\pi)}$ the orthogonal martingale $M^{(\pi)}$ vanishes.

Proposition 3.2. *If $F^{(\pi)}$ is independent of $W^{(\pi)}$ then $M^{(\pi)} \equiv 0$.*

In particular, it follows that in the pure jump case, the martingale $M^{(\pi)}$ is zero and the representation property holds true.

Proof. The assertion follows directly from (3.9) once we have shown that in the case that $F^{(\pi)} \in L^2(\mathcal{F}^{(\pi)}_{t_{i+1}})$ is independent of $W^{(\pi)}$ then $Z^{(\pi)}_{t_i}$ and $\tilde{Z}^{(\pi)}_{t_i}$ defined in (3.7) and (3.8) are such that $Z^{(\pi)}_{t_i} = 0$ and

$$F^{(\pi)} = \mathbb{E}_{t_i}[F^{(\pi)}] + \left\{ \tilde{Z}^{(\pi)}_{t_i}(\Delta X^{(\pi)}_{t_i}) - \mathbb{E}_{t_i}\left[\tilde{Z}^{(\pi)}_{t_i}(\Delta X^{(\pi)}_{t_i}) \right] \right\}. \tag{3.11}$$

That $Z^{(\pi)}_{t_i} = 0$ follows directly from (3.7) [with $[Y^{(\pi)}_{t_{i+1}} = F^{(\pi)}]$, since $W^{(\pi)}_{t_i}$ has zero mean and is independent of F. To see that the identity (3.11) holds we first note that, as $F^{(\pi)} \in L^2(\mathrm{d}\mathbb{P}, \mathcal{F}^{(\pi)}_{t_{i+1}})$ and $F^{(\pi)}$ is independent of $W^{(\pi)}$ there exists a function f in $L^2(G^{(\pi)}(\mathrm{d}x) \times \mathrm{d}\mathbb{P}, \mathcal{B}(\mathbb{R}^{d_2}) \otimes \mathcal{F}^{(\pi)}_{t_i})$ satisfying $F^{(\pi)} = f(\Delta X^{(\pi)}_{t_i})$. Inserting the forms of $F^{(\pi)}$ and $\tilde{Z}^{(\pi)}_{t_i}$ in the rhs of (3.11) and performing straightforward manipulations (similar to those in the proof of Proposition 3.1) shows that the rhs and lhs in (3.11) coincide. □

3.2. **Numerical stability.** In this section we turn to the numerical stability of the BSΔEs in L^2 sense. We start by specifying uniform conditions for the collection of drivers $(f^{(\pi)})$ of the BSΔEs.

Assumption 1. (i) For some $K > 0$, the drivers $f^{(\pi)}$ are uniformly K-Lipschitz continuous (*i.e.*, $f^{(\pi)}$ satisfies (3.2)).

(ii) $f^{(\pi)}(t, 0, 0, 0)$ is bounded uniformly over all $t \in \pi$ and partitions π.

(iii) For every $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^{d_1}$ and uniformly Lipschitz continuous function $\tilde{z}$ (*i.e.*, $\tilde{z}$ for which $|\tilde{z}(x)|/|x|$ is bounded over all $x \in \mathbb{R}^{d_2} \backslash \{0\}$), we have

$$\lim_{\Delta \to 0} f^{(\pi)}(t, y, z, \tilde{z}) = f(t, y, z, \tilde{z}). \tag{3.12}$$

Remarks 3.3. (i) Note that the functions $f^{(\pi)}(t, y, z, \tilde{z})$ in (3.12) are well-defined since every Lipschitz continuous function $\tilde{z}$ is square-integrable with respect to the measures $\nu^{(\pi)}$ and ν.

(ii) In Assumption 1 (iii) it suffices to require the convergence of the drivers only for uniformly Lipschitz continuous functions $\tilde{z}$ as these functions form a dense subset in $L^2(\nu^{(\pi)}, \mathcal{B}(\mathbb{R}^{d_2} \backslash \{0\}))$.

(iii) When the driver $f(t, y, z, \cdot)$ is distribution-invariant under the measure $\nu(\mathrm{d}x)$, *i.e.*, there exists a function $\hat{f}$ such that $f(t, y, z, \tilde{z}) = \hat{f}(t, y, z, \nu \circ \tilde{z}^{-1})$, a natural first candidate for $f^{(\pi)}$ would be to set $f^{(\pi)}(t, y, z, \tilde{z}) := \hat{f}(t, y, z, \nu^{(\pi)} \circ \tilde{z}^{-1})$.

We have the following estimate for BSΔEs as in (3.1) with drivers $f^{(\pi),0}, f^{(\pi),1}$ and terminal conditions $F^{(\pi),0}, F^{(\pi),1}$ and corresponding solution quadruples denoted by $(Y^{(\pi),k}, Z^{(\pi),k}, \tilde{Z}^{(\pi),k}, M^{(\pi),k})$, $k = 0, 1$, respectively.

Theorem 3.4. *There exists an $n_0 \in \mathbb{N}$ and a constant $\bar{C}$ such that for all $\pi = \pi_N$ with $N \geq n_0$, all drivers $f^{(\pi),0}, f^{(\pi),1}$ satisfying Assumption 1(i)-(ii), and square integrable terminal conditions $F^{(\pi),0}, F^{(\pi),1}$, and $t_i \in \pi$,* *we have*

$$\mathbb{E}\left[\max_{t_j\le t_i, t_j\in\pi}|\delta Y_{t_j}^{(\pi)}|^2+\sum_{j=0}^{i-1}\left\{|\delta Z_{t_j}^{(\pi)}|^2\Delta+|\delta M_{t_j}^{(\pi)}|^2+|\delta\tilde{Z}_{t_j}^{(\pi)}(\Delta X_{t_j}^{(\pi)})-\mathbb{E}_{t_j}[\delta\tilde{Z}_{t_j}^{(\pi)}(\Delta X_{t_j}^{(\pi)})]|^2\right\}\right] \le \bar{C}\mathbb{E}\left[|\delta Y_{t_i}^{(\pi)}|^2+\sum_{j=0}^{i-1}|\delta f^{(\pi)}(t_j,Y_{t_j}^{(\pi),0},Z_{t_j}^{(\pi),0},\tilde{Z}_{t_j}^{(\pi),0})|^2\Delta\right], \tag{3.13}$$

with $\delta Y^{(\pi)}=Y^{(\pi),0}-Y^{(\pi),1}$*, etc.*

Remark 3.5. In continuous-time the following analogous estimate holds true for some constant $\bar{c}>0$:

$$\begin{aligned}\mathbb{E}&\left[\sup_{0\le t\le t'}|\delta Y_t|^2+\int_0^{t'}|\delta Z_s|^2\mathrm{d}s+\int_{[0,t']\times\mathbb{R}^{d_2}\setminus\{0\}}|\delta\tilde{Z}_s(x)|^2\nu(\mathrm{d}x)\mathrm{d}s\right]\\ \le\ &\bar{c}\,\mathbb{E}\left[|\delta Y_{t'}|^2+\int_0^{t'}|\delta f(s,Y_s^0,Z_s^0,\tilde{Z}_s^0)|^2\mathrm{d}s\right],\qquad t'\in[0,T].\end{aligned} \tag{3.14}$$

For a proof of (3.14), see for instance to Proposition 3.3 in Becherer (2006) or Lemma 3.1.1 in Delong (2013).

In the proof of Theorem 3.4, which is provided in the Appendix, the following estimate is deployed which is a consequence of the zero-jump-condition (2.7):

Lemma 3.6. *There exist* $\delta_0>0$ *and* $C'>0$ *such that for all* $\Delta\le\delta_0$*, for all functions* $(\tilde{U}_{t_j})_j$ *with* $\tilde{U}_{t_j}(0)=0$ *and* $\tilde{U}_{t_j}\in L^2(\nu^{(\pi)}(\mathrm{d}x)\times\mathrm{d}\mathbb{P},\mathcal{B}(\mathbb{R}^{d_2})\otimes\mathcal{F}_{t_j}^{(\pi)})$*, and for any* $j=0,\dots,n-1$ *we have*

$$\sum_{i=j}^{n-1}\left(\mathbb{E}_{t_i}\left[|\tilde{U}_{t_i}(\Delta X_{t_i}^{(\pi)})|^2\right]-\left|\mathbb{E}_{t_i}\left[\tilde{U}_{t_i}(\Delta X_{t_i}^{n})\right]\right|^2\right)\ge C'\sum_{i=j}^{n-1}\left|\mathbb{E}_{t_i}\left[\tilde{U}_{t_i}(\Delta X_{t_i}^{(\pi)})\right]\right|^2. \tag{3.15}$$

Proof. Assume without loss of generality that $j=0$. Using Hölder's inequality we have

$$\begin{aligned}\sum_{i=0}^{n-1}\left|\mathbb{E}_{t_i}\left[\tilde{U}_{t_i}(\Delta X_{t_i}^{(\pi)})\right]\right|^2&=\sum_{i=0}^{n-1}\left|\mathbb{E}_{t_i}\left[\tilde{U}_{t_i}(\Delta X_{t_i}^{(\pi)})I_{\{\Delta X_{t_i}^{(\pi)}\neq 0\}}\right]\right|^2\\ &\le\left(\max_i\mathbb{P}[\Delta X_{t_i}^{(\pi)}\neq 0]\right)\sum_{i=0}^{n-1}\mathbb{E}_{t_i}\left[\left|\tilde{U}_{t_i}(\Delta X_{t_i}^{(\pi)})\right|^2\right].\end{aligned} \tag{3.16}$$

Since $X^{(\pi)}$ has stationary increments the first factor in the final line is equal to $\mathbb{P}[\Delta X_{t_1}^{(\pi)}\neq 0]$, which is bounded above by $(1-a+\delta)$ for all partitions with mesh $\Delta\le\delta_0$, where δ is some number small enough such that $a-\delta>0$, and δ_0 is chosen sufficiently small using (2.7). By combining the upper bound with (3.16) we obtain (3.15) (with $C'=a-\delta$). □

3.3. **Solution of the BSΔE via Picard iteration.** The process $(Y^{(\pi)},Z^{(\pi)},\tilde{Z}^{(\pi)})$ satisfying the BSΔE can be obtained as the limit of an recursively defined Picard sequence $(Y^{(\pi,p)},Z^{(\pi,p)},\tilde{Z}^{(\pi,p)})_{p\in\mathbb{N}^*}$, which is initialised with $(Y^{(\pi,0)},Z^{(\pi,0)},\tilde{Z}^{(\pi,0)})\equiv(0,0,0)$ and is defined for $p\in\mathbb{N}$ and $t_i\in\pi$ by the right-hand sides of formulas (3.6), (3.7) and (3.8) respectively, with $Y_{t_j}^{(\pi)}$, $Z_{t_j}^{(\pi)}$ and $\tilde{Z}_{t_j}^{(\pi)}$ replaced by $Y_{t_j}^{(\pi,p-1)}$, $Z_{t_j}^{(\pi,p-1)}$ and $\tilde{Z}_{t_j}^{(\pi,p-1)}$. We may associate to the sequence $(Y^{(\pi),p},Z^{(\pi),p},\tilde{Z}^{(\pi),p})_{p\in\mathbb{N}^*}$ a sequence of square-integrable orthogonal martingales $(M^{(\pi),p})_{p\in\mathbb{N}^*}$ defined by $M^{(\pi),0}\equiv 0$ and for $p\in\mathbb{N}$ by $M^{(\pi),p}=\{M_{t_i}^{(\pi),p},t_i\in\pi\}$ with

$$\Delta M_{t_i}^{(\pi),p}=Y_{t_{i+1}}^{(\pi),p}-\mathbb{E}_{t_i}\left[Y_{t_{i+1}}^{(\pi),p}\right]-Z_{t_i}^{(\pi),p}\Delta W_{t_i}^{(\pi)}\ -\ \left\{\tilde{Z}_{t_i}^{(\pi),p}(\Delta X_{t_i}^{(\pi)})-\mathbb{E}_{t_i}\left[\tilde{Z}_{t_i}^{(\pi),p}(\Delta X_{t_i}^{(\pi)})\right]\right\}.$$

We also note that we have

$$
\begin{aligned}
(3.17)\qquad m_{t_i}^{(\pi)} &:= \mathbb{E}_{t_i}\left[F^{(\pi)} + \sum_{t_j\in\pi} f^{(\pi)}(t_j, Y_{t_j}^{(\pi),p}, Z_{t_j}^{(\pi),p}, \tilde{Z}_{t_j}^{(\pi),p})\Delta\right] \\
&= Y_0^{(\pi),p+1} + \sum_{t_j\in\pi, j<i} Z_{t_j}^{(\pi),p+1}\Delta W_{t_j}^{(\pi)} + \sum_{t_j\in\pi,j<i}\left\{\tilde{Z}_{t_j}^{(\pi),p+1}(\Delta X_{t_j}^{(\pi)}) - \mathbb{E}_{t_i}\left[\tilde{Z}_{t_j}^{(\pi),p+1}(\Delta X_{t_j}^{(\pi)})\right]\right\} + M_{t_i}^{(\pi),p+1}.
\end{aligned}
$$

It is well-known that, as p tends to infinity, the Picard sequence $(Y^{(\pi,p)}, Z^{(\pi,p)}, \tilde{Z}^{(\pi,p)}, M^{(\pi,p)})$ converges to $(Y^{(\pi)}, Z^{(\pi)}, \tilde{Z}^{(\pi)}, M^{(\pi)})$. In particular, it follows from Theorem 3.4 (by reasoning analogously as in Corollary 10 in Briand *et al.* (2002)) that for some $n_0 \in \mathbb{N}$ it holds

$$
\begin{aligned}
(3.18)\qquad &\sup_{\pi_N: N\geq n_0} \mathbb{E}\Bigg[\sup_{t_i\in\pi_N} |Y_{t_i}^{(\pi)} - Y_{t_i}^{(\pi),p}|^2 + \sum_{t_i\in\pi_N}\{|Z_{t_i}^{(\pi)} - Z_{t_i}^{(\pi),p}|^2\Delta + \Delta(M^{(\pi)} - M^{(\pi),p})_{t_i}^2\} \\
&+ \sum_{t_i\in\pi_N}\left\{\tilde{Z}_{t_i}^{(\pi)}(\Delta X_{t_i}^{(\pi)}) - \tilde{Z}_{t_i}^{(\pi),p}(\Delta X_{t_i}^{(\pi)}) - \mathbb{E}_{t_i}\left[\tilde{Z}_{t_i}^{(\pi)}(\Delta X_{t_i}^{(\pi)}) - \tilde{Z}_{t_i}^{(\pi),p}(\Delta X_{t_i}^{(\pi)})\right]\right\}^2\Bigg] \to 0 \quad \text{as } p\to\infty.
\end{aligned}
$$

4. Convergence

With the results concerning the convergence of the approximating random walks and the properties of the discrete time BSDEs in hand, we turn next to the question of weak convergence of BSΔEs to the limiting BSDE as the mesh size tends to zero. Let $Y_t^{(\pi)} = Y_{t_i}^{(\pi)}$ for $t_i \leq t < t_{i+1}$ and define $(Z_t^{(\pi)}, \tilde{Z}_t^{(\pi)}, M_t^{(\pi)})$ similarly.

Theorem 4.1. *Let (π) be a sequence of partitions π with the mesh Δ tending to zero. If $F^{(\pi)}$ converges to F in L^2, then $Y^{(\pi)} \xrightarrow{\mathcal{L}} Y$ and in particular*

$$Y_0^{(\pi)} \to Y_0.$$

Moreover, with d_S denoting the Skorokhod metric, we have

$$\mathbb{E}[d_S^2(Y^{(\pi)}, Y)] \to 0.$$

Proof. The idea, inspired by Briand *et al.* (2001,2002), is to reduce the question of weak convergence of the solutions of the BSΔEs to the solution of BSDE to that of the Picard sequences by using the fact that both the solutions of the BSDE and of the BSΔEs are equal to limits of Picard sequences.

Define the sequence $(Y^{\infty,p}, Z^{\infty,p}, \tilde{Z}^{\infty,p})_{p\in\mathbb{N}\cup\{0\}}$ recursively by $(Y^{\infty,0}, Z^{\infty,0}, \tilde{Z}^{\infty,0}) = (0,0,0)$ and

$$Y_t^{\infty,p+1} := F + \int_t^T f(s, Y_s^{\infty,p}, Z_s^{\infty,p}, \tilde{Z}_s^{\infty,p})\mathrm{d}s - \int_t^T Z_s^{\infty,p+1}\mathrm{d}W_s - \int_{(t,T]\times\mathbb{R}^{d_2}\setminus\{0\}} \tilde{Z}_s^{\infty,p+1}(x)\tilde{N}(\mathrm{d}s\times\mathrm{d}x)$$

for $p \in \mathbb{N}\cup\{0\}$, where $(Z^{\infty,p+1}, \tilde{Z}^{\infty,p+1})$ are the unique coefficients in the martingale representation of the square-integrable martingale $N^p = \{N_t^p, t\in[0,T]\}$:

$$
\begin{aligned}
N_t^p &:= \mathbb{E}\left[F + \int_0^T f(s, Y_{s-}^{\infty,p}, Z_s^{\infty,p}, \tilde{Z}_s^{\infty,p})\mathrm{d}s\middle|\mathcal{F}_t\right] - \mathbb{E}\left[F + \int_0^T f(s, Y_{s-}^{\infty,p}, Z_s^{\infty,p}, \tilde{Z}_s^{\infty,p})\mathrm{d}s\right] \\
(4.1)\qquad &= \int_0^t Z_s^{\infty,p+1}\mathrm{d}W_s + \int_0^t \tilde{Z}_s^{\infty,p+1}\tilde{N}(\mathrm{d}s\times\mathrm{d}x).
\end{aligned}
$$

Furthermore, recall that we denote by $(Y^{(\pi),p}, Z^{(\pi),p}, \tilde{Z}^{(\pi),p}, M^{(\pi),p})_{p\in\mathbb{N}\cup\{0\}}$ the Picard sequences corresponding to the BSΔEs defined on the grid π. In the remainder of the proof we will deploy the continuous-time extensions of $(Y^{(\pi),p}, Z^{(\pi),p}, \tilde{Z}^{(\pi),p}, M^{(\pi),p})_{p\in\mathbb{N}\cup\{0\}}$ defined by taking paths to be piecewise constant; we denote these extensions also by $(Y^{(\pi),p}, Z^{(\pi),p}, \tilde{Z}^{(\pi),p}, M^{(\pi,p)})_{p\in\mathbb{N}\cup\{0\}}$.

In view of the decomposition

$$Y^{(\pi)} - Y = Y^{(\pi)} - Y^{(\pi),p} + Y^{(\pi),p} - Y^{\infty,p} + Y^{\infty,p} - Y$$

and the fact that $Y^{\infty,p}$ converges to Y and $Y^{(\pi),p}$ to $Y^{(\pi)}$ in $\mathcal{S}^2$-norm as $p \to \infty$ (see Tang & Li (1994) and (3.18) above, respectively), we have that the convergence of $Y^{(\pi)}$ to Y in the Skorokhod metric in L^2 will follow once we show that $Y^{(\pi),p}$ converges to $Y^{\infty,p}$ in the latter sense, for any fixed p:

Lemma 4.1. *Let* $p \in \mathbb{N}$. *Then we have*

$$\mathbb{E}[d_S^2(Y^{(\pi),p}, Y^{\infty,p})] \to 0, \quad \text{as } \Delta \to 0. \tag{4.2}$$

To establish Lemma 4.1 we first provide a proof in the case of 'smooth' drivers and terminal conditions (in Section 4.1), and use subsequently density arguments to show that the convergence carries over to the general case (in Section 4.2). □

4.1. **The smooth case.** In order to show convergence we first restrict to the case that the terminal conditions and driver functions are bounded infinitely (Fréchet-)differentiable functionals, in the following sense

Definition 4.2. Let $\mathcal{H}$ be a Hilbert space. (i) A function $f : \mathcal{H} \to \mathbb{R}$ is differentiable if it is Fréchet-differentiable in every $l \in \mathcal{H}$, *i.e.*, there exists a bounded linear operator $A_l : \mathcal{H} \to \mathbb{R}$ satisfying

$$\lim_{h\to 0} \frac{f(h+l) - f(l) - A_l(h)}{|h|} = 0.$$

We set $D^{(1)}f(l) = A_l$.

(ii) A function $f : \mathcal{H} \to \mathbb{R}$ is k-times differentiable in l, $k \in \mathbb{N}$, if there exists a bounded k-linear map $A_l : \mathcal{H}^k \to \mathbb{R}$ such that for every $h_1, \ldots, h_{k-1} \in \mathcal{H}$

$$\lim_{h_k\to 0} \frac{D^{(k-1)}f(h_k + l)(h_1 \ldots, h_{k-1}) - D^{(k-1)}f(l)(h_1, \ldots, h_{k-1}) - A_l(h_1, \ldots, h_k)}{|h_k|} = 0.$$

(iii) A function $f : \mathcal{H} \to \mathbb{R}$ is element of $C_b^\infty(\mathcal{H})$ if all its higher derivatives are bounded, *i.e.*, for every $k \in \mathbb{N}$ there exists $\bar{C}_k > 0$ such that for all $h_i \in \mathcal{H}$

$$\sup_{l\in\mathcal{H}} |D^{(k)}H(l)(h_1, \ldots, h_k)| \leq \bar{C}_k \prod_{i=1}^{k} |h_i|.$$

Given these definitions the formulation of the smoothness condition that is in force throughout this subsection is as follows:

Assumption 2. (i) For some $k \in \mathbb{N}$ and $H \in C_b^\infty(\mathbb{R}^{2k})$ the terminal conditions F and $F^{(\pi)}$ are given by

$$\begin{aligned} F^{(\pi)} &= H(W_{s_1}^{(\pi)}, \ldots, W_{s_k}^{(\pi)}, X_{s_1}^{(\pi)}, \ldots, X_{s_k}^{(\pi)}), \\ F &= H(W_{s_1}, \ldots, W_{s_k}, X_{s_1}, \ldots, X_{s_k}), \end{aligned} \qquad \text{for some } s_1, \ldots, s_k \in [0, T].$$

Moreover, $F^{(\pi)}$ converges to F as $\Delta \to 0$ in $L^2(\mathbb{P})$.

(ii) The drivers f and $f^{(\pi)}$ satisfy $f(t, \cdot) \in C_b^\infty(\mathbb{R} \times \mathbb{R}^{d_1} \times L^2(\nu(\mathrm{d}x), \mathcal{B}(\mathbb{R}^{d_2} \setminus \{0\})))$ and $f^{(\pi)}(t, \cdot) \in C_b^\infty(\mathbb{R} \times \mathbb{R}^{d_1} \times L^2(\nu^{(\pi)}(\mathrm{d}x), \mathcal{B}(\mathbb{R}^{d_2} \setminus \{0\})))$ where, for each k, the k-th derivative of $f^{(\pi)}$ is bounded uniformly in t and Δ, the mesh of π.

Under the smoothness conditions given in Assumption 2 the corresponding Picard sequences obey a number of properties that play an important role in the proof of Proposition 4.4:

Lemma 4.3. (i) *Let $p \in \mathbb{N}$. There exists a constant $\bar{K}_p > 0$ satisfying for all partitions π*

$$|\tilde{Z}_s^{(\pi),p}(x)| \leq \bar{K}_p |x| \text{ for all } x \in \mathbb{R}^{d_2},\ \ s \in [0,T], \tag{4.3}$$

where $\tilde{Z}_s^{(\pi),p}$ denotes a continuous version (in x). Furthermore, $Y^{(\pi),p}$ and $Z^{(\pi),p}$ are uniformly bounded over partitions π.

(ii) *$\tilde{Z}^{\infty,p}$ are uniformly Lipschitz-continuous in x,* i.e., *there exists a constant $K'_p > 0$ such that $|\tilde{Z}_t^{\infty,p}(x)| \leq K'_p|x|$ for all $x \in \mathbb{R}^{d_2}$ and every $t \in [0,T]$, where $\tilde{Z}_t^{\infty,p}$ denotes again a continuous version (in x).*

Given these properties, which proof is given in the Appendix, we show the convergence of $Y^{(\pi)}$ as stated in Lemma 4.1 and in addition the convergence in mean-square of the triplet $(Z^{(\pi)}, \tilde{Z}^{(\pi)}, M^{(\pi)})$ to $(Z, \tilde{Z}, 0)$:

Proposition 4.4. *For any $p \in \mathbb{N}$ we have as $\Delta \searrow 0$*

$$\mathbb{E}[d_S^2(Y^{(\pi),p}, Y^{\infty,p})] \to 0, \tag{4.4}$$

$$\mathbb{E}\left[\int_0^T \left\{|Z_s^{(\pi),p} - Z_s^{\infty,p}|^2 + \int_{\mathbb{R}^{d_2}\setminus\{0\}} |\tilde{Z}_s^{(\pi),p}(x) - \tilde{Z}_s^{\infty,p}(x)|^2 \nu(\mathrm{d}x)\right\} \mathrm{d}s + |M_T^{(\pi),p}|^2\right] \to 0. \tag{4.5}$$

Proof. The proof is based on an induction with respect to p. We note that the assertions are trivially satisfied for $p = 0$. Assuming that the assertion is satisfied for a certain p we show next that (4.4) and (4.5) are satisfied for $p + 1$.

Proof of (4.4) *with p replaced by $p+1$:* In view of the uniform Lipschitz continuity of the driver functions $f^{(\pi)}$ and since these are piecewise constant we have

$$\begin{aligned}
&\limsup_{\Delta\to 0} \sup_{t\in[0,T]} \Big| \sum_{j:t_j\in\pi\cap[0,t]} f^{(\pi)}(t_j, Y_{t_j}^{(\pi),p}, Z_{t_j}^{(\pi),p}, \tilde{Z}_{t_j}^{(\pi),p})\Delta - \int_0^t f(s, Y_s^{\infty,p}, Z_s^{\infty,p}, \tilde{Z}_s^{\infty,p})\mathrm{d}s\Big| \\
&\quad\leq \limsup_{\Delta\to 0} \int_0^T |f^{(\pi)}(s, Y_s^{(\pi),p}, Z_s^{(\pi),p}, \tilde{Z}_s^{(\pi),p}) - f^{(\pi)}(s, Y_s^{\infty,p}, Z_s^{\infty,p}, \tilde{Z}_s^{\infty,p})|\mathrm{d}s \\
&\quad\quad + \limsup_{\Delta\to 0} \int_0^T |f^{(\pi)}(s, Y_s^{\infty,p}, Z_s^{\infty,p}, \tilde{Z}_s^{\infty,p}) - f(s, Y_s^{\infty,p}, Z_s^{\infty,p}, \tilde{Z}_s^{\infty,p})|\mathrm{d}s \\
&\quad\leq \limsup_{\Delta\to 0} K\Big(\int_0^T \Big[|Y_s^{(\pi),p} - Y_s^{\infty,p}| + |Z_s^{(\pi),p} - Z_s^{\infty,p}| + \sqrt{\mathbb{E}_{\nu^{(\pi)}}([\tilde{Z}_s^{(\pi),p}(\xi) - \tilde{Z}_s^{\infty,p}(\xi)]^2)}\Big]\mathrm{d}s\Big),
\end{aligned} \tag{4.6}$$

where in the third line the limsup vanishes in view of Assumption 1 and Lemma 4.3(ii). Using Lemmas 2.1 and 4.3 we find for any $s \in [0,T]$

$$\lim_{\Delta\to 0} \mathbb{E}_{\nu^{(\pi)}}([\tilde{Z}_s^{(\pi),p}(\xi) - \tilde{Z}_s^{\infty,p}(\xi)]^2) = \lim_{\Delta\to 0} \int_{\mathbb{R}^{d_2}\setminus\{0\}} |\tilde{Z}_s^{(\pi),p}(x) - \tilde{Z}_s^{\infty,p}(x)|^2 \nu(\mathrm{d}x). \tag{4.7}$$

The induction assumption implies that the right-hand sides of (4.6) and (4.7) are equal to zero, where the limits are in L^2. By combining the convergence in $\mathcal{H}^2$-norm of the drivers and the extended convergence in Proposition 2.4 (see also the remark after Definition 2.2) we find that as $\Delta \searrow 0$

$$m_t^{(\pi)} := \mathbb{E}\left[F^{(\pi)} + \sum_{j:t_j\in\pi} f^{(\pi)}(t_j, Y_{t_j}^{(\pi),p}, Z_{t_j}^{(\pi),p}, \tilde{Z}_{t_j}^{(\pi),p})\Delta \middle| \mathcal{F}_t^{(\pi)}\right]$$

$$\longrightarrow m_t := \mathbb{E}\left[F + \int_0^T f(s, Y_s^{\infty,p}, Z_s^{\infty,p}, \tilde{Z}_s^{\infty,p})\mathrm{d}s \middle| \mathcal{F}_t\right],$$

and as a consequence also

$$Y_t^{(\pi),p+1} = m_t^{(\pi)} - \sum_{j:t_j\in\pi\cap[0,t]} f^{(\pi)}(t_j, Y_{t_j}^{(\pi),p}, Z_{t_j}^{(\pi),p}, \tilde{Z}_{t_j}^{(\pi),p})\Delta \to Y_t^{\infty,p+1} = m_t - \int_0^t f(s, Y_s^{\infty,p}, Z_s^{\infty,p}, \tilde{Z}_s^{\infty,p})\mathrm{d}s,$$

where the convergence is in probability in Skorokhod J_1-topology.

As $Y^{(\pi),p+1}$ is uniformly bounded over partitions π (Lemma 4.3(i)), we deduce that $\mathbb{E}[d_S^2(Y^{(\pi),p+1}, Y^{\infty,p+1})]$ tends to zero as $\Delta \to 0$, so that (4.4) holds with p replaced by $p+1$.

Proof of (4.5) *with p replaced by $p+1$:* The argument consists of a number of steps that are listed in the following auxiliary result:

Lemma 4.5. *The following convergence holds in the supremum norm in probability as $\Delta \to 0$:*

$$\int_0^{\cdot} |Z_s^{(\pi),p+1}|^2 \mathrm{d}s + \int_{[0,\cdot]\times\mathbb{R}^{d_2}\setminus\{0\}} |\tilde{Z}_s^{(\pi),p+1}(x)|^2 \nu(\mathrm{d}x)\mathrm{d}s + \langle M^{(\pi),p+1}\rangle. \longrightarrow \int_0^{\cdot} |Z_s^{\infty,p+1}|^2 \mathrm{d}s + \int_{[0,\cdot]\times\mathbb{R}^{d_2}\setminus\{0\}} |\tilde{Z}_s^{\infty,p+1}(x)|^2 \nu(\mathrm{d}x)\mathrm{d}s, \tag{4.8}$$

$$\int_0^{\cdot} Z_s^{(\pi),p+1} \mathrm{d}s \longrightarrow \int_0^{\cdot} Z_s^{\infty,p+1} \mathrm{d}s, \tag{4.9}$$

$$\int_{[0,\cdot]\times\mathbb{R}^{d_2}\setminus\{0\}} \tilde{Z}_s^{(\pi),p+1}(x)\bar{Z}_s(x)\nu(\mathrm{d}x)\mathrm{d}s \longrightarrow \int_{[0,\cdot]\times\mathbb{R}^{d_2}\setminus\{0\}} \tilde{Z}_s^{\infty,p+1}(x)\bar{Z}_s(x)\nu(\mathrm{d}x)\mathrm{d}s, \tag{4.10}$$

for any function $\bar{Z} : [0,T]\times\mathbb{R}^{d_2}\setminus\{0\} \to \mathbb{R}$ that is bounded, jointly continuous, and zero in an environment around zero. Furthermore, we have the following convergence in L^1:

$$\int_0^T \left\{ |Z_s^{(\pi),p+1} - Z_s^{\infty,p+1}|^2 + \int_{\mathbb{R}^{d_2}\setminus\{0\}} |\tilde{Z}_s^{(\pi),p+1}(x) - \tilde{Z}_s^{\infty,p+1}(\omega,x)|^2 \nu(\mathrm{d}x) \right\} \mathrm{d}s + |M_T^{(\pi),p+1}|^2 \longrightarrow 0, \tag{4.11}$$

It follows from (4.11) that (4.5) is valid with p replaced by $p+1$, and thus the proof of the proposition is complete. □

Proof of Lemma 4.5: The proof is given in four parts (corresponding to the different equations):

Proof of (4.8)*:* We show that the assertion follows from the convergence of the compensators of the martingales $L^{(\pi)}$, defined by

$$L_t^{(\pi)} = m_t^{(\pi)} - Y_0^{(\pi),p+1}, \tag{4.12}$$

to the compensator of the martingale $L = \{L_t = m_t - Y_0^{\infty,p+1}\}$, by verifying that the conditions of Theorem 2.5 are satisfied. We first show

$$\mathbb{E}[d_S^2(L^{(\pi)}, L)] \to 0, \quad \text{as } \Delta \to 0. \tag{4.13}$$

Since the processes $L^{(\pi)}$ converge to L in probability in the Skorokhod J_1-topology (see the end of the proof of (4.4)), the convergence in (4.13) follows by the lemma de la Vallée-Poussin and the fact that the collection $(L^{(\pi)})_\pi$ is uniformly bounded, as

$$\begin{aligned}\sup_\pi \|L_T^{(\pi)}\|_\infty \;\le\; & \sup_\pi \|F^{(\pi)}\|_\infty + \sup_\pi |Y_0^{(\pi)}| + T\sup_{\pi,t} |f^{(\pi)}(t,0,0,0)| + B^{(\pi)} \\ & + KT\Big(\sup_\pi \Big\| \sup_t |Y_t^{(\pi),p}| \Big\|_\infty + \sup_\pi \Big\| \sup_t |Z_t^{(\pi),p}| \Big\|_\infty \Big),\end{aligned} \tag{4.14}$$

where $||\cdot||_\infty$ denotes the essential-supremum norm and where by Jensen's inequality, the independence of increments and the conditions (2.4) and (2.5), we have

$$\begin{aligned} B^{(\pi)} &:= \sum_i \Delta K \bar{K}_{p+1} \sqrt{\int_{\mathbb{R}^{d_2}\setminus\{0\}} |x|^2 \nu^{(\pi)}(\mathrm{d}x)} \le \sqrt{T} K \bar{K}_{p+1} \sqrt{\sum_i \int_{\mathbb{R}^{d_2}\setminus\{0\}} |x|^2 \nu^{(\pi)}(\mathrm{d}x)\Delta} \\ &= \sqrt{T} K \bar{K}_{p+1} \sqrt{\mathbb{E}[|X_T^{(\pi)}|^2]} \to \sqrt{T} K \bar{K}_{p+1} \sqrt{\mathbb{E}[|X_T|^2]}. \end{aligned}$$

Since we have the convergence of $L_T^{(\pi)}$ to L_T in L^2 (from (4.13)) and the extended convergence $(L^{(\pi)}, \mathcal{F}^{(\pi)}) \to (L, \mathcal{F})$ (from (4.13), Proposition 2.4 and Remark 2.3) it follows from Theorem 2.5 that $(\langle L^{(\pi)}, L^{(\pi)}\rangle_t)$ converges to $(\langle L, L\rangle_t)$ in probability in the Skorokhod J_1-topology. By the orthogonality of the martingales $M^{(\pi)}$, $W^{(\pi)}$ and the point process induced by $X^{(\pi)}$ on the one hand and the orthogonality of W and $\tilde{N}$ on the other hand we find

$$\sum_{t_i \in \pi\setminus\{T\}\cap[0,\cdot]} \left\{ |Z_{t_i}^{(\pi),p+1}|^2 \Delta + \mathbb{E}_{t_i}[|\tilde{Z}_{t_i}^{(\pi),p+1}(\Delta X_{t_i}^{(\pi)})|^2] \right\} - \sum_{t_i \in \pi\setminus\{T\}\cap[0,\cdot]} \left| \mathbb{E}_{t_i}\left[\tilde{Z}_{t_i}^{(\pi),p+1}(\Delta X_{t_i}^{(\pi)}) \right] \right|^2 + \langle M^{(\pi),p+1}\rangle.$$

$$\longrightarrow \int_0^{\cdot} |Z_s^{\infty,p+1}|^2 \mathrm{d}s + \int_{[0,\cdot]\times\mathbb{R}^{d_2}\setminus\{0\}} |\tilde{Z}_s^{\infty,p+1}(x)|^2 \nu(\mathrm{d}x)\mathrm{d}s \tag{4.15}$$

in the supremum norm in probability. In this display we note that the second sum vanishes as Δ tends to zero. More specifically, in view of Lemma 4.3 and the condition in (2.3) we have

$$\sum_{t_i \in \pi\setminus\{T\}} \left| \mathbb{E}_{t_i}\left[\tilde{Z}_{t_i}^{(\pi),p+1}(\Delta X_{t_i}^{(\pi)}) \right] \right|^2 \leq \bar{K}_{p+1}^2 \sum_{t_i \in \pi\setminus\{T\}} \left| \mathbb{E}_{t_i}\left[|\Delta X_{t_i}^{(\pi)}| \right] \right|^2 \to 0 \tag{4.16}$$

as Δ tends to zero, where we used that $\mathbb{E}_{t_i}[|\Delta X_{t_i}^{(\pi)}|] = \mathbb{E}[|\Delta X_{t_i}^{(\pi)}|]$ by the independence of the increments of $X^{(\pi)}$. The assertion in (4.8) follows by combining (4.15) and (4.16) with the fact that $Z_s^{(\pi),p+1}$ is piecewise constant as function of s and with Lemma 2.1(ii), which is applicable as $(\tilde{Z}^{(\pi),p+1})_\pi$ is uniformly Lipschitz-continuous.

Proof of (4.9)*:* It follows from Corollary 2.6(i) and the representation (3.17) of the square-integrable martingale $L^{(\pi)}$ defined in (4.12) that as $\Delta \to 0$

$$\langle W^{(\pi)}, L^{(\pi)}\rangle_{\cdot} = \sum_{i=0}^{\lfloor \cdot N\rfloor - 1} |Z_{t_i}^{(\pi),p+1}|\Delta \to \langle W, L\rangle_{\cdot} = \int_0^{\cdot} Z_s^{\infty,p+1}\mathrm{d}s,$$

in the supremeum norm in probability which implies the assertion in (4.9).

Proof of (4.10)*:* We conclude from Corollary 2.6(ii) and the representation of the martingale $L^{(\pi)}$

$$\lim_{\Delta\to 0} \sum_{t_i \in \pi\setminus\{T\}\cap[0,\cdot]} \left\{ \mathbb{E}_{t_i}[\tilde{Z}_{t_i}^{(\pi),p+1}(\Delta X_{t_i}^{(\pi)})\bar{Z}_{t_i}(\Delta X_{t_i}^{(\pi)})] - \mathbb{E}_{t_i}[\tilde{Z}_{t_i}^{(\pi),p+1}(\Delta X_{t_i}^{(\pi)})]\mathbb{E}_{t_i}[\bar{Z}_{t_i}(\Delta X_{t_i}^{(\pi)})] \right\}$$
$$\to \int_{[0,\cdot]\times\mathbb{R}^{d_2}\setminus\{0\}} \tilde{Z}_s^{\infty,p+1}(x)\bar{Z}_s(x)\nu(\mathrm{d}x)\mathrm{d}s,$$

in probability in the Skorokhod J_1-topology. As the limit is continuous, this convergence also holds in the supremum norm. Moreover, as $\bar{Z}$ is bounded, continuous, and zero in an environment around zero, it is clear that there exists $\hat{K} > 0$ such that $|\bar{Z}_s(x)| \leq \hat{K}|x|$. It follows then from Lemma 2.1(ii) that we have the convergence in (4.10) in the supremum norm in probability.

Proof of (4.11)*:* Next let us switch to a subsequence and assume that all previous convergence results derived in the proofs of (4.8)–(4.10) hold for a.e. $\omega \in \Omega$. Fix such an $\omega \in \Omega$. By Lemma 4.3 there exists constants $\bar{K}_{p+1} > 0$ such that

$$\sup_\pi \int_{[0,T]\times\mathbb{R}^{d_2}\setminus\{0\}} \left| \tilde{Z}_s^{(\pi),p+1}(\omega, x) \right|^2 \nu(\mathrm{d}x)\mathrm{d}s \leq \bar{K}_{p+1}^2 \int_{[0,T]\times\mathbb{R}^{d_2}\setminus\{0\}} |x|^2\nu(\mathrm{d}x)\mathrm{d}s = T\bar{K}_{p+1}^2 \int_{\mathbb{R}^{d_2}\setminus\{0\}} |x|^2\nu(\mathrm{d}x) < \infty.$$

Hence, $\tilde{Z}_s^{(\pi),p+1}(\omega, x)$ is uniformly bounded in $L^2(\nu(\mathrm{d}x)\times\mathrm{d}s)$. By switching to a subsequence, we may assume that $\tilde{Z}_s^{(\pi),p+1}(\omega, x)$ converges weakly in $L^2(\nu(\mathrm{d}x)\times\mathrm{d}s, \mathcal{B}(\mathbb{R}^{d_2}\setminus\{0\})\otimes\mathcal{B}([0,T]))$ to a limiting function. From (4.10) it follows that this limit is equal to $\tilde{Z}_{\cdot}^{\infty,p+1}(\omega,\cdot)$. Furthermore, by (4.9) we also know that for a.e. ω we have that $Z_{\cdot}^{(\pi),p+1}(\omega)$ converges weakly to $Z_{\cdot}^{\infty,p+1}(\omega)$ in $L^2_{d_1}(\mathrm{d}s)$. We also have that the pairs $(Z^{n,p+1}(\omega), \tilde{Z}^{n,p+1}(\omega))$

converge weakly $(Z^{\infty,p+1}(\omega), \tilde{Z}^{\infty,p+1}(\omega))$ in $L^2_{d_1}(ds) \times L^2(\nu(dx) \times ds)$ equipped with the inner product

$$\langle (z^1, \tilde{z}^1), (z^2, \tilde{z}^2) \rangle_* = \int_0^T z^1_s z^2_s \mathrm{d}s + \int_{[0,T]\times\mathbb{R}^{d_2}\setminus\{0\}} \tilde{z}^1_s(x)\tilde{z}^2_s(x)\nu(\mathrm{d}x)\mathrm{d}s.$$

Denoting by $\|\cdot\|_*$ the norm associated to this inner product we have by (4.8)

$$\limsup_{\Delta\to 0} \|I_\Delta\|^2_* := \limsup_{\Delta\to 0} \left\|(Z^{(\pi),p+1}(\omega), \tilde{Z}^{(\pi),p+1}(\omega))\right\|^2_* \leq \left\|(Z^{\infty,p+1}(\omega), \tilde{Z}^{\infty,p+1}(\omega))\right\|^2_* =: ||I||^2_*.$$

Therefore, we have

$$0 \leq \limsup_{\Delta\to 0} \langle I_\Delta - I, I_\Delta - I\rangle = \limsup_{\Delta\to 0} (\langle I_\Delta, I_\Delta\rangle - 2\langle I, I_\Delta\rangle + \langle I, I\rangle) \leq \langle I, I\rangle - 2\langle I, I\rangle + \langle I, I\rangle = 0.$$

Hence, all inequalities must be equalities and we get that

$$\int_0^T |Z^{(\pi),p+1}_s(\omega) - Z^{\infty,p+1}_s(\omega)|^2 \mathrm{d}s + \int_{[0,T]\times\mathbb{R}^{d_2}\setminus\{0\}} |\tilde{Z}^{(\pi),p+1}_s(\omega,x) - \tilde{Z}^{\infty,p+1}_s(\omega,x)|^2 \nu(\mathrm{d}x)\mathrm{d}s \to 0 \text{ as } \Delta \to 0.$$

By (4.8) it follows that also $|M^{(\pi),p+1}_t(\omega)|^2$ converges to zero as well. Therefore, for a.e. ω, for any subsequence on the left-hand side in (4.11) we can find a subsubsequence converging to zero. Thus, we must have convergence in probability in (4.11). We note that, for any p, $(M^{(\pi),p}_T)^2$ is uniformly integrable over partitions π, since we have the bound

$$\sup_\pi \mathbb{E}[\langle M^{(\pi),p}, M^{(\pi),p}\rangle^2_T] \leq \sup_\pi \mathbb{E}[\langle L^{(\pi),p}, L^{(\pi),p}\rangle^2_T] \leq \sup_\pi \bar{C}\|L^{(\pi),p}_T\|^4_\infty < \infty, \text{ for a } \bar{C} > 0,$$

which follow by the definitions of $M^{(\pi),p}$ and $L^{(\pi),p}$, the BDG and Doob inequalities, and the fact that $L^{(\pi)}_T$ is bounded uniformly in π (by (4.14)).

That the convergence in (4.11) also holds true in L^1 may be seen from another application of the Lemma de la Vallée-Poussin, which is applicable as the integral on the left-hand side is bounded uniformly in π (by Lemma 4.3), in combination with the uniform integrability of $(M^{(\pi),p+1}_T)^2$. □

4.2. **Density argument.** We complete the proof of Theorem 4.1 by combining Proposition 4.4 with a density argument.

Proof of Theorem 4.1. Let $k \in \mathbb{N}$ be arbitrary. By standard density results we can find functions $H_k \in C^\infty_b(\mathbb{R}^{2k})$ and uniformly K-Lipschitz-continuous functions f^k and $f^{(\pi),k}$ such that $f^k(t,\cdot) \in C^{0,\infty}_b(\mathbb{R}\times\mathbb{R}^{d_1}\times L^2(\nu(\mathrm{d}x))$, and $f^{(\pi),k}(t,\cdot) \in C^{0,\infty}_b(\mathbb{R}\times\mathbb{R}^{d_1}\times L^2(\nu^{(\pi)}(\mathrm{d}x))$ converging to f^k with

$$\begin{aligned} & T \sup_{t,y,z,\tilde{z}} (|f(t,y,z,\tilde{z}) - f^k(t,y,z,\tilde{z})| + |f^{(\pi),k}(t,y,z,\tilde{z}) - f^{(\pi)}(t,y,z,\tilde{z})|) \\ & + \quad \mathbb{E}[|F - H_k(W_{s_1}, X_{s_1}, \ldots, W_{s_k}, X_{s_k})|^2] \leq \frac{1}{k}. \end{aligned} \tag{4.17}$$

The triangle inequality for the Skorokhod metric d_S and the inequality $(x+y+z)^2 \leq 3(x^2+y^2+z^2)$ imply

$$\begin{aligned} & \mathbb{E}[d^2_S(Y^{(\pi)}, Y)] \\ \leq \quad & 3\mathbb{E}[d^2_S(Y^{(\pi)}, \tilde{Y}^{(\pi)})] + 3\mathbb{E}[d^2_S(\tilde{Y}^{(\pi)}, \tilde{Y})] + 3\mathbb{E}[d^2_S(\tilde{Y}, Y)] := 3d^2_1(k,\pi) + 3d^2_2(k,\pi) + 3d^2_3(k), \end{aligned} \tag{4.18}$$

where $\tilde{Y}$ and $\tilde{Y}^{(\pi)}$ denote the solutions of the BSDE and BSΔE with terminal conditions $\tilde{F} = H_k(W,X)$, $\tilde{F}^{(\pi)} = H_k(W^{(\pi)}, X^{(\pi)})$ and drivers $\tilde{f} = f^k$ and $\tilde{f}^{(\pi)} = f^{(\pi),k}$, respectively.

We first estimate the distances between $Y^{(\pi)}$ and $\tilde{Y}^{(\pi)}$ and between Y and $\tilde{Y}$ in the supremum norm. By applying Theorem 3.4 and Remark 3.5 we see that the following bounds hold true:

$$\mathbf{d}_1^2(k,\pi) := \mathbb{E}\left[\sup_{t\in[0,T]} |Y_t^{(\pi)} - \tilde{Y}_t^{(\pi)}|^2\right] \leq \bar{C}\mathbb{E}\left[|F^{(\pi)} - \tilde{F}^{(\pi)}|^2 + \int_0^T |\delta f^{(\pi),k}(s, Z_s^{(\pi)}, \tilde{Z}_s^{(\pi)})|^2 \mathrm{d}s\right],$$

$$\mathbf{d}_3^2(k) := \mathbb{E}\left[\sup_{t\in[0,T]} |Y_t - \tilde{Y}_t|^2\right] \leq \bar{c}\,\mathbb{E}\left[|F - \tilde{F}|^2 + \int_0^T |\delta f^k(s, Z_s, \tilde{Z}_s)|^2 \mathrm{d}s\right],$$

where we denote $\delta f^k := f - f^k$ and $\delta f^{(\pi),k} := f^{(\pi)} - f^{(\pi),k}$. By deploying the bound (4.17) and Proposition 4.4 and using that $F^{(\pi)} \in L^2(\mathcal{F}^{(\pi)})$ converges to $F \in L^2(\mathcal{F})$ in L^2, we find

$$\limsup_{\Delta\to 0} \mathbf{d}_1^2(k,\pi) \leq \bar{C}\mathbb{E}[|F - \tilde{F}|^2] + \frac{\bar{C}}{k} \leq \frac{2\bar{C}}{k}, \quad \limsup_{\Delta\to 0} d_2^2(k,\pi) = 0, \quad \mathbf{d}_3^2(k) \leq \frac{2\bar{c}}{k}. \tag{4.19}$$

Since the Skorokhod metric is dominated by the supremum norm (see *e.g.* Eqn. VI.1.26 in Jacod & Shiryaev (2003)) we conclude from (4.18) and (4.19) that $\limsup_{\Delta\to 0} \mathbb{E}[d_S^2(Y^{(\pi)}, Y)] \leq 6(\bar{C} + \bar{c})/k$ for arbitrary k. Hence the proof is complete. □

5. Example

By way of illustration we specify in this section a sequence of approximating BSΔEs driven by a discrete-valued approximating sequence $(X^{(\pi)})_\pi$. We consider the BSDE

$$Y_t = F + \int_t^T f(s, Y_s, \tilde{Z}_s)\mathrm{d}s - \int_{(t,T]\times\mathbb{R}\backslash\{0\}} \tilde{Z}_s(x)\tilde{N}(\mathrm{d}s \times \mathrm{d}x), \quad t \in [0,T], \tag{5.1}$$

which is driven by the compensated Poisson random measure $\tilde{N}$ associated to a square-integrable zero-mean real-valued Lévy process X ($d_2 = 1$). Here, $f : [0,T] \times \mathbb{R} \times L^2(\nu(\mathrm{d}x), \mathcal{B}(\mathbb{R}\backslash\{0\})) \to \mathbb{R}$ is the driver function. As usual we assume that f is continuous as function of t, and uniformly Lipschitz continuous (as in (1.3) without the Brownian term). We consider final conditions F of the form

$$F = H(X_{s_0}, \ldots, X_{s_D}) \quad \text{with } s_i - s_{i-1} = \Delta_0,\ s_0 = 0,\ s_D = T, \tag{5.2}$$

for some Lipschitz function $H : \mathbb{R}^{D+1} \to \mathbb{R}$ (with Lipschitz constant K say). We also suppose that the Lévy measure ν of X has Blumenthal-Getoor index[‡] $\beta < 2$ and admits a strictly positive density g_ν on $\mathbb{R}\backslash\{0\}$ satisfying the integrability condition

$$\int_{\{|x|>1\}} g_\nu(x)|x|^{2+\epsilon}\mathrm{d}x < \infty, \quad \text{for some } \epsilon > 0. \tag{5.3}$$

We mention that, under the integrability condition (5.3), $\mathbb{E}[|X_t|^{2+\epsilon}]$ is finite for any t (see Sato (1999)), so that in particular F is square-integrable.

For the ease of presentation we consider BSΔEs defined on grids that are refinements of $\pi_0 = \{s_0, s_1, \ldots, s_D\}$. We next specify the final value $F^{(\pi)}$, the driver $f^{(\pi)}$ and the random walk $X^{(\pi)}$ and denote the corresponding BSΔE on the uniform time-grid $\pi \subset \pi_0$ by

$$Y_{t_i}^{(\pi)} = F^{(\pi)} + \sum_{t_j: t_i \leq t_j < T} f^{(\pi)}(t_j, Y_{t_j}^{(\pi)}, \tilde{Z}_{t_j}^{(\pi)})\Delta - \sum_{t_j: t_i \leq t_j < T} \left\{\tilde{Z}_{t_j}^{(\pi)}(\Delta X_{t_j}^{(\pi)}) - \mathbb{E}_{t_j}[\tilde{Z}_{t_j}^{(\pi)}(\Delta X_{t_j}^{(\pi)})]\right\}. \tag{5.4}$$

Define the spatial mesh size h by $h^2 = 3\Delta\Sigma^2$, where $\Sigma^2 = \int_{\mathbb{R}\backslash\{0\}} x^2\nu(\mathrm{d}x)$ and, as before, Δ denotes the mesh of the partition π. Then we have $\nu(\{x : |x| > h\})\,\Delta < 1/3$ as

$$\nu(\{x : |x| > h\})\,\Delta = (3\Sigma^2)^{-1}h^2\nu(\{x : |x| > h\}) < (3\Sigma^2)^{-1}\int_{\{|x|>h\}} x^2\nu(\mathrm{d}x) < \frac{1}{3}.$$

[‡]The Blumenthal-Getoor index β of X is $\beta = \inf\{p > 0 : \int_{\{|x|<1\}} |x|^p\nu(\mathrm{d}x) < \infty\}$.

We define the distribution of the increments of $X^{(\pi)}$ in terms of the averages of the Lévy measure ν over certain sets:

$$\alpha(A) := \frac{1}{\nu(A)} \int_A x\nu(\mathrm{d}x), \quad A \in \mathcal{B}(\mathbb{R}), \nu(A) > 0.$$

If $\alpha([-h,h]^c) \geq 0$ then set $h_- := h$ and $h_+ := \inf\{u \geq h : \alpha([-h,u]^c) = 0\}$ and similarly if $\alpha([-h,h]^c) < 0$ then set $h_+ := h$ and $h_- := \inf\{\ell \geq h : \alpha([-\ell,h]^c) = 0\}$.

Setting $B_{i+1} := (h_+(i), h_+(i+1)]$ and $B_{-i-1} := [-h_-(i+1), -h_-(i))$ for $i \in \mathbb{N}$ for some strictly increasing sequences $(h_\pm(i))_{i\in\mathbb{N}}$ with $h_+(1) = h_+$ and $h_-(1) = h_-$ and mesh size going to zero, we define for any integer $|i| \geq 2$

$$\mathbb{P}(X_{t_1}^{(\pi)} = x_i) = p_i \quad \text{with} \quad p_i = \Delta\, \nu(B_i), \quad x_i = \frac{1}{\nu(B_i)} \int_{B_i} x\nu(\mathrm{d}x).$$

Note that with this choice we have

$$\mathbb{P}\left(X_{t_1}^{(\pi)} \notin [h_-, h_+]\right) = \sum_{i:|i|\geq 2} p_i = \Delta\, \nu(\{x : x \notin [h_-, h_+]\}),$$

$$\mathbb{E}\left[X_{t_1}^{(\pi)} I_{\{X_{t_1}^{(\pi)} \notin [h_-,h_+]\}}\right] = \sum_{i:|i|\geq 2} x_i p_i = 0 = \alpha([h_-, h_+]^c).$$

The description of the distribution of $X_{t_1}^{(\pi)}$ is completed by setting $\mathbb{P}(X_{t_1}^{(\pi)} = \pm h) = p_{\pm 1}$ and $P(X_{t_1}^{(\pi)} = 0) = p_0$, where p_0 and $p_{\pm 1}$ are chosen so as to satisfy the conditions of unit mass and zero mean and to match the instantaneous variance:

$$\sum_{i:|i|\geq 0} p_i = 1, \quad \sum_{i:|i|\geq 0} x_i p_i = 0, \quad \sum_{i:|i|\geq 0} p_i (x_i)^2 = \Delta \int_{\mathbb{R}\setminus\{0\}} x^2 \nu(\mathrm{d}x),$$

or equivalently, $p_{-1} + p_0 + p_1 = 1 - \nu(\{x : x \notin [h_-, h_+]\})\,\Delta$, and

$$(p_1 - p_{-1})h = 0, \quad (p_1 + p_{-1})h^2 = \Delta \int_{\{x\in[h_-,h_+]\}} x^2 \nu(\mathrm{d}x) + \Delta\, V(h_-, h_+)$$

$$\Rightarrow p_{-1} = p_1 = \frac{1}{6\Sigma^2} \left\{S(h_-, h_+) + V(h_-, h_+)\right\} \leq \frac{1}{6}, \quad p_0 > \frac{1}{3},$$

with

$$S(h_-, h_+) := \int_{[h_-,h_+]} x^2 \nu(\mathrm{d}x) \;\text{ and }\; V(h_-, h_+) := \int_{\{x\notin[h_-,h_+]\}} x^2 \nu(\mathrm{d}x) - \sum_{i:|i|\geq 2} \frac{1}{\nu(B_i)} \left\{\int_{B_i} x\nu(\mathrm{d}x)\right\}^2,$$

which is non-negative as a consequence of the Cauchy-Schwarz inequality. In particular, we see that the zero-jump-condition (2.7) is satisfied. We note that the approximating processes $(X^{(\pi)})_\pi$ also satisfy conditions (2.3) and (2.4), since by construction $\mathbb{E}[|X_{t_1}^{(\pi)}|^2] = \Delta \int_{\mathbb{R}\setminus\{0\}} x^2 \nu(\mathrm{d}x)$, while the expectation of $\left|X_{t_1}^{(\pi)}\right|$ is $o(\sqrt{\Delta})$, since we have

$$\mathbb{E}\left[\left|X_{t_1}^{(\pi)}\right|\right] = (p_1 + p_{-1})h + \Delta \int_{[h_-,h_+]^c} |x|\nu(\mathrm{d}x),$$

where $p_1 + p_{-1}$ tends to zero when $\Delta \to 0$ and the second term is bounded by $c \cdot h^{2-\beta/2}$ (which is $o(\sqrt{\Delta})$ as $\Delta \to 0$ since $\beta < 2$ by assumption) with $c = \int |x|^{1+\beta/2}\nu(\mathrm{d}x)/(3\Sigma^2)$ (which is finite by definition of β and $\int x^2\nu(\mathrm{d}x) < \infty \Leftrightarrow \mathbb{E}[X_t^2] < \infty$).

Furthermore, it is easily checked that the sequence $(X^{(\pi)})_\pi$ also satisfies the conditions in (2.5) and (2.6). In particular, $X^{(\pi)} \xrightarrow{\mathcal{L}} X$ as $\Delta \to 0$, and on a suitably chosen probability space, $X^{(\pi)}$ converges to X in probability in the Skorokhod J_1-topology, and $X_T^{(\pi)}$ converges to X_T in L^2.

Next we define $F^{(\pi)} = H(X_{s_0}^{(\pi)}, \ldots, X_{s_D}^{(\pi)})$. By the Lipschitz continuity of H and the convergence of $X^{(\pi)}$ to X in $\mathcal{S}^2$ (by Doob's maximal inequality) it follows that also $F^{(\pi)}$ converges to F in L^2.

Finally, we specify $f^{(\pi)}$ in terms of f by $f^{(\pi)}(t,y,\tilde z) = f(t,y,Q\tilde z)$ with

$$(Q\tilde z)(x) = \begin{cases} \tilde z(x_i), & x\in B_i, i\neq 1, i\neq 0,\\ \frac{x}{\sqrt{2}h}(\tilde z(-h)+\tilde z(+h)), & x\in[h_-,h_+]\backslash\{0\},\\ 0, & x=0.\end{cases}$$

It is straightforward to verify that the drivers $f^{(\pi)}$ satisfy the required regularity conditions. In particular, the uniform Lipschitz-continuity of $f^{(\pi)}$ (as in (3.2)) can be derived as follows: for any $y_1, y_0\in\mathbb{R}$, $\tilde z_1,\tilde z_0\in L^2(\nu^{(\pi)},\mathcal{B}(\mathbb{R}\backslash\{0\}))$ the Lipschitz continuity of f implies

$$|f^{(\pi)}(t,y_1,\tilde z_1) - f^{(\pi)}(t,y_0,\tilde z_0)| \le K\left(|y_1-y_0| + \sqrt{I_Q}\right), \quad \text{with} \quad I_Q := \int_{\mathbb{R}\backslash\{0\}} |Q\tilde z_1(x) - Q\tilde z_0(x)|^2\nu(\mathrm{d}x).$$

Inserting the definitions of $Q\tilde z_0$ and $Q\tilde z_1$ shows

$$\begin{aligned} I_Q &= \sum_{i:|i|\ge 2} |\tilde z_1(x_i)-\tilde z_0(x_i)|^2\nu(B_i) + \frac{(\tilde z_1(h)-\tilde z_0(h)+\tilde z_1(-h)-\tilde z_0(-h))^2}{2}\frac{S(h_-,h_+)}{h^2}\\ &\le \sum_{i:|i|\ge 2}\frac{p_i}{\Delta}|\tilde z_1(x_i)-\tilde z_0(x_i)|^2 + (\tilde z_1(h)-\tilde z_0(h))^2\frac{p_1}{\Delta} + (\tilde z_1(-h)-\tilde z_0(-h))^2\frac{p_{-1}}{\Delta}\\ &= \int_{\mathbb{R}\backslash\{0\}} |\tilde z_1(x)-\tilde z_0(x)|^2\nu^{(\pi)}(\mathrm{d}x). \end{aligned}$$

We next move to the description of the solution of the BSΔE. Specifically, we have from Proposition 3.1 that the solution is given by

$$Y^{(\pi)}_{t_i} = v_i\left(X^{(\pi)}_{t_0},\ldots,X^{(\pi)}_{t_i}\right), \quad i=0,\ldots,N-1, \tag{5.5}$$

$$\tilde Z^{(\pi)}_{t_i}(x) = w_i\left(X^{(\pi)}_{t_0},\ldots,X^{(\pi)}_{t_i};x\right) - w_i\left(X^{(\pi)}_{t_0},\ldots,X^{(\pi)}_{t_i};0\right), \tag{5.6}$$

with $Y^{(\pi)}_{t_N} = F^{(\pi)}$, for certain functions $v_i:\mathbb{R}^i\to\mathbb{R}$ and $w_i:\mathbb{R}^{i+1}\to\mathbb{R}$ that are specified recursively as follows:

$$\begin{aligned} w_i(\underline{z}_{0,i};x) &= \mathbb{E}[Y^{(\pi)}_{t_{i+1}}|X^{(\pi)}_{t_0}=z_0,\ldots,X^{(\pi)}_{t_i}=z_i,\Delta X^{(\pi)}_{t_i}=x] = v_{i+1}(\underline{z}_{0,i},z_i+x), \quad x\in E^{(\pi)},\\ v_i(\underline{z}_{0,i}) &= f^{(\pi)}(t_i,v_i(\underline{z}_{0,i}),w_i(\underline{z}_{0,i};\cdot)-w_i(\underline{z}_{0,i};0)) + \sum_{x_j\in E^{(\pi)}} v_{i+1}(\underline{z}_{0,i},z_i+x_j)p_j, \end{aligned} \tag{5.7}$$

where $\underline{z}_{0,i} = (z_0,\ldots,z_i)$ and $E^{(\pi)} = \{x_i : i\in\mathbb{Z}\}$ denotes the support of the step-size distribution of $X^{(\pi)}$. While in general v_i is only implicitly defined by (5.7), the recursion in (5.7) has an explicit solution when the driver $f(t,y,\tilde z)$ (and thus $f^{(\pi)}(t,y,\tilde z)$) is constant as function of y. In this case, the solution Y of the BSDE is translation invariant in the sense that $Y(F+a) = a+Y(F)$ for $a\in\mathbb{R}$, where $Y(F)$ denotes the solution of the BSDE with final condition F (see Royer (2006)). By Theorem 4.1, the solution $Y^{(\pi)}$ of the BSΔE (5.4) specified in (5.5) converges to the solution Y of the BSDE (5.1) in L^2 in the Skorokhod J_1-topology.

Appendix A. Proof of Theorem 3.4

The structure of the proof is inspired by that of an analogous estimate derived in a Wiener setting in Proposition 7 in Briand *et al.* (2002).

Assuming without loss of generality $t_i = T$ and that $f^{(\pi)}(t,0,0,0) = 0$, and simplifying notation by dropping in the proof the superscripts (π) in the solution $(Y^{(\pi)}, Z^{(\pi)}, \tilde{Z}^{(\pi)}, M^{(\pi)})$, we have for $t_j, t_k \in \pi$ with $t_j < t_k$

$$\begin{aligned}\delta Y_{t_j} &= \delta Y_{t_k} + \sum_{r=j}^{k-1}\left(\delta f^{(\pi)}(t_r, Y^0_{t_r}, Z^0_{t_r}, \tilde{Z}^0_{t_r}) + f^{(\pi),1}(t_r, Y^0_{t_r}, Z^0_{t_r}, \tilde{Z}^0_{t_r}) - f^{(\pi),1}(t_r, Y^1_{t_r}, Z^1_{t_r}, \tilde{Z}^1_{t_r})\right)\Delta \\ &\quad - \sum_{r=j}^{k-1}\{\delta Z_{t_r}\Delta W^{(\pi)}_{t_r} + \delta\tilde{Z}_{t_r}(\Delta X^{(\pi)}_{t_r}) - E_{t_r}[\delta\tilde{Z}_{t_r}(\Delta X^{(\pi)}_{t_r})]\} - (\delta M_{t_k} - \delta M_{t_j}).\end{aligned} \tag{A.1}$$

Since the functions $f^{(\pi)}$ are K-Lipschitz and assuming without loss of generality $K > 1$, we have

$$|\delta Y_{t_j}| \le \mathbb{E}_{t_j}\left[|\delta Y_{t_k}| + K\sum_{r=j}^{k-1}\left\{\left|\delta f^{(\pi)}(t_r, Y^0_{t_r}, Z^0_{t_r}, \tilde{Z}^0_{t_r})\right| + |\delta Y_{t_r}| + |\delta Z_{t_r}| + \sqrt{\mathbb{E}_{\nu^{(\pi)}}([\delta\tilde{Z}_{t_r}(\xi)]^2)}\right\}\Delta\right]$$

and an application Doob's inequality yields for $t_m < t_k$ with $t_m, t_k \in \pi\backslash\{0\}$

$$\begin{aligned}\mathbb{E}\left[\sup_{m\le j<k}|\delta Y_{t_j}|^2\right] &\le 4\mathbb{E}\Bigg[\Bigg(|\delta Y_{t_k}| + K\sum_{r=m}^{k-1}\left\{\left|\delta f^{(\pi)}(t_r, Y^0_{t_r}, Z^0_{t_r}, \tilde{Z}^0_{t_r})\right|\right. \\ &\qquad \left.+|\delta Y_{t_r}| + |\delta Z_{t_r}| + \sqrt{\mathbb{E}_{\nu^{(\pi)}}([\delta\tilde{Z}_{t_r}(\xi)]^2)}\right\}\Delta\Bigg)^2\Bigg].\end{aligned} \tag{A.2}$$

Since $W^{(\pi)}$, $X^{(\pi)}$ and $\delta M^{(\pi)}$ are orthogonal martingales, we have

$$\begin{aligned}&\mathbb{E}\left[\left|\sum_{r=m}^{k-1}\left\{\delta Z_{t_r}\Delta W^{(\pi)}_{t_r} + \delta\tilde{Z}_{t_r}(\Delta X^{(\pi)}_{t_r}) - \mathbb{E}_{t_r}[\delta\tilde{Z}_{t_r}(\Delta X^{(\pi)}_{t_r})]\right\} + \delta M_{t_k} - \delta M_{t_m}\right|^2\right] \\ &\quad = \mathbb{E}\left[\sum_{r=m}^{k-1}|\delta Z_{t_r}|^2\Delta + \langle\delta\tilde{M}\rangle_{t_k} - \langle\delta\tilde{M}\rangle_{t_m} + \langle\delta M\rangle_{t_k} - \langle\delta M\rangle_{t_m}\right]\end{aligned} \tag{A.3}$$

where $\delta\tilde{M}$ is the martingale that is piecewise constant (outside the partition π) and has increment $\delta\tilde{M}_{t_{r+1}} - \delta\tilde{M}_{t_r}$ given by $\ell_{t_r}(\Delta X^{(\pi)}_{t_r}) := \delta\tilde{Z}_{t_r}(\Delta X^{(\pi)}_{t_r}) - \mathbb{E}_{t_r}[\delta\tilde{Z}_{t_r}(\Delta X^{(\pi)}_{t_r})]$, and $\langle\delta\tilde{M}\rangle$ and $\langle\delta M\rangle$ denote the predictable compensators of δM and $\delta\tilde{M}$, which are equal to

$$\langle\delta M\rangle_{t_i} = \sum_{t_j\le t_{i-1}}\mathbb{E}_{t_j}[|\Delta M^{(\pi)}_{t_j}|^2], \quad \langle\delta\tilde{M}\rangle_{t_i} = \sum_{t_j\le t_{i-1}}\mathbb{E}_{t_j}[|\Delta\ell_{t_j}(X^{(\pi)}_{t_j})|^2].$$

Using the fact that $f^{(\pi),1}$ is K-Lipschitz using (A.1) we obtain

$$\begin{aligned}&\left|\sum_{r=m}^{k-1}\left\{\delta Z_{t_r}\Delta W^{(\pi)}_{t_r} + \delta\tilde{Z}_{t_r}(\Delta X^{(\pi)}_{t_r}) - \mathbb{E}_{t_r}[\delta\tilde{Z}_{t_r}(\Delta X^{(\pi)}_{t_r})]\right\} + \delta M_{t_k} - \delta M_{t_m}\right| \\ &\quad\le |\delta Y_{t_k}| + K\sum_{r=m}^{k-1}\left\{\left|\delta f^{(\pi)}(t_r, Y^0_{t_r}, Z^0_{t_r}, \tilde{Z}^0_{t_r})\right| + |\delta Y_{t_r}|\right. \\ &\qquad\left.+|\delta Z_{t_r}| + \sqrt{\mathbb{E}_{\nu^{(\pi)}}([\delta\tilde{Z}_{t_r}(\xi)]^2)}\right\}\Delta + \sup_{m\le r<k}|\delta Y_{t_r}|.\end{aligned} \tag{A.4}$$

By combining the estimates in (A.2), (A.3), (A.4) we get

$$\mathbb{E}\left[\sup_{m\le r<k}|\delta Y_{t_r}|^2+\sum_{r=m}^{k-1}|\delta Z_{t_r}|^2\Delta+\langle\delta\tilde M\rangle_{t_k}-\langle\delta\tilde M\rangle_{t_m}+\langle\delta M\rangle_{t_k}-\langle\delta M\rangle_{t_m}\right]$$
$$\le 14\,\mathbb{E}\Bigg[\Bigg(|\delta Y_{t_k}|+K\sum_{r=m}^{k-1}\Bigg\{\Big|\delta f^{(\pi)}(t_r,Y^0_{t_r},Z^0_{t_r},\tilde Z^0_{t_r})\Big|$$
$$+|\delta Y_{t_r}|+|\delta Z_{t_r}|+\sqrt{\mathbb{E}_{\nu^{(\pi)}}([\delta\tilde Z_{t_r}(\xi)]^2)}\Bigg\}\Delta\Bigg)^2\Bigg].$$

An application of Hölder's inequality leads then to the estimate

$$\mathbb{E}\left[\sup_{m\le r<k}|\delta Y_{t_r}|^2+\sum_{r=m}^{k-1}|\delta Z_{t_r}|^2\Delta+\langle\delta\tilde M\rangle_{t_k}-\langle\delta\tilde M\rangle_{t_m}+\langle\delta M\rangle_{t_k}-\langle\delta M\rangle_{t_m}\right]$$
$$\le\quad C(t_k-t_m)\mathbb{E}\Bigg[\max_{m\le r<k}|\delta Y_{t_r}|^2+\sum_{r=m}^{k-1}\Bigg\{\Big|\delta f^{(\pi)}(t_r,Y^0_{t_r},Z^0_{t_r},\tilde Z^0_{t_r})\Big|^2+|\delta Z_{t_r}|^2+\mathbb{E}_{\nu^{(\pi)}}([\delta\tilde Z_{t_r}(\xi)]^2)\Bigg\}\Delta\Bigg]$$
$$+42\,\mathbb{E}[|\delta Y_{t_k}|^2] \tag{A.5}$$

with $C(u)=126K^2\max\{u^2,u\}$ independent of π.

Next we let $r_0\in(0,T)$ be such that $C(r)\le\frac{1}{6}\min\{1,C'\}$ for all $r\le r_0$, where C' is the constant from Lemma 3.6 [with $\tilde U$ taken equal to the function $\delta\tilde Z^{(\pi)}$]. Let us fix $b=[T/r_0]+1$ and consider the regular partition of $[0,T]$ into b intervals. We set for $0\le\ell\le b-1, I_\ell=\{k:t_k\in\pi\cap[\ell T/b,(\ell+1)T/b)\}$, $\ell_*=\min I_\ell$, $\ell^*=\max I_\ell+1$. Then we obtain from (A.5) and Lemma 3.6 that for every ℓ^*

$$\mathbb{E}\left[\sup_{\ell_*\le r<\ell^*}|\delta Y_{t_r}|^2+\sum_{r=\ell_*}^{\ell^*-1}|\delta Z_{t_r}|^2\Delta+\langle\delta\tilde M\rangle_{t_{\ell^*-1}}-\langle\delta\tilde M\rangle_{t_{\ell_*}}+\langle\delta M\rangle_{t_{\ell_*}}-\langle\delta M\rangle_{t_{\ell^*-1}}\right]$$
$$\le\quad 42\cdot\frac{6}{5}\,\mathbb{E}\Bigg[|\delta Y_{t^*_\ell}|+\frac{1}{5}\sum_{r=\ell_*}^{\ell^*-1}\Big|\delta f^{(\pi)}(t_r,Y^0_{t_r},Z^0_{t_r},\tilde Z^0_{t_r})\Big|^2\Delta\Bigg].$$

The proof is completed by a repeated application of this inequality.

Appendix B. Proof of Lemma 4.3

Proof of part (i). Recall that $\tilde Z^{(\pi),p}_{t_i}(0)=0$ for all i. Thus, to prove (4.3), it is enough to show that $\tilde Z^{(\pi),p}_{t_i}(x)$ is uniformly Lipschitz in $x\in\mathbb{R}^{d_2}$. Given the assumed form of F, it is possible to find a function $y^{(\pi),p}_{t_i}:\mathbb{R}^{(d_1+d_2)i}\to\mathbb{R}$ such that $y^{(\pi),p}_{t_i}(\Delta W^{(\pi)}_{t_0},\Delta X^{(\pi)}_{t_0},\dots,\Delta W^{(\pi)}_{t_{i-1}},\Delta X^{(\pi)}_{t_{i-1}}):=Y^{(\pi),p}_{t_i}$. Subsequently, we will suppress the arguments $\Delta W^{(\pi)}_{t_0},\Delta X^{(\pi)}_{t_0},\dots,\Delta W^{(\pi)}_{t_{i-2}},\Delta X^{(\pi)}_{t_{i-2}}$ whenever there is no ambiguity and write $y^{(\pi),p}_{t_i}(\Delta W^{(\pi)}_{t_{i-1}},\Delta X^{(\pi)}_{t_{i-1}})$.

Fix $t\in[0,T]$ and for every mesh size Δ choose i such that $i\Delta\le t<(i+1)\Delta$ and denote $w=\Delta w_{t_i}$ and $x=\Delta x_{t_i}$. Denote by $Y^{(\pi),p,w,x}_{t_j}$ for $j\ge i+1$ the process $(Y^{(\pi),p}_{t_j})_{j\ge i+1}$ conditional on $\Delta W^{(\pi)}_{t_i}=w$ and $\Delta X^{(\pi)}_{t_i}=x$. For $j\ge i+1$ the conditioned BSΔE with solution $(Y^{(\pi),p,w,x}_t,Z^{(\pi),p,w,x}_t,\tilde Z^{(\pi),p,w,x}_t,M^{(\pi),p,w,x}_t)$ can be written as

$$Y^{(\pi),p,w,x}_{t_j}=F^{(\pi),w,x}+\sum_{t_u\ge t_j}f(t_u,Y^{(\pi),p-1,w,x}_{t_u},Z^{(\pi),p-1,w,x}_{t_u},\tilde Z^{(\pi),p-1,w,x}_{t_u})\Delta-\sum_{t_u\ge t_j}Z^{(\pi),p,w,x}_{t_u}\Delta W^{(\pi)}_u$$
$$-\sum_{t_u\ge t_j}\Big\{\tilde Z^{(\pi),p,w,x}_{t_u}(\Delta X^{(\pi)}_{t_i})-\mathbb{E}_{t_{i-1}}[\tilde Z^{(\pi),p,w,x}_{t_u}(\Delta X^{(\pi)}_{t_i})]\Big\}-(M^{(\pi),p,w,x}_T-M^{(\pi),p,w,x}_{t_j}), \tag{B.1}$$

with

$$F^{(\pi),w,x} = H\Big(w_{t_1}, \dots, w_{t_i} + w, \dots, w_{t_i} + w + W_T^{(\pi)} - W_{t_{i+1}}^{(\pi)},$$
$$x_{t_1}, \dots, x_{t_i} + x, \dots, x_{t_i} + x + X_T^{(\pi)} - X_{t_{i+1}}^{(\pi)}\Big).$$

Clearly, $y_{t_{i+1}}^{(\pi),p}(w,x)$ has the same law as $Y_{t_{i+1}}^{(\pi),p,w,x}$. To simplify notation let us assume for the rest of the proof that W and X (and hence $W^{(\pi)}$ and $X^{(\pi)}$) are one-dimensional. The lemma would follow if we could prove through an induction over p that for every p the following holds: For every $l \in \mathbb{N}_0$ there exist constants $\tilde{K}_{Y,l,p}, \tilde{K}_{Z,l,p}, \tilde{K}_{\tilde{Z},l,p} > 0$ such that for every $m, k = 0, \dots, l$, and for all t :

(a) For the mappings $(w,x) \to Y_j^{(\pi),p,w,x}$ we have that $\sup_{w,x,\Delta,t_j>t} \left|\frac{\partial^{m+k}}{\partial w^m \partial x^k} Y_{t_j}^{(\pi),p,w,x}\right| \le \tilde{K}_{Y,l,p}$.
(b) For the mappings $(w,x) \to \tilde{Z}_{t_j}^{(\pi),p,w,x}$ we have that $\sup_{w,x,\Delta,t_j>t} \left|\frac{\partial^{m+k}}{\partial w^m \partial x^k} \tilde{Z}_{t_j}^{(\pi),p,w,x}\right| \le \tilde{K}_{\tilde{Z},l,p}$.
(c) For the mappings $(w,x) \to Z_{t_j}^{(\pi),p,w,x}$ we have that $\sup_{w,x,\Delta,t_j>t} \left|\frac{\partial^{m+k}}{\partial w^m \partial x^k} Z_{t_j}^{(\pi),p,w,x}\right| \le \tilde{K}_{Z,l,p}$.

Notice that (b) implies in particular that $\sup_{w,x,\Delta,t_j>t} \left|\tilde{Z}_{t_j}^{(\pi),p,w,x}(x')\right| \le (\tilde{K}_{\tilde{Z},1,p} \vee \tilde{K}_{\tilde{Z},0,p})(1 \wedge |x'|)$.

Let us prove (a)—(c). As $Y^{(\pi),0} = Z^{(\pi),0} = \tilde{Z}^{(\pi),0} = 0$, (a)–(c) clearly hold for $p = 0$ with $\tilde{K}_{Y,l,0} = \tilde{K}_{Z,l,0} = \tilde{K}_{\tilde{Z},l,0} = 0$ for all l. Now assume that we have shown the induction for $p-1$. Let us next show (a)-(c) for p.

By the induction assumption for all $j\Delta \ge t$ all higher derivatives of the processes $Y_j^{(\pi),p-1,w,x}$, $Z_j^{(\pi),p-1,w,x}$, and $\tilde{Z}_j^{(\pi),p-1,w,x}$ with respect to w and x satisfy (a)—(c). As by assumption also all higher derivatives of $f^{(\pi)}(t_j, \cdot, \cdot, \cdot)$ are bounded as well uniformly in t, j and Δ with $t_j > t$ we have that

$$\frac{\partial^{m+k}}{\partial w^m \partial x^k} f^{(\pi)}(t_j, Y_{t_j}^{(\pi),p-1,w,x}, Z_{t_j}^{(\pi),p-1,w,x}, \tilde{Z}_{t_j}^{(\pi),p-1,w,x})$$

is uniformly bounded by a constant, say $\hat{K}_{l,p-1}$. Now (B.1) entails that

$$\begin{aligned}
&\sup_{w,x,\Delta,t_j>t} \Big|\frac{\partial^{m+k}}{\partial w^m \partial x^k} Y_{t_j}^{(\pi),p,w,x}\Big| \\
&= \sup_{w,x,\Delta,t_j>t} \Big|\mathbb{E}_{t_j}\left[\frac{\partial^{m+k}}{\partial w^m \partial x^k} F^{(\pi),w,x} + \sum_{u \ge j} \frac{\partial^{m+k}}{\partial w^m \partial x^k} f^{(\pi)}(t_u, Y_{t_u}^{(\pi),p-1,w,x}, Z_{t_u}^{(\pi),p-1,w,x}, \tilde{Z}_{t_u}^{(\pi),p-1,w,x})\Delta\right]\Big| \\
&\le \sup_{w,x,\Delta} \Big|\Big|\frac{\partial^{m+k}}{\partial w^m \partial x^k} F^{(\pi),w,x}\Big|\Big|_\infty + T \sup_{w,x,\Delta,t_j>t} \Big|\Big|\frac{\partial^{m+k}}{\partial w^m \partial x^k} f^n((j+1)/n, Y_{t_j}^{(\pi),p-1,w,x}, Z_{t_j}^{(\pi),p-1,w,x}, \tilde{Z}_{t_j}^{(\pi),p-1,w,x})\Big|\Big|_\infty \\
&\le \tilde{K}_{H,l} + T\hat{K}_{l,p-1} =: \tilde{K}_{Y,l,p},
\end{aligned}$$

where $\tilde{K}_{H,l}$ is the uniform bound of the derivatives of the function H up to order l for every $l \in \mathbb{N}_0$. This shows that (a) holds. The validity of (b) follows immediately from that of (a) and the form (3.8) of $\tilde{Z}$.

To see that (c) holds true note that for every $t_j > t$

$$\begin{aligned}
&|\frac{\partial^{m+k}}{\partial w^m \partial x^k} Z_{t_j}^{(\pi),p,w,x}| \\
&= \Delta^{-1}\Big|\mathbb{E}_{t_j}\left[\frac{\partial^{m+k}}{\partial w^m \partial x^k} Y_{t_j}^{(\pi),p,w,x} \Delta W_{t_j}^{(\pi)}\right]\Big| \\
&= \Delta^{-1}\Big|\mathbb{E}_{t_j}\left[\Big(\frac{\partial^{m+k}}{\partial w^m \partial x^k} y_{t_j}^{(\pi),p,w,x}(\Delta W_{t_j}^{(\pi)}, \Delta X_{t_j}^{(\pi)}) - \frac{\partial^{m+k}}{\partial w^m \partial x^k} y_{t_j}^{(\pi),p,w,x}(0, \Delta X_{t_j}^{(\pi)})\Big)\Delta W_{t_j}^{(\pi)}\right]\Big| \\
&\le \Delta^{-1}\mathbb{E}_{t_j}\left[\Big|\frac{\partial^{m+k}}{\partial w^m \partial x^k} y_{t_j}^{(\pi),p,w,x}(\Delta W_{t_j}^{(\pi)}, \Delta X_{t_j}^{(\pi)}) - \frac{\partial^{m+k}}{\partial w^m \partial x^k} y_{t_j}^{(\pi),p,w,x}(0, \Delta X_{t_j}^{(\pi)})\Big||\Delta W_{t_{j+1}}^{(\pi)}|\right] \\
&\le \Delta^{-1}\tilde{K}_{Y,l+1,p}\mathbb{E}\left[|\Delta W_{t_j}^{(\pi)}|' \ |\Delta W_{t_j}^{(\pi)}|\right] = \Delta^{-1}\tilde{K}_{Y,l+1,p}\mathbb{E}\left[\Big|\Delta W_{t_j}^{(\pi)}\Big|^2\right] = \tilde{K}_{Y,l+1,p}.
\end{aligned}$$

This establishes that (c) holds with $\tilde{K}_{Z,l,p} := \tilde{K}_{Y,l+1,p}$. The proof of the induction is complete. □

Proof of part (ii). The proof is analogous to the proof of part (i). Denote by $Y_s^{p,\infty,x}$ for $s \geq t$ the process $(Y_t^{p,\infty})_{s\geq t}$ conditional on $\Delta X_t = X_t - X_{t^-} = x$. For $s \geq t$ the conditioned BSDE with solution $(Y_s^{p,\infty,x}, Z_s^{p,\infty,x}, \tilde{Z}_s^{p,\infty,x})$ can be written as

$$Y_s^{p,\infty,x} = F^x + \int_s^T f(u, Y_u^{p-1,\infty,x}, Z_u^{p-1,\infty,x}, \tilde{Z}_u^{p-1,\infty,x})\mathrm{d}u - \int_s^T Z_u^{p,\infty,x}\mathrm{d}W_u$$
$$- \int_{(s,T]\times\mathbb{R}^{d_2}\setminus\{0\}} \tilde{Z}_u^{p,\infty,x}(z)\tilde{N}_p(\mathrm{d}u \times \mathrm{d}z).$$

One may check directly that we have $\tilde{Z}_t^{p,\infty,x}(z) = Y_t^{p,\infty,x+z} - Y_t^{p,\infty,x}$, so that we only have to show that $\frac{\partial Y_t^{p,\infty,x}}{\partial x}$ is uniformly bounded. Using the assumptions on H (Assumption 2) this follows by a line of reasoning that is analogous to the one followed in part (i). □

References

[1] S. Aazizi (2013). Discrete time approximation of decoupled Forward-Backward SDE driven by pure jump Lévy-processes. *Advances in Applied Probability* 45, 791-821.

[2] S. Asmussen, and J. Rosiński (2001). Approximations of small jumps of Lévy processes with a view towards simulation. *Journal of Applied Probability* 38, 482-493.

[3] G. Barles, R. Buckdahn, and E. Pardoux (1997). Backward stochastic differential equations and integral-partial differential equations. *Stochastics: An International Journal of Probability and Stochastic Processes* 60, 57-83.

[4] D. Becherer (2006). Bounded solutions to backward SDE's with jumps for utility optimization and indifference hedging. *Annals of Applied Probability* 16, 2027-2054.

[5] B. Bouchard, and R. Elie (2008). Discrete-time approximation of decoupled forward-backward SDE with jumps. *Stochastic Processes and Their Applications* 118, 53-75.

[6] B. Bouchard, R. Elie, and N. Touzi (2009). Discrete-time approximation of BSDEs and probabilistic schemes for fully nonlinear PDEs. *Radon Series on Computational and Applied Mathematics* 8, 91-124.

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

[8] P. Briand, B. Delyon, and J. Mémin (2002). On the robustness of backward stochastic differential equations. *Stochastic Processes and Their Applications* 97, 229-253.

[9] P. Briand, B. Delyon, and J. Mémin (2001). Donsker-type theorem for BSDEs. *Electronic Communications in Probabability* 6, 1–14.

[10] P. Cheridito, and M. Stadje (2013). BS∆Es and BSDEs with non-Lipschitz drivers: comparison, convergence and robustness. *Bernoulli* 19, 1047-1085.

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

[12] F. Coquet, J. Mémin, and L. Słominski (2004). On weak convergence of filtrations. *Séminaire de Probabilités XXXV*, 306–328.

[13] L. Delong (2013) *Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications*, Springer.

[14] N. El Karoui, and R. Rouge (2000). Pricing via utility maximization and entropy. *Mathematical Finance* 10, 259-276.

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

[16] A. Eyraud-Loisel (2005). Backward stochastic differential equations with enlarged filtration. Option hedging of an insider trader in a financial market with jumps. *Stochastic Processes and Their Applications* 115, 1745-1763.

[17] Y. Hu, P. Imkeller, and M. Müller (2005). Utility maximization in incomplete markets. *Annals of Applied Probability* 15, 1691-1712.

[18] J. Jacod, and A.N. Shiryaev (2003). *Limit theorems for stochastic processes*, Springer Berlin.

[19] M. Jeanblanc, A. Matoussi, and A. Ngoupeyou (2010). Robust utility maximization in a discontinuous filtration. *Preprint*, University of Evry-Val d'Essonne, Évry.

[20] S. Klöppel, and M. Schweizer (2007). Dynamic indifference valuation via convex risk measures. *Mathematical Finance* 17, 599-627.

[21] T.G. Kurtz, and P.E. Protter (1991). Weak limit theorems for stochastic integrals and stochastic differential equations. *Annals of Probability* 1035-1070.

[22] R.J.A. Laeven, and M. Stadje (2012). Robust portfolio choice and indifference valuation. Forthcoming in: *Mathematics of Operations Research*.

[23] A.E.B. Lim (2006). Mean-variance hedging when there are jumps. *SIAM Journal on Control and Optimization* 44, 1893-1922.

[24] A. Lejay, E. Mordecki, and S. Torres (2007). *Numerical approximation of Backward Stochastic Differential Equations with Jumps. Preprint.*

[25] H. Luschgy, and G. Pagès (2008). Moment estimates for Lévy processes. *Electronic Journal of Probability* 13, 422–434.

[26] J. Ma, P. Protter, J. San Martin, and S. Torres (2002). Numerical methods for backward SDEs. *Annals of Applied Probabability* 12, 302–316.

[27] J. Mémin, and L. Słominski (1991). Condition UT et stabilité en loi des solutions d'équations differéntielles stochastiques. *Séminaire de Probabilités XXV*, 162–177, Springer

[28] J. Mémin (2003). Stability of Doob-Meyer decomposition under extended convergence. *Acta Mathematicae Applicatae Sinica (English Series)* 19, 177–190.

[29] M. Mania, and M. Schweizer (2005). Dynamic exponential utility indifference valuation. *Annals of Applied Probability* 15, 2113-2143.

[30] M. A. Morlais (2009a). Quadratic BSDEs driven by a continuous martingale and applications to the utility maximization problem. *Finance and Stochastics* 13, 121-150.

[31] M.A. Morlais (2009b). Utility maximization in a jump market model. *Stochastics: An International Journal of Probability and Stochastics Processes* 81, 1-27.

[32] M.A. Morlais (2010). A new existence result for BSDEs with jumps and application to the utility maximization problem. *Stochastic Processes and Their Applications* 120, 1966-1995.

[33] A. Pelsser, and M. Stadje (2014). Time-consistent and market-consistent evaluations. *Mathematical Finance* 24, 25-65.

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

[35] K. Sato (1999). *Lévy processes and infinitely divisible distributions.* Cambridge university press.

[36] R. Sircar, and S. Sturm (2011). From smile asymptotics to market risk measures. Forthcoming in: *Mathematical Finance.*

[37] S. Tang, and X. Li (1994). Necessary conditions for optimal control of stochastic systems with random jumps. *SIAM Journal of Control and Optimization* 32 1447-1475.

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

Department of Mathematics, Imperial College London
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