Math-Net.Ru

Общероссийский математический портал

R. Buckdahn, H. J. Engelbert, A. Rascanu, On weak solutions of backward stochastic differential equations, *Теория вероятн. и ее примен.*, 2004, том 49, выпуск 1, 70–108

DOI: <https://doi.org/10.4213/tvp237>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.162.85.209

26 апреля 2023 г., 15:05:10

© 2004 г.

BUCKDAHN R.*, ENGELBERT H.-J.**

RĂȘCANU A.***

ON WEAK SOLUTIONS OF BACKWARD STOCHASTIC
DIFFERENTIAL EQUATIONS¹⁾*En hommage à Paul-André Meyer*

Основная цель статьи — обсудить понятие слабого решения одного типа обратных стохастических дифференциальных уравнений. Используя слабую сходимость в топологии Мейера–Чжэна, мы получаем общий результат о существовании. Терминальное условие H функционально зависит от непрерывного справа и имеющего пределы слева управляющего процесса X , а коэффициент f зависит от времени t и функционально от X и процесса решения Y . Функционал $f(t, x, y)$, $(t, x, y) \in [0, T] \times D([0, T]; \mathbf{R}^{d+m})$ предполагается ограниченным и непрерывным по (x, y) на пространстве Скорохода $D([0, T]; \mathbf{R}^{d+m})$ в топологии Мейера–Чжэна. С помощью примеров типа примера Цирельсона мы показываем, что, действительно, существуют слабые решения, не являющиеся сильными, т.е. не являющиеся решениями в обычном смысле. Мы обсуждаем также потраекторную единственность решения и единственность по распределению и заключаем, что (подобно теореме Ямада–Ватанабэ) потраекторная единственность и слабое существование гарантируют существование (единственного) сильного решения. Применение этих идей дало нам возможность установить существование (единственного) сильного решения в случае, когда, в дополнение к описанным выше предположениям, f удовлетворяет некоторому обобщенному условию липшицева типа.

Ключевые слова и фразы: обратные стохастические дифференциальные уравнения, слабые решения, сильные решения, пример Цирельсона, потраекторная единственность, единственность по распределению, топология Мейера–Чжэна, слабая сходимость.

* Laboratoire des Mathématiques, CNRS-FRE 2218, Université de Bretagne Occidentale, Brest, France; e-mail: Rainer.Buchdahn@univ-brest.fr

** Institut für Stochastik, Fakultät für Mathematik, Friedrich Schiller-Universität, Jena, Germany; e-mail: englbirt@minet.uni-jena.de

*** Facultatea de matematică, Universitatea Alexandru Ioan Cuza, Iași, Romania; e-mail: rascanu@uaic.ro

¹⁾ This work was supported in part by the European Community's Human Potential Programme under contract HPRN-CT-2002-00281 (Evolution Equations).

1. Introduction and basic definitions. The notion of a weak solution of a (forward) stochastic differential equation (SDE) has been well-understood for a long time and plays a basic role in the development of the theory in the last decades. After the pioneering work of K. Itô [13], [14] who investigated SDEs in the case of Lipschitz continuous coefficients, A. V. Skorokhod [26], [27] was the first who constructed solutions of SDEs for only continuous coefficients. It turned out that these solutions, in general, lost the property to be adapted to the filtration generated by the driving Brownian motion. Henceforth they were called «weak» solutions to distinguish them from «strong» solutions which are adapted to the Brownian filtration. Much progress was achieved in the first half of the 70th, above all by Japanese and Soviet probabilists, in clarifying the relation between weak existence, strong existence, uniqueness in law and pathwise uniqueness. We only mention the celebrated Yamada–Watanabe theorem (cf. [29]), Tanaka's example (cf., e.g., [30]) and Tsirelson's example (cf. [28]). A milestone was put by N. V. Krylov [17], [18] who proved the existence of weak solutions of SDEs for only measurable and bounded coefficients in the case of diffusion coefficients satisfying a strict ellipticity condition. This result could even be improved by A. Rozkosz and L. Słomiński [24], [25] weakening significantly the nondegeneracy condition and allowing linear growth of the coefficients.

In the one-dimensional case one can say even more. For vanishing drift coefficient and time-independent diffusion coefficient depending only on the present state, H.-J. Engelbert and W. Schmidt [9]–[11] gave necessary and sufficient conditions for the existence of weak solutions as well as for uniqueness in law. For equations with nonvanishing drift coefficient (depending only on the present state), there were found simple sufficient conditions characterizing the «regular» case. Extensions to «singular» SDEs were investigated by A. S. Cherny and H.-J. Engelbert in [6].

The theory of backward stochastic differential equations (BSDEs) has been developed only recently, beginning with the 90th. In a general form, BSDEs were introduced by E. Pardoux and S. Peng [21], who proved existence and uniqueness of the solution (adapted to the Brownian filtration) under conditions including basically the Lipschitz continuity of the driver f . Their aim was to describe a solution of a second order quasilinear partial differential equation in probabilistic terms.

Linear BSDEs related to the stochastic version of Pontryagin's maximum principle had been studied long time before the pioneering work on general nonlinear BSDEs by Pardoux and Peng [21]: In 1973, J. M. Bismut [4] was the first who introduced such a linear BSDE; other works on the maximum principle in stochastic control were written by Yu. M. Kabanov [16] in 1978 and V. I. Arkin and M. T. Saksonov [2] in 1979. In 1978, J. M. Bismut [5] introduced a BSDE with nonlinearity (Riccati equation) for which he proved existence and uniqueness of the solution.

In the last decade, the theory of BSDEs has found further important applications and has become a powerful tool in many fields, above all financial mathematics, optimal control and stochastic games, partial differential equations and homogenization. The collected texts [8] edited by N. El Karoui and L. Mazliak give a useful introduction into the theory of BSDEs and their applications.

From the beginning, many authors attempted to improve the existence and uniqueness result of E. Pardoux and S. Peng [21] by weakening the Lipschitz continuity of f . One recent paper is [3], where existence and uniqueness is proven for coefficients satisfying a certain local monotonicity condition. In the one-dimensional case, J.-P. Lepeltier and J. San Martin [19] were able to prove existence of the solution for only continuous f (with linear growth). Another far reaching result was obtained by S. Hamadène, J.-P. Lepeltier and S. Peng [12]. They could prove the existence of a solution if the driving process X is a (strong) solution of a (forward) SDE with Lipschitz coefficients and if the driver $f(s, X_s, Y_s, Z_s)$ (where (Y, Z) denotes the solution process) depends on X in a Markovian way. They only assume that f is continuous in (y, z) and satisfies a certain growth condition. Additionally, the transition probabilities of the Markov process X must satisfy a certain domination condition. As an example, we can choose $X = W$, a Brownian motion, or, more generally, a solution X of a forward SDE with diffusion matrix which is bounded and strictly nondegenerated. However, as far as we know, all solutions of BSDEs which were constructed previously are adapted to the filtration generated by the driving process X . A solution of this type should be called strong solution. F. Antonelli and J. Ma [1] introduce a certain type of weak solutions of forward-backward SDEs. But it should be noted that all solutions constructed in this paper, although not adapted to a certain Brownian filtration, always are adapted with respect to the filtration generated by the driving process X . In our understanding, these are not weak, but strong solutions.

One may wonder that in the theory of BSDEs a notion what should be called weak solution and its investigation is not well developed. This may be caused by the special feature of BSDEs resulting in conceptional difficulties. On the other side, and this was our main motivation for working in this topic, one could expect that, similar to forward equations, an appropriate extension of the notion of a solution would be an enrichment in many aspects, above all in handling existence and uniqueness of solutions.

The main purpose of the present paper is to make an attempt in this direction. For our better understanding of the essential features, we concentrate on BSDEs rather than to include forward-backward SDEs or other generalizations. However, the driving process X (usually a Brownian motion) is a general càdlàg process and the terminal condition H and the driver f are of functional type. To explain the situation in more detail,

let us introduce some definitions and notation.

Let $T > 0$ be a real number (the terminal time). For any natural number $p \geq 1$, by $D(\mathbf{R}^p) := D([0, T]; \mathbf{R}^p)$ we denote the Skorokhod space of càdlàg (right continuous with left-hand limits) functions x on $[0, T]$ with values in $\mathbf{R}^p$ such that $x(T-) := \lim_{t \uparrow T} x(t) = x(T)$, and by convention, $x(0-) = 0$. Let ζ be the coordinate mapping on $D(\mathbf{R}^p)$ defined by

$$\zeta_t(x) = x(t), \quad t \in [0, T], \quad x \in D(\mathbf{R}^p),$$

and introduce the σ -algebras of subsets of $D(\mathbf{R}^p)$,

$$\mathcal{D}_t^s := \mathcal{D}_t^s(\mathbf{R}^p) := \sigma(\{\zeta_u : u \in [s, t]\}), \quad 0 \leq s \leq t \leq T,$$

$$\mathcal{D} := \mathcal{D}(\mathbf{R}^p) := \mathcal{D}_T^0(\mathbf{R}^p).$$

Suppose that there are given the following objects:

(H.1) a measurable $\mathbf{R}^m$ -valued functional H on $(D(\mathbf{R}^d), \mathcal{D}(\mathbf{R}^d))$;

(H.2) a measurable $\mathbf{R}^m$ -valued functional f on the product space $([0, T] \times D(\mathbf{R}^d) \times D(\mathbf{R}^m), \mathcal{B}([0, T]) \otimes \mathcal{D}(\mathbf{R}^d) \otimes \mathcal{D}(\mathbf{R}^m))$, where $\mathcal{B}([0, T])$ is the σ -algebra of Borel sets of $[0, T]$.

Definition 1.1. Let $X = (X_t)_{t \in [0, T]}$ be an $\mathbf{R}^d$ -valued càdlàg process defined on a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$, i.e., X is a random variable with values in $(D(\mathbf{R}^d), \mathcal{D}(\mathbf{R}^d))$ defined on $(\Omega, \mathcal{F}, \mathbf{P})$. One denotes by $\mathbb{F}^X = (\mathcal{F}_t^X)_{t \in [0, T]}$ the smallest filtration satisfying the usual conditions, which is such that X is $\mathbb{F}^X$ -adapted. An $\mathbf{R}^m$ -valued càdlàg process Y is called a *strong solution of BSDE* (H, f, X) if the following conditions are satisfied:

(i) $\mathbf{E}[|H(X)|] < +\infty$, $\mathbf{E}[\int_0^T |f(s, X, Y)| ds] < +\infty$

(ii) the BSDE

$$Y_t = \mathbf{E} \left[H(X) + \int_t^T f(s, X, Y) ds \middle| \mathcal{F}_t^X \right], \quad t \in [0, T], \quad (1.1)$$

holds.

The triplet (H, f, X) is called the *data set* of BSDE (1.1), it determines the *terminal value* $H(X)$. The functional f is called the *driver* and X the *driving process*.

Remark 1.1. (i) It is an immediate consequence of the definition of a strong solution Y of BSDE (H, f, X) that Y is always $\mathbb{F}^X$ -adapted.

(i) Usually the driving process X is a d -dimensional Brownian motion W . In the case of a (noncoupled) forward-backward SDE, X is a (strong) solution of a forward SDE.

Now we come to the definition of a weak solution of the BSDE.

Definition 1.2. Let μ be a probability measure on $(D(\mathbf{R}^d), \mathcal{D}(\mathbf{R}^d))$. Suppose that $(\Omega, \mathcal{F}, \mathbf{P})$ is a complete probability space, $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$ a filtration satisfying the usual conditions, and X, Y are

càdlàg processes defined on $(\Omega, \mathcal{F}, \mathbf{P})$ with values in $\mathbf{R}^d$ and $\mathbf{R}^m$, respectively. Then $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y)$ is called a *weak solution of BSDE* (H, f, μ) if the following conditions are satisfied:

- (A.i) *The distribution of X on $D(\mathbf{R}^d)$ is μ ;*
- (A.ii) *X is $\mathbb{F}$ -adapted;*
- (A.iii) $\mathbf{E}[|H(X)|] < +\infty$, $\mathbf{E}[\int_0^T |f(s, X, Y)| ds] < +\infty$;
- (A.iv) *every $\mathbb{F}^X$ -martingale is also an $\mathbb{F}$ -martingale;*
- (A.v) *the BSDE*

$$Y_t = \mathbf{E}\left[H(X) + \int_t^T f(s, X, Y) ds \mid \mathcal{F}_t\right], \quad t \in [0, T], \quad (1.2)$$

holds.

R e m a r k 1.2. (i) If $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y)$ is a weak solution of BSDE (H, f, μ) , then, obviously, Y is $\mathbb{F}$ -adapted.

(ii) If $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y)$ is a weak solution of BSDE (H, f, μ) such that Y is $\mathbb{F}^X$ -adapted, then Y is clearly a strong solution of BSDE (H, f, X) on $(\Omega, \mathcal{F}, \mathbf{P})$.

(iii) Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y)$ be a weak solution of BSDE (H, f, μ) . By $\mathbb{F}^{X,Y}$ we denote the smallest filtration satisfying the usual conditions, such that the process (X, Y) is adapted. Then $(\Omega, \mathcal{F}, \mathbb{F}^{X,Y}, \mathbf{P}, X, Y)$ is again a weak solution of BSDE (H, f, μ) . The filtration $\mathbb{F}^{X,Y}$ is the smallest with this property.

(iv) Obviously, every strong solution Y of BSDE (H, f, X) on $(\Omega, \mathcal{F}, \mathbf{P})$ defines a weak solution $(\Omega, \mathcal{F}, \mathbb{F}^X, \mathbf{P}, X, Y)$ of BSDE (H, f, μ) , where μ is the distribution of X .

(v) Condition (A.iv) of the above definition means that the filtration $\mathbb{F}$ is not too large in comparison with $\mathbb{F}^X$. More precisely, this condition is equivalent to the following condition:

(A.iv') $\mathcal{F}_t$ and $\mathcal{F}_T^X$ are conditionally independent given $\mathcal{F}_t^X$, $t \in [0, T]$.

This amounts to saying that the filtration $\mathbb{F}$ consists, additionally to $\mathbb{F}^X$, only of randomizations (i.e., independent experiments).

(vi) An important case is that of $X = W$, a Brownian motion. Let now $\mathbb{F}$ be a filtration such that $(W, \mathbb{F})$ is still a Brownian motion. The representation property of $(W, \mathbb{F})$ yields that $\mathbb{F}^X$ and $\mathbb{F}$ always satisfy condition (A.iv) of the above definition. Hence, in this case, looking for a weak solution $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y)$ of BSDE (H, f, μ) amounts to looking for a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$ endowed with a Brownian motion $(W, \mathbb{F})$ and with $\mathbb{F}$ -adapted, càdlàg processes X, Y such that (1.2) holds.

We are now going to discuss the BSDEs (1.1) and (1.2). To this end, we transform them in another but equivalent form. For this we state the following lemma which will also be used in the sequel.

Given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$, by $D_{\mathbb{F}, \mathbf{R}^m} := D_{\mathbb{F}, \mathbf{R}^m}(\Omega, \mathcal{F}, \mathbf{P})$ we denote the set of $\mathbf{R}^m$ -valued, $\mathbb{F}$ -adapted càdlàg processes.

The set of $\mathbf{R}^m$ -valued $\mathbb{F}$ -semimartingales U which are p -integrable, i.e., are such that

$$\mathbf{E} \left[\sup_{s \in [0, T]} |U_s|^p \right] < +\infty,$$

is denoted by $\mathcal{S}_{\mathbb{F}, \mathbf{R}^m}^p$.

Lemma 1.1. *Let $X \in D_{\mathbb{F}, \mathbf{R}^d}$, $Y \in D_{\mathbb{F}, \mathbf{R}^m}$ and $p \geq 1$ be such that $\mathbf{E}[\int_0^T |f(s, X, Y)|^p ds] < +\infty$. Furthermore, let $\xi \in L^p(\Omega, \mathcal{F}_T, \mathbf{P}; \mathbf{R}^m)$. Then the process*

$$Y_t^1 := \mathbf{E} \left[\xi + \int_t^T f(s, X, Y) ds \middle| \mathcal{F}_t \right], \quad t \in [0, T],$$

possesses a càdlàg version (which will again be denoted by Y^1). Moreover, Y^1 is an $\mathbb{F}$ -semimartingale and has the semimartingale decomposition

$$Y_t^1 = Y_0^1 + M_t + A_t, \quad t \in [0, T], \quad (1.3)$$

where M is an $\mathbf{R}^m$ -valued càdlàg martingale vanishing in $t = 0$, with the representation

$$\begin{aligned} M_t &= \mathbf{E} \left[\xi + \int_0^T f(s, X, Y) ds \middle| \mathcal{F}_t \right] - \mathbf{E} \left[\xi + \int_0^T f(s, X, Y) ds \middle| \mathcal{F}_0 \right] \\ &\quad - \mathbf{E} \left[\int_0^t \left(f(s, X, Y) - \{f(\cdot, X, Y)\}_s^{\mathbb{F}} \right) ds \middle| \mathcal{F}_t \right], \quad t \in [0, T], \end{aligned} \quad (1.4)$$

and A is an $\mathbf{R}^m$ -valued $\mathbb{F}$ -predictable process with p -integrable variation, given by

$$A_t = - \int_0^t \{f(\cdot, X, Y)\}_s^{\mathbb{F}} ds, \quad t \in [0, T]. \quad (1.5)$$

Here $\{f(\cdot, X, Y)\}^{\mathbb{F}}$ denotes the optional projection of the process $f(\cdot, X, Y)$ with respect to the filtration $\mathbb{F}$.¹⁾

Moreover, if $p > 1$, then $Y^1 \in \mathcal{S}_{\mathbb{F}, \mathbf{R}^m}^p$ and the càdlàg martingale M is p -integrable.

P r o o f. The proof is straight-forward and therefore omitted.

Proposition 1.1. *Suppose that $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y)$ satisfies the conditions (A.i)–(A.iv) of Definition 1.2. Then $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y)$ is a weak solution of BSDE(H, f, μ) if and only if Y satisfies the following equation:*

$$Y_t = H(X) + \int_t^T \{f(\cdot, X, Y)\}_s^{\mathbb{F}} ds - (M_T - M_t), \quad t \in [0, T], \quad (1.6)$$

where M is an $\mathbb{F}$ -martingale.

¹⁾ Denoting the Lebesgue measure on $[0, T]$ by λ , for any $\lambda \otimes \mathbf{P}$ -integrable, measurable process Z its optional projection $\{Z\}^{\mathbb{F}}$ is defined as $\{Z\}^{\mathbb{F}} = \mathbf{E}_{\lambda \otimes \mathbf{P}}[Z | \mathcal{O}]$ (the «conditional expectation» of Z with respect to the σ -finite measure $\lambda \otimes \mathbf{P}$), where $\mathcal{O}$ is the σ -algebra of $\mathbb{F}$ -optional subsets of $[0, T] \times \Omega$.

P r o o f. The necessity of relation (1.6) easily follows from Lemma 1.1, computing $Y_t - Y_T$. For the converse, we notice that

$$\mathbf{E} \left[\int_t^T \{f(\cdot, X, Y)\}_s^{\mathbb{F}} ds \mid \mathcal{F}_t \right] = \mathbf{E} \left[\int_t^T f(s, X, Y) ds \mid \mathcal{F}_t \right], \quad \mathbf{P}\text{-a.s.}$$

implies the desired result.

Replacing $\mathbb{F}$ by $\mathbb{F}^X$, Proposition 1.1 also admits a characterization of strong solutions Y of BSDE(H, f, X). Equation (1.6) is closer to the traditional form of a BSDE. Suppose, for example, that $p = 2$ and $\mathbb{F} = \mathbb{F}^W$, where $W = (W^1, \dots, W^d)$ is an $\mathbf{R}^d$ -valued martingale satisfying the predictable representation property (cf. [15]). Then we get

$$M_T - M_t = \int_t^T Z_s dW_s, \quad t \in [0, T],$$

for some $\mathbb{F}^W$ -predictable $\mathbf{R}^{m \times d}$ -valued process $Z = (Z^{i,j})$ with

$$\mathbf{E} \left[\sum_{i=1}^m \sum_{j=1}^d \int_0^T |Z_s^{i,j}|^2 d\langle W^j \rangle_s \right] < +\infty,$$

and equation (1.6) transforms into

$$Y_t = H(X) + \int_t^T \{f(\cdot, X, Y)\}_s^{\mathbb{F}^W} ds - \int_t^T Z_s dW_s, \quad t \in [0, T]. \quad (1.7)$$

As a special case, W can be chosen as a Brownian motion. Then, in the case of equation (1.7), it is commonly used that the driver f also depends on the unknown process $Z \in L_{\mathbb{F}^W, \mathbf{R}^{m \times d}}^2 := L^2([0, T] \times \Omega, \mathcal{P}, \lambda \otimes \mathbf{P}; \mathbf{R}^{m \times d})$, where $\mathcal{P}$ is the σ -algebra of the $\mathbb{F}^W$ -predictable subsets of $[0, T] \times \Omega$ and λ is the Lebesgue measure. To be more precise, we can write

$$Y_t = H(X) + \int_t^T \{f(\cdot, X, Y, Z)\}_s^{\mathbb{F}^W} ds - \int_t^T Z_s dW_s, \quad t \in [0, T]; \quad (1.8)$$

here the driver f is a measurable functional on $[0, T] \times D(\mathbf{R}^d) \times D(\mathbf{R}^m) \times L^2([0, T]; \mathbf{R}^{m \times d})$ with values in $\mathbf{R}^m$. Finally, noting that under some additional, more restrictive measurability properties on the driver f (f should depend on (s, x) in a nonanticipating way and on (s, x, y, z) only through $(s, x, y(s), z(s))$) we have

$$\int_t^T \{f(\cdot, X, Y, Z)\}_s^{\mathbb{F}^W} ds = \int_t^T f(s, X, Y, Z) ds,$$

which recovers the classical form of a BSDE.

At this place we emphasize that if one looks for a (possibly weak) solution of a backward equation $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y)$ with data (H, f, μ) , which has the form (1.8) or (1.7), where $(W, \mathbb{F})$ is some Brownian motion, one needs the *a priori* assumption $\mathbb{F} = \mathbb{F}^W$ to ensure the representation property for $(W, \mathbb{F})$. But this assumption has the immediate consequence that W can be used as driving process instead of X (i.e., one can suppose that μ is the

distribution of W) and Y is $\mathbb{F}^W$ -adapted. But then Y is a strong solution of $\text{BSDE}(H, f, W)$. So, in the context of weak solutions, equation (1.8), or also (1.7), makes not a lot of sense because it defines a too restrictive frame for the study of weak solutions.

This explains that, in investigating weak solutions for BSDEs, an integral representation of the martingale M in equation (1.6), giving the possibility to the driver f to depend on the integrand Z , has, in general, no meaning. However, there is a quite natural extension of equation (1.2) and (1.6). For this we consider, instead of f described in (H.2),

(H.2') a measurable $\mathbf{R}^m$ -valued functional f' on the product space

$$([0, T] \times D(\mathbf{R}^d) \times D(\mathbf{R}^m) \times D(\mathbf{R}^m), \mathcal{B}([0, T]) \otimes \mathcal{D}(\mathbf{R}^d) \otimes \mathcal{D}(\mathbf{R}^m) \otimes \mathcal{D}(\mathbf{R}^m)).$$

Definition 1.3. On a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$, let $X \in D_{\mathbb{F}, \mathbf{R}^d}$, and let Y be an $\mathbb{F}$ -adapted $\mathbf{R}^m$ -valued semimartingale with martingale part $M \in D_{\mathbb{F}, \mathbf{R}^m}$. Furthermore, let $\widetilde{M}$ be the process defined by $\widetilde{M}_t = M_T - M_t$, $t \in [0, T]$. The six-tuple $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y)$ is then called a *weak solution* of $\text{BSDE}(H, f', \mu)$ if (A.i), (A.ii), and (A.iv) of Definition 1.2 and the following conditions are satisfied:

(A.iii') $\mathbf{E}|H(X)| < +\infty$, $\mathbf{E} \int_0^T |f'(s, X, Y, \widetilde{M})| ds < +\infty$;

(A.v') the BSDE

$$Y_t = \mathbf{E} \left[H(X) + \int_t^T f'(s, X, Y, \widetilde{M}) ds \mid \mathcal{F}_t \right], \quad t \in [0, T], \quad (1.9)$$

holds.

Observe that, for (1.9), an equivalent formulation is

$$Y_t = \mathbf{E}H(X) + \int_t^T \{f'(\cdot, X, Y, \widetilde{M})\}_s^{\mathbb{F}} ds - (M_T - M_t), \quad t \in [0, T], \quad (1.10)$$

for some $\mathbb{F}$ -martingale $M \in D_{\mathbb{F}, \mathbf{R}^m}$ with (A.iii').

For simplicity of the presentation, we will concentrate on $\text{BSDE}(H, f, \mu)$ rather than to investigate $\text{BSDE}(H, f', \mu)$. The more general equation does not provide any serious difficulties and the results are analogous.

The paper is organized as follows.

In Section 2, we collect, as a recall, some results on the existence and uniqueness of strong solutions of $\text{BSDE}(H, f, X)$.

Section 3 includes examples of weak solutions to $\text{BSDE}(H, f, \mu)$ which are not strong. These examples are of Tsirelson type: They are constructed from the famous Tsirelson example.

In Section 4, we prove a general existence result for weak solutions of $\text{BSDE}(H, f, \mu)$. The main tool is weak convergence in the Meyer–Zheng topology. For this reason, our assumptions on the driver f include continuity on the Skorokhod space $D(\mathbf{R}^{d+m})$ in the Meyer–Zheng topology.

The use of the Meyer–Zheng topology is crucial for our purposes. It was an idea developed by D. Dellacherie and P.-A. Meyer [7, IV.40–46] to consider the Skorokhod space $D([0, T]; \mathbf{R}^p)$ as a subspace of the compact metric space $\mathcal{P}$ of all probability measures on $[0, T] \times [-\infty, +\infty]^p$, equipped with the topology of weak convergence: To every path associating the so-called pseudo-path (a probability measure on $[0, T] \times \mathbf{R}^p$), they created a concrete embedding which enabled to give effective and simple criteria for a probability measure on $\mathcal{P}$ being already supported by $D([0, T]; \mathbf{R}^p)$. Such criteria were just given by P.-A. Meyer and W. A. Zheng [20] in several important cases. Actually, this lead to effective and simple criteria for relative compactness of a sequence of probability measures on $D([0, T]; \mathbf{R}^p)$. This possibility depends on the fact that the Meyer–Zheng topology is much coarser than the topology of uniform convergence or the Skorokhod topology. The obvious disadvantage that now the continuity of functionals on $D([0, T]; \mathbf{R}^p)$ is a stronger condition, turns out to be not too restrictive.

In Section 5, we discuss pathwise uniqueness and uniqueness in law of weak solutions. Similar to the Yamada–Watanabe theorem [29], we conclude that pathwise uniqueness implies uniqueness in law and that pathwise uniqueness and existence of a weak solution ensure the existence of a (pathwise unique) strong solution. As for forward SDE, this conclusion provides an efficient tool for dividing the often hard problem of the existence of a strong solution into two simpler tasks: To prove existence of a weak solution and to prove pathwise uniqueness.

In Section 6 we give an application of this idea. As a result, we state the existence of a strong solution of BSDE(H, f, X) if, additionally to the assumptions of the existence theorem of Section 4, the driver f satisfies a certain generalized Lipschitz type condition.

2. Strong solutions of BSDE: A recall. Suppose that we are given an $\mathbf{R}^d$ -valued càdlàg process X on a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$, a terminal time $T > 0$, a terminal functional H , and a driver f . We always assume that H and f satisfy the conditions (H.1) and (H.2). Furthermore, we shall assume the following hypotheses.

- C o n d i t i o n 2.1.** (i) $\mathbf{E}|H(X)|^2 < +\infty$;
(ii) $f(s, x, \cdot)$ is $\mathcal{D}_T^s(\mathbf{R}^m)$ -measurable, $(s, x) \in [0, T] \times D(\mathbf{R}^d)$;
(iii) (uniform Lipschitz condition)

$$|f(s, x, y) - f(s, x, y')| \leq C \sup_{t \in [0, T]} |y_t - y'_t|,$$

$$(s, x) \in [0, T] \times D(\mathbf{R}^d), \quad y, y' \in D(\mathbf{R}^m);$$

- (iv) (linear growth condition)

$$|f(s, x, y)| \leq C \left(1 + \sup_{t \in [0, T]} |y_t| \right), \quad (s, x) \in [0, T] \times D(\mathbf{R}^d), \quad y \in D(\mathbf{R}^m).$$

Here and in the following, C denotes a constant which may change from line to line.

Proposition 2.1. *Suppose that the conditions (H.1), (H.2), and 2.1 (i)–(iv) are satisfied. Then there exists a strong solution $Y \in \mathcal{S}_{\mathbb{F}^X, \mathbf{R}^m}^2$ of BSDE(H, f, X). This solution is unique.*

P r o o f. We define a sequence $Y^n \in \mathcal{S}_{\mathbb{F}^X, \mathbf{R}^m}^2 \subset \mathcal{D}_{\mathbb{F}^X, \mathbf{R}^m}$ by induction. For $n = 0$, we set $Y^0 \equiv 0$. Given $Y^n \in \mathcal{S}_{\mathbb{F}^X, \mathbf{R}^m}^2$ for some $n \geq 0$, we note that

$$\begin{aligned} \mathbf{E} \left[\int_0^T |f(s, X, Y^n)|^2 ds \right] &\leq C \mathbf{E} \left[\int_0^T \left(1 + \sup_{t \in [0, T]} |Y_t^n| \right)^2 ds \right] \\ &\leq 2CT \left(1 + \mathbf{E} \left[\sup_{t \in [0, T]} |Y_t^n|^2 \right] \right) < +\infty, \end{aligned}$$

and define

$$Y_t^{n+1} = \mathbf{E} \left[H(X) + \int_t^T f(s, X, Y^n) ds \middle| \mathcal{F}_t^X \right], \quad t \in [0, T]. \quad (2.1)$$

Lemma 1.1 yields $Y^{n+1} \in \mathcal{S}_{\mathbb{F}^X, \mathbf{R}_m^2}$. Moreover in view of Condition 2.1 (ii), $f(s, x, y)$ only depends on y_u , $u \geq s$, for all $(s, x) \in [0, T] \times D(\mathbf{R}^d)$ and, consequently,

$$|f(s, X, Y^n) - f(s, X, Y^{n-1})| \leq C \sup_{u \in [s, T]} |Y_u^n - Y_u^{n-1}|, \quad s \in [0, T], \quad n \geq 1.$$

This implies

$$\begin{aligned} |Y_t^{n+1} - Y_t^n| &\leq \mathbf{E} \left[\int_t^T |f(s, X, Y^n) - f(s, X, Y^{n-1})| ds \middle| \mathcal{F}_t^X \right] \\ &\leq C \mathbf{E} \left[\int_t^T \sup_{u \in [s, T]} |Y_u^n - Y_u^{n-1}| ds \middle| \mathcal{F}_t^X \right], \quad t \in [0, T]. \end{aligned}$$

Setting $\eta_s^n = \sup_{u \in [s, T]} |Y_u^n - Y_u^{n-1}|$, $s \in [0, T]$, yields

$$\eta_t^{n+1} \leq C \sup_{u \in [t, T]} \mathbf{E} \left[\int_u^T \eta_s^n ds \middle| \mathcal{F}_u^X \right] \leq C \sup_{u \in [t, T]} \mathbf{E} \left[\int_t^T \eta_s^n ds \middle| \mathcal{F}_u^X \right],$$

$t \in [0, T]$. By Doob's inequality, we obtain

$$\begin{aligned} \mathbf{E}(\eta_t^{n+1})^2 &\leq C^2 \mathbf{E} \left[\left(\sup_{u \in [t, T]} \mathbf{E} \left[\int_u^T \eta_s^n ds \middle| \mathcal{F}_u^X \right] \right)^2 \right] \\ &\leq 4C^2 \mathbf{E} \left[\left(\int_t^T \eta_s^n ds \right)^2 \right] \leq 4C^2 T \int_t^T \mathbf{E}(\eta_s^n)^2 ds, \quad t \in [0, T], \end{aligned}$$

for all $n \geq 1$. By standard arguments, it now follows

$$\lim_{n, m \rightarrow +\infty} \mathbf{E} \left[\sup_{t \in [0, T]} |Y_t^n - Y_t^m|^2 \right] = 0.$$

Consequently, there exists $Y \in D_{\mathbb{F}^X, \mathbf{R}^m}$ with $\mathbf{E}[\sup_{t \in [0, T]} |Y_t|^2] < +\infty$ and such that $\lim_{n \rightarrow +\infty} \mathbf{E}[\sup_{t \in [0, T]} |Y_t^n - Y_t|^2] = 0$. Taking the limit in (2.1) as $n \rightarrow +\infty$, we get

$$Y_t = \mathbf{E} \left[H(X) + \int_t^T f(s, X, Y) ds \middle| \mathcal{F}_t^X \right], \quad \mathbf{P}\text{-a.s.}, \quad t \in [0, T].$$

Finally, Lemma 1.1 shows that $Y \in \mathcal{S}_{\mathbb{F}^X, \mathbf{R}^m}^2$. For proving the uniqueness, let $Y^1, Y^2 \in \mathcal{S}_{\mathbb{F}^X, \mathbf{R}^m}^2$ be two solutions, and put

$$g(t) = \mathbf{E} \left[\sup_{s \in [t, T]} |Y_s^1 - Y_s^2|^2 \right], \quad t \in [0, T].$$

Similar as above, we find

$$g(t) \leq 4C^2 T \int_t^T g(s) ds, \quad t \in [0, T].$$

Since g is bounded, Gronwall's lemma yields $g \equiv 0$ and hence $Y^1 = Y^2$, $\mathbf{P}$ -a.s.

R e m a r k 2.1. Let $X = W$ be a d -dimensional Brownian motion, and let $Y \in \mathcal{S}_{\mathbb{F}^W, \mathbf{R}^m}^2$ be a unique strong solution of BSDE(H, f, W) according to Proposition 2.1. In view of Lemma 1.1, the process M given by

$$M_t = Y_t - Y_0 + \int_0^t \{f(\cdot, X, Y)\}_s^{\mathbb{F}^W} ds, \quad t \in [0, T],$$

is an m -dimensional square integrable $\mathbb{F}^W$ -martingale. Because of the representation property of W , there is a unique $Z \in L_{\mathbb{F}^W, \mathbf{R}^{m \times d}}^2 := L^2([0, T] \times \Omega, \mathcal{P}, \lambda \otimes \mathbf{P}; \mathbf{R}^{m \times d})$ such that $M_t = \int_0^t Z_s dW_s$, $t \in [0, T]$, $\mathbf{P}$ -a.s. Consequently,

$$Y_t = H(W) + \int_t^T \{f(\cdot, W, Y)\}_s^{\mathbb{F}^W} ds - \int_t^T Z_s dW_s, \quad t \in [0, T], \quad \mathbf{P}\text{-a.s.},$$

or, equivalently,

$$dY_t = -\{f(\cdot, W, Y)\}_t^{\mathbb{F}^W} dt + Z_t dW_t, \quad t \in [0, T], \quad Y_T = H(W).$$

In the following proposition, this result will be generalized admitting drivers f which may depend, additionally, on the unknown process Z .

Proposition 2.2. *Suppose that the driving process $X = W$ is a d -dimensional Brownian motion, H satisfies the conditions (H.1) and 2.1 (i), and the driver f satisfies the following conditions:*

(B.i) *On $[0, T] \times D(\mathbf{R}^d) \times D(\mathbf{R}^m) \times L^2([0, T]; \mathbf{R}^{m \times d})$, f is an $\mathbf{R}^m$ -valued measurable functional;*

(B.ii) *there exists a submarkovian kernel $\mu(s, A)$, $s \in [0, T]$, $A \in \mathcal{B}([0, T])$, with the property*

$$\int_0^T \mu(s, A) ds \leq \lambda(A), \quad A \in \mathcal{B}([0, T]), \quad (2.2)$$

where λ denotes the Lebesgue measure on $[0, T]$, and such that

$$\begin{aligned} & |f(s, x, y, z) - f(s, x, y', z')| \\ & \leq C \left(\sup_{u \in [s, T]} |y(u) - y'(u)| + \int_s^T |z(u) - z'(u)| \mu(s, du) \right), \quad (2.3) \\ & |f(s, x, y, z)| \leq C \left(1 + \sup_{u \in [s, T]} |y(u)| + \int_s^T |z(u)| \mu(s, du) \right), \end{aligned}$$

for all $(s, x) \in [0, T] \times D(\mathbf{R}^d)$, $y, y' \in D(\mathbf{R}^m)$, $z, z' \in L^2([0, T]; \mathbf{R}^{m \times d})$.

Then there exists a unique $(Y, Z) \in \mathcal{S}_{\mathbb{F}^W, \mathbf{R}^m}^2 \times L_{\mathbb{F}^W, \mathbf{R}^{m \times d}}^2$ which is a solution of the BSDE

$$Y_t = H(W) + \int_t^T \{f(\cdot, W, Y, Z)\}_s^{\mathbb{F}^W} ds - \int_t^T Z_s dW_s, \quad t \in [0, T], \quad \mathbf{P}\text{-a.s.} \quad (2.4)$$

P r o o f. 1) Let $(Y, Z) \in \mathcal{S}_{\mathbb{F}^W, \mathbf{R}^m}^2 \times L_{\mathbb{F}^W, \mathbf{R}^{m \times d}}^2$. Then we define

$$Y_t^1 = \mathbf{E} \left[H(W) + \int_t^T f(s, W, Y, Z) ds \middle| \mathcal{F}_t^W \right], \quad t \in [0, T].$$

Similar to the statement of Lemma 1.1 and Remark 2.1, $Y^1 \in \mathcal{S}_{\mathbb{F}^W, \mathbf{R}^m}^2$ and $Z^1 \in L_{\mathbb{F}^W, \mathbf{R}^{m \times d}}^2$ can be found such that the operator F defined by $F(Y, Z) = (Y^1, Z^1)$ is a contraction in the Banach space $\mathcal{S}_{\mathbb{F}^W, \mathbf{R}^m}^2 \times L_{\mathbb{F}^W, \mathbf{R}^{m \times d}}^2$ restricted to a small time interval $[T - \delta, T]$. From the Banach fixed point theorem, we obtain a unique solution of BSDE (2.4) on $[T - \delta, T]$. A repeated application of the argument will then prove the proposition.

2) To verify the contraction property, let $(Y, Z), (\bar{Y}, \bar{Z}) \in \mathcal{S}_{\mathbb{F}^W, \mathbf{R}^m}^2 \times L_{\mathbb{F}^W, \mathbf{R}^{m \times d}}^2$. Then we have

$$Y_t^1 - \bar{Y}_t^1 = \int_t^T \left\{ f(\cdot, W, Y, Z) - f(\cdot, W, \bar{Y}, \bar{Z}) \right\}_s^{\mathbb{F}^W} ds - \int_t^T (Z_s^1 - \bar{Z}_s^1) dW_s, \quad (2.5)$$

$t \in [0, T]$. Hence,

$$Y_t^1 - \bar{Y}_t^1 = \mathbf{E} \left[\int_t^T \left\{ f(\cdot, W, Y, Z) - f(\cdot, W, \bar{Y}, \bar{Z}) \right\}_s^{\mathbb{F}^W} ds \middle| \mathcal{F}_t^W \right]$$

and for all $t \in [0, T]$ we can estimate $|Y_t^1 - \bar{Y}_t^1| \leq C \mathbf{E} \left[\int_t^T (\sup_{u \in [s, T]} |Y_u - \bar{Y}_u| + \int_s^T |Z_u - \bar{Z}_u| \mu(s, du)) ds \middle| \mathcal{F}_t^W \right]$. Using the property of the kernel μ and Doob's inequality, we now obtain

$$\begin{aligned} & \mathbf{E} \left[\sup_{t \in [r, T]} |Y_t^1 - \bar{Y}_t^1|^2 \right] \\ & \leq C^2 \mathbf{E} \left[\sup_{t \in [r, T]} \mathbf{E} \left[\int_r^T \left(\sup_{u \in [s, T]} |Y_u - \bar{Y}_u| + \int_s^T |Z_u - \bar{Z}_u| \mu(s, du) \right) ds \middle| \mathcal{F}_t^W \right]^2 \right] \\ & \leq C^2 \mathbf{E} \left[\sup_{t \in [r, T]} \mathbf{E} \left[\int_r^T \left(\sup_{u \in [s, T]} |Y_u - \bar{Y}_u| + |Z_s - \bar{Z}_s| \right) ds \middle| \mathcal{F}_t^W \right]^2 \right] \end{aligned}$$

$$\begin{aligned}
&\leq 4C^2 \mathbf{E} \left[\left(\int_r^T \left(\sup_{u \in [s, T]} |Y_u - \bar{Y}_u| + |Z_s - \bar{Z}_s| \right) ds \right)^2 \right] \\
&\leq 4C^2 (T-r) \mathbf{E} \left[\int_r^T \left(\sup_{u \in [s, T]} |Y_u - \bar{Y}_u| + |Z_s - \bar{Z}_s| \right)^2 ds \right] \\
&\leq 8C^2 (T-r) \left(\int_r^T \mathbf{E} \left[\sup_{u \in [s, T]} |Y_u - \bar{Y}_u|^2 \right] ds + \mathbf{E} \left[\int_r^T |Z_s - \bar{Z}_s|^2 ds \right] \right).
\end{aligned}$$

Consequently, for $T-r \leq 1$, we get

$$\begin{aligned}
&\mathbf{E} \left[\sup_{t \in [r, T]} |Y_t^1 - \bar{Y}_t^1| \right] \\
&\leq 8C^2 (T-r) \left(\mathbf{E} \left[\sup_{u \in [r, T]} |Y_u - \bar{Y}_u|^2 \right] + \mathbf{E} \left[\int_r^T |Z_s - \bar{Z}_s|^2 ds \right] \right). \quad (2.6)
\end{aligned}$$

Next we give a corresponding estimate for Z . Using (2.5) for $t = r$ and again assuming that $T-r \leq 1$, we conclude

$$\begin{aligned}
&\mathbf{E} \left[\int_r^T |Z_s - \bar{Z}_s|^2 ds \right] \leq \mathbf{E} [|Y_r - \bar{Y}_r|^2] + \mathbf{E} \left[\int_r^T |Z_s - \bar{Z}_s|^2 ds \right] \\
&= \mathbf{E} \left[\left| \int_r^T \left\{ f(\cdot, W, Y, Z) - f(\cdot, W, \bar{Y}, \bar{Z}) \right\}_s^{\mathbb{F}^W} ds \right|^2 \right] \\
&\leq (T-r) \mathbf{E} \left[\int_r^T \left| \left\{ f(\cdot, W, Y, Z) - f(\cdot, W, \bar{Y}, \bar{Z}) \right\}_s^{\mathbb{F}^W} \right|^2 ds \right] \\
&\leq (T-r) \mathbf{E} \left[\int_r^T \left\{ |f(\cdot, W, Y, Z) - f(\cdot, W, \bar{Y}, \bar{Z})|^2 \right\}_s^{\mathbb{F}^W} ds \right] \\
&\leq (T-r) \mathbf{E} \left[\int_r^T |f(\cdot, W, Y, Z) - f(\cdot, W, \bar{Y}, \bar{Z})|^2 ds \right] \\
&\leq C^2 (T-r) \mathbf{E} \left[\int_r^T \left(\sup_{u \in [s, T]} |Y_u - \bar{Y}_u| + \int_s^T |Z_u - \bar{Z}_u| \mu(s, du) \right)^2 ds \right] \\
&\leq 2C^2 (T-r) \mathbf{E} \left[\int_r^T \left(\sup_{u \in [s, T]} |Y_u - \bar{Y}_u|^2 + \left(\int_s^T |Z_u - \bar{Z}_u| \mu(s, du) \right)^2 \right) ds \right] \\
&\leq 2C^2 (T-r) \mathbf{E} \left[\sup_{u \in [r, T]} |Y_u - \bar{Y}_u|^2 + \int_r^T \int_s^T |Z_u - \bar{Z}_u|^2 \mu(s, du) ds \right] \\
&\leq 2C^2 (T-r) \left(\mathbf{E} \left[\sup_{u \in [r, T]} |Y_u - \bar{Y}_u|^2 \right] + \mathbf{E} \left[\int_r^T |Z_s - \bar{Z}_s|^2 ds \right] \right).
\end{aligned}$$

Together with (2.6) this implies

$$\begin{aligned}
&\mathbf{E} \left[\sup_{t \in [r, T]} |Y_t^1 - \bar{Y}_t^1| \right] + \mathbf{E} \left[\int_r^T |Z_s - \bar{Z}_s|^2 ds \right] \\
&\leq 10C^2 (T-r) \left(\mathbf{E} \left[\sup_{u \in [r, T]} |Y_u - \bar{Y}_u|^2 \right] + \mathbf{E} \left[\int_r^T |Z_s - \bar{Z}_s|^2 ds \right] \right). \quad (2.7)
\end{aligned}$$

We now choose $\delta > 0$ such that $\delta < \frac{1}{10}C^{-2}$ and, for some $N \geq 1$, $N\delta = T$. For $r = T - \delta$, (2.7) is the desired contraction property on $[T - \delta, T]$.

3) Now let (Y^δ, Z^δ) be a unique solution of BSDE (2.4) on $[T - \delta, T]$. We consider the subspace of $\mathcal{S}_{\mathbb{F}^W, \mathbf{R}^m}^2 \times L_{\mathbb{F}^W, \mathbf{R}^m \times d}^2$ of all (Y, Z) which coincide with (Y^δ, Z^δ) on $[T - \delta, T]$. Proceeding precisely as in step 2), it can be verified that, for all (Y, Z) and $(\bar{Y}, \bar{Z})$ from this subspace,

$$\begin{aligned} & \mathbf{E} \left[\sup_{t \in [T-2\delta, T]} |Y_t^1 - \bar{Y}_t^1| \right] + \mathbf{E} \left[\int_{T-2\delta}^T |Z_s - \bar{Z}_s|^2 ds \right] \\ & \leq 10C^2\delta \left(\mathbf{E} \left[\sup_{u \in [T-2\delta, T]} |Y_u - \bar{Y}_u|^2 \right] + \mathbf{E} \left[\int_{T-2\delta}^T |Z_s - \bar{Z}_s|^2 ds \right] \right) \end{aligned}$$

with the same constant $10C^2\delta < 1$ as in (2.7). This allows one to prove that (Y^δ, Z^δ) can be extended uniquely to a solution of BSDE (2.4) on the interval $[T - 2\delta, T]$. A repeated application of this step leads to a unique strong solution (Y, Z) of BSDE (2.4) on $[0, T]$. Proposition 2.2 is proved.

One may wonder which are the submarkovian kernels $\mu(s, A)$ on $[0, T] \times \mathcal{B}([0, T])$ which satisfy the condition stated in Proposition 2.2. Here the property that μ is submarkovian is not important at all, it could be replaced by the boundedness condition

$$\mu(s, [0, T]) \leq K, \quad s \in [0, T].$$

Example 2.1. Every of the following kernels μ satisfy the relation 2.2 in (B.ii) and, therefore, can serve as example for Proposition 2.2:

(i) $\mu(s, \cdot) = \varepsilon_s$, $s \in [0, T]$, where ε_s denotes the Dirac measure in $s \in [0, T]$. (In this case, $f(s, x, y, z)$ depends on z only through z_s and is Lipschitz continuous in this variable.)

(ii) Let $u_1, \dots, u_n \in [0, T]$ be fixed and set

$$\mu(s, \cdot) = \frac{1}{n} \sum_{1 \leq \ell \leq n, s+u_\ell \leq T} \varepsilon_{s+u_\ell}, \quad s \in [0, T].$$

(iii) With the Lebesgue measure λ on $[0, T]$, suppose

$$\mu(s, \cdot) \leq \frac{1}{T} \lambda, \quad s \in [0, T], \quad \lambda\text{-a.e.}$$

(iv) More generally, let m be an arbitrary probability measure on $[0, T]$ and define m_s as the image measure on $\mathbf{R}$ by the s -translation:

$$m_s(B) = m((B - s) \cap [0, T]), \quad B \in \mathcal{B}(\mathbf{R}), \quad s \in [0, T].$$

We now put

$$\mu(s, A) = m_s(A \cap [0, T]), \quad A \in \mathcal{B}(\mathbf{R}), \quad s \in [0, T].$$

The second term of the right-hand side of relation (2.3) in (B.ii) then transforms into

$$\int_s^T |z(u) - z'(u)| \mu(s, du) = \int_0^{T-s} |z(s+u) - z'(s+u)| m(du).$$

(v) Let $(K, \mathcal{K}, \kappa)$ be a probability space and $\mu(k, s, A)$, $(k, s) \in K \times [0, T]$, $A \in \mathcal{B}([0, T])$ a uniformly bounded kernel from $K \times [0, T]$ into $[0, T]$ such that, for every $k \in K$, $\mu(k, \cdot, \cdot)$ satisfies relation (2.2) in (B.ii). Then the kernel formed by «convex combination»

$$\mu(s, A) = \int_K \mu(k, s, A) \kappa(dk), \quad s \in [0, T], \quad A \in \mathcal{B}([0, T]),$$

possesses this property, too.

3. Tsirelson type examples. In his famous example, in 1975 B. Tsirelson [28] gave a (forward) SDE for which, although a weak solution can be constructed, there does not exist any strong solution (also see [23, IX.3.6]). In this section, we adapt the Tsirelson equation to our framework in order to get a BSDE which admits a weak solution which is not strong. Moreover, there is even an infinite number of weak (but not strong) solutions with different law. But, it may be surprising that, contrary to the Tsirelson SDE, one can also find a strong solution.

Following the idea of the construction of Tsirelson's example, we restrict ourselves to one-dimensional equations ($d = m = 1$) and choose a strictly decreasing sequence of real numbers $T := 1 = t_0 > t_1 > \dots > t_n > \dots > 0$ such that $t_n \rightarrow 0$ as $n \rightarrow +\infty$. For every $x \in D(\mathbf{R}) := D([0, 1]; \mathbf{R})$ we put

$$b(t, x) = \left\{ \frac{x(t_j) - x(t_{j+1})}{t_j - t_{j+1}} \right\}, \quad t_j < t \leq t_{j-1}, \quad j \geq 1, \quad (3.1)$$

and $b(0, x) = 0$, where $\{r\}$ denotes the fractional part of the real r . Obviously, $\{r\}$ is always a number from $[0, 1)$. It is well known from the theory of forward SDE that the equation

$$X_t = W_t + \int_0^t b(s, X) ds, \quad t \in [0, T], \quad (3.2)$$

driven by a standard Brownian motion W admits a weak solution (cf. [23, IX.3.6]), i.e., there are a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$ endowed with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$ satisfying the usual conditions, an $\mathbb{F}$ -Brownian motion W and an $\mathbb{F}$ -adapted process X , both defined on $(\Omega, \mathcal{F}, \mathbf{P})$, such that (3.2) is satisfied $\mathbf{P}$ -a.s. However, as Tsirelson has shown, this equation does not admit a strong solution. In the sequel, we will make use of the following properties of the filtrations $\mathbb{F}^W$ and $\mathbb{F}^X$ generated by W and X , respectively (see, e.g., [23, IX.3.6]).

Lemma 3.1. *For every $k \geq 1$, the random variable $U_k = \{(X_{t_k} - X_{t_{k+1}})/(t_k - t_{k+1})\}$ is independent of $\mathcal{F}_1^W$ and uniformly distributed over $[0, 1]$. Moreover,*

$$\mathcal{F}_t^X = \mathcal{F}_t^W \vee \sigma(U_k) \quad \text{for all } 0 < t_k \leq t.$$

We now transform the forward SDE (3.2) into a BSDE. To this end, we define the process $\tilde{Y} = X - W$ that satisfies the SDE

$$\tilde{Y}_t = \int_0^t b(s, \tilde{Y} + W) ds, \quad t \in [0, 1],$$

and we put $Y_t = \tilde{Y}_{1-t}$, $t \in [2^{-1}, 1]$. Obviously, $(Y_t)_{t \in [2^{-1}, 1]}$ is a continuous process consisting of random variables Y_t which are $\mathcal{F}_{1-t}(\subset \mathcal{F}_{2^{-1}} \subset \mathcal{F}_t)$ -measurable, for all $t \in [2^{-1}, 1]$, and it satisfies the equation

$$Y_t = \int_t^1 b(s, \tilde{Y} + W) ds, \quad t \in [2^{-1}, 1].$$

We now extend the process Y onto the whole time interval $[0, 1]$, putting

$$Y_t = \mathbf{E}[Y_{2^{-1}} | \mathcal{F}_t^W], \quad t \in [0, 2^{-1}).$$

Finally, the functional f , the driver of the BSDE we are constructing, is defined as follows:

$$f(s, w, y) = \begin{cases} b(1-s, \tilde{y} + w), & s \in [2^{-1}, 1], \\ 0, & s \in [0, 2^{-1}), \end{cases}$$

for every $w, y \in D(\mathbf{R})$, where $\tilde{y}$ is defined by $\tilde{y}(s) = y((1-s)-)$, $s \in [0, 1]$, $\tilde{y}(1) = \tilde{y}(1-)$. (Here we used the notation $y(u-) = \lim_{u > r \uparrow u} y(r)$, $u \in (0, 1]$, $y(0-) = 0$, for every $y \in D(\mathbf{R})$.)

Now we collect some properties of f . For this we remark that, for every $s \in [2^{-1}, 1]$, we have $1-s \leq s$.

i) For every $s \in [0, 1]$, $f(s, \cdot, \cdot)$ is measurable and bounded on the product space $(D(\mathbf{R}) \times D(\mathbf{R}), \mathcal{D}_s^0(\mathbf{R}) \otimes \mathcal{D}_1^s(\mathbf{R}))$.

ii) For every $(w, y) \in D(\mathbf{R}) \times D(\mathbf{R})$, $f(\cdot, w, y)$ is right-continuous on $[0, 1]$.

iii) f is measurable on the product space $([0, 1] \times D(\mathbf{R}) \times D(\mathbf{R}), \mathcal{B}([0, 1]) \otimes \mathcal{D}(\mathbf{R}) \otimes \mathcal{D}(\mathbf{R}))$.

Property iii) already follows from i)-ii) by standard arguments. This is more than we need for the BSDE. In particular, for every $y \in D(\mathbf{R})$, the process $f(\cdot, W, y)$ is $\mathbb{F}^W$ -adapted and right-continuous and, hence is $\mathbb{F}^W$ -optional.

We now introduce the filtration $\mathbb{G} = (\mathcal{G}_t)_{t \in [0, 1]}$ by

$$\mathcal{G}_t = \begin{cases} \mathcal{F}_t^W, & t \in [0, 2^{-1}), \\ \mathcal{F}_t & t \in [2^{-1}, 1]. \end{cases}$$

Proposition 3.1. (i) *The six-tuple $(\Omega, \mathcal{F}, \mathbb{G}, \mathbf{P}, W, Y)$ is a weak solution of BSDE(0, f, μ), where μ denotes the Wiener measure (lifted on $D(\mathbf{R})$).*
(ii) *The process Y on $(\Omega, \mathcal{F}, \mathbf{P})$ is no strong solution of BSDE(0, f, W).*

P r o o f. From the definition of f , it can easily be seen that the process Y solves the equation

$$Y_t = \begin{cases} \int_t^1 f(s, W, Y) ds, & t \in [2^{-1}, 1], \\ \mathbf{E}[Y_{2^{-1}} | \mathcal{F}_t^W] = \mathbf{E}\left[\int_t^1 f(s, W, Y) ds \middle| \mathcal{F}_t^W\right], & t \in [0, 2^{-1}). \end{cases}$$

Obviously, the process Y is $\mathbb{G}$ -adapted and, consequently,

$$Y_t = \mathbf{E}\left[\int_t^1 f(s, W, Y) ds \middle| \mathcal{G}_t\right], \quad t \in [0, 1].$$

To finish the proof of (i), it still remains to verify condition (A.iv) that every $\mathbb{F}^W$ -martingale on $(\Omega, \mathcal{F}, \mathbf{P})$ is also a $\mathbb{G}$ -martingale. But W is an $\mathbb{F}$ -Brownian motion and $\mathbb{F}^W \subset \mathbb{G} \subset \mathbb{F}$. The assertion now follows from the martingale representation property of $(W, \mathbb{F}^W)$.

For proving the assertion (ii), we assume that Y would be a strong solution of $\text{BSDE}(0, f, W)$. Then Y would be $\mathbb{F}^W$ -adapted and, consequently, $X_{1-t} = Y_t + W_{1-t}$ would be $\mathcal{F}_1^W$ -measurable for all $t \in [\frac{1}{2}, 1]$. In particular, $\{U_k\}$ would be $\mathcal{F}_1^W$ -measurable for all k such that $t_k \leq 2^{-1}$. But this is impossible, since these random variables are uniformly distributed over $[0, 1]$ (and, hence, they are not constant) and independent of $\mathcal{F}_1^W$, because of Lemma 3.1.

To give a more detailed description of $\text{BSDE}(0, f, \mu)$, we are now going to show that it does not have the property of uniqueness in law, and that we can even construct a strong solution. To this end, let us consider the weak solution X of equation (3.2) defined on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$ with an $\mathbb{F}$ -Brownian motion W . (Note that X is unique in law.) Starting from W , $k \geq 1$, and $\lambda \in [0, 1]$, we construct a new process W^λ as follows:

$$W_t^\lambda = \begin{cases} W_t, & t \in [0, t_k], \\ W_t - \frac{t - t_k}{t_{k-1} - t_k} (W_{t_{k-1}} - W_{t_k}) + \frac{t - t_k}{t_{k-1} - t_k} \\ \quad \times (\sqrt{1 - \lambda} (W_{t_{k-1}} - W_{t_k}) + \sqrt{\lambda} \Phi_{0, t_{k-1} - t_k}^{-1}(U_k)), & t \in [t_k, t_{k-1}], \\ W_{t_{k-1}}^\lambda + W_t - W_{t_{k-1}}, & t \in [t_{k-1}, 1]. \end{cases} \quad (3.3)$$

Here $\Phi_{0,t}$ denotes the distribution function of the Gaussian law with mean zero and variance t , and U_k is defined in Lemma 3.1. Obviously, the increments of W^λ coincide with those of W outside (t_k, t_{k-1}) . Furthermore, we have

$$W_{t_{k-1}}^\lambda - W_{t_k}^\lambda = \sqrt{1 - \lambda} (W_{t_{k-1}} - W_{t_k}) + \sqrt{\lambda} \Phi_{0, t_{k-1} - t_k}^{-1}(U_k). \quad (3.4)$$

Lemma 3.2. *For every $k \geq 1$ and $\lambda \in [0, 1]$, W^λ is a Brownian motion.*

P r o o f. For better understanding, note that the process written in the first line of the definition of $W_t^\lambda - W_{t_k}^\lambda$, $t \in [t_k, t_{k-1}]$, is a Brownian bridge

on $[t_k, t_{k-1}]$. Also note that, in view of Lemma 3.1, U_k is independent of W and $\Phi_{0, t_{k-1}-t_k}^{-1}(U_k)$ has a normal distribution with mean zero and variance $t_{k-1} - t_k$. It is sufficient to verify that $\widetilde{W}_t^\lambda = W_{t+t_k}^\lambda - W_{t_k}^\lambda$, $t \in [0, t_{k-1} - t_k]$, is a Brownian motion. But, clearly, this is a continuous Gaussian process with mean zero, and the covariance function $\text{cov}(\widetilde{W}_s^\lambda, \widetilde{W}_t^\lambda)$ can be computed as $\min\{s, t\}$.

To the Brownian motion W^λ , we now associate the process X^λ defined by

$$X_t^\lambda = \begin{cases} X_t, & t \in [0, t_k], \\ X_{t_\ell}^\lambda + (W_t - W_{t_\ell}) + (t - t_\ell) \left\{ \frac{X_{t_\ell}^\lambda - X_{t_{\ell+1}}^\lambda}{t_\ell - t_{\ell+1}} \right\}, & t \in [t_\ell, t_{\ell+1}], \quad 1 \leq \ell \leq k. \end{cases} \quad (3.5)$$

Obviously, the process X^λ satisfies the equation

$$X_t^\lambda = W_t^\lambda + \int_0^t b(s, X_s^\lambda) ds, \quad t \in [0, 1], \quad \mathbf{P}\text{-a.s.} \quad (3.6)$$

However, in view of (3.4) and $U_k = \{(X_{t_k} - X_{t_{k+1}})/(t_k - t_{k+1})\}$, $W_{t_{k-1}}^\lambda - W_{t_k}^\lambda$ is not independent of $\mathcal{F}_{t_k}^{X^\lambda} = \mathcal{F}_{t_k}^X$. Consequently, W^λ cannot be an $\mathbb{F}^{X^\lambda}$ -Brownian motion and, therefore, X^λ is not a weak solution of SDE (3.2) in the classical sense, but only in the generalized sense according to the definition below.

Definition 3.1. Let $\widetilde{\mathbb{F}}$ -adapted processes $\widetilde{X}$, $\widetilde{W}$, where $\widetilde{W}$ is a Brownian motion, be given on a filtered probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{F}}, \widetilde{\mathbf{P}})$ (with the usual conditions). Then we call $\widetilde{X}$ a weak solution of SDE (3.2) in the generalized sense if $\widetilde{X}$ and $\widetilde{W}$ satisfy (3.2), i.e.,

$$\widetilde{X}_t = \widetilde{W}_t + \int_0^t b(s, \widetilde{X}) ds, \quad t \in [0, 1], \quad \mathbf{P}\text{-a.s.}$$

It is remarkable that for weak solutions of SDE (3.2) in the generalized sense the uniqueness in law is lost, as the following proposition shows.

Proposition 3.2. For different $k \geq 1$ or for different $\lambda \in (0, 1]$, the laws of X^λ are different, and they are different from the law of X .

Proof. The statement is obvious in the case of different k . Let us verify the statement for the same $k \geq 1$ and different $\lambda \in [0, 1]$. Using (3.3), (3.4) and (3.5), we obtain

$$X_{t_{k-1}}^\lambda - X_{t_k}^\lambda = \sqrt{1-\lambda}(W_{t_{k-1}} - W_{t_k}) + \sqrt{\lambda}\Phi_{0, t_{k-1}-t_k}^{-1}(U_k) + (t_{k-1} - t_k)U_k$$

and, given $U_k = u$, the increment $X_{t_{k-1}}^\lambda - X_{t_k}^\lambda$ follows a normal distribution with mean $\sqrt{\lambda}\Phi_{0, t_{k-1}-t_k}^{-1}(u) + (t_{k-1} - t_k)u$ and variance $(1-\lambda)(t_{k-1} - t_k)$.

We now use this family of weak solutions in the generalized sense in order to construct a family of weak solutions of BSDE(0, f , μ). For this we

assume $t_{k-1} \leq 2^{-1}$ and follow the same scheme as developed above for X . We introduce the process $\tilde{Y}^\lambda = X^\lambda - W^\lambda$, and put

$$Y_t^\lambda = \begin{cases} \tilde{Y}_{1-t}^\lambda, & t \in [2^{-1}, 1], \\ \mathbf{E}[\tilde{Y}_{2^{-1}}^\lambda | \mathcal{F}_t^{W^\lambda}], & t \in [0, 2^{-1}). \end{cases}$$

We also define the filtration $\mathbb{G}^\lambda = (\mathcal{G}_t^\lambda)_{t \in [0,1]}$ by

$$\mathcal{G}_t^\lambda = \begin{cases} \mathcal{F}_t^{X^\lambda}, & t \in [2^{-1}, 1], \\ \mathcal{F}_t^{W^\lambda}, & t \in [0, 2^{-1}). \end{cases}$$

Proposition 3.3. (i) For all $\lambda \in [0, 1]$, the six-tuple $(\Omega, \mathcal{F}, \mathbb{G}^\lambda, \mathbf{P}, W^\lambda, Y^\lambda)$ is a weak solution of BSDE(0, f, μ).

(ii) If $\lambda, \lambda' \in [0, 1]$ with $\lambda \neq \lambda'$, then the laws of Y^λ and $Y^{\lambda'}$ are different.

P r o o f. The proof of (i) follows the same lines as above for the case $\lambda = 0$. We have only to take care that condition (A.iv) is fulfilled. For this it is sufficient that W^λ is a $\mathbb{G}^\lambda$ -Brownian motion or, equivalently, that $W_t^\lambda - W_s^\lambda$ is independent of $\mathcal{G}_s^\lambda$ for all $0 \leq s < t \leq 1$. This is clear for $s < 2^{-1}$, because $\mathcal{G}_s^\lambda = \mathcal{F}_s^{W^\lambda}$ in this case. If $s \geq 2^{-1}$ ($\geq t_{k-1}$), then $W_t^\lambda - W_s^\lambda = W_t - W_s$ which is independent of $\mathcal{F}_s^X$. But

$$\mathcal{F}_s^{X^\lambda} \subset \mathcal{F}_s^X, \quad s \in [t_{k-1}, 1], \quad (3.7)$$

from which the assertion follows. To verify (3.7), we remark that we have the relations

$$\begin{aligned} \mathcal{F}_s^{X^\lambda} &= \mathcal{F}_s^{W^\lambda} \vee \sigma \left(\left\{ \frac{X_{t_k}^\lambda - X_{t_{k+1}}^\lambda}{t_k - t_{k+1}} \right\} \right) \\ &= \mathcal{F}_s^{W^\lambda} \vee \sigma \left(\left\{ \frac{X_{t_k} - X_{t_{k+1}}}{t_k - t_{k+1}} \right\} \right) \subset \mathcal{F}_s^W \vee \sigma \left(\left\{ \frac{X_{t_k} - X_{t_{k+1}}}{t_k - t_{k+1}} \right\} \right) = \mathcal{F}_s^X \end{aligned}$$

for all $s \in [t_{k-1}, 1]$ (cf. Lemma 3.1 for the last equality; the first equality can be got in the same manner as the last equality, cf., [23, IX.3.6]).

For proving (ii), we use (3.6) and observe that

$$\begin{aligned} Y_{1-t_{k-1}}^\lambda - Y_{1-t_{k-2}}^\lambda &= (X_{t_{k-1}}^\lambda - X_{t_{k-1}}^\lambda) - (W_{t_{k-1}}^\lambda - W_{t_{k-1}}^\lambda) \\ &= (t_{k-1} - t_{k-2}) \left\{ \frac{X_{t_{k-1}}^\lambda - X_{t_k}^\lambda}{t_{k-1} - t_k} \right\}. \end{aligned}$$

By Proposition 3.2 (more precisely, by its proof) it follows that this increment has different laws for different λ .

Proposition 3.4. The process Y^1 on $(\Omega, \mathcal{F}, \mathbf{P})$ is even a strong solution of BSDE(0, f, W).

P r o o f. It is sufficient to verify that the filtrations $\mathbb{G}^1$ and $\mathbb{F}^{W^1}$ coincide. By definition, $\mathcal{G}_t^1 = \mathcal{F}_t^{W^1}$ for all $t \in [0, 2^{-1})$. On the other hand, the

inclusion $\mathcal{F}_t^{W^1} \subset \mathcal{F}_t^{X^1} = \mathcal{G}_t^1$, $t \in [2^{-1}, 1]$, is obvious. The converse inclusions follow from $\mathcal{G}_t^1 = \mathcal{F}_t^{X^1}$, $t \in [2^{-1}, 1]$, and

$$\mathcal{F}_t^{X^1} = \mathcal{F}_t^{W^1} \vee \sigma \left(\left\{ \frac{X_{t_k}^1 - X_{t_{k+1}}^1}{t_k - t_{k+1}} \right\} \right) = \mathcal{F}_t^{W^1} \vee \sigma \left(\left\{ \frac{X_{t_k} - X_{t_{k+1}}}{t_k - t_{k+1}} \right\} \right) \subset \mathcal{F}_t^{W^1}$$

for all $t \in [t_{k-1}, 1]$. (Note that $t_{k-1} \leq 2^{-1}$.) The last inclusion is a consequence of

$$W_{t_{k-1}}^1 - W_{t_k}^1 = \Phi_{0, t_{k-1} - t_k}^{-1} \left(\left\{ \frac{X_{t_k} - X_{t_{k+1}}}{t_k - t_{k+1}} \right\} \right),$$

as special case of (3.4) for $\lambda = 1$.

Remark 3.1. (i) For $t_k \leq t < t_{k-1}$, $\{(X_{t_k} - X_{t_{k+1}})/(t_k - t_{k+1})\}$ is not $\mathcal{F}_t^{W^1}$ -measurable and, hence, $\mathcal{F}_t^{W^1} \subsetneq \mathcal{F}_t^{X^1}$, $t \in [t_k, t_{k-1})$. The last relation is also valid for $t \in [0, t_k)$, because W^1 and X^1 coincide with W and X , respectively, on $[0, t_k]$.

(ii) For $\lambda \in (0, 1)$, it is not difficult to understand that $\{(X_{t_k}^\lambda - X_{t_{k+1}}^\lambda)/(t_k - t_{k+1})\} = \{(X_{t_k} - X_{t_{k+1}})/(t_k - t_{k+1})\}$ is not $\mathcal{F}_1^{W^\lambda}$ -measurable. Concluding as in the proof of Proposition 3.1, we can observe that Y^λ is not a strong solution of BSDE(0, f , W) on $(\Omega, \mathcal{F}, \mathbf{P})$, as for the case $\lambda = 0$.

(iii) For $\lambda \neq 1$, it can be shown that the filtration $(\mathcal{F}_t^{X^\lambda})_{t \in [t_{k-1}, 1]}$ does not depend on λ . We always have

$$\mathcal{F}_t^{X^\lambda} = \mathcal{F}_t^W \vee \sigma \left(\left\{ \frac{X_{t_k} - X_{t_{k+1}}}{t_k - t_{k+1}} \right\} \right), \quad t \in [t_{k-1}, 1],$$

and therefore

$$\mathcal{F}_t^{X^\lambda} = \mathcal{F}_t^X, \quad t \in [t_{k-1}, 1], \quad \lambda \in [0, 1).$$

4. Existence of weak solutions. In this section, we shall prove a general existence theorem for weak solutions of equation BSDE(H, f, μ) (cf. Definition 1.2). We assume that ν is an arbitrary probability measure on $(D(\mathbf{R}^d), \mathcal{D}(\mathbf{R}^d))$. Without mentioning it every time, the terminal functional H and the driver f always satisfy the conditions (H.1)–(H.2). Furthermore, we assume that H is μ -integrable:

$$\int_{D(\mathbf{R}^d)} |H(x)| \mu(dx) < +\infty. \quad (4.1)$$

To begin with, we establish the connection of solutions of BSDE(H, f, μ) to quasimartingales.

Lemma 4.1. *Let $(\Omega, \mathcal{F}, \mathbb{G}, \mathbf{P}, X, Y)$ be a weak solution of BSDE(H, f, μ). Then $(Y, \mathbb{G})$ is a quasimartingale with conditional variation*

$$V(Y) \leq \mathbf{E} \left[|H(X)| + \int_0^T |f(s, X, Y)| ds \right] < +\infty.$$

P r o o f. Given a subdivision $\tau: 0 = t_0 < t_1 < \dots < t_n = T$ we get

$$\begin{aligned} V_\tau(Y) &:= \mathbf{E}|Y_T| + \sum_{k=0}^{n-1} \mathbf{E} \left| \mathbf{E} \left[Y_{t_{k+1}} - Y_{t_k} \mid \mathcal{G}_{t_k} \right] \right| \\ &= \mathbf{E}|Y_T| + \sum_{k=0}^{n-1} \mathbf{E} \left| \mathbf{E} \left[\int_{t_k}^{t_{k+1}} f(s, X, Y) ds \mid \mathcal{G}_{t_k} \right] \right| \\ &\leq \mathbf{E}|Y_T| + \mathbf{E} \left[\int_0^T |f(s, X, Y)| ds \right] < +\infty. \end{aligned}$$

Noting that $V(Y) = \sup_\tau V_\tau(Y)$, the result follows.

Given any natural $p \geq 1$, we introduce the following metric ρ on $D(\mathbf{R}^p)$:

$$\rho(x, y) = \int_0^T (|x(s) - y(s)| \wedge 1) ds, \quad x, y \in D(\mathbf{R}^p).$$

The topology induced by this metric is the Meyer–Zheng topology on $D(\mathbf{R}^p)$ (cf. [20]). The convergence in the metric ρ is just the convergence in Lebesgue measure. More precisely, let h be an $\mathbf{R}^p$ -valued Borel function on $[0, T]$. The pseudo-path $\psi(h)$ of h is defined as the image probability law on $[0, T] \times \mathbf{R}^p$ of the uniform distribution on $[0, T]$ by the mapping $t \rightarrow (t, h(t))$. The mapping ψ defines an embedding of $D(\mathbf{R}^p)$ in the compact metric space $\mathcal{P}$ of all probability measures on $[0, T] \times \bar{\mathbf{R}}^p$ (where $\bar{\mathbf{R}} = [-\infty, +\infty]$). The topology induced on $D(\mathbf{R}^p)$ by this embedding is the Meyer–Zheng topology. This means that $D(\mathbf{R}^p)$ is considered as a topological subspace of $\mathcal{P}$. Clearly, $(D(\mathbf{R}^p), \rho)$ is a separable metric space, but it is not complete. However, $(D(\mathbf{R}^p), \rho)$ is a Lusin space: For every embedding in a compact metric space K , $D(\mathbf{R}^p)$ is a Borel set in K ; cf. [7, IV.40–46] and [20] for details.

In the sequel, our main tool is weak convergence in the Meyer–Zheng topology of sequences of laws corresponding to solutions of BSDE(H, f, μ). To prepare the main theorem, we start with the following two lemmata.

Lemma 4.2. *Let $\mathcal{B}(D(\mathbf{R}^p))$ be the σ -algebra of Borel subsets of $D(\mathbf{R}^p)$ in the Meyer–Zheng topology. Then $\mathcal{B}(D(\mathbf{R}^p)) = \mathcal{D}(\mathbf{R}^p)$.*

P r o o f. We recall that $\mathcal{D}(\mathbf{R}^p)$ is generated by the coordinate mappings ζ_t , $t \in [0, T]$. It is clear that the mapping $x \mapsto \rho(x, a)$, where $a \in D(\mathbf{R}^p)$ is fixed, is $\mathcal{D}(\mathbf{R}^p)$ -measurable. (Approximate $x, a \in D(\mathbf{R}^p)$ by step functions x_n, a_n .) Hence, every ball in the metric ρ belongs to $\mathcal{D}(\mathbf{R}^p)$ and, since $(D(\mathbf{R}^p), \rho)$ is separable, $G \in \mathcal{D}(\mathbf{R}^p)$ for every ρ -open set G . This proves $\mathcal{B}(D(\mathbf{R}^p)) \subset \mathcal{D}(\mathbf{R}^p)$. For proving the converse, we define

$$\zeta_t^n(x) = \frac{1}{n} \int_t^{T \wedge (t + \frac{1}{n})} \zeta_s(x) ((n - |\zeta_s(x)|)^+ \wedge 1) ds, \quad x \in D(\mathbf{R}^p), \quad t \in [0, T].$$

We then have that ζ_t^n is ρ -continuous and hence $\mathcal{B}(D(\mathbf{R}^p))$ -measurable. This yields that ζ_t , as pointwise limit of ζ_t^n as $n \rightarrow +\infty$, is $\mathcal{B}(D(\mathbf{R}^p))$ -measurable for all $t \in [0, T]$. Finally, from the definition of $D(\mathbf{R}^p)$, it follows that also

$\zeta_T = \lim_{n \rightarrow +\infty} \zeta_{T-1/n}$ (pointwise) is $\mathcal{B}(D(\mathbf{R}^p))$ -measurable. Consequently, $\mathcal{D}(\mathbf{R}^p) \subset \mathcal{B}(D(\mathbf{R}^p))$.

Lemma 4.2 shows that the Meyer–Zheng topology is consistent with the given measurable structure on $D(\mathbf{R}^p)$.

Lemma 4.3. *For every $n \geq 1$, let (H^n, f^n, μ^n) satisfy the conditions stated at the beginning of this section. Suppose that*

- i) *there exists a probability measure μ on $(D(\mathbf{R}^d), \mathcal{D}(\mathbf{R}^d))$ such that μ^n converges weakly in the Meyer–Zheng topology to μ as $n \rightarrow +\infty$;*
- ii) *for all $n \geq 1$, there exists a weak solution $(\Omega^n, \mathcal{F}^n, \mathbb{G}^n, \mathbf{P}^n, X^n, Y^n)$ of equation BSDE(H^n, f^n, μ^n) such that*

$$\sup_{n \geq 1} \mathbf{E}_{\mathbf{P}^n} [|H(X^n)| + \int_0^T |f(s, X^n, Y^n)| ds] < +\infty. \quad (4.2)$$

Let $Q^n := P_{(X^n, Y^n)}^n$ be the distribution of (X^n, Y^n) on $D(\mathbf{R}^d) \times D(\mathbf{R}^m)$ with respect to $\mathbf{P}^n$. Then there exist a subsequence $(Q^{n_k})_{k \geq 1}$ of $(Q^n)_{n \geq 1}$ and a probability law Q on $D(\mathbf{R}^d) \times D(\mathbf{R}^m)$ such that Q^{n_k} converges weakly in the Meyer–Zheng topology to Q as $k \rightarrow +\infty$. Moreover,

$$Q(A \times D(\mathbf{R}^m)) = \mu(A), \quad A \in \mathcal{D}(\mathbf{R}^d).$$

P r o o f. Since $D(\mathbf{R}^d)$ is a separable metric space, there exists a compact metric space K_1 such that $D(\mathbf{R}^d)$ is a subset of K_1 . As mentioned above, $D(\mathbf{R}^d)$ is a Lusin space and hence a Borel subset of K_1 : $D(\mathbf{R}^d) \in \mathcal{B}(K_1)$. Analogously, there is a compact metric space K_2 such that $D(\mathbf{R}^m)$ is a subspace of K_2 and $D(\mathbf{R}^m) \in \mathcal{B}(K_2)$. On the compact metric space $K = K_1 \times K_2$ we set

$$\tilde{Q}^n(A) = Q^n(A \cap (D(\mathbf{R}^d) \times D(\mathbf{R}^m))), \quad A \in \mathcal{B}(K).$$

Of course, $A \cap (D(\mathbf{R}^d) \times D(\mathbf{R}^m))$ belongs to $\mathcal{B}(D(\mathbf{R}^d)) \otimes \mathcal{B}(D(\mathbf{R}^m)) = \mathcal{D}(\mathbf{R}^d) \otimes \mathcal{D}(\mathbf{R}^m)$, the last equality being true in view of Lemma 4.2. The set of probability measures on the compact metric space K is compact in the weak convergence. Hence we can choose a subsequence $(n_k)_{k \geq 1}$ and a probability measure $\tilde{Q}$ on K such that $\tilde{Q}^{n_k} \rightarrow \tilde{Q}$ weakly on K . We now show that

$$\tilde{Q}(D(\mathbf{R}^d) \times D(\mathbf{R}^m)) = 1. \quad (4.3)$$

To this end, let $\tilde{Q}_2^{n_k}, \tilde{Q}_2$ be the marginal distributions with respect to the second coordinate:

$$\tilde{Q}_2^{n_k}(A_2) = \tilde{Q}^{n_k}(K_1 \times A_2), \quad \tilde{Q}_2(A_2) = \tilde{Q}(K_1 \times A_2), \quad A_2 \in \mathcal{B}(K_2).$$

We notice that $\tilde{Q}_2^{n_k}$ is the distribution of Y^{n_k} considered as a random variable with values in $(K_2, \mathcal{B}(K_2))$. Clearly, $\tilde{Q}_2^{n_k} \rightarrow \tilde{Q}_2$ weakly, as $k \rightarrow +\infty$. The proof of Theorem 4 in [20], in combination with Lemma 4.1 and condition (4.2), shows that

$$\tilde{Q}_2(D(\mathbf{R}^m)) = 1. \quad (4.4)$$

Next we consider the marginal distributions $\tilde{Q}_1^{n_k}$ and $\tilde{Q}_1$ of $\tilde{Q}^{n_k}$ and $\tilde{Q}$, respectively, with respect to the first coordinate:

$$\tilde{Q}_1^{n_k}(A_1) = \tilde{Q}^{n_k}(A_1 \times K_2), \quad \tilde{Q}_1(A_1) = \tilde{Q}(A_1 \times K_2), \quad A_1 \in \mathcal{B}(K_1).$$

Obviously, $\tilde{Q}_1^{n_k}$ is the law of X^{n_k} considered as a random variable in $(K_1, \mathcal{B}(K_1))$. We also have $\tilde{Q}_1^{n_k} \rightarrow \tilde{Q}_1$ weakly on K_1 as $k \rightarrow +\infty$. But $\tilde{Q}_1^{n_k} = \tilde{\mu}^{n_k}$, where

$$\tilde{\mu}^{n_k}(A) = \mu^{n_k}(A \cap D(\mathbf{R}^d)), \quad \tilde{\mu}(A) = \mu(A \cap D(\mathbf{R}^d)), \quad A \in \mathcal{B}(K_1).$$

Since $\mu^{n_k} \rightarrow \mu$ weakly on $D(\mathbf{R}^d)$ in the Meyer–Zheng topology, for all $\varphi \in C_b(K_1)$ we can conclude

$$\begin{aligned} \int_{K_1} \varphi d\tilde{Q}_1 &= \lim_{k \rightarrow +\infty} \int_{K_1} \varphi d\tilde{Q}_1^{n_k} = \lim_{k \rightarrow +\infty} \int_{K_1} \varphi d\tilde{\mu}^{n_k} = \lim_{k \rightarrow +\infty} \int_{D(\mathbf{R}^d)} \varphi d\mu^{n_k} \\ &= \int_{D(\mathbf{R}^d)} \varphi d\mu = \int_{K_1} \varphi d\tilde{\mu} \end{aligned}$$

and, therefore, $\tilde{Q}_1 = \tilde{\mu}$. This implies

$$\tilde{Q}_1(D(\mathbf{R}^d)) = 1. \quad (4.5)$$

From (4.4) and (4.5) it follows immediately that

$$\tilde{Q}(D(\mathbf{R}^d) \times K_2) = 1, \quad \tilde{Q}(K_1 \times D(\mathbf{R}^m)) = 1.$$

This, obviously, implies (4.3), proving the claim. Due to (4.3), we define Q as restriction of $\tilde{Q}$ to $D(\mathbf{R}^d) \times D(\mathbf{R}^m)$ and it follows that $Q^{n_k} \rightarrow Q$ weakly on $D(\mathbf{R}^d) \times D(\mathbf{R}^m)$ in the Meyer–Zheng topology (cf. [7, III.58]). This finishes the proof of the lemma.

Now we come to the main result of the present paper.

Theorem 4.1. *Let (H, f, μ) be a data set described at the beginning of this section. Additionally, suppose that the following conditions are satisfied:*

- (C.i) *f is bounded on $[0, T] \times D(\mathbf{R}^d) \times D(\mathbf{R}^m)$;*
- (C.ii) *for every $(t, x) \in [0, T] \times D(\mathbf{R}^d)$, $f(t, x, \cdot)$ is $\mathcal{D}_T^t(\mathbf{R}^m)$ -measurable;*
- (C.iii) *the mapping F on $D(\mathbf{R}^{d+m})$ defined by*

$$F(x, y) = f(\cdot, x, y), \quad (x, y) \in D(\mathbf{R}^{d+m}),$$

with values in $L^1([0, T]; \mathbf{R}^m)$ is continuous in the sense of Meyer–Zheng.

Then BSDE(H, f, μ) possesses a weak solution.

R e m a r k 4.1. (i) Condition (C.iii) means the following: If $z_n, z \in D(\mathbf{R}^{d+m})$ are such that

$$\lim_{n \rightarrow +\infty} \int_0^T (|z_n(s) - z(s)| \wedge 1) ds = 0,$$

then

$$\lim_{n \rightarrow +\infty} \int_0^T |f(s, z_n) - f(s, z)| ds = 0.$$

(ii) Condition (C.iii) is obviously satisfied if, ds -a.e. on $[0, T]$, the functional $f(s, \cdot, \cdot)$ on $D(\mathbf{R}^{d+m})$ is continuous with respect to the Meyer–Zheng topology.

(iii) Let φ be a measurable and bounded function on $[0, T] \times \mathbf{R}^d \times \mathbf{R}^m$ such that, for almost every $s \in [0, T]$, $\varphi(s, \cdot, \cdot)$ is continuous. We set

$$f(t, x, y) = \varphi(t, x(t), y(t)), \quad (t, x, y) \in [0, T] \times D(\mathbf{R}^{d+m}).$$

Then f satisfies the assumptions of the theorem.

We are now going to prove Theorem 4.1.

P r o o f. Let $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mathbf{P}) = (D(\mathbf{R}^d), \mathcal{D}^\mu(\mathbf{R}^d), \mu)$ and ζ be the identity map. Here $\mathcal{D}^\mu(\mathbf{R}^d)$ denotes the completion of $\mathcal{D}(\mathbf{R}^d)$ with respect to μ . As always, $\mathbb{F}^\zeta$ is the smallest of all filtrations $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$ such that the process ζ is $\mathbb{F}$ -adapted and $\mathbb{F}$ satisfies the usual conditions. For every $n \geq 1$, we consider the driver

$$f_n(s, x, y) = \begin{cases} f(s + n^{-1}, x, y), & s \in [0, T - n^{-1}], \\ 0, & s \in (T - n^{-1}, T], \end{cases} \quad (x, y) \in D(\mathbf{R}^{d+m}),$$

and we introduce $\text{BSDE}(H, f_n, \mu)$ which is «time advanced by n^{-1} »:

$$Y_t^n = \mathbf{E} \left[H(\zeta) + \int_t^T f_n(s, \zeta, Y^n) ds \middle| \mathcal{F}_t^\zeta \right], \quad t \in [0, T].$$

Because of (C.i), (C.ii) it is clear that $\text{BSDE}(H, f_n, \zeta)$ has a unique strong solution Y^n on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mathbf{P})$ (which can be computed step by step beginning with the interval $[T - n^{-1}, T]$). Furthermore,

$$\sup_{n \geq 1} \mathbf{E} \left[|H(\zeta)| + \int_0^T |f_n(s, \zeta, Y^n)| ds \right] \leq \mathbf{E} |H(\zeta)| + K T < +\infty,$$

where $K = \sup_{(s, x, y) \in [0, T] \times D(\mathbf{R}^{d+m})} |f(s, x, y)|$.

Consequently, denoting the law of (ζ, Y^n) on $D(\mathbf{R}^{d+m})$ (with respect to $\mathbf{P}$) by Q^n , from Lemma 4.3 it follows that there exist a probability measure Q on $D(\mathbf{R}^{d+m})$ and a subsequence $(n_k)_{k \geq 1}$ such that $Q^{n_k} \rightarrow Q$ weakly in the Meyer–Zheng topology. Moreover, we have $Q(A \times D(\mathbf{R}^m)) = \mu(A)$, $A \in \mathcal{D}(\mathbf{R}^d)$. For simplicity of the presentation, we assume $n_k = k$, $k \geq 1$.

Now we define the projections X and Y on $D(\mathbf{R}^{d+m})$:

$$X(x, y) = x, \quad Y(x, y) = y, \quad (x, y) \in D(\mathbf{R}^d) \times D(\mathbf{R}^m),$$

and by X_t, Y_t we denote the coordinate mappings

$$X_t(x, y) = x(t), \quad Y_t(x, y) = y(t), \quad (x, y) \in D(\mathbf{R}^d) \times D(\mathbf{R}^m),$$

for all $t \in [0, T]$. We consider the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}^{X, Y}, Q)$, where $\Omega = D(\mathbf{R}^{d+m})$ and $\mathcal{F} = \mathcal{D}^Q(\mathbf{R}^{d+m})$, the completion of $\mathcal{D}(\mathbf{R}^{d+m})$ with respect to Q . Note that the filtration $\mathbb{F}^{X, Y}$ generated by (X, Y) satisfies the usual conditions by definition.

We are now going to verify that the six-tuple $(\Omega, \mathcal{F}, \mathbb{F}^{X,Y}, Q, X, Y)$ is, indeed, a weak solution of BSDE(H, f, μ).

1) Let F be an arbitrary bounded and continuous function on $D(\mathbf{R}^{d+m})$. For every $t \in [0, T]$, we put

$$F(t, x, y) = F(x^t, y^t), \quad (x, y) \in D(\mathbf{R}^d) \times D(\mathbf{R}^m),$$

where x^t, y^t are the paths $x^t(s) = x(s)I_{[0,t]}(s)$, $y^t(s) = y(s)I_{[0,t]}(s)$, $s \in [0, T]$. It is then clear that, for every $t \in [0, T]$, $F(t, \cdot, \cdot)$ is again bounded and continuous on $D(\mathbf{R}^{d+m})$. Furthermore, $F(t, X, Y)$ is $\mathcal{F}_t^{X,Y}$ -measurable.

2) Since $C_b(D(\mathbf{R}^d), \rho)$ is dense in $L^1(D(\mathbf{R}^d), \mathcal{D}(\mathbf{R}^d), \mu)$ (cf. Lemma 4.2), for any $\varepsilon > 0$, we choose an $H^\varepsilon \in C_b(D(\mathbf{R}^d), \rho)$ such that

$$\mathbf{E}_{Q^n} |H(X) - H^\varepsilon(X)| = \mathbf{E}_Q |H(X) - H^\varepsilon(X)| \leq \varepsilon. \quad (4.6)$$

3) Let $t \in [0, T]$ and set $Y_t^k = Y_t((k+1 - |Y_t|)^+ \wedge 1)$, $k \in [0, +\infty)$. We will estimate

$$\mathbf{E}_{Q^n} |Y_t^k - Y_t| = \mathbf{E}_{Q^n} [|Y_t|((|Y_t| - k)^+ \wedge 1)] \leq \mathbf{E}_{Q^n} [|Y_t| I_{\{|Y_t| > k\}}]. \quad (4.7)$$

For this we notice that Y is a (strong) solution of BSDE(H, f_n, X) on $(\Omega, \mathcal{F}, Q^n)$ and, hence

$$\begin{aligned} |Y_t| &\leq \mathbf{E}_{Q^n} \left[|H(X)| + \int_0^T |f_n(s, X, Y)| ds \mid \mathcal{F}_t^{X,Y} \right] \\ &\leq \mathbf{E}_{Q^n} [|H(X)| + KT \mid \mathcal{F}_t^{X,Y}] \leq \mathbf{E}_{Q^n} [h(X) \mid \mathcal{F}_t^{X,Y}], \quad \mathbf{P}\text{-a.s.}, \end{aligned}$$

with $h(x) = |H(x)| + KT$, $x \in D(\mathbf{R}^d)$. Using (4.7), for every $a > 0$, it follows that

$$\begin{aligned} \mathbf{E}_{Q^n} |Y_t - Y_t^k| &\leq \mathbf{E}_{Q^n} [h(X) I_{\{|Y_t| > k\}}] \\ &\leq a Q^n(\{|Y_t| > k\}) + \mathbf{E}_{Q^n} [h(X) I_{\{h(X) > a\}}] \\ &\leq \frac{a}{k} \mathbf{E}_{Q^n} |Y_t| + \mathbf{E}_{Q^n} [h(X) I_{\{h(X) > a\}}] \\ &\leq \frac{a}{k} \mathbf{E}_Q h(X) + \mathbf{E}_Q [h(X) I_{\{h(X) > a\}}]. \end{aligned} \quad (4.8)$$

(Note that the law of X with respect to Q^n and Q is always μ .) Since $h(X)$ is Q -integrable, this allows one to find a constant $k(\varepsilon)$ such that

$$\sup_{n \geq 1, t \in [0, T]} \mathbf{E}_{Q^n} |Y_t^k - Y_t| \leq \varepsilon, \quad k \geq k(\varepsilon). \quad (4.9)$$

4) In this step we shall show that the coordinate process $Y = (Y_t)_{t \in [0, T]}$ is uniformly integrable under Q and that, by enlarging $k(\varepsilon)$ if necessary, we have the following estimate:

$$\sup_{t \in [0, T]} \mathbf{E}_Q |Y_t^k - Y_t| \leq \varepsilon, \quad k \geq k(\varepsilon). \quad (4.10)$$

For this we notice that the functional $\delta^{-1} \int_t^{T \wedge (t+\delta)} |Y_s^\ell - Y_s^k| ds$ on $D(\mathbf{R}^{d+m})$ is bounded and continuous (in the Meyer-Zheng topology).

Taking into account (4.8), we have, for all $t \in [0, T]$, $\ell \geq k \geq 0$, $a > 0$ and $\delta > 0$,

$$\begin{aligned} \mathbf{E}_Q \left[\frac{1}{\delta} \int_t^{t+\delta} |Y_s^\ell - Y_s^k| ds \right] &= \lim_{n \rightarrow +\infty} \mathbf{E}_{Q^n} \left[\frac{1}{\delta} \int_t^{t+\delta} |Y_s^\ell - Y_s^k| ds \right] \\ &\leq \liminf_{n \rightarrow +\infty} \frac{1}{\delta} \int_t^{t+\delta} \mathbf{E}_{Q^n} |Y_s - Y_s^k| ds \\ &\leq \frac{a}{k} \mathbf{E}_Q h(X) + \mathbf{E}_Q [h(X) I_{\{h(X) > a\}}]. \end{aligned}$$

The monotone convergence theorem now allows one to take the limit as $\ell \rightarrow +\infty$, and then, using Fatou's lemma, we get

$$\mathbf{E}_Q |Y_t - Y_t^k| \leq \frac{a}{k} \mathbf{E}_Q h(X) + \mathbf{E}_Q [h(X) I_{\{h(X) > a\}}], \quad t \in [0, T], \quad k \geq 0.$$

Consequently,

$$\begin{aligned} \lim_{k \rightarrow +\infty} \sup_{t \in [0, T]} \mathbf{E}_Q [|Y_t| I_{\{|Y_t| \geq k\}}] &= \lim_{k \rightarrow +\infty} \sup_{t \in [0, T]} \mathbf{E}_Q |Y_t - Y_t^k| \\ &\leq \mathbf{E}_Q [h(X) I_{\{h(X) > a\}}], \quad a > 0. \end{aligned}$$

To get the desired result it now suffices to let a tend to $+\infty$.

5) Let $t \in [0, T]$. Taking into account the right continuity of Y^k at time t , we have $\lim_{\delta \downarrow 0} \delta^P - 1 \int_t^{t+\delta} Y_s^k ds = Y_t^k$, and by Lebesgue's theorem

$$\lim_{\delta \downarrow 0} \mathbf{E}_Q \left| \frac{1}{\delta} \int_t^{t+\delta} Y_s^k ds - Y_t^k \right| = 0.$$

Hence, there is a $\delta(\varepsilon; k) > 0$ (depending on t) such that

$$\mathbf{E}_Q \left| \frac{1}{\delta} \int_t^{t+\delta} Y_s^k ds - Y_t^k \right| \leq \varepsilon, \quad \delta \leq \delta(\varepsilon; k), \quad k \in [0, +\infty). \quad (4.11)$$

6) In this step, we verify a property similar to (4.11), replacing Q by Q^n , uniformly in n . To this end, for $t \leq s \leq T$, we estimate

$$\begin{aligned} |\mathbf{E}_{Q^n} [Y_s - Y_t \mid \mathcal{F}_t^{X,Y}]| &= \left| \mathbf{E}_{Q^n} \left[\int_t^s f_n(u, X, Y) du \mid \mathcal{F}_t^{X,Y} \right] \right| \\ &\leq (s - t) K, \quad Q^n\text{-a.s.} \end{aligned}$$

This yields

$$\begin{aligned} &\left| \mathbf{E}_{Q^n} \left[\frac{1}{\delta} \int_t^{t+\delta} Y_s^k ds - Y_t^k \right] \right| \\ &\leq \frac{1}{\delta} \int_t^{t+\delta} \mathbf{E}_{Q^n} |Y_s^k - Y_s| ds + \mathbf{E}_{Q^n} |Y_t^k - Y_t| + \left| \mathbf{E}_{Q^n} \left[\frac{1}{\delta} \int_t^{t+\delta} Y_s ds - Y_t \right] \right| \\ &\leq 2\varepsilon + \frac{1}{\delta} \int_t^{t+\delta} |\mathbf{E}_{Q^n} [Y_s - Y_t]| ds \\ &\leq 2\varepsilon + \frac{1}{\delta} \int_t^{t+\delta} \mathbf{E}_{Q^n} |\mathbf{E}_{Q^n} [Y_s - Y_t \mid \mathcal{F}_t^{X,Y}]| ds \\ &\leq 2\varepsilon + K \delta, \quad k \geq k(\varepsilon), \quad \delta > 0, \end{aligned}$$

because of (4.9) and the above estimate. Consequently,

$$\sup_{n \geq 1} \left| \mathbf{E}_{Q^n} \left[\frac{1}{\delta} \int_t^{t+\delta} Y_s^k ds - Y_t^k \right] \right| \leq 3\varepsilon, \quad k \geq k(\varepsilon), \quad \delta \leq \frac{\varepsilon}{K}. \quad (4.12)$$

7) Obviously, we have $\int_t^T f_n(s, X, Y) ds = \int_{t+n^{-1}}^T f(s, X, Y) ds$ and, therefore,

$$\left| \int_t^T f(s, X, Y) ds - \int_t^T f_n(s, X, Y) ds \right| \leq \int_t^{t+n^{-1}} |f(s, X, Y)| ds \leq \frac{K}{n}. \quad (4.13)$$

8) In this step, we will show that Y satisfies the BSDE with respect to Q . We denote

$$L = \sup_{(x,y) \in D(\mathbf{R}^{d+m})} |F(x, y)|.$$

Recall that $t \in [0, T]$ is fixed and let $\varepsilon > 0$ be arbitrary but fixed. We choose $k \geq k(\varepsilon)$ and $\delta \leq \delta(\varepsilon, k) \wedge \varepsilon/K$ as in the estimates (4.9)–(4.12). Using that $\delta^{-1} \int_t^{t+\delta} Y_s^k ds$, $H^\varepsilon(X)$, $\int_t^T f(s, X, Y) ds$, $F(t, X, Y)$ are bounded and continuous on $D(\mathbf{R}^{d+m})$, we obtain

$$\begin{aligned} & \left| \mathbf{E}_Q \left[\left(Y_t - H(X) - \int_t^T f(s, X, Y) ds \right) F(t, X, Y) \right] \right| \\ & \leq 2\varepsilon L + \left| \mathbf{E}_Q \left[\left(Y_t^k - H^\varepsilon(X) - \int_t^T f(s, X, Y) ds \right) F(t, X, Y) \right] \right| \\ & \leq 3\varepsilon L + \left| \mathbf{E}_Q \left[\left(\frac{1}{\delta} \int_t^{t+\delta} Y_s^k ds - H^\varepsilon(X) - \int_t^T f(s, X, Y) ds \right) F(t, X, Y) \right] \right| \\ & = 3\varepsilon L + \lim_{n \rightarrow +\infty} \left| \mathbf{E}_{Q^n} \left[\left(\frac{1}{\delta} \int_t^{t+\delta} Y_s^k ds - H^\varepsilon(X) - \int_t^T f(s, X, Y) ds \right) F(t, X, Y) \right] \right| \\ & \leq 6\varepsilon L + \liminf_{n \rightarrow +\infty} \left| \mathbf{E}_{Q^n} \left[\left(Y_t^k - H^\varepsilon(X) - \int_t^T f(s, X, Y) ds \right) F(t, X, Y) \right] \right| \\ & \leq 8\varepsilon L + \liminf_{n \rightarrow +\infty} \left| \mathbf{E}_{Q^n} \left[\left(Y_t - H(X) - \int_t^T f_n(s, X, Y) ds \right) F(t, X, Y) \right] \right| \\ & = 8\varepsilon L + \liminf_{n \rightarrow +\infty} \left| \mathbf{E}_{Q^n} \left[\mathbf{E}_{Q^n} \left[Y_t - H(X) - \int_t^T f_n(s, X, Y) ds \mid \mathcal{F}_t^{X,Y} \right] F(t, X, Y) \right] \right| \\ & = 8\varepsilon L. \end{aligned}$$

This implies

$$\mathbf{E}_Q \left[\left(Y_t - H(X) - \int_t^T f(s, X, Y) ds \right) F(t, X, Y) \right] = 0.$$

By the monotone class theorem, this equality extends to every bounded measurable F on $D(\mathbf{R}^{d+m})$, where again $F(t, \cdot, \cdot)$ is introduced as in Step 1). But

$F(t, X, Y)$ runs over all bounded and $\mathcal{F}_{t-}^0 := \sigma(\{X_s, Y_s: s < t\})$ -measurable functions and we can conclude

$$\mathbf{E}_Q \left[Y_t - H(X) - \int_t^T f(s, X, Y) ds \middle| \mathcal{F}_{t-}^0 \right] = 0, \quad Q\text{-a.s.},$$

or, equivalently,

$$\mathbf{E}_Q \left[Y_t - H(X) - \int_t^T f(s, X, Y) ds \middle| \mathcal{F}_u^0 \right] = 0, \quad Q\text{-a.s.}, \quad 0 \leq u < t \leq T.$$

Then, in particular, we obtain

$$\mathbf{E}_Q \left[Y_{t+2n-1} - H(X) - \int_{t+2n-1}^T f(s, X, Y) ds \middle| \mathcal{F}_{t+n-1}^0 \right] = 0, \quad Q\text{-a.s.}, \quad n \geq 1. \quad (4.14)$$

Since the Q -uniformly integrable process Y is right-continuous and f is bounded, in $L^1(Q)$ we have

$$Y_{t+2n-1} - H(X) - \int_{t+2n-1}^T f(s, X, Y) ds \xrightarrow{n \rightarrow +\infty} Y_t - H(X) - \int_t^T f(s, X, Y) ds.$$

Consequently, in $L^1(Q)$

$$\begin{aligned} & \mathbf{E}_Q \left[Y_{t+2n-1} - H(X) - \int_{t+2n-1}^T f(s, X, Y) ds \middle| \mathcal{F}_{t+n-1}^0 \right] \\ & \xrightarrow{n \rightarrow +\infty} \mathbf{E}_Q \left[Y_t - H(X) - \int_t^T f(s, X, Y) ds \middle| \mathcal{F}_{t+}^0 \right], \end{aligned}$$

and from (4.14) we conclude

$$Y_t = \mathbf{E}_Q \left[H(X) + \int_t^T f(s, X, Y) ds \middle| \mathcal{F}_{t+}^0 \right], \quad Q\text{-a.s.}, \quad t \in [0, T].$$

Since $\mathcal{F}_t^{X,Y}$ is the completion of $\mathcal{F}_{t+}^0$ by all Q -negligible sets from $\mathcal{F}$, it follows that

$$Y_t = \mathbf{E}_Q \left[H(X) + \int_t^T f(s, X, Y) ds \middle| \mathcal{F}_t^{X,Y} \right], \quad Q\text{-a.s.}, \quad t \in [0, T].$$

9) To complete the proof that $(\Omega, \mathcal{F}, \mathbb{F}^{X,Y}, Q, X, Y)$ is a weak solution of equation BSDE(H, f, μ) it remains to show that every $\mathbb{F}^X$ -martingale is also an $\mathbb{F}^{X,Y}$ -martingale on $(\Omega, \mathcal{F}, \mathbf{P})$.

Let N be an arbitrary $\mathbb{F}^X$ -martingale. Then, for every $t \in [0, T]$, there is a $\sigma(X)$ -measurable $\tilde{N}_t$ such that $N_t = \tilde{N}_t$, Q -a.s. Consequently, we can choose a measurable function G_t on $(D(\mathbf{R}^d), \mathcal{D}(\mathbf{R}^d))$ satisfying $\tilde{N}_t = G_t(X)$ and, hence $N_t = G_t(X)$, Q -a.s. Clearly, $G_t \in L^1(D(\mathbf{R}^d), \mathcal{D}(\mathbf{R}^d), \mu)$. Since $C_b(D(\mathbf{R}^d), \rho)$ is dense in this space (cf. Lemma 4.2), for every $\varepsilon > 0$, we can find $G_t^\varepsilon \in C_b(D(\mathbf{R}^d), \rho)$ such that

$$\mathbf{E}_Q |G_t^\varepsilon(X) - G_t(X)| (= \mathbf{E}_{Q^n} |G_t^\varepsilon(X) - G_t(X)|) \leq \varepsilon.$$

Let now F and $F(t, \cdot, \cdot)$ be as in Step 1) and recall

$$L = \sup_{(x,y) \in D(\mathbf{R}^{d+m})} |F(x, y)|.$$

We then obtain, for every $0 \leq t < s \leq T$,

$$\begin{aligned} \left| \mathbf{E}_Q[(N_s - N_t) F(t, X, Y)] \right| &= \left| \mathbf{E}_Q[(G_s(X) - G_t(X)) F(t, X, Y)] \right| \\ &\leq 2\varepsilon L + \left| \mathbf{E}_Q[(G_s^\varepsilon(X) - G_t^\varepsilon(X)) F(t, X, Y)] \right| \\ &\leq 2\varepsilon L + \lim_{n \rightarrow +\infty} \left| \mathbf{E}_{Q^n}[(G_s^\varepsilon(X) - G_t^\varepsilon(X)) F(t, X, Y)] \right| \\ &\leq 4\varepsilon L + \lim_{n \rightarrow +\infty} \left| \mathbf{E}_{Q^n}[(G_s(X) - G_t(X)) F(t, X, Y)] \right|. \end{aligned}$$

But Y is a strong solution of BSDE(H, f_n, X) with respect to Q^n . Hence, there exists a measurable mapping S^n on $D(\mathbf{R}^d)$ with values in $D(\mathbf{R}^m)$ such that $Y = S^n(X)$, Q^n -a.s. Taking into account that the law of X with respect to Q^n coincides with the law of X with respect to Q , that $F(t, X, S^n(X))$ is $\mathcal{F}_t^X$ -measurable and that N is an $\mathbb{F}^X$ -martingale on $(\Omega, \mathcal{F}, Q)$, we can compute

$$\begin{aligned} &\mathbf{E}_{Q^n}[(G_s(X) - G_t(X)) F(t, X, Y)] \\ &= \mathbf{E}_{Q^n}[(G_s(X) - G_t(X)) F(t, X, S^n(X))] \\ &= \mathbf{E}_Q[(G_s(X) - G_t(X)) F(t, X, S^n(X))] = 0. \end{aligned}$$

Letting $\varepsilon \downarrow 0$, this yields

$$\mathbf{E}_Q[(N_s - N_t) F(t, X, Y)] = 0. \quad (4.15)$$

An application of the monotone class theorem implies that (4.15) is true for all bounded and measurable F on $D(\mathbf{R}^{d+m})$ (with $F(t, X, Y) = F(X^t, Y^t)$, cf. Step 1)). Since $F(t, X, Y)$ runs over the set of all bounded $\mathcal{F}_{t-}^0 = \sigma(\{X_u, Y_u: 0 \leq u < t\})$ -measurable functions, this means

$$\mathbf{E}_Q[N_s - N_t \mid \mathcal{F}_{t-}^0] = 0, \quad Q\text{-a.s.}$$

By standard arguments (cf. the end of Step 8)), from this it follows that N is an $\mathbb{F}^{X,Y}$ -martingale on $(\Omega, \mathcal{F}, Q)$. The proof of Theorem 4.1 is completed.

Before closing this section, we state a more general version of Theorem 4.1 in which the driver may also depend on the martingale part of the solution.

Theorem 4.2. *Let (H, f', μ) be a data set, where H, μ are described at the beginning of this section. Suppose that the driver f' satisfies the following conditions:*

- (D.i) f' is a bounded and measurable $\mathbf{R}^m$ -valued functional on the product space $([0, T] \times D(\mathbf{R}^d) \times D(\mathbf{R}^{2m}), \mathcal{B}([0, T]) \otimes \mathcal{D}(\mathbf{R}^d) \otimes \mathcal{D}(\mathbf{R}^{2m}))$;
- (D.ii) for every $(t, x) \in [0, T] \times D(\mathbf{R}^d)$, $f'(t, \cdot)$ is $\mathcal{D}_T^t(\mathbf{R}^{2m})$ -measurable;

(D.iii) the mapping F on $D(\mathbf{R}^{d+2m})$ defined by

$$F(x, y, m) = f'(\cdot, x, y, m), \quad x \in D(\mathbf{R}^d), \quad y, m \in D(\mathbf{R}^m),$$

with values in $L^1([0, T]; \mathbf{R}^m)$ is continuous in the sense of Meyer–Zheng.

Then $\text{BSDE}(H, f', \mu)$ possesses a weak solution.

P r o o f. For $\text{BSDE}(H, f', \mu)$ cf. Definition 1.3. The proof of Theorem 4.2 is similar to that of the preceding theorem. Now one has to work with weak convergence of the triplet (X, Y, M) on the space $D(\mathbf{R}^{d+2m})$. Defining a sequence (Y^n, M^n) of solutions Y^n and martingale parts M^n as in the proof of Theorem 4.1 and Lemma 1.1, one has to ensure, additionally, that the laws of M^n are relatively compact in the Meyer–Zheng topology. For this it is sufficient to compute the conditional variation of M^n , which is just $\mathbf{E}[M_T^n]$, and to verify that it is bounded in n (cf. [20]). The details are left to the reader.

5. Pathwise uniqueness and uniqueness in law of weak solutions. In this section, we introduce the notions of pathwise uniqueness and uniqueness in law of weak solutions of $\text{BSDE}(H, f, \mu)$. In analogy to the celebrated theorem of Yamada–Watanabe [29] we then prove two results: 1) Pathwise uniqueness implies uniqueness in law; 2) if there is a weak solution and if pathwise uniqueness holds then there exists a (pathwise unique) strong solution.

In Section 6, we will give an application, showing existence of a strong solution under quite general assumptions.

We recall $\text{BSDE}(H, f, \mu)$ from Definition 1.2.

Definition 5.1. (i) We say that *pathwise uniqueness* holds for $\text{BSDE}(H, f, \mu)$, if whenever the six-tuples $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y^i)$, $i = 1, 2$, are weak solutions of $\text{BSDE}(H, f, \mu)$ with the same filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$ and the same driving process X then $Y^1 = Y^2$, $\mathbf{P}$ -a.s.

(ii) We say that *joint uniqueness in law* holds for $\text{BSDE}(H, f, \mu)$, if whenever the six-tuples $(\Omega^i, \mathcal{F}^i, \mathbb{F}^i, \mathbf{P}^i, X^i, Y^i)$, $i = 1, 2$, are weak solutions of $\text{BSDE}(H, f, \mu)$, then $P_{(X^1, Y^1)}^1 = P_{(X^2, Y^2)}^2$, i.e., the laws of (X^i, Y^i) with respect to $\mathbf{P}^i$ coincide.

(iii) We say that *uniqueness in law* holds for $\text{BSDE}(H, f, \mu)$, if whenever the six-tuples $(\Omega^i, \mathcal{F}^i, \mathbb{F}^i, \mathbf{P}^i, X^i, Y^i)$, $i = 1, 2$, are weak solutions of $\text{BSDE}(H, f, \mu)$, then $P_{Y^1}^1 = P_{Y^2}^2$, i.e., the laws of Y^i with respect to $\mathbf{P}^i$ coincide.

Obviously, the notion of the joint uniqueness in law is more restrictive than that of uniqueness in law.

Proposition 5.1. *Pathwise uniqueness implies joint uniqueness in law (and, hence, uniqueness in law).*

P r o o f. 1) The proof follows the same scheme as that of the Yamada–Watanabe theorem. Therefore, we shall give only the basic steps, leaving the details to the reader.

We can always assume that H is μ -integrable, since otherwise there is no weak solution of $\text{BSDE}(H, f, \mu)$ and, hence, joint uniqueness in law is fulfilled automatically.

2) Now let $(\Omega^i, \mathcal{F}^i, \mathbb{F}^i, \mathbf{P}^i, X^i, Y^i)$ be weak solutions of $\text{BSDE}(H, f, \mu)$, $i = 1, 2$. By $Q^i = P^i_{(X^i, Y^i)}$ we denote the law of (X^i, Y^i) with respect to $\mathbf{P}^i$ on $D(\mathbf{R}^{d+m}) = D(\mathbf{R}^d) \times D(\mathbf{R}^m)$, $i = 1, 2$. We notice that $D(\mathbf{R}^m)$ endowed with the Skorokhod topology is a Polish space and its Borel σ -algebra coincides with $\mathcal{D}(\mathbf{R}^m)$. It is well known that then there exist conditional laws of Y^i given $X^i = x$, denoted by $q^i(x, B)$, $x \in D(\mathbf{R}^d)$, $B \in \mathcal{D}(\mathbf{R}^m)$, $i = 1, 2$. These are probability kernels from $(D(\mathbf{R}^d), \mathcal{D}(\mathbf{R}^d))$ into $(D(\mathbf{R}^m), \mathcal{D}(\mathbf{R}^m))$ satisfying

$$Q^i(A \times B) = \mathbf{P}^i\{X^i \in A, Y^i \in B\} = \int_A q^i(x, B) \mu(dx)$$

for all $A \in \mathcal{D}(\mathbf{R}^d)$, $B \in \mathcal{D}(\mathbf{R}^m)$. (Note that the law of X^i with respect to $\mathbf{P}^i$ is μ .) We now define

$$\Omega = D(\mathbf{R}^d) \times D(\mathbf{R}^m) \times D(\mathbf{R}^m), \quad \mathcal{F}^0 = \mathcal{D}(\mathbf{R}^d) \otimes \mathcal{D}(\mathbf{R}^m) \otimes \mathcal{D}(\mathbf{R}^m).$$

On $(\Omega, \mathcal{F}^0)$ we introduce the probability measure Q as

$$Q(C) = \int_{\Omega} I_C(x, y_1, y_2) q^1(x, dy^1) q^2(x, dy^2) \mu(dx), \quad C \in \mathcal{F}^0.$$

By $\mathcal{F}$ we denote the completion of $\mathcal{F}^0$ with respect to Q . We consider the projections ζ, η^1, η^2 of Ω onto $D(\mathbf{R}^d)$ and $D(\mathbf{R}^m)$, respectively:

$$\zeta(x, y^1, y^2) = x, \quad \eta^1(x, y^1, y^2) = y^1, \quad \eta^2(x, y^1, y^2) = y^2, \quad (x, y^1, y^2) \in \Omega.$$

As random elements on $(\Omega, \mathcal{F}, Q)$ in $D(\mathbf{R}^d)$ and $D(\mathbf{R}^m)$, respectively, ζ and η^i , $i = 1, 2$, are càdlàg processes, and it can easily be checked that

$$Q_{(\zeta, \eta^i)} = Q^i = P^i_{(X^i, Y^i)}, \quad i = 1, 2. \quad (5.1)$$

Furthermore, the conditional law of η^i given $\zeta = x$ with respect to Q is $q^i(x, \cdot)$, $i = 1, 2$.

Next we introduce the σ -algebras $\mathcal{G}_t^{\zeta, \eta^1, \eta^2} = \sigma(\{\zeta_s, \eta_s^1, \eta_s^2: s \in [0, t]\})$ and the completions $\mathcal{F}_t^{\zeta, \eta^1, \eta^2}$ of $\mathcal{G}_t^{\zeta, \eta^1, \eta^2}$ by all Q -negligible sets from $\mathcal{F}$, and let $\mathbb{F}^{\zeta, \eta^1, \eta^2} = (\mathcal{F}_t^{\zeta, \eta^1, \eta^2})_{t \in [0, T]}$. In the same way, $\mathcal{G}_t^{\zeta}, \mathcal{G}_t^{\eta^1}, \mathcal{G}_t^{\eta^2}$ and $\mathcal{F}_t^{\zeta}, \mathcal{F}_t^{\eta^1}, \mathcal{F}_t^{\eta^2}$ as well as the corresponding filtrations are defined.

3) Using (5.1), it can easily be derived that $(\Omega, \mathcal{F}, \mathbb{F}^{\zeta, \eta^i}, Q, \zeta, \eta^i)$, $i = 1, 2$, are weak solutions of $\text{BSDE}(H, f, \mu)$. For this we notice that it is sufficient to verify equation (1.2) for $\mathcal{G}_t^{\zeta, \eta^i}$ (instead of $\mathcal{F}_t^{\zeta, \eta^i}$). But $(\Omega, \mathcal{F}, \mathbb{F}^{\zeta, \eta^i}, Q, \zeta, \eta^i)$ actually lives on the image space of (X^i, Y^i) and the problem reduces to integral transformation by passing to image measures.

To prove that every $\mathbb{F}^\zeta$ -martingale is again an $\mathbb{F}^{\zeta, \eta^1}$ -martingale, similar arguments apply.

4) In this step, we show the following property of the probability kernels q^i : For every $B \in \mathcal{D}_t^0(\mathbf{R}^m)$, $q^i(\zeta, B)$ is Q -a.s. equal to a $\mathcal{G}_t^\zeta$ -measurable random variable. For the proof it is sufficient to verify

$$\mathbf{E}_Q[q^i(\zeta, B) \mid \mathcal{G}_t^\zeta] = q^i(\zeta, B), \quad Q\text{-a.s.}$$

In view of the definition of q^i this is equivalent to $\int_A \mathbf{E}_Q[q^i(\zeta, B) \mid \mathcal{G}_t^\zeta] dQ = Q(A \cap \{\eta^i \in B\})$, for all $A \in \mathcal{G}_t^\zeta$. But

$$\begin{aligned} \int_A \mathbf{E}_Q[q^i(\zeta, B) \mid \mathcal{G}_t^\zeta] dQ &= \int_\Omega \mathbf{E}_Q[I_A \mid \mathcal{G}_t^\zeta] \mathbf{E}_Q[q^i(\zeta, B) \mid \mathcal{G}_t^\zeta] dQ \\ &= \int_\Omega \mathbf{E}_Q[I_A \mid \mathcal{G}_t^\zeta] q^i(\zeta, B) dQ = \int_\Omega \mathbf{E}_Q[I_A \mid \mathcal{G}_t^\zeta] I_{\{\eta^i \in B\}} dQ. \end{aligned}$$

Now we have $\{\eta^i \in B\} \in \mathcal{G}_t^{\eta^1}$ and, moreover, $\mathbf{E}_Q[I_A \mid \mathcal{G}_t^\zeta] = \mathbf{E}_Q[I_A \mid \mathcal{G}_t^{\zeta, \eta^1}]$, Q -a.s., the last identity following from the property that the $\mathbb{F}^\zeta$ -martingale $\mathbf{E}_Q[I_A \mid \mathcal{G}_t^\zeta]$ is also an $\mathbb{F}^{\zeta, \eta^1}$ -martingale (see Step 3)). Consequently,

$$\begin{aligned} \int_A \mathbf{E}_Q[q^i(\zeta, B) \mid \mathcal{G}_t^\zeta] dQ &= \int_\Omega \mathbf{E}_Q[I_A \mid \mathcal{G}_t^{\zeta, \eta^1}] I_{\{\eta^i \in B\}} dQ \\ &= \int_\Omega \mathbf{E}_Q[I_{A \cap \{\eta^i \in B\}} \mid \mathcal{G}_t^{\zeta, \eta^1}] dQ = Q(A \cap \{\eta^i \in B\}), \end{aligned}$$

which proves the claim.

5) We now show that $(\Omega, \mathcal{F}, \mathbb{F}^{\zeta, \eta^1, \eta^2}, Q, \zeta, \eta^i)$ are weak solutions of $\text{BSDE}(H, f, \mu)$, $i = 1, 2$. Regarding Step 3), it still remains to prove that every $\mathbb{F}^\zeta$ -martingale N over $(\Omega, \mathcal{F}, Q)$ is also an $\mathbb{F}^{\zeta, \eta^1, \eta^2}$ -martingale. We notice that N is Q -a.s. equal to a $\sigma(\zeta)$ -measurable random variable, for all $t \in [0, T]$. Let now $t \in [0, T]$ and let the sets $A \in \mathcal{G}_t^\zeta$, $A^i \in \mathcal{G}_t^{\eta^i}$, $i = 1, 2$, be given. Then there are $B^i \in \mathcal{D}_t^0(\mathbf{R}^m)$ such that $A^i = \{\eta^i \in B^i\}$, $i = 1, 2$. It now follows

$$\begin{aligned} \mathbf{E}_Q[N_T I_{A \cap A^1 \cap A^2}] &= \mathbf{E}_Q[N_T I_A Q(\{\eta^1 \in B^1, \eta^2 \in B^2\} \mid \sigma(\zeta))] \\ &= \mathbf{E}_Q[N_T I_A q^1(\zeta, B^1) q^2(\zeta, B^2)] \\ &= \mathbf{E}_Q[\mathbf{E}_Q[N_T \mid \mathcal{G}_t^\zeta] I_A q^1(\zeta, B^1) q^2(\zeta, B^2)] \\ &= \cdots = \mathbf{E}_Q[\mathbf{E}_Q[N_T \mid \mathcal{G}_t^\zeta] I_{A \cap A^1 \cap A^2}], \end{aligned}$$

since $q^i(\cdot, B^i)$ is $\mathcal{G}_t^{\eta^i}$ -measurable by Step 4) and N is an $\mathbb{F}^\zeta$ -martingale. By the monotone class theorem, we observe that

$$\mathbf{E}_Q[N_T I_C] = \mathbf{E}_Q[\mathbf{E}_Q[N_T \mid \mathcal{G}_t^\zeta] I_C], \quad C \in \mathcal{G}_t^{\zeta, \eta^1, \eta^2}.$$

From this, obviously, it follows that $\mathbf{E}_Q[N_T \mid \mathcal{G}_t^\zeta] = \mathbf{E}_Q[N_T \mid \mathcal{G}_t^{\zeta, \eta^1, \eta^2}]$, Q -a.s. Using this formula for $t + n^{-1}$ (instead of t) and passing to the limit as

$n \rightarrow +\infty$, yields $\mathbf{E}_Q[N_T | \mathcal{G}_{t+}^\zeta] = \mathbf{E}_Q[N_T | \mathcal{G}_{t+}^{\zeta, \eta^1, \eta^2}]$, Q -a.s., proving that N is an $\mathbb{F}^{\zeta, \eta^1, \eta^2}$ -martingale.

3) Finally, from the pathwise uniqueness for BSDE(H, f, μ) we conclude that we have $Q(\{\eta^1 = \eta^2\}) = 1$ and, hence, $Q(\{(\zeta, \eta^1) = (\zeta, \eta^2)\}) = 1$. This, obviously, implies

$$P_{(X^1, Y^1)}^1 = Q^1 = Q_{(\zeta, \eta^1)} = Q_{(\zeta, \eta^2)} = Q^2 = P_{(X^2, Y^2)}^2,$$

and the proof of Proposition 5.1 is complete.

Proposition 5.2. *Existence of a weak solution and pathwise uniqueness for BSDE(H, f, μ) imply existence of a strong solution. In this case, for every weak solution $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y)$ of BSDE(H, f, μ), Y is also a strong solution on $(\Omega, \mathcal{F}, \mathbf{P})$ of BSDE(H, f, X).*

Proof. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y)$ be a weak solution of BSDE(H, f, μ). From the pathwise uniqueness we can conclude that joint uniqueness in law holds. Hence the joint solution measure $P_{(X, Y)}$ on $D(\mathbf{R}^d) \times D(\mathbf{R}^m)$ is uniquely determined. Let $q(x, \cdot)$ be the conditional law of Y given $X = x$ on $D(\mathbf{R}^m)$, $x \in D(\mathbf{R}^d)$. We also notice that $q(x, \cdot)$ is μ -a.s. uniquely determined. Moreover, in view of Step 4) in the proof of Proposition 5.1, $q(\cdot, B)$ is $\mathcal{D}_t^{0, \mu}(\mathbf{R}^d)$ -measurable, for all $B \in \mathcal{D}_t^0(\mathbf{R}^m)$ and $t \in [0, T]$. Here $\mathcal{D}_t^{0, \mu}(\mathbf{R}^d)$ denotes the completion of $\mathcal{D}_t^0(\mathbf{R}^d)$ by all μ -negligible sets from $\mathcal{D}^\mu(\mathbf{R}^d)$.

The proof of Proposition 5.2 essentially depends on the following lemma.

Lemma 5.1. *There exists a functional F on $D(\mathbf{R}^d)$ into $D(\mathbf{R}^m)$ such that, for all $t \in [0, T]$, the mapping*

$$F: (D(\mathbf{R}^d), \mathcal{D}_t^{0, \mu}(\mathbf{R}^d)) \longrightarrow (D(\mathbf{R}^m), \mathcal{D}_t^{0, \mu}(\mathbf{R}^m))$$

is measurable and satisfies the condition $q(X, \{F(X)\}) = 1$, $\mathbf{P}$ -a.s. Moreover,

$$\mathbf{P}\{Y = F(X)\} = 1.$$

From this lemma, it follows at once that $Y = F(X)$ is a strong solution of BSDE(H, f, X) on $(\Omega, \mathcal{F}, \mathbf{P})$, and this obviously proves Proposition 5.2.

It still remains to prove Lemma 5.1.

Proof. The second statement is an immediate consequence of the first one and the definition of $q(x, \cdot)$ as conditional law of Y given $X = x$.

For proving the first statement, we introduce $(\tilde{\Omega}, \tilde{\mathcal{F}}, Q)$ as in the proof of Proposition 5.1 ($\Omega, \mathcal{F}, Q$) was defined, with $q^1(x, B) = q^2(x, B) = q(x, B)$, $x \in D(\mathbf{R}^d)$, $B \in \mathcal{D}(\mathbf{R}^m)$. The processes η^1, η^2 on $(\tilde{\Omega}, \tilde{\mathcal{F}}, Q)$ are defined precisely as there. In the proof of Proposition 5.1 we have seen that $Q(\{\eta^1 = \eta^2\}) = 1$. Denoting the diagonal $\{(y^1, y^2) \in D(\mathbf{R}^m) \times D(\mathbf{R}^m): y^1 = y^2\}$ in the product space $D(\mathbf{R}^m) \times D(\mathbf{R}^m)$ by D , we can easily deduce that $q(x, \cdot) \otimes q(x, \cdot)(D) = 1$ for all $x \in N^c$, where $N \subset D(\mathbf{R}^d)$, $N \in \mathcal{D}(\mathbf{R}^d)$ such that $\mu(N) = 0$. Using the Fubini theorem, we observe $1 = q(x, \cdot) \otimes q(x, \cdot)(D) = \int_{D(\mathbf{R}^m)} q(x, \{y\}) q(x, dy) = \sum_{\substack{y \in D(\mathbf{R}^m) \\ q(x, \{y\}) > 0}} q(x, \{y\})^2$, $x \in N^c$, which implies that

there exists a unique $y = F(x) \in D(\mathbf{R}^m)$ such that $q(x, \{F(x)\}) = 1$, $x \in N^c$. For any $x \in N$, we put $F(x) = 0$. Now, for $x \in N^c$, $F(x)$ can be detected by

$$F(x)_j(t) = \inf\{\alpha \text{ rational: } q(x, \{y: y_j(t) \leq \alpha\}) = 1\},$$

$j = 1, \dots, m$, $t \in [0, T]$. Here we used the notation $F(x) = (F(x)_1, \dots, F(x)_m)$ and $y = (y_1, \dots, y_m) \in D(\mathbf{R}^m)$. From this identity, the measurability property of F stated in Lemma 5.1 follows from that of the conditional law $q(x, \cdot)$ (see beginning of the proof of Proposition 5.2).

6. An application to strong existence. In this section, we apply the results of Sections 4 and 5 in order to give sufficient conditions for the existence of a pathwise unique strong solution of BSDE(H, f, X). An important ingredient is the pathwise uniqueness of weak solutions of BSDE(H, f, μ), which will be proven in a first step, exploiting a certain generalized Lipschitz type condition of the driver f .

In the following, we always assume that μ is a probability measure on $(D(\mathbf{R}^d); \mathcal{D}(\mathbf{R}^d))$, the terminal functional H satisfies condition (H.1) and the driver f satisfies the two conditions (H.2) and (C.ii). Note that $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbf{R}^m$.

Proposition 6.1. *Let $\kappa: (0, +\infty) \rightarrow [0, +\infty)$ be an increasing and concave function such that $\int_0^\varepsilon (\kappa(u))^{-1} du = +\infty$ for all $\varepsilon > 0$. Suppose that the following condition holds:*

$$\langle f(s, x, y) - f(s, x, y'), y(s) - y'(s) \rangle \leq \kappa \left(\sup_{u \in [0, T]} |y(u) - y'(u)|^2 \right), \quad (6.1)$$

for all $s \in [0, T]$, $x \in D(\mathbf{R}^d)$, $y, y' \in D(\mathbf{R}^m)$. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y)$ and $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y')$ be two solutions of BSDE(H, f, μ), defined on the same filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$, with the same driving process X . If $Y, Y' \in \mathcal{S}_{\mathbb{F}, \mathbf{R}^m}^2$, then $Y = Y'$, $\mathbf{P}$ -a.s.

Remark 6.1. (i) If the property stated in Proposition 6.1 is satisfied, we will say that the weak solution $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y)$ is pathwise unique in the class $\mathcal{S}_{\mathbb{F}, \mathbf{R}^m}^2$.

(ii) Condition (6.1) together with condition (C.ii) is equivalent to the following condition:

$$\langle f(s, x, y) - f(s, x, y'), y(s) - y'(s) \rangle \leq \kappa \left(\sup_{u \in [s, T]} |y(u) - y'(u)|^2 \right), \quad (6.2)$$

for all $s \in [0, T]$, $x \in D(\mathbf{R}^d)$, $y, y' \in D(\mathbf{R}^m)$. This is due to the fact that $D(\mathbf{R}^m)$ is a Polish space in the Skorokhod topology, with $\mathcal{D}(\mathbf{R}^m)$ as its Borel σ -algebra, and, therefore, a measurable function on $(D(\mathbf{R}^m), \mathcal{D}(\mathbf{R}^m))$ is $\mathcal{D}_T^s(\mathbf{R}^m)$ -measurable if and only if it is constant on the atoms of $\mathcal{D}_T^s(\mathbf{R}^m)$, as a countably generated σ -algebra.

(iii) The following functions κ are examples for which Theorem 6.1 can be applied:

a) $\kappa(u) = u$, $u \in (0, +\infty)$; b) $\kappa(u) = |\ln u| \cdot u$, $u \in (0, +\infty)$;

c) $\kappa(u) = |\ln(|\ln u|)| \cdot |\ln u| \cdot u$, $u \in (0, +\infty)$, etc.

(iv) We notice that the conditions on κ already imply $\kappa(0+) = \lim_{u \downarrow 0} \kappa(u) = 0$.

(v) Let us consider the condition

$$\begin{aligned} & \langle f(s, x, y) - f(s, x, y'), y(s) - y'(s) \rangle \\ & \leq \kappa_1 \left(\sup_{u \in [s, T]} |y(u) - y'(u)|^2 \right) + C |y(s) - y'(s)|^2, \end{aligned} \quad (6.3)$$

for all $s \in [0, T]$, $x \in D(\mathbf{R}^d)$, $y, y' \in D(\mathbf{R}^m)$, where κ_1 is a function as in Proposition 6.1 and C is some positive constant. We will show that this condition is included in (6.1) as special case. As example, let f have a decomposition $f(s, x, y) = f_1(s, x, y) + f_2(s, x, y(s))$, where f_1 satisfies the conditions of Proposition 6.1 and f_2 is a measurable function on $[0, T] \times D(\mathbf{R}^d) \times \mathbf{R}^m$ such that $\langle f_2(s, x, u) - f_2(s, x, u'), u - u' \rangle \leq C|u - u'|^2$, $s \in [0, T]$, $x \in D(\mathbf{R}^d)$, $u, u' \in \mathbf{R}^m$. We remark that the roles of κ_1 and C can also be changed so that f_2 now satisfies the condition

$$\langle f_2(s, x, u) - f_2(s, x, u'), u - u' \rangle \leq \kappa_1(|u - u'|^2),$$

$s \in [0, T]$, $x \in D(\mathbf{R}^d)$, $u, u' \in \mathbf{R}^m$, and f_1 is Lipschitz:

$$|f_1(s, x, y) - f_1(s, x, y')| \leq C \sup_{u \in [s, T]} |y(u) - y'(u)|,$$

for all $s \in [0, T]$, $x \in D(\mathbf{R}^d)$, $y, y' \in D(\mathbf{R}^m)$.

Hence, we see that Proposition 6.1 includes Lipschitz and monotonicity conditions commonly used, cf., for example, [3].

To prove the claim, let $\kappa(u) = \kappa_1(u) \vee \kappa_2(u)$ with $\kappa_2(u) = Cu$, $u \in (0, +\infty)$. Then κ is increasing and concave, and (6.1) holds for it. Thus it remains to verify that

$$\int_0^\varepsilon (\kappa(u))^{-1} du = +\infty, \quad \text{for all } \varepsilon > 0.$$

As concave function, κ_1 has right derivatives $\kappa_1'(u)$ which are decreasing. Using L'Hospital's rule, we see that there exists the limit

$$a = \lim_{u \downarrow 0} \frac{\kappa_1(u)}{\kappa_2(u)} = \lim_{u \downarrow 0} \frac{\kappa_1'(u)}{\kappa_2'(u)} = \lim_{u \downarrow 0} \frac{\kappa_1'(u)}{C} \in [0, +\infty].$$

In the case that $a \leq 1$, given $\delta > 0$, for $\varepsilon > 0$ sufficiently small, we see that

$$\int_0^\varepsilon (\kappa(u))^{-1} du \geq (1 + \delta)^{-1} \int_0^\varepsilon (\kappa_2(u))^{-1} du = +\infty.$$

The case that $a > 1$ is treated similarly, proving the assertion. Instead of $\kappa_2(u) = Cu$, we can also use a more general κ_2 , however, it is needed that $\lim_{u \downarrow 0} \kappa_2'(u) < +\infty$.

We now give the proof of Proposition 6.1.

P r o o f. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y)$, $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}, X, Y')$ be weak solutions of BSDE(H, f, μ), defined on the same filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$, with the same driving process X and such that $Y, Y' \in \mathcal{S}_{\mathbb{F}, \mathbf{R}^m}^2$. Let M and M' be the martingale parts of Y and Y' , respectively, with $M_0 = M'_0 = 0$ (cf. Lemma 1.1). We notice that M and M' are square integrable. In view of Proposition 1.1, we have

$$\begin{aligned} Y_t &= \int_t^T \{f(\cdot, X, Y)\}_s^{\mathbb{F}} ds - (M_T - M_t), \\ Y'_t &= \int_t^T \{f(\cdot, X, Y')\}_s^{\mathbb{F}} ds - (M'_T - M'_t) \end{aligned}$$

for all $t \in [0, T]$. Setting $\bar{Y} = Y - Y'$, $\bar{M} = M - M'$, this yields

$$\bar{Y}_t = \int_t^T \{f(\cdot, X, Y) - f(\cdot, X, Y')\}_s^{\mathbb{F}} ds - (\bar{M}_T - \bar{M}_t), \quad t \in [0, T].$$

Denoting $\bar{Y} = (\bar{Y}^1, \dots, \bar{Y}^m)$, $\bar{M} = (\bar{M}^1, \dots, \bar{M}^m)$, we consider the square variation processes

$$[\bar{Y}] := \sum_{i=1}^m [\bar{Y}^i], \quad [\bar{M}] := \sum_{i=1}^m [\bar{M}^i]$$

where $[\bar{Y}^i]$, $[\bar{M}^i]$ are the square variations of the components $\bar{Y}^i$ and $\bar{M}^i$, respectively, $i = 1, \dots, m$. Note that $[\bar{Y}] = [\bar{M}]$. An application of Itô's formula now shows $|\bar{Y}_t|^2 = |\bar{Y}_0|^2 + 2 \int_0^t \langle \bar{Y}_{s-}, d\bar{Y}_s \rangle + [\bar{Y}]_t = |\bar{Y}_0|^2 + 2 \int_0^t \langle \bar{Y}_{s-}, d\bar{Y}_s \rangle + [\bar{M}]_t$, $t \in [0, T]$. From this, it can easily be seen that

$$\begin{aligned} |\bar{Y}_t|^2 + [\bar{M}]_T - [\bar{M}]_t &= 2 \int_t^T \langle \bar{Y}_s, \{f(\cdot, X, Y) - f(\cdot, X, Y')\}_s^{\mathbb{F}} \rangle ds \\ &\quad - 2 \int_{t+}^T \langle \bar{Y}_{s-}, d\bar{M}_s \rangle, \quad t \in [0, T]. \end{aligned} \quad (6.4)$$

Taking into account the Burkholder–Davis–Gandhi inequality (cf. [22]), we obtain

$$\begin{aligned} \mathbf{E} \left[\sup_{r \in [t, T]} \left| \int_{t+}^r \langle \bar{Y}_{s-}, d\bar{M}_s \rangle \right| \right] &\leq C \mathbf{E} \left[\left(\sum_{i=1}^m \int_{t+}^T |Y_{s-}^i|^2 d[\bar{M}^i]_s \right)^{1/2} \right] \\ &\leq C \mathbf{E} \left[\left(\sum_{i=1}^m ([\bar{M}^i]_T - [\bar{M}^i]_t) \sup_{s \in (t, T)} |Y_{s-}^i|^2 \right)^{1/2} \right] \\ &\leq C \mathbf{E} \left[\sup_{s \in [t, T]} |Y_s| ([\bar{M}]_T - [\bar{M}]_t)^{1/2} \right] \\ &\leq C \left(\mathbf{E} \left[\sup_{s \in [t, T]} |Y_s|^2 \right] \right)^{1/2} (\mathbf{E} [[\bar{M}]_T - [\bar{M}]_t])^{1/2} < +\infty. \end{aligned} \quad (6.5)$$

For $t = 0$, from (6.5) we see that the local martingale $\int_0^\cdot \langle \bar{Y}_{s-}, d\bar{M}_s \rangle$ is even a martingale and, in particular, $\mathbf{E}[\int_t^T \langle \bar{Y}_{s-}, d\bar{M}_s \rangle] = 0$.

From (6.4) and the assumptions on f , we now derive

$$\begin{aligned} \mathbf{E}[\bar{M}_T - \bar{M}_t] &\leq 2\mathbf{E}\left[\int_t^T \left|\langle \bar{Y}_s, \{f(\cdot, X, Y) - f(\cdot, X, Y')\}_s^{\mathbb{F}}\rangle\right| ds\right] \\ &\leq 2\mathbf{E}\left[\int_t^T \left|\langle \bar{Y}_s, f(s, X, Y) - f(s, X, Y')\rangle\right| ds\right] \\ &\leq 2\mathbf{E}\left[\int_t^T \kappa\left(\sup_{u \in [s, T]} |\bar{Y}_u|^2\right) ds\right], \quad t \in [0, T]. \end{aligned} \quad (6.6)$$

Using (6.4) once again, we obtain $\sup_{u \in [t, T]} |\bar{Y}_u|^2 \leq 2 \int_t^T |\langle \bar{Y}_s, \{f(\cdot, X, Y) - f(\cdot, X, Y')\}_s^{\mathbb{F}}\rangle| ds + 4 \sup_{r \in [t, T]} |\int_{t+}^r \langle \bar{Y}_{s-}, d\bar{M}_s \rangle|$ and, by virtue of estimates (6.5) and (6.6) and the inequality $ab \leq 2^{-1}(a^2 + b^2)$,

$$\begin{aligned} \mathbf{E}\left[\sup_{u \in [t, T]} |\bar{Y}_u|^2\right] &\leq 2\mathbf{E}\left[\int_t^T \kappa\left(\sup_{u \in [s, T]} |\bar{Y}_u|^2\right) ds\right] \\ &\quad + 4C\left(\mathbf{E}\left[\sup_{u \in [t, T]} |Y_u|^2\right]\right)^{1/2} (\mathbf{E}[\bar{M}_T - \bar{M}_t])^{1/2} \\ &\leq 2\mathbf{E}\left[\int_t^T \kappa\left(\sup_{u \in [s, T]} |\bar{Y}_u|^2\right) ds\right] + 2^{-1}\mathbf{E}\left[\sup_{u \in [t, T]} |Y_u|^2\right] \\ &\quad + 8C^2\mathbf{E}[\bar{M}_T - \bar{M}_t] \leq (16C^2 + 2)\mathbf{E}\left[\int_t^T \kappa\left(\sup_{u \in [s, T]} |\bar{Y}_u|^2\right) ds\right] \\ &\quad + 2^{-1}\mathbf{E}\left[\sup_{u \in [t, T]} |Y_u|^2\right]. \end{aligned}$$

This yields

$$\begin{aligned} \mathbf{E}\left[\sup_{u \in [t, T]} |Y_u|^2\right] &\leq C_1\mathbf{E}\left[\int_t^T \kappa\left(\sup_{u \in [s, T]} |\bar{Y}_u|^2\right) ds\right] \\ &\leq C_1\int_t^T \kappa\left(\mathbf{E}\left[\sup_{u \in [s, T]} |\bar{Y}_u|^2\right]\right) ds \end{aligned}$$

(with $C_1 = 32C^2 + 4$), where we used that κ is concave. Setting $\varphi(t) = \mathbf{E}[\sup_{u \in [t, T]} |\bar{Y}_u|^2]$, we arrived at the inequality

$$\varphi(t) \leq C_1 \int_t^T \kappa(\varphi(s)) ds, \quad t \in [0, T].$$

Note that φ is right-continuous, decreasing and such that $\varphi(T) = 0$. Taking into account that $\int_0^\varepsilon (\kappa(u))^{-1} du = +\infty$, $\varepsilon > 0$, it can be verified that the inequality for φ derived above must imply $\varphi(t) = 0$, $t \in [0, T]$. Consequently, we have $Y = Y'$, $\mathbf{P}$ -a.s., completing the proof of the proposition.

Now we are well-prepared to state a result on the existence of a (path-wise unique) strong solution of BSDE(H, f, X). Using the ideas developed

in Section 5, solving the problem of the existence of a strong solution of $\text{BSDE}(H, f, X)$ can be divided into two subtasks: to prove the existence of a weak solution of $\text{BSDE}(H, f, \mu)$ and to prove pathwise uniqueness for $\text{BSDE}(H, f, \mu)$. The subtasks being already solved by Theorem 4.1 and Proposition 6.1, we are able to give the next proposition, as an instructive application. As for forward SDE, this method provides an efficient tool for finding strong solutions of $\text{BSDE}(H, f, X)$.

Proposition 6.2. *Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space and $\mathbb{F}$ a filtration satisfying the usual conditions. Let $X \in D_{\mathbb{F}, \mathbf{R}^d}$ and H be a measurable functional on $(D(\mathbf{R}^d), \mathcal{D}(\mathbf{R}^d))$ such that $\mathbf{E}[|H(X)|^2] < +\infty$. Suppose, moreover, that the functional f on $[0, T] \times D(\mathbf{R}^d) \times D(\mathbf{R}^m)$ satisfies the conditions (C.i)–(C.iii) and (6.1). Then there exists a strong solution Y of $\text{BSDE}(H, f, X)$ on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$ such that $Y \in \mathcal{S}_{\mathbb{F}, \mathbf{R}^m}^2$. The solution is pathwise unique.*

Proof. According to Theorem 4.1, there exists a weak solution $(\Omega, \mathcal{F}, \mathbb{F}^{X,Y}, \mathbf{P}, X, Y)$ of equation $\text{BSDE}(H, f, \mu)$, where μ is the law of X . The assumptions on H and f guarantee that $Y \in \mathcal{S}_{\mathbb{F}, \mathbf{R}^m}^2$ (cf. Lemma 1.1).

Proposition 6.1 now yields that $(\Omega, \mathcal{F}, \mathbb{F}^{X,Y}, \mathbf{P}, X, Y)$ is pathwise unique in the class $\mathcal{S}_{\mathbb{F}, \mathbf{R}^m}^2$ and, because $\tilde{Y} \in \mathcal{S}_{\mathbb{F}, \mathbf{R}^m}^2$ for every weak solution $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}^{X,Y}, \tilde{\mathbf{P}}, \tilde{X}, \tilde{Y})$ of $\text{BSDE}(H, f, \mu)$ with $\tilde{P}_{\tilde{X}} = \mu$, it is also pathwise unique (without restriction to a subclass). By Proposition 5.2, Y is a strong solution of $\text{BSDE}(H, f, X)$ on $(\Omega, \mathcal{F}, \mathbf{P})$. Since pathwise uniqueness holds, the proof of the proposition is finished.

REFERENCES

1. Antonelli F., Ma J. Weak solutions of forward-backward SDE's. — Stochastic Anal. Appl., 2003, v. 21, № 3, p. 493–514.
2. Аркин В. И., Саксонов М. Т. Необходимые условия оптимальности в задачах управления стохастическими дифференциальными уравнениями. — Докл. АН СССР, 1979, т. 244, № 1, с. 11–15.
3. Bahlali K., Essaky E. H., Hassani M., Pardoux E. Existence, uniqueness and stability of backward stochastic differential equations with locally monotone coefficient. — C. R. Acad. Sci. Paris Sér. I, 2002, v. 335, p. 757–762.
4. Bismut J. M. Conjugate convex functions in optimal stochastic control. — J. Math. Anal. Appl., 1973, v. 44, p. 384–404.
5. Bismut J. M. Contrôle des systèmes linéaires quadratiques: applications de l'intégrale stochastique. — Lecture Notes in Math., 1978, v. 649, p. 180–264.
6. Cherny A. S., Engelbert H.-J. Isolated singular points of stochastic differential equations. — Stochastic Processes and Related Topics. Proceedings of the 12th Winter School (Siegmundsborg) 2000. Ed. by R. Buckdahn et al., London: Taylor & Francis, 2002, p. 55–80.
7. Dellacherie C., Meyer P.-A. Probabilités et potentiel. Chapitres I à IV. Paris: Hermann, 1975, 291 p.
8. El Karoui N., Mazliak L. Backward Stochastic Differential Equations. Harlow: Longman, 1997, 221 p.

9. Engelbert H.-J., Schmidt W. On solutions of one-dimensional stochastic differential equations without drift. — *Z. Wahrscheinlichkeitstheor. verw. Geb.*, 1985, v. 68, № 3, p. 287–314.
10. Engelbert H.-J., Schmidt W. On one-dimensional stochastic differential equations with generalized drift. — *Lecture Notes in Control and Inform. Sci.*, 1985, v. 69, p. 143–155.
11. Engelbert H.-J., Schmidt W. Strong Markov continuous local martingales and solutions of one-dimensional stochastic differential equations, I, II, III. — *Math. Nachr.*, 1989, v. 143, p. 167–184; 1989, v. 144, p. 241–281; 1991, v. 151, p. 149–197.
12. Hamadène S., Lepeltier J.-P., Peng S. BSDEs with continuous coefficients and stochastic differential games. — *Backward Stochastic Differential Equations*. Ed by N. El Karoui et al. Harlow: Longman, 1997, p. 115–128.
13. Itô K. Differential equations determining Markov processes. — *Zenkoku Shijō Sūgaku Danwakai*, 1942, v. 244, № 1077, p. 1352–1400 (in Japanese).
14. Itô K. On stochastic differential equations. — *Mem. Amer. Math. Soc.*, 1951, № 4, p. 1–51.
15. Jacod J. Calcul stochastique et problèmes de martingales. — *Lecture Notes in Math.*, 1979, v. 714, p. 1–539.
16. Кабанов Ю. М. О принципе максимума Понтрягина для линейных стохастических дифференциальных уравнений. — *Вероятностные модели и управление экономическими процессами*. М.: ЦЭМИ, 1978.
17. Крылов Н. В. О стохастических интегральных уравнениях Ито. — *Теория вероятн. и ее примен.*, 1969, т. 14, в. 2, с. 340–348; исправл.: 1972, т. 17, в. 2, с. 392–393.
18. Крылов Н. В. Управляемые диффузионные процессы. М.: Наука, 1977, 388 с.
19. Lepeltier J.-P., San Martin J. Backward stochastic differential equations with continuous coefficient. — *Statist. Probab. Lett.*, 1997, v. 32, № 4, p. 425–430.
20. Meyer P. A., Zheng W. A. Tightness criteria for laws of semimartingales. — *Ann. Inst. Henri Poincaré, Probab. Statist.*, 1984, v. 20, p. 353–372.
21. Pardoux E., Peng S. Adapted solution of a backward stochastic differential equation. — *Systems Control Lett.*, 1990, v. 14, № 1, p. 55–61.
22. Protter Ph. *Stochastic Integration and Differential Equations. A New Approach*. Berlin: Springer-Verlag, 1990, 302 p.
23. Revuz D., Yor M. *Continuous Martingales and Brownian Motion*. Berlin: Springer-Verlag, 1991, 533 p.
24. Rozkosz A., Słomiński L. On existence and stability of weak solutions of multidimensional stochastic differential equations with measurable coefficients. — *Stochastic Process. Appl.*, 1991, v. 37, № 2, p. 187–197.
25. Rozkosz A., Słomiński L. On stability and existence of SDEs with reflection at the boundary. — *Stochastic Process. Appl.*, 1997, v. 68, № 2, p. 285–302.
26. Скороход А. В. Исследование по теории случайных процессов. Киев: Изд-во Киевского ун-та, 1961, 216 с.
27. Скороход А. В. Стохастические уравнения для процессов диффузии с границами. I, II. — *Теория вероятн. и ее примен.*, 1961, т. 6, в. 3, с. 287–298; 1962, т. 7, в. 1, с. 5–25.
28. Цирельсон Б. С. Один пример стохастического дифференциального уравнения, не имеющего сильного решения. — *Теория вероятн. и ее примен.*, 1975, т. 20, в. 2, с. 427–430.
29. Watanabe S., Yamada T. On the uniqueness of solutions of stochastic differential equations. — *J. Math. Kyoto Univ.*, 1971, v. 11, p. 155–167.
30. Звонкин А. К. Преобразование фазового пространства диффузионного процесса, «уничтожающее» снос. — *Матем. сб.*, 1974, т. 93, № 1, с. 129–149.

Поступила в редакцию
24.XI.2002