Set-Valued Backward Stochastic Differential Equations

Çağın Ararat[*], Jin Ma[†], and Wenqian Wu[‡]

Abstract

In this paper, we establish an analytic framework for studying *set-valued backward stochastic differential equations (set-valued BSDE)*, motivated largely by the current studies of dynamic set-valued risk measures for multi-asset or network-based financial models. Our framework will make use of the notion of *Hukuhara difference* between sets, in order to compensate the lack of "inverse" operation of the traditional Minkowski addition, whence the vector space structure in set-valued analysis. While proving the well-posedness of a class of set-valued BSDEs, we shall also address some fundamental issues regarding generalized Aumann-Itô integrals, especially when it is connected to the martingale representation theorem. In particular, we propose some necessary extensions of the integral that can be used to represent set-valued martingales with non-singleton initial values. This extension turns out to be essential for the study of set-valued BSDEs.

Keywords. Set-valued stochastic analysis, set-valued stochastic integral, integrably bounded set-valued process, set-valued backward stochastic differential equation, Picard iteration, convex compact set, Hukuhara difference.

2020 AMS Mathematics subject classification: 60H05,10; 60G44; 28B20; 47H04.

[*]Department of Industrial Engineering, Bilkent University, Ankara, 06800, Turkey. E-mail: cararat@bilkent.edu.tr. This author is supported in part by Turkish NSF (TÜBİTAK) 3501-CAREER project #117F438. The author acknowledges the additional support of the University of Southern California during a research visit for this work in January 2019.

[†]Department of Mathematics, University of Southern California, Los Angeles, CA, 90089, USA. Email: jinma@usc.edu. This author is supported in part by US NSF grant #1106853.

[‡]Department of Mathematics, University of Southern California, Los Angeles, CA, 90089, USA. E-mail: wenqian@usc.edu.

1 Introduction

Set-valued analysis, both deterministic and stochastic, has found many applications over the years, most of them are in optimization and optimal control theory. Recently more applications have been studied in economics and finance, the one that particularly motivated this work is the so-called *set-valued dynamic risk measures*, which we now briefly describe.

The risk measure of a financial position ξ at a specific time t, often denoted by $\rho_t(\xi)$, is defined as a convex functional of the (bounded) real-valued random variable ξ satisfying certain axioms such as monotonicity and translativity (cash-additivity) (cf., e.g., [3, 5, 33]). A *dynamic risk measure* is a family of risk measures $\{\rho_t\}_{t\in[0,T]}$ such that, for each financial position ξ, $\{\rho_t(\xi)\}_{t\in[0,T]}$ is an adapted stochastic process satisfying the so-called *time-consistency*, in the sense that the following "tower property" holds (cf. [5, 7, 33]):

$$\rho_s(\xi) = \rho_s(-\rho_t(\xi)), \qquad \xi \in \mathbb{L}^\infty_{\mathcal{F}_T}(\Omega, \mathbb{R}), \ 0 \le s \le t \le T, \tag{1.1}$$

where $\mathbb{L}^\infty_{\mathcal{F}_T}(\Omega, \mathbb{R})$ is the space of $\mathcal{F}_T$-measurable essentially bounded random variables with values in $\mathbb{R}$. A monumental result in the theory of dynamic risk measures is that, any coherent or even convex risk measure satisfying certain "dominating" conditions can be represented as the solution of the following *backward stochastic differential equation* (BSDE):

$$\rho_t(\xi) = -\xi + \int_t^T g(s, \rho_s(\xi), Z_s)ds - \int_t^T Z_s dB_s, \quad t \le T, \tag{1.2}$$

where the driver g is determined completely by the properties of $\{\rho_t\}_{t\in[0,T]}$ (cf. [7, 25, 33]).

There has been a tremendous effort to extend univariate risk measures to the case where the risk appears in the form of a random vector $\xi = (\xi_1, \ldots, \xi_d) \in \mathbb{L}^\infty_{\mathcal{F}_T}(\Omega, \mathbb{R}^d)$ with $d \in \mathbb{N}$, typically known as *systemic risk* in the context of default contagion (see, e.g., [10] for another application in the context of multi-asset markets with transaction costs). For example, one can consider the contagion of (default) risks in a financial market with large number of institutions as a *network*, in which each institution's future asset value can be viewed as a "random shock", to be assessed by its ability to meet its obligations to other members of the network. As a result, it is natural to evaluate these random shocks collectively, which leads to a multivariate setting of a risk measure, often referred to as "systemic risk measures" (cf., e.g., [1, 4, 9]).

One way to characterize a systemic risk measure is to consider it as a multivariate but scalar-valued function. In a static framework, one can define an aggregation function $\Lambda \colon \mathbb{R}^d \to \mathbb{R}$, so as to essentially reduce the problem to a one-dimensional risk measure. For

example, a systemic risk measure can be defined as (cf. [4])

$$\rho^{\text{sys}}(\xi) = \rho(\Lambda(\xi)) = \inf\{k \in \mathbb{R} \colon \Lambda(\xi) + k \in \mathcal{A}\}, \qquad (1.3)$$

where $\xi \in \mathbb{L}^{\infty}_{\mathcal{F}_T}(\Omega, \mathbb{R}^d)$ is the wealth vector of the institutions, $\mathcal{A}$ is a certain *acceptance set*, and ρ is a standard risk measure. Such a definition of a systemic risk measure is convenient but have some fundamental deficiencies, especially when one seeks a dynamic version. For example, it would be almost impossible to define the tower property (1.1), due to the mis-match of the dimensionality. Furthermore, in practice one is often interested in the individual contribution of each institution, and assess the risk for each institution, thus a better way would be to allocate risks individually, so that the value of a systemic risk measure is defined as a set of vectors.

It is worth noting that the set-valued risk measure for a random vector $\xi \in \mathbb{R}^d$ $(d \geq 2)$ can no longer be defined as the "smallest" capital requirement vector, as it may not exist, for instance, with respect to the componentwise ordering of vectors. One remedy is to define it as the set $R_0(\xi)$ (say, at $t = 0$) of all the risk compensating portfolio vectors of ξ so that the risk measure R_0 is a set-valued functional (see, e.g., [8]). Similarly, one can also define a dynamic set-valued risk measure $\{R_t\}_{t \in [0,T]}$. The tower property (1.1) can be defined by

$$R_s(\xi) = \bigcup_{\eta \in R_t(\xi)} R_s(-\eta) =: R_s[-R_t(\xi)], \qquad 0 \leq s \leq t \leq T. \qquad (1.4)$$

However, a BSDE-type mechanism to construct or characterize time-consistent set-valued dynamic risk measures is a widely open problem, and is the main motivation of this paper.

Measurability and integration of set-valued functions can be traced back to 1960s. The commonly used notion of integral is provided by the celebrated work of Aumann [2], where the (Aumann) integral of a set-valued function is defined as the set of all (Lebesgue) integrals of its integrable selections. On the other hand, stochastic integrals of set-valued processes (with respect to Brownian motion or other semimartingales) are relatively new in the literature (see [16]). *Set-valued stochastic differential equations (SDE)*, whose solutions are set-valued stochastic processes, and *stochastic differential inclusions (SDI)*, whose solutions are vector-valued processes, have been studied recently (cf., e.g., [23, 26]). However, while backward SDIs have been around for some time (see, e.g., [17, 18]), to the best of our knowledge, the systematic study of set-valued BSDEs, especially in the general form

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s, \qquad t \in [0, T], \qquad (1.5)$$

is still widely open. (Here Y, Z are set-valued processes that form the solution, ξ is a set-valued terminal condition, f is a set-valued driver function, B is a standard Brownian

motion; the set operations are understood in the Minkowski sense, and the precise definitions of the set-valued integrals will be given in Section 3.3.)

We should point out that the first major difficulty in set-valued analysis, particularly, for studying set-valued BSDEs, is the lack of a vector space structure. More precisely, the (Minkowski) addition for sets does not have an "inverse" (e.g., $A + (-1)A \neq \{0\}(!)$). Consequently, the equivalence between the BSDE (1.5) and its more popular form (cf., e.g., [17, 18] for differential inclusions and [30] for vector-valued BSDEs)

$$Y_t = \mathbb{E}\Big[\xi + \int_t^T f(s, Y_s, Z_s)ds \ \Big| \ \mathcal{F}_t\Big], \qquad t \in [0, T], \tag{1.6}$$

is actually not clear at all. To overcome this difficulty, we shall explore the notion of the so-called *Hukuhara difference* between sets, originated by M. Hukuhara in 1967 [14] and establish some fundamental results on stochastic analysis using Hukuhara difference. It turns out that the seemingly simple additional algebraic structure causes surprisingly subtle technicalities in all aspects of the stochastic analysis, we shall therefore focus on the most basic properties (e.g., Proposition 2.2, Lemmas 2.10, 3.3) and some key estimates (Lemma 4.1, Proposition 4.2) that will be useful for further development.

Another important issue in the study of set-valued BSDEs is related to the choice of Z. For BSDE (1.5), one would naturally search for a solution (Y, Z) that is a set-valued process. However, it is known that (see Remark 3.2(iv)) when Z is a set-valued stochastic process, the corresponding set-valued stochastic integral $\int Z_s dB_s$ *cannot* be *square-integrably bounded* (see §2.3 for the precise definition) unless Z is singleton-valued. This rules out the possibility that both components of (Y, Z) are *set-valued processes*. We shall thus work with *generalized* Aumann-Itô integrals $\int \mathcal{Z} dB$ in which $\mathcal{Z}$ is only a *set of matrix-valued processes*; and consider the simpler version of the BSDE (1.5) and (1.6) in which f is free of Z (or $\mathcal{Z}$).

It turns out that there are some special technical issues in set-valued stochastic analysis involving the generalized Aumann-Itô integral $\int \mathcal{Z} dB$. These issues are subtle, and only occur in the truly set-valued scenarios. For example, in the set-valued framework, a *martingale representation theorem* (with zero initial value) was shown in [21], using the generalized Aumann-Itô integral $\int \mathcal{Z} dB$. However, as was pointed out in the recent work [35], if a set-valued stochastic integral is both a martingale and null at zero, then it must essentially reduce to a singleton. Such an observation nullifies any possible role of the martingale representation theorem in the study of set-valued BSDE, unless some modification on the definition of the stochastic integral is adopted. We shall therefore propose an extension of the generalized Aumann-Itô integral that allows non-singleton initial values, and prove a new martingale representation theorem for a fairly general class of set-valued martingales

4

(Theorem 5.6). With our new notion of set-valued stochastic integrals, we shall also prove the well-posedness of the set-valued BSDE (1.6) (Theorem 6.8) when f is free of Z. By a careful treatment of the existence of Hukuhara differences, we shall gradually work towards obtaining a form of the set-valued BSDE that is similar to (1.5) (Theorems 6.11, 6.14), and discuss the equivalence between the different forms of the set-valued BSDEs (Corollaries 6.12, 6.16). However, some technical issues listed in various remarks will be addressed in our future publications in order not to disturb the main purpose of the paper.

The rest of the paper is organized as follows. In §2, we give the necessary preliminaries on set-valued analysis, introduce the notion of Hukuhara difference and its properties, and extend the existing results (mostly in the book [20]) to those that involve Hukuhara difference. In §3, we revisit set-valued stochastic analysis, again with an eye on these that involve Hukuhara difference. In §4, we establish some key estimates on set-valued conditional expectations and set-valued Lebesgue integrals. In §5, we study set-valued martingales and their representations as generalized stochastic integrals. Finally, in §6, we study the well-posedness of a class of BSDEs of the form (1.5) in the case where f is free of Z and compare it with the BSDE of the form (1.6).

2   Basics of Set-Valued Analysis

In this section, we give a brief introduction to set-valued analysis and all the necessary notations associated to it. The interested reader is referred to the books [20, 22] for many of the definitions but we shall present all the results in a self-contained way.

2.1   Spaces of Sets

In many discussions on set-valued analysis, it is customary to work with closed sets; this covers singletons as a special case, see [28, §1.1.1] for a discussion.

Although most of our discussion applies to more general Hausdorff locally convex topological vector spaces, throughout this paper, we let $\mathbb{X}$ be a separable Banach space with norm $|\cdot|$. We shall denote $\mathscr{P}(\mathbb{X})$ to be the set of all nonempty subsets of $\mathbb{X}$, $\mathscr{C}(\mathbb{X})$ to be the set of all *closed* sets in $\mathscr{P}(\mathbb{X})$, and $\mathscr{K}(\mathbb{X})$ to be the set of all compact convex sets in $\mathscr{P}(\mathbb{X})$, with respect to the norm topology on $\mathbb{X}$. We further denote $\mathscr{K}_w(\mathbb{X})$ to be the set of all weakly compact convex sets in $\mathscr{P}(\mathbb{X})$.

Algebraic Structure on $\mathscr{K}(\mathbb{X})$. Let $A, B \in \mathscr{K}(\mathbb{X})$ and $\alpha \in \mathbb{R}$. We define

$$A + B := \{a + b \colon a \in A,\ b \in B\}, \quad \alpha A := \{\alpha a \colon a \in A\}. \tag{2.1}$$

5

We note that the operations in (2.1) are often referred to as the *Minkowski addition* and *multiplication by scalars.* It can be checked that $\mathscr{K}(\mathbb{X})$ is closed under these operations. It is important to note that the so-called *cancellation law* (cf., e.g., [29, 34]), namely,

$$A + C = B + C \quad \implies \quad A = B. \tag{2.2}$$

holds for $A, B, C \in \mathscr{K}(\mathbb{X})$. Clearly, multiplying A by $\alpha = -1$ gives the "opposite" of A, as $-A := (-1)A$, which leads to the "Minkowski difference"

$$A - B := A + (-1)B = \{a - b \colon a \in A, \ b \in B\}.$$

But in general, $A + (-1)A \neq \{0\}$, that is, the opposite of A is not the "inverse" of A under the Minkowski addition (unless A is a singleton). Consequently, these operations do not establish a vector space structure on $\mathscr{K}(\mathbb{X})$. An early effort to address the inverse operation of Minkowski addition, often still referred to as the *Minkowski difference*, is the so-called "geometric difference" or "inf-residuation" (see [11] and [12]), defined by

$$A \dot{-} B := \{x \in \mathbb{X} \colon x + B \subset A\},$$

with $x + B := \{x\} + B$. Such a difference satisfies $A \dot{-} A = \{0\}$, and can be defined for all $A, B \in \mathscr{K}(\mathbb{X})$. However, one only has $(A \dot{-} B) + B \subset A$; the reverse inclusion usually fails.

In 1967, M. Hukuhara introduced a definition of set difference that has since been referred to as *Hukuhara difference* (cf. [14]) as follows: for $A, B \in \mathscr{K}(\mathbb{X})$,

$$A \ominus B = C \quad \Longleftrightarrow \quad A = B + C. \tag{2.3}$$

The Hukuhara difference has many convenient properties, the only subtlety is that it does not always exist(!). The following result characterizes the existence and gives an explicit expression of $A \ominus B$, which will be used frequently in our future discussions. Recall that, for $A \in \mathscr{K}(\mathbb{X})$ and $a \in A$, a is called an *extreme point* of A if it cannot be written as a strict convex combination of two points in A, that is, for every $x_1, x_2 \in A$ and $\lambda \in (0, 1)$, we have $a \neq \lambda x_1 + (1 - \lambda)x_2$. We denote $\text{ext}(A)$ to be the set of all extreme points of A.

Proposition 2.1. *Let $A, B \in \mathscr{K}(\mathbb{X})$. The Hukuhara difference $A \ominus B$ exists if and only if for every $a \in \text{ext}(A)$, there exists $x \in \mathbb{X}$ such that $a \in x + B \subset A$. In this case, $A \ominus B$ is unique, closed, convex, and we have*

$$A \ominus B = A \dot{-} B = \{x \in \mathbb{X} \mid x + B \subset A\}.$$

Proof: Since this is an infinite-dimensional version of [14, Proposition 4.2] combined with a simple application of the Krein-Milman theorem, we omit the proof. ∎

The Hukuhara difference facilitates set-valued analysis greatly, without the vector space structure on $\mathscr{K}(\mathbb{X})$. We list some properties that will be used often in this paper.

Proposition 2.2. *Let* $A, B, A_1, A_2, B_1, B_2 \in \mathscr{K}(\mathbb{X})$. *Then, the following identities hold:*

(i) $A \ominus A = \{0\}$, $A \ominus \{0\} = A$;

(ii) $(A_1 + B_1) \ominus (A_2 + B_2) = (A_1 \ominus A_2) + (B_1 \ominus B_2)$ *if* $A_1 \ominus A_2$ *and* $B_1 \ominus B_2$ *exist;*

(iii) $(A_1 + B_1) \ominus B_2 = A_1 + (B_1 \ominus B_2)$ *if* $B_1 \ominus B_2$ *exists;*

(iv) $A_1 + (B_1 \ominus B_2) = (A_1 \ominus B_2) + B_1$ *if* $B_1 \ominus B_2$ *and* $A_1 \ominus B_2$ *exist; and*

(v) $A = B + (A \ominus B)$ *if* $A \ominus B$ *exists.*

Proof: (i) Both relations are immediate since $A = A + \{0\} = \{0\} + A$.

(ii) Denote $X := A_1 \ominus A_2$ and $Y := B_1 \ominus B_2$, that is, $A_1 = A_2 + X$ and $B_1 = B_2 + Y$. Adding the identities, we get $A_1 + B_1 = A_2 + X + B_2 + Y = A_2 + B_2 + X + Y$. By (2.3), we see that $(A_1 + B_1) \ominus (A_2 + B_2) = X + Y = (A_1 \ominus A_2) + (B_1 \ominus B_2)$, proving (ii).

(iii) Let $A_2 = \{0\}$ in (ii). By the second equality in (i), we obtain (iii).

(iv) Denote $X := B_1 \ominus B_2$ and $Y := A_1 \ominus B_2$, that is, $B_1 = X + B_2$ and $A_1 = Y + B_2$. Then, $A_1 + X = Y + B_2 + X = Y + B_1$. This is exactly (iv).

(v) This follows immediately by taking $A_1 = A$ and $B_1 = B_2 = B$ in (iv). ∎

Topological Structure on $\mathscr{K}(\mathbb{X})$. We note that both $\mathscr{K}(\mathbb{X})$ and $\mathscr{K}_w(\mathbb{X})$ are closed under the Minkowski addition and multiplication by scalars since $\mathbb{X}$ is a locally convex topological vector space under both the strong and weak topologies. Moreover, the cancellation law (2.2) and Propositions 2.1, 2.2 are valid for both spaces.

For $A, B \in \mathscr{K}(\mathbb{X})$, let us define $\bar{h}(A, B) := \sup_{a \in A} d(a, B)$, where $d(x, B) := \inf_{b \in B} |x - b|$ for $x \in \mathbb{X}$. Then, the *Hausdorff distance* between A and B is given by

$$h(A, B) := \bar{h}(A, B) \vee \bar{h}(B, A) = \inf\{\varepsilon > 0 : A \subset V_\varepsilon(B), \ B \subset V_\varepsilon(A)\},$$

where $V_\varepsilon(C) := \{x \in \mathbb{X} : d(x, C) \leq \varepsilon\}$, $C \in \mathscr{K}(\mathbb{X})$, $\varepsilon > 0$ (cf. [20, Corollary 1.1.3]). Moreover, since $\mathbb{X}$ is a Polish space, so is $(\mathscr{K}(\mathbb{X}), h)$ ([6, Corollary II.9]), and we can define

$$\|A\| := h(A, \{0\}) = \sup\{|a| : a \in A\}, \qquad A \in \mathscr{K}(\mathbb{X}).$$

We have the following easy results.

Proposition 2.3. *(i) The mapping* $\|\cdot\| : \mathscr{K}(\mathbb{X}) \to \mathbb{R}_+$ *satisfies the properties of a norm.*

(ii) If $A, B \in \mathscr{K}(\mathbb{X})$ *and* $A \ominus B$ *exists, then* $h(A, B) = \|A \ominus B\|$.

7

Proof. (i) Clearly, $\|A\| = 0$ implies $A = \{0\}$, and we have $\|\lambda A\| = h(\lambda A, \{0\}) = \sup\{|\lambda y| : y \in A\} = |\lambda| \sup\{|y| : y \in A\} = |\lambda|\|A\|$ for every $\lambda \in \mathbb{R}$. Finally, the "triangle inequality", in the sense that $\|A + B\| \leq \|A\| + \|B\|$, is trivial by definition of $\|\cdot\|$.

(ii) Since $A, B \in \mathscr{K}(\mathbb{X})$, applying the translation invariance property of Hausdorff distance (cf. [15, Proposition 1.3.2]), we see that

$$\|A \ominus B\| = h(A \ominus B, \{0\}) = h((A \ominus B) + B, \{0\} + B) = h(A, B), \tag{2.4}$$

whenever $A \ominus B$ exists. ∎

Remark 2.4. It should be noted that the fact that $\|\cdot\|$ satisfies the properties of a norm *does not* imply that $(\mathscr{K}(\mathbb{X}), \|\cdot\|)$ is a normed space, since $\mathscr{K}(\mathbb{X})$ is not a vector space. It is particularly worth noting that, although the Hausdorff distance is symmetric, the identity (2.4) does not render $(A, B) \mapsto \|A \ominus B\|$ a metric on $\mathscr{K}(\mathbb{R}^d)$ in the usual sense, since the existence of $A \ominus B$ by no means implies that of $B \ominus A$. In fact, it can be checked that both $A \ominus B$ and $B \ominus A$ exist if and only if A is a translation of B, i.e., $A = x + B$ for some $x \in \mathbb{X}$. Nevertheless, the relation in Proposition 2.3(ii) is useful and sufficient for our purposes. ∎

2.2   Set-Valued Measurable Functions and Decomposable Sets

We now consider set-valued functions. Let $(E, \mathcal{E}, \mu)$ be a finite measure space. If E is a topological space, then we take $\mathcal{E} = \mathscr{B}(E)$, the Borel σ-algebra on E. We shall make use of the following definition of set-valued "measurable" function.

Definition 2.5 ([28, Definition 1.3.1])**.** *A set-valued function* $F \colon E \to \mathscr{C}(\mathbb{X})$ *is said to be (strongly) measurable if* $\{e \in E \colon F(e) \cap B \neq \emptyset\} \in \mathcal{E}$ *for every closed set* $B \subset \mathbb{X}$.

The following *selection/representation theorems* for set-valued functions are well-known and will be useful in later sections. We shall denote $\mathrm{cl}(A)$ to be the closure of a set $A \subset \mathbb{X}$.

Theorem 2.6. *Let* $F \colon E \to \mathscr{C}(\mathbb{X})$ *be a set-valued function.*
(i) (Kuratowski and Ryll-Nardzewski, [22, Theorem 2.2.2]) If F is measurable, then there exists an $\mathcal{E}/\mathscr{B}(\mathbb{X})$-measurable function $f \colon E \to \mathbb{X}$, called a measurable selection of F, such that $f(e) \in F(e)$ for each $e \in E$.
(ii) (Castaing, [22, Theorem 2.2.3]) F is measurable if and only if there exists a sequence $\{f_n\}_{n \in \mathbb{N}}$ of measurable selections of F such that $F(e) = \mathrm{cl}\{f_n(e) \colon n \in \mathbb{N}\}$, $e \in E$.

Let us denote $\mathbb{L}^0(E, \mathbb{X}) = \mathbb{L}^0_{\mathcal{E}}(E, \mathbb{X})$ to be the set of all $\mathcal{E}/\mathscr{B}(\mathbb{X})$-measurable functions $f \colon E \to \mathbb{X}$ that are distinguished up to μ-a.e. (almost everywhere) equality. For

$p \in [1, +\infty)$, let $\mathbb{L}^p(E, \mathbb{X}) = \mathbb{L}^p_{\mathcal{E}}(E, \mathbb{X})$ be the set of all $f \in \mathbb{L}^0(E, \mathbb{X})$ such that $\|f\|_p^p := \int_E |f(e)|^p \mu(de) < \infty$. Then, $\mathbb{L}^p(E, \mathbb{X})$ is a Banach space. For $p \in (1, +\infty)$ and $\mathbb{X} = \mathbb{R}^d$, $\mathbb{L}^p(E, \mathbb{X})$ is also reflexive.

Next, let $\mathscr{L}^0(E, \mathscr{C}(\mathbb{X})) = \mathscr{L}^0_{\mathcal{E}}(E, \mathscr{C}(\mathbb{X}))$ be the set of all measurable set-valued mappings $F \colon E \to \mathscr{C}(\mathbb{X})$ distinguished up to μ-a.e. equality. For $F \in \mathscr{L}^0(E, \mathscr{C}(\mathbb{X}))$, we define the set

$$S(F) := S_{\mathcal{E}}(F) := \{f \in \mathbb{L}^0(E, \mathbb{X}) \colon f(e) \in F(e) \ \mu\text{-a.e. } e \in E\}$$

of its measurable selections, which is nonempty by Theorem 2.6(i). Moreover, by Theorem 2.6(ii), two measurable set-valued functions F and G are identical in $\mathscr{L}^0(E, \mathscr{C}(\mathbb{X}))$ if and only if $S(F) = S(G)$. An interesting and crucial question in set-valued analysis is whether a given set of measurable functions in $\mathbb{L}^0(E, \mathbb{R}^d)$ can be seen as the set of measurable selections of a measurable set-valued function. It turns out that this is a highly non-trivial question, for which the following notion is fundamental.

Definition 2.7. *A set $V \subset \mathbb{L}^0(E, \mathbb{X})$ is said to be decomposable with respect to $\mathcal{E}$ if it holds $\mathbf{1}_D f_1 + \mathbf{1}_{D^c} f_2 \in V$ for every $f_1, f_2 \in V$ and $D \in \mathcal{E}$.*

It is easy to check that the intersection of an arbitrary number of decomposable subsets of $\mathbb{L}^0(E, \mathbb{X})$ is again decomposable. With this observation, given a set $V \subset \mathbb{L}^p(E, \mathbb{X})$ with $p \in [1, +\infty)$, we define the *decomposable hull* of V, denoted by $\mathrm{dec}(V) = \mathrm{dec}_{\mathcal{E}}(V)$, to be the smallest decomposable subset of $\mathbb{L}^p(E, \mathbb{X})$ containing V. It can be checked that $\mathrm{dec}(V)$ precisely consists of functions of the form $f = \sum_{i=1}^m \mathbf{1}_{D_i} f_i$, where $\{D_1, \ldots, D_m\}$ is an $\mathcal{E}$-measurable partition of E with $m \in \mathbb{N}$ and $f_1, \ldots, f_m \in V$. We shall often consider $\overline{\mathrm{dec}}(V) = \overline{\mathrm{dec}}_{\mathcal{E}}(V)$, the closure of $\mathrm{dec}(V)$ in $\mathbb{L}^p(E, \mathbb{X})$. It is readily seen that $\overline{\mathrm{dec}}(V)$ is the smallest decomposable and closed subset of $\mathbb{L}^p(E, \mathbb{X})$ containing V.

For $p \in [1, +\infty)$ and $F \in \mathscr{L}^0(E, \mathscr{C}(\mathbb{X}))$, we define $S^p(F) := S^p_{\mathcal{E}}(F) := S(F) \cap \mathbb{L}^p(E, \mathbb{X})$. It is easy to check that $S^p(F)$ is a closed decomposable subset of $\mathbb{L}^p(E, \mathbb{X})$. But it is possible that $S^p(F) = \emptyset$, i.e., F may have no p-integrable selections. We thus consider the set

$$\mathscr{A}^p(E, \mathscr{C}(\mathbb{X})) := \mathscr{A}^p_{\mathcal{E}}(E, \mathscr{C}(\mathbb{X})) := \{F \in \mathscr{L}^0(E, \mathscr{C}(\mathbb{X})) : S^p(F) \neq \emptyset\}, \qquad (2.5)$$

and say that F is p-integrable if $F \in \mathscr{A}^p(E, \mathscr{C}(\mathbb{X}))$. By [20, Corollary 2.3.1], for $F, G \in \mathscr{A}^p(E, \mathscr{C}(\mathbb{X}))$, F and G are identical if and only if $S^p(F) = S^p(G)$. Moreover, we have the following important theorem.

Theorem 2.8 ([20, Theorem 2.3.2]). *Let V be a nonempty closed subset of $\mathbb{L}^p(E, \mathbb{X})$, $p \geq 1$. Then, there exists $F \in \mathscr{A}^p(E, \mathscr{C}(\mathbb{X}))$ such that $V = S^p(F)$ if and only if V is decomposable.*

2.3 Set-Valued Integrals

We shall now assume that $\mathbb{X} = \mathbb{R}^d$, and define the *Aumann integral* of a set-valued function $F \colon E \to \mathscr{C}(\mathbb{R}^d)$ through its measurable selections. To begin with, for $f \in \mathbb{L}^1(E, \mathbb{R}^d)$, we define $I(f) := \int_E f(e)\mu(de)$ and, for a set $M \subset \mathbb{L}^1(E, \mathbb{R}^d)$, we define $I[M] := \{I(f) : f \in M\}$. Then, one can check (see [20, Lemma II.3.9]) that $I[M]$ is a convex subset of $\mathbb{R}^d$ whenever M is decomposable. Now, for a set-valued function $F \in \mathscr{A}^1(E, \mathscr{C}(\mathbb{R}^d))$, we define

$$\int_E F(e)\mu(de) := \mathrm{cl}(I[S^1(F)]). \tag{2.6}$$

Clearly, the "integral" $\int_E F(e)\mu(de)$ is a nonempty closed convex set, and is called the *(closed version of the) Aumann integral* of F.

Let $p \in [1, +\infty)$. We say that $F \in \mathscr{L}^0(E, \mathscr{C}(\mathbb{R}^d))$ is *p-integrably bounded* if there exists $\ell \in \mathbb{L}^p(E, \mathbb{R}_+)$ such that $\|F(e)\| = h(F(e), \{0\}) \leq \ell(e)$ μ-a.e. $e \in E$. Let $\mathscr{L}^p(E, \mathscr{C}(\mathbb{R}^d)) = \mathscr{L}^p_{\mathcal{E}}(E, \mathscr{C}(\mathbb{R}^d))$ be the set of all p-integrably bounded set-valued functions in $\mathscr{L}^0(E, \mathscr{C}(\mathbb{R}^d))$. It is readily seen that $\mathscr{L}^p(E, \mathscr{C}(\mathbb{R}^d)) \subset \mathscr{A}^p(E, \mathscr{C}(\mathbb{R}^d))$. Moreover, by [20, Theorem 2.4.1-(ii)], a set-valued function $F \in \mathscr{A}^p(E, \mathscr{C}(\mathbb{R}^d))$ is p-integrably bounded if and only if $S^p(F)$ is a bounded subset of $\mathbb{L}^p(E, \mathbb{R}^d)$. In this case, it is even true that $S^p(F) = S^{p'}(F) = S(F)$ for every $p' \in [1, p]$ (cf. [28, Proposition 2.1.4]). In what follows, we shall consider mostly the cases $p = 1$ and $p = 2$; and say that F is *integrably bounded* if $F \in \mathscr{L}^1(E, \mathscr{C}(\mathbb{R}^d))$, and *square-integrably bounded* if $F \in \mathscr{L}^2(E, \mathscr{C}(\mathbb{R}^d))$. Clearly, $\mathscr{L}^2(E, \mathscr{C}(\mathbb{R}^d)) \subset \mathscr{L}^1(E, \mathscr{C}(\mathbb{R}^d))$.

We have the following result on integrably bounded set-valued functions. For a subset A of a vector space, $\mathrm{co}(A)$ denotes the convex hull of A.

Theorem 2.9 ([20, Theorem 2.3.4])**.** *Let* $F \in \mathscr{L}^1(E, \mathscr{C}(\mathbb{R}^d))$*. Then,*

$$\int_E F(e)\mu(de) = \int_E \mathrm{co}(F(e))\mu(de).$$

In view of Theorem 2.9, in the integrably bounded case, it is enough to consider the Aumann integrals of convex-valued functions. On the other hand, if $F \in \mathscr{L}^p(E, \mathscr{C}(\mathbb{R}^d))$, then it is immediate that $F(e)$ is a bounded (hence compact) set for μ-a.e. $e \in E$. In what follows, we mostly restrict our attention to the case $F : E \to \mathscr{K}(\mathbb{R}^d)$ and define the spaces $\mathscr{A}^p(E, \mathscr{K}(\mathbb{R}^d))$, $\mathscr{L}^p(E, \mathscr{K}(\mathbb{R}^d))$, and so on in an obvious manner.

Let $F \in \mathscr{L}^p(E, \mathscr{K}(\mathbb{R}^d))$, $p \geq 1$. By [28, Theorem 2.1.18], $S^p(F) = S(F) \in \mathscr{K}_w(\mathbb{L}^p(E, \mathbb{R}^d))$. Moreover, since I is a (weakly) continuous linear mapping on $\mathbb{L}^p(E, \mathbb{R}^d)$, $I[S^p(F)] = I[S(F)]$ is a nonempty compact convex set and one can remove the closure in (2.6), that is,

$$\int_E F(e)\mu(de) = I[S(F)] \in \mathscr{K}(\mathbb{R}^d).$$

The following lemma will be helpful in some later calculations.

Lemma 2.10. *Let $F_1, F_2 \in \mathscr{L}^p(E, \mathscr{K}(\mathbb{R}^d))$, $p \geq 1$. Then, $F_1 + F_2 \in \mathscr{L}^p(E, \mathscr{K}(\mathbb{R}^d))$ and*

$$S(F_1 + F_2) = S(F_1) + S(F_2). \tag{2.7}$$

Furthermore, if $F_1 \ominus F_2$ exists, then $F_1 \ominus F_2 \in \mathscr{L}^p(E, \mathscr{K}(\mathbb{R}^d))$. In this case, we have

$$S(F_1 \ominus F_2) = S(F_1) \ominus S(F_2).$$

Proof. The relation (2.7) is known (see, e.g., [20, Lemma 2.4.1]). In particular, $S^p(F_1 + F_2) \neq \emptyset$ so that $F_1 + F_2 \in \mathscr{A}^p(E, \mathscr{K}(\mathbb{R}^d))$. Moreover, since $S^p(F_1 + F_2)$ is clearly bounded, we have $F_1 + F_2 \in \mathscr{L}^p(E, \mathscr{K}(\mathbb{R}^d))$ and $S^p(F_1 + F_2) = S(F_1 + F_2)$.

Assume now $F_1 \ominus F_2$ exists. We first claim that it is measurable. Indeed, for $e \in E$ and $x \in \mathbb{R}^d$, it is easy to check that (cf. [11, Proposition 4.16]) $x \in F_1(e) \ominus F_2(e)$ holds if and only if there exists a countable dense set $D \subset \mathbb{R}^d$ (independent of the choice of e) such that

$$\langle w, x \rangle \geq \sup_{x_1 \in F_1(e)} \langle w, x_1 \rangle - \sup_{x_2 \in F_2(e)} \langle w, x_2 \rangle, \qquad w \in D.$$

In other words, we can write

$$F_1(e) \ominus F_2(e) = \bigcap_{w \in D} \{ x \in \mathbb{R}^d : \langle w, x \rangle \geq \sup_{x_1 \in F_1(e)} \langle w, x_1 \rangle - \sup_{x_2 \in F_2(e)} \langle w, x_2 \rangle \}. \tag{2.8}$$

For fixed $w \in D$, the mappings $e \mapsto \sup_{x_1 \in F_1(e)} \langle w, x_1 \rangle$, $\sup_{x_2 \in F_2(e)} \langle w, x_2 \rangle$ are measurable ([31, Example 14.51]). Since halfspace-valued mappings inside the intersection in (2.8) are measurable, so is the countable intersection $F_1 \ominus F_2$ ([31, Proposition 14.11-(a)]).

Next, note that $\|F_1(e) \ominus F_2(e)\| \leq \|F_1(e)\| + \|F_2(e)\|$ for every $e \in E$. Since F_1, F_2 are p-integrably bounded, we see that $\|F_1 \ominus F_2\| \in \mathbb{L}^p(E, \mathbb{R})$ and $F_3 := F_1 \ominus F_2$ is p-integrably bounded. Finally, since $F_2, F_3 \in \mathscr{L}^p(E, \mathscr{K}(\mathbb{R}^d))$ and $F_1 = F_2 + F_3$, (2.7) yields $S(F_1) = S(F_2) + S(F_3)$, which then implies that $S(F_1 \ominus F_2) = S(F_3) = S(F_1) \ominus S(F_2)$. $\blacksquare$

To end this section we state a famous result concerning the minimization of integral functionals, which will be particularly useful in Section 4.

Proposition 2.11 ([13, Theorem 2.2]). *Let $F \in \mathscr{A}^p(E, \mathscr{C}(\mathbb{R}^d))$ and $p \geq 1$. Let $\phi \colon E \times \mathbb{R}^d \to \mathbb{R}$ be a jointly measurable function. Suppose that the integral $T_\phi(f) := \int_E \phi(e, f(e))\mu(de)$ is well-defined for each $f \in S^p(F)$, and $T_\phi(f) < +\infty$ for at least one $f \in S^p(F)$. Suppose further that one of the following conditions holds: (i) $x \mapsto \phi(e, x)$ is upper semicontinuous for every $e \in E$. (ii) $(E, \mathscr{E}, \mu)$ is a complete measure space and $x \mapsto \phi(e, x)$ is lower semicontinuous for every $e \in E$. Then,*

$$\inf_{f \in S^p(F)} T_\phi(f) = \int_E \inf_{x \in F(e)} \phi(e, x)\mu(de).$$

3 Set-Valued Stochastic Analysis Revisited

In this section, we review some basics of set-valued stochastic analysis, and establish some fine results that will be useful for our discussion but not covered by the existing literature. Throughout the rest of the paper, we shall consider a given complete, filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F} = \{\mathcal{F}_t\}_{t\in[0,T]})$, on which is defined a standard m-dimensional Brownian motion $B = \{B_t\}_{t\in[0,T]}$, where $T > 0$ is a given time horizon. We shall denote $\mathbb{L}^p_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^d)$ to be the space of all $\mathbb{F}$-progressively measurable d-dimensional processes $\{\phi_t\}_{t\in[0,T]}$ with $\mathbb{E}[\int_0^T |\phi_t|^p dt] < +\infty$. The space $\mathbb{L}^p_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^{d\times m})$ of matrix-valued processes can be defined similarly.

3.1 Set-Valued Conditional Expectations

A set-valued random variable $X : \Omega \to \mathscr{C}(\mathbb{R}^d)$ is an $\mathcal{F}$-measurable set-valued function. If $X \in \mathscr{A}^1(\Omega, \mathscr{C}(\mathbb{R}^d))$, then we define its expectation, denoted by $\mathbb{E}[X]$ as usual, by its Aumann integral $\int_\Omega X(\omega)\mathbb{P}(d\omega)$. Given $p \geq 1$, if $X \in \mathscr{A}^p(\Omega, \mathscr{C}(\mathbb{R}^d))$, then $S^p(X)$ is a closed decomposable subset of $\mathbb{L}^p(\Omega, \mathbb{R}^d)$ and $S^p(\overline{\mathrm{co}}(X)) = \overline{\mathrm{co}}(S^p(X))$ (see [20, Lemma 2.3.3]). Further, X is p-integrably bounded if and only if $S^p(X)$ is a bounded set in $\mathbb{L}^p(\Omega, \mathbb{R}^d)$, that is, $\mathbb{E}[\|X\|^p] = \int_\Omega \sup\{|x|^p : x \in X(\omega)\}\mathbb{P}(d\omega) = \int_\Omega h^p(X(\omega), \{0\})\mathbb{P}(d\omega) < \infty$. In particular, if $X \in \mathscr{L}^p(\Omega, \mathscr{K}(\mathbb{R}^d))$, then $S^p(X) = S(X)$ is a weakly compact convex subset of $\mathbb{L}^p(\Omega, \mathbb{R}^d)$.

Let $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$. We denote $\mathbb{L}^p_{\mathcal{G}}(\Omega, \mathbb{R}^d)$, $\mathscr{A}^p_{\mathcal{G}}(\Omega, \mathscr{C}(\mathbb{R}^d))$, $\mathscr{L}^p_{\mathcal{G}}(\Omega, \mathscr{K}(\mathbb{R}^d))$, $S^p_{\mathcal{G}}(X)$ to be the same as those in §2.2 and §2.3, on the probability space $(\Omega, \mathcal{G}, \mathbb{P})$. Further, for $X \in \mathscr{A}^1_{\mathcal{F}}(\Omega, \mathscr{C}(\mathbb{R}^d))$, the *conditional expectation* of X given $\mathcal{G}$ is defined as the $\mathbb{P}$-a.s. (almost surely) unique set-valued random variable $\mathbb{E}[X|\mathcal{G}] \in \mathscr{A}^1_{\mathcal{G}}(\Omega, \mathscr{C}(\mathbb{R}^d))$ that satisfies

$$S^1_{\mathcal{G}}(\mathbb{E}[X|\mathcal{G}]) = \mathrm{cl}\{\mathbb{E}[f|\mathcal{G}] : f \in S^1(X)\}, \tag{3.1}$$

where the closure is evaluated in $\mathbb{L}^1_{\mathcal{G}}(\Omega, \mathbb{R}^d)$. The existence of $\mathbb{E}[X|\mathcal{G}]$ follows by Theorem 2.8 since the set on the right in (3.1) is decomposable. Moreover, for $p \geq 1$, if $X \in \mathscr{L}^p_{\mathcal{F}}(\Omega, \mathscr{K}(\mathbb{R}^d))$, then it can be shown that the closure in (3.1) is not needed and $\mathbb{E}[X|\mathcal{G}] \in \mathscr{L}^p_{\mathcal{G}}(\Omega, \mathscr{K}(\mathbb{R}^d))$. In this case, $\mathbb{E}[X|\mathcal{G}]$ satisfies the usual identity

$$\int_D \mathbb{E}[X|\mathcal{G}](\omega)\mathbb{P}(d\omega) = \int_D X(\omega)\mathbb{P}(d\omega), \qquad D \in \mathcal{G}.$$

It can be easily checked that $\mathbb{E}[\cdot|\mathcal{G}]$ satisfies all the natural properties of a conditional expectation, except that the "linearity" should be interpreted in terms of the Minkowski addition and multiplication by scalars. Furthermore, we note that the conditional expectation of a set $V \subset \mathbb{L}^1_{\mathcal{F}}(\Omega, \mathbb{R}^d)$ of random variables can also be defined in a generalized sense even if it is not

the set of selections of a set-valued random variable. To be more precise, if $V \subset \mathbb{L}^1_{\mathcal{F}}(\Omega, \mathbb{R}^d)$ is a nonempty closed decomposable set, then there exists a unique $\mathbb{E}[V|\mathcal{G}] \in \mathscr{A}^1_{\mathcal{G}}(\Omega, \mathscr{C}(\mathbb{R}^d))$ (by a slight abuse of notation) such that

$$S^1_{\mathcal{G}}(\mathbb{E}[V|\mathcal{G}]) = \mathrm{cl}\{\mathbb{E}[f|\mathcal{G}] : f \in V\}.$$

The following is a seemingly obvious fact regarding set-valued conditional expectations.

Corollary 3.1. *Let* $X_1, X_2 \in \mathscr{L}^p(\Omega, \mathscr{K}(\mathbb{R}^d))$ *with* $p \in [1, +\infty)$. *Suppose that* $X_1 \ominus X_2$ *exists. Then,* $\mathbb{E}[X_1 \ominus X_2|\mathcal{G}]$ *exists in* $\mathscr{L}^p_{\mathcal{G}}(\Omega, \mathscr{K}(\mathbb{R}^d))$ *and it holds that*

$$\mathbb{E}[X_1 \ominus X_2|\mathcal{G}] = \mathbb{E}[X_1|\mathcal{G}] \ominus \mathbb{E}[X_2|\mathcal{G}]. \tag{3.2}$$

Proof. By Lemma 2.10, $X_1 \ominus X_2 \in \mathscr{L}^p(\Omega, \mathscr{K}(\mathbb{R}^d))$ so that $\mathbb{E}[X_1 \ominus X_2|\mathcal{G}]$ exists in $\mathscr{L}^p_{\mathcal{G}}(\Omega, \mathscr{K}(\mathbb{R}^d))$. By the definition of conditional expectation and repeated applications of Lemma 2.10, we have

$$
\begin{aligned}
S_{\mathcal{G}}(\mathbb{E}[X_1 \ominus X_2|\mathcal{G}] + \mathbb{E}[X_2|\mathcal{G}]) &= S_{\mathcal{G}}(\mathbb{E}[X_1 \ominus X_2|\mathcal{G}]) + S_{\mathcal{G}}(\mathbb{E}[X_2|\mathcal{G}]) \\
&= \{\mathbb{E}[f_1|\mathcal{G}] : f_1 \in S(X_1 \ominus X_2)\} + \{\mathbb{E}[f_2|\mathcal{G}] : f_2 \in S(X_2)\} \\
&= \{\mathbb{E}[f|\mathcal{G}] : f \in S(X_1 \ominus X_2) + S(X_2)\} \\
&= \{\mathbb{E}[f|\mathcal{G}] : f \in S((X_1 \ominus X_2) + X_2)\} = S_{\mathcal{G}}(\mathbb{E}[X_1|\mathcal{G}]).
\end{aligned}
$$

This is equivalent to having $\mathbb{E}[X_1|\mathcal{G}] = \mathbb{E}[X_1 \ominus X_2|\mathcal{G}] + \mathbb{E}[X_2|\mathcal{G}]$, whence (3.2). ∎

3.2 Set-Valued Stochastic Processes

A set-valued stochastic process $\Phi = \{\Phi_t\}_{t \in [0,T]}$ is a family of set-valued random variables taking values in $\mathscr{C}(\mathbb{R}^d)$. We call Φ measurable if it is $\mathscr{B}([0,T]) \otimes \mathcal{F}$-measurable as a single set-valued function on $[0,T] \times \Omega$. The notions such as "adaptedness" or "progressive measurability" can be defined accordingly in the obvious ways. We denote $\mathscr{L}^0_{\mathbb{F}}([0,T] \times \Omega, \mathscr{C}(\mathbb{R}^d))$ to be the space of all set-valued, $\mathbb{F}$-progressively measurable processes taking values in $\mathscr{C}(\mathbb{R}^d)$. For $\Phi \in \mathscr{L}^0_{\mathbb{F}}([0,T] \times \Omega, \mathscr{C}(\mathbb{R}^d))$, we denote $S_{\mathbb{F}}(\Phi)$ to be the set of all $\mathbb{F}$-progressively measurable selections of Φ, which is nonempty by Theorem 2.6. For $p \in [1, +\infty)$, we define $S^p_{\mathbb{F}}(\Phi) := S_{\mathbb{F}}(\Phi) \cap \mathbb{L}^p_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^d)$ and denote $\mathscr{L}^p_{\mathbb{F}}([0,T] \times \Omega, \mathscr{C}(\mathbb{R}^d))$ to be the set of all $\mathbb{F}$-progressively measurable, $\mathscr{C}(\mathbb{R}^d)$-valued processes Φ with $\mathbb{E}[\int_0^T \|\Phi_t\|^p\, dt] < +\infty$ (i.e., p-integrably bounded). The notations $\mathscr{L}^p_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d)), \mathscr{L}^p_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^{d \times m}))$ for set-valued processes with compact convex values are defined similarly for $p = 0$ and $p \geq 1$. It is worth pointing out that the space $\mathscr{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$ is not a Hilbert space, but only a complete metric space, with the metric $d_H(\Phi, \Psi) := (\mathbb{E}[\int_0^T h^2(\Phi_t, \Psi_t) dt])^{1/2}$.

3.3 Set-Valued Stochastic Integrals

In this section, we assume that $\mathbb{F} = \mathbb{F}^B$, the natural filtration generated by B, augmented by all the $\mathbb{P}$-null sets of $\mathcal{F}$ so that it satisfies the *usual hypotheses.*

Let us consider the two linear mappings $J : \mathbb{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^d) \to \mathbb{L}^2_{\mathcal{F}_T}(\Omega, \mathbb{R}^d)$, and $\mathcal{J} : \mathbb{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^{d \times m}) \to \mathbb{L}^2_{\mathcal{F}_T}(\Omega, \mathbb{R}^d)$ defined by

$$J(\phi) := \int_0^T \phi_t dt, \qquad \mathcal{J}(\psi) := \int_0^T \psi_t dB_t, \tag{3.3}$$

for $\phi \in \mathbb{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^d)$, $\psi \in \mathbb{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^{d \times m})$, respectively. For $K \subset \mathbb{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^d)$ (resp. $K' \subset \mathbb{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^{d \times m})$), the set $J[K]$ (resp. $\mathcal{J}[K']$) is defined in an obvious way.

Let $\Phi \in \mathscr{L}^0_{\mathbb{F}}([0,T] \times \Omega, \mathscr{C}(\mathbb{R}^d))$ and $\Psi \in \mathscr{L}^0_{\mathbb{F}}([0,T] \times \Omega, \mathscr{C}(\mathbb{R}^{d \times m}))$ such that $S^2_{\mathbb{F}}(\Phi) \neq \emptyset$, $S^2_{\mathbb{F}}(\Psi) \neq \emptyset$. Then, one can show that there exist unique set-valued random variables $\int_0^T \Phi_t dt \in \mathscr{A}^2_{\mathcal{F}_T}(\Omega, \mathscr{C}(\mathbb{R}^d))$ and $\int_0^T \Psi_t dB_t \in \mathscr{A}^2_{\mathcal{F}_T}(\Omega, \mathscr{C}(\mathbb{R}^d))$ such that

$$S^2_{\mathcal{F}_T}\left(\int_0^T \Phi_t dt\right) = \overline{\mathrm{dec}}_{\mathcal{F}_T}(J[S^2_{\mathbb{F}}(\Phi)]), \qquad S^2_{\mathcal{F}_T}\left(\int_0^T \Psi_t dB_t\right) = \overline{\mathrm{dec}}_{\mathcal{F}_T}(\mathcal{J}[S^2_{\mathbb{F}}(\Psi)]). \tag{3.4}$$

We call $\int_0^T \Phi_t dt$ and $\int_0^T \Psi_t dB_t$ *set-valued stochastic integrals.* As usual, for $t \in [0,T]$, we define the *indefinite* stochastic integrals as $\int_0^t \Phi_s ds := \int_0^T \mathbf{1}_{(0,t]}(s) \Phi_s ds$ and $\int_0^t \Psi_s dB_s := \int_0^T \mathbf{1}_{(0,t]}(s) \Psi_s dB_s$. Equivalently, one can define them via the relations $S^2_{\mathcal{F}_t}(\int_0^t \Phi_s ds) = \overline{\mathrm{dec}}_{\mathcal{F}_t}(J_{0,t}[S^2_{\mathbb{F}}(\Phi)])$, $S^2_{\mathcal{F}_t}(\int_0^t \Psi_s dB_s) = \overline{\mathrm{dec}}_{\mathcal{F}_t}(\mathcal{J}_{0,t}[S^2_{\mathbb{F}}(\Psi)])$, where $J_{0,t}(\phi) := \int_0^t \phi_s ds$, $\mathcal{J}_{0,t}(\psi) := \int_0^t \psi_s dB_s$. The integrals $\int_t^T \Phi_s ds$ and $\int_t^T \Psi_s dB_s$, and the mappings $J_{t,T}$, $\mathcal{J}_{t,T}$ can be defined similarly for $t \in [0,T]$.

Remark 3.2. The set-valued Itô stochastic integrals have many interesting properties, we refer the interested reader to the books [20, 22] for the exhaustive explorations. Here, we mention a few that will be useful for our discussion.

(i) The definition (3.4) implies that both $\int_0^T \Phi_t dt$ and $\int_0^T \Psi_t dB_t$ are $\mathcal{F}_T$-measurable set-valued random variables. However, neither of the sets $J[S^2_{\mathbb{F}}(\Phi)], \mathcal{J}[S^2_{\mathbb{F}}(\Psi)] \subset \mathbb{L}^2_{\mathcal{F}_T}(\Omega, \mathbb{R}^d)$ is necessarily decomposable (see [20, p.105] for counterexamples). Thus, by virtue of Theorem 2.8, they cannot be seen as the selections of any $\mathcal{F}_T$-measurable set-valued random variables.

(ii) One can actually show that $\{\mathbb{E}[x] : x \in \mathcal{J}[S^2_{\mathbb{F}}(\Psi)]\} = \{0\}$, and $\mathcal{J}[S^2_{\mathbb{F}}(\Psi)]$ is decomposable if and only if it is a singleton(!).

(iii) By [20, Theorem 3.1.1], we have $\overline{\mathrm{dec}}_{\mathcal{F}_T}(\mathcal{J}[S^2_{\mathbb{F}}(\Psi)]) = \mathbb{L}^2_{\mathcal{F}_T}(\Omega, \mathbb{R}^d)$ if and only if $\overline{\mathrm{dec}}(\mathcal{J}[S^2_{\mathbb{F}}(\Psi)]) \neq \emptyset$.

(iv) If Φ and Ψ are convex-valued, then so are $\int_0^T \Phi_t dt$ and $\int_0^T \Psi_t dB_t$. If $\Phi \in \mathscr{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$, then it is known that $\int_0^T \Phi_t dt \in \mathscr{L}^2_{\mathcal{F}_T}(\Omega, \mathscr{K}(\mathbb{R}^d))$, that is, the stochastic time

14

integral of a square-integrably bounded process is a square-integrably bounded set-valued random variable (see [19, Theorem 3.2]). However, the Itô integral $\int_0^T \Psi_t dB_t$ fails to be square-integrably bounded in general even if $\Psi \in \mathscr{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^{d \times m}))$ (see [27]).

(v) The set-valued stochastic integrals $\int_0^t \Phi_s ds$, $\int_0^t \Psi_s dB_s$ are defined, almost surely, for each $t \in [0,T]$, and they are $\mathbb{F}$-*adapted*, in the usual sense. Furthermore, when $\Phi \in \mathscr{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$, by [24, Theorem 2.4], the process $\{\int_0^t \Phi_s ds\}_{t \in [0,T]}$ has a continuous (with respect to h), whence progressively measurable, modification. We can define the indefinite integral $\int_0^{\cdot} \Phi_s ds$ by this progressively measurable set-valued process. However, the continuity of the Itô integral $\{\int_0^t \Psi_s dB_s\}_{t \in [0,T]}$ is much more involved, and so is the progressive measurability issue (see [22, §5.5] for a special case). ∎

The following lemma shows that the additivity holds for both integrals, which also allows to calculate the integrals of the Hukuhara difference of two processes.

Lemma 3.3. *Suppose that $\mathbb{P}$ is a nonatomic probability measure. Let $\Phi^1, \Phi^2 \in \mathscr{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$ and $\Psi^1, \Psi^2 \in \mathscr{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^{d \times m}))$. Then, for every $t \in [0,T]$,*

$$\int_0^t (\Phi_s^1 + \Phi_s^2)ds = \int_0^t \Phi_s^1 ds + \int_0^t \Phi_s^2 ds, \quad \int_0^t (\Psi_s^1 + \Psi_s^2)dB_s = \int_0^t \Psi_s^1 dB_s + \int_0^t \Psi_s^2 dB_s \ (3.5)$$

hold almost surely. If $\Phi^1 \ominus \Phi^2$ and $\Psi^1 \ominus \Psi^2$ exist ($dt \times d\mathbb{P}$-a.e.), then we have $\Phi^1 \ominus \Phi^2 \in \mathscr{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$, $\Phi^1 \ominus \Phi^2 \in \mathscr{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^{d \times m}))$ and, for every $t \in [0,T]$,

$$\int_0^t (\Phi_s^1 \ominus \Phi_s^2)ds = \int_0^t \Phi_s^1 ds \ominus \int_0^t \Phi_s^2 ds, \quad \int_0^t (\Psi_s^1 \ominus \Psi_s^2)dB_s = \int_0^t \Psi_s^1 dB_s \ominus \int_0^t \Psi_s^2 dB_s \ (3.6)$$

hold almost surely.

Proof: The relations in (3.5) are by [19, Theorem 3.1-3.2]. Suppose that $\Phi^1 \ominus \Phi^2$ exists. It is clear that $\Phi^1 \ominus \Phi^2$ takes values in $\mathscr{K}(\mathbb{R}^d)$. Since $\|\Phi_t^1 \ominus \Phi_t^2\| \leq \|\Phi_t^1\| + \|\Phi_t^2\|$,

$$\mathbb{E}\left[\int_0^T \|\Phi_t^1 \ominus \Phi_t^2\|^2 dt\right] \leq 2\mathbb{E}\left[\int_0^T \|\Phi_t^1\|^2 dt\right] + 2\mathbb{E}\left[\int_0^T \|\Phi_t^2\|^2 dt\right] < +\infty.$$

This and Lemma 2.10 imply that $\Phi^1 \ominus \Phi^2 \in \mathscr{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$. We have $\Phi^1 = \Phi^2 + (\Phi^1 \ominus \Phi^2)$. Let $t \in [0,T]$. By the first relation in (3.5), we obtain $\int_0^t \Phi_s^1 ds = \int_0^t \Phi_s^2 ds + \int_0^t (\Phi_s^1 \ominus \Phi_s^2)ds$. By the definition of Hukuhara difference, the first relation in (3.6) follows. The proofs of the claims related to $\Psi^1 \ominus \Psi^2$ are similar, hence omitted. ∎

Remark 3.4. The nonatomic assumption on probability $\mathbb{P}$ in Lemma 3.3 is inherited from [19, Theorem 3.1-3.2], where the argument relies on the fact that a bounded weakly closed decomposable subset of $\mathbb{L}^2(\Omega, \mathbb{R}^d)$ is convex under such assumption, thanks to Lyapunov's theorem (cf. [19, Theorem 2.1(viii)] and [22, Theorem 3.2.1, Lemma 3.2.5]). ∎

15

Corollary 3.5. *Suppose that $\mathbb{P}$ is a nonatomic probability measure. Let $\Phi \in \mathscr{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$, $\Psi \in \mathscr{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^{d \times m}))$. For each $t \in [0,T]$,*

$$\int_0^T \Phi_s ds = \int_0^t \Phi_s ds + \int_t^T \Phi_s ds, \quad \int_0^T \Psi_s dB_s = \int_0^t \Psi_s dB_s + \int_t^T \Psi_s dB_s$$

and

$$\int_t^T \Phi_s ds = \int_0^T \Phi_s ds \ominus \int_0^t \Phi_s ds, \quad \int_t^T \Psi_s dB_s = \int_0^T \Psi_s dB_s \ominus \int_0^t \Psi_s dB_s.$$

hold almost surely.

Proof. This is immediate from Lemma 3.3 and the definitions of the integrals since $\mathbf{1}_{(0,T]}(s)\xi_s = \mathbf{1}_{(0,t]}(s)\xi_s + \mathbf{1}_{(t,T]}(s)\xi_s$, $\xi \in \{\Phi, \Psi\}$, for all $s \in [0,T]$. ∎

The notion of stochastic integral can be extended to the case where the integrand is only a set of processes, instead of a set-valued process. We briefly describe the idea (cf. [21]). Let $\mathcal{Z} \in \mathscr{P}(\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m}))$ be a nonempty set and consider the sets $\mathcal{J}_t[\mathcal{Z}] = \{\int_0^t z_s dB_s : z \in \mathcal{Z}\}$, $t \in [0,T]$. Due to lack of decomposability, $\mathcal{J}_t[\mathcal{Z}]$ is not equal to the set of square-integrable selections of a set-valued random variable, in general. But similar to the stochastic integral discussed above, one can show that, for each $t \in [0,T]$, there exists a unique $\int_0^t \mathcal{Z}dB \in \mathscr{A}_{\mathcal{F}_t}^2(\Omega, \mathscr{C}(\mathbb{R}^d))$ such that

$$S_{\mathcal{F}_t}^2 \left(\int_0^t \mathcal{Z}dB \right) = \overline{\mathrm{dec}}_{\mathcal{F}_t}(\mathcal{J}_t[\mathcal{Z}]). \tag{3.7}$$

We call $\int_0^t \mathcal{Z}dB$ the *generalized (indefinite) Aumann-Itô stochastic integral* (cf. [21]). If $\mathcal{Z}$ is convex, then $\int_0^t \mathcal{Z}dB$ is convex-valued (see [21, Theorem 2.2]).

We have the following analogue of Lemma 3.3.

Lemma 3.6. *Assume that $\mathbb{P}$ is nonatomic, and let $\mathcal{Z}^1, \mathcal{Z}^2 \in \mathscr{K}_w(\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m}))$. Then, the following statements are true:*

(i) $\mathcal{Z}^1 + \mathcal{Z}^2 \in \mathscr{K}_w(\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m}))$ and for every $t \in [0,T]$, it holds that

$$\int_0^t (\mathcal{Z}^1 + \mathcal{Z}^2)dB = \int_0^t \mathcal{Z}^1 dB + \int_0^t \mathcal{Z}^2 dB \qquad \mathbb{P}\text{-a.s.} \tag{3.8}$$

(ii) If $\mathcal{Z}^1 \ominus \mathcal{Z}^2$ exists, then $\mathcal{Z}^1 \ominus \mathcal{Z}^2 \in \mathscr{K}_w(\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m}))$ and for every $t \in [0,T]$,

$$\int_0^t (\mathcal{Z}^1 \ominus \mathcal{Z}^2)dB = \int_0^t \mathcal{Z}^1 dB \ominus \int_0^t \mathcal{Z}^2 dB \qquad \mathbb{P}\text{-a.s.} \tag{3.9}$$

(iii) If $\mathcal{Z}^1 \ominus \mathcal{Z}^2$ exists and $\int_0^t \mathcal{Z}^1 dB = \int_0^t \mathcal{Z}^2 dB$, $\mathbb{P}$-a.s., for all $t \in [0,T]$, then $\mathcal{Z}^1 = \mathcal{Z}^2$ as subsets of $\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m})$.

Proof: (i) The additivity result (3.8) is given in [21, Theorem 2.2]. (ii) Since $\mathcal{Z}^1, \mathcal{Z}^2$ are bounded subsets of $\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m})$, it can be checked that $\mathcal{Z}^1 \ominus \mathcal{Z}^2$ is also bounded. Moreover, $\mathcal{Z}^1 \ominus \mathcal{Z}^2$ is convex and closed as a Hukuhara difference. Since $\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m})$ is reflexive, we may conclude that $\mathcal{Z}^1 \ominus \mathcal{Z}^2$ is weakly compact. Hence, $\mathcal{Z}^1 \ominus \mathcal{Z}^2 \in \mathscr{K}_w(\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m}))$. The proof of the identity (3.9) follows from the additivity of integral as in the proof of Lemma 3.3.

It remains to prove (iii). First, by the property of the Hukuhara difference and the assertion (ii), it suffices to show that having $\int_0^t \mathcal{Z} dB = \{0\}$ $\mathbb{P}$-a.s. for every $t \in [0,T]$ implies $\mathcal{Z} = \{0\}$. To see this, we observe that, for a fixed $t \in [0,T]$, having $\int_0^t \mathcal{Z} dB = \{0\}$ $\mathbb{P}$-a.s. amounts to saying, by definition, that $S_{\mathcal{F}_t}^2(\int_0^t \mathcal{Z} dB) = \overline{\mathrm{dec}}_{\mathcal{F}_t}(\mathcal{J}_t[\mathcal{Z}]) = \{0\}$, which is obviously equivalent to $\mathcal{J}_t[\mathcal{Z}] = \{0\}$. In other words, we have $\int_0^t z_s dB_s = \{0\}$ $\mathbb{P}$-a.s. for all $z \in \mathcal{Z}$. But since the integral $M_t^z := \int_0^t z_s dB_s$, $t \in [0,T]$, is a continuous martingale, we can conclude that $\mathbb{P}\{M_t^z = 0 \text{ for all } t \in [0,T]\} = 1$ for each $z \in \mathcal{Z}$. This leads to that $z \equiv 0$, $\mathbb{P}$-a.s. for all $z \in \mathcal{Z}$, that is, $\mathcal{Z} = \{0\}$. ∎

In the proof of Lemma 3.6(iii), the existence of the Hukuhara difference $\mathcal{Z}^1 \ominus \mathcal{Z}^2$ is needed in order to obtain the conclusion $\mathcal{Z}^1 = \mathcal{Z}^2$, and hence $\mathcal{Z}^1 \ominus \mathcal{Z}^2 = \{0\}$, by using Lemma 3.6(ii). To remove this assumption, we will pass to a quotient space of $\mathscr{K}_w(\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m}))$ in which two sets of processes are considered identical if they yield the same Itô integral. To make this idea precise, let us define a relation $\cong$ on $\mathscr{K}_w(\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m}))$ by

$$\mathcal{Z}^1 \cong \mathcal{Z}^2 \quad \Leftrightarrow \quad \int_0^t \mathcal{Z}^1 dB = \int_0^t \mathcal{Z}^2 dB \quad \mathbb{P}\text{-a.s. for all } t \in [0,T]. \tag{3.10}$$

It is easy to see that $\cong$ is an equivalence relation on $\mathscr{K}_w(\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m}))$; let us denote $\mathbb{K}_w(\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m}))$ to be the set of all equivalence classes of $\cong$. For a class $\mathcal{Z} \in \mathbb{K}_w(\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m}))$, we define its stochastic integral $\{\int_0^t \mathcal{Z} dB\}_{t \in [0,T]}$, as the stochastic integral of any member of $\mathcal{Z}$, which is uniquely defined up to modifications. Hence, for $\mathcal{Z}^1, \mathcal{Z}^2 \in \mathbb{K}_w(\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m}))$, if $\int_0^t \mathcal{Z}^1 dB = \int_0^t \mathcal{Z}^2 dB$ $\mathbb{P}$-a.s. for all $t \in [0,T]$, then $\mathcal{Z}^1 = \mathcal{Z}^2$ in $\mathbb{K}_w(\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m}))$.

For future use, let us extend the definition of Minkowski addition for the new space. For $\mathcal{Z}, \hat{\mathcal{Z}} \in \mathbb{K}_w(\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m}))$, we define

$$\mathcal{Z} + \hat{\mathcal{Z}} := \{\mathcal{Z}^1 + \hat{\mathcal{Z}}^1 : \mathcal{Z}^1 \in \mathcal{Z}, \hat{\mathcal{Z}}^1 \in \hat{\mathcal{Z}}\}, \tag{3.11}$$

which is well-defined since $\mathcal{Z}^1 + \hat{\mathcal{Z}}^1 \cong \mathcal{Z}^2 + \hat{\mathcal{Z}}^2$ whenever $\mathcal{Z}^1, \mathcal{Z}^2 \in \mathcal{Z}$ and $\hat{\mathcal{Z}}^1, \hat{\mathcal{Z}}^2 \in \hat{\mathcal{Z}}$. Then, $\ominus$ has an obvious definition on $\mathbb{K}_w(\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m}))$ by (2.3). With these definitions, Lemma 3.6 can be rewritten for $\mathbb{K}_w(\mathbb{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathbb{R}^{d \times m}))$ except that, in (iii), the existence of the Hukuhara difference is not needed.

The next corollary is an important observation.

Corollary 3.7. *Suppose that $\mathbb{P}$ is a nonatomic probability measure. Let $\mathcal{Z} \in \mathscr{K}_w(\mathbb{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^{d \times m}))$ be a nonempty set of processes and $t \in [0,T]$. Then, it holds almost surely that*

$$\int_0^T \mathcal{Z} dB \subset \int_0^t \mathcal{Z} dB + \int_t^T \mathcal{Z} dB. \tag{3.12}$$

Moreover, if $\mathcal{Z}$ is decomposable, then it holds almost surely that

$$\int_0^T \mathcal{Z} dB = \int_0^t \mathcal{Z} dB + \int_t^T \mathcal{Z} dB, \quad \int_t^T \mathcal{Z} dB = \int_0^T \mathcal{Z} dB \ominus \int_0^t \mathcal{Z} dB. \tag{3.13}$$

Proof: Let $t \in [0,T]$. By [22, Lemma 3.3.4], we have $\mathcal{Z} \subset \mathbf{1}_{[0,t]}\mathcal{Z} + \mathbf{1}_{(t,T]}\mathcal{Z}$, and equality holds when $\mathcal{Z}$ is decomposable. Applying Lemma 3.6(i) together with the monotonicity of the integral with respect to $\subset$, the relations in (3.12) and (3.13) hold. ∎

Remark 3.8. The essence of Corollary 3.7 is that, unlike Corollary 3.5, the temporal additivity $\int_0^T = \int_0^t + \int_t^T$ is *not* necessarily true in the case of generalized stochastic integrals for lack of decomposability of the integrand. In particular, the Hukuhara difference $\int_0^T \mathcal{Z} dB \ominus \int_0^t \mathcal{Z} dB$ may not exist in general. This peculiar feature of generalized stochastic integrals will be particularly felt when we study the set-valued BSDEs in §6. ∎

4 Some Important Estimates

In this section, we establish some important estimates regarding set-valued conditional expectations and stochastic integrals. These estimates, albeit conceivable, need justifications given the special natures of the set-valued stochastic analysis, as well as the lack of a vector space structure in general. Some of the arguments are following those in [20] closely, but we nevertheless provide the details for the sake of completeness.

Recall the set $\mathscr{K}(\mathbb{R}^d)$, the collection of all nonempty convex compact subsets of $\mathbb{R}^d$. For $p \in [1, +\infty)$ and $X_1, X_2 \in \mathscr{L}^p_{\mathcal{F}}(\Omega, \mathscr{K}(\mathbb{R}^d))$, define

$$\mathcal{H}_p(X_1, X_2) := (\mathbb{E}[h^p(X_1, X_2)])^{\frac{1}{p}}. \tag{4.1}$$

The following result is a strengthened version of [20, Theorem 2.4.1] in the $\mathbb{L}^2$ sense.

Lemma 4.1. *Let $X_1, X_2 \in \mathscr{L}^2_{\mathcal{F}}(\Omega, \mathscr{K}(\mathbb{R}^d))$, and $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra. Then, one has*

$$h^2(\mathbb{E}[X_1|\mathcal{G}], \mathbb{E}[X_2|\mathcal{G}]) \le \mathbb{E}[h^2(X_1, X_2)|\mathcal{G}] \qquad \mathbb{P}\text{-}a.s. \tag{4.2}$$

18

In particular, the following inequalities hold:

$$\mathcal{H}_2(\mathbb{E}[X_1|\mathcal{G}], \mathbb{E}[X_2|\mathcal{G}]) \ \leq \ \mathcal{H}_2(X_1, X_2), \tag{4.3}$$

$$\|\mathbb{E}[X_1|\mathcal{G}]\|^2 \ \leq \ \mathbb{E}[\|X_1\|^2 \,|\mathcal{G}] \qquad \mathbb{P}\text{-}a.s. \tag{4.4}$$

Proof. Let us introduce the notation $\mathbb{E}[\xi : D] := \mathbb{E}[\xi \mathbf{1}_D]$ for $\xi \in \mathbb{L}^1_{\mathcal{F}}(\Omega, \mathbb{R})$ and $D \in \mathcal{F}$. Note that (4.2) is equivalent to

$$\mathbb{E}\big[h^2(\mathbb{E}[X_1|\mathcal{G}], \mathbb{E}[X_2|\mathcal{G}]) : D\big] \leq \mathbb{E}\big[h^2(X_1, X_2) : D\big], \qquad D \in \mathcal{G}. \tag{4.5}$$

Let $D \in \mathcal{G}$, and set $C := \{\omega \in \Omega : \bar{h}(\mathbb{E}[X_1|\mathcal{G}](\omega), \mathbb{E}[X_2|\mathcal{G}](\omega)) \geq \bar{h}(\mathbb{E}[X_1|\mathcal{G}](\omega), \mathbb{E}[X_2|\mathcal{G}](\omega))\}$. Clearly, by the definition of conditional expectation, $C \in \mathcal{G}$. Now we can write

$$\begin{aligned}
\mathbb{E}[h^2(\mathbb{E}[X_1|\mathcal{G}], \mathbb{E}[X_2|\mathcal{G}]) : D] &= \ \mathbb{E}[\bar{h}^2(\mathbb{E}[X_1|\mathcal{G}], \mathbb{E}[X_2|\mathcal{G}]) : D \cap C] \\
&\quad + \mathbb{E}[\bar{h}^2(\mathbb{E}[X_1|\mathcal{G}], \mathbb{E}[X_2|\mathcal{G}]) : D \cap C^c].
\end{aligned} \tag{4.6}$$

Repeatedly applying Proposition 2.11, we obtain

$$
\mathbb{E}\left[\bar{h}^2\left(\mathbb{E}\left[X_1|\mathcal{G}\right], \mathbb{E}\left[X_2|\mathcal{G}\right]\right) : D \cap C\right] = \int_{D \cap C} \sup_{x \in \mathbb{E}[X_1|\mathcal{G}](\omega)} d^2(x, \mathbb{E}\left[X_2|\mathcal{G}\right](\omega)) \mathbb{P}(d\omega)
$$

$$
= \sup_{\eta \in S(\mathbb{E}[X_1|\mathcal{G}])} \mathbb{E}\left[d^2(\eta, \mathbb{E}\left[X_2|\mathcal{G}\right]) : D \cap C\right] = \sup_{\eta \in \{\mathbb{E}[\varphi|\mathcal{G}] : \varphi \in S(X_1)\}} \mathbb{E}\left[d^2(\eta, \mathbb{E}\left[X_2|\mathcal{G}\right]) : D \cap C\right]
$$

$$
= \sup_{\varphi \in S(X_1)} \mathbb{E}\left[d^2(\mathbb{E}\left[\varphi|\mathcal{G}\right], \mathbb{E}\left[X_2|\mathcal{G}\right]) : D \cap C\right]
$$

$$
= \sup_{\varphi \in S(X_1)} \int_{D \cap C} \inf_{y \in \mathbb{E}[X_2|\mathcal{G}](\omega)} |\mathbb{E}[\varphi|\mathcal{G}](\omega) - y|^2 \mathbb{P}(d\omega)
$$

$$
= \sup_{\varphi \in S(X_1)} \inf_{\psi \in S(X_2)} \mathbb{E}\left[|\mathbb{E}\left[\varphi|\mathcal{G}\right] - \mathbb{E}\left[\psi|\mathcal{G}\right]|^2 : D \cap C\right]
$$

$$
= \sup_{\varphi \in S(X_1)} \inf_{\psi \in S(X_2)} \mathbb{E}\left[|\mathbb{E}\left[\varphi - \psi|\mathcal{G}\right]|^2 : D \cap C\right]
$$

$$
\leq \sup_{\varphi \in S(X_1)} \inf_{\psi \in S(X_2)} \mathbb{E}\left[\mathbb{E}\left[|\varphi - \psi|^2 \mid \mathcal{G}\right] : D \cap C\right] = \sup_{\varphi \in S(X_1)} \inf_{\psi \in S(X_2)} \mathbb{E}\left[|\varphi - \psi|^2 : D \cap C\right]
$$

$$
= \ \mathbb{E}\left[\bar{h}^2(X_1, X_2) : D \cap C\right] \leq \mathbb{E}\left[h^2(X_1, X_2) : D \cap C\right]. \tag{4.7}
$$

Here in the above, the inequality is due to the conditional version of Jensen's inequality. Similarly, we also have $\mathbb{E}\left[\bar{h}^2\left(\mathbb{E}\left[X_1|\mathcal{G}\right], \mathbb{E}\left[X_2|\mathcal{G}\right]\right) : D \cap C^c\right] \leq \mathbb{E}\left[h^2(X_1, X_2) : D \cap C^c\right]$. Combining the two inequalities with (4.6), we obtain (4.5) and hence (4.2). Then, (4.3) is immediate from (4.2). Finally, (4.4) follows from (4.2) by taking $X_2 \equiv \{0\}$. ∎

Next, we present a Hölder-type inequality regarding the Aumann integral. A similar inequality appears in [23, Theorem 2.1] for a special class of integrands. For completeness, we provide a full proof here for our version.

Proposition 4.2. *Let $\Phi^1, \Phi^2 \in \mathscr{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$, and $t \in [0,T]$. Then, it holds that*

$$h^2\left(\int_t^T \Phi_s^1 ds, \int_t^T \Phi_s^2 ds\right) \leq (T-t) \int_t^T h^2(\Phi_s^1, \Phi_s^2) ds \qquad \mathbb{P}\text{-a.s.}$$

Proof. Recalling the definition of the Hausdorff metric h, it suffices to show that

$$\begin{cases} \bar{h}^2\left(\int_t^T \Phi_s^1 ds, \int_t^T \Phi_s^2 ds\right) \leq (T-t)\int_t^T \bar{h}^2(\Phi_s^1, \Phi_s^2)ds & \mathbb{P}\text{-a.s.}, \\ \bar{h}^2\left(\int_t^T \Phi_s^2 ds, \int_t^T \Phi_s^1 ds\right) \leq (T-t)\int_t^T \bar{h}^2(\Phi_s^2, \Phi_s^1)ds & \mathbb{P}\text{-a.s.} \end{cases} \tag{4.8}$$

By symmetry, we shall check only the first inequality in (4.8). To begin with, we first note that the statement is equivalent to showing, for every $D \in \mathcal{F}_T$, that

$$\mathbb{E}\left[\bar{h}^2\left(\int_t^T \Phi_s^1 ds, \int_t^T \Phi_s^2 ds\right) : D\right] \leq (T-t)\mathbb{E}\left[\int_t^T \bar{h}^2(\Phi_s^1, \Phi_s^2)ds : D\right], \tag{4.9}$$

where we keep using the notation introduced in the proof of Lemma 4.1.

To see (4.9), we first note that, similar to (4.7), we have

$$\mathbb{E}\left[\bar{h}^2\left(\int_t^T \Phi_s^1 ds, \int_t^T \Phi_s^2 ds\right) : D\right] = \sup_{\eta_1 \in \overline{\mathrm{dec}} J_{t,T}(S^2_{\mathbb{F}}(\Phi^1))} \inf_{\eta_2 \in \overline{\mathrm{dec}} J_{t,T}(S^2_{\mathbb{F}}(\Phi^2))} \mathbb{E}[|\eta_1 - \eta_2|^2 : D]. \tag{4.10}$$

Next, by the standard Hölder's inequality, we have

$$\sup_{\eta_1 \in J_{t,T}(S^2_{\mathbb{F}}(\Phi^1))} \inf_{\eta_2 \in J_{t,T}(S^2_{\mathbb{F}}(\Phi^2))} \mathbb{E}[|\eta_1 - \eta_2|^2 : D] = \sup_{\varphi^1 \in S^2_{\mathbb{F}}(\Phi^1)} \inf_{\varphi^2 \in S^2_{\mathbb{F}}(\Phi^2)} \mathbb{E}\left[\left|J_{t,T}(\varphi^1) - J_{t,T}(\varphi^2)\right|^2 : D\right]$$

$$\leq (T-t) \sup_{\varphi^1 \in S^2_{\mathbb{F}}(\Phi^1)} \inf_{\varphi^2 \in S^2_{\mathbb{F}}(\Phi^2)} \mathbb{E}\left[\int_t^T |\varphi_s^1 - \varphi_s^2|^2 ds : D\right]. \tag{4.11}$$

Now, for given $D \in \mathcal{F}_T$, we consider the probability space $(D, \mathcal{F}_T^D, \mathbb{P}^D)$, where $\mathcal{F}_T^D := \{C \cap D : C \in \mathcal{F}_T\}$ and $\mathbb{P}^D(C) = [\mathbb{P}(C)/\mathbb{P}(D)]\mathbf{1}_{\{\mathbb{P}(D)>0\}}$, $C \in \mathcal{F}_T^D$. We also define the filtration $\mathbb{F}^D = \{\mathcal{F}_t^D\}_{t \in [0,T]}$ in a similar way. Applying Proposition 2.11 again, we have

$$\sup_{\varphi^1 \in S^2_{\mathbb{F}}(\Phi^1)} \inf_{\varphi^2 \in S^2_{\mathbb{F}}(\Phi^2)} \mathbb{E}\left[\int_t^T |\varphi_s^1 - \varphi_s^2|^2 ds : D\right] = \sup_{\varphi^1 \in S^2_{\mathbb{F}}(\Phi^1)} \inf_{\varphi^2 \in S^2_{\mathbb{F}}(\Phi^2)} \mathbb{E}^{\mathbb{P}^D}\left[\int_t^T |\varphi_s^1 - \varphi_s^2|^2 ds\right]$$

$$= \sup_{\varphi^1 \in S^2_{\mathbb{F}^D}(\Phi^1)} \inf_{\varphi^2 \in S^2_{\mathbb{F}^D}(\Phi^2)} \int_{D \times [t,T]} \left|\varphi_s^1(\omega) - \varphi_s^2(\omega)\right|^2 \mathbb{P}^D(d\omega)ds \tag{4.12}$$

$$= \int_{D \times [t,T]} \sup_{x \in \Phi_s^1(\omega)} \inf_{y \in \Phi_s^2(\omega)} |x - y|^2 \mathbb{P}^D(d\omega)ds = \int_{D \times [t,T]} \bar{h}^2(\Phi_s^1(\omega), \Phi_s^2(\omega))\mathbb{P}^D(d\omega)ds$$

$$= \mathbb{E}^{\mathbb{P}^D}\left[\int_t^T \bar{h}^2(\Phi_s^1, \Phi_s^2)ds\right] = \mathbb{E}\left[\int_t^T \bar{h}^2(\Phi_s^1, \Phi_s^2)ds : D\right].$$

Let $\alpha_D := (T - t)\mathbb{E}\Big[\int_t^T \bar{h}^2(\Phi_s^1, \Phi_s^2)ds : D\Big]$. Combining (4.11) and (4.12), we have

$$\sup_{\eta_1 \in J_{t,T}(S_\mathbb{F}^2(\Phi^1))} \inf_{\eta_2 \in J_{t,T}(S_\mathbb{F}^2(\Phi^2))} \mathbb{E}[|\eta_1 - \eta_2|^2 : D] \leq \alpha_D. \tag{4.13}$$

Next, we show that (4.13) implies that

$$\sup_{\eta_1 \in \operatorname{dec} J_{t,T}(S_\mathbb{F}^2(\Phi^1))} \inf_{\eta_2 \in J_{t,T}(S_\mathbb{F}^2(\Phi^2))} \mathbb{E}\Big[|\eta_1 - \eta_2|^2 : D\Big] \leq \alpha_D. \tag{4.14}$$

For every $\eta_1 \in \operatorname{dec} J_{t,T}(S_\mathbb{F}^2(\Phi^1))$, we may write $\eta_1 = \sum_{i=1}^m \mathbf{1}_{D_i}\eta_{1,i}$ for some $D_1, \ldots, D_m \in \mathcal{F}_T$ partitioning Ω, and $\eta_{1,1}, \ldots, \eta_{1,m} \in J_{t,T}(S_\mathbb{F}^2(\Phi^1))$. Then, for $\eta_2 \in J_{t,T}(S_\mathbb{F}^2(\Phi^2))$, we can apply Jensen's inequality to get

$$
\begin{aligned}
\mathbb{E}[|\eta_1 - \eta_2|^2 : D] &= \mathbb{E}^{\mathbb{P}^D}\Big[\Big|\sum_{i=1}^m \mathbf{1}_{D_i}(\eta_{1,i} - \eta_2)\Big|^2\Big] \leq \mathbb{E}^{\mathbb{P}^D}\Big[\sum_{i=1}^m \mathbf{1}_{D_i}|\eta_{1,i} - \eta_2|^2\Big] \\
&= \sum_{i=1}^m \mathbb{E}\big[\mathbf{1}_{D\cap D_i}|\eta_{1,i} - \eta_2|^2\big].
\end{aligned}
\tag{4.15}
$$

Since η_1 and η_2 are arbitrary, we deduce from (4.15) and (4.13) that

$$\sup_{\eta_1 \in \operatorname{dec} J_{t,T}(S_\mathbb{F}^2(\Phi^1))} \inf_{\eta_2 \in J_{t,T}(S_\mathbb{F}^2(\Phi^2))} \mathbb{E}[|\eta_1 - \eta_2|^2 : D]$$

$$\leq \sum_{i=1}^m \sup_{\eta_{1,i} \in J_{t,T}(S_\mathbb{F}^2(\Phi^1))} \inf_{\eta_2 \in J_{t,T}(S_\mathbb{F}^2(\Phi^2))} \mathbb{E}\big[|\eta_{1,i} - \eta_2|^2 : D \cap D_i\big] \leq \sum_{i=1}^m \alpha_{D\cap D_i} = \alpha.$$

This proves (4.14). Noting that $\overline{\operatorname{dec}}J_{t,T}(S_\mathbb{F}^2(\Phi^2)) \supset J_{t,T}(S_\mathbb{F}^2(\Phi^2))$, (4.14) implies that

$$\sup_{\eta_1 \in \operatorname{dec} J_{t,T}(S_\mathbb{F}^2(\Phi^1))} \inf_{\eta_2 \in \overline{\operatorname{dec}}J_{t,T}(S_\mathbb{F}^2(\Phi^2))} \mathbb{E}\Big[|\eta_1 - \eta_2|^2 : D\Big] \leq \alpha_D. \tag{4.16}$$

Finally, we claim that (4.16) implies

$$\sup_{\eta_1 \in \overline{\operatorname{dec}}J_{t,T}(S_\mathbb{F}^2(\Phi^1))} \inf_{\eta_2 \in \overline{\operatorname{dec}}J_{t,T}(S_\mathbb{F}^2(\Phi^2))} \mathbb{E}\Big[|\eta_1 - \eta_2|^2 : D\Big] \leq \alpha_D, \tag{4.17}$$

which, together with (4.10), would lead to (4.9). Indeed, let $\eta_1 \in \overline{\operatorname{dec}}J_{t,T}(S_\mathbb{F}^2(\Phi^1))$, and let $\{\eta_1^n\}_{n\in\mathbb{N}} \subset \operatorname{dec} J_{t,T}(S_\mathbb{F}^2(\Phi^1))$ be a sequence that converges to η_1 (strongly) in $\mathbb{L}^2_{\mathcal{F}_T}(\Omega, \mathbb{R}^d)$. Let $\varepsilon > 0$. For each $n \in \mathbb{N}$, thanks to (4.16), we may find $\eta_2^n \in \overline{\operatorname{dec}}J_{t,T}(S_\mathbb{F}^2(\Phi^2))$ such that

$$\mathbb{E}[|\eta_1^n - \eta_2^n|^2 : D] < \alpha_D + \varepsilon. \tag{4.18}$$

By Remark 3.2, $\{\eta_2^n\}_{n\in\mathbb{N}}$ is a bounded sequence in $\mathbb{L}^2_{\mathcal{F}_T}(\Omega, \mathbb{R}^d)$; hence, by Banach-Saks theorem, it has a subsequence $\{\eta_2^{n_k}\}_{k\in\mathbb{N}}$ for which the sequence $\{\bar{\eta}_2^k\}_{k\in\mathbb{N}}$ converges to some

21

$\overline{\eta}_2 \in \mathbb{L}^2_{\mathcal{F}_T}(\Omega, \mathbb{R}^d)$ strongly, where $\overline{\eta}_2^k := \frac{1}{k} \sum_{\ell=1}^k \eta_2^{n_\ell}$ is the Cesàro average, for $k \in \mathbb{N}$. Moreover, since $\overline{\mathrm{dec}} J_{t,T}(S^2_{\mathbb{F}}(\Phi^2))$ is a closed convex set, all Cesàro averages and their limit $\overline{\eta}_2$ belong to $\overline{\mathrm{dec}} J_{t,T}(S^2_{\mathbb{F}}(\Phi^2))$. The strong convergence of $\{\eta_1^n\}_{n \in \mathbb{N}}$ implies that $\{\overline{\eta}_1^k\}_{k \in \mathbb{N}} \subset \overline{\mathrm{dec}} J_{t,T}(S^2_{\mathbb{F}}(\Phi^1))$ converges to η_1 strongly in $\mathbb{L}^2_{\mathcal{F}_T}(\Omega, \mathbb{R}^d)$, where $\overline{\eta}_1^k := \frac{1}{k} \sum_{\ell=1}^k \eta_1^{n_\ell}$, $k \in \mathbb{N}$. By (4.18), we have

$$\mathbb{E}[|\overline{\eta}_1^k - \overline{\eta}_2^k|^2 : D] \leq \left(\frac{1}{k} \sum_{\ell=1}^k \left(\mathbb{E}[|\eta_1^{n_\ell} - \eta_2^{n_\ell}|^2 : D] \right)^{\frac{1}{2}} \right)^2 < \alpha_D + \varepsilon, \quad k \in \mathbb{N}.$$

Thus,

$$
\begin{aligned}
(\mathbb{E}[|\eta_1 - \overline{\eta}_2|^2 : D])^{\frac{1}{2}} &\leq (\mathbb{E}[|\eta_1 - \overline{\eta}_1^k|^2 : D])^{\frac{1}{2}} + (\mathbb{E}[|\overline{\eta}_1^k - \overline{\eta}_2^k|^2 : D])^{\frac{1}{2}} + (\mathbb{E}[|\overline{\eta}_2^k - \overline{\eta}_2|^2 : D])^{\frac{1}{2}} \\
&\leq (\mathbb{E}[|\eta_1 - \overline{\eta}_1^k|^2])^{\frac{1}{2}} + (\alpha_D + \varepsilon)^{\frac{1}{2}} + (\mathbb{E}[|\overline{\eta}_2^k - \overline{\eta}_2|^2])^{\frac{1}{2}},
\end{aligned}
$$

and letting $k \to \infty$ yields

$$\inf_{\eta_2 \in \overline{\mathrm{dec}} J_{t,T}(S^2_{\mathbb{F}}(\Phi^2))} \mathbb{E}[|\eta_1 - \eta_2|^2 : D] \leq \mathbb{E}[|\eta_1 - \overline{\eta}_2|^2 : D] \leq \alpha + \varepsilon.$$

Since $\varepsilon > 0$ and $\eta_1 \in \overline{\mathrm{dec}} J_{t,T}(S_{\mathbb{F}}(\Phi^1))$ are arbitrary, (4.17) follows, concluding the proof. ∎

5 Set-Valued Martingales and their Integral Repsentations

Using the notion of *conditional expectation* in §3.1, one can define *set-valued martingales* as follows. We say that a set-valued process $M = \{M_t\}_{t \in [0,T]}$ is a *set-valued $\mathbb{F}$-martingale* if $M \in \mathscr{L}^0_{\mathbb{F}}([0,T] \times \Omega, \mathscr{C}(\mathbb{R}^d))$, $M_t \in \mathscr{A}^1_{\mathcal{F}_t}(\Omega, \mathscr{C}(\mathbb{R}^d))$, and $M_s = \mathbb{E}[M_t | \mathcal{F}_s]$ for all $0 \leq s \leq t$. M is called *square-integrable* if $M_t \in \mathscr{A}^2_{\mathcal{F}_t}(\Omega, \mathscr{C}(\mathbb{R}^d))$, $0 \leq t \leq T$, and *uniformly square-integrably bounded* if there exists $\ell \in \mathbb{L}^2(\Omega, \mathbb{R}_+)$ such that $\sup_{t \in [0,T]} \|M_t\| \leq \ell$ $\mathbb{P}$-a.s.

We note that, if M is a square-integrable set-valued martingale, then for each $t \in [0,T]$, the set of square-integrable selections, $S^2_{\mathcal{F}_t}(M_t)$, is decomposable. On the other hand, we consider the set of all square-integrable martingale selections, that is, all d-dimensional $\mathbb{F}$-martingales $f = \{f_t\}_{t \in [0,T]}$ such that $f_t \in S^2_{\mathcal{F}_t}(M_t)$, $t \in [0,T]$, and denote it by $MS(M)$. If M is convex-valued, then it is known that $MS(M) \neq \emptyset$ (cf. [21, §3]). For $t \in [0,T]$, consider the t-section of $MS(M)$, defined as $P_t[MS(M)] := \{f_t : f \in MS(M)\} \subset \mathbb{L}^2_{\mathcal{F}_t}(\Omega, \mathbb{R}^d)$. We remark that the two sets $S^2_{\mathcal{F}_t}(M_t)$ (the selections of the t-section) and $P_t[MS(M)]$ (the t-section of the selections) are quite different. In particular, the former is known to be decomposable, but the latter is not. However, the following relation holds (see [21, Proposition 3.1]):

$$S^2_{\mathcal{F}_t}(M_t) = \overline{\mathrm{dec}}_{\mathcal{F}_t}(P_t[MS(M)]), \qquad t \in [0,T], \tag{5.19}$$

22

where $\overline{\mathrm{dec}}_{\mathcal{F}_t}$ denotes the closed decomposable hull with respect to $\mathbb{L}^2_{\mathcal{F}_t}(\Omega, \mathbb{R}^d)$.

5.1 Representation of Martingales with Trivial Initial Value

In what follows, we assume that $\mathbb{F} = \mathbb{F}^B$, for some m-dimensional standard Brownian motion $B = \{B_t\}_{t \in [0,T]}$. The fundamental building block of the theory of BSDEs is the celebrated *martingale representation theorem*, which states that every square-integrable $\mathbb{F}$-martingale can be written, uniquely, as a stochastic integral against B, whence continuous. There is a similar result for set-valued martingales (see §3.2), which we now describe.

Let M be a convex set-valued $\mathbb{F}$-martingale that is square-integrable, i.e., we have $M_t \in \mathscr{A}^2_{\mathcal{F}_t}(\Omega, \mathscr{C}(\mathbb{R}^d))$ for each $t \in [0,T]$. Then, for each $y \in MS(M)$, by the standard martingale representation theorem, there exists unique $z^y \in \mathbb{L}^2_{\mathbb{F}}([0,T], \mathbb{R}^{d \times m})$ such that $y_t = \int_0^t z^y_s dB_s$, $t \in [0,T]$, $\mathbb{P}$-a.s. Denote $\mathcal{Z}^M := \{z^y : y \in MS(M)\} \in \mathscr{P}(\mathbb{L}^2_{\mathbb{F}}([0,T], \mathbb{R}^{d \times m}))$.

Remark 5.1. We should note that while a set-valued martingale always gives rise to a set of vector-valued martingales, i.e., stochastic integrals, not every set of vector-valued martingales can be realized as $MS(M)$ for some set-valued martingale M. ∎

The following *set-valued martingale representation theorem* is due to [21].

Theorem 5.2 (Kisielewicz [21, Proposition 4.1, Theorem 4.2]). *For every convex square-integrable set-valued martingale $M = \{M_t\}_{t \in [0,T]}$ with $M_0 = \{0\}$, there exists $\mathcal{Z}^M \in \mathscr{P}(\mathbb{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^{d \times m}))$ such that $M_t = \int_0^t \mathcal{Z}^M dB$ $\mathbb{P}$-a.s. for every $t \in [0,T]$. If M is also uniformly square-integrably bounded, then $\mathcal{Z}^M$ is a convex weakly compact set, that is, $\mathcal{Z}^M \in \mathscr{K}_w(\mathbb{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^{d \times m}))$.*

Remark 5.3. (i) We first note that in the set-valued martingale representation, the "martingale integrand" $\mathcal{Z}^M$ may not correspond to a measurable set-valued process. In fact, if the set-valued martingale is square-integrably bounded, then the integrand $\mathcal{Z}^M$ *cannot* be decomposable unless it is a singleton (see [22, Corollary 5.3.2]). Thus the stochastic integral can only be in the generalized sense. But on the other hand, if $\mathcal{Z}^M$ is not decomposable, then the temporal additivity of the set-valued stochastic integral fails in general (see, Corollary 3.7, Remark 3.8), i.e., we may not write $M_T = M_t + \int_t^T \mathcal{Z}^M dB$, and $M_T \ominus M_t$ may not even exist in general. Such a conflict leads to some fundamental difficulties for the study of set-valued BSDEs, and it does not seem to be amendable unless some more general framework of set-valued stochastic integrals is established.

(ii) If Ω is separable, then there exists a sequence $\{z_n\}_{n \geq 1} \subset \mathbb{L}^2_{\mathbb{F}}([0.T], \mathbb{R}^{d \times m})$ such that $M_t = \mathrm{cl}\{\int_0^t z^n_s dB_s\}_{n \geq 1}$, $S^2_{\mathcal{F}_t}(M_t) = \overline{\mathrm{dec}}_{\mathcal{F}_t}\{\int_0^t z^n_s dB_s\}_{n \geq 1}$, $t \in [0,T]$ (see [21, Theorem 4.3]).

(iii) If M is a uniformly square-integrably bounded martingale and $\mathbb{P}$ is nonatomic, then there exists a sequence $\{z_n\}_{n\geq 1} \subset \mathbb{L}^2_{\mathbb{F}}([0.T], \mathbb{R}^{d\times m})$ such that $M_t = \overline{\text{co}}\{\int_0^t z_s^n dB_s\}_{n\geq 1}$ for all $t \in [0, T]$ (see [21, Theorem 4.3]).

(iv) In light of the equivalence relation $\cong$ in (3.10), in the last part of Theorem 5.2, we can easily conclude that such $\mathcal{Z}^M$ is unique in $\mathbb{K}_w(\mathbb{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^{d\times m}))$. Indeed, if there exist $\mathcal{Z}_1^M$ and $\mathcal{Z}_2^M$ in $\mathscr{K}_w(\mathbb{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^{d\times m}))$ such that $\int_0^t \mathcal{Z}_1^M dB = \int_0^t \mathcal{Z}_2^M dB = M_t$, $t \in [0, T]$, then $\mathcal{Z}_1^M \cong \mathcal{Z}_2^M$, that is, they correspond to the same element of $\mathbb{K}_w(\mathbb{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^{d\times m}))$ and we may denote this element by $\mathcal{Z}^M$ with a slight abuse of notation.

(v) Unlike usual stochastic integrals, set-valued stochastic integrals do not always generate set-valued martingales. In fact, given a nonempty set $\mathcal{Z} \in \mathscr{K}_w(\mathbb{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^{d\times m}))$ (or $\mathcal{Z} \in \mathbb{K}_w(\mathbb{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^{d\times m}))$) of processes, the set-valued process $\{\int_0^t \mathcal{Z} dB\}_{t\in[0,T]}$ forms a set-valued *submartingale* in the sense that $\int_0^u \mathcal{Z} dB \subset \mathbb{E}[\int_0^t \mathcal{Z} dB_s | \mathcal{F}_u]$ for every $0 \leq u \leq t \leq T$ (see [24, Theorem 4.2]). Nevertheless, the stochastic integrals that appear in Theorem 5.2 are naturally martingales. ∎

5.2 Representation of Martingales with General Initial Value

We would like to point out that, in Theorem 5.2, it is assumed that $M_0 = \{0\}$. Such a seemingly benign assumption actually has some severe consequences. In particular, as it was pointed out recently in [35], a set-valued martingale whose initial value is a singleton is essentially a vector-valued martingale. Therefore, Theorem 5.2 actually is not a suitable tool for the study of set-valued BSDEs with non-singleton terminal values. The main purpose of this subsection is to establish a refined version of set-valued martingale representation theorem for set-valued martingales with general (non-singleton) initial values.

Our idea is to extend the notion of Aumann-Itô integral so that it is a martingale but its expectation is not necessarily zero (see [35, Example 3.1] for the set-valued dilemma). To this end, for every $t \in [0, T]$, we consider the space $\mathbb{R}_t := \mathbb{L}^2_{\mathcal{F}_t}(\Omega, \mathbb{R}^d) \times \mathbb{L}^2_{\mathbb{F}}([t, T] \times \Omega, \mathbb{R}^{d\times m})$.

Given a process $z = \{z_u\}_{u\in[0,T]}$, we denote $z^{t,T} := (z_u)_{u\in[t,T]}$ to be the restriction of z onto the interval $[t, T]$, and define a mapping $F^t : \mathbb{R}_0 \to \mathbb{R}_t$ by

$$F^t(x, z) := \left(x + \int_0^t z_s dB_s, z^{t,T} \right), \quad (x, z) \in \mathbb{R}_0.$$

We have the following result.

Lemma 5.4. *For given $t \in [0, T]$ and $(\xi, z^t) \in \mathbb{R}_t$, define a process $\mathcal{J}^t(\xi, z^t) = \{\mathcal{J}^t_u(\xi, z^t)\}_{u\in[0,T]}$:*

$$\mathcal{J}^t_u(\xi, z^t) := \mathbb{E}[\xi|\mathcal{F}_u]\mathbf{1}_{[0,t)}(u) + \left(\xi + \int_t^u z_s^t dB_s \right)\mathbf{1}_{[t,T]}(u), \quad u \in [0, T]. \quad (5.20)$$

Then, $\mathcal{J}^t(\xi, z^t)$ is an $\mathbb{F}$-martingale on $[0, T]$. Moreover, it holds that $\mathcal{J}^t \circ F^t = \mathcal{J}^0$ on $\mathbb{R}_0$.

Proof. That $\mathcal{J}^t(\xi, z^t)$ is a martingale is obvious. To check the identity, let $(x, z) \in \mathbb{R}_0$. Following the definitions of $\mathcal{J}^t, F^t, \mathcal{J}^0$, we have

$$
\begin{aligned}
\mathcal{J}_u^t(F^t(x, z)) &= \mathcal{J}_u^t\Big(x + \int_0^t z_s dB_s, z^{t,T}\Big) \\
&= \mathbb{E}\Big[x + \int_0^t z_s dB_s \Big| \mathcal{F}_u\Big] \mathbf{1}_{[0,t)}(u) + \Big(x + \int_0^t z_s dB_s + \int_t^u z_s dB_s\Big) \mathbf{1}_{[t,T]}(u) \\
&= \Big(x + \int_0^u z_s dB_s\Big) \mathbf{1}_{[0,t)}(u) + \Big(x + \int_0^u z_s dB_s\Big) \mathbf{1}_{[t,T]}(u) = \mathcal{J}_u^0(x, z)
\end{aligned}
$$

for every $u \in [0, T]$. Hence, $\mathcal{J}^t(F^t(x, z)) = \mathcal{J}^0(x, z)$. ∎

Next, let $\mathcal{R} \subset \mathbb{R}_0$ be a nonempty set and $t \in [0, T]$. By virtue of Theorem 2.8, there exists a set-valued random variable in $\mathbb{L}_{\mathcal{F}_t}^2(\Omega, \mathscr{C}(\mathbb{R}^d))$, denoted by $\int_{0-}^t \mathcal{R} dB$, such that

$$
S_{\mathcal{F}_t}^2\Big(\int_{0-}^t \mathcal{R} dB\Big) = \overline{\mathrm{dec}}_{\mathcal{F}_t}(\mathcal{J}_t^0[\mathcal{R}]). \tag{5.21}
$$

We call $\int_{0-}^t \mathcal{R} dB$ the *stochastic integral* of $\mathcal{R}$. Clearly, such a stochastic integral is an extended version of the generalized Aumann-Itô stochastic integral $\int_0^t \mathcal{Z} dB$ defined by (3.7), and in particular, the integrand $\mathcal{R}$ consists of pairs (x, z), which keeps track of the initial values x of the martingales in $\mathcal{J}^0[\mathcal{R}]$, motivating the choice of the notation $\int_{0-}^t$.

To see how the integral $\int_{0-}^t \mathcal{R} dB$ (or more precisely, $\mathcal{R}$) can be defined through a set-valued martingale, let $M = \{M_u\}_{u \in [0,T]}$ be a convex uniformly square-integrably bounded set-valued martingale with respect to $\mathbb{F} = \mathbb{F}^B$, and M_0 is a non-singleton convex set. Let $MS(M)$ be the set of all $\mathbb{L}^2$-martingale selections of M. By standard martingale representation theorem, for fixed $t \in [0, T]$, each $y \in MS(M)$ can be written as $y = \mathcal{J}^t(\xi, z)$ for a unique pair $(\xi, z) \in \mathbb{R}_t$. We shall define, for each $t \in [0, T]$,

$$
\mathcal{R}_t^M := \big\{(\xi, z) \in \mathbb{R}_t : \mathcal{J}^t(\xi, z) \in MS(M)\big\}; \quad \text{and} \quad \mathcal{R}^M := \mathcal{R}_0^M. \tag{5.22}
$$

In what follows, for $(\xi, z) \in \mathbb{R}_t$, we write $\pi_\xi(\xi, z) := \xi$ and $\pi_z(\xi, z) := z$. (For convenience, we suppress the dependence of the mappings π_ξ, π_z on t.) Furthermore, we define $\mathcal{Z}_t^M := \pi_z[\mathcal{R}_t^M]$. The following theorem collects various forms of "time-consistency" properties of the collection $\{\mathcal{R}_t^M\}_{t \in [0,T]}$, which will be useful in our future discussion.

Proposition 5.5 (Time-consistency)**.** *Let $t_1 \in [0, T]$. Then, it holds*

$$
F^{t_1}[\mathcal{R}_0^M] = \mathcal{R}_{t_1}^M. \tag{5.23}
$$

Furthermore, the following relations hold for every $t_2 \in (t_1, T]$:

(i) $\mathcal{J}^{t_1}[\mathcal{R}_{t_1}^M] = \mathcal{J}^{t_2}[\mathcal{R}_{t_2}^M] = MS(M)$.

(ii) $\pi_\xi[\mathcal{R}_0^M] = M_0$.

(iii) $\pi_\xi[\mathcal{R}_{t_1}^M] = J_{t_1}^{t_1}[\mathcal{R}_{t_1}^M] = J_{t_1}^{t_2}[\mathcal{R}_{t_2}^M] = P_{t_1}[MS(M)]$.

(iv) $\pi_\xi[\mathcal{R}_{t_1}^M] = \{\mathbb{E}[\xi \mid \mathcal{F}_{t_1}] \colon \xi \in \pi_\xi[\mathcal{R}_{t_2}^M]\}$.

(v) $\mathcal{Z}_{t_1}^M \mathbf{1}_{[t_2, T]} = \mathcal{Z}_{t_2}^M \mathbf{1}_{[t_2, T]} = \mathcal{Z}_0^M \mathbf{1}_{[t_2, T]}$.

Proof. We first prove (5.23). Fix $t_1 \in [0, T]$ and let $(x, z) \in \mathcal{R}_0^M$. By Lemma 5.4 and the definition of $\mathcal{R}_0^M$, we have $\mathcal{J}^{t_1}(F^{t_1}(x, z)) = \mathcal{J}^0(x, z) \in MS(M)$. On the other hand, since $(x, z) \in \mathbb{R}_0$, we have $F^{t_1}(x, z) \in \mathbb{R}_{t_1}$, which implies $F^{t_1}(x, z) \in \mathcal{R}_{t_1}^M$. Namely, $F^{t_1}[\mathcal{R}_0^M] \subset \mathcal{R}_{t_1}^M$.

Conversely, let $(\hat{\xi}, \hat{z}) \in \mathcal{R}_{t_1}^M$. Define a martingale $y_s := \mathbb{E}[\hat{\xi}|\mathcal{F}_s]$, $s \in [0, t_1]$. By martingale representation theorem, there exists a unique pair $(x, \bar{z}) \in \mathbb{R}_0$ such that

$$y_u = x + \int_0^u \bar{z}_s dB_s, \quad u \in [0, t_1].$$

Let $z := \bar{z}\mathbf{1}_{[0, t_1)} + \hat{z}\mathbf{1}_{[t_1, T]} \in \mathbb{L}_\mathbb{F}^2([0, T] \times \Omega, \mathbb{R}^{d \times m})$. Then, $(x, z) \in \mathbb{R}_0$ and for every $u \in [0, T]$,

$$
\begin{aligned}
\mathcal{J}_u^0(x, z) &= x + \int_0^u z_s dB_s = \left(x + \int_0^u \bar{z}_s dB_s\right) \mathbf{1}_{[0, t_1)}(u) + \left(\hat{\xi} + \int_{t_1}^u \hat{z}_s dB_s\right) \mathbf{1}_{[t_1, T]}(u) \\
&= \mathbb{E}[\hat{\xi} \mid \mathcal{F}_u]\mathbf{1}_{[0, t_1)}(u) + \left(\hat{\xi} + \int_{t_1}^u \hat{z}_s dB_s\right) \mathbf{1}_{[t_1, T]}(u) = \mathcal{J}_u^{t_1}(\hat{\xi}, \hat{z}).
\end{aligned}
$$

Hence, $\mathcal{J}^0(x, z) = \mathcal{J}^{t_1}(\hat{\xi}, \hat{z}) \in MS(M)$, that is, $(x, z) \in \mathcal{R}_0^M$. Finally,

$$F^{t_1}(x, z) = \left(x + \int_0^{t_1} z_s dB_s, z^{t_1, T}\right) = (\hat{\xi}, \hat{z}).$$

So $(\hat{\xi}, \hat{z}) \in F^{t_1}[\mathcal{R}_0^M]$. Consequently, we have $\mathcal{R}_{t_1}^M \subset F^0[\mathcal{R}_0^M]$, proving (5.23).

We now turn to properties (i)–(v). The proof of (i) is immediate since $\mathcal{J}^{t_i}[\mathcal{R}_{t_i}^M] = MS(M)$ by the definition of $\mathcal{R}_{t_i}^M$ for $i \in \{1, 2\}$.

To see (ii), let $(x, z) \in \mathcal{R}_0^M$. Since $\mathcal{J}^0(x, z) \in MS(M)$, we have $\pi_\xi(x, z) = x = \mathcal{J}_0^0(x, z) \in M_0$. Conversely, since M is a set-valued martingale, $M_0 = \mathbb{E}[M_T|\mathcal{F}_0] = \mathbb{E}[M_T]$, thanks to Blumenthal's 0-1 law. Hence, by the definition of set-valued expectation, for every $x \in M_0$, there exists $\xi \in S_{\mathcal{F}_T}^2(M_T)$ such that $x = \mathbb{E}[\xi]$. Furthermore, by the martingale representation theorem, there exists $z \in \mathbb{L}_\mathbb{F}^2([0, T] \times \Omega, \mathbb{R}^{d \times m})$ such that

$$\mathbb{E}[\xi|\mathcal{F}_u] = x + \int_0^u z_s dB_s = \mathcal{J}_u^0(x, z), \quad u \in [0, T].$$

Note that, as M is a set-valued martingale, we have $\mathbb{E}[\xi|\mathcal{F}_u] \in S_{\mathcal{F}_u}^2(M_u)$, $u \in [0, T]$. Hence, $J^0(x, z) = \{\mathbb{E}[\xi|\mathcal{F}_u]\}_{u \in [0, T]} \in MS(M)$, that is, $(x, z) \in \mathcal{R}_0^M$, or $x \in \pi_\xi[\mathcal{R}_0^M]$, proving (ii).

26

To prove (iii), first note that, for every $(x, z) \in \mathbb{R}_0$,

$$\pi_\xi(F^{t_1}(x, z)) = \pi_\xi\left(x + \int_0^{t_1} z_s dB_s, z^{t_1, T}\right) = x + \int_0^{t_1} z_s dB_s = \mathcal{J}_{t_1}^0(x, z).$$

This implies that

$$\pi_\xi[\mathcal{R}_{t_1}^M] = \pi_\xi[F^t[\mathcal{R}_0^M]] = \left\{\mathcal{J}_{t_1}^0(x, z) \colon (x, z) \in \mathcal{R}_0^M\right\} = \mathcal{J}_{t_1}^0[\mathcal{R}_0^M] = \mathcal{J}_{t_1}^{t_1} \circ F^{t_1}[\mathcal{R}_0^M] = \mathcal{J}_{t_1}^{t_1}[\mathcal{R}_{t_1}^M],$$

where the first and last equalities are by (5.23) and the fourth equality is due to Lemma 5.4. On the other hand, by the definitions of $P_{t_1}, \mathcal{J}^{t_1}$, we see that $P_{t_1} \circ \mathcal{J}^{t_1} = \mathcal{J}_{t_1}^{t_1}$. Therefore,

$$\pi_\xi[\mathcal{R}_{t_1}^M] = \mathcal{J}_{t_1}^{t_1}[\mathcal{R}_{t_1}^M] = P_{t_1}[\mathcal{J}^{t_1}[\mathcal{R}_{t_1}^M]] = P_{t_1}[\mathcal{J}^{t_2}[\mathcal{R}_{t_2}^M]] = P_{t_1}[MS(M)]$$

thanks to (i), which concludes the proof of (iii).

To prove (iv), first note that $\mathbb{E}[P_{t_2}(y)|\mathcal{F}_{t_1}] = P_{t_1}(y)$ whenever $y = \{y_u\}_{u \in [0,T]}$ is a martingale. Hence, applying (iii) twice, we obtain

$$\{\mathbb{E}[\xi|\mathcal{F}_{t_1}] \colon \xi \in \pi_\xi[\mathcal{R}_{t_2}^M]\} = \{\mathbb{E}[\xi|\mathcal{F}_{t_1}] \colon \xi \in P_{t_2}[MS(M)]\} = P_{t_1}[MS(M)] = \pi_\xi[\mathcal{R}_{t_1}^M].$$

Finally, to prove (v), note that, for every $(x, z) \in \mathbb{R}_0$,

$$\pi_z(F^{t_1}(x, z)) = \pi_z\left(x + \int_0^{t_1} z_s dB_s, z^{t_1, T}\right) = z^{t_1, T}.$$

Hence,

$$
\begin{aligned}
\mathcal{R}_{t_1}^M \mathbf{1}_{[t_2, T]} &= \pi_z[\mathcal{R}_{t_1}^M]\mathbf{1}_{[t_2, T]} = \pi_z[F^{t_1}[\mathcal{R}_0^M]]\mathbf{1}_{[t_2, T]} \\
&= \{z^{t_1, T} \colon (x, z) \in \mathcal{R}_0^M\}\mathbf{1}_{[t_2, T]} = \{z\mathbf{1}_{[t_2, T]} \colon z \in \mathcal{Z}_0^M\} = \mathcal{Z}_0^M \mathbf{1}_{[t_2, T]}.
\end{aligned}
$$

Taking $t_1 = t_2$ above, we also obtain $\mathcal{Z}_{t_2}^M \mathbf{1}_{[t_2, T]} = \mathcal{Z}_0^M \mathbf{1}_{[t_2, T]}$. $\blacksquare$

The following theorem is a martingale representation theorem for set-valued martingales with possibly nontrivial initial values, i.e., M_0 is a non-singleton deterministic set.

Theorem 5.6. *Let $M = \{M_u\}_{u \in [0,T]}$ be a convex uniformly square-integrably bounded set-valued martingale with respect to $\mathbb{F} = \mathbb{F}^B$. Then, for each $u \in [0, T]$, it holds*

$$M_u = \int_{0-}^u \mathcal{R}^M dB \quad \mathbb{P}\text{-}a.s.,$$

where $\mathcal{R}^M$ is defined by (5.22). Moreover, for each $t \in [0, u]$, it holds that $S_{\mathcal{F}_u}^2(M_u) = \overline{\text{dec}}_{\mathcal{F}_u}(\mathcal{J}_u^t[\mathcal{R}_t^M])$.

27

Proof: By Lemma 5.5(ii), we have $\mathcal{J}_u^0[\mathcal{R}^M] = P_u[\mathcal{J}^0[\mathcal{R}^M]] = P_u[MS(M)]$, $u \in [0, T]$. On the other hand, by [21, Proposition 3.1], we have $\overline{\text{dec}}_{\mathcal{F}_u}(P_u[MS(M)]) = S_{\mathcal{F}_u}^2(M_u)$. Combining these with the definition of stochastic integral in (5.21), we get

$$S_{\mathcal{F}_u}^2\Big(\int_{0-}^u \mathcal{R}^M dB\Big) = \overline{\text{dec}}_{\mathcal{F}_u}(\mathcal{J}_u^0[\mathcal{R}^M]) = \overline{\text{dec}}_{\mathcal{F}_u}(P_u[MS(M)]) = S_{\mathcal{F}_u}^2(M_u).$$

This shows that $M_u = \int_{0-}^u \mathcal{R}^M dB$ $\mathbb{P}$-a.s. The second part of the theorem is an immediate consequence of Lemma 5.4. ∎

Remark 5.7. It is interesting to note the relationship between the new stochastic integral $\int_{0-}^u \mathcal{R}^M dB$ (5.21) and the generalized Aumann-Itô stochastic integral $\int_0^u \mathcal{Z}^M dB$ (3.7), where $\mathcal{Z}^M := \mathcal{Z}_0^M$. Recalling (3.3) and (5.20), we have

$$\mathcal{J}_u^0(x, z) = x + \mathcal{J}_u(z) \in M_0 + \mathcal{J}_u[\mathcal{Z}^M]$$

for every $(x, z) \in \mathcal{R}^M$. Hence, $\mathcal{J}_u^0[\mathcal{R}^M] \subset M_0 + \mathcal{J}_u[\mathcal{Z}^M]$. After taking closed decomposable hulls, it follows that

$$S_u^2\Big(\int_{0-}^u \mathcal{R}^M dB\Big) = \overline{\text{dec}}_{\mathcal{F}_u}(\mathcal{J}_u^0[\mathcal{R}^M]) \subset M_0 + \overline{\text{dec}}_{\mathcal{F}_u}(\mathcal{J}_u[\mathcal{Z}^M]) = M_0 + S_{\mathcal{F}_u}^2\Big(\int_0^u \mathcal{Z}^M dB\Big).$$

Therefore,

$$\int_{0-}^u \mathcal{R}^M dB \subset M_0 + \int_0^u \mathcal{Z}^M dB \quad \mathbb{P}\text{-a.s.}$$

and the reverse inclusion fails to hold in general. When $M_0 = \{0\}$, we have $\int_{0-}^u \mathcal{R}^M dB = \int_0^u \mathcal{Z}^M dB$ since $\mathcal{R}^M = \{0\} \times \mathcal{Z}^M$ in this case. ∎

Remark 5.8. In view of Remark 5.7 and Theorem 5.6, the new integral $\int_{0-}^u \mathcal{R}^M dB$ is a non-trivial and necessary extension of the generalized Aumann-Itô stochastic integral, which can be used for the integral representation of any truly set-valued martingale M with a non-zero (non-singleton) initial value M_0. ∎

6   Set-Valued BSDEs

We are now ready to study the set-valued BSDEs. Assume from now on that $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ is a filtered probability space on which is defined an m-dimensional standard Brownian motion $B = \{B_t\}_{t \in [0,T]}$. We assume further that $\mathbb{F} = \mathbb{F}^B$, the natural filtration generated by B, augmented by all the $\mathbb{P}$-null sets of $\mathcal{F}$ so that it satisfies the *usual hypotheses*. In particular, we may assume without loss of generality that $(\Omega, \mathcal{F}) = (\mathbb{C}([0, T]), \mathscr{B}(\mathbb{C}([0, T])))$ is the canonical space with $\mathcal{F}_t = \sigma\{\omega(\cdot \wedge t), \omega \in \Omega\}$, $t \in [0, T]$, and $\mathbb{P}$ is the Wiener measure on $(\Omega, \mathcal{F})$. Hence, Ω is separable and $\mathbb{P}$ is nonatomic.

6.1 Set-Valued BSDEs in Conditional Expectation Form

In this section, we shall focus on the following simplest form of set-valued BSDE:

$$Y_t = \mathbb{E}\Big[\xi + \int_t^T f(s, Y_s)ds \ \Big| \ \mathcal{F}_t\Big], \qquad t \in [0, T]. \tag{6.1}$$

where $\xi \in \mathscr{L}^2_{\mathcal{F}_T}(\Omega, \mathscr{K}(\mathbb{R}^d))$, $f \colon [0, T] \times \Omega \times \mathscr{K}(\mathbb{R}^d) \to \mathscr{K}(\mathbb{R}^d)$ is a set-valued function to be specified later. We first give the definition of the solution to the set-valued BSDE (6.1).

Definition 6.1. *A set-valued process $Y \in \mathscr{L}^2_{\mathbb{F}}([0, T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$ is called an adapted solution to the set-valued BSDE (6.1) if, for each $t \in [0, T]$, we have*

$$Y_t = \mathbb{E}\Big[\xi + \int_t^T f(s, Y_s)ds \ \Big| \ \mathcal{F}_t\Big] \qquad \mathbb{P}\text{-a.s.}$$

We shall make use of the following assumptions on the coefficient f.

Assumption 6.2. *The function $f : [0, T] \times \Omega \times \mathscr{K}(\mathbb{R}^d) \to \mathscr{K}(\mathbb{R}^d)$ enjoys the following properties:*

(i) for fixed $A \in \mathscr{K}(\mathbb{R}^d)$, $f(\cdot, \cdot, A) \in \mathscr{L}^0_{\mathbb{F}}([0, T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$;

(ii) $f(\cdot, \cdot, \{0\}) \in \mathscr{L}^2_{\mathbb{F}}([0, T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$, that is,

$$\mathbb{E}\Big[\int_0^T \|f(t, \{0\})\|^2 dt\Big] = \mathbb{E}\Big[\int_0^T h^2(f(t, \{0\}), \{0\})dt\Big] < \infty; \tag{6.2}$$

(iii) $f(t, \omega, \cdot)$ is Lipschitz, uniformly in $(t, \omega) \in [0, T] \times \Omega$, in the following sense: there exists $K > 0$ such that

$$h(f(t, \omega, A), f(t, \omega, B)) \le K h(A, B), \qquad A, B \in \mathscr{K}(\mathbb{R}^d), \ (t, \omega) \in [0, T] \times \Omega. \tag{6.3}$$

Remark 6.3. Note that a multifunction f satisfying Assumption 6.2 must be a Carathéodory multifunction (see [22, §2.2]), which requires only continuity on the spatial variable. ∎

Remark 6.4. By Assumption 6.2, it is easy to check that $\{f(t, Y_t)\}_{t \in [0, T]} \in \mathscr{L}^2_{\mathbb{F}}([0, T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$ whenever $\{Y_t\}_{t \in [0, T]} \in \mathscr{L}^2_{\mathbb{F}}([0, T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$. ∎

We shall consider the following standard Picard iteration. Let $Y^{(0)} \equiv \{0\}$ and for $n \in \mathbb{N}$, we define $Y^{(n)}$ recursively by

$$Y_t^{(n)} = \mathbb{E}\Big[\xi + \int_t^T f(s, Y_s^{(n-1)})ds \ \Big| \ \mathcal{F}_t\Big], \qquad t \in [0, T]. \tag{6.4}$$

We should point out that the set-valued random variable $Y_t^{(n)}$ is defined almost surely for each fixed $t \in [0, T]$. An immediate question is whether $\{Y_t^{(n)}\}_{t \in [0, T]}$ makes sense, as a (jointly) measurable set-valued process, which, as usual, requires justification as we have seen frequently in the set-valued case. The following lemma is important for this purpose.

Lemma 6.5. *Let $X \in \mathscr{L}^2(\Omega, \mathscr{K}(\mathbb{R}^d))$ and define $F_t := \mathbb{E}[X|\mathcal{F}_t]$, $t \in [0, T]$. Then, $\{F_t\}_{t \in [0,T]}$ has an optional modification that is a uniformly $\mathbb{L}^2$-bounded martingale.*

Proof. Consider the (trivial) set-valued process $G_t \equiv X$, $t \in [0, T]$, which is clearly (jointly) measurable and $G_\tau = X$ is integrable for every $\mathbb{F}$-stopping time $\tau \colon \Omega \to [0, T]$. By [32, Theorem 3.7], there exists a unique *optional projection* $\{^oG_t\}_{t \in [0,T]}$ of process $\{G_t\}_{t \in [0,T]}$, such that $\mathbb{E}[G_\tau|\mathcal{F}_\tau] = {}^oG_\tau$, $\mathbb{P}$-a.s. for every $\mathbb{F}$-stopping time τ. In particular, $\{^oG_t\}_{t \in [0,T]}$ is an optional modification of $\{F_t\}_{t \in [0,T]}$.

It is easy to check that $\{^oG_t\}_{t \in [0,T]}$ is a square-integrable set-valued martingale, and by an $\mathbb{L}^1$-version of Lemma 4.1, it holds that

$$\|^oG_t\| = \|\mathbb{E}[X|\mathcal{F}_t]\| \leq \mathbb{E}[\|X\| \,|\mathcal{F}_t], \qquad t \in [0, T].$$

Finally, note that $X \in \mathscr{L}^2(\Omega, \mathscr{K}(\mathbb{R}^d))$; applying Doob's $\mathbb{L}^2$-maximal inequality to the (real-valued) martingale $M_t := \mathbb{E}[\|X\| \,|\mathcal{F}_t]$, $t \in [0, T]$, we obtain

$$\mathbb{E}\Big[\sup_{t \in [0,T]} \|^oG_t\|^2 \Big] \leq \mathbb{E}\Big[\sup_{t \in [0,T]} |M_t|^2 \Big] \leq 4\mathbb{E}[\|X\|^2] < +\infty.$$

That is, $\{^oG_t\}_{t \in [0,T]}$ is uniformly square-integrably bounded. ∎

The next proposition establishes the desired measurability for the Picard iteration.

Proposition 6.6. *For each $n \in \mathbb{N}$, $Y^{(n)}$ has a progressively measurable modification.*

Proof. Note that $Y^{(0)} \equiv \{0\}$ is progressively measurable itself. Let $n \in \mathbb{N}$ and suppose that $Y^{(n-1)}$ has a progressively measurable modification, which we denote by $Y^{(n-1)}$ for ease of notation, and interpret (6.4) accordingly.

For each $t \in [0, T]$, using Corollaries 3.1 and 3.5, we have

$$Y_t^{(n)} = \mathbb{E}\Big[\xi + \int_0^T f(s, Y_s^{(n-1)})ds \,\Big|\, \mathcal{F}_t\Big] \ominus \int_0^t f(s, Y_s^{(n-1)})ds. \tag{6.5}$$

By Remark 3.2(v), $\{\int_0^t f(s, Y_s^{(n-1)})ds\}_{t \in [0,T]}$ has a progressively measurable modification. Moreover, by Lemma 6.5, $\{\mathbb{E}[\xi + \int_0^T f(s, Y_s^{(n-1)})ds|\mathcal{F}_t]\}_{t \in [0,T]}$ has an optional, hence progressively measurable, modification. Replacing the original processes with such modifications in (6.5), and using Lemma 2.10, we have that $Y^{(n)}$ is progressively measurable. ∎

In view of Proposition 6.6, we will assume without loss of generality that $Y^{(n)}$ is progressively measurable, in particular, $Y^{(n)} \in \mathscr{L}_{\mathbb{F}}^2([0, T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$ for each $n \in \mathbb{N}$.

In order to guarantee the convergence of the sequence $\{Y^{(n)}\}_{n \in \mathbb{N}}$ constructed in (6.4), we will use a recursive estimate on $\{\mathbb{E}h^2(Y_t^{(n)}, Y_t^{(n-1)})\}_{n \in \mathbb{N}}$, which is provided by the following lemma. We note that unlike the vector-valued BSDEs, this lemma is non-trivial because of the lack of standard tools, in particular a set-valued Itô's formula.

Lemma 6.7. *For each $n \in \mathbb{N}$, it holds that*

$$\mathbb{E}[h^2(Y_t^{(n)}, Y_t^{(n-1)})] \leq TK^2 \int_t^T \mathbb{E}[h^2(Y_s^{(n-1)}, Y_s^{(n-2)})]ds, \quad t \in [0, T]. \tag{6.6}$$

Proof. By Proposition 2.3, Lemma 4.1 and the properties of Hausdorff distance, we get

$$\mathbb{E}[h^2(Y_t^{(n)}, Y_t^{(n-1)})] = \mathbb{E}\left[h^2\left(\mathbb{E}\left[\xi + \int_t^T f(s, Y_s^{(n-1)})ds \Big| \mathcal{F}_t\right], \mathbb{E}\left[\xi + \int_t^T f(s, Y_s^{(n-2)})ds \Big| \mathcal{F}_t\right]\right)\right]$$

$$\leq \mathbb{E}\left[\mathbb{E}\left[h^2\left(\xi + \int_t^T f(s, Y_s^{(n-1)})ds, \xi + \int_t^T f(s, Y_s^{(n-2)})ds\right) \Big| \mathcal{F}_t\right]\right]$$

$$\leq \mathbb{E}\left[h^2\left(\int_t^T f(s, Y_s^{(n-1)})ds, \int_t^T f(s, Y_s^{(n-2)})ds\right)\right]. \tag{6.7}$$

Then, combining (6.7) with Proposition 4.2, Assumption 6.2(iv), and Proposition 2.3, we derive (6.6). ∎

We are now ready to establish the well-posedness of the set-valued BSDE (6.1).

Theorem 6.8. *Assume Assumption 6.2. Then, the set-valued BSDE* (6.1) *has a solution* $Y \in \mathscr{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$. *Moreover, the solution is unique up to modifications: if* $Y' \in \mathscr{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$ *is another solution of* (6.1), *then* $Y_t = Y_t'$ *$\mathbb{P}$-a.s., $t \in [0,T]$.*

Proof. Recall that $(\mathscr{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d)), d_H)$ is a complete metric space, where the metric is defined by $d_H(\Phi, \Psi) = (\mathbb{E}[\int_0^T h^2(\Phi_t, \Psi_t)dt])^{\frac{1}{2}}$ for $\Phi, \Psi \in \mathscr{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$. We shall argue that the sequence $\{Y^{(n)}\}_{n \in \mathbb{N}}$ of the Picard iteration is Cauchy in $\mathscr{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$. To this end, for fixed $t \in [0,T]$, we note that $Y_t^{(0)} = \{0\}$. Thus, by repeatedly applying Lemma 4.1, we have

$$\mathbb{E}h^2(Y_t^{(1)}, Y_t^{(0)}) = \mathbb{E}h^2\left(\mathbb{E}\left[\xi + \int_t^T f(s, \{0\})ds \Big| \mathcal{F}_t\right], \{0\}\right) \tag{6.8}$$

$$\leq \mathbb{E}h^2\left(\xi + \int_t^T f(s, \{0\})ds, \{0\}\right) \leq 2\left[\mathbb{E}\|\xi\|^2 + \mathbb{E}h^2\left(\int_t^T f(s, \{0\})ds, \{0\}\right)\right]$$

$$\leq 2\left[\mathbb{E}\|\xi\|^2 + T\int_0^T \mathbb{E}\|f(s, \{0\})\|^2 ds\right] =: C.$$

Note that C is free of the choice of t. We claim that, for $n \in \mathbb{N}$, it holds that

$$\mathbb{E}h^2(Y_t^{(n)}, Y_t^{(n-1)}) \leq \frac{C(TK^2)^{n-1}(T-t)^{(n-1)}}{(n-1)!}. \tag{6.9}$$

Indeed, for $n = 1$, (6.9) is just (6.8). Now assume that (6.9) holds for $n-1$, then by Lemma 6.7 we have

$$\mathbb{E}h^2(Y_t^{(n)}, Y_t^{(n-1)}) \leq TK^2 \int_t^T \mathbb{E}[h^2(Y_s^{(n-1)}, Y_s^{(n-2)})]ds \leq TK^2 \int_t^T \frac{C(TK^2)^{n-2}(T-s)^{(n-2)}}{(n-2)!}ds$$

$$= \frac{C(TK^2)^{n-1}(T-t)^{(n-1)}}{(n-1)!} \leq \frac{CK^{2(n-1)}T^{2(n-1)}}{(n-1)!} =: a_n^2. \tag{6.10}$$

31

Since $\mathcal{H}_2$ is a metric on $\mathscr{L}^2_{\mathcal{F}_t}(\Omega, \mathscr{K}(\mathbb{R}^d))$ (see (4.1)), the estimate in (6.10) then yields for $m > n \geq 1$:

$$\mathcal{H}_2(Y_t^{(n)}, Y_t^{(m)}) \leq \sum_{k=n}^{m-1} \mathcal{H}_2(Y_t^{(k+1)}, Y_t^{(k)}) \leq \sum_{k=n}^{m-1} a_{k+1}, \tag{6.11}$$

where $a_k = \frac{\sqrt{C} K^{(k-1)} T^{(k-1)}}{\sqrt{(k-1)!}}$, $k \geq 1$, by (6.10). Hence,

$$d_H^2(Y^{(m)}, Y^{(n)}) = \int_0^T \mathcal{H}_2^2(Y_t^{(n)}, Y_t^{(m)}) dt \leq T \Big(\sum_{k=n}^{m-1} a_{k+1} \Big)^2. \tag{6.12}$$

Now note that

$$\frac{a_{k+1}}{a_k} = \frac{\frac{\sqrt{C} K^k T^k}{\sqrt{k!}}}{\frac{\sqrt{C} K^{(k-1)} T^{(k-1)}}{\sqrt{(k-1)!}}} = \frac{TK}{\sqrt{k}} \to 0, \qquad \text{as } k \to \infty.$$

By ratio test, $\sum_{k=1}^\infty a_k$ converges. That is, $\{Y^{(n)}\}_{n \in \mathbb{N}}$ is a Cauchy sequence in $\mathscr{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$, thanks to (6.11); whence converges to some $Y \in \mathscr{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$.

Next, we show that the limit process $Y = \{Y_t\}_{t \in [0,T]}$ indeed leads to a solution to the BSDE (6.1). Since $d_H(Y, Y^{(n)}) \to 0$ as $n \to \infty$, there exists a subsequence $\{Y^{(n_\ell)}\}_{\ell \in \mathbb{N}}$ such that $h(Y_t(\omega), Y_t^{(n_\ell)}(\omega)) \to 0$ as $\ell \to \infty$ for $dt \times d\mathbb{P}$-a.e. $(t, \omega) \in [0,T] \times \Omega$. By Lemma 4.1, Proposition 4.2, and Assumption 6.2(iii,iv), we have

$$\mathbb{E} h^2 \Big(\mathbb{E}\Big[\int_t^T f(s, Y_s) ds \mid \mathcal{F}_t \Big], \mathbb{E}\Big[\int_t^T f(s, Y_s^{(n_\ell)}) ds \mid \mathcal{F}_t \Big] \Big) \tag{6.13}$$

$$\leq \mathbb{E} h^2 \Big(\int_t^T f(s, Y_s) ds, \int_t^T f(s, Y_s^{(n_\ell)}) ds \Big)$$

$$\leq (T-t) \mathbb{E}\Big[\int_t^T h^2(f(s, Y_s), f(s, Y_s^{(n_\ell)})) ds \Big] \leq TK^2 \int_0^T \mathcal{H}_2^2(Y_s, Y_s^{(n_\ell)}) ds.$$

By the construction of the limit Y, we have $\int_0^T \mathcal{H}_2^2(Y_s, Y_s^{(n_\ell)}) ds \to 0$ as $\ell \to \infty$. Now, (6.13) shows that

$$\sup_{t \in [0,T]} \mathbb{E} h^2 \Big(\mathbb{E}\Big[\int_t^T f(s, Y_s^{(n_\ell)}) ds \mid \mathcal{F}_t \Big], \mathbb{E}\Big[\int_t^T f(s, Y_s) ds \mid \mathcal{F}_t \Big] \Big) \to 0 \quad \text{as } n \to \infty.$$

It follows that Y satisfies the BSDE

$$Y_t = \mathbb{E}\Big[\xi + \int_t^T f(s, Y_s) ds \mid \mathcal{F}_t \Big], \quad t \in [0,T].$$

In fact, a similar argument as above (using Assumption 6.2(iii)) also shows that Y is actually unique, as the solution of (6.1) in the space $\mathscr{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$ up to modifications. This proves the theorem. ∎

6.2 Set-Valued BSDEs with Martingale Terms

The set-valued BSDE (6.1) considered in §6.1 is formulated using set-valued conditional expectations. In this section, we aim to remove the conditional expectation by introducing an additional term to the BSDE, that is, a set-valued martingale. More specifically, we consider a set-valued BSDE of the form

$$Y_t + M_T = \xi + \int_t^T f(s, Y_s)ds + M_t, \quad t \in [0, T], \tag{6.14}$$

where $\{M_t\}_{t\in[0,T]}$ is a set-valued martingale and $\{Y_t\}_{t\in[0,T]}$, ξ, $f \colon [0,T] \times \Omega \to \mathscr{K}(\mathbb{R}^d)$ are as in §6.1.

Definition 6.9. *A pair $(Y, M) \in (\mathscr{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d)))^2$ of set-valued process is called an adapted solution to the set-valued BSDE (6.14) if M is a uniformly square-integrably bounded set-valued martingale with $M_0 = Y_0$ and, for each $t \in [0,T]$,*

$$Y_t + M_T = \xi + \int_t^T f(s, Y_s)ds + M_t \qquad \mathbb{P}\text{-a.s.}$$

Remark 6.10. *In light of Hukuhara difference, the set-valued BSDE (6.14) is equivalent to*

$$Y_t = \left(\xi + \int_t^T f(s, Y_s)ds + M_t \right) \ominus M_T, \quad t \in [0, T].$$

Furthermore, if $M_t \ominus M_T$ exists, the same BSDE is also equivalent to

$$Y_t = \xi + \int_t^T f(s, Y_s)ds + (M_t \ominus M_T), \quad t \in [0, T],$$

thanks to Proposition 2.2. However, the existence of $M_t \ominus M_T$ is a tall order due to the lack of temporal additivity of generalized stochastic integral (see Remark 6.15 for more details).

We now establish the well-posedness of the set-valued BSDE (6.14).

Theorem 6.11. *Suppose that Assumption 6.2 holds. Then, the set-valued BSDE (6.14) has a solution $(Y, M) \in (\mathscr{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d)))^2$. Moreover, the solution is unique up to modifications: if $(Y', M') \in (\mathscr{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d)))^2$ is another solution of (6.14), then $Y_t = Y'_t$ and $M_t = M'_t$ $\mathbb{P}$-a.s. for every $t \in [0, T]$.*

Proof. By Theorem 6.8, the set-valued BSDE (6.1) in conditional expectation form has a solution $Y \in \mathscr{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$. Define a process $M = \{M_t\}_{t\in[0,T]}$ by

$$M_t := \mathbb{E}\Big[\xi + \int_0^T f(s, Y_s)ds \mid \mathcal{F}_t \Big], \quad t \in [0, T].$$

By Remark 6.4, $\{f(t, Y_t)\}_{t \in [0,T]} \in \mathscr{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$; by Lemma 6.5, M is a uniformly square-integrably bounded set-valued martingale. On the other hand, by Corollary 3.5,

$$\int_0^T f(s, Y_s)ds = \int_0^t f(s, Y_s)ds + \int_t^T f(s, Y_s)ds, \quad t \in [0,T]. \tag{6.15}$$

By the linearity of set-valued conditional expectation and (6.15), we have

$$M_t = \mathbb{E}\left[\xi + \int_0^T f(s, Y_s)ds \;\Big|\; \mathcal{F}_t\right] = \mathbb{E}\left[\xi + \int_0^t f(s, Y_s)ds + \int_t^T f(s, Y_s)ds \;\Big|\; \mathcal{F}_t\right] \tag{6.16}$$

$$= \mathbb{E}\left[\xi + \int_t^T f(s, Y_s)ds \;\Big|\; \mathcal{F}_t\right] + \int_0^t f(s, Y_s)ds = Y_t + \int_0^t f(s, Y_s)ds.$$

Using the definitions of M_T, M_t, and combining (6.15) and (6.16) give

$$Y_t + M_T = Y_t + \xi + \int_0^t f(s, Y_s)ds + \int_t^T f(s, Y_s)ds = \xi + \int_t^T f(s, Y_s)ds + M_t.$$

Finally, $M_0 = \mathbb{E}[\xi + \int_0^T f(s, Y_s)ds] = Y_0$ by the definitions of M_0 and Y_0. Hence, the pair (Y, M) is a solution to the set-valued BSDE (6.14).

To prove uniqueness, let (Y', M') be another solution to (6.14). Let $t \in [0, T]$. Hence,

$$Y'_t + M'_T = \xi + \int_t^T f(s, Y'_s)ds + M'_t \qquad \mathbb{P}\text{-a.s.} \tag{6.17}$$

Since M' is a martingale, taking conditional expectation in (6.17) under $\mathcal{F}_t$ gives

$$Y'_t + M'_t = \mathbb{E}\left[\xi + \int_t^T f(s, Y'_s)ds)ds \;\Big|\; \mathcal{F}_t\right] + M'_t.$$

Hence, by cancellation law, Y' is a solution to the BSDE (6.1). By the uniqueness part of Theorem 6.8, $Y_t = Y'_t$ $\mathbb{P}$-a.s. Using this, we may rewrite (6.17) as

$$Y_t + M'_T = \xi + \int_t^T f(s, Y_s)ds + M'_t \qquad \mathbb{P}\text{-a.s.}$$

In particular, when $t = 0$, we have

$$Y_0 + M'_T = \xi + \int_0^T f(s, Y_s)ds + M'_0. \tag{6.18}$$

On the other hand, we have $M_0 = Y_0 = Y'_0 = M'_0$. Hence, cancellation law in (6.18) and the definition of M_T yield

$$M'_T = \xi + \int_0^T f(s, Y_s)ds = M_T.$$

Finally, since M, M' are both martingales with the same terminal value, we obtain $M_t = M'_t$ $\mathbb{P}$-a.s. for each $t \in [0, T]$. Hence, (Y, M) and (Y', M') coincide up to modifications. ∎

We conclude this section by proving that the forms of the set-valued BSDE given by (6.1) and (6.14) are indeed equivalent.

Corollary 6.12. *(i) If $Y \in \mathscr{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$ is a solution of (6.1), then there exists a unique $M \in \mathscr{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$ such that (Y, M) is a solution of (6.14).*

(ii) If $(Y, M) \in (\mathscr{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d)))^2$ is a solution of (6.14), then Y solves (6.1).

Proof. (i) Let Y be a solution of (6.1). Following the construction in the proof of Theorem 6.11, one can find a set-valued martingale M such that (Y, M) solves (6.14). The uniqueness of such M is also a consequence of the uniqueness part of Theorem 6.11.

(ii) Let (Y, M) be a solution of (6.14). For each $t \in [0, T]$, taking conditional expectation with respect to $\mathcal{F}_t$ in (6.14) gives

$$Y_t + M_t = \mathbb{E}\Big[\xi + \int_t^T f(s, Y_s)ds \ \Big| \ \mathcal{F}_t\Big] + M_t$$

since M is a set-valued martingale. By cancellation law, it follows that Y solves (6.1). ∎

6.3   Set-Valued BSDEs with Generalized Stochastic Integrals

We can now combine the martingale representation theorem developed in §5.2 with the set-valued BSDE (6.14) to study the equation of the form

$$Y_t + \int_{0-}^T \mathcal{R}dB = \xi + \int_t^T f(s, Y_s)ds + \int_{0-}^t \mathcal{R}dB, \quad t \in [0, T], \tag{6.19}$$

where $\mathcal{R} \subset \mathbb{R}_0 = \mathbb{R}^d \times \mathbb{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathbb{R}^{d \times m})$ is a set of martingale representer pairs.

Definition 6.13. *A pair $(Y, \mathcal{R})$ with $Y \in \mathscr{L}^2_{\mathbb{F}}([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d))$ and $\mathcal{R} \in \mathbb{R}_0$ is called a solution of the set-valued BSDE (6.19) if $Y_0 = \pi_\xi[\mathcal{R}]$, $\{\int_{0-}^t \mathcal{R}dB\}_{t \in [0,T]}$ is a uniformly square-integrably bounded set-valued martingale and, for each $t \in [0, T]$,*

$$Y_t + \int_{0-}^T \mathcal{R}dB = \xi + \int_t^T f(s, Y_s)ds + \int_{0-}^t \mathcal{R}dB \qquad \mathbb{P}\text{-a.s.}$$

The following theorem provides a well-posedness result for the set-valued BSDE (6.19).

Theorem 6.14. *Under Assumption 6.2, the set-valued BSDE (6.19) has a solution $(Y, \mathcal{R})$. Moreover, the solution is unique in the following sense: if $(Y', \mathcal{R}')$ is another solution of (6.19), then $Y_t = Y'_t$ and $\int_{0-}^t \mathcal{R}dB = \int_{0-}^t \mathcal{R}'dB$ $\mathbb{P}$-a.s. for every $t \in [0, T]$.*

Proof. By Theorem 6.11, there exists a solution (Y, M) of the set-valued BSDE (6.14). By Definition 6.9, M is a uniformly square-integrably bounded set-valued martingale. Hence, by Theorem 5.6, we may write $M_t = \int_{0-}^t \mathcal{R}^M dB$ $\mathbb{P}$-a.s. for each $t \in [0, T]$. Hence, $(Y, \mathcal{R}^M)$ is a solution of (6.19). The uniqueness claim is an immediate consequence of the uniqueness part of Theorem 6.11. ∎

Remark 6.15. The indefinite integral $M = \{\int_{0-}^{t} \mathcal{R}^M dB\}_{t \in [0,T]}$ in the proof of Theorem 6.14 is a uniformly square-integrably bounded set-valued martingale. Following similar arguments as in [22, Corollary 5.3.2], it can be shown that if $\mathcal{Z}^M$ is decomposable, then, $\mathcal{Z}^M$ as well as $\mathcal{R}^M$ are singletons. Hence, in all cases where $\mathcal{Z}^M$ contains more than one processes, the set $\mathcal{Z}^M$ is not decomposable. Similar to Corollary 3.7 for the generalized Aumann-Itô stochastic integral, the indefinite integral considered here does not have time-additivity in general, that is, the Hukuhara difference $M_T \ominus M_t = \int_{0-}^{T} \mathcal{R}^M dB \ominus \int_{0-}^{t} \mathcal{R}^M dB$ does not exist. In particular, in view of Remark 5.7, the inclusion $\int_{0-}^{t} \mathcal{R}^M dB \subset M_0 + \int_{0}^{t} \mathcal{Z}^M dB$ is generally a strict one. ∎

We conclude the paper by showing that the equivalence between the set-valued BSDE (6.19) and the one in (6.14).

Corollary 6.16. *(i) If $(Y, \mathcal{R})$ is a solution of* (6.19), *then* $(Y, \{\int_{0-}^{t} \mathcal{R} dB\}_{t \in [0,T]})$ *is a solution of* (6.14).
(ii) If $(Y, M) \in (\mathscr{L}_{\mathbb{F}}^2([0,T] \times \Omega, \mathscr{K}(\mathbb{R}^d)))^2$ is a solution of (6.14), *then there exists $\mathcal{R} \in \mathbb{R}_0$ such that $(Y, \mathcal{R})$ solves* (6.19), *and such $\mathcal{R}$ is unique in the following sense: if $(Y, \mathcal{R}')$ solves* (6.19) *for some $\mathcal{R}' \in \mathbb{R}_0$, then $\int_{0-}^{t} \mathcal{R} dB = \int_{0-}^{t} \mathcal{R}' dB$ $\mathbb{P}$-a.s. for every $t \in [0,T]$.*

Proof. (i) is an immediate consequence of Definitions 6.9 and 6.13. The existence part of (ii) follows by the set-valued martingale representation theorem (Theorem 5.6) as in the proof of Theorem 6.14. The uniqueness part follows from the uniqueness claim in Theorem 6.11. ∎

Corollaries 6.12 and 6.16 together yield that the three forms of the set-valued BSDE, namely, (6.1), (6.14), and (6.13), are essentially equivalent.

Acknowledgments

References

[1] Ararat, Ç., Rudloff, B., (2020), *Dual representations for systemic risk measures*, Math. Financ. Econ., **14**(1), 139–174.

[2] Aumann, R. J., (1965), *Integrals of set-valued functions*, J. Math. Anal. Appl., **12**(1), 1–12.

[3] Artzner, P., Delbaen, F., Eber, J., Heath, D., (1999), *Coherent measures of risk*, Math. Financ., **9**(3), 203–228.

[4] Biagini, F., Fouque, J.-P., Frittelli, M., Meyer-Brandis, T., (2019), *A unified approach to systemic risk measures via acceptance sets*, Math. Financ., **29**(1), 329–367.

[5] Bion-Nadal, J., (2009), *Time consistent dynamic risk processes*, Stoch. Proc. Appl., **119**(2), 633–654.

[6] Castaing, C., Valadier, M., (1977), *Convex Analysis and Measurable Multifunctions*, Springer-Verlag, Berlin Heidelberg.

[7] Coquet, F., Hu, Y., Mémin, J., Peng, S., (2002), *Filtration-consistent nonlinear expectations and related g-expectations*, Probab. Theory Rel., **123**(1), 1–27.

[8] Feinstein, Z., Rudloff, B., (2015), *Multi-portfolio time consistency for set-valued convex and coherent risk measures*, Financ. Stoch., **19**(1), 67–107.

[9] Feinstein, Z., Rudloff, B., Weber, S., (2017), *Measures of systemic risk*, SIAM J. Financ. Math., **8**(1), 672--708.

[10] Hamel, A. H., Heyde, F., Rudloff, B., (2011), *Set-valued risk measures for conical market models*, Math. Financ. Econ., **5**(1), 1–28.

[11] Hamel, A. H., Heyde, F., Löhne A., Rudloff B., Schrage C., (2015), *Set optimization - a rather short introduction*, in: A. H. Hamel, F. Heyde, A. Löhne, B. Rudloff, C. Schrage (eds.), *Set Optimization and Applications - the state of the art. From set relations to set-valued risk measures*, 65–141, Springer-Verlag, Berlin.

[12] Hamel, A. H., Schrage C., (2014), *Directional derivatives, subdifferentials and optimality conditions for set-valued convex functions*, Pac. J. Optim., **10**(4), 667–689.

[13] Hiai. F., Umegaki. H., (1977), *Conditional expectations, and martingales of multivalued functions*, J. Multivariate Anal., **7**(1), 149–182.

[14] Hukuhara, M., (1967), *Integration des applications measurables dont la valeur est un compact convexe*, Funkcialaj Ekvacioj, **10**, 205-223.

[15] Lakshmikantham, V., Gnana Bhaskar, T., Vasundhara Devi, J., (2006), *Theory of Set Differential Equations in Metric Spaces*, Cambridge Scientific Publishers, Cambridge.

[16] Kisielewicz, M., (1997), *Set-valued stochastic integrals and stochastic inclusions*, Stoch. Anal. Appl., **15**(5), 783–800.

[17] Kisielewicz, M., (2007), *Backward stochastic differential inclusions*, Dyn. Syst. Appl., **16**, 121–140.

[18] Kisielewicz, M., (2008), *Weak compactness of weak solutions to backward stochastic differential inclusions*, Dyn. Syst. Appl., **17**, 351–370.

[19] Kisielewicz, M., (2012), *Some properties of set-valued stochastic integrals*, J. Math. Anal. Appl., **388**(2), 984–995.

[20] Kisielewicz, M., (2013), *Stochastic Differential Inclusions and Applications*, Springer, New York.

[21] Kisielewicz, M., (2014), *Martingale representation theorem for set-valued martingales*, J. Math. Anal. Appl., **409**(1), 111–118.

[22] Kisielewicz, M., (2020), *Set-valued Stochastic Integrals and Applications*, Springer, Switzerland.

[23] Kisielewicz, M., Michta, M., (2016), *Properties of set-valued stochastic differential equations*, Optimization, **65**(12), 2153–2169.

[24] Kisielewicz, M., Michta, M., (2017), *Integrably bounded set-valued stochastic integrals*, J. Math. Anal. Appl., **449**(2), 1892–1910.

[25] Ma, J., Yao, S., (2010), *On quadratic g-evaluations/expectations and related analysis*, Stoch. Anal. Appl., **28**(4), 711–734.

[26] Malinowski, M., Michta, M., (2013), *The interrelation between stochastic differential inclusions and set-valued stochastic differential equations*, J. Math. Anal. Appl., **408**(2), 733–743.

[27] Michta, M., (2015), *Remarks on unboundedness of set-valued Itô stochastic integrals*, J. Math. Anal. Appl., **424**(1), 651–663.

[28] Molchanov, I., (2018), *Theory of Random Sets*, second edition, Springer-Verlag, London.

[29] Pallaschke, D., Urbański, R., (2010), *On the separation and order law of cancellation for bounded sets*, Optimization, **51**(3), 487–496.

[30] Buckdahn, R., Engelber, H.-J., Răşcanu, A., (2005), *On weak solutions of backward stochastic differential equations*, Theory Probab. its Appl., **49**(1), 16–50.

[31] Rockafellar, R. T., Wets, R. J.-B., (1998), *Variational Analysis*, Springer-Verlag, Berlin Heidelberg.

[32] Wang, R., (2001), *Optional and predictable projections of set-valued measurable processes*, Appl. Math. J. Chinese Univ. Ser. B, **16**(3), 323–329.

[33] Rosazza Gianin, E., (2006), *Risk measures via g-expectations*, Insur. Math. Econ., **39**(1), 19–34.

[34] Urbański, R., (1976), *A generalization of the Minkowski-Radström-Hörmander theorem*, Bulletin of the Polish Academy of Sciences, Mathematics, Astronomy, and Physics, **24**, 709–715.

[35] Zhang, J., Yano, K., (2020), *Remarks on martingale representation theorem for set-valued martingales*, arXiv preprint 2012.06988.