Chapter 4
Stochastic Differential Inclusions

This chapter is devoted to the theory of stochastic differential inclusions. The main results deal with stochastic functional inclusions defined by set-valued functional stochastic integrals. Subsequent sections discuss properties of stochastic and backward stochastic differential inclusions.

1 Stochastic Functional Inclusions

Throughout this section, by $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ we shall denote a complete filtered probability space and assume that $F : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ satisfy the following conditions $(\mathcal{H})$:

(i) F and G are measurable,
(ii) F and G are uniformly square integrably bounded.

For set-valued mappings F and G as given above, by stochastic functional inclusions $SFI(F,G)$, $SFI(\overline{F},G)$, and $\overline{SFI}(F,G)$ we mean relations of the form

$$x_t - x_s \in J_{st}[S_{\mathbb{F}}(F \circ x)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G \circ x)],$$

$$x_t - x_s \in \mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F \circ x)]\} + \mathcal{J}_{st}[S_{\mathbb{F}}(G \circ x)]\},$$

and

$$x_t - x_s \in \mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F \circ x)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G \circ x)]\},$$

respectively, which have to be satisfied for every $0 \le s \le t \le T$ by a system $(\mathcal{P}_{\mathbb{F}}, X, B)$ consisting of a complete filtered probability space $\mathcal{P}_{\mathbb{F}}$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ satisfying the usual conditions, an d-dimensional $\mathbb{F}$-adapted continuous stochastic process $X = (X_t)_{0\le t\le T}$, and an m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{0\le t\le T}$ defined on $\mathcal{P}_{\mathbb{F}}$. Such systems $(\mathcal{P}_{\mathbb{F}}, X, B)$ are said to be weak solutions of $SFI(F,G)$, $SFI(\overline{F},G)$, and $\overline{SFI}(F,G)$, respectively. If

M. Kisielewicz, *Stochastic Differential Inclusions and Applications*,
Springer Optimization and Its Applications 80, DOI 10.1007/978-1-4614-6756-4_4,

μ is a given probability measure on $\beta(\mathbb{R}^d)$, then a system $(\mathcal{P}_{\mathbb{F}}, X, B)$ is said to be a weak solution of the initial value problems $SFI(F,G,\mu)$, $SFI(\overline{F},G,\mu)$, or $\overline{SFI}(F,G,\mu)$, respectively, if it satisfies condition (1)–(1), respectively, and $PX_0^{-1} = \mu$. The set of all weak solutions of $SFI(F,G,\mu)$, $SFI(\overline{F},G,\mu)$, and $\overline{SFI}(F,G,\mu)$ (equivalence classes $[(\mathcal{P}_{\mathbb{F}}, X, B)]$ consisting of all systems $(\mathcal{P}'_{\mathbb{F}}, X', B,)$ satisfying (1)–(1), respectively and such that $PX_0^{-1} = P(X'_0)^{-1} = \mu$ and $PX^{-1} = P(X')^{-1}$) are denoted by $\mathcal{X}_\mu(F,G)$, $\mathcal{X}_\mu(\overline{F},G)$, and $\overline{\mathcal{X}}_\mu(F,G)$,, respectively. By $\mathcal{X}^0_\mu(F,G)$ we denote the set of all $[(\mathcal{P}_{\mathbb{F}}, X, B)] \in \mathcal{X}_\mu(F,G)$ with a separable filtered probability space $\mathcal{P}_{\mathbb{F}}$.

Remark 1.1. We can also consider initial value problems for $SFI(F,G)$, $SFI(\overline{F},G)$, and $\overline{SFI}(F,G)$ with an initial condition $x_s = x$ a.s. for a fixed $0 \le s \le T$ and $x \in \mathbb{R}^d$. The sets of all weak solutions for such initial value problems are denoted by $\mathcal{X}_{s,x}(F,G)$, $\mathcal{X}_{s,x}(\overline{F},G)$, and $\overline{\mathcal{X}}_{s,x}(F,G)$, respectively. □

Remark 1.2. The following inclusions follow immediately from Lemma 1.6 of Chap. 3: $\mathcal{X}_\mu(F,G) \subset \mathcal{X}_\mu(\overline{F},G) \subset \overline{\mathcal{X}}_\mu(F,G) \subset \mathcal{X}_\mu(\text{co}\,F, \text{co}\,G)$ for all measurable set-valued functions $F : [0,T] \times \mathbb{R}^d \to \text{Cl}(\mathbb{R}^d)$ and $G : [0,T] \times \mathbb{R}^d \to \text{Cl}(\mathbb{R}^{d\times m})$ and probability measure μ on $\beta(\mathbb{R}^d)$. □

Remark 1.3. In what follows, we shall identify weak solutions (equivalence classes $[(\mathcal{P}_{\mathbb{F}}, X, B)]$) of $SFI(F,G)$, $SFI(\overline{F},G)$, and $\overline{SFI}(F,G)$, respectively, with pairs (X,B) of stochastic processes X and B defined on $\mathcal{P}_{\mathbb{F}}$ or with a process X. □

If apart from the set-valued mappings F and G, we are also given a filtered probability space $\mathcal{P}_{\mathbb{F}}$ and an m-dimensional $\mathbb{F}$-Brownian motion on $\mathcal{P}_{\mathbb{F}}$, then a continuous $\mathbb{F}$-adapted process X such that $(\mathcal{P}_{\mathbb{F}}, X, B)$ satisfies (1)–(1), respectively, is called a strong solution for $SFI(F,G)$, $SFI(\overline{F},G)$, and $\overline{SFI}(F,G)$, respectively. For a given $\mathcal{F}_0$-measurable random variable $\xi : \Omega \to \mathbb{R}^d$, the sets of all strong solutions of the above stochastic functional inclusions corresponding to a filtered probability space $\mathcal{P}_{\mathbb{F}}$ and an m-dimensional $\mathbb{F}$-Brownian motion B satisfying an initial condition $X_0 = \xi$ a.s. will be denoted by $\mathcal{S}_\xi(F,G,B,\mathcal{P}_{\mathbb{F}})$, $\mathcal{S}_\xi(\overline{F},G,B,\mathcal{P}_{\mathbb{F}})$, and $\overline{\mathcal{S}}_\xi(F,G,B,\mathcal{P}_{\mathbb{F}})$, respectively. Immediately from Lemma 1.6 of Chap. 3, it follows that $\mathcal{S}_\xi(F,G,B,\mathcal{P}_{\mathbb{F}}) \subset \mathcal{S}_\xi(\overline{F},G,B,\mathcal{P}_{\mathbb{F}}) \subset \overline{\mathcal{S}}_\xi(F,G,B,\mathcal{P}_{\mathbb{F}}) \subset \mathcal{S}_\xi(\overline{\text{co}}\,F, \overline{\text{co}}\,G, B, \mathcal{P}_{\mathbb{F}}) \subset \mathcal{S}(\mathbb{F},\mathbb{R}^d)$, where $\mathcal{S}(\mathbb{F},\mathbb{R}^d)$ denotes the Banach space of all d-dimensional $\mathbb{F}$-semimartingales $(X_t)_{0\le t\le T}$ on $\mathcal{P}_{\mathbb{F}}$ such that $E[\sup_{0\le t\le T} |X_t|^2] < \infty$. If $\mathcal{P}_{\mathbb{F}}$ is separable, then by virtue of Lemma 1.7 of Chap. 3, one has $\mathcal{S}_\xi(\overline{F},G,B,\mathcal{P}_{\mathbb{F}}) = \mathcal{S}_\xi(\text{co}\,F,G,B,\mathcal{P}_{\mathbb{F}})$.

In what follows, norms of $\mathbb{R}^r$, $\mathbb{L}^2(\Omega,\mathcal{F},\mathbb{R}^r)$, and $\mathbb{L}^2([0,T]\times\Omega,\Sigma_{\mathbb{F}},\mathbb{R}^r)$ with $r = d$ and $r = d\times m$ will be denoted by $|\cdot|$. It will be clear from the context which of the above normed space is considered.

Theorem 1.1. *Let $B = (B_t)_{0\le t\le T}$ be an m-dimensional $\mathbb{F}$-Brownian motion on $\mathcal{P}_{\mathbb{F}}$, and $\xi : \Omega \to \mathbb{R}^d$ an $\mathcal{F}_0$-measurable random variable. If F and G satisfy conditions $(\mathcal{H})$ and are such that $F(t,\cdot)$ and $G(t,\cdot)$ are Lipschitz continuous with* *a Lipschitz function* $k \in \mathbb{L}^2([0,T],\mathbb{R}^+)$ *such that* $K(\sqrt{T}+1) < 1$*, where* $K = (\int_0^T k^2(t)\mathrm{d}t)^{1/2}$*, then* $\mathcal{S}_\xi(F,G,B,\mathcal{P}_\mathbb{F}) \neq \emptyset$*.*

Proof. Let $\mathcal{X} = \mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^d)\times\mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^{d\times m})$ and put $X_t^{fg} = \xi + \int_0^t f_\tau \mathrm{d}\tau + \int_0^t g_\tau \mathrm{d}B_\tau$ a.s. for $0 \le t \le T$ and $(f,g) \in \mathcal{X}$. It is clear that $X^{fg} = (X_t^{fg})_{0\le t\le T} \in \mathcal{S}(\mathbb{F},\mathbb{R}^d)$. Define on $\mathcal{X}$ a set-valued mapping Q by setting $Q(f,g) = S_\mathbb{F}(F\circ X^{fg})\times S_\mathbb{F}(G\circ X^{fg})$ for every $(f,g)\in\mathcal{X}$. It is clear that for every $(f,g)\in\mathcal{X}$, we have $Q(f,g)\in \mathrm{Cl}(\mathcal{X})$.

Let $\lambda(A\times C, B\times D) = \max\{H(A,B),H(C,D)\}$, for $A,B\in \mathrm{Cl}(\mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^d)$ and $C,D\in \mathrm{Cl}(\mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^{d\times m})$, where for simplicity, H denotes the Hausdorff metric on $\mathrm{Cl}(\mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^d)$ and $\mathrm{Cl}(\mathbb{L}^2([0,T]\times\Omega,\Sigma_\mathbb{F},\mathbb{R}^{d\times m})$. It is clear that λ is a metric on $\mathrm{Cl}(\mathcal{X})$. By virtue of Lemma 3.7 of Chap. 2, we have $H(S_\mathbb{F}(F\circ X^{fg}),S_\mathbb{F}(F\circ X^{f'g'})) \le K\|X^{fg}-X^{f'g'}\|_c$ and $H(S_\mathbb{F}(G\circ X^{fg}),S_\mathbb{F}(G\circ X^{f'g'})) \le K\|X^{fg}-X^{f'g'}\|_c$ for every $(f,g),(f',g')\in\mathcal{X}$, where $\|\cdot\|_c$ denotes the norm of $\mathcal{S}(\mathbb{F},\mathbb{R}^d)$ defined by $\|x\|_c^2 = E[\sup_{0\le t\le T}|x_t|^2]$ for $x=(x_t)_{0\le t\le T}\in\mathcal{S}(\mathbb{F},\mathbb{R}^d)$. But

$$
\begin{aligned}
\|X^{fg}-X^{f'g'}\|_c &= \left(E\sup_{0\le t\le T}\left|\int_0^t (f_\tau - f'_\tau)\mathrm{d}\tau + \int_0^t (g_\tau - g'_\tau)\mathrm{d}B_\tau\right|^2\right)^{1/2}\\
&\le \left(E\sup_{0\le t\le T}\left|\int_0^t (f_\tau - f'_\tau)\mathrm{d}\tau\right|^2\right)^{1/2}\\
&\quad + \left(E\sup_{0\le t\le T}\left|\int_0^t (g_\tau - g'_\tau)\mathrm{d}B_\tau\right|^2\right)^{1/2}\\
&\le \sqrt{T}\left(E\sup_{0\le t\le T}\int_0^t |f_\tau - f'_\tau|^2\mathrm{d}\tau\right)^{1/2}\\
&\quad + \left(E\sup_{0\le t\le T}\int_0^t |g_\tau - g'_\tau|^2\mathrm{d}\tau^2\right)^{1/2}\\
&= \sqrt{T}\,|f-f'| + |g-g'| \le (\sqrt{T}+1)\,\|(f,g)-(f',g')\|,
\end{aligned}
$$

where $\|\cdot\|$ denotes the norm on $\mathcal{X}$. Therefore,

$$
\lambda(Q(f,g),Q(f',g')) \le K(\sqrt{T}+1)\,\|(f,g)-(f',g')\|
$$

for every $(f,g),(f',g')\in\mathcal{X}$, which by th Covitz–Nadler fixed-point theorem, implies the existence of $(f,g)\in\mathcal{X}$ such that $(f,g)\in Q(f,g)$. Hence it follows that $\int_s^t f_\tau \mathrm{d}\tau + \int_s^t g_\tau \mathrm{d}B_\tau \in J_{st}[S_{\mathbb{F}}(F \circ X^{fg})] + \mathcal{J}_{st}[S_{\mathbb{F}}(G \circ X^{fg})]$ for every $0 \le s \le t \le T$. This, by the definition of X^{fg}, implies that $X^{fg} \in \mathcal{S}_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$. □

Remark 1.4. By an appropriate changing the norm (see Remark 1.1 of Chap. 7) of the space $\mathcal{X}$, the result of Theorem 1.1 can be obtained for every $T > 0$ and $k \in \mathbb{L}^2([0,T], \mathbb{R}^+)$ without the constraint $K(\sqrt{T} + 1) < 1$. □

Let us denote by $\Lambda_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$ the set of all fixed points of the set-valued mapping Q defined in the proof of Theorem 1.1.

Theorem 1.2. *If the assumptions of Theorem 1.1 are satisfied, then*

(i) $\Lambda_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$ *is a closed subset of* $\mathcal{X}$;
(ii) $\mathcal{S}_\xi(\mathrm{co}\, F, \mathrm{co}\, G, B, \mathcal{P}_{\mathbb{F}}) \neq \emptyset$ *if and only if* $\Lambda_\xi(\mathrm{co}\, F, \mathrm{co}\, G, B, \mathcal{P}_{\mathbb{F}}) \neq \emptyset$;
(iii) $\mathcal{S}_\xi(\mathrm{co}\, F, \mathrm{co}\, G, B, \mathcal{P}_{\mathbb{F}})$ *is a closed subset of* $\mathcal{S}(\mathbb{F}, \mathbb{R}^d)$;
(iv) *for every* $x \in \overline{\mathcal{S}}_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$ *and* $\varepsilon > 0$, *there exists* $x^\varepsilon \in \mathcal{S}(\mathbb{F}, \mathbb{R}^d)$ *such that* $\sup_{0 \le t \le t}(E|x - x^\varepsilon|^2)^{1/2} \le \varepsilon$ *and* $\mathrm{dist}(x_t - x_s, J_{st}[S_{\mathbb{F}}(F \circ x)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G \circ x)]) \le \varepsilon$;
(v) $\mathcal{X}_\mu(F, G) \neq \emptyset$ *for every probability measure* μ *on* $\beta(\mathbb{R}^d)$.

Proof. (i) The closedness of $\Lambda_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$ follows immediately from the properties of the set-valued mappings $\mathcal{X} \ni (f, g) \to S_{\mathbb{F}}(F \circ X^{fg}) \subset \mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^d)$ and $\mathcal{X} \ni (f, g) \to S_{\mathbb{F}}(G \circ X^{fg}) \subset \mathbb{L}^2([0,T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^{d \times m})$. Indeed, if $\{(f^n, g^n)\}_{n=1}^\infty$ is a sequence of $\Lambda_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$ converging to (f, g), then $\mathrm{dist}(f, S_{\mathbb{F}}(F \circ X^{fg})) = 0$, because

$$\begin{aligned}\mathrm{dist}(f, S_{\mathbb{F}}(F \circ X^{fg})) &\le |f - f^n| + \mathrm{dist}(f^n, S_{\mathbb{F}}(F \circ X^{f^n g^n})) \\ &\quad + H(S_{\mathbb{F}}(F \circ X^{fg}), S_{\mathbb{F}}(F \circ X^{f^n g^n})),\end{aligned}$$

and by virtue of Lemma 3.7 of Chap. 2, for every $n \ge 1$ one has

$$H(S_{\mathbb{F}}(F \circ X^{fg}), S_{\mathbb{F}}(F \circ X^{f^n g^n}) \le K(\sqrt{T} + 1)\|(f, g) - (f^n, g^n)\|.$$

In a similar way, we also get $\mathrm{dist}(g, S_{\mathbb{F}}(G \circ x^{fg})) = 0$. Hence, by the closedness of $S_{\mathbb{F}}(F \circ x^{fg})$ and $S_{\mathbb{F}}(G \circ x^{fg})$, it follows that $(f, g) \in Q(f, g)$. Then $(f, g) \in \Lambda_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$.

(ii) The implication $\Lambda_\xi(\mathrm{co}\, F, \mathrm{co}\, G, B, \mathcal{P}_{\mathbb{F}}) \neq \emptyset \Rightarrow \mathcal{S}_\xi(\mathrm{co}\, F, \mathrm{co}\, G, B, \mathcal{P}_{\mathbb{F}}) \neq \emptyset$ follows immediately from the proof of Theorem 1.1. The converse implication follows immediately from Theorem 1.5 of Chap. 3.

(iii) Let $(u^n)_{n=1}^\infty$ be a sequence of $\mathcal{S}_\xi(\mathrm{co}\, F, \mathrm{co}\, G, B, \mathcal{P}_{\mathbb{F}})$ converging to $u \in \mathcal{S}(\mathbb{F}, \mathbb{R}^d)$. By Theorem 1.5 of Chap. 3, there exists a sequence $\{(f^n, g^n)\}_{n=1}^\infty$ of $S_{\mathbb{F}}(\mathrm{co}\, F \circ u^n) \times S_{\mathbb{F}}(\mathrm{co}\, G \circ u^n)$ such that $u_t^n = \xi + J_{0t}(f^n) + \mathcal{J}_{0t}(g^n)$ for $n \ge 1$ and $t \in [0,T]$. By Remark 3.1 of Chap. 2, there is a subsequence $\{(f^{n_k}, g^{n_k})\}_{k=1}^\infty$ of $\{(f^n, g^n)\}_{n=1}^\infty$ weakly converging to (f, g), which implies that $J_{0t}(f^{n_k}) + \mathcal{J}_{0t}(g^{n_k}) \to J_{0t}(f) + \mathcal{J}_{0t}(g)$ for every $t \in [0,T]$ in the weak topology of $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$ as $k \to \infty$. But for every $t \in [0, T]$, a sequence $(u_t^{n_k})_{k=1}^{\infty}$ also converges weakly in $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$ to u_t. Therefore, $u_t = \xi + J_{0t}(f) + \mathcal{J}_{0t}(g)$ for every $t \in [0, T]$. Then $u \in \mathcal{S}_\xi(\mathrm{co}\, F, \mathrm{co}\, G, B, \mathcal{P}_{\mathbb{F}})$.

(iv) For every $x \in \overline{\mathcal{S}}_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$ and $\varepsilon > 0$, there exists $x^\varepsilon \in \mathcal{S}_\xi(F, G, B, \mathcal{P}_{\mathbb{F}})$ such that $\sup_{0 \le t \le t}(E|x - x^\varepsilon|^2)^{1/2} \le \varepsilon/[2 + L(\sqrt{T} + 1)]$, where $L = (\int_0^T k^2(t)\mathrm{d}t)^{1/2}$. Similarly as in the proof of Lemma 3.7 of Chap. 2 (see Lemma 1.3 of Chap. 5), it follows that set-valued mappings $\mathcal{S}(\mathbb{F}, \mathbb{R}^d) \ni x \to J_{st}[S_{\mathbb{F}}(F \circ x)] \subset \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$ and $\mathcal{S}(\mathbb{F}, \mathbb{R}^d) \ni x \to J_{st}[S_{\mathbb{F}}(G \circ x)] \subset \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$ are Lipschitz continuous with Lipschitz constants $\sqrt{T}L$ and L, respectively. Therefore,

$$\begin{aligned}
&\mathrm{dist}(x_t - x_s, J_{st}[S_{\mathbb{F}}(F \circ x)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G \circ x)]) \\
&\quad \le |(x_t - x_s) - (x_t^\varepsilon - x_s^\varepsilon)| \\
&\qquad + \mathrm{dist}(x_t^\varepsilon - x_s^\varepsilon, J_{st}[S_{\mathbb{F}}(F \circ x^\varepsilon)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G \circ x^\varepsilon)]) \\
&\qquad + H(J_{st}[S_{\mathbb{F}}(F \circ x^\varepsilon)], J_{st}[S_{\mathbb{F}}(F \circ x)]) \\
&\qquad + H(\mathcal{J}_{st}[S_{\mathbb{F}}(G \circ x^\varepsilon)], \mathcal{J}_{st}[S_{\mathbb{F}}(G \circ x)]) \\
&\quad \le [2 + L(\sqrt{T} + 1)]\|x - x^\varepsilon\|_c \le \varepsilon.
\end{aligned}$$

(v) If μ is a given probability measure on $\beta(\mathbb{R}^d)$, then taking an $\mathcal{F}_0$-measurable random variable ξ such that $P\xi^{-1} = \mu$, we obtain the existence of a strong solution X for $SFD(F, G)$ such that $PX_0^{-1} = \mu$, which implies that $\mathcal{X}_\mu(F, G) \neq \emptyset$, because $(\mathcal{P}_{\mathbb{F}}, X, B) \in \mathcal{X}_\mu(F, G)$. □

We associate now with $SFI(F, G)$ and its weak solution $(\mathcal{P}_{\mathbb{F}}, x, B)$ a set-valued partial differential operator $\mathbb{L}_{FG}^x$ defined on the space $C_b^2(\mathbb{R}^d)$ of all real-valued continuous bounded functions $h : \mathbb{R}^d \to \mathbb{R}$ having continuous bounded partial derivatives h'_{x_i} and $h''_{x_i x_j}$ for $i, j = 1, 2, \ldots$. Assume that $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d \times m})$ are measurable and uniformly square integrably bounded such that $F(t, \cdot)$ and $G(t, \cdot)$ are continuous for fixed $t \in [0, T]$. Let G be diagonally convex and $x = (x_t)_{0 \le t \le T}$ a d-dimensional continuous process on a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$. For every $(f, g) \in S_{\mathbb{F}}(\mathrm{co}\, F \circ x) \times S_{\mathbb{F}}(G \circ x)$, we define a linear operator $\mathbb{L}_{fg}^x : C_b^2(\mathbb{R}^d) \to \mathbb{L}^2([0, T] \times \Omega, \mathbb{R}^d)$ by setting

$$(\mathbb{L}_{fg}^x h)_t = \sum_{i=1}^n h'_{x_i}(x_t) f_t^i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n h''_{x_i x_j}(x_t) \sigma_t^{ij}$$

a.s. for $0 \le t \le T$ and $h \in C_b^2(\mathbb{R}^d)$, where $f_t = (f_t^1, \ldots, f_t^n)$, and $\sigma = g \cdot g^* = (\sigma^{ij})_{n \times m}$. For a process x as given above and sets $A \subset \mathbb{L}^2([0, T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^d)$ and $B \subset \mathbb{L}^2([0, T] \times \Omega, \Sigma_{\mathbb{F}}, \mathbb{R}^{d \times m})$, by $\mathbb{L}_{AB}^x$ we denote a family $\{\mathbb{L}_{fg}^x : (f, g) \in A \times B\}$.

We say that $\mathbb{L}^x_{fg} \in \mathbb{L}^x_{AB}$ generates on $C_b^2(\mathbb{R}^d)$ a continuous local $\mathbb{F}$-martingale if the process $[(\varphi^x_{fg}h)_t]_{0\le t\le T}$ defined by

$$(\varphi^x_{fg}h)_t = h(x_t) - h(x_0) - \int_0^t (\mathbb{L}^x_{fg}h)_\tau \mathrm{d}\tau \quad \text{with} \quad (P.1) \tag{1.1}$$

for $t \in [0,T]$ is for every $h \in C_b^2(\mathbb{R}^d)$ a continuous local $\mathbb{F}$-martingale on $\mathcal{P}_{\mathbb{F}}$. The family of all $\mathbb{L}^x_{fg} \in \mathbb{L}^x_{AB}$ generating on $C_b^2(\mathbb{R}^d)$ a family of continuous local $\mathbb{F}$-martingales is denoted by $\mathcal{M}^x_{AB}$. In what follows, for the set-valued mappings $F : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ as given above, the families $\mathbb{L}^x_{S_{\mathbb{F}}(\mathrm{co}\,F\circ x)S_{\mathbb{F}}(G\circ x)}$ and $\mathcal{M}^x_{S_{\mathbb{F}}(\mathrm{co}\,F\circ x)S_{\mathbb{F}}(G\circ x)}(C_b^2)$ will be denoted by $\mathbb{L}^x_{FG}$ and $\mathcal{M}^x_{FG}$, respectively.

Lemma 1.1. *Assume that $F : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ are measurable and uniformly square integrably bounded such that $F(t,\cdot)$ and $G(t,\cdot)$ are continuous for fixed $t \in [0,T]$. Let G be diagonally convex, and let $x = (x_t)_{0\le t\le T}$ and $\tilde{x} = (\tilde{x}_t)_{0\le t\le T}$ be d-dimensional continuous $\mathbb{F}$- and $\tilde{\mathbb{F}}$-adapted processes on $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$, respectively, such that $Px^{-1} = P\tilde{x}^{-1}$. Then $\mathcal{M}^x_{FG} \ne \emptyset$ if and only if $\mathcal{M}^{\tilde{x}}_{FG} \ne \emptyset$.*

Proof. Let $\mathcal{M}^x_{FG} \ne \emptyset$. There exist $f \in S_{\mathbb{F}}(\mathrm{co}\,F \circ x)$ and $g \in S_{\mathbb{F}}(G \circ x)$ such that for every $h \in C_b^2(\mathbb{R}^d)$, the process $[(\varphi^x_h)_t]_{0\le t\le T}$ defined by (1.1) is a continuous local $\mathbb{F}$-martingale on $\mathcal{P}_{\mathbb{F}}$. Therefore, there exists a sequence $(r_k)_{k=1}^\infty$ of $\mathbb{F}$-stopping times on $\mathcal{P}_{\mathbb{F}}$ such that $r_{k-1} \le r_k$ for $k = 1,2,\dots$ with $r_0 = 0$, $\lim_{k\to\infty} r_k = +\infty$ with $(P.1)$ and such that for every $k = 1,2,\dots$, the process $[(\varphi^x_h)_{t\wedge r_k}]_{0\le t\le T}$ is a continuous square integrable $\mathbb{F}$-martingale on $\mathcal{P}_{\mathbb{F}}$. In particular, it follows that for every $0 \le s < t \le T$, one has $E[(\varphi^x_h)_{t\wedge r_k}|\mathcal{F}_s] = (\varphi^x_h)_{s\wedge r_k}$ with $(P.1)$. Thus for every $0 \le s < t \le T$ and $h \in C_b^2(\mathbb{R}^d)$, we have $E\{[(\varphi^x_h)_{t\wedge r_k}) - (\varphi^x_h)_{s\wedge r_k}]|\mathcal{F}_s\} = 0$ with $(P.1)$. Let $l \in \mathcal{C}_1$. By the continuity of $l \in \mathcal{C}_1$ and the $\mathcal{F}_s$-measurability of x_s, the random variable $l(x_s)$ is also $\mathcal{F}_s$-measurable. Therefore, $E\{(l(x_s)[(\varphi^x_h)_{t\wedge r_k}) - (\varphi^x_h)_{s\wedge r_k}]|\mathcal{F}_s\} = 0$ with $(P.1)$ for every $0 \le s < t \le T$, which, in particular, implies that $E(l(x_s)[(\varphi^x_h)_{t\wedge r_k}) - (\varphi^x_h)_{s\wedge r_k}]) = 0$. Thus in the limit $k \to \infty$, we obtain $E\{(l(x_s)[(\varphi^x_h)_t - (\varphi^x_h)_s]) = 0$. Then

$$E\left(l(x_s)[(h(x_t) - h(x_s)]\right) = E\left(l(x_s)\int_s^t (\mathbb{L}^x_{fg}h)_\tau \mathrm{d}\tau\right)$$

for every $0 \le s < t \le T$, $l \in \mathcal{C}_1$, and $h \in C_b^2(\mathbb{R}^d)$. By virtue of Theorem 4.2 of Chap. 3, there exist $\tilde{f} \in S_{\tilde{\mathbb{F}}}(\mathrm{co}F \circ \tilde{x})$ and $\tilde{g} \in S_{\tilde{\mathbb{F}}}(G \circ \tilde{x})$ such that

$$E\int_s^t l(x_s)(\mathbb{L}^x_{fg}h)_\tau \mathrm{d}\tau = \tilde{E}\int_s^t l(\tilde{x}_s)(\mathbb{L}^{\tilde{x}}_{\tilde{f}\tilde{g}}h)_\tau \mathrm{d}\tau$$

for every $0 \le s < t \le T$, $l \in \mathcal{C}_1$, and $h \in C_b^2(\mathbb{R}^r)$. But

$$E\int_s^t l(x_s)(\mathbb{L}_{fg}^x h)_\tau \mathrm{d}\tau = E\left[l(x_s)\int_s^t (\mathbb{L}_{fg}^x h)_\tau \mathrm{d}\tau\right],$$

$$\tilde{E}\int_s^t l(\tilde{x}_s)(\mathbb{L}_{\tilde{f}\tilde{g}}^{\tilde{x}} h)_\tau \mathrm{d}\tau = \tilde{E}\left[l(\tilde{x}_s)\int_s^t (\mathbb{L}_{\tilde{f}\tilde{g}}^{\tilde{x}} h)_\tau \mathrm{d}\tau\right] \quad \text{and}$$

$$E\{l(x_s)[h(x_t) - h(x_s)]\} = \tilde{E}\{l(\tilde{x}_s)[h(\tilde{x}_t) - h(\tilde{x}_s)]\}$$

for every $0 \le s < t \le T$, because $l \in \mathcal{C}_1$ and $h \in C_b^2(\mathbb{R}^d)$ are continuous and $Px^{-1} = P\tilde{x}^{-1}$. Therefore,

$$\tilde{E}\left\{l(\tilde{x}_s)[h(\tilde{x}_t) - h(\tilde{x}_s)]\right\} = \tilde{E}\left\{l(x_s)\int_s^t (\mathbb{L}_{\tilde{f}\tilde{g}}^{\tilde{x}} h)_\tau \mathrm{d}\tau\right\}$$

for $0 \le s < t \le T$, $l \in \mathcal{C}_1$, and $h \in C_b^2(\mathbb{R}^d)$. Then $\tilde{E}\{l(\tilde{x}_s)[(\varphi_h^{\tilde{x}})_t - (\varphi_h^{\tilde{x}})_s]\} = 0$, which, in particular, implies that $\tilde{E}[l(\tilde{x}_s)\cdot E\{[(\varphi_n^{\tilde{x}})_t - (\varphi_n^{\tilde{x}})_s]|\tilde{\mathcal{F}}_s\}] = 0$ for $0 \le s < t \le T$, $l \in C_1$, and $h \in C_b^2(\mathbb{R}^d)$. By the monotone class theorem, it follows that the above equality is also true for every measurable bounded function $l : \mathbb{R}^d \to \mathbb{R}$. Taking in particular l such that $l(\tilde{x}_s) = \tilde{E}\{[(\varphi_h^{\tilde{x}})_t - (\varphi_h^{\tilde{x}})_s]|\tilde{\mathcal{F}}_s\}$ with $(\tilde{P}.1)$, we get $\tilde{E}|\tilde{E}\{[(\varphi_n^{\tilde{x}})_t - (\varphi_n^{\tilde{x}})_s]|\tilde{\mathcal{F}}_s\}|^2 = 0$ for $0 \le s < t \le T$ and $h \in C_b^2(\mathbb{R}^d)$. Therefore, $\tilde{E}\{[(\varphi_n^{\tilde{x}})_t - (\varphi_n^{\tilde{x}})_s]|\tilde{\mathcal{F}}_s\} = 0$ with $(\tilde{P}.1)$ for every $0 \le s < t \le T$ and $h \in C_b^2(\mathbb{R}^d)$. Then $\mathbb{L}_{\tilde{f}\tilde{g}}^{\tilde{x}} \in \mathcal{M}_{FG}^{\tilde{x}}(C_b^2)$. In a similar way, we can verify that $\mathcal{M}_{FG}^{\tilde{x}} \neq \emptyset$ implies that $\mathcal{M}_{FG}^{x} \neq \emptyset$. □

Lemma 1.2. *Assume that* $F : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ *and* $G : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ *are measurable and uniformly square integrably bounded such that* $F(t,\cdot)$ *and* $G(t,\cdot)$ *are continuous for fixed* $t \in [0,T]$. *Let* G *be diagonally convex and let* $(x_t)_{0\le t\le T}$ *and* $(x_t^k)_{0\le t\le T}$ *be* d*-dimensional continuous stochastic processes on* $(\Omega, \mathcal{F}, \mathbb{F}, P)$ *for every* $k = 1,2,\ldots$ *such that* $\lim_{k\to\infty} P(\{\sup_{0\le t\le T}|x_t - x^k| > \varepsilon\}) = 0$ *for every* $\varepsilon > 0$ *and* $\mathcal{M}_{FG}^{x^k} \neq \emptyset$ *for every* $k = 1,2,\ldots$. *Then* $\mathcal{M}_{FG}^{x} \neq \emptyset$.

Proof. Let $f^k \in S_{\mathbb{F}}(\mathrm{co}F \circ x^k)$ and $g^k \in S_{\mathbb{F}}(G \circ x^k)$ be such that $\mathbb{L}_{f_k g_k}^{x^k} \in \mathcal{M}_{FG}^{x^k}$ for every $k = 1,2,\ldots$. Let $(x^{k_r})_{r=1}^{\infty}$ be a subsequence of $(x^k)_{k=1}^{\infty}$ such that $\lim_{r\to\infty}\sup_{0\le t\le T}|x_t - x_t^{k_r}| = 0$ with $(P.1)$. By the uniform square integrably boundedness of $F \circ x^k$, it follows that the sequence $(f^{k_r})_{r=1}^{\infty}$ is weakly compact. Then there exist a d-dimensional $\mathbb{F}$-nonanticipative process f and a subsequence, still denoted by $(f^{k_r})_{r=1}^{\infty}$, of $(f^{k_r})_{r=1}^{\infty}$ weakly converging to f. For every $A \in \Sigma_{\mathbb{F}}$ and $k = 1,2,\ldots$, one has

$$\begin{aligned}
&\mathrm{dist}\left(\int_A f_t(\omega)\mathrm{d}t\mathrm{d}P, \int_A \mathrm{co}F(t,x_t(\omega))\mathrm{d}t\mathrm{d}P\right)\\
&\le \left|\int_A f_t(\omega)\mathrm{d}t\mathrm{d}P - \int_A f_t^{k_r}\mathrm{d}t\mathrm{d}P\right|\\
&\quad + h\left(\int_A \mathrm{co}F(t,x_t^{k_r}(\omega))\mathrm{d}t\mathrm{d}P, \int_A \mathrm{co}F(t,x_t(\omega))\mathrm{d}t\mathrm{d}P)\right).
\end{aligned}$$

Then $\int_A f_t(\omega)\mathrm{d}t\mathrm{d}P \in \int_A \mathrm{co}F(t,x_t(\omega))\mathrm{d}t\mathrm{d}P$ for every $A \in \Sigma_{\mathbb{F}}$, which implies that $f \in S(\mathrm{co}\,F \circ x)$. Hence, by the properties of the set-valued mapping $\Phi(\varphi,\cdot)$ defined in Sect. 4 of Chap. 3, it follows that

$$\lim_{r\to\infty} E\left(l(x_s^{k_r})\int_s^t \Phi(\varphi(h), f_\tau^{k_r})(x_\tau^{k_r})\mathrm{d}\tau\right) = E\left(l(x_s)\int_s^t \Phi(\varphi(h), f_\tau)(x_\tau)\mathrm{d}\tau\right)$$

for every $0 \le s < t \le T$, $l \in \mathcal{C}^1$, and $h \in C_b^2(\mathbb{R}^d)$. In a similar way, we can verify the existence of $g \in S_{\mathbb{F}}(G \circ x)$ such that

$$\lim_{r\to\infty} E\left(l(x_s^{k_r})\int_s^t \Psi(\psi(h), \sigma_\tau^{k_r})(x_\tau^{k_r})\mathrm{d}\tau\right) = E\left(l(x_s)\int_s^t \Psi(\psi(h), \sigma_\tau)(x_\tau)\mathrm{d}\tau\right)$$

for every $0 \le s < t \le T$, $l \in \mathcal{C}^1$, and $h \in C_b^2(\mathbb{R}^d)$, where $\Psi(\psi,\cdot)$ is defined in Sect. 4 of Chap. 3, $\sigma^{k_r} = g^{k_r}\cdot(g^{k_r})^*$, and $\sigma = g\cdot g^*$. By the definitions of $\mathbb{L}^x_{fg}$ and mappings $\Phi(\varphi,\cdot)$ and $\Psi(\psi,\cdot)$, it follows that

$$\lim_{r\to\infty} E\left(l(x_s^{k_r})\int_s^t (\mathbb{L}^{x^{k_r}}_{f^{k_r}g^{k_r}}h)_\tau \mathrm{d}\tau\right) = E\left(l(x_s)\int_s^t (\mathbb{L}^x_{fg}h)_\tau\right)\mathrm{d}\tau,$$

for every $0 \le s < t \le T$, $l \in \mathcal{C}^1$, and $h \in C_b^2(\mathbb{R}^d)$. But $\mathbb{L}^{x^k}_{f^kg^k} \in \mathcal{M}^{x^k}_{FG}$ for $k = 1,2,\ldots$. Then

$$E\left(l(x_s^{k_r})[h(x_t^{k_r}) - h(x_s^{k_r})]\right) = E\left(l(x_s^{k_r})\int_s^t (\mathbb{L}^{x^{kr}}_{f^{kr}g^{kr}}h)_\tau \mathrm{d}\tau\right)$$

for every $0 \le s < t \le T$, $k = 1,2,\ldots$, $l \in \mathcal{C}_1$, and $h \in C_b^2(\mathbb{R}^d)$. Passing to the limit as $r \to \infty$, we obtain $E\{l(x_s)[(\varphi_h^x)_t - (\varphi_h^x)_s]\} = 0$ for $0 \le s < t \le T$, $l \in \mathcal{C}_1$, and $h \in C_b^2(\mathbb{R}^d)$. Similarly as in the proof of Lemma 1.1, it follows that $\mathbb{L}^x_{fg} \in \mathcal{M}^x_{FG}$. Then $\mathcal{M}^x_{FG} \ne \emptyset$. □

Remark 1.5. In a similar way, it can be verified that by the assumptions of Lemma 1.2, without the continuity of $F(t,\cdot)$ and $G(t,\cdot)$ for fixed $t \in [0,T]$ the, nonemptiness of $\mathcal{M}^{x^k}_{FG}$ for every $k = 1,2,\ldots$ implies that $\mathcal{M}^x_{FG} \ne \emptyset$. □

Lemma 1.3. *Assume that $F : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ are measurable and uniformly square integrably bounded such that $F(t,\cdot)$ and $G(t,\cdot)$ are continuous for fixed $t \in [0,T]$. Let G be diagonally convex and let $(x_t^k)_{0\le t\le T}$ be for every $k = 1,2,\ldots$, a d-dimensional continuous $\mathbb{F}^k$-adapted stochastic process on $(\Omega^k,\mathcal{F}^k,\mathbb{F}^k,P^k)$ such that $\mathcal{M}^{x^k}_{FG} \ne \emptyset$ for every $k = 1,2,\ldots$. Let $\tilde{x}^k = (\tilde{x}_t^k)_{0\le t\le T}$ and $\tilde{x} = (\tilde{x}_t)_{0\le t\le T}$ be for $k = 1,2,\ldots$, continuous d-dimensional $\tilde{\mathbb{F}}$-adapted processes on $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{F}},\tilde{P})$ such that $P(\tilde{x}^k)^{-1} = P(x^k)^{-1}$ for $k = 1,2,\ldots$ and $\lim_{k\to\infty}\tilde{P}(\{\sup_{0\le t\le T}|\tilde{x}_t - \tilde{x}^k| > \varepsilon\}) = 0$ for every $\varepsilon > 0$. Then $\mathcal{M}^{\tilde{x}}_{FG} \ne \emptyset$.*

Proof. By virtue of Lemma 1.1, one has $\mathcal{M}_{FG}^{\tilde{x}^k} \neq \emptyset$ for every $k = 1, 2, \ldots$, which by Lemma 1.2, implies that $\mathcal{M}_{FG}^{\tilde{x}} \neq \emptyset$. □

Lemma 1.4. *Let $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ be measurable and uniformly square integrably bounded. If $(x_t, B_t)_{0\leq t\leq T}$ is a weak solution of $SFI(\mathrm{co}\, F, G)$ on a complete probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$, then there is a sequence $(x^k)_{k=1}^{\infty}$ of Itô processes $x^k = (x_t^k)_{0\leq t\leq T}$ of the form $x_t^k = x_0 + \int_0^t f_\tau^k \mathrm{d}\tau + \int_0^t g_\tau^k \mathrm{d}B_\tau$ a.s. for $t \in [0, T]$ with $f^k \in S_{\mathbb{F}}(\mathrm{co} F \circ x))$ and $g^k \in S_{\mathbb{F}}(G \circ x))$ such that $\lim_{k\to\infty} P(\{\sup_{0\leq t\leq T} |x_t - x_t^k| > \varepsilon\}) = 0$ for every $\varepsilon > 0$.*

Proof. By virtue of Theorem 1.4 of Chap. 3, there are sequences $(f^k)_{k=1}^{\infty}$ and $(g^k)_{k=1}^{\infty}$ of $S_{\mathbb{F}}(\mathrm{co}\, F \circ x)$ and $S_{\mathbb{F}}(G \circ x)$, respectively, such that $\sup_{0\leq t\leq T} E|x_t - x_t^k|^2 \to 0$ as $k \to \infty$, where $x_t^k = x_0 + \int_0^t f_\tau^k \mathrm{d}\tau + \int_0^t g_\tau^k \mathrm{d}B_\tau$ with $(P.1)$ for $t \in [0, T]$ and $k = 1, 2, \ldots$. By Theorem 3.4 of Chap. 1, we can assume that $(x_t)_{0\leq t\leq T}$ and $(x_t^k)_{0\leq t\leq T}$ are continuous for $k \geq 1$ because for $\alpha = 2r$, and $\beta = r$ with $r \geq 1$, there is a positive number M such that $E|x_t - x_s|^\alpha \leq M|t - s|^{1+\beta}$ and $E|x_t^k - x_s^k|^\alpha \leq M|t - s|^{1+\beta}$ for every $0 \leq s < t \leq T$ and $k = 1, 2, \ldots$. For every $\varepsilon > 0$, $0 \leq s < t \leq T$, and $k = 1, 2, \ldots$, we have

$$P(\{|x_t - x_t^k| > \varepsilon\}) \leq \frac{1}{\varepsilon^\alpha} E|x_t - x_t^k|^\alpha, \quad P(\{|x_t - x_s| > \varepsilon\}) \leq \frac{1}{\varepsilon^\alpha} E|x_t - x_s|^\alpha$$

and

$$P(\{|x_t^k - x_s^k| > \varepsilon\}) \leq \frac{1}{\varepsilon^\alpha} E|x_t^k - x_s^k|^\alpha.$$

Then for every $m = 1, 2, \ldots$, there is a positive integer k_m such that

$$\max\,[P(\{|x_{i/2^m} - x_{i/2^m}^k| > 1/2^{ma}\}),$$
$$P(\{|x_{(i+1)/2^m} - x_{i/2^m}| > 1/2^{ma}\}),$$
$$P(\{|x_{(i+1)/2^m}^k - x_{i/2^m}^k| > 1/2^{ma}\})] \leq M\frac{2^{m\alpha}}{2^{m(1+\beta)}}$$

for $k \geq k_m$ and $0 \leq i \leq 2^m T - 1$, where $a > 0$ is such that $a < \beta/\alpha$.

Hence in particular, it follows that

$$\max\left[P\left(\left\{\max_{0\leq i\leq 2^m T-1} |x_{i/2^m} - x_{i/2^m}^k| > i/2^{ma}\right\}\right),\right.$$
$$P\left(\left\{\max_{0\leq i\leq 2^m T-1} |x_{(i+1)/2^m} - x_{i/2^m}^k| > i/2^{ma}\right\}\right),$$
$$\left.P\left(\left\{\max_{0\leq i\leq 2^m T-1} |x_{(i+1)/2^m}^k - x_{i/2^m}^k| > i/2^{ma}\right\}\right)\right]$$
$$\leq MT2^{-m(\beta-a\alpha)}$$

for $k \geq k_m$ and $m = 1, 2, \ldots$. For $\varepsilon > 0$ and $\delta > 0$ select $\nu = \nu(\varepsilon, \delta)$ such that $(1 + 2/(2^a - 1))/2^{\nu a} \leq \varepsilon$ and $\sum_{m=\nu}^{\infty} 2^{-m(\beta - a\alpha)} \leq \frac{\delta}{3MT}$. For every $m \geq \nu$ and $k \geq k_m$, one gets

$$P\left(\bigcup_{m=\nu}^{\infty}\left\{\max_{0\leq i\leq 2^m T-1}|x_{i/2^m} - x^k_{i/2^m}| > 1/2^{ma}\right\}\right)$$

$$\leq \delta/\varepsilon,\ P\left(\bigcup_{m=\nu}^{\infty}\left\{\max_{0\leq i\leq 2^m T-1}|x_{(i+1)/2^m} - x_{i/2^m}| > 1/2^{ma}\right\}\right) \leq \delta/\varepsilon$$

$$\text{and } P\left(\bigcup_{m=\nu}^{\infty}\left\{\max_{0\leq i\leq 2^m T-1}|x^k_{(i+1)/2^m} - x^k_{i/2^m}| > 1/2^{ma}\right\}\right) \leq \delta\varepsilon.$$

Let

$$\Omega_\nu^{1,k} = \bigcup_{m=\nu}^{\infty}\left\{\max_{0\leq i\leq 2^m T-1}|x_{i/2} - x^k_{i/2^m}| > 1/2^{ma}\right\},$$

$$\Omega_\nu^{2} = \bigcup_{m=\nu}^{\infty}\left\{\max_{0\leq i\leq 2^m T-1}|x_{(i+1)/2^m} - x_{i/2^m}| > 1/2^{ma}\right\}$$

$$\text{and } \Omega_\nu^{3,k} = \bigcup_{m=\nu}^{\infty}\left\{\max_{0\leq i\leq 2^m T-1}|x^k_{(i+1)/2^m} - x^k_{i/2^m}| > 1/2^{ma}\right\}$$

for $k \geq k_\nu$. Taking $\Omega_\nu^k = \Omega_\nu^{1,k} \cup \Omega_\nu^2 \cup \Omega_\nu^{3,k}$, one obtains $P(\Omega_\nu^k) \leq \delta$ for every $k \geq k_\nu$. By the definition of Ω_ν^k, for every $\omega \notin \Omega_\nu^k$, $k \geq k_\nu$, and $0 \leq i \leq 2^\nu T - 1$, we get

$$|x_{i/2^\nu} - x^k_{i/2^\nu}| \leq \frac{1}{2^{\nu a}},\ |x_{(i+1)/2^\nu} - x_{i/2^\nu}| \leq \frac{1}{2^{\nu a}} \text{ and } |x^k_{(i+1)/2^\nu} - x^k_{i/2^\nu}| \leq \frac{1}{2^{\nu a}}.$$

Let D_T be the set of dyadic numbers of $[0, T]$. For every $t \in D_T \cap [i/2^\nu, (i+1)/2^\nu]$, one has $t = i/2^\nu + \sum_{i=1}^{j} \alpha_l/2^{\nu+1}$ with $\alpha_l \in \{0, 1\}$ for $l = 1, 2, \ldots, j$. For every $k \geq k_\nu$, $\omega \notin \Omega_\nu^k$ and i fixed above, we get

$$\begin{aligned}
|x_t - x_t^k| &\leq |x_t - x_{i/2^\nu}| + |x_{i/2^\nu} - x^k_{i/2^\nu}| + |x^k_{i/2^\nu} - x^k_t|\\
&\leq \sum_{r=1}^{j}|x_{i/2^\nu + \sum_{l=1}^{r}\alpha_l/2^{\nu+l}} - x_{i/2^\nu + \sum_{l=1}^{r-1}\alpha_l/2^{\nu+l}}| + |x_{i/2^\nu} - x^k_{i/2^\nu}|\\
&\quad + \sum_{r=1}^{j}|x_{i/2^\nu + \sum_{l=1}^{r}\alpha_l/2^{\nu+l}} - x_{i/2^\nu + \sum_{l=1}^{r}\alpha_l/2^{\nu+l}}| \leq 2\sum_{r=1}^{j} 1/2^{(\nu+r)a} + \frac{1}{2^{\nu a}}
\end{aligned}$$

$$\leq 2\sum_{r=1}^{\infty} 1/2^{(\nu+r)a} + \frac{1}{2^{\nu a}} = \frac{2}{(2^a-1)2^{\nu a}} + \frac{1}{2^{\nu a}}$$
$$= (1 + 2/(2^a-1)2^{\nu a}) \leq \varepsilon.$$

But D_T is dense in $[0,T]$, and $(x_t)_{0\leq t\leq T}$ and $(x_t^k)_{0\leq t\leq T}$ are continuous. Then for every $k \geq k_\nu$ and $\omega \notin \Omega_\nu^k$, one obtains $|x_t(\omega) - x_t^k(\omega)| \leq \varepsilon$ for $t \in [0,T]$, which implies that

$$P(\{\max_{0\leq t\leq T} |x_t - x_t^k| > \varepsilon\}) \leq P(\Omega_\nu^k) < \delta$$

for every $k \geq k_\nu$. Thus for every $\varepsilon > 0$ and $\delta > 0$, there is $k_\nu = k_{\nu(\varepsilon,\delta)}$ such that

$$P\left(\left\{\sup_{0\leq t\leq T} |x_t - x_t^k| > \varepsilon\right\}\right) \leq \delta$$

for $k \geq k_\nu$, i.e., $\lim_{k\to\infty} P(\{\sup_{0\leq t\leq T} |x_t - x_t^k| > \varepsilon\}) = 0$ for every $\varepsilon > 0$. □

Theorem 1.3. *Let $F : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ be measurable and uniformly square integrably bounded and let G be diagonally convex. For every probability measure μ on $\beta(\mathbb{R}^d)$, the problem $SFI(\mathrm{co}\,F, G, \mu)$ possesses at least one weak solution with an initial distribution μ if and only if there exist a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ and a d-dimensional continuous $\mathbb{F}$-adaptive stochastic process $x = (x_t)_{0\leq t\leq T}$ on $\mathcal{P}_{\mathbb{F}}$ such that $Px_0^{-1} = \mu$ and $\mathcal{M}_{FG}^x \neq \emptyset$.*

Proof. ($\Rightarrow$) Let $(\mathcal{P}_{\mathbb{F}}, x, B)$ be a weak solution of $SFI(\mathrm{co}\,F, G, \mu)$ with $x = (x_t)_{0\leq t\leq T}$. By virtue of Lemma 1.4, there exist sequences $(f^k)_{k=1}^{\infty}$ and $(g^k)_{k=1}^{\infty}$ of $S_{\mathbb{F}}(\mathrm{co}\,F \circ x))$ and $S_{\mathbb{F}}(G \circ x)$, respectively, such that the sequence $(x^k)_{k=1}^{\infty}$ of continuous $\mathbb{F}$-adapted processes $x^k = (x_t^k)_{0\leq t\leq T}$ defined by $x_t^k = x_0 + \int_0^t f_\tau^k + \int_0^t g_\tau^k \mathrm{d}B_\tau$ a.s. for $0 \leq t \leq T$ is such that $\lim_{k\to\infty} P(\{\sup_{0\leq t\leq T} |x_t - x_t^k| > \varepsilon\}) = 0$ for every $\varepsilon > 0$. By Itô's formula, for every $h \in C_b^2(\mathbb{R}^d)$ and $k = 1, 2\ldots$ one obtains

$$h(x_t^k) - h(x_0^k) - \int_0^t (\mathbb{L}_{f^k g^k}^{x^k} h)_\tau \mathrm{d}\tau = \sum_{i=1}^n \sum_{j=1}^n \int_0^t h'_{x_i}(x_\tau^k)(g^k)_\tau^{ij} \mathrm{d}B_\tau^j$$

with $(P.1)$ for $t \in [0,T]$, where $B_t = (B_t^1, \ldots, B_t^m)^*$ and $g_t^k = [(g^k)_t^{ij}]_{d\times m}$ for $0 \leq t \leq T$. By the definition of $[\varphi_{f^k g^k}^{x^k} h]_t$, the above equality can be written in the form

$$[\varphi_{f^k g^k}^{x^k} h]_t = \sum_{i=1}^n \sum_{j=1}^n \int_0^t h'_{x_i}(x_\tau^k)(g^k)_\tau^{ij} \mathrm{d}B_\tau^j$$

with $(P.1)$ for $t \in [0,T]$. Hence, by the properties of Itô integrals, it follows that $[(\varphi^{x^k}_{f^k g^k} h)_t]_{0\le t\le T}$ is a continuous local $\mathbb{F}$-martingale on $\mathcal{P}_{\mathbb{F}}$ for every $k = 1,2,\dots$ and $h \in C_b^2(\mathbb{R}^d)$. Therefore, $\mathcal{M}^{x^k}_{FG} \neq \emptyset$ for $k = 1,2,\dots$, which by Remark 1.5 implies that $\mathcal{M}^x_{FG} \neq \emptyset$.

($\Leftarrow$) Let $\mathcal{P}_{\mathbb{F}} = (\Omega,\mathcal{F},\mathbb{F},P)$ be a filtered probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ and $(x_t)_{0\le t\le T}$ a d-dimensional continuous $\mathbb{F}$-adapted process on $\mathcal{P}_{\mathbb{F}}$ such that $x_0^{-1} = \mu$ and $\mathcal{M}^x_{FG} \neq \emptyset$. Then there exist $f \in S_{\mathbb{F}}(\mathrm{co}F \circ x)$ and $g \in S_{\mathbb{F}}(G \circ x)$ such that $\mathbb{L}^x_{fg} \in \mathcal{M}^x_{FG}$. Let $(\tau_k)_{k=1}^{\infty}$ be a sequence of stopping times $\tau_k = \inf\{t \in [0,T] : x_t \notin K_k\}$, where $K_k = \{x \in \mathbb{R}^d : |x| \le k\}$ for $k = 1,2,\dots$. Select now, in particular, $h_i \in C_b^2(\mathbb{R}^d)$ such that $h_i(x) = x_i$ for $x \in K_k$, where $x = (x^1,\dots,x^n)$. For such $h_i \in C_b^2(\mathbb{R}^d)$, we have

$$\int_0^{t\wedge\tau_k} (\mathbb{L}^x_{fg} h_i)_\tau \mathrm{d}\tau = \int_0^{t\wedge\tau_k} f^i_\tau \mathrm{d}\tau \ \text{ and hence } \ (\varphi^x_{h_i})_{t\wedge\tau_k} = x^i_{t\wedge\tau_l} - x^i_0 - \int_0^{t\wedge\tau_k} f^i_\tau \mathrm{d}\tau$$

a.s. for $k \ge 1$ and $i = 1,2,\dots,d$ and $t \in [0,T]$. But $\mathbb{L}^x_{fg} \in \mathcal{M}^x_{FG}(C_b^2)$. Then $[(\varphi^x_{h_i})_{t\wedge\tau_l}]_{0\le t\le T}$ is for every $i = 1,\dots,d$ and $k = 1,2,\dots$ a continuous local $\mathbb{F}$-martingale on $\mathcal{P}_{\mathbb{F}}$. Let $M^i_t = (\varphi^x_{h_i})_t$ for $i = 1,\dots,d$ and $t \in [0,T]$. Taking, in particular, $h_{ij} \in C_b^2(\mathbb{R}^d)$ such that $h_{ij}(x) = x^i x^j$ for $x \in K_k$ and $i,j = 1,2,\dots,d$, we obtain a family $(M^{ij}_t)_{0\le t\le T}$ for $i,j = 1,\dots,d$ of continuous local $\mathbb{F}$-martingales on $\mathcal{P}_{\mathbb{F}}$ such that

$$M^{ij}_t = x^i_t x^j_t - x^i_0 x^j_0 - \int_0^t [x^i_\tau f^j_\tau + x^j_\tau f^i_\tau x_\tau) + \sigma^{ij}_\tau]\mathrm{d}\tau$$

a.s. for $i,j = 1,2,\dots,n$ and $t \in [0,T]$, where $\sigma = g \cdot g^*$. Let $\sigma = (\sigma^{ij})_{d\times d}$. Similarly as in the proof of Theorem 9.1 of Chap. 1, it follows that

$$\langle M^i, M^j\rangle_t = \int_0^t \sigma^{ij}_\tau \mathrm{d}\tau$$

a.s. for $i,j = 1,2,\dots,d$ and $t \in [0,T]$, which similarly as in the proof of Theorem 9.1 of Chap. 1, implies that there exist a standard extension $\hat{\mathcal{P}}_{\hat{\mathbb{F}}} = (\hat\Omega,\hat{\mathcal{F}},\hat{\mathbb{F}},\hat P)$ of $(\Omega,\mathcal{F},\mathbb{F},P)$ and an m-dimensional $\hat{\mathbb{F}}$-Brownian motion $\hat B = (\hat B_t)_{0\le t\le T}$ on $(\hat\Omega,\hat{\mathcal{F}},\hat{\mathbb{F}},\hat P)$ such that

$$M^i_t = \sum_{j=1}^m \int_0^t \hat g^{ij}_\tau \mathrm{d}\hat B^j_\tau$$

$\hat P$-a.s. for $i = 1,2\dots,d$ and $t \in [0,T]$, with $\hat g_t(\hat\omega) = g_t(\pi(\hat\omega))$ for $\hat\omega \in \hat\Omega$, where $\pi : \hat\Omega \to \Omega$ is the $(\hat{\mathcal{F}},\mathcal{F})$-measurable mapping described in the definition of the extension of $\mathcal{P}_{\mathbb{F}} = (\Omega,\mathcal{F},\mathbb{F},P)$ because a standard extension $\hat{\mathcal{P}}_{\hat{\mathbb{F}}}$ of $\mathcal{P}_{\mathbb{F}}$ is also an extension of it. Let $\hat x_t(\hat\omega) = x_t(\pi(\hat\omega))$ for $\hat\omega \in \hat\Omega$. For every $A \in \beta_T$, we get $(P\hat x_0^{-1})(A) = \hat P[\hat x_0^{-1}(A)] = \hat P[(x\circ\pi)^{-1}(A)] = (\hat P\circ\pi^{-1})[(x_0^{-1}(A)] = P[x_0^{-1}(A)] = (Px_0^{-1})(A) = \mu(A)$, which implies that $P\hat x_0^{-1} = \mu$. By the definition of M^i_t, it follows that

$$\hat{x}_t^i = \hat{x}_0^i + \int_0^t \hat{f}_\tau^i \mathrm{d}\tau + \sum_{j=1}^m \int_0^t \hat{g}^{ij}(\tau, \hat{x}_\tau)\mathrm{d}\hat{B}_\tau^j$$

$\hat{P}$-a.s. for $i = 1, 2, \ldots, d$ and $t \in [0, T]$, where $\hat{f}_t(\hat{\omega}) = f_t(\pi(\hat{\omega}))$ for $\hat{\omega} \in \hat{\Omega}$. Then

$$\hat{x}_t = \hat{x}_0 + \int_0^t \hat{f}_\tau \mathrm{d}\tau + \int_0^t \hat{g}_\tau \mathrm{d}\hat{B}_\tau$$

$\hat{P}$-a.s. for $0 \leq t \leq T$. Therefore, $\hat{x}_t - \hat{x}_s \in J_{st}[S_{\hat{\mathbb{F}}}(\mathrm{co}\, F \circ \hat{x})] + \mathcal{J}_{st}[S_{\hat{\mathbb{F}}}(G \circ \hat{x})]$ for every $0 \leq s < t \leq T$ and $P\hat{x}_0^{-1} = \mu$. Thus $(\hat{\mathcal{P}}_{\hat{\mathbb{F}}}, \hat{x}, \hat{B})$ is a weak solution of $SFI(\mathrm{co}F, G, \mu)$. □

Theorem 1.4. *Let $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ be measurable and uniformly square integrably bounded, and let G be diagonally convex. For every probability measure μ on $\beta(\mathbb{R}^n)$, the problem $SFI(\mathrm{co}\, F, G, \mu)$ possesses a weak solution $(\mathcal{P}_{\mathbb{F}}, x, B)$ with a separable filtered probability space $\mathcal{P}_{\mathbb{F}}$ if and only if there exist a separable filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ and a d-dimensional continuous $\mathbb{F}$-adaptive stochastic process $x = (x_t)_{0\leq t\leq T}$ on $\mathcal{P}_{\mathbb{F}}$ such that $Px_0^{-1} = \mu$ and $\mathcal{M}_{FG}^x \neq \emptyset$.*

Proof. Similarly as of the proof of Theorem 1.3, we can verify that if $(\mathcal{P}_{\mathbb{F}}, x, B)$ is a weak solution of $SFI(\mathrm{co}\, F, G, \mu)$ with a separable filtered probability space $\mathcal{P}_{\mathbb{F}}$, then $\mathcal{M}_{FG}^x \neq \emptyset$. Let $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ be a separable filtered probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$, and $(x_t)_{0\leq t\leq T}$ a d-dimensional continuous $\mathbb{F}$-adapted process on $\mathcal{P}_{\mathbb{F}}$ such that $\mathcal{M}_{FG}^x \neq \emptyset$. Then there exist $f \in S_{\mathbb{F}}(\mathrm{co}F \circ x)$ and $g \in S_{\mathbb{F}}(G \circ x)$ such that $\mathbb{L}_{fg}^x \in \mathcal{M}_{FG}^x$. Similarly as in the proof of Theorem 1.3, we can define a local $\mathbb{F}$-martingale $(M_t^i)_{0\leq t\leq T}$, on $\mathcal{P}_{\mathbb{F}}$ such that $\langle M^i, M^j\rangle_t = \int_0^t \sigma_\tau^{ij} \mathrm{d}\tau$ with $(P.1)$ for $i, j = 1, \ldots, d$ and $t \in [0, T]$. Therefore, by virtue of Theorem 8.2 of Chap. 1 and Remark 8.2 of Chap. 1, there exist a standard separable extension $\hat{\mathcal{P}}_{\hat{\mathbb{F}}} = (\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$ of $(\Omega, \mathcal{F}, \mathbb{F}, P)$ and an $\hat{\mathbb{F}}$-Brownian motion $\hat{B} = (\hat{B}_t^1, \ldots, \hat{B}_t^m)_{0\leq t\leq T}$ on $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$ such that

$$M_t^i = \sum_{j=1}^m \int_0^t \hat{g}_\tau^{ij} \mathrm{d}\hat{B}^j{}_\tau$$

$\hat{P}$-a.s. for $i = 1, 2, \ldots, d$ and $t \in [0, T]$, where $\hat{x}$ and $\hat{g}$ denote extensions of x and g on $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$ defined in the usual way. It is clear that $P\hat{x}_0^{-1} = \mu$. Hence it follows that

$$\hat{x}_t^i = \hat{x}_0^i + \int_0^t \hat{f}_\tau^i \mathrm{d}\tau + \sum_{j=1}^m \int_0^t \hat{g}_\tau^{ij} \mathrm{d}\hat{B}_\tau^j$$

$\hat{P}$-a.s. for $i = 1, 2, \dots, d$ and $t \in [0, T]$, where $\hat{f}$ denotes an extension of f on $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$. Then

$$\hat{x}_t = \hat{x}_0 + \int_0^t \hat{f}_\tau \mathrm{d}\tau + \int_0^t \hat{g}_\tau \mathrm{d}\hat{B}_\tau$$

$\hat{P}$-a.s. for $0 \le t \le T$ with $P\hat{x}_0^{-1} = \mu$. Therefore, $(\hat{\mathcal{P}}_{\hat{\mathbb{F}}}, \hat{x}, \hat{B})$ is a weak solution of $SFI(co\ F, G, \mu)$ with a separable filtered probability space $\hat{\mathcal{P}}_{\hat{\mathbb{F}}}$. □

It follows immediately from Theorem 1.2 that if F and G satisfy the assumptions of Theorem 1.1, then $\mathcal{X}_\mu(F, G) \neq \emptyset$ for every probability measure μ on $\beta(\mathbb{R}^d)$. We shall show that if F and G are convex-valued and G is diagonally convex, then for nonemptiness of $\mathcal{X}_\mu(F, G)$, it is enough to assume that $F(t, \cdot)$ and $G(t, \cdot)$ are continuous instead of Lipschitz continuous.

Theorem 1.5. *Let $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ be measurable, uniformly square integrably bounded, and convex-valued such that $F(t, \cdot)$ and $G(t, \cdot)$ are continuous for a.e. fixed $t \in [0, T]$. If G is diagonally convex, then $\mathcal{X}_\mu(F, G) \neq \emptyset$ for every probability measure μ on $\beta(\mathbb{R}^d)$.*

Proof. Let $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ be a complete filtered probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ such that there exists an m-dimensional $\mathbb{F}$-Brownian motion $(B_t)_{0\le t\le T}$ on $\mathcal{P}_{\mathbb{F}}$. Assume that x_0 is an $\mathcal{F}_0$-measurable random variable such that $Px_0^{-1} = \mu$. By virtue of Lemma 3.8 of Chap. 2, there exist $\beta_T \otimes \beta(\mathbb{R}^d)$-measurable selectors f and g of F and G, respectively, such that $\int_0^t f(\tau, \cdot)\mathrm{d}\tau$ and $\int_0^t g(\tau\ \cdot)\mathrm{d}\tau$ are continuous on $\mathbb{R}^d$ for every $t \in [0, T]$. Define for every $k = 1, 2, \dots$ a continuous process $(x_t^k)_{0\le t\le T}$ by setting

$$x_t^k = \begin{cases} x_0 \ \text{ a.s. for } \ -\frac{T}{k} \le t \le 0, \\ x_0 + \int_0^t f(\tau, x^k_{\tau-\frac{T}{k}})\mathrm{d}\tau + \int_0^t g(\tau, x^k_{\tau-\frac{T}{k}})\mathrm{d}B_\tau \\ \text{a.s. for } \ t \in [0, T]. \end{cases} \tag{1.2}$$

It is clear that x^k is continuous and $\mathbb{F}$-adapted for every $k = 1, 2, \dots,$. it follows immediately from (1.2) that $P(\{|x_0^k| > N\}) = P(\{|x_0| > N\})$ for every $k \ge 1$ and $N \ge 1$. Then $\lim_{N\to\infty} \sup_{k\ge 1} P(\{|x_0^k| > N\}) = \lim_{N\to\infty} P(\{|x_0| > N\}) = 0$. For every λ and $k \ge 1$, we get

$$P\left(\{|x_t^k - x_s^k| > \lambda\}\right) \le P\left(\left\{\left|\int_s^t f(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}\tau\right| > \lambda\right\}\right) \\ + P\left(\left\{\left|\int_s^t g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right| > \lambda\right\}\right).$$

By Chebyshev's inequality, it follows that

$$P\left(\left\{\left|\int_s^t f(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}\tau\right| > \lambda\right\}\right) \le \frac{1}{\lambda^4} E\left[\left|\int_s^t f(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}\tau\right|^4\right]$$
$$\le \frac{T^2}{\lambda^4}\left(\int_s^t K^2(t)\mathrm{d}t\right)^2,$$

where $K \in \mathbb{L}^2([0,T], \mathbb{R}^+)$ is such that $\max(\|F(t,x)\|, \|G(t,x)\|) \le K(t)$ for a.e. $t \in [0,T]$ and every $x \in \mathbb{R}^d$. Similarly, we obtain

$$P\left(\left\{\left|\int_s^t g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right| > \lambda\right\}\right) \le \frac{1}{\lambda^4} E\left[\left|\int_s^t g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right|^4\right].$$

By the definition of $\int_s^t g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau$, one has

$$\left|\int_s^t g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right| = \max_{1\le i\le d}\left|\sum_{j=1}^m \int_s^t g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|$$
$$\le \max_{1\le i\le d}\sum_{j=1}^m \left|\int_s^t g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|$$
$$\le \sum_{j=1}^m \left[\max_{1\le i\le d, 1\mathrm{j}\le m}\left|\int_s^t g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|\right]$$
$$= m \cdot \max_{1\le i\le d, 1\le j\le m}\left|\int_s^t g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|.$$

Then

$$\left|\int_s^t g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right|^4 \le m^4 \cdot \max_{1\le i\le d, 1\le j\le m}\left|\int_s^t g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|^4.$$

By Itô's formula, we obtain

$$E\left[\left|\int_s^{t\wedge\tau_N} g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|^4\right]$$
$$= 6\,E\left[\int_s^{t\wedge\tau_N}\left(\left|\int_s^\tau g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|^2 \cdot \left|g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\right|^2\right)\mathrm{d}\tau\right]$$
$$\le 6\,E\left[\int_s^t\left(\left|\int_s^\tau g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|^2 \cdot K^2(\tau)\right)\mathrm{d}\tau\right]$$

$$= 6\int_s^t \left[K^2(\tau)\cdot E\left|\int_s^\tau g^{ij}(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B^j_\tau\right|^2\right]\mathrm{d}\tau \le 6\left(\int_s^t K^2(t)\mathrm{d}t\right)^2,$$

for every $1 \le i \le d$ and $1 \le j \le m$, where

$$\tau_N = \inf\left\{t > 0 : \sup_{s\le\tau\le t}\left|\int_s^\tau g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right| \ge N\right\} \wedge T.$$

Then

$$E\left[\left|\int_s^{t\wedge\tau_N} g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right|^4\right] \le 6m^4\left(\int_s^t K^2(t)\mathrm{d}t\right)^2$$

for every $N \ge 1$, which implies that

$$E\left[\left|\int_s^{t} g(\tau, x^k_{\tau-\frac{1}{k}})\mathrm{d}B_\tau\right|^4\right] \le 6m^4\left(\int_s^t K^2(t)\mathrm{d}t\right)^2.$$

Hence it follows that

$$P\left(\{|x^k_t - x^k_s| > \lambda\}\right) \le \frac{T^2}{\lambda^4}\left(\int_s^t K^2(t)\mathrm{d}t\right)^2 + \frac{6m^4}{\lambda^4}\left(\int_s^t K^2(t)\mathrm{d}t\right)^2$$
$$\le \frac{1}{\lambda^4}|\Gamma(t) - \Gamma(s)|^2$$

for $s, t \in [0, T]$, where

$$\Gamma(t) = \sqrt{T^2 + 6m^4}\int_0^t K^2(\tau)\mathrm{d}\tau \quad \text{for} \quad 0 \le t \le T.$$

This, by virtue of Theorem 3.6 of Chap. 1, Theorem 2.2 of Chap. 1, and Theorem 2.3 of Chap. 1, implies that there exist an increasing sequence $(k_r)_{r=1}^\infty$ of positive integers, a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, and d-dimensional continuous stochastic processes $\tilde{x}$ and $\tilde{x}^{k_r}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ for $r = 1, 2, \ldots$, such that $P(x^{k_r})^{-1} = P(\tilde{x}^{k_r})^{-1}$ for $1, 2, \ldots$ and $\sup_{0\le t\le T}|\tilde{x}^{k_r}_t - \tilde{x}_t| \to 0$ with $(\tilde{P}.1)$ as $r \to \infty$. By Corollary 3.3 of Chap. 1, it follows that $P\tilde{x}_0^{-1} = \mu$, because $P(x_0^{k_r})^{-1} = \mu$ for $r = 1, 2, \ldots$ and $P(x_0^{k_r})^{-1} \Rightarrow P\tilde{x}_0^{-1}$ as $r \to \infty$. Let $\tilde{\mathbb{F}}$ be a filtration defined by a process $\tilde{x}$. Similarly as in the proof of Theorem 1.3, immediately from (1.2), it follows that $\mathbb{L}^{x^{k_r}}_{fg}$ generates on $C_b^2(\mathbb{R}^d)$ a family of continuous local $\mathbb{F}$-martingales for every $r = 1, 2, \ldots$, i.e., that $\mathcal{M}^{x^{k_r}}_{FG} \ne \emptyset$ for every $r = 1, 2, \ldots$, which by Lemma 1.3, implies that $\mathcal{M}^{\tilde{x}}_{FG} \ne \emptyset$. Thus there exist a filtered probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$ and a continuous $\tilde{\mathbb{F}}$-adapted process $\tilde{x}$ such that $P\tilde{x}_0^{-1} = \mu$ and $\mathcal{M}^{\tilde{x}}_{FG} \ne \emptyset$. Therefore, by virtue of Theorem 1.3, for every probability measure μ on $\beta(\mathbb{R}^d)$, one has $\mathcal{X}_\mu(F, G) \ne \emptyset$. □

Remark 1.6. If the assumptions of Theorem 1.5 are satisfied without the convexity of values of F, then $\mathcal{X}_\mu^0(\overline{F}, G) \neq \emptyset$.

Proof. By Lemma 1.7 of Chap. 3, one has $\mathcal{X}_\mu^0(\overline{F}, G) = \mathcal{X}_\mu^0(\mathrm{co}\, F, G)$. Similarly as in the proof of Theorem 1.5, by virtue of Theorem 1.4, one gets $\mathcal{X}_\mu^0(\mathrm{co}\, F, G) \neq \emptyset$. Then $\mathcal{X}_\mu^0(\overline{F}, G) \neq \emptyset$. □

2 Stochastic Differential Inclusions

Assume that $F : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ satisfy conditions $(\mathcal{H})$. By stochastic differential inclusions $SDI(F,G)$ and $\overline{SDI}(F,G)$, we mean relations of the form

$$x_t - x_s \in \int_s^t F(\tau, x_\tau)\mathrm{d}\tau + \int_s^t G(\tau, x_\tau)\mathrm{d}B_\tau\,, \quad \text{a.s.} \tag{2.1}$$

and

$$x_t - x_s \in \mathrm{cl}\left(\int_s^t F(\tau, x_\tau)\mathrm{d}\tau + \int_s^t G(\tau, x_\tau)\mathrm{d}B_\tau\right)\,, \quad \text{a.s.,} \tag{2.2}$$

which have to be satisfied for every $0 \leq s \leq t \leq T$ by a system $(\mathcal{P}_\mathbb{F}, x, B)$ consisting of a complete filtered probability space $\mathcal{P}_\mathbb{F}$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ satisfying the usual conditions, a d-dimensional $\mathbb{F}$-adapted continuous stochastic process $x = (x_t)_{0\leq t\leq T}$, and an m-dimensional $\mathbb{F}$-Brownian motion $B = (B_t)_{0\leq t\leq T}$ on $\mathcal{P}_\mathbb{F}$, where $\int_s^t F(\tau, x_\tau)\mathrm{d}\tau$ and $\int_s^t G(\tau, x_\tau)\mathrm{d}B_\tau$ denote Aumann and Itô set-valued integrals of set-valued processes $F \circ x = (F(t, x_t))_{0\leq t\leq T}$ and $G \circ x = (G(t, x_t))_{0\leq t\leq T}$, respectively. Similarly as above, systems $(\mathcal{P}_\mathbb{F}, x, P)$ are said to be weak solutions of $SDI(F,G)$ and $\overline{SDI}(F,G)$, respectively. If μ is a given probability measure on $\beta(\mathbb{R}^d)$, then a system $(\mathcal{P}_\mathbb{F}, x, B)$ is said to be a weak solution of the initial value problems $SDI(F,G,\mu)$ or $\overline{SDI}(F,G,\mu)$, if it satisfies conditions (2.1) or (2.2) and $Px_0^{-1} = \mu$. If apart from the set-valued mappings F and G, we are also given a filtered probability space $\mathcal{P}_\mathbb{F}$ and an m-dimensional $\mathbb{F}$-Brownian motion B on $\mathcal{P}_\mathbb{F}$, then a continuous $\mathbb{F}$-adapted process X such that the system $(\mathcal{P}_\mathbb{F}, X, B)$ satisfies (2.1) or (2.2) is said to be a strong solution of $SDI(F,G)$ or $\overline{SDI}(F,G)$, respectively.

Corollary 2.1. *For every measurable set-valued mappings $F : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ every weak (strong) solution of $\overline{SFI}(F,G)$ is a weak (strong) solution of $\overline{SDI}(F,G)$.*

Proof. If $(\mathcal{P}_\mathbb{F}, x, B)$ is a weak solution of $\overline{SFI}(F,G)$, then $S_\mathbb{F}(F \circ x) \neq \emptyset$ and $S_\mathbb{F}(G \circ x) \neq \emptyset$. A set $\mathrm{cl}_\mathbb{L}\{J_{st}[S_\mathbb{F}(F \circ x)] + \mathcal{J}_{st}[S_\mathbb{F}(G \circ x)]\}$ is a subset of $\mathrm{cl}_\mathbb{L}\{\overline{\mathrm{dec}}\{J_{st}[S_\mathbb{F}(F \circ x)]\} + \overline{\mathrm{dec}}\{\mathcal{J}_{st}[S_\mathbb{F}(G \circ x)]\}\}$ for every $0 \leq s \leq t \leq T$ and every continuous $\mathbb{F}$-adapted d-dimensional stochastic process $x = (x_t)_{0\leq t\leq T}$. From this and Theorem 2.1 of Chap. 3, it follows that every weak solution of $\overline{SFI}(F,G)$ is a weak solution of $\overline{SDI}(F,G)$. In a similar way, the above result for strong solutions can be obtained. □

Corollary 2.2. *For set-valued mappings $F : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ satisfying conditions ($\mathcal{H}$), every weak (strong) solution of $SFI(F,G)$ is a weak (strong) solution of $SDI(F,G)$.*

Proof. By (iv) of Theorem 2.1 of Chap. 3, a system $(\mathcal{P}_{\mathbb{F}}, x, B)$ is a weak solution of $\overline{SDI}(F,G)$ if and only if $x_t - x_s \in \overline{\mathrm{dec}}\{J_{st}[S_{\mathbb{F}}(F \circ x)]\} + \overline{\mathrm{dec}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(G \circ x)]\}$ for every $0 \le s < t \le T$. But $J_{st}[S_{\mathbb{F}}(F \circ x)] + J_{st}[S_{\mathbb{F}}(G \circ x)] \subset \overline{\mathrm{dec}}\{J_{st}[S_{\mathbb{F}}(F \circ x)]\} + \overline{\mathrm{dec}}\{\mathcal{J}_{st}[S_{\mathbb{F}}(G \circ x)]\}$ for every $0 \le s < t \le T$. Then every weak solution of $SFI(F,G)$ is a weak solution of $\overline{SDI}(F,G)$. But for every F and G satisfying conditions ($\mathcal{H}$), a stochastic differential inclusion $\overline{SDI}(F,G)$ is reduced to the form $SDI(F,G)$, because in this case, $\int_s^t F(\tau, x_\tau)\mathrm{d}\tau + \int_s^t G(\tau, x_\tau)\mathrm{d}B_\tau$ is a closed subset of $\mathbb{R}^d$. Therefore, every weak solution of $SFI(F,G)$ is a weak solution of $SDI(F,G)$. In a similar way, the above result for strong solutions of the above inclusions can be obtained. □

It is natural to expect that for every strong solution $(\mathcal{P}_{\mathbb{F}}, x, B)$ of $SDI(F,G)$ and every $\varepsilon > 0$, there exist a partition $(A_k)_{k=1}^N \in \Pi(\Omega, \mathcal{F}_T)$ and a family $(\mathcal{P}_{\mathbb{F}}, x^k, B)_{k=1}^N$ of strong solutions of $SFI(F,G)$ such that $\|(x_t - x_s) - \sum_{k=1}^N \mathbb{1}_{A_k}(x_t^k - x_s^k)\| \le \varepsilon$ for every $0 \le s < t \le T$, where $\|\cdot\|$ is the norm of $\mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$. It seems that the proof of such a result depends in an essential way on the $\mathbb{L}^2$-continuity of the mapping $[0,T] \ni t \to x_t \in \mathbb{L}^2(\Omega, \mathcal{F}_t, \mathbb{R}^d)$. By the definition of solutions of $SDI(F,G)$, it follows that the mapping $[0,T] \ni t \to x_t(\omega) \in \mathbb{R}^d$ is continuous for a.e. $\omega \in \Omega$. Therefore, a family $(x_t)_{0\le t\le T}$ of random variables $x_t : \Omega \to \mathbb{R}^d$ has to be uniformly square integrably bounded. But this depends, among other things, on the uniform square integrable boundedness of $(\int_0^t G(\tau, x_\tau)\mathrm{d}B_\tau)_{0\le t\le T}$. From the properties of set-valued integrals $\int_0^t G(\tau, x_\tau)\mathrm{d}B_\tau$, it follows that such a property of the family $(\int_0^t G(\tau, x_\tau)\mathrm{d}B_\tau)_{0\le t\le T}$ is difficult to obtain. Therefore, the desired above property is difficult to obtain. We can prove the following theorem.

Theorem 2.1. *Let $B = (B_t)_{t\ge 0}$ be an m-dimensional $\mathbb{F}$-Brownian motion on a filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with a filtration $\mathbb{F}$ satisfying the usual conditions and Hölder continuous with exponential $\alpha = 3$. Assume that $F : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ are measurable, uniformly square integrably bounded, and Lipschitz continuous with respect to the second variable for every fixed $t \in [0,T]$ with a Lipschitz function $k \in \mathbb{L}^2([0,T], \mathbb{R})$. Then for every $\varepsilon > 0$ and every strong solution x of $\overline{SDI}(F,G)$, there exist a number $\lambda_\varepsilon > 0$ and a strong $\varepsilon\lambda_\varepsilon$-approximating solution x^ε of $SFI(F,G)$ such that $\sup_{0\le t\le T} \|x_t - x_t^\varepsilon\| \le \varepsilon\lambda_\varepsilon$, i.e., there exists a continuous $\mathbb{F}$-adapted stochastic process x^ε such that $x_t^\varepsilon - x_s^\varepsilon \in \{J_{st}[S_{\mathbb{F}}(F \circ x^\varepsilon)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G \circ x^\varepsilon)]\} + \varepsilon\lambda_\varepsilon\mathcal{B}$ for every $0 \le s < t \le T$, where $\mathcal{B}$ denotes the closed unit ball of $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{R}^d)$.*

Proof. Let $x = (x_t)_{0\le t\le T}$ be a strong solution of $\overline{SDI}(F,G)$ and $\varepsilon > 0$. By virtue of Remark 2.3 of Chap. 3, for $\bar{\varepsilon} = \varepsilon/L(1+\sqrt{T})$ there exist a number $\lambda_\varepsilon = 1 + m_\varepsilon\beta\left[3\sqrt{6}d(T+2\delta_\varepsilon) + T + \delta_\varepsilon^3\sqrt{\delta_\varepsilon}\right]$ and processes $f^\varepsilon \in S_{\mathbb{F}}(F\circ x)$ and $g^\varepsilon \in S_{\mathbb{F}}(G\circ x)$ such that $\sup_{0\le t\le T}\|x_t - x_t^\varepsilon\| \le \lambda_\varepsilon\varepsilon/L(1+\sqrt{T})$, where $L^2 = \int_0^T k_t^2\mathrm{d}t$ and $x_t^\varepsilon = x_0 + \int_0^t f_\tau^\varepsilon\mathrm{d}\tau + \int_0^t g_\tau^\varepsilon\mathrm{d}B_\tau$ a.s. for $0\le t\le T$. Hence in particular, it follows that $x_t^\varepsilon - x_s^\varepsilon \in J_{st}[S_{\mathbb{F}}(F\circ x)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)]$ for every $0\le s<t\le T$. Similarly as in the proof of Remark 4.1 of Chap. 2, we obtain

$$
\begin{aligned}
&H\,(\mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F\circ x)]+\mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)]\}, \mathrm{cl}_{\mathbb{L}}\{J_{st}[S_{\mathbb{F}}(F\circ x^\varepsilon)]+\mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x^\varepsilon)]\})\\
&\quad = H\,(J_{st}[S_{\mathbb{F}}(F\circ x)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)], J_{st}[S_{\mathbb{F}}(F\circ x^\varepsilon)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x^\varepsilon)])\\
&\quad \le L(1+\sqrt{T})\sup_{0\le t\le T}\|x_t - x_t^\varepsilon\|
\end{aligned}
$$

for every $0\le s\le t\le T$. Therefore, for every $0\le s\le t\le T$, we get

$$
\begin{aligned}
&\mathrm{dist}\left(x_t^\varepsilon - x_s^\varepsilon, J_{st}[S_{\mathbb{F}}(F\circ x^\varepsilon)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x^\varepsilon)]\right)\\
&\quad \le H\,(J_{st}[S_{\mathbb{F}}(F\circ x] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x)], J_{st}[S_{\mathbb{F}}(F\circ x^\varepsilon)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x^\varepsilon)])\\
&\quad \le L(1+\sqrt{T})\sup_{0\le t\le T}\|x_t - x_t^\varepsilon\|.
\end{aligned}
$$

Then $x_t^\varepsilon - x_s^\varepsilon \in \{J_{st}[S_{\mathbb{F}}(F\circ x^\varepsilon)] + \mathcal{J}_{st}[S_{\mathbb{F}}(G\circ x^\varepsilon)]\} + \varepsilon\lambda_\varepsilon\mathcal{B}$ for every $0\le s<t\le T$, where $\mathcal{B}$ denotes the closed unit ball of $\mathbb{L}^2(\Omega,\mathcal{F},\mathbb{R}^d)$. □

Remark 2.1. It is difficult to obtain better properties of $SDI(F,G)$, because up to now, we have not been able to prove that the uniform integrable boundedness of G and continuity of $G(t,\cdot)$ imply the integrable boundedness and continuity of the Itô integral $\int_0^T G(t,\cdot)\mathrm{d}B_t$. □

3 Backward Stochastic Differential Inclusions

We shall consider now a special case of stochastic differential inclusions. They are written as relations of the form $x_s \in E[x_t + \int_s^t F(\tau,x_\tau)\mathrm{d}\tau|\mathcal{F}_s]$ a.s., where $F:[0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ is a given measurable set-valued mapping and $E[x_t + \int_s^t F(\tau,x_\tau)\mathrm{d}\tau|\mathcal{F}_s]$ denotes the set-valued conditional expectation of $x_t + \int_s^t F(\tau,x_\tau)\mathrm{d}\tau$. Such relations are considered together with a terminal condition $x_T \in H(x_T)$ a.s. for a given set-valued mapping $H:\mathbb{R}^d\to\mathrm{Cl}(\mathbb{R}^d)$. In what follows, the terminal problem presented above will be denoted by $BSDI(F,H)$ and called a backward stochastic differential inclusion. By a weak solution of $BSDI(F,H)$, we mean a system $(\mathcal{P}_{\mathbb{F}},x)$ consisting of a complete filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ satisfying the usual conditions and a càdlàg d-dimensional stochastic process $x = (x_t)_{0\le t\le T}$ such that the following conditions are satisfied:

$$\begin{cases} x_s \in E[x_t + \int_s^t F(\tau, x_\tau)\mathrm{d}\tau|\mathcal{F}_s] \ \text{ a.s. for } 0 \le s < t \le T, \\ x_T \in H(x_T) \text{ a.s.} \end{cases} \tag{3.1}$$

Similarly as in the theory of stochastic differential inclusions, we can consider the terminal problem $BSDI(F, H)$ if apart from F and H, a filtered probability space $\mathcal{P}_{\mathbb{F}}$ is also given. In such a case, a d-dimensional càdlég process x on $\mathcal{P}_{\mathbb{F}}$ such that a system $(\mathcal{P}_{\mathbb{F}}, x)$ satisfies (3.1) is said of be a strong solution of $BSDI(F, H)$ on $\mathcal{P}_{\mathbb{F}}$. It is clear that if x is a strong solution of $BSDI(F, H)$ on $\mathcal{P}_{\mathbb{F}}$, then the pair $(\mathcal{P}_{\mathbb{F}}, x)$ is a weak solution. The set of all weak solutions of $BSDI(F, H)$ is denoted by $\mathcal{B}(F, H)$, and a subset containing all $(\mathcal{P}_{\mathbb{F}}, x) \in \mathcal{B}(F, H)$ with a continuous process x is denoted by $\mathcal{CB}(F, H)$. We obtain the following result immediately from Theorem 3.1 of Chap. 3.

Corollary 3.1. *If* $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ *and* $H : \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ *are measurable and uniformly integrably bounded, then* $(\mathcal{P}_{\mathbb{F}}, x) \in \mathcal{B}(F, H)$ *if and only if* $x_T \in H(x_T)$ *a.s. and there exists* $f \in S(\mathrm{co}\, F \circ x)$, *a measurable selector of* $\mathrm{co}\, F \circ x$, *such that* $x_t = E[x_T + \int_t^T f_\tau \mathrm{d}\tau|\mathcal{F}_t]$ *a.s. for every* $0 \le t \le T$. □

Backward stochastic differential inclusions can be regarded as generalizations of backward stochastic differential equations:

$$x_t = E\left[h(x) + \int_t^T f(\tau, x_\tau, z_\tau)\mathrm{d}\tau|\mathcal{F}_t\right] \text{ a.s.}, \tag{3.2}$$

where the triplet (h, f, z) is called the data set of such an equation. Usually, if we consider strong solutions of (3.2) apart from (h, f, z), a probability space $\mathcal{P} = (\Omega, \mathcal{F}, P)$ is also given, and the filtration $\mathbb{F}^z$ is defined to be the smallest filtration satisfying the usual conditions and such that the process z is $\mathbb{F}^z$-adapted. The process z is called the driving process. In practical applications, the driving process z is taken as a d-dimensional Brownian motion or a strong solution of a forward stochastic differential equation. In the case of weak solutions of (3.2) apart from h and f, a probability measure μ on the space $\mathcal{D}_T(\mathbb{R}^d)$ of d-dimensional càdlàg functions on $[0, T]$ is also given, a weak solution of which with an initial distribution μ is defined as a system $(\mathcal{P}_{\mathbb{F}}, x, z)$ satisfying (3.2) and $Pz^{-1} = \mu$, and such that every $\mathbb{F}^z$-martingale is also an $\mathbb{F}$-martingale. Let us observe that in a particular case, for a given weak solution $(\mathcal{P}_{\mathbb{F}}, x)$ of $BSDI(F, H)$ with $H(x) = \{h(x)\}$ and $F(t, x) = \{f(t, x, z) : z \in \mathcal{Z}\}$ for $(t, x) \in [0, T] \times \mathbb{R}^m$, where f and h are given measurable functions and $\mathcal{Z}$ is a nonempty compact subset of the space $\mathcal{D}_T(\mathbb{R}^d)$, there exists a measurable $\mathbb{F}$-adapted stochastic process $(z_t)_{0\le t\le T}$ with values in $\mathcal{Z}$ such that

$$x_t = E\left[h(x) + \int_t^T f(\tau, x_\tau, z_\tau)\mathrm{d}\tau|\mathcal{F}_t\right] \text{ a.s.} \tag{3.3}$$

For given probability measures μ_0 and μ_T on $\mathbb{R}^d$, we can look for a weak solution $(\mathcal{P}_{\mathbb{F}}, x)$ for $BSDI(F, H)$ such that $Px_0^{-1} = \mu_0$ and $Px_T^{-1} = \mu_T$. If F and H are as above, then there exists a measurable and $\mathbb{F}$-adapted stochastic process $(z_t)_{0\leq t\leq T}$ such that (3.3) is satisfied and such that $E[h(x) + \int_0^T f(\tau, x_\tau, z_\tau)\mathrm{d}\tau] = \int_{\mathbb{R}^d} u\mathrm{d}\mu_0$. If $f(t, x, z) = f(t, x) + g(z)$ with $g \in C(\mathcal{D}_T(\mathbb{R}^d), \mathbb{R}^d)$, then

$$\int_0^T \int_{\mathcal{D}_T(\mathbb{R}^d)} g(v)\mathrm{d}\lambda_\tau \mathrm{d}\tau = \int_{\mathbb{R}^d} u\mathrm{d}\mu_0 - \int_{\mathbb{R}^d} h(u)\mathrm{d}\mu_T - E\int_0^T f(\tau, x_\tau)\mathrm{d}\tau,$$

where $\lambda_t = Pz_t^{-1}$ for $t \in [0, T]$.

In some special cases, weak solutions of $BSDI(F, H)$ describe a class of recursive utilities under uncertainty. To verify this, suppose $(\mathcal{P}_{\mathbb{F}}, x)$ is a weak solution of $BSDI(F, H)$ with $H(x) = \{h(x)\}$ and $F(t, x) = \{f(t, x, c, z) : (c, z) \in \mathcal{C} \times \mathcal{Z}\}$, where h and f are measurable functions and $\mathcal{C}$, $\mathcal{Z}$ are nonempty compact subsets of $C([0, T], \mathbb{R}^+)$ and $\mathcal{D}_T(\mathbb{R}^d)$, respectively. Similarly as above, we can find a pair of measurable $\mathbb{F}$-adapted stochastic processes $(c_t)_{0\leq t\leq T}$ and $(z_t)_{0\leq t\leq T}$ with values in $\mathcal{C}$ and $\mathcal{Z}$, respectively, such that

$$x_t = E\left[h(x) + \int_t^T f(\tau, x_\tau, c_\tau, z_\tau)\mathrm{d}\tau | \mathcal{F}_t\right] \quad \text{a.s.} \tag{3.4}$$

for $0 \leq t \leq T$. In such a case, (3.4) describes a certain class of recursive utilities under uncertainty, where $(c_t(s, \cdot))_{0\leq s\leq T}$ denotes for fixed $t \in [0, T]$ the future consumption. Let us observe that in some special cases, a strong solution x of $BSDI(F, H)$ on a filtered probability space $\mathcal{P}_{\mathbb{F}}$ with the "constant" filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$, i.e., such that $\mathcal{F}_t = \mathcal{F}$ for $0 \leq t \leq T$, is a solution of a backward random differential inclusion $-x'_t \in \overline{\mathrm{co}}\, F(t, x_t)$ with a terminal condition $x_T \in H(x_T)$ that has to be satisfied a.s. for a.e. $t \in [0, T]$.

Throughout this section, we assume that $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ is a complete filtered probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ satisfying the usual hypotheses, and by $\mathbb{D}(\mathbb{F}, \mathbb{R}^d)$ and $\mathbf{C}(\mathbb{F}, \mathbb{R}^d)$, we denote the spaces of all d-dimensional $\mathbb{F}$-adapted càdlàg and continuous, respectively, processes X on $\mathcal{P}_{\mathbb{F}}$ such that $\|X\|^2 = E[\sup_{s\in[0,T]}|X_s|^2] < \infty$. Similarly as above, we denote by $\mathcal{S}(\mathbb{F}, \mathbb{R}^d)$ the set of all d-dimensional $\mathbb{F}$-semimartingales X on $\mathcal{P}_{\mathbb{F}}$ such that $\|X\|^2 = E[\sup_{s\in[0,T]}|X_s|^2] < \infty$. We have $\mathbf{C}(\mathbb{F}, \mathbb{R}^d) \subset \mathbb{D}(\mathbb{F}, \mathbb{R}^d)$ and $\mathcal{S}(\mathbb{F}, \mathbb{R}^d) \subset \mathbb{D}(\mathbb{F}, \mathbb{R}^d)$. It can be proved that $(\mathcal{S}(\mathbb{F}, \mathbb{R}^d), \|\cdot\|)$ is a Banach space. In what follows, we shall assume that $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $H : \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ satisfy the following conditions $(\mathcal{A})$:

(i) F is measurable and uniformly square integrably bounded;
(ii) H is measurable and bounded;
(iii) $F(t, \cdot)$ is Lipschitz continuous for a.e. fixed $t \in [0, T]$;
(iv) there is a random variable $\xi \in \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ such that $\xi \in H(\xi)$ a.s.

We shall prove that conditions $(\mathcal{A})$ are sufficient for the existence of strong solutions for $BSDI(F, H)$, which implies that $\mathcal{B}(F, H)$ is nonempty. It is natural to look for weaker conditions implying the nonemptiness of $\mathcal{B}(F,H)$. The problem is quite complicated. It needs new sufficient conditions for tightness of sets of probability measures. We do not consider it in this book.

Lemma 3.1. *Let $F : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $H : \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ satisfy conditions $(\mathcal{A})$. For every filtered probability space $\mathcal{P}_{\mathbb{F}} = (\Omega,\mathcal{F},\mathbb{F},P)$ and a random variable $\xi : \Omega \to \mathbb{R}^d$, there exists a sequence $(x^n)_{n=0}^{\infty}$ of $\mathcal{S}(\mathbb{F},\mathbb{R}^d)$ defined by $x_t^n = E[\xi + \int_t^T f_\tau^{n-1}\mathrm{d}\tau|\mathcal{F}_t]$ a.s. and $0\le t\le T$ with $x^0\in\mathcal{S}(\mathbb{F},\mathbb{R}^d)$ satisfying $x_T^0=\xi$ a.s. and $f^{n-1}\in S_{\mathbb{F}}(\mathrm{co}F\circ x^{n-1})$ for $n=1,2,\dots$ such that*

$$E[\sup_{t\le u\le T}|x_u^{n+1}-x_u^n|^2]\le 4E\left[\int_t^T K(\tau)\sup_{\tau\le s\le T}|x_s^n-x_s^{n-1}|\mathrm{d}\tau\right]^2$$

for $n=1,2\dots$ and $0\le t\le T$, with $K(t)=K_d\cdot k(t)$ for $0\le t\le T$, where $k\in\mathbb{L}^2([0,T],\mathbb{R}^+)$ is a Lipschitz function of $F(t,\cdot)$ and K_d is the number defined in Remark 2.6 of Chap. 2.

Proof. Let $\mathcal{P}_{\mathbb{F}}$ be a filtered probability space and let $x^0=(x_t^0)_{0\le t\le T}\in\mathcal{S}(\mathbb{F},\mathbb{R}^d)$ be such that $x_T^0=\xi$ a.s. Put $f_t^0=s(\overline{\mathrm{co}}\,F(t,x_t^0))$ a.s. for $0\le t\le T$, where s is the Steiner point mapping defined by formula (2.1) of Chap. 2. It is clear that $f^0\in S_{\mathbb{F}}(\overline{\mathrm{co}}\,F\circ x^0)$, because by virtue of Corollary 2.2 of Chap. 2, the function $s(\overline{\mathrm{co}}\,F(t,\cdot))$ is Lipschitz continuous for a.e. fixed $0\le t\le T$, and x^0 is $\mathbb{F}$-adapted. We now define a sequence $(x^n)_{n=1}^{\infty}$ by the successive approximation procedure, i.e., by taking $x_t^n = E[\xi+\int_t^T f_\tau^{n-1}\mathrm{d}\tau|\mathcal{F}_t]$ a.s. for $n=1,2,\dots$ and $0\le t\le T$, where $f_t^{n-1}=s(\overline{\mathrm{co}}\,F(t,x_t^{n-1}))$ a.s. for $0\le t\le T$. Similarly as above, we have $f^{n-1}\in S_{\mathbb{F}}(\overline{\mathrm{co}}\,F\circ x^{n-1})$. By Corollary 3.2 of Chap. 3, we have $x^n\in\mathcal{S}(\mathbb{F},\mathbb{R}^d)$. Immediately from the above definitions and Corollary 2.2 of Chap. 2, it follows that $|f_t^n-f_t^{n-1}|\le K(t)\sup_{t\le s\le T}|x_s^n-x_s^{n-1}|$ a.s. for a.e. $0\le t\le T$ and $n=1,2,\dots$. Hence it follows that

$$|x_t^{n+1}-x_t^n|\le E\left[\int_t^T|f_\tau^n-f_\tau^{n-1}|\mathrm{d}\tau|\mathcal{F}_t\right]\le E\left[\int_t^T K(\tau)\sup_{\tau\le s\le T}|x_s^n-x_s^{n-1}|\mathrm{d}\tau|\mathcal{F}_t\right]$$

a.s. for $0\le t\le T$. Therefore,

$$\begin{aligned}\sup_{t\le u\le T}|x_u^{n+1}-x_u^n| &\le \sup_{t\le u\le T}E\left[\int_u^T K(\tau)\sup_{\tau\le s\le T}|x_s^n-x_s^{n-1}|\mathrm{d}\tau|\mathcal{F}_u\right]\\ &\le \sup_{t\le u\le T}E\left[\int_t^T K(\tau)\sup_{\tau\le s\le T}|x_s^n-x_s^{n-1}|\mathrm{d}\tau|\mathcal{F}_u\right]\end{aligned}$$

a.s. for $0\le t\le T$ and $n=1,2,\dots$. By Doob's inequality, we obtain

$$E\left(\sup_{t\le u\le T}E\left[\int_t^T K(\tau)\sup_{\tau\le s\le T}|x_s^n-x_s^{n-1}|\mathrm{d}\tau|\mathcal{F}_u\right]\right)^2\le 4E\left(\int_t^T K(\tau)\sup_{\tau\le s\le T}|x_s^n-x_s^{n-1}|\mathrm{d}\tau\right)^2$$

for $0 \leq t \leq T$. Therefore, for every $n = 1, 2, \ldots$ and $0 \leq t \leq T$, we have

$$E\left[\sup_{t\leq u\leq T} |x_u^{n+1} - x_u^n|^2\right] \leq 4E\left(\int_t^T K(\tau) \sup_{\tau\leq t\leq T} |x_s^n - x_s^{n-1}|\mathrm{d}\tau\right)^2. \qquad \square$$

We obtain the following result immediately from the properties of multivalued conditional expectations.

Lemma 3.2. *If F satisfies conditions $(\mathcal{A})$, then for every $x, y \in \mathcal{S}(\mathbb{F}, \mathbb{R}^d)$, one has*

$$E\left[h\left(E\left[\int_s^t F(\tau, x_\tau)\,\mathrm{d}\tau|\mathcal{F}_s\right], E\left[\int_s^t F(\tau, y_\tau)\mathrm{d}\tau|\mathcal{F}_s\right]\right)\right] \leq \int_s^t k(\tau)E|x_\tau - y_\tau|\mathrm{d}\tau$$

for every $0 \leq s \leq t \leq T$, where h is the Hausdorff metric on $\mathrm{Cl}(\mathbb{R}^d)$.

We can now prove the following existence theorem.

Theorem 3.1. *If $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $H : \mathbb{R}^m \to \mathrm{Cl}(\mathbb{R}^m)$ satisfy conditions $(\mathcal{A})$, then for every complete filtered probability space $\mathcal{P}_{\mathbb{F}}$ and fixed point ξ of H, there exists a strong solution of (3.1).*

Proof. Let $\mathcal{P}_{\mathbb{F}}$ be given and assume that $\xi \in \mathbb{L}^2(\Omega, \mathcal{F}_T, \mathbb{R}^d)$ is such that $\xi \in H(\xi)$. By virtue of Lemma 3.1, there exists a sequence $(x^n)_{n=1}^{\infty}$ of $\mathcal{S}(\mathbb{F}, \mathbb{R}^d)$ such that $x_T^n = \xi$, $x_s^n \in E[x_t^n + \int_s^t F(\tau, x_\tau^{n-1}\mathrm{d}\tau|\mathcal{F}_t]$ a.s. for $0 \leq s \leq t \leq T$ and

$$E\left[\sup_{t\leq u\leq T} |x_u^{n+1} - x_u^n|^2\right] \leq 4E\left(\int_t^T K(\tau) \sup_{\tau\leq s\leq T} |x_s^n - x_s^{n-1}|\mathrm{d}\tau\right)^2$$

for $n = 1, 2, \ldots$ and $0 \leq t \leq T$. By properties of F and H, one has $E[\sup_{t\leq u\leq T} |x_u^1 - x_u^0|^2] \leq L$, where $L = 4[E|\xi|^2 + \int_0^T m^2(\tau)\mathrm{d}\tau] + 2E[\sup_{0\leq t\leq T} |x_t^0|^2]$ with $m \in \mathbb{L}^2([0, T], \mathbb{R}^+)$ such that $\|F(t, x)\| \leq m(t)$ for every $x \in \mathbb{R}^d$ and a.e. $0 \leq t \leq T$. Therefore,

$$E\left[\sup_{t\leq u\leq T} |x_u^2 - x_u^1|^2\right] \leq 4TL\int_t^T K^2(\tau)\mathrm{d}\tau.$$

Hence it follows that

$$\begin{aligned} E\left[\sup_{t\leq u\leq T} |x_u^3 - x_u^2|^2\right] &\leq (4T)^2L\int_t^T \left(K^2(\tau)\int_\tau^T K^2(s)\mathrm{d}s\right)\mathrm{d}\tau \\ &= \frac{(4T)^2L}{2}\left(\int_t^T K^2(\tau)\mathrm{d}\tau\right)^2. \end{aligned}$$

By the inductive procedure, for every $n = 1, 2, \dots$ and $0 \le t \le T$, we get

$$E\left[\sup_{t\le u\le T} |x_u^{n+1} - x_u^n|^2\right] \le \frac{(4T)^n L^{n-1}}{n!}\left(\int_t^T K^2(\tau)\mathrm{d}\tau\right)^n.$$

Then $(x^n)_{n=1}^{\infty}$ is a Cauchy sequence of $\mathcal{S}(\mathbb{F}, \mathbb{R}^d)$. Therefore, there exists a process $(x_t)_{0\le t\le T} \in \mathcal{S}(\mathbb{F}, \mathbb{R}^d)$ such that $E[\sup_{0\le t\le T} |x_t^n - x_t|^2] \to 0$ as $n \to \infty$. By Lemma 3.2, it follows that

$$\begin{aligned}
&E \operatorname{dist}\left(x_s, E\left[x_t + \int_s^t F(\tau, x_\tau)\mathrm{d}\tau|\mathcal{F}_s\right]\right)\\
&\le E|x_s - x_s^n|] + E\left[\operatorname{dist}\left(x_s^n, E\left[x_t^n + \int_s^t F(\tau, x_\tau^{n-1})\mathrm{d}\tau|\mathcal{F}_s\right]\right)\right]\\
&\quad + E\left[h\left(E\left[x_t^n + \int_s^t F(\tau, x_\tau^{n-1})\mathrm{d}\tau|\mathcal{F}_s\right], E\left[x_t + \int_s^t F(\tau, x_\tau)\mathrm{d}\tau|\mathcal{F}_s\right]\right)\right]\\
&\le E|x_s^n - x_s| + E|x_t^n - x_t| + \int_s^t K(\tau)E|x_\tau^{n-1} - x_\tau|\mathrm{d}\tau\\
&\le 2\|x^n - x\| + \left(\int_0^T K^2(\tau)\mathrm{d}\tau\right)^{\frac{1}{2}} \|x^{n-1} - x\|
\end{aligned}$$

for every $0 \le s \le t \le T$ and $n = 1, 2, \dots$. Therefore, $\operatorname{dist}(x_s, E[x_t + \int_s^t F(\tau, x_\tau)\mathrm{d}\tau|\mathcal{F}_s]) = 0$ a.s. for every $0 \le s \le t \le T$, which implies that $x_s \in E\left[x_t + \int_s^t F(\tau, x_\tau)\mathrm{d}\tau|\mathcal{F}_s\right]$ a.s for every $0 \le s \le t \le T$. By the definition of $(x_t^n)_{0\le t\le T}$, we have $x_T^n = \xi \in H(\xi)$ a.s. for every $n = 1, 2, \dots$. Therefore, we also have $x_T = \xi$ a.s. Thus $x_T \in H(x_T)$ a.s. Then x satisfies (3.1). □

4 Weak Compactness of Solution Sets

For given measurable multifunctions $F : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$, $G : [0, T] \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ and a probability measure μ on $\beta(\mathbb{R}^d)$, by $\mathcal{X}_\mu(F, G)$ we denote, similarly as above, the set of all weak solutions (equivalence classes defined in Sect. 1) of $SFI(F, G, \mu)$. Elements $[(\mathcal{P}_{\mathbb{F}}, X, B)]$ of $\mathcal{X}_\mu(F, G)$ will be identified with equivalence classes $[X]$ of all d-dimensional continuous processes Z such that $PX^{-1} = PZ^{-1}$. In what follows, $[X]$ will be denoted simply by X. It is clear that we can associate with every $[(\mathcal{P}_{\mathbb{F}}, X, B)] \in \mathcal{X}_\mu(F, G)$ a probability measure PX^{-1}, a distribution of X, defined on a Borel σ-algebra $\beta(C_T)$ of the space $C_T =: C(([0, T], \mathbb{R}^d)$. The family of all such probability measures, corresponding to all classes belonging to $\mathcal{X}_\mu(F, G)$, is denoted by $\mathcal{X}_\mu^P(F, G)$. It is a subset of the space $\mathcal{M}(C_T)$ of probability measures on C_T. The set $\mathcal{X}_\mu(F, G)$ is said to be weakly compact, or weakly compact in distribution, if $\mathcal{X}_\mu^P(F,G)$ is a weakly compact subset of $\mathcal{M}(C_T)$. We now present sufficient conditions for the weak compactness of $\mathcal{X}_\mu(F,G)$.

Theorem 4.1. *Let $F : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : [0,T]\times\mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times m})$ be measurable, uniformly square integrably bounded, and convex-valued such that $F(t,\cdot)$ and $G(t,\cdot)$ are continuous for fixed $t \in [0,T]$. If G is diagonally convex, then for every probability measure μ on $\beta(\mathbb{R}^d)$, the set $\mathcal{X}_\mu(F,G)$ is nonempty and weakly compact.*

Proof. The nonemptiness of $\mathcal{X}_\mu(F,G)$ follows from Theorem 1.5. To show that $\mathcal{X}_\mu(F,G)$ is relatively weakly compact in the sense of distributions, let us note that by virtue of Theorem 1.5 of Chap. 3, for every $(\mathcal{P}_{\mathbb{F}},x,B)\in\mathcal{X}_\mu(F,G)$ there are $f\in S_{\mathbb{F}}(F\circ x)$ and $g\in S_{\mathbb{F}}(G\circ x)$ such that $Px_0^{-1}=\mu$ and $x_t = x_0+\int_0^t f_\tau \mathrm{d}\tau + \int_0^t g_\tau \mathrm{d}B_\tau$ for every $t\in[0,T]$. Similarly as in the proof of Theorem 1.5, we can verify that every sequence $(\mathcal{P}^n_{\mathbb{F}^n},x^n,B^n)_{n=1}^\infty$ of $\mathcal{X}_\mu(F,G)$ satisfies the conditions of Theorem 3.6 of Chap. 1. Therefore, for every sequence $(\mathcal{P}^n_{\mathbb{F}^n},x^n,B^n)$ of $\mathcal{X}_\mu(F,G)$, there exists an increasing subsequence $(n_k)_{k=1}^\infty$ of $(n)_{n=1}^\infty$ such that the sequence $\{P(x^{n_k})^{-1}\}_{n=1}^\infty$ is weakly convergent in distribution. Then the sequence $(x^n)_{n=1}^\infty$ possesses a subsequence converging in distribution.

Let $(x^r)_{r=1}^\infty$ be a sequence of $\mathcal{X}_\mu(F,G)$ convergent in distribution. Then there exists a probability measure $\mathcal{P}$ on $\beta(C_T)$ such that $P(x^r)^{-1}\Rightarrow\mathcal{P}$ as $r\to\infty$. By virtue of Theorem 2.3 of Chap. 1, there exist a probability space $(\tilde\Omega,\tilde{\mathcal{F}},\tilde P)$ and random variables $\tilde x^r : \tilde\Omega\to C_T$ and $\tilde x : \tilde\Omega \to C_T$ for $r=1,2,\dots$ such that $P(x^r)^{-1}=P(\tilde x^r)^{-1}$ for $r=1,2,\dots$, $\tilde P(\tilde x)^{-1}=\mathcal{P}$ and $\lim_{r\to\infty}\sup_{0\le t\le T}|\tilde x^r_t - \tilde x_t| = 0$ with $(\tilde P.1)$. Immediately from Corollary 3.3 of Chap. 1, it follows that $x_0^r\Rightarrow\tilde x_0$ as $r\to\infty$, because $P(x^r)^{-1}\Rightarrow P(\tilde x)^{-1}$ as $r\to\infty$. But $P(x_0^r)^{-1}=\mu$ for every $r\ge 1$. Then $P\tilde x_0^{-1}=\mu$. By Theorem 1.3, we have $\mathcal{M}_{FG}^{x_r}\ne\emptyset$ for every $r\ge1$, which by Lemma 1.3, implies that $\mathcal{M}_{FG}^{\tilde x}\ne\emptyset$. Therefore, by virtue of Theorem 1.3, there exist a standard extension $\hat{\mathcal{P}}_{\hat{\mathbb{F}}}=(\hat\Omega,\hat{\mathcal{F}},\hat{\mathbb{F}},\hat P)$ of $(\tilde\Omega,\tilde{\mathcal{F}},\tilde{\mathbb{F}},\tilde P)$ and an m-dimensional Brownian motion $\hat B$ such that $(\hat{\mathcal{P}}_{\hat{\mathbb{F}}},\hat x,\hat B)$, with $\hat x(\hat\omega)=\tilde x(\pi(\hat\omega))$ for every $\hat\omega\in\hat\Omega$, is a weak solution of $SFI(F,G,\mu)$, where $\pi:\hat\Omega\to\tilde\Omega$ is an $(\hat{\mathcal{F}},\tilde{\mathcal{F}})$-measurable mapping as described in the definition of the extension of $(\tilde\Omega,\tilde{\mathcal{F}},\tilde{\mathbb{F}},\tilde P)$, because its standard extension $\hat{\mathcal{P}}_{\hat{\mathbb{F}}}$ is also its extension. Let $\hat x^r(\hat\omega)=\tilde x^r(\pi(\hat\omega))$ for $\hat\omega\in\hat\Omega$. For every $A\in\beta(C)$, one has $P(\hat x^r)^{-1}(A) = \hat P[(\hat x^r)^{-1}(A)] = \hat P[(\tilde x^r\circ\pi)^{-1}(A)] = (\hat P\circ\pi^{-1})[(\tilde x^r)^{-1}(A)] = \tilde P[(\tilde x^r)^{-1}(A)] = P(\tilde x^r)^{-1}(A)$. Therefore, $P(\hat x^r)^{-1}=P(\tilde x^r)^{-1}=P(x^r)^{-1}$ for every $r\ge1$. By the properties of the sequence $(\tilde x^r)_{r=1}^\infty$, it follows that $\tilde x^r_t(\tilde\omega)\to\tilde x_t(\tilde\omega)$ with $(\tilde P.1)$ as $r\to\infty$ uniformly with respect to $0\le t\le T$. Hence in particular, it follows that $\tilde x^r_t(\pi(\hat\omega))\to\tilde x_t(\pi(\hat\omega))$ with $(\hat P.1)$ as $r\to\infty$ uniformly with respect to $0\le t\le T$. Therefore, for every $f\in C_b(C)$, one has $f(\hat x^r(\hat\omega))\to f(\hat x(\hat\omega))$ with $(\hat P.1)$ as $r\to\infty$. By the boundedness of $f\in C_b(C)$, this implies that $\hat E\{f(\hat x^r)\}\to\hat E\{f(\hat x)\}$ as $r\to\infty$, which by Corollary 2.1 of Chap. 1, is equivalent to $P(\hat{x}^r)^{-1} \Rightarrow P\hat{x}^{-1}$. But $P(\hat{x}^r)^{-1} = P(x^r)^{-1}$ for every $r \geq 1$. Then $x^r \Rightarrow \hat{x}$, which implies that $\mathcal{X}_\mu(F, G)$ is weakly closed. □

In a similar way, we can prove the following theorem.

Theorem 4.2. *Let $F : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^d)$ and $G : \mathbb{R}^+ \times \mathbb{R}^d \to \mathrm{Cl}(\mathbb{R}^{d\times d})$ be measurable and uniformly square integrably bounded such that $F(t, \cdot)$ and $G(t, \cdot)$ are continuous for fixed $t \in [0, T]$. If G is convex-valued and diagonally convex, then for every probability measure μ on $\beta(\mathbb{R}^d)$, the set $\mathcal{X}^0_\mu(\overline{F}, G)$ is nonempty and weakly compact.*

Proof. The nonemptiness of $\mathcal{X}^0_\mu(\overline{F}, G)$ follows from Remark 1.6. In a similar way as above, we can verify that the set $\mathcal{X}^0_\mu(\mathrm{co}\, F, G)$ of all weak solutions $(\mathcal{P}_\mathbb{F}, x, B)$ of $SFI(\mathrm{co}\, F, G)$ with a separable filtered probability space $\mathcal{P}_\mathbb{F}$ is weakly compact in distribution. By virtue of Lemma 1.7 of Chap. 3, one has $\mathcal{X}^0_\mu(\overline{F}, G) = \mathcal{X}^0_\mu(\mathrm{co}\, F, G)$. Then $\mathcal{X}^0_\mu(\overline{F}, G)$ is nonempty and weakly compact. □

5 Some Properties of Exit Times of Continuous Processes

Let D be a domain in $\mathbb{R}^d$ and $(s, x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(\cdot, t))_{t\geq 0}$ and $X^n = (X^n(\cdot, t))_{t\geq 0}$ are continuous stochastic processes on a stochastic base $\mathcal{P}_\mathbb{F} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ such that $X(\cdot, s) = X^n(\cdot, s) = x$ a.s. for $n = 1, 2, \ldots$ and $\sup_{t\geq 0} |X^n(\cdot, t) - X(\cdot, t)| \to 0$ a.s. as $n \to \infty$. Let $\tau = \inf\{r > s : X(\cdot, r) \notin D\}$ and $\tau_n = \inf\{r > s : X^n(\cdot, r) \notin D\}$ for $n = 1, 2, \ldots$. We shall show that if $\tau_n < \infty$ a.s. for every $n \geq 1$, then $\tau_n \to \tau$ a.s. as $n \to \infty$. We begin with the following lemmas.

Lemma 5.1. *Let D be a domain in $\mathbb{R}^d$, $(s, x) \in \mathbb{R}^+ \times D$, and $X = (X(\cdot, t))_{t\geq 0}$ a continuous d-dimensional stochastic process on $\mathcal{P}_\mathbb{F} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ such that $X(\cdot, s) = x$ a.s. and $\tau = \inf\{r > s : X(\cdot, r) \notin D\} < \infty$ a.s. If $T : \Omega \to \mathbb{R}$ is such that $T > \tau$ a.s., then $\tau = \inf\{r \in (s, T) : X(\cdot, r) \notin D\}$ a.s.*

Proof. For simplicity, assume that the above relations are satisfied for every $\omega \in \Omega$ and let us observe that $\tau(\omega) = \inf X^{-1}(\omega, \cdot)(D^\sim)$, where $D^\sim = \mathbb{R}^d \setminus D$. We have $X^{-1}(\omega, \cdot)(D^\sim) = X^{-1}(\omega, \cdot)(D^\sim) \cap (s, T(\omega)) \cup X^{-1}(\omega, \cdot)(D^\sim) \cap [T(\omega), \infty)$. Therefore, $\inf X^{-1}(\omega, \cdot)(D^\sim) \leq \inf(X^{-1}(\omega, \cdot)(D^\sim) \cap (s, T(\omega)))$. For every $\omega \in \Omega$, there exists $t(\omega) \in X^{-1}(\omega, \cdot)(D^\sim)$ such that $s < t(\omega) < T(\omega)$, because $\tau(\omega) < T(\omega)$ for $\omega \in \Omega$. Therefore, $X^{-1}(\omega, \cdot)(D^\sim) \cap (s, T(\omega)) \neq \emptyset$ and $\inf(X^{-1}(\omega, \cdot)(D^\sim) \cap (s, T(\omega))) \leq T(\omega)$ for a.e. $\omega \in \Omega$. Suppose $\tau = \inf X^{-1}(\omega, \cdot)(D^\sim) < \tau_T(\omega) =: \inf(X^{-1}(\omega, \cdot)(D^\sim) \cap (s, T(\omega)))$ on a set $\Omega_0 \in \mathcal{F}$ such that $P(\Omega_0) > 0$. Then for every $\omega \in \Omega_0$, there exists $\bar{t}(\omega) \in X^{-1}(\omega, \cdot)(D^\sim)$ such that $s < \bar{t}(\omega) < \tau_T(\omega) < T(\omega)$, which is a contradiction, because for every $\omega \in \Omega$ and $t \in X^{-1}(\omega, \cdot)(D^\sim) \cap (s, T(\omega))$, we have $\tau_T(\omega) \leq t$. Then $\tau(\omega) = \inf\{X^{-1}(\omega, \cdot)(D^\sim) \cap (s, T(\omega))\}$ for a.e. $\omega \in \Omega$. □

Lemma 5.2. *Let D be a domain in $\mathbb{R}^d$ and $(s,x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(\cdot,t))_{t\geq 0}$ and $X^n = (X^n(\cdot,t))_{t\geq 0}$ are continuous d-dimensional stochastic processes on a stochastic base $\mathcal{P}_{\mathbb{F}} = (\Omega,\mathcal{F},\mathbb{F},P)$ such that $X(\cdot,s) = X^n(\cdot,s) = x$ for $n = 1,2,\ldots$ and $\sup_{t\geq 0}|X^n(\cdot,t) - X(\cdot,t)| \to 0$ a.s. as $n \to \infty$. Then $\mathrm{Li}\, X_n^{-1}(\omega,\cdot)(D^\sim) = X^{-1}(\omega,\cdot)(D^\sim) = \mathrm{Ls}\, X_n^{-1}(\omega,\cdot)(D^\sim)$ for a.e. $\omega \in \Omega$.*

Proof. For simplicity, assume that $X(\omega,\cdot)$ and $X_n(\omega,\cdot)$ for $n = 1,2,\ldots$ are continuous and $\lim_{n\to\infty}\sup_{t\geq 0}|X^n(\omega,t) - X(\omega,t)| = 0$ for every $\omega \in \Omega$. For every $\omega \in \Omega$ and $\varepsilon > 0$, there exists $N_\varepsilon(\omega) \geq 1$ such that $X_n(\omega,t) \in X(\omega,t) + \varepsilon B$ and $X(\omega,t) \in X_n(\omega,t) + \varepsilon B$ for $t \geq s$ and $n \geq N_\varepsilon(\omega)$, where B is a closed unit ball of $\mathbb{R}^d$. Then $X_n^{-1}(\omega,\cdot)(\{X_n(\omega,t)\}) \subset X_n^{-1}(\omega,\cdot)(\{X(\omega,t) + \varepsilon B\})$ and $X^{-1}(\omega,\cdot)(\{X(\omega,t)\}) \subset X^{-1}(\omega,\cdot)(\{X_n(\omega,t) + \varepsilon B\})$ a.s. for $n \geq N_\varepsilon(\omega)$. Let us observe that for every $A \subset \mathbb{R}^+$ and $C \subset \mathbb{R}^d$, one has $A \subset X_n^{-1}(\omega,\cdot)(X_n(\omega,A))$, $A \subset X^{-1}(\omega,\cdot)(X(\omega,A))$, $X_n(\omega,X^{-1}(\omega,\cdot)(C)) \subset C + \varepsilon B$, and $X(\omega,X^{-1}(\omega,\cdot)(C)) \subset C$ for $n = 1,2,\ldots$. Taking in particular $A = X^{-1}(\omega,\cdot)(D^\sim)$ and $C = D^\sim$ in the above inclusions, we obtain $X^{-1}(\omega,\cdot)(D^\sim) \subset X_n^{-1}(\omega,\cdot)(X_n(\omega,X^{-1}(\omega,\cdot)(D^\sim))) \subset X_n^{-1}(\omega,\cdot)(D^\sim + \varepsilon B)$ a.s. for $n \geq N_\varepsilon(\omega)$. Similarly, taking $A = X_n^{-1}(\omega,\cdot)(D^\sim)$ and $C = D^\sim$, we obtain $X_n^{-1}(\omega,\cdot)(D^\sim) \subset X^{-1}(\omega,\cdot)(X(\omega,X_n^{-1}(\omega,\cdot)(D^\sim))) \subset X^{-1}(\omega,\cdot)(X_n(\omega,X_n^{-1}(\omega,\cdot)(D^\sim)) + \varepsilon B) \subset X^{-1}(\omega,\cdot)(D^\sim + \varepsilon B)$ a.s. for $n \geq N_\varepsilon(\omega)$. Hence it follows that

$$
\begin{aligned}
X^{-1}(\omega,\cdot)(D^\sim) &\subset \bigcap_{k=0}^{\infty} X^{-1}_{k+N_\varepsilon}(\omega,\cdot)(D^\sim + \varepsilon B)\\
&\subset \bigcup_{n=1}^{N_\varepsilon-1}\bigcap_{k=0}^{\infty} X^{-1}_{k+n}(\omega,\cdot)(D^\sim + \varepsilon B)\\
&\cup \bigcap_{k=0}^{\infty} X^{-1}_{k+N_\varepsilon}(\omega,\cdot)(D^\sim + \varepsilon B)\\
&\cup \bigcup_{n=N_\varepsilon+1}^{\infty}\bigcap_{k=0}^{\infty} X^{-1}_{k+n}(\omega,\cdot)(D^\sim + \varepsilon B)\\
&= \bigcup_{n=1}^{\infty}\bigcap_{k=0}^{\infty} X^{-1}_{k+n}(\omega,\cdot)(D^\sim + \varepsilon B)\\
&= \mathrm{Lim\,inf}\, X_n^{-1}(\omega,\cdot)(D^\sim + \varepsilon B)
\end{aligned}
$$

a.s. for every $\varepsilon > 0$, which by virtue of Corollary 1.1 of Chap. 2, implies $X^{-1}(\omega,\cdot)(D^\sim) \subset \bigcap_{\varepsilon>0}\mathrm{Lim\,inf}\, X^{-1}(\omega,\cdot)(D^\sim + \varepsilon B) = \mathrm{Lim\,inf}\, X_n^{-1}(\omega,\cdot)(D^\sim)$ a.s. Hence, by virtue of (ii) of Lemma 1.2 of Chap. 2, we obtain $X^{-1}(\omega,\cdot)(D^\sim) \subset \mathrm{Li}\, X_n^{-1}(\omega,\cdot)(D^\sim)$. In a similar way, we get $\overline{\bigcup_{k=0}^{\infty} X^{-1}_{k+N_\varepsilon}(\omega,\cdot)(D^\sim + \varepsilon B)} \subset X^{-1}(\omega,\cdot)(D^\sim + \varepsilon B)$. Then

$$\overline{\bigcap_{n=1}^{N_\varepsilon-1}\bigcup_{k=0}^{\infty} X_{k+n}^{-1}(\omega,\cdot)(D^{\sim})} \cup \overline{\bigcup_{k=0}^{\infty} X_{k+N_\varepsilon}^{-1}(\omega,\cdot)(D^{\sim})} \cup \bigcap_{n=N_\varepsilon+1}^{\infty}\overline{\bigcup_{k=0}^{\infty} X_{k+n}^{-1}(\omega,\cdot)(D^{\sim})}$$
$$\subset X^{-1}(\omega,\cdot)(D^{\sim}+\varepsilon B)$$

for every $\varepsilon > 0$. Hence, by virtue of (v) of Lemma 1.2 of Chap. 2, it follows that

$$\mathrm{Ls}\ X_n^{-1}(\omega,\cdot)(D^{\sim}) = \bigcap_{n=1}^{\infty}\overline{\bigcup_{k=0}^{\infty} X_{k+n}^{-1}(\omega,\cdot)(D^{\sim})} \subset X^{-1}(\omega,\cdot)(D^{\sim}+\varepsilon B)$$

for every $\varepsilon > 0$. Thus $\mathrm{Ls}\ X_n^{-1}(\omega,\cdot)(D^{\sim}) \subset \bigcap_{\varepsilon>0} X^{-1}(\omega,\cdot)(D^{\sim}+\varepsilon B) = X^{-1}(\omega,\cdot)(D^{\sim})$ a.s. From the above inclusions, we obtain $X^{-1}(\omega,\cdot)(D^{\sim}) \subset \mathrm{Li}\ X_n^{-1}(\omega,\cdot)(D^{\sim}) \subset \mathrm{Ls}\ X_n^{-1}(\omega,\cdot)(D^{\sim}) \subset X^{-1}(\omega,\cdot)(D^{\sim})$ a.s. Then

$$\mathrm{Ls}\ X_n^{-1}(\omega,\cdot)(D^{\sim}) \subset X^{-1}(\omega,\cdot)(D^{\sim}) \subset \mathrm{Li}\ X_n^{-1}(\omega,\cdot)(D^{\sim}),$$

which by (i) of Corollary 1.2 of Chap. 2, implies that $\mathrm{Li}\ X_n^{-1}(\omega,\cdot)\ (D^{\sim}) = \mathrm{Ls}\ X_n^{-1}(\omega,\cdot)(D^{\sim}) = X^{-1}(\omega,\cdot)(D^{\sim})$. □

Lemma 5.3. *Let D be a domain in $\mathbb{R}^d$ and $(s,x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(\cdot,t))_{t\geq 0}$ and $X^n = (X^n(\cdot,t))_{t\geq 0}$ are continuous d-dimensional stochastic processes on a stochastic base $\mathcal{P}_{\mathbb{F}} = (\Omega,\mathcal{F},\mathbb{F},P)$ such that $X(\cdot,s) = X^n(\cdot,s) = x$ for $n = 1,2,\ldots$ and $\sup_{t\geq 0}|X^n(\cdot,t) - X(\cdot,t)| \to 0$ a.s. as $n \to \infty$. If there exists a mapping $T : \Omega \to \mathbb{R}^+$ such that $\max(\tau,\tau_n) < T$ a.s. for $n = 1,2,\ldots$, where $\tau = \inf\{r > s : X(\cdot,r) \notin D\}$ and $\tau_n = \inf\{r > s : X_n(\cdot,r) \notin D\}$, then $(X^{-1}(\omega,\cdot)(D^{\sim})) \cap [s,T(\omega)) = \mathrm{Li}\,(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega))) = \mathrm{Ls}\,(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega)))$ for a.e. $\omega \in \Omega$.*

Proof. Assume that $X(\omega,\cdot)$ and $X_n(\omega,\cdot)$ for $n = 1,2,\ldots$ are continuous, $\max(\tau(\omega),\tau_n(\omega)) < T(\omega)$ for $n = 1,2,\ldots$, and $\lim_{n\to\infty}\sup_{t\geq 0}|X^n(\omega,t) - X(\omega,t)| = 0$ for every $\omega \in \Omega$. By virtue of (iv) and (vi) of Lemma 1.2 of Chap. 2 and Lemma 5.2, we get

$$\begin{aligned}\mathrm{Ls}\,(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega))) &\subset \mathrm{Ls}\,X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap [s,T(\omega))\\ &= X^{-1}(\omega,\cdot)(D^{\sim}) \cap [s,T(\omega)).\end{aligned}$$

Similarly, by virtue of (iii) and (vi) of Lemma 1.2 of Chap. 2, we also have

$$\mathrm{Li}\,(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega))) \subset (\mathrm{Li}\,(X_n^{-1}(\omega,\cdot)(D^{\sim})) \cap [s,T(\omega)).$$

By virtue of (ii) of Corollary 1.2 of Chap. 2, for every $t \in (\mathrm{Li}\,(X_n^{-1}(\omega,\cdot)(D^{\sim})) \cap [s,T(\omega))$, there exists $\bar{n} \geq 1$ such that for every $n > \bar{n}$, there is $t_n \in X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap [s,T(\omega))$ such that $t_n \to t$ as $n \to \infty$. Then $\mathrm{dist}(t, X_n^{-1}(\omega,\cdot)(D^{\sim})) \to 0$ as $n \to \infty$. Therefore, for every $\varepsilon > 0$, there exists $N_\varepsilon > \bar{n}$ such that $t \in X_n^{-1}(\omega,\cdot)(D^{\sim}) + \varepsilon B$ for $n \geq N_\varepsilon$. Hence, similarly as in the proof of

Lemma 5.2, it follows that for every $t \in \mathrm{Li}\,(X_n^{-1}(\omega,\cdot)(D^{\sim})) \cap [s, T(\omega))$ and $\varepsilon > 0$, one has

$$\begin{aligned} t &\in \bigcap_{k=0}^{\infty}\{(X_{k+N_\varepsilon}^{-1}(\omega,\cdot)(D^{\sim}) + \varepsilon B) \cap [s, T(\omega))\} \\ &\subset \bigcup_{n=1}^{\infty}\bigcap_{k=0}^{\infty}\{(X_{k+n}^{-1}(\omega,\cdot)(D^{\sim}) + \varepsilon B) \cap [s, T(\omega))\} \\ &= \mathrm{Lim\,inf}\{(X_n^{-1}(\omega,\cdot)(D^{\sim}) + \varepsilon B) \cap [s, T(\omega))\}. \end{aligned}$$

Then

$$\mathrm{Li}(X_n^{-1}(\omega,\cdot)(D^{\sim})) \cap [s, T(\omega)) \subset \mathrm{Lim\,inf}\{X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap [s, T(\omega))\}$$

$$\subset \mathrm{Li}\,(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap [s, T(\omega))) \subset \mathrm{Li}(X_n^{-1}(\omega,\cdot)(D^{\sim})) \cap [s, T(\omega)).$$

Thus

$$\begin{aligned} X^{-1}(\omega,\cdot)(D^{\sim}) \cap [s, T(\omega)) &= (\mathrm{Li}\,X_n^{-1}(\omega,\cdot)(D^{\sim})) \cap [s, T(\omega)) \\ &= \mathrm{Li}\,(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap [s, T(\omega))). \end{aligned}$$

Therefore, by (iv) and (vi) of Lemma 1.2 of Chap. 2 and Lemma 5.2, one has

$$\begin{aligned} &\mathrm{Ls}\,(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap [s, T(\omega))) \\ &\quad\subset X^{-1}(\omega,\cdot)(D^{\sim}) \cap [s, T(\omega)) \\ &\quad= \mathrm{Li}\,(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap [s, T(\omega))). \end{aligned}$$ ☐

Lemma 5.4. *Let D be a domain in $\mathbb{R}^d$ and $(s, x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(\cdot,t))_{t\geq 0}$ and $X^n = (X^n(\cdot,t))_{t\geq 0}$ are continuous d-dimensional stochastic processes on a stochastic base $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ such that $X(\cdot,s) = X^n(\cdot,s) = x$ for $n = 1, 2, \ldots$ and $\sup_{t\geq 0}|X^n(\cdot,t) - X(\cdot,t)| \to 0$ a.s. as $n \to \infty$. If $\inf X_n^{-1}(\omega,\cdot)(D^{\sim}) < \infty$ for a.e. $\omega \in \Omega$ for $n = 1, 2, \ldots$, then $\inf X^{-1}(\omega,\cdot)(D^{\sim}) < \infty$ for a.e. $\omega \in \Omega$.*

Proof. Let $\tau_n(\omega) = \inf X_n^{-1}(\omega,\cdot)(D^{\sim}) < \infty$ and $\tau(\omega) = \inf X^{-1}(\omega,\cdot)(D^{\sim})$ for $\omega \in \Omega$. Put $\Lambda = \{\omega \in \Omega : \tau(\omega) = \infty\}$ and $\Lambda_n = \{\omega \in \Omega : \tau_n(\omega) = \infty\}$ for $n = 1, 2, \ldots$. For every $\omega \in \Lambda$, one has $X(\omega, t) \in D$ for $t \geq s$. By the properties of the sequence $(X_n)_{n=1}^{\infty}$ for a.e. fixed $\omega \in \Lambda$, there exists a positive integer $N(\omega) \geq 1$ such that $X_n(\omega, t) \in D$ for $t \geq s$ and every $n \geq N(\omega)$. Then for a.e. $\omega \in \Lambda$ and every $n \geq N(\omega)$, we have $\tau_n(\omega) = \infty$. For simplicity, assume that $\tau_n(\omega) = \infty$ for every $n \geq N(\omega)$ and $\omega \in \Lambda$. By the assumption that $\tau_n < \infty$ a.s. and the definition of Λ_n, we have $P(\Lambda_n) = 0$ for every $n \geq 1$. Then $P(\bigcup_{n=1}^{\infty}\Lambda_n) = 0$. But for every $\omega \in \Lambda$ and $n \geq N(\omega)$, we have $\tau_n(\omega) = \infty$. Therefore, $\Lambda \subset \bigcup_{n=1}^{\infty}\Lambda_n$. Then $P(\Lambda) = 0$. ☐

Lemma 5.5. *Let D be a domain in $\mathbb{R}^d$ and $(s,x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(\cdot,t))_{t\geq 0}$ and $X^n = (X^n(\cdot,t))_{t\geq 0}$ are continuous d-dimensional stochastic processes on a stochastic base $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ such that $X(\cdot,s) = X^n(\cdot,s) = x$ for $n = 1,2,\ldots$ and $\sup_{t\geq 0}|X^n(\cdot,t) - X(\cdot,t)| \to 0$ a.s. as $n \to \infty$ and let $\tau_n(\omega) = \inf X_n^{-1}(\omega,\cdot)(\bar{D}^{\sim})$ and $\tau(\omega) = \inf X^{-1}(\omega,\cdot)(D^{\sim})$ for $\omega \in \Omega$. If $\max(\tau_n,\tau) < \infty$ a.s. for $n = 1,2,\ldots$, then there is a mapping $T : \Omega \to \mathbb{R}^+$ such that $\max(\tau_n,\tau) < T$ a.s. for $n \geq 1$.*

Proof. By virtue of Lemma 5.2, we have $\tau(\omega) = \inf(\mathrm{Li}\, X_n^{-1}(\omega,\cdot)(D^{\sim}))$ for a.e. $\omega \in \Omega$. By virtue of (ii) of Corollary 1.2 of Chap. 2, for a.e. $\omega \in \Omega$ there is $\bar{n} \geq 1$ such that for every $n > \bar{n}$, there exists $t_n \in X_n^{-1}(\omega,\cdot)(D^{\sim})$ such that $t_n \to \tau$ a.s. as $n \to \infty$. For every $n > \bar{n}$, we have $\tau_n \leq \tau$, because $X^{-1}(\omega,\cdot)(D^{\sim}) \subset X_n^{-1}(\omega,\cdot)(D^{\sim} + \varepsilon B)$, $\tau_n^{\varepsilon} \leq \tau$ and $\tau_n^{\varepsilon} \to \tau_n$ a.s. as $\varepsilon \to 0$, where $\tau_n^{\varepsilon}(\omega) = \inf X_n^{-1}(\omega,\cdot)(D^{\sim} + \varepsilon B)$ for $n \geq \bar{n}$. Then $\limsup \tau_n \leq \tau$ a.s., which implies that for a.e. $\omega \in \Omega$, there exists a positive integer $N(\omega) \geq 1$ such that $\tau_n(\omega) < \tau(\omega)$ for $n \geq N(\omega)$. Taking $T(\omega) = \max\{\tau_1(\omega)+1, \tau_2(\omega)+1, \ldots, \tau_{N(\omega)}(\omega)+1, \tau(\omega)+1\}$ for a.e. $\omega \in \Omega$, we have defined a mapping $T : \Omega \to \mathbb{R}^+$ such that $\max(\tau_n,\tau) < T$ a.s. for $n \geq 1$. □

Now we can prove the following convergence theorem.

Theorem 5.1. *Let D be a domain in $\mathbb{R}^d$ and $(s,x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(\cdot,t))_{t\geq 0}$ and $X^n = (X^n(\cdot,t))_{t\geq 0}$ are continuous d-dimensional stochastic processes on a stochastic base $\mathcal{P}_{\mathbb{F}} = (\Omega, \mathcal{F}, \mathbb{F}, P)$ such that $X(\cdot,s) = X^n(\cdot,s) = x$ for $n = 1,2,\ldots$ and $\sup_{t\geq 0}|X^n(\cdot,t) - X(\cdot,t)| \to 0$ a.s. as $n \to \infty$. If $\tau_n = \inf\{r > s : X_n(\cdot,r) \notin D\} < \infty$ a.s. for $n = 1,2,\ldots$, then $\lim_{n\to\infty}\tau_n = \tau$ a.s., where $\tau = \inf\{r > s : X(\cdot,r) \notin D\}$.*

Proof. By virtue of Lemma 5.4, we have $\max(\tau_n,\tau) < \infty$ a.s. for $n = 1,2,\ldots$. Therefore, by virtue of Lemma 5.5, there is a mapping $T : \Omega \to \mathbb{R}^+$ such that $\max(\tau_n,\tau) < T$ a.s. for $n = 1,2,\ldots$. Then by virtue of Lemma 5.1, we have $\tau_n(\omega) = \inf(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega)))$ and $\tau(\omega) = \inf(X^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega)))$ for $\omega \in \Omega$ and $n = 1,2,\ldots$. By virtue of Lemma 5.3, Remark 1.2 of Chap. 2, and Theorem 1.1 of Chap. 2, we get

$$\begin{aligned}
&\lim_{n\to\infty} h((X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega))), X^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega))))\\
&\quad = \lim_{n\to\infty} h(\overline{(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega)))}, \overline{X^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega))}))\\
&\quad = 0
\end{aligned}$$

for a.e. $\omega \in \Omega$, where h is the Hausdorff metric on $\mathrm{Cl}([s,T(\omega)])$ for every fixed $\omega \in \Omega$. Let $\varepsilon > 0$ and $t_\varepsilon(\omega) \in X^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega)))$ be such that $t_\varepsilon(\omega) < \tau(\omega) + \varepsilon$ for fixed $\omega \in \Omega$. By the above property of the sequence $(X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega)))_{n=1}^{\infty}$ and the definition of the Hausdorff metric h, we have $\mathrm{dist}(t_\varepsilon(\omega), X_n^{-1}(\omega,\cdot)(D^{\sim}) \cap (s,T(\omega))) \to 0$ for fixed $\omega \in \Omega$ and every $\varepsilon > 0$ as $n \to \infty$. Therefore, for every fixed $\omega \in \Omega$, there exists a sequence $(t_\varepsilon^n(\omega))_{n=1}^{\infty}$

such that $t_\varepsilon^n(\omega) \in X_n^{-1}(\omega,\cdot)(D^\sim) \cap (s, T(\omega))$ for $n \geq 1$ and $|t_\varepsilon^n(\omega) - t_\varepsilon(\omega)| \to 0$ as $n \to \infty$. Hence it follows that

$$\tau_n(\omega) \leq t_\varepsilon^n(\omega) \leq |t_\varepsilon^n(\omega) - t_\varepsilon(\omega)| + t_\varepsilon(\omega) < |t_\varepsilon^n(\omega) - t_\varepsilon(\omega)| + \tau(\omega) + \varepsilon$$

for $\varepsilon > 0$ and $n \geq 1$. Then $\limsup_{n\to\infty} \tau_n(\omega) \leq \tau(\omega)$.

Similarly, for fixed $\omega \in \Omega$ and every $\varepsilon > 0$ and $n \geq 1$, we can select $t_\varepsilon^n(\omega) \in X_n^{-1}(\omega,\cdot)(D^\sim) \cap [s, T(\omega))$ and $\bar{t}_\varepsilon^n \in X^{-1}(\omega,\cdot)(D^\sim) \cap [s, T(\omega))$ such that $t_\varepsilon^n(\omega) \leq \tau_n(\omega) + \varepsilon$ and $|\bar{t}_\varepsilon^n(\omega) - t_\varepsilon^n(\omega)| \to 0$ as $n \to \infty$. Hence it follows that

$$\tau(\omega) \leq \bar{t}_\varepsilon^n(\omega) \leq |\bar{t}_\varepsilon^n(\omega) - t_\varepsilon^n(\omega)| + t_\varepsilon^n(\omega) \leq |\bar{t}_\varepsilon^n(\omega) - t_\varepsilon^n(\omega)| + \tau_n(\omega) + \varepsilon$$

for every $\varepsilon > 0$ and $n \geq 1$. Therefore, $\tau(\omega) \leq \liminf_{n\to\infty} \tau_n(\omega)$. Then $\limsup_{n\to\infty} \tau_n(\omega) \leq \tau(\omega) \leq \liminf_{n\to\infty} \tau_n(\omega)$ for a.e. $\omega \in \Omega$, which implies that $\lim_{n\to\infty} \tau_n = \tau$ a.s. □

Let D be a domain in $\mathbb{R}^d$ and $(s,x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(t))_{t\geq 0}$ and $\tilde{X} = (\tilde{X}(t))_{t\geq 0}$ are continuous d-dimensional stochastic processes on $(\Omega, \mathcal{F}, P)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, respectively, such that $X(s) = x$ a.s. and $PX^{-1} = P\tilde{X}^{-1}$. We shall show that $P(\tau_D)^{-1} = P(\tilde{\tau}_D)^{-1}$, $P(X \circ \tau_D)^{-1} = P(\tilde{X} \circ \tilde{\tau}_D)^{-1}$, and $P(\tau_D, X \circ \tau_D)^{-1} = P(\tilde{\tau}_D, \tilde{X} \circ \tilde{\tau}_D)^{-1}$, where $\tau_D = \inf\{t > s : X_t \notin D\}$ and $\tilde{\tau}_D = \inf\{t > s : \tilde{X}_t \notin D\}$.

The next results will follow from the following fundamental lemma, similar to Lemma 2.1 of Chap. 1.

Lemma 5.6. *Let X and $\tilde{X}$ be as above, $(Y, \mathcal{G})$ a measurable space, and $C =: C(\mathbb{R}^+, \mathbb{R}^d)$. If $\Phi : C \to Y$ is $(\beta, \mathcal{G})$-measurable, where β is a Borel σ-algebra on C, then $P(\Phi \circ X)^{-1} = P(\Phi \circ \tilde{X})^{-1}$.*

Proof. Let $Z = \Phi \circ X$ and $\tilde{Z} = \Phi \circ \tilde{X}$. For every $A \in \mathcal{G}$, one has $P(\{Z \in A\}) = P(\{\Phi \circ X \in A\}) = P(X^{-1}(\Phi^{-1}(A))) = \tilde{P}(\tilde{X}^{-1}(\Phi^{-1}(A))) = \tilde{P}(\{\Phi \circ \tilde{X} \in A\}) = \tilde{P}(\{\tilde{Z} \in A\})$. Then $P(\Phi \circ X)^{-1} = P(\Phi \circ \tilde{X})^{-1}$. □

The following theorem can be derived immediately from the above result.

Theorem 5.2. *Let D be a domain in $\mathbb{R}^d$ and $(s,x) \in \mathbb{R}^+ \times D$. Assume that $X = (X(t))_{t\geq 0}$ and $\tilde{X} = (\tilde{X}(t))_{t\geq 0}$ are continuous d-dimensional stochastic processes on $(\Omega, \mathcal{F}, P)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, respectively, such that $X(s) = x$ a.s. and $PX^{-1} = P\tilde{X}^{-1}$. Then $P(\tau_D)^{-1} = P(\tilde{\tau}_D)^{-1}$, $P(X \circ \tau_D)^{-1} = P(\tilde{X} \circ \tilde{\tau}_D)^{-1}$, and $P(\tau_D, X \circ \tau_D)^{-1} = P(\tilde{\tau}_D, \tilde{X} \circ \tilde{\tau}_D)^{-1}$, where $\tau_D = \inf\{t > s : X_t \notin D\}$ and $\tilde{\tau}_D = \inf\{t > s : \tilde{X}_t \notin D\}$.*

Proof. Let $\eta : C \to \mathbb{R}^+$ be defined by $\eta(x) = \inf\{t > s : x(t) \notin D\}$ for $x \in C$. It is clear that η is (β, β_+)-measurable, where β_+ denotes the Borel σ-algebra on $\mathbb{R}^+$. Taking $Y = \mathbb{R}^+$, $\mathcal{G} = \beta_+$, and $\Phi = \eta$, we get $\tau_D = \Phi \circ X$ and $\tilde{\tau}_D = \Phi \circ \tilde{X}$. Therefore, by virtue of Lemma 5.6, we obtain $P(\tau_D)^{-1} = P(\tilde{\tau}_D)^{-1}$. Let $\psi(t,x) = x(t)$ for $x \in C$ and $t \in \mathbb{R}^+$ and put $\Phi(x) = \psi(\eta(x), x))$ for $x \in C$. It is clear that the mapping Φ satisfies the conditions of Lemma 5.6 with $Y = \mathbb{R}^d$

and $\mathcal{G} = \beta$, where β denotes the Borel σ-algebra on $\mathbb{R}^d$. Furthermore, we have $\Phi \circ X = X \circ \tau_D$ and $\Phi \circ \tilde{X} = \tilde{X} \circ \tilde{\tau}_D$. Therefore, by virtue of Lemma 5.6, we obtain $P(X \circ \tau_D)^{-1} = P(\tilde{X} \circ \tilde{\tau}_D)^{-1}$. Finally, let $\Phi(x) = (\eta(x), \psi(\eta(x), x))$ for $x \in C$. Immediately from the properties of the mappings ψ and η, it follows that Φ satisfies the conditions of Lemma 5.6 with $Y = \mathbb{R}^+ \times \mathbb{R}^d$ and $\mathcal{G} = \beta_+ \times \beta$, where β_+ denotes the Borel σ-algebra of $\mathbb{R}^+$. Furthermore, $\Phi \circ X = (\tau_D, X \circ \tau_D)$ and $\Phi \circ \tilde{X} = (\tilde{\tau}_D, \tilde{X} \circ \tilde{\tau}_D)$, which by virtue of Lemma 5.6, implies $P(\tau_D, X \circ \tau_D)^{-1} = P(\tilde{\tau}_D, \tilde{X} \circ \tilde{\tau}_D)^{-1}$. □

Corollary 5.1. *If the assumptions of Theorem 5.2 are satisfied, then for every continuous bounded function* $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$, *one has* $E[f(\tau_D, X \circ \tau_D)] = \tilde{E}[f(\tilde{\tau}_D, \tilde{X} \circ \tilde{\tau}_D)]$, *where* E *and* $\tilde{E}$ *denote the mean value operators with respect to probability measures* P *and* $\tilde{P}$, *respectively.* □

6 Notes and Remarks

The first papers concerning stochastic functional inclusions written in the set-valued integral form are due to Hiai [38] and Kisielewicz [51, 55], where stochastic functional inclusions containing set-valued stochastic integrals were independently investigated. In the above papers, only strong solutions were considered. An extension of the Fillipov theorem for stochastic differential inclusions was given by Da Prato and Frankowska [23]. Existence and stability of solutions of stochastic differential inclusions were considered by Motyl in [82] and [83], resp. Weak solutions of stochastic functional inclusions have been considered by Aubin and Da Prato [9], Kisielewicz [53] and Levakov [71]. Weak compactness with respect to convergence in distribution of solution sets of weak solutions of stochastic differential inclusions was considered in Kisielewicz [56, 58, 60]. Also, Levakov in [71] considered weak compactness of all distributions of weak solutions of some special type of stochastic differential inclusions. Compactness of solutions of second order dynamical systems was considered by Michta and Motyl in [78]. The results of the last three sections of this chapter are based on Kisielewicz [56, 58], where stochastic functional inclusions in the finite intervals $[0, T]$ are considered. The results dealing with backward stochastic differential inclusions were first considered in the author's paper [59]. The results contained in Sect. 5 are taken entirely from Kisielewicz [55]. The properties of stochastic differential inclusions presented in Sect. 2 are the first dealing with such inclusions. By Theorem 2.1 of Chap. 3, stochastic differential inclusions $SDI(F, G)$ are equivalent to stochastic functional inclusions of the form $x_t - x_s \in \overline{\text{dec}}\{J(F \circ x)\} + \overline{\text{dec}}\{\mathcal{J}(G \circ x)\}$. Therefore, for multifunctions F and G satisfying the assumptions of Theorem 1.5, the set $S_w(F, G, \mu)$ of all weak solutions of $SDI(F, G)$ with an initial distribution μ contains a set considered in optimal control problems described by $SDI(F, G)$.

For the existence of solutions of such optimal control problems, it is necessary to have some sufficient conditions implying weak compactness of a solution set $S_w(F, G, \mu)$. Such results are difficult to obtain by the methods used in the proof of Theorem 4.1, because boundedness or square integrable boundedness of $\overline{\text{dec}}\{J(F \circ x)\}$ and $\overline{\text{dec}}\{\mathcal{J}(G \circ x)\}$ is necessary in such a proof.