Weak and strong solutions of general stochastic models

Thomas G. Kurtz *
Departments of Mathematics and Statistics
University of Wisconsin - Madison
480 Lincoln Drive
Madison, WI 53706-1388
kurtz@math.wisc.edu
http://www.math.wisc.edu/~kurtz/

April 25, 2013

Abstract

Typically, a stochastic model relates stochastic “inputs” and, perhaps, controls to stochastic “outputs”. A general version of the Yamada-Watanabe and Engelbert theorems relating existence and uniqueness of weak and strong solutions of stochastic equations is given in this context. A notion of compatibility between inputs and outputs is critical in relating the general result to its classical forebears.

Key words: weak solution, strong solution, stochastic models pointwise uniqueness, pathwise uniqueness, stochastic differential equations, stochastic partial differential equations

MSC 2010 Subject Classification: Primary: 60G05 Secondary: 60H10, 60H15, 60H20, 60H25.

1 Introduction and main theorem

This paper is essentially a rewrite of Kurtz (2007) following a realization that the general, abstract theorem in that paper was neither as abstract as it could be nor as general as it should be. The reader familiar with the earlier paper may not be pleased by the greater abstraction, but an example indicating the value of the greater generality will be given in Section 2. To simplify matters for the reader, proofs of several lemmas that originally appeared in the earlier paper are included, but the reader should refer to the earlier paper for more examples and additional references.

As with the results of the earlier paper, the main theorem given here generalizes the famous theorem of Yamada and Watanabe (1971) giving the relationship between weak and strong solutions of an Itô equation for a diffusion and their existence and uniqueness. A

*Research supported in part by NSF grant DMS 11-06424
second reason for this rewrite is that the main observation that ensures that the main theorem gives the Yamada-Watanabe result is buried in a proof in the earlier paper. Here it is stated separately as Lemma 2.8.

The motivation of the original Yamada-Watanabe result arises naturally in the process of proving existence of solutions of a stochastic differential equation or, in the context of the present paper, existence of a stochastic model determined by constraints that may but need not be equations. The basic existence argument starts by identifying a sequence of approximations to the equation (or model) for which existence of solutions is simple to prove, proving relative compactness of the sequence of approximating solutions, and then verifying that any limit point is a solution of the original equation (model). The issue addressed by the Yamada-Watanabe theorem is that frequently, the kind of compactness verified is weak or distributional compactness. Consequently, what can be claimed about the limit is that there exists a probability space on which processes are defined that satisfy the original equation. Such solutions are called weak solutions, and their existence leaves open the question of whether there exists a solution on every probability space that supports the stochastic inputs of the model, that is, the Brownian motion and initial position in the original Itô equation context. The assertion of the Yamada-Watanabe theorem and Theorem 1.5 below is that if a strong enough form of uniqueness can be verified, then existence of a weak solution implies existence on every such probability space.

A stochastic model describes the relationship between stochastic inputs and stochastic outputs. For example, in the case of the Itô equation,

\[ X(t) = X(0) + \int_0^t \sigma(X(s))dW(s) + \int_0^t b(X(s))ds, \]

where \( X(0) \) and \( W \) are the stochastic inputs and the solution \( X \) gives the outputs. Typically, the distribution of the inputs is specified (for example, the initial distribution is given and \( X(0) \) is assumed independent of the Brownian motion \( W \)), and the model is determined by a set of constraints (possibly, but not necessarily, equations) that relate the inputs to the outputs. In the general setting here, the inputs will be given by a random variable \( Y \) with values in a complete, separable metric space \( S_2 \) and the outputs \( X \) will take values in a complete, separable metric space \( S_1 \). For the Itô equation, we could take \( S_2 = \mathbb{R}^d \times \mathbb{C} \mathbb{R}^d[0,\infty) \) and \( S_1 = \mathbb{C} \mathbb{R}^d[0,\infty) \).

Let \( P(S_1 \times S_2) \) be the space of probability measures on \( S_1 \times S_2 \), and for random variables \((X,Y)\) in \( S_1 \times S_2 \), let \( \mu_{XY} \in P(S_1 \times S_2) \) denote their joint distribution. Our model is determined by specifying a distribution \( \nu \) for the inputs \( Y \) and a set of constraints \( \Gamma \) relating \( X \) and \( Y \). Let \( P_{\nu}(S_1 \times S_2) \) be the set of \( \mu \in P(S_1 \times S_2) \) such that \( \mu(S_1 \times \cdot) = \nu \), and let \( S_{\Gamma,\nu} \) be the subset of \( P_{\nu}(S_1 \times S_2) \) such that \( \mu_{XY} \in S_{\Gamma,\nu} \) implies \((X,Y)\) meets the constraints in \( \Gamma \). Of course, since we are not placing any restriction on the nature of the constraints, \( S_{\Gamma,\nu} \) could be any subset of \( P_{\nu}(S_1 \times S_2) \).

For a second example, consider a typical stochastic optimization problem.

**Example 1.1** Suppose \( \Gamma_0 \) is a collection of constraints of the form

\[ E[\psi(X,Y)] < \infty \text{ and } E[f_i(X,Y)] = 0, \quad i \in \mathcal{I}, \]

where \( \psi \geq 0 \) and \( |f_i(x,y)| \leq \psi \).

2
Let \( 0 \leq c(x,y) \leq \psi(x,y) \), and let \( \Gamma \) be the set of constraints obtained from \( \Gamma_0 \) by adding the requirement
\[
\int c(x,y)\mu(dx,dy) = \inf_{\mu' \in \mathcal{S}_{\Gamma_0,\nu}} \int c(x,y)\mu'(dx,dy).
\]

In the terminology of Engelbert (1991) and Jacod (1980), \( \mu \in \mathcal{S}_{\Gamma,\nu} \) is a joint solution measure for our model \((\Gamma,\nu)\). Any distribution \( \mu \in \mathcal{S}_{\Gamma,\nu} \) determines a weak solution for our model \((\Gamma,\nu)\), that is, there exists a probability space on which are defined random variables \((X,Y)\) such that \(Y\) has distribution \(\nu\) and \((X,Y)\) meet the constraints in \(\Gamma\). We have the following definition for a strong solution.

**Definition 1.2** A solution \((X,Y)\) for \((\Gamma,\nu)\) is a strong solution if there exists a Borel measurable function \(F : S_2 \to S_1\) such that \(X = F(Y)\) a.s.

If a strong solution exists on some probability space, then a strong solution exists for any \(Y\) with distribution \(\nu\). It is important to note that being a strong solution is a distributional property, that is, the joint distribution of \((X,Y)\) is determined by \(\nu\) and \(F\). The following lemma helps to clarify the difference between a strong solution and a weak solution that does not correspond to a strong solution.

**Lemma 1.3** Let \(\mu \in \mathcal{P}_\nu(S_1 \times S_2)\) and \(\mu(S_1 \times \cdot) = \nu\).

a) There exists a transition function \(\eta\) such that \(\mu(dx \times dy) = \eta(y,dx)\nu(dy)\).

b) There exists a Borel measurable \(G : S_2 \times [0,1] \to S_1\) such that if \(Y\) has distribution \(\nu\) and \(\xi\) is independent of \(Y\) and uniformly distributed on \([0,1]\), \((G(Y,\xi),Y)\) has distribution \(\mu\).

c) \(\mu\) corresponds to a strong solution if and only if \(\eta(y,dx) = \delta_{F(y)}(dx)\).

**Proof.** Statement (a) is a standard result on the disintegration of measures. A particularly nice construction that gives the desired \(G\) in Statement (b) can be found in Blackwell and Dubins (1983). Statement (c) is immediate. \(\square\)

We have the following notions of uniqueness.

**Definition 1.4** Pointwise (pathwise for stochastic processes) uniqueness holds, if \(X_1, X_2,\) and \(Y\) defined on the same probability space with \(\mu_{X_1,Y}, \mu_{X_2,Y} \in \mathcal{S}_{\Gamma,\nu}\) implies \(X_1 = X_2\) a.s.

Joint uniqueness in law (or weak joint uniqueness) holds, if \(\mathcal{S}_{\Gamma,\nu}\) contains at most one measure.

Uniqueness in law (or weak uniqueness) holds if all \(\mu \in \mathcal{S}_{\Gamma,\nu}\) have the same marginal distribution on \(S_1\).

We have the following generalization of the theorems of Yamada and Watanabe (1971) and Engelbert (1991).

**Theorem 1.5** The following are equivalent:
a) $S_{\Gamma,\nu} \neq \emptyset$, and pointwise uniqueness holds.

b) There exists a strong solution, and joint uniqueness in law holds.

**Remark 1.6** In the special case that all constraints are given by simple equations, for example,

$$f_i(X,Y) = 0 \quad \text{a.s.} \quad i \in \mathcal{I},$$

then Proposition 2.10 of [Kurtz (2007)] shows that pointwise uniqueness, joint uniqueness in law, and uniqueness in law are equivalent. Note that stochastic differential equations are not of the form (1.1) (see Section 2), and the equivalence of uniqueness in law and joint uniqueness in law does not follow from this proposition in that setting; however, [Cherny (2003)] has shown the equivalence of uniqueness in law and joint uniqueness in law for Itô equations for diffusion processes.

**Proof.** Assume (a). If $\mu_1, \mu_2 \in S_{\Gamma,\nu}$, then there exist $G_1(y,u)$ and $G_2(y,u)$ such that for $Y$ with distribution $\nu$ and $\xi_1, \xi_2$ uniform on $[0,1]$, all independent, $(G_1(Y,\xi_1), Y)$ has distribution $\mu_1$ and $(G_2(Y,\xi_2), Y)$ has distribution $\mu_2$. By pointwise uniqueness,

$$G_1(Y,\xi_1) = G_2(Y,\xi_2) \quad \text{a.s.}$$

From the independence of $\xi_1$ and $\xi_2$, it follows that there exists $F$ such that $F(Y) = G_1(Y,\xi_1) = G_2(Y,\xi_2)$. (See Lemma A.2 of [Kurtz (2007)].)

Assume (b). Suppose $X_1, X_2, Y$ are defined on the same probability space and $\mu_{X_1,Y}, \mu_{X_2,Y} \in S_{\Gamma,\nu}$. By Lemma 1.3 the unique $\mu \in S_{\Gamma,\nu}$ must satisfy $\mu(dx \times dy) = \delta_{F(y)}(dx)\nu(dy)$, so $X_1 = F(Y) = X_2$ almost surely giving pointwise uniqueness.

The main result in [Kurtz (2007)], Theorem 3.14, was stated assuming the compatibility condition to be discussed in the next section and under the assumption that $S_{\Gamma,\nu}$ was convex. Neither assumption is needed for Theorem 1.5. The compatibility condition is critical to showing that Theorem 1.5 implies the classical Yamada-Watanabe result as well as a variety of more recent results for other kinds of stochastic equations. (See [Kurtz (2007)] for references.) The convexity assumption is useful in giving the following additional result.

**Corollary 1.7** Suppose $S_{\Gamma,\nu}$ is nonempty and convex. Then every solution is a strong solution if and only if pointwise uniqueness holds.

**Proof.** By Theorem 1.5 pointwise uniqueness implies $S_{\Gamma,\nu}$ contains only one distribution and the corresponding solution is strong. Conversely, suppose every solution is a strong solution. If $\mu_1, \mu_2 \in S_{\Gamma,\nu}$, then $\mu_0 = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 \in S_{\Gamma,\nu}$. Let $Y$ have distribution $\nu$. Then there exist Borel Functions $F_1$ and $F_2$ such that $(F_1(Y), Y)$ has distribution $\mu_1$ and $(F_2(Y), Y)$ has distribution $\mu_2$. Let $\xi$ be uniformly distributed on $[0,1]$ and independent of $Y$. Define

$$X = \begin{cases} F_1(Y) & \xi > 1/2 \\ F_2(Y) & \xi \leq 1/2 \end{cases}.$$

Then $(X,Y)$ has distribution $\mu_0$ and must satisfy $X = F(Y)$ for some $F$. Since $\xi$ is independent of $Y$, we must have $F_1(Y) = F(Y) = F_2(Y)$ a.s., giving pointwise uniqueness. \qed

4
2 Compatibility

It is not immediately obvious that Theorem 1.5 gives the classical Yamada-Watanabe theorem since proofs of pathwise uniqueness require appropriate adaptedness conditions in order to compare two solutions. This leads us to introduce the notion of compatibility. In what follows, if \( S \) is a metric space, then \( \mathcal{B}(S) \) will denote the Borel \( \sigma \)-algebra and \( B(S) \) will denote the space of bounded, Borel measurable functions; if \( \mathcal{M} \) is a \( \sigma \)-algebra, \( B(\mathcal{M}) \) will denote the space of bounded, \( \mathcal{M} \)-measurable functions.

Let \( E_1 \) and \( E_2 \) be complete, separable metric spaces, and let \( D_{E_1}[0, \infty) \), be the Skorohod space of cadlag \( E_1 \)-valued functions. Let \( Y \) be a process in \( D_{E_1}[0, \infty) \). By \( \mathcal{F}_t^Y \), we mean the completion of \( \sigma(Y(s), s \leq t) \).

**Definition 2.1** A process \( X \) in \( D_{E_1}[0, \infty) \) is temporally compatible with \( Y \) if for each \( t \geq 0 \) and \( h \in B(D_{E_2}(0, \infty)) \),

\[
E[h(Y)|\mathcal{F}_t^{X,Y}] = E[h(Y)|\mathcal{F}_t^Y] \quad \text{a.s.}
\]

where \( \{\mathcal{F}_t^{X,Y}\} \) denotes the complete filtration generated by \( (X, Y) \) and \( \{\mathcal{F}_t^Y\} \) denotes the complete filtration generated by \( Y \).

This definition is essentially (4.5) of Jacod (1980). If \( Y \) has independent increments, then \( X \) is compatible with \( Y \) if \( Y(t + \cdot) - Y(t) \) is independent of \( \mathcal{F}_t^{X,Y} \) for all \( t \geq 0 \).

We will consider a more general notion of compatibility. If \( \mathcal{B}_{\alpha_1}^{S_1} \) is a sub-\( \sigma \)-algebra of \( \mathcal{B}(S_1) \) and \( X \) is an \( S_1 \)-valued random variable on a complete probability space \( (\Omega, \mathcal{F}, P) \), then \( \mathcal{F}_\alpha^X \equiv \) the completion of \( \{\{X \in D\} : D \in \mathcal{B}_{\alpha_1}^{S_1}\} \) is the complete, sub-\( \sigma \)-algebra of \( \mathcal{F} \) generated by \( \{h(X) : h \in B(\mathcal{B}_{\alpha_1}^{S_1})\} \), where \( B(\mathcal{B}_{\alpha_1}^{S_1}) \) is the collection of \( h \in B(S_1) \) that are \( \mathcal{B}_{\alpha_1}^{S_1} \)-measurable.

**Definition 2.2** Let \( \mathcal{A} \) be an index set, and for each \( \alpha \in \mathcal{A} \), let \( \mathcal{B}_{\alpha_1}^{S_1} \) be a sub-\( \sigma \)-algebra of \( \mathcal{B}(S_1) \) and \( \mathcal{B}_{\alpha_2}^{S_2} \) be a sub-\( \sigma \)-algebra of \( \mathcal{B}(S_2) \). The collection \( \mathcal{C} \equiv \{(\mathcal{B}_{\alpha_1}^{S_1}, \mathcal{B}_{\alpha_2}^{S_2}) : \alpha \in \mathcal{A}\} \) will be referred to as a compatibility structure.

Let \( Y \) be an \( S_2 \)-valued random variable. An \( S_1 \)-valued random variable \( X \) is \( \mathcal{C} \)-compatible with \( Y \) if for each \( \alpha \in \mathcal{A} \) and each \( h \in B(S_2) \),

\[
E[h(Y)|\mathcal{F}_\alpha^X \vee \mathcal{F}_\alpha^Y] = E[h(Y)|\mathcal{F}_\alpha^Y] \quad \text{a.s.,}
\]

where \( \mathcal{F}_\alpha^X \equiv \) the completion of \( \{\{X \in D\} : D \in \mathcal{B}_{\alpha_1}^{S_1}\} \) and \( \mathcal{F}_\alpha^Y \equiv \) the completion of \( \{\{Y \in D\} : D \in \mathcal{B}_{\alpha_2}^{S_2}\} \).

**Lemma 2.3** If \( \mathcal{A} \) is partially ordered and \( \{\mathcal{F}_\alpha^X, \alpha \in \mathcal{A}\} \) and \( \{\mathcal{F}_\alpha^Y, \alpha \in \mathcal{A}\} \) are filtrations \((\alpha_1 \prec \alpha_2 \text{ implies } \mathcal{F}_{\alpha_1}^X \subseteq \mathcal{F}_{\alpha_2}^X \text{ and } \mathcal{F}_{\alpha_1}^Y \subseteq \mathcal{F}_{\alpha_2}^Y)\), then \( \mathcal{C} \)-compatibility is equivalent to the assertion that every \( \{\mathcal{F}_\alpha^Y, \alpha \in \mathcal{A}\} \)-martingale is a \( \{\mathcal{F}_\alpha \vee \mathcal{F}_\alpha^X, \alpha \in \mathcal{A}\} \)-martingale.

**Remark 2.4** In the temporally ordered setting, Buckdahn, Engelbert, and Rașcanu (2003) employ a similar martingale assumption.
Proof. Let \( \{M_\alpha, \alpha \in \mathcal{A}\} \) be a \( \{\mathcal{F}_t\} \)-martingale. For each \( \alpha \in \mathcal{A} \), there must exist a Borel function \( h_\alpha \) such that \( M_\alpha = h_\alpha(Y) \). Suppose \( \alpha_1 < \alpha_2 \). Then

\[
E[M_{\alpha_2} | \mathcal{F}_{\alpha_1}^{X} \vee \mathcal{F}_{\alpha_1}^{Y}] = E[h_{\alpha_2}(Y) | \mathcal{F}_{\alpha_1}^{X} \vee \mathcal{F}_{\alpha_1}^{Y}] = E[h_{\alpha_2}(Y) | \mathcal{F}_{\alpha_1}^{Y}] = M_{\alpha_1}.
\]

The proof of the converse is similar. \( \square \)

Note that (2.2) is equivalent to requiring that for each \( h \in B(S_2) \),

\[
\inf_{f \in B(B(S_1^1 \times B(S_2^2))} E[(h(Y) - f(X,Y))^2] = \inf_{f \in B(S_2)} E[(h(Y) - f(Y))^2],
\]

so compatibility is a property of the joint distribution of \((X,Y)\). Consequently, compatibility is a constraint on joint distributions. To emphasize the special role of compatibility, \( S_{\Gamma, \mathcal{C}, \nu} \) will denote the collection of joint distributions that satisfy the constraints in \( \Gamma \) and the \( \mathcal{C} \)-compatibility constraint.

Example 2.5 Let \( U \) be a process in \( D_{\mathbb{R}^d}[0, \infty) \), \( V \) an \( \mathbb{R}^m \)-valued semimartingale with respect to the filtration \( \{\mathcal{F}_t^{U,V}\} \), and \( H : D_{\mathbb{R}^d}[0, \infty) \to D_{\mathbb{M}^{d \times m}}[0, \infty) \) (\( \mathbb{M}^{d \times m} \) the space of \( d \times m \)-dimensional matrices) be Borel measurable and satisfy \( H(x, t) = H(x(\cdot \land t), t) \) for all \( x \in D_{\mathbb{R}^d}[0, \infty) \) and \( t \geq 0 \). Then \( X \) is a solution of

\[
X(t) = U(t) + \int_0^t H(X, s-)dV(s)
\]

if \( X \) is temporally compatible with \( Y = (U, V) \) and

\[
\lim_{n \to \infty} E[1 \land |X(t) - U(t) - \sum_k H(X, \frac{k}{n})(V(\frac{k+1}{n} \land t) - V(\frac{k}{n} \land t))] = 0, \quad t \geq 0.
\]

To prove pointwise (pathwise) uniqueness, we still need some way of comparing compatible solutions.

Definition 2.6 Let the random variables \( X_1, X_2, \) and \( Y \) be defined on the same probability space with \( X_1 \) and \( X_2 \) \( S_1 \)-valued and \( Y \) \( S_2 \)-valued. \( (X_1, X_2) \) are jointly \( \mathcal{C} \)-compatible with \( Y \) if

\[
E[h(Y) | \mathcal{F}_{\alpha_1}^{X_1} \vee \mathcal{F}_{\alpha_2}^{X_2} \vee \mathcal{F}_{\alpha_1}^{Y}^\prime] = E[h(Y) | \mathcal{F}_{\alpha_1}^{Y}], \quad \alpha \in \mathcal{A}, h \in B(S_2).
\]

(Note that if \( (X_1, X_2) \) are jointly \( \mathcal{C} \)-compatible with \( Y \), then each of \( X_1 \) and \( X_2 \) is \( \mathcal{C} \)-compatible with \( Y \).)

Pointwise uniqueness for jointly \( \mathcal{C} \)-compatible solutions holds if for every triple of processes \((X_1, X_2, Y)\) defined on the same probability space such that \( \mu_{X_1,Y}, \mu_{X_2,Y} \in S_{\Gamma, \mathcal{C}, \nu} \) and \((X_1, X_2)\) is jointly compatible with \( Y \), \( X_1 = X_2 \) a.s.

Uniqueness for jointly temporally compatible solutions is the usual kind of uniqueness considered for stochastic differential equations. The following lemma ensures that this kind of uniqueness is equivalent to the notion of pointwise uniqueness used in Theorem 1.5 and hence, for example, Theorem 1.5 implies the classical Yamada-Watanable theorem.
Lemma 2.7 Pointwise uniqueness for jointly $C$-compatible solutions in $S_{\Gamma,C,\nu}$ is equivalent to pointwise uniqueness in $S_{\Gamma,C,\nu}$.

Recall that for $\mu_1, \mu_2 \in S_{\Gamma,C,\nu}$ and $Y, \xi_1$, and $\xi_2$ independent, $Y$ with distribution $\nu$ and $\xi_1$ and $\xi_2$ uniform on $[0,1]$, there exist $G_1 : S_2 \times [0,1] \to S_1$ and $G_2 : S_2 \times [0,1] \to S_1$ such that $(G_1(Y, \xi_1), Y)$ has distribution $\mu_1$ and $(G_2(Y, \xi_2), Y)$ has distribution $\mu_2$.

Clearly pointwise uniqueness in $S_{\Gamma,C,\nu}$ implies pointwise uniqueness for jointly $C$-compatible solutions. The converse is a consequence of the following lemma.

Lemma 2.8 If $\mu_1, \mu_2 \in S_{\Gamma,C,\nu}$ and $(G_1(Y, \xi_1), Y)$ has distribution $\mu_1$ and $(G_2(Y, \xi_2), Y)$ has distribution $\mu_2$, where $\xi_1$ and $\xi_2$ are independent and independent of $Y$, then $G_1(Y, \xi_1), G_2(Y, \xi_2)$ are jointly compatible with $Y$.

In order to prove Lemma 2.8 we need the following technical lemma.

Lemma 2.9 $X$ is $C$-compatible with $Y$ if and only if for each $\alpha \in A$ and each $g \in B(B^S_{\alpha_1})$,

$$E[g(X)|Y] = E[g(X)|\mathcal{F}^Y_\alpha] \quad \text{a.s.}$$

(2.4)

Proof. Suppose that $X$ is $C$-compatible with $Y$. Then for $f \in B(S_2)$ and $g \in B(B^S_{\alpha_1})$,

$$E[f(Y)g(X)] = E[E[f(Y)|\mathcal{F}^X_\alpha \vee \mathcal{F}^Y_\alpha]g(X)]$$

$$= E[E[f(Y)|\mathcal{F}^Y_\alpha]g(X)]$$

$$= E[E[f(Y)|\mathcal{F}^Y_\alpha]E[g(X)|\mathcal{F}^Y_\alpha]]$$

$$= E[f(Y)E[g(X)|\mathcal{F}^Y_\alpha]],$$

and (2.4) follows. Conversely, for $f \in B(S_2)$, $g \in B(B^S_{\alpha_2})$, and $h \in B(B^S_{\alpha_2})$, we have

$$E[E[f(Y)|\mathcal{F}^Y_\alpha]g(X)h(Y)] = E[E[f(Y)|\mathcal{F}^Y_\alpha]E[g(X)|\mathcal{F}^Y_\alpha]h(Y)]$$

$$= E[f(Y)E[g(X)|Y]h(Y)]$$

$$= E[f(Y)g(X)h(Y)],$$

and compatibility follows. □

Proof.[of Lemma 2.8] For $f \in B(B^S_{\alpha_1})$, by the independence of $\xi_2$ from $(Y, \xi_1)$ and Lemma 2.9,

$$E[f(G_1(Y, \xi_1))|Y, \xi_2] = E[f(G_1(Y, \xi_1))|Y] = E[f(G_1(Y, \xi_1))|\mathcal{F}^Y_\alpha].$$

Consequently, for $f \in B(S_2)$, $g_1, g_2 \in B(B^S_{\alpha_2})$, and $h \in B(B^S_{\alpha_2})$,

$$E[f(Y)g_1(X_1)g_2(X_2)h(Y)]$$

$$= E[f(Y)E[g_1(X_1)|Y, \xi_2]g_2(X_2)h(Y)]$$

$$= E[f(Y)E[g_1(X_1)|\mathcal{F}^X_\alpha]g_2(X_2)h(Y)]$$

$$= E[E[f(Y)|\mathcal{F}^X_\alpha \vee \mathcal{F}^Y_\alpha]E[g_1(X_1)|\mathcal{F}^Y_\alpha]g_2(X_2)h(Y)]$$

$$= E[E[f(Y)|\mathcal{F}^Y_\alpha]E[g_1(X_1)|Y, \xi_2]g_2(X_2)h(Y)]$$

$$= E[E[f(Y)|\mathcal{F}^Y_\alpha]g_1(X_1)g_2(X_2)h(Y)],$$

giving the joint compatibility. □

Lemma 2.9 also gives the following result.
**Proposition 2.10** If $X$ is a strong, compatible solution, then $F^X_\alpha \subset F^Y_\alpha$ for each $\alpha \in \mathcal{A}$. In particular, in the temporal compatibility setting, $X$ is adapted to the filtration $\{F^Y_t\}$.

**Proof.** Since $X = F(Y)$, by (2.4), for each $g \in B(B^S_\alpha)$,

$$g(X) = g(F(Y)) = E[g(F(Y))|Y] = E[g(X)|Y] = E[g(X)|F^Y_t] \quad a.s.$$  

Consequently, $g(X)$ is $F^Y_\alpha$-measureable and hence $F^X_\alpha \subset F^Y_\alpha$. □

**Example 2.11** McKean-Vlasov limits lead naturally to stochastic differential equations of the form

$$X(t) = X(0) + \int_0^t \sigma(X(s), \mu_{X(s)})dW(s) + \int_0^t b(X(s), \mu_{X(s)})ds$$

where $\mu_{X(s)}$ is required to be the distribution of $X(s)$. Alexander Veretennikov raised the question of a Yamada-Watanabe type result for equations of this form. Setting $Y = (X(0), W)$ and requiring temporal compatibility, the set of joint solution measures $S_{\Gamma,\mathcal{C},\nu}$ may not be convex. Consequently, the results of [Kurtz (2007)] may not apply. Theorem 1.5, however, does not assume convexity of $S_{\Gamma,\mathcal{C},\nu}$ and consequently gives the desired result.

**References**

David Blackwell and Lester E. Dubins. An extension of Skorohod’s almost sure representation theorem. *Proc. Amer. Math. Soc.*, 89(4):691–692, 1983. ISSN 0002-9939. doi: 10.2307/2044607. URL http://dx.doi.org/10.2307/2044607

R. Buckdahn, H.-J. Engelbert, and A. Răşcanu. On weak solutions of backward stochastic differential equations. *Theory Probab. Appl.*, 49(1):16–50, 2005. URL http://dx.doi.org/10.1137/S0040585X97980877

A. S. Cherny. On the uniqueness in law and the pathwise uniqueness for stochastic differential equations. *Theory Probab. Appl.*, 46(3):406–419, 2003. ISSN 0040-361X.

H. J. Engelbert. On the theorem of T. Yamada and S. Watanabe. *Stochastics Stochastics Rep.*, 36(3-4):205–216, 1991. ISSN 1045-1129.

Jean Jacod. Weak and strong solutions of stochastic differential equations. *Stochastics*, 3 (3):171–191, 1980. ISSN 0090-9491.

Thomas G. Kurtz. The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities. *Electron. J. Probab.*, 12:951–965, 2007. ISSN 1083-6489. doi: 10.1214/EJP.v12-431. URL http://dx.doi.org/10.1214/EJP.v12-431

Toshio Yamada and Shinzo Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11:155–167, 1971.