Weak and strong solutions of general stochastic models[*]

Thomas G. Kurtz[†]

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. The usual formulation of stochastic differential equations driven by semimartingales does not require compatibility, so a notion of *partial compatibility* is introduced which does hold. Since compatibility implies partial compatibility, classical strong uniqueness results imply strong uniqueness for compatible solutions. Weak existence arguments typically give existence of compatible solutions (not just partially compatible solutions), and as in the original Yamada-Watanabe theorem, existence of strong solutions follows.

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 second reason for this rewrite is that the main observation ensuring 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 2.11.

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

[†]University of Wisconsin-Madison, USA.
 E-mail: kurtz@math.wisc.edu http://www.math.wisc.edu/~kurtz/

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 the kind of compactness verified is frequently 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,

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

$X(0)$ 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 $S_2 = \mathbb{R}^d \times C_{\mathbb{R}^d}[0, \infty)$ and $S_1 = C_{\mathbb{R}^d}[0, \infty)$.

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

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

Example 1.1. Suppose Γ_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 $|f_i(x, y)| \leq \psi$.

Let $0 \leq c(x, y) \leq \psi(x, y)$, and let Γ be the set of constraints obtained from Γ_0 by adding the requirement

$$\int c(x, y)\mu(dx \times dy) = \inf_{\mu' \in \mathcal{S}_{\Gamma_0,\nu}} \int c(x, y)\mu'(dx \times dy).$$

It is natural to ask if the infimum is achieved with X of the form $X = F(Y)$. □

In the terminology of Engelbert (1991) and Jacod (1980), $\mu \in \mathcal{S}_{\Gamma,\nu}$ is a *joint solution measure* for our model (Γ, ν). A *weak solution* (or simply a *solution*) for (Γ, ν) is any pair of random variables (X, Y) defined on any probability space such that Y has distribution ν and (X, Y) meets the constraints in Γ, that is, $\mu_{X,Y} \in \mathcal{S}_{\Gamma,\nu}$. We have the following definition for a strong solution.

Definition 1.2. *A solution (X, Y) for (Γ, ν) 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 ν. 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 ν 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 $\mu \in \mathcal{P}_\nu(S_1 \times S_2)$.*

  a) There exists a transition function η 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 ν and ξ is independent of Y and uniformly distributed on $[0, 1]$, $(G(Y, \xi), Y)$ has distribution μ.
  c) μ 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. □

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) $\mathcal{S}_{\Gamma,\nu} \neq \emptyset$, 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,*

$$f_i(X, Y) = 0 \quad a.s. \quad i \in \mathcal{I}, \tag{1.1}$$

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) since (1.1) does not involve any adaptedness requirements. Consequently, 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 $\mu_1, \mu_2 \in \mathcal{S}_{\Gamma,\nu}$, then there exist Borel measurable functions $G_1(y, u)$ and $G_2(y, u)$ on $S_2 \times [0, 1]$ such that for Y with distribution ν and ξ_1, ξ_2 uniform on $[0, 1]$, all independent, $(G_1(Y, \xi_1), Y)$ has distribution μ_1 and $(G_2(Y, \xi_2), Y)$ has distribution μ_2. By pointwise uniqueness,

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

From the independence of ξ_1 and ξ_2, it follows that there exists a Borel measurable F on S_2 such that $F(Y) = G_1(Y, \xi_1) = G_2(Y, \xi_2)$ a.s. (See Lemma A.2 of Kurtz (2007).)

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

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 $\mathcal{S}_{\Gamma,\nu}$ 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 $\mathcal{S}_{\Gamma,\nu}$ 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 $\mathcal{S}_{\Gamma,\nu}$ contains only one distribution and the corresponding solution is strong. Conversely, suppose every solution is a strong solution. If $\mu_1, \mu_2 \in \mathcal{S}_{\Gamma,\nu}$, then $\mu_0 = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 \in \mathcal{S}_{\Gamma,\nu}$. Let Y have distribution ν. Then there exist Borel Functions F_1 and F_2 such that $(F_1(Y), Y)$ has distribution μ_1 and $(F_2(Y), Y)$ has distribution μ_2. Let ξ 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 μ_0 and must satisfy $X = F(Y)$ a.s. for some F. Since ξ is independent of Y, we must have $F_1(Y) = F(Y) = F_2(Y)$ *a.s.*, giving pointwise uniqueness. $\square$

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 σ-algebra and $B(S)$ will denote the space of bounded, Borel measurable functions; if $\mathcal{M}$ is a σ-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 $D_{E_i}[0, \infty)$, be the Skorohod space of cadlag E_i-valued functions. Let Y be a process in $D_{E_2}[0, \infty)$. 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 $h \in B(D_{E_2}[0, \infty))$,*

$$E[h(Y)|\mathcal{F}_t^{X,Y}] = E[h(Y)|\mathcal{F}_t^Y] \tag{2.1}$$

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) which is basic to the statement of Theorem 8.3 of that paper which gives a version of the Yamada-Watanabe theorem for general stochastic differential equations driven by semimartingales. 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 $t \geq 0$. (See Lemma 2.4 below.)

We will consider a more general notion of compatibility. If $\mathcal{B}_\alpha^{S_1}$ is a sub-σ-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 $\{\{X \in D\} : D \in \mathcal{B}_\alpha^{S_1}\}$ is the complete sub-σ-algebra of $\mathcal{F}$ generated by $\{h(X) : h \in B(\mathcal{B}_\alpha^{S_1})\}$. $\mathcal{F}_\alpha^Y$ is defined similarly for a sub-σ-algebra $\mathcal{B}_\alpha^{S_2} \subset \mathcal{B}(S_2)$.

Definition 2.2. *Let $\mathcal{A}$ be an index set, and for each $\alpha \in \mathcal{A}$, let $\mathcal{B}_\alpha^{S_1}$ be a sub-σ-algebra of $\mathcal{B}(S_1)$ and $\mathcal{B}_\alpha^{S_2}$ be a sub-σ-algebra of $\mathcal{B}(S_2)$. The collection $\mathcal{C} \equiv \{(\mathcal{B}_\alpha^{S_1}, \mathcal{B}_\alpha^{S_2}) : \alpha \in \mathcal{A}\}$ 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 $h \in B(S_2)$ (or equivalently, each $h \in L^1(\nu)$),

$$E[h(Y)|\mathcal{F}_\alpha^X \vee \mathcal{F}_\alpha^Y] = E[h(Y)|\mathcal{F}_\alpha^Y] \tag{2.2}$$

Remark 2.3. *Temporal compatibility, as defined above, is a special case of compatibility, and we will reserve this terminology for the case in which $\{\mathcal{F}_t^X\}$ and $\{\mathcal{F}_t^Y\}$ are the complete filtrations generated by X and Y. Of course, in this setting $\mathcal{F}_t^{X,Y} = \mathcal{F}_t^X \vee \mathcal{F}_t^Y$.*

Compatibility conditions do arise that have index set $\mathcal{A} = [0, \infty)$ but which are not temporal compatibility. For example, for a time-change equation

$$X(t) = Y\left(\int_0^t \beta(X(s))ds\right),$$

the natural compatibility condition sets

$$\mathcal{F}_\alpha^Y = \text{ the completion of } \sigma(Y(u) : 0 \le u \le \alpha)$$

but takes

$$\mathcal{F}_\alpha^X = \text{ the completion of } \sigma(\{\int_0^t \beta(X(s))ds \le r\} : r \le \alpha, t \ge 0),$$

so that compatibility ensures $\tau(t) = \int_0^t \beta(X(s))ds$ is a stopping time with respect to the filtration $\{\mathcal{F}_\alpha^X \vee \mathcal{F}_\alpha^Y, \alpha \ge 0\}$.

Lemma 2.4. *Suppose that for each $\alpha \in \mathcal{A}$ there exist random variables (Y_α, Y^α) with values in some measurable space $R_\alpha \times R^\alpha$ such that $\sigma(Y) = \sigma(Y_\alpha, Y^\alpha)$, Y_α is $\mathcal{F}_\alpha^Y$-measurable, and Y^α is independent of $\mathcal{F}_\alpha^X \vee \mathcal{F}_\alpha^Y$. Then X is compatible with Y.*

Proof. If $h \in B(S_2)$, then there exist $h_\alpha \in B(R_\alpha \times R^\alpha)$ such that $h(Y) = h_\alpha(Y_\alpha, Y^\alpha)$ a.s. Then

$$
\begin{aligned}
E[h(Y)|\mathcal{F}_\alpha^X \vee \mathcal{F}_\alpha^Y] &= E[h_\alpha(Y_\alpha, Y^\alpha)|\mathcal{F}_\alpha^X \vee \mathcal{F}_\alpha^Y] \\
&= E[\int_{R^\alpha} h_\alpha(Y_\alpha, y)\mu_{Y^\alpha}(dy)|\mathcal{F}_\alpha^X \vee \mathcal{F}_\alpha^Y] \\
&= \int_{R^\alpha} h_\alpha(Y_\alpha, y)\mu_{Y^\alpha}(dy) \\
&= E[\int_{R^\alpha} h_\alpha(Y_\alpha, y)\mu_{Y^\alpha}(dy)|\mathcal{F}_\alpha^Y].
\end{aligned}
$$

$\square$

In the temporal setting, Buckdahn, Engelbert, and Răşcanu (2005) employ a condition that requires every $\{\mathcal{F}_t^Y\}$-martingale to be a $\{\mathcal{F}_t^{X,Y}\}$-martingale. More generally, $\{\mathcal{F}_\alpha^Y, \alpha \in \mathcal{A}\}$ is a *filtration* if $\mathcal{A}$ is partially ordered and $\alpha_1 \prec \alpha_2$ implies $\mathcal{F}_{\alpha_1}^Y \subset \mathcal{F}_{\alpha_2}^Y$. We consider the following condition.

Condition 2.5. $\{\mathcal{F}_\alpha^Y, \alpha \in \mathcal{A}\}$ and $\{\mathcal{F}_\alpha^X, \alpha \in \mathcal{A}\}$ are filtrations and every $\{\mathcal{F}_\alpha^Y\}$-martingale is a $\{\mathcal{F}_\alpha^Y \vee \mathcal{F}_\alpha^X\}$ martingale.

Lemma 2.6. If $\{\mathcal{F}_\alpha^Y, \alpha \in \mathcal{A}\}$ and $\{\mathcal{F}_\alpha^X, \alpha \in \mathcal{A}\}$ are filtrations, then $\mathcal{C}$-compatibility implies Condition 2.5.

Remark 2.7. *The earlier paper (Kurtz (2007)) and the original version of the current paper casually claimed equivalence of the martingale condition and compatibility. A referee has pointed out that the claim was not only casual, but false. Condition 2.5 gives an example of what we will call* partial *compatibility conditions, that is, (2.2) holds for a subset of $h \in L^1(\nu)$. Partial compatibility conditions will be discussed further in Section 3.*

Proof. Let $\{M(\alpha), \alpha \in \mathcal{A}\}$ be a $\{\mathcal{F}_\alpha^Y\}$-martingale. For each $\alpha \in \mathcal{A}$, there exists a Borel function h_α such that $M(\alpha) = h_\alpha(Y)$ a.s. Suppose $\alpha_1 \prec \alpha_2$. Then

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

$\square$

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

$$\inf_{f \in B(\mathcal{B}_\alpha^{S_1} \times \mathcal{B}_\alpha^{S_2})} E[(h(Y) - f(X,Y))^2] = \inf_{f \in B(\mathcal{B}_\alpha^{S_2})} E[(h(Y) - f(Y))^2], \tag{2.3}$$

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, $\mathcal{S}_{\Gamma, \mathcal{C}, \nu}$ will denote the collection of joint distributions that satisfy the constraints in Γ and the $\mathcal{C}$-compatibility constraint.

Example 2.8. 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 $H(x, t) = H(x(\cdot \wedge t), t)$ for all $x \in D_{\mathbb{R}^d}[0, \infty)$ and $t \geq 0$. Then X is defined to be a solution of

$$X(t) = U(t) + \int_0^t H(X, s-)dV(s) \tag{2.4}$$

if X is temporally compatible with $Y = (U, V)$ (ensuring that the stochastic integral exists) and

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

Note that this definition assumes more regularity than is necessary or is assumed in Jacod (1980).

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

Definition 2.9. *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*

$$E[h(Y)|\mathcal{F}_\alpha^{X_1} \vee \mathcal{F}_\alpha^{X_2} \vee \mathcal{F}_\alpha^Y] = E[h(Y)|\mathcal{F}_\alpha^Y], \quad \alpha \in \mathcal{A}, h \in B(S_2). \tag{2.5}$$

(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 $\mu_{X_1,Y}, \mu_{X_2,Y} \in \mathcal{S}_{\Gamma, \mathcal{C}, \nu}$ and (X_1, X_2) is jointly $\mathcal{C}$-compatible with Y, $X_1 = X_2$ a.s.*

With reference to Lemma 2.4, uniqueness for jointly temporally compatible solutions is the usual kind of uniqueness considered for stochastic differential equations driven by Brownian motion, Lévy processes, and/or Poisson random measures. For example, let $Y = (X(0), Z)$, where Z is a Lévy process. Consider the equation

$$X(t) = X(0) + \int_0^t H(X(s-))dZ(s),$$

where we require X and Z to be adapted to a filtration $\{\mathcal{F}_t\}$ such that $Z(t + \cdot) - Z(t)$ is independent of $\mathcal{F}_t$, $t \geq 0$. If there exist two such solutions with $X_1(0) = X_2(0) = X(0)$ adapted to $\{\mathcal{F}_t\}$, then since $\mathcal{F}_t^{X_1} \vee \mathcal{F}_t^{X_2} \vee \mathcal{F}_t^Z \subset \mathcal{F}_t$,

$$
\begin{aligned}
E[h(Z(t + \cdot) &- Z(t), Z(\cdot \wedge t))|\mathcal{F}_t^{X_1} \vee \mathcal{F}_t^{X_2} \vee \mathcal{F}_t^Z] \\
&= E[E[h(Z(t + \cdot) - Z(t), Z(\cdot \wedge t))|\mathcal{F}_t]|\mathcal{F}_t^{X_1} \vee \mathcal{F}_t^{X_2} \vee \mathcal{F}_t^Z] \\
&= E[\int h(z, Z(\cdot \wedge t))\mu_{Z(t+\cdot)-Z(t)}(dz)|\mathcal{F}_t^{X_1} \vee \mathcal{F}_t^{X_2} \vee \mathcal{F}_t^Z] \\
&= \int h(z, Z(\cdot \wedge t))\mu_{Z(t+\cdot)-Z(t)}(dz) \\
&= E[\int h(z, Z(\cdot \wedge t))\mu_{Z(t+\cdot)-Z(t)}(dz)|\mathcal{F}_t^Z \vee \sigma(X(0))],
\end{aligned}
$$

which gives the joint compatibility of X_1 and X_2 with $(X(0), Z)$.

The following lemma ensures that pointwise uniqueness of jointly compatible solutions is equivalent to the notion of pointwise uniqueness used in Theorem 1.5 and hence, for example, Theorem 1.5 implies the classical Yamada-Watanabe theorem.

Lemma 2.10. *Pointwise uniqueness for jointly $\mathcal{C}$-compatible solutions in $\mathcal{S}_{\Gamma,\mathcal{C},\nu}$ is equivalent to pointwise uniqueness in $\mathcal{S}_{\Gamma,\mathcal{C},\nu}$.*

Recall that for $\mu_1, \mu_2 \in \mathcal{S}_{\Gamma,\mathcal{C},\nu}$ and Y, ξ_1, and ξ_2 independent, Y with distribution ν and ξ_1 and ξ_2 uniform on $[0,1]$, there exist Borel measurable $G_1 : S_2 \times [0,1] \to S_1$ and $G_2 : S_2 \times [0,1] \to S_1$ such that $(G_1(Y,\xi_1), Y)$ has distribution μ_1 and $(G_2(Y,\xi_2), Y)$ has distribution μ_2.

Clearly pointwise uniqueness in $\mathcal{S}_{\Gamma,\mathcal{C},\nu}$ implies pointwise uniqueness for jointly $\mathcal{C}$-compatible solutions. The converse follows by repeating the reasoning in the proof of Theorem 1.5 now using the following lemma.

Lemma 2.11. *If $\mu_1, \mu_2 \in \mathcal{S}_{\Gamma,\mathcal{C},\nu}$ and $(G_1(Y,\xi_1), Y)$ has distribution μ_1 and $(G_2(Y,\xi_2), Y)$ has distribution μ_2, where ξ_1 and ξ_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 2.11, we need the following technical lemma.

Lemma 2.12. *X is $\mathcal{C}$-compatible with Y if and only if for each $\alpha \in \mathcal{A}$ and each $g \in B(\mathcal{B}_\alpha^{S_1})$,*

$$E[g(X)|Y] = E[g(X)|\mathcal{F}_\alpha^Y] \tag{2.6}$$

Proof. Suppose that X is $\mathcal{C}$-compatible with Y. Then for $f \in B(S_2)$ and $g \in B(\mathcal{B}_\alpha^{S_1})$,

$$
\begin{aligned}
E[f(Y)g(X)] &= E[E[f(Y)|\mathcal{F}_\alpha^X \vee \mathcal{F}_\alpha^Y]g(X)] \\
&= E[E[f(Y)|\mathcal{F}_\alpha^Y]g(X)] \\
&= E[E[f(Y)|\mathcal{F}_\alpha^Y]E[g(X)|\mathcal{F}_\alpha^Y]] \\
&= E[f(Y)E[g(X)|\mathcal{F}_\alpha^Y]],
\end{aligned}
$$

and (2.6) follows. Conversely, for $f \in B(S_2)$, $g \in B(\mathcal{B}_\alpha^{S_1})$, and $h \in B(\mathcal{B}_\alpha^{S_2})$, we have

$$
\begin{aligned}
E[E[f(Y)|\mathcal{F}_\alpha^Y]g(X)h(Y)] &= E[E[f(Y)|\mathcal{F}_\alpha^Y]E[g(X)|\mathcal{F}_\alpha^Y]h(Y)] \\
&= E[f(Y)E[g(X)|Y]h(Y)] \\
&= E[f(Y)g(X)h(Y)],
\end{aligned}
$$

and compatibility follows. $\qquad\square$

Proof. **[of Lemma 2.11]** For $g \in B(\mathcal{B}_\alpha^{S_1})$, by the independence of ξ_2 from (Y, ξ_1) (and hence from $X_1 = G_1(Y, \xi_1)$) and Lemma 2.12,

$$
E[g(X_1)|Y, \xi_2] = E[g(X_1)|Y] = E[g(X_1)|\mathcal{F}_\alpha^Y]. \tag{2.7}
$$

Consequently, for $X_1 = G_1(Y, \xi_1)$, $X_2 = G_2(Y, \xi_2)$, $f \in B(S_2)$, $g_1, g_2 \in B(\mathcal{B}_\alpha^{S_1})$, and $g_3 \in B(\mathcal{B}_\alpha^{S_2})$,

$$
\begin{aligned}
&E[f(Y)g_1(X_1)g_2(X_2)g_3(Y)] \\
&\quad = E[f(Y)E[g_1(X_1)|Y, \xi_2]g_2(X_2)g_3(Y)] \\
&\quad = E[f(Y)E[g_1(X_1)|\mathcal{F}_\alpha^Y]g_2(X_2)g_3(Y)] \\
&\quad = E[E[f(Y)|\mathcal{F}_\alpha^{X_2} \vee \mathcal{F}_\alpha^Y]E[g_1(X_1)|\mathcal{F}_\alpha^Y]g_2(X_2)g_3(Y)] \\
&\quad = E[E[f(Y)|\mathcal{F}_\alpha^Y]E[g_1(X_1)|Y, \xi_2]g_2(X_2)g_3(Y)] \\
&\quad = E[E[f(Y)|\mathcal{F}_\alpha^Y]g_1(X_1)g_2(X_2)g_3(Y)],
\end{aligned}
$$

giving the joint compatibility. $\qquad\square$

Lemma 2.12 also gives the following result.

Proposition 2.13. *If X is a strong, compatible solution, then $\mathcal{F}_\alpha^X \subset \mathcal{F}_\alpha^Y$ for each $\alpha \in \mathcal{A}$. (In particular, in the temporal compatibility setting, X is adapted to the filtration $\{\mathcal{F}_t^Y\}$.) Conversely, if $\mathcal{F}_\alpha^X \subset \mathcal{F}_\alpha^Y$ for each $\alpha \in \mathcal{A}$ and $\sigma(X) \subset \vee_{\alpha \in \mathcal{A}}\mathcal{F}_\alpha^X$, then X is a strong, compatible solution.*

Proof. Since $X = F(Y)$, by (2.6), for each $g \in B(\mathcal{B}_\alpha^{S_1})$,

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

Consequently, $g(X)$ is $\mathcal{F}_\alpha^Y$-measurable and hence $\mathcal{F}_\alpha^X \subset \mathcal{F}_\alpha^Y$.

Conversely, the assumption that $\mathcal{F}_\alpha^X \subset \mathcal{F}_\alpha^Y$ for each $\alpha \in \mathcal{A}$ implies X is compatible with Y, and the additional assumption implies

$$
\sigma(X) \subset \vee_{\alpha \in \mathcal{A}}\mathcal{F}_\alpha^X \subset \vee_{\alpha \in \mathcal{A}}\mathcal{F}_\alpha^Y \subset \sigma(Y),
$$

so there exists a Borel measurable function F such that $X = F(Y)$ a.s. $\qquad\square$

Example 2.14. McKean-Vlasov limits lead naturally to stochastic differential equations of the form

$$
X(t) = X(0) + \int_0^t \sigma(X(s), \mu_{X(s)})dW(s) + \int_0^t b(X(s), \mu_{X(s)})ds \tag{2.8}
$$

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 $\mathcal{S}_{\Gamma, \mathcal{C}, \nu}$ may not be convex. Consequently, the results of Kurtz (2007) may not apply. Theorem 1.5, however, does not assume convexity of $\mathcal{S}_{\Gamma, \mathcal{C}, \nu}$, and hence weak existence and pathwise uniqueness imply the existence of a strong solution of (2.8).

3 Partial compatibility and existence of compatible solutions.

Let $\mathcal{H} \subset B(S_2)$ (or $\mathcal{H} \subset L^1(\nu)$). We will say that a random variable X is $(\mathcal{C}, \mathcal{H})$-*partially compatible* with Y if (2.2) holds for each $h \in \mathcal{H}$ but not necessarily for all $h \in B(S_2)$. Similarly, we define *joint* $(\mathcal{C}, \mathcal{H})$-*partial compatibility* by requiring the identity in (2.5) to hold for $h \in \mathcal{H}$. As above, let $\mathcal{S}_{\Gamma, (\mathcal{C}, \mathcal{H}), \nu}$ denote the collection of joint distributions for partially compatible solutions. The analog of the Engelbert theorem ((b) implies (a) in Theorem 1.5) follows as before.

Theorem 3.1. *Suppose there exists a strong solution in $\mathcal{S}_{\Gamma, (\mathcal{C}, \mathcal{H}), \nu}$ and joint uniqueness in law holds. Then pointwise uniqueness holds.*

We could prove the analog of the Yamada-Watanabe theorem ((a) implies (b) in Theorem 1.5) for $\mathcal{S}_{\Gamma, (\mathcal{C}, \mathcal{H}), \nu}$ the same way we handled $\mathcal{S}_{\Gamma, \mathcal{C}, \nu}$ if the analog of Lemma 2.11 held. Unfortunately, that is not in general the case.

Example 3.2. Let $\zeta_1, \ldots, \zeta_4$ be independent with distribution $P\{\zeta_i = 1\} = P\{\zeta_i = -1\} = \frac{1}{2}$, and let
$$Y = (Y_1, Y_2, Y_3, Y_4) = (\zeta_1\zeta_2, \zeta_2\zeta_3, \zeta_3\zeta_4, \zeta_4\zeta_1).$$
Note that any three of the components are independent but the four are not. Assume that the index set $\mathcal{A}$ consists of a single element α. Let $\mathcal{F}_\alpha^Y = \sigma(Y_1)$, and for ξ independent of Y and uniformly distributed on $[0, 1]$, let
$$X = G(Y, \xi) = \mathbf{1}_{\{\xi < \frac{1}{2}\}} Y_2 + \mathbf{1}_{\{\xi \geq \frac{1}{2}\}} Y_3,$$
and $\mathcal{F}_\alpha^X = \sigma(G(Y, \xi))$. For $h_0(Y) = Y_4$,
$$E[h_0(Y)|\mathcal{F}_\alpha^Y \vee \mathcal{F}_\alpha^X] = 0 = E[h_0(Y)|\mathcal{F}_\alpha^Y],$$
so X is $(\mathcal{C}, \mathcal{H})$-partially compatible with Y for $\mathcal{H} = \{h_0\}$. However, if ξ_1 and ξ_2 are independent, uniform $[0, 1]$ random variables and we define
$$X_1 = G(Y, \xi_1) \text{ and } X_2 = G(Y, \xi_2),$$
then
$$E[h_0(Y)|\mathcal{F}_\alpha^Y \vee \mathcal{F}_\alpha^{X_1} \vee \mathcal{F}_\alpha^{X_2}] = \mathbf{1}_{\{X_1 \neq X_2\}} Y_1 X_1 X_2 + \frac{1}{3} \mathbf{1}_{\{X_1 = X_2\}} Y_1,$$
and the corresponding joint partial compatibility condition fails.

We do have the following modification of Lemma 2.11 that gives the desired coupling if one of the solutions is compatible.

Lemma 3.3. *If $\mu_1 \in \mathcal{S}_{\Gamma, \mathcal{C}, \nu}$ and $\mu_2 \in \mathcal{S}_{\Gamma, (\mathcal{C}, \mathcal{H}), \nu}$, and $(G_1(Y, \xi_1), Y)$ has distribution μ_1 and $(G_2(Y, \xi_2), Y)$ has distribution μ_2, where ξ_1 and ξ_2 are independent and independent of Y, then $X_1 = G_1(Y, \xi_1)$ and $X_2 = G_2(Y, \xi_2)$ are jointly $(\mathcal{C}, \mathcal{H})$-partially compatible with Y.*

Proof. Since X_1 is compatible with Y, (2.7) still holds. Consequently, for $h \in \mathcal{H}$, $g_1, g_2 \in B(\mathcal{B}_\alpha^{S_1})$, and $g_3 \in B(\mathcal{B}_\alpha^{S_2})$, as in the proof of Lemma 2.11,

$$
\begin{aligned}
E[h(Y)&g_1(X_1)g_2(X_2)g_3(Y)] \\
&= E[h(Y)E[g_1(X_1)|Y, \xi_2]g_2(X_2)g_3(Y)] \\
&= E[h(Y)E[g_1(X_1)|\mathcal{F}_\alpha^Y]g_2(X_2)g_3(Y)] \\
&= E[E[h(Y)|\mathcal{F}_\alpha^{X_2} \vee \mathcal{F}_\alpha^Y]E[g_1(X_1)|\mathcal{F}_\alpha^Y]g_2(X_2)g_3(Y)] \\
&= E[E[h(Y)|\mathcal{F}_\alpha^Y]E[g_1(X_1)|Y, \xi_2]g_2(X_2)g_3(Y)] \\
&= E[E[h(Y)|\mathcal{F}_\alpha^Y]g_1(X_1)g_2(X_2)g_3(Y)],
\end{aligned}
$$

giving the joint $(\mathcal{C}, \mathcal{H})$-partial compatibility. $\qquad \square$

The construction in Lemma 3.3 gives the proof of the following theorem which states that weak existence of a compatible solution and pointwise uniqueness for jointly $(\mathcal{C}, \mathcal{H})$-partially compatible solutions implies that the only solution is a strong, compatible solution.

Theorem 3.4. *Suppose that if X_1 and X_2 are jointly $(\mathcal{C}, \mathcal{H})$-partially compatible with Y and $\mu_{X_1 Y}, \mu_{X_2, Y} \in \mathcal{S}_{\Gamma, (\mathcal{C}, \mathcal{H}), \nu}$, then $X_1 = X_2$ a.s. and that there exists a compatible solution, that is, $\mathcal{S}_{\Gamma, \mathcal{C}, \nu} \neq \emptyset$. Then there exists a unique, partially compatible solution and it is strong and compatible.*

Proof. The uniqueness assumption for jointly partially compatible solutions implies uniqueness for jointly compatible solutions. Consequently, there exists a unique, strong, compatible solution, $X = F(Y)$. But Lemma 3.3 implies that every partially compatible solution and the unique strong, compatible solution can be constructed to be jointly $(\mathcal{C}, \mathcal{H})$-partially compatible and hence the partially compatible solution must also be $F(Y)$. $\square$

Theorem 3.4 is relevant not only under Condition 2.5 but also for the general stochastic differential equation given in Example 2.8. Uniqueness results for equations of the form (2.4) are usually proved under the assumption that solutions X_1 and X_2 and $Y = (U, V)$ are adapted to a filtration $\{\mathcal{F}_t\}$ under which V is a semimartingale. V can always be written as $V = M + A$, where M is a local martingale with jumps bounded by 1 and A is a finite variation process. The localizing sequence for M can be taken to be $\tau_n = \inf\{t : \sup_{s \leq t} |M(s)| \geq n\}$, and an appropriate joint partial compatibility condition follows from the observation that for $t > s$,

$$
\begin{aligned}
E[M(t \wedge \tau_n)|\mathcal{F}_s^{X_1} \vee \mathcal{F}_s^{X_2} \vee \mathcal{F}_s^Y] &= E[E[M(t \wedge \tau_n)|\mathcal{F}_s]|\mathcal{F}_s^{X_1} \vee \mathcal{F}_s^{X_2} \vee \mathcal{F}_s^Y] \\
&= M(s \wedge \tau_n) \\
&= E[M(t \wedge \tau_n)|\mathcal{F}_s^Y].
\end{aligned}
$$

Consequently, pathwise uniqueness results in settings of this form imply pathwise uniqueness for jointly compatible solutions.

To apply Theorem 3.4 when pointwise uniqueness is known under partial compatibility conditions still requires existence of a compatible solution. The following lemma gives a general approach to the required existence.

Lemma 3.5. *Suppose there exist $C_\alpha^{S_2} \subset C_b(S_2)$ and $C_\alpha^{S_1} \subset C_b(S_1)$ such that $\mathcal{B}_\alpha^{S_2} = \sigma(g \in C_\alpha^{S_2})$ and $\mathcal{B}_\alpha^{S_1} = \sigma(g \in C_\alpha^{S_1})$. (Without loss of generality, we can assume $C_\alpha^{S_1}$ and $C_\alpha^{S_2}$ are algebras.) Suppose $(X_n, Y) \in S_1 \times S_2$, X_n is $\mathcal{C}$-compatible with Y, $(X_n, Y) \Rightarrow (X, Y)$. Then X is $\mathcal{C}$-compatible with Y.*

Remark 3.6. *With reference to the continuous mapping theorem (for example, Ethier and Kurtz (1986), Corollary 3.1.9), the continuity assumption on the functions generating $\mathcal{B}_\alpha^{S_1}$ and $\mathcal{B}_\alpha^{S_2}$ can be weakened. For $\mathcal{B}_\alpha^{S_1}$, it is enough for the functions g to be continuous almost everywhere with respect to μ_X, and for $\mathcal{B}_\alpha^{S_2}$, the functions g only need to be continuous almost everywhere with respect to μ_Y. This observation is particularly relevant for cadlag processes since the evaluation function $x \in D_E[0, \infty) \to x(t) \in E$ is not continuous, but it will be almost everywhere continuous for the process of interest provided t is not a fixed point of discontinuity, that is, provided $P\{X(t) \neq X(t-)\} = 0$.*

In many settings, natural approximations for a solution will satisfy $X_n = F_n(Y)$ and $\mathcal{F}_\alpha^{X_n} \subset \mathcal{F}_\alpha^Y$ and hence will be strong, compatible solutions of approximating models. (See Proposition 2.13.)

Proof. For $f \in C_b(S_2)$, $g_1 \in C_\alpha^{S_1}$ and $g_2 \in C_\alpha^{S_2}$

$$E[f(Y)g_1(X_n)g_2(Y)] = E[E[f(Y)|\mathcal{F}_\alpha^Y]g_1(X_n)g_2(Y)].$$

Since $C_b(S_2)$ is dense in $L^1(\nu)$, for each α and $\epsilon > 0$, there exists $f_{\alpha,\epsilon} \in C_b(S_2)$ such that

$$E[|E[f(Y)|\mathcal{F}_\alpha^Y] - f_{\alpha,\epsilon}(Y)|] \leq \epsilon.$$

Consequently, it follows that

$$
\begin{aligned}
E[f(Y)g_1(X)g_2(Y)] &= \lim_{n\to\infty} E[f(Y)g_1(X_n)g_2(Y)] \\
&= \lim_{n\to\infty} E[E[f(Y)|\mathcal{F}_\alpha^Y]g_1(X_n)g_2(Y)] \\
&= \lim_{\epsilon\to 0} \lim_{n\to\infty} E[f_{\alpha,\epsilon}(Y)g_1(X_n)g_2(Y)] \\
&= \lim_{\epsilon\to 0} E[f_{\alpha,\epsilon}(Y)g_1(X)g_2(Y)] \\
&= E[E[f(Y)|\mathcal{F}_\alpha^Y]g_1(X)g_2(Y)]
\end{aligned}
$$

verifying compatibility. □

Note that in the proof of the above lemma, we use the fact that Y, or more precisely, the distribution of Y, does not depend on n in order to obtain the $f_{\alpha,\epsilon}$.

Problems do arise in which input processes have fixed points of discontinuity and the application of Lemma 3.5 is problematic even with the observation made in Remark 3.6. The following definition of RC-compatibility (or more precisely, RC-temporal compatibility) avoids this problem. It looks strange, but Lemma 3.8 shows that it is equivalent to a more natural assumption. $M_E[0,\infty)$ denotes the collection of Borel measurable functions $x : [0,\infty) \to E$. S_i could be $D_{E_i}[0,\infty)$ under the usual Skorohod topology, but other spaces can be useful. (See Example 3.10.)

Definition 3.7. *Let $\mathcal{A} = \{(t,\epsilon) : t \in [0,\infty), \epsilon > 0\}$, $S_1 \subset M_{E_1}[0,\infty)$, and $S_2 \subset M_{E_2}[0,\infty)$. For $\alpha = (t,\epsilon)$, define*

$$C_\alpha^{S_2} = \{\int_s^{s+r} g(x(u))du : s \leq t, 0 < r < \epsilon, g \in C_b(E_2)\}$$

and

$$C_\alpha^{S_1} = \{\int_{(s-r)\vee 0}^s g(x(u))ds : s \leq t, 0 < r < \epsilon, g \in C_b(E_1)\},$$

and set $\mathcal{B}_\alpha^{S_2} = \sigma(g \in C_\alpha^{S_2})$ and $\mathcal{B}_\alpha^{S_1} = \sigma(g \in C_\alpha^{S_1})$. Then $\mathcal{C}_{RC} \equiv \{(\mathcal{B}_\alpha^{S_1}, \mathcal{B}_\alpha^{S_2}) : \alpha \in \mathcal{A}\}$ defines the RC-compatibility *structure (RC for "right continuous") on (S_1, S_2).*

Note that $C_\alpha^{S_1}$ and $C_\alpha^{S_2}$ differ not only in the choice of range spaces E_1 and E_2 but also in the collections of time intervals determining the integrals. If $S_1 = D_{E_1}[0,\infty)$ and $S_2 = D_{E_2}[0,\infty)$, then $C_\alpha^{S_1}$ and $C_\alpha^{S_2}$ are collections of continuous functions and Lemma 3.5 applies to RC-compatibility.

Assume that X and Y are right continuous, and let $\{\mathcal{F}_t^X\}$ and $\{\mathcal{F}_t^Y\}$ denote their natural filtrations. Note that for $t > 0$, $\mathcal{F}_{(t,\epsilon)}^X = \mathcal{F}_{t-}^X \equiv \vee_{s<t}\mathcal{F}_s^X$, $\cap_{\epsilon>0}\mathcal{F}_{(t,\epsilon)}^Y = \mathcal{F}_{t+}^Y \equiv \cap_{s>t}\mathcal{F}_s^Y$, and $\mathcal{F}_{(t,\epsilon)}^Y = \mathcal{F}_{(t+\epsilon)-}^Y$. We have the following lemma.

Lemma 3.8. *Let X be a right continuous, E_1-valued process and Y be a right continuous, E_2-valued process. Then X is RC-compatible with Y if and only if*

$$E[h(Y)|\mathcal{F}_{t+}^Y \vee \mathcal{F}_{t-}^X] = E[h(Y)|\mathcal{F}_{t+}^Y] \tag{3.1}$$

for all $t > 0$.

Proof. Since $\mathcal{F}^X_{(t,\epsilon)} = \mathcal{F}^X_{t-}$, RC-compatibility implies

$$E[h(Y)|\mathcal{F}^Y_{(t,\epsilon)} \vee \mathcal{F}^X_{t-}] = E[h(Y)|\mathcal{F}^Y_{(t,\epsilon)}].$$

Taking the limit $\epsilon \to 0$, we have

$$E[h(Y)| \cap_{\epsilon>0} (\mathcal{F}^Y_{(t,\epsilon)} \vee \mathcal{F}^X_{t-})] = E[h(Y)|\mathcal{F}^Y_{t+}].$$

Since $\cap_{\epsilon>0}(\mathcal{F}^Y_{(t,\epsilon)} \vee \mathcal{F}^X_{t-}) \supset \mathcal{F}^Y_{t+} \vee \mathcal{F}^X_{t-} \supset \mathcal{F}^Y_{t+}$, conditioning both sides on $\mathcal{F}^Y_{t+} \vee \mathcal{F}^X_{t-}$ gives (3.1).

Now assuming (3.1) holds for all $t > 0$, we have

$$E[h(Y)|\mathcal{F}^Y_{(t+s)+} \vee \mathcal{F}^X_{(t+s)-}] = E[h(Y)|\mathcal{F}^Y_{(t+s)+}],$$

and letting $s \to \epsilon-$, we have

$$E[h(Y)| \vee_{s<\epsilon} (\mathcal{F}^Y_{(t+s)+} \vee \mathcal{F}^X_{(t+s)-})] = E[h(Y)|\mathcal{F}^Y_{(t+\epsilon)-}] = E[h(Y)|\mathcal{F}^Y_{(t,\epsilon)}]. \qquad (3.2)$$

Since

$$\vee_{s<\epsilon}\mathcal{F}^Y_{(t+s)+} \vee \mathcal{F}^X_{(t+s)-} \supset \mathcal{F}^Y_{(t+\epsilon)-} \vee \mathcal{F}^X_{(t+\epsilon)-} \supset \mathcal{F}^Y_{(t,\epsilon)} \vee \mathcal{F}^X_{(t,\epsilon)},$$

conditioning both sides of (3.2) on $\mathcal{F}^Y_{(t,\epsilon)} \vee \mathcal{F}^X_{(t,\epsilon)}$ gives the desired result. $\qquad\square$

Example 3.9. An Euler approximation gives a natural approach to proving existence of compatible or RC-compatible solutions for

$$X(t) = U(t) + \int_0^t H(X, s-)dV(s). \qquad (3.3)$$

Set $\eta_n(t) = \frac{[nt]}{n}$, and et $U_n = U \circ \eta_n$ and $V_n = V \circ \eta_n$. Then existence of a solution X_n of

$$X_n(t) = U_n(t) + \int_0^t H(X_n, s-)dV_n(s), \qquad (3.4)$$

is immediate and X_n is adapted to $\{\mathcal{F}^Y_t\}$. It follows that X_n is both temporally compatible and RC-compatible with Y. Theorem 5.4 of Kurtz and Protter (1991) gives conditions on H that ensure the convergence of (U_n, V_n, X_n) to (U, V, X) satisfying (3.3). Lemma 3.5 then ensures that X is temporally compatible with $Y = (U, V)$, if Y has no fixed points of discontinuity, or at least RC-compatible with Y. The constructions of Jacod and Mémin (1980/81) and Lebedev (1983) should give compatible solutions under different assumptions.

Example 3.10. Let $T > 0$ and $Y = (U, V)$ be a process in $D_{\mathbb{R}^m \times \mathbb{R}^d}[0, T]$. Let f be a measurable function

$$f : [0, T] \times D_{\mathbb{R}^m}[0, T] \times D_{\mathbb{R}^d}[0, T] \to \mathbb{R}^m$$

satisfying $f(t, x, v) = f(t, x(\cdot \vee t), v)$ for each $(t, x, v) \in [0, T] \times D_{\mathbb{R}^m}[0, T] \times D_{\mathbb{R}^d}[0, T]$. Following Buckdahn, Engelbert, and Răşcanu (2005), we consider the backward stochastic differential equation

$$X(t) = U(t) + E[\int_t^T f(s, X, V)ds|\mathcal{F}^Y_t \vee \mathcal{F}^X_t],$$

where Buckdahn et al. (2005) requires Condition 2.5. We will require X to be temporally compatible with Y, or if Y has fixed points of discontinuity, that X be RC-compatible

with Y. Setting $X_n(t) = U(T)$ for $t \geq T$, there exist solutions to the approximating problems

$$X_n(t) = U(t) + E[\int_t^T f(s, X_n(\cdot + \frac{1}{n}), V) ds | \mathcal{F}_t^Y].$$

Assume that $|f(s, x, v)| \leq g(s, v)$ and $E[\int_0^T g(s, V) ds] < \infty$. Set

$$Z_n(t) = E[\int_t^T f(s, X_n(\cdot + \frac{1}{n}), V) ds | \mathcal{F}_t^Y].$$

Recalling the definition of conditional variation, we have

$$\mathcal{V}_T(Z_n) \equiv \sup_{\{t_i\}} E[\sum_i |E[Z_n(t_{i+1}) - Z_n(t_i) | \mathcal{F}_{t_i}^Y]|] \leq E[\int_0^T g(s, V) ds],$$

where the $\sup$ is over all partitions of $[0, T]$. We also have

$$\sup_{0 \leq t \leq T} |Z_n(t)| \leq \sup_{0 \leq t \leq T} E[\int_0^T g(s, V) ds | \mathcal{F}_t^Y] < \infty \quad a.s.,$$

so the sequence $\{Z_n\}$ satisfies the Meyer-Zheng conditions (see Meyer and Zheng (1984); Kurtz (1991)), or more precisely, $\{Z_n\}$ is relatively compact in the Jakubowski topology (see Jakubowski (1997)). The Jakubowski topology is not metrizable, but versions of the Prohorov theorem and the Skorohod representation theorem still hold. See Theorem 1.1 of Jakubowski (1997). We will denote the space of cadlag functions under the Jakubowski topology by $D_E^{\mathcal{J}}[0, T]$.

Convergence in the Jakubowski topology implies convergence in measure, that is convergence in the metric $d_m(x, y) = \int_0^T |x(s) - y(s)| \wedge 1 ds$ which is used in the original paper, Meyer and Zheng (1984), and in Buckdahn et al. (2005). Relative compactness of $\{Z_n\}$ in $D_{\mathbb{R}^m}^{\mathcal{J}}[0, T]$ implies relative compactness of (Z_n, Y) in $D_{\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^d}^{\mathcal{J}}[0, T]$. In contrast to the Skorohod topology (that is, the Skorohod J_1 topology),

$$D_{\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^d}^{\mathcal{J}}[0, T] = D_{\mathbb{R}^m}^{\mathcal{J}}[0, T] \times D_{\mathbb{R}^m \times \mathbb{R}^d}^{\mathcal{J}}[0, T].$$

Addition is continuous in the Jakubowski topology, so if (Z_n, Y) converges, then setting $X_n = U + Z_n$, (X_n, Z_n, Y) converges. If X_n converges to X, then $X_n(\cdot + \frac{1}{n})$ converges to X and for all but at most countably many t, $X_n(t)$ converges to $X(t)$.

For each $t \in [0, T]$, assume that the mapping

$$(x, v) \in D_{\mathbb{R}^m \times \mathbb{R}^d}^{\mathcal{J}}[0, T] \to \int_t^T f(s, x, v) ds \in \mathbb{R}$$

is continuous. Assume that we have selected a subsequence such that $(X_n, Y) \Rightarrow (X, Y)$. By Theorem 3.11 of Jakubowski (1997) there exists a countable set D such that for $\{t_i\} \subset [0, T] \setminus D$

$$(X_n(t_1), \ldots, X_n(t_k), Y(t_1), \ldots, Y(t_k), X_n, Y) \Rightarrow (X(t_1), \ldots, X(t_k), Y(t_1), \ldots, Y(t_k), X_n, Y)$$

in $(\mathbb{R}^m)^k \times (\mathbb{R}^{m+d})^k \times D_{\mathbb{R}^m \times \mathbb{R}^{m+d}}^{\mathcal{J}}[0, T]$.

Let $g_i \in C_b(\mathbb{R}^{2m+d})$. Then for $0 \leq t_1 < \cdots < t_k \leq t$, $\{t_i\}, t \in [0, T] \setminus D$,

$$
\begin{aligned}
0 &= E[(X_n(t) - U(t) - \int_t^T f(s, X_n, V) ds) \prod_{i=1}^k g_i(X_n(t_i), Y(t_i))] \\
&\to E[(X(t) - U(t) - \int_t^T f(s, X, V) ds) \prod_{i=1}^k g_i(X(t_i), Y(t_i))].
\end{aligned}
$$

Note that since

$$|X_n(t) - U(t)| \le E[\int_0^T g(s, V)ds|\mathcal{F}_t^Y],$$

$\{X_n(t) - U(t)\}$ is uniformly integrable justifying the convergence of the expectations. It follows that for each $t \in [0, T] \setminus D$,

$$X(t) = U(t) + E[\int_t^T f(s, X, V)ds|\mathcal{F}_t^X \vee \mathcal{F}_t^Y],$$

and the identity extends to all $t \in [0, T]$ by the right continuity of X and U.

If Y has no fixed points of discontinuity, then X has no fixed points of discontinuity and X is temporally compatible with Y. In any case, X is RC-compatible with Y.

Example 3.11. The multiple time-change equation

$$X(t) = X(0) + \sum_{k=1}^m W_k(\int_0^t \beta_k(X(s))ds)\zeta_k + \int_0^t F(X(s))ds, \qquad (3.5)$$

arises naturally in the derivation of diffusion approximations for continuous time Markov chains. (See, for example, Ethier and Kurtz (1986), Chapter 11.) Here the W_k are independent, scalar, standard Brownian motions, $X(0)$ is a $\mathbb{R}^d$-valued random variable independent of the W_k, $\zeta_k \in \mathbb{R}^d$, and the β_k and F are measurable functions (typically continuous) satisfying $\beta_k : \mathbb{R}^d \to [0, \infty)$ and $F : \mathbb{R}^d \to \mathbb{R}^d$. Setting $Y = (X(0), W_1, \ldots, W_m)$ and $\tau_k(t) = \int_0^t \beta_k(X(s))ds$, for $\alpha \in [0, \infty)^m$, define

$$\mathcal{F}_\alpha^Y = \sigma(W_k(s_k) : 0 \le s_k \le \alpha_k, k = 1, \ldots, m) \vee \sigma(X(0))$$

and

$$\mathcal{F}_\alpha^X = \sigma(\{\tau_1(t) \le s_1, \tau_2(t) \le s_2, \ldots\} : s_i \le \alpha_i, i = 1, \ldots, m, t \ge 0).$$

If the β_k are continuous, $\{\mathcal{F}_\alpha^Y\}$ and $\{\mathcal{F}_\alpha^X\}$ determine a compatibility condition satisfying the conditions of Lemma 3.5.

If X is a compatible solution, then $\tau(t) = (\tau_1(t), \ldots, \tau_m(t))$ is a stopping time with respect to $\{\mathcal{F}_\alpha^X \vee \mathcal{F}_\alpha^Y\}$ and $W_k(\int_0^t \beta_k(X(s))ds)$, $k = 1, \ldots, m$, are $\{\mathcal{F}_{\tau(t)}\}$-martingales. It follows that X is a solution of the martingale problem for

$$Af(x) = \frac{1}{2}\sum_{i,j} a_{ij}(x)\partial_i\partial_j f(x) + F(x) \cdot \nabla f(x),$$

$a(x) = \sum_{k=1}^m \beta_k(x)\zeta_k\zeta_k^T$. (Note that m may be infinity provided $\sum_{k=1}^\infty \beta_k(x)|\zeta_k|^2 < \infty$.)

Setting $\eta_n(t) = \frac{[nt]}{n}$,

$$X_n(t) = X(0) + \sum_{k=1}^m W_k(\int_0^{\eta_n(t)} \beta_k(X_n(s))ds)\zeta_k + \int_0^{\eta_n(t)} F(X_n(s))ds$$

has a unique piecewise constant solution that has the same distribution as the usual Euler approximation to the corresponding Itô equation. Under appropriate growth conditions on the β_k and F (for example, if the β_k and F are bounded), $\{X_n\}$ is relatively compact for convergence in distribution in $D_{\mathbb{R}^d}[0, \infty)$, and if the β_k and F are continuous, any limit point X of $\{X_n\}$ will satisfy (3.5). Lemma 3.5 gives that X is compatible with Y.

Uniqueness of the distribution of X would follow from uniqueness for the corresponding martingale problem; however, except for $m = 1$, no pathwise uniqueness

result of any generality is known. Let $\tau_k(t) = \int_0^t \beta_k(X(s))ds$ and $\gamma(t) = \int_0^t F(X(s))ds$. Then

$$\begin{aligned}
\dot{\tau}_l(t) &= \beta_l(X(0) + \sum_k W_k(\tau_k(t))\zeta_k + \gamma(t)) \\
\dot{\gamma}(t) &= F(X(0) + \sum_k W_k(\tau_k(t))\zeta_k + \gamma(t)),
\end{aligned}$$

which is a random ordinary differential equation. Except in the case β_k all constant, however, the right side is at best Hölder of order 1/2.

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. MR-718998

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. MR-2141331

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. MR-1978664

H. J. Engelbert. On the theorem of T. Yamada and S. Watanabe. *Stochastics Stochastic Rep.*, 36(3-4):205–216, 1991. ISSN 1045-1129. MR-1128494

Stewart N. Ethier and Thomas G. Kurtz. *Markov processes: Characterization and Convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. ISBN 0-471-08186-8. MR-0838085

Jean Jacod. Weak and strong solutions of stochastic differential equations. *Stochastics*, 3(3):171–191, 1980. ISSN 0090-9491. MR-0573202

Jean Jacod and Jean Mémin. Existence of weak solutions for stochastic differential equations with driving semimartingales. *Stochastics*, 4(4):317–337, 1980/81. ISSN 0090-9491. doi: 10.1080/17442508108833169. URL http://dx.doi.org/10.1080/17442508108833169. MR-609691

Adam Jakubowski. A non-Skorohod topology on the Skorohod space. *Electron. J. Probab.*, 2:no. 4, 21 pp. (electronic), 1997. ISSN 1083-6489. doi: 10.1214/EJP.v2-18. URL http://dx.doi.org/10.1214/EJP.v2-18. MR-1475862

Thomas G. Kurtz. Random time changes and convergence in distribution under the Meyer-Zheng conditions. *Ann. Probab.*, 19(3):1010–1034, 1991. ISSN 0091-1798. URL http://www.jstor.org/stable/2244471. MR-1112405

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. MR-2336594

Thomas G. Kurtz and Philip Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.*, 19(3):1035–1070, 1991. ISSN 0091-1798. MR-1112406

V. A. Lebedev. On the existence of weak solutions for stochastic differential equations with driving martingales and random measures. *Stochastics*, 9(1-2):37–76, 1983. ISSN 0090-9491. doi: 10.1080/17442508308833247. URL http://dx.doi.org/10.1080/17442508308833247. MR-703847

P.-A. Meyer and W. A. Zheng. Tightness criteria for laws of semimartingales. *Ann. Inst. H. Poincaré Probab. Statist.*, 20(4):353–372, 1984. ISSN 0246-0203. URL http://www.numdam.org/item?id=AIHPB_1984__20_4_353_0. MR-0771895

Toshio Yamada and Shinzo Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11:155–167, 1971. MR-0278420