The Annals of Probability
1995, Vol. 23, No. 3, 1234–1261

CONSTRUCTING GAMMA-MARTINGALES WITH PRESCRIBED LIMIT, USING BACKWARDS SDE

BY R. W. R. DARLING

University of South Florida

Let V_W and V_Y be Euclidean vector spaces and let $V_Z \equiv L(V_W \to V_Y)$. Given a Wiener process W on V_W, with natural filtration $\{\mathfrak{T}_t\}$, and a $\mathfrak{T}_T$-measurable random variable U in V_Y, we seek adapted processes (Y, Z) in $V_Y \times V_Z$ satisfying the SDE

$$U = Y(t) + \int_{(t,T]} Z\,dW - \int_{(t,T]} \Gamma(Y, ZZ^*)\,ds/2, \qquad 0 \le t \le T,$$

under local Lipschitz and convexity conditions on the map $(y, A) \to \Gamma(y, A)$. These conditions apply in particular in the case $\Gamma(y, A) = \sum \Gamma^i_{jk}(y) A^{jk}$, where Γ is a linear connection on V_Y whose Christoffel symbols Γ^i_{jk} are bounded and Lipschitz, and Γ has certain convexity properties. In that case the solution Y above is known as a Γ-martingale with terminal value U. The solution (Y, Z) is constructed explicitly using the Pardoux–Peng theory of backwards SDE's. Applications include the Dirichlet problem and the heat equation for harmonic mappings, and other PDE's.

0. Introduction. Pardoux and Peng (1990, 1992, 1994) have considered existence and uniqueness of adapted solutions (Y, Z) in $R^m \times R^{m \times l}$ to stochastic differential equations of the form

$$(1)\quad U = Y(t) + \int_{(t,T]} Z_s\,dW - \int_{(t,T]} F(x, Y_s, Z_s)\,ds/2, \qquad 0 \le t \le T,$$

where W is a multidimensional Wiener process, $F(t,\cdot,\cdot)$ is progressively measurable and U is a given $\mathfrak{T}_T$-measurable random variable known as the *terminal value*; see Section 1.4 below for more technical details.

An interesting example of an equation of this type is the case where $T = \infty$, $\{\Gamma^i_{jk}\}$ are the Christoffel symbols of a connection Γ on R^m and

$$F^i(s, y, z) \equiv \sum_{j,k,q} \Gamma^i_{jk}(y) z^j_q z^k_q,$$

in which case Y becomes a Γ-martingale with prescribed limit $Y_\infty = U$. As explained in Section 2.1 below, a Γ-martingale is a kind of stochastic process with values in a manifold M, with connection Γ, which generalizes the notion of continuous local martingale on Euclidean space. Solutions of (1) are closely related to the following problems.

Received February 1994.
AMS 1991 *subject classifications*. 58G32, 60G48, 60H20.
Key words and phrases. Stochastic differential equation, terminal value, Brownian motion, martingale, gamma-martingale, connection, harmonic map, quasilinear partial differential equation.

TABLE 1
Literature on existence of a gamma-martingale converging to a prescribed limit

	Levi–Civita Γ	**Linear Γ**	**Nonlinear Γ, or "connector"**
Wiener filtration $\{F_t\}$	Kendall (1990) Picard (1989, 1991)	—	This article
More general $\{F_t\}$	—	Arnaudon (1993) Darling (1993)	Picard (1994)

1. The Dirichlet problem for harmonic mappings with values in (M, Γ), as discussed in Kendall (1990, 1993).
2. The construction of a barycenter with associativity under conditioning; see Emery and Mokobodzki (1991), Picard (1994) and Peng (1993).
3. Duffie and Epstein's model (1992a, b) of stochastic differential utility.

The problem of uniqueness of such a Y was studied by Emery (1985). Table 1 gives a brief guide to the literature on the problem of existence. In every case except Picard (1989, 1991), some kind of "convex geometry" is assumed, meaning that there exists a continuous, convex function from $M \times M$ to $[0, \infty)$ which is zero precisely on the diagonal; this is a stronger condition than uniqueness of solutions to the geodesic equation [see Kendall (1992)]. In every case except Darling (1995), M is compact, with some additional convexity properties. In Picard (1989, 1991, 1994) there are some further restrictions on the terminal value.

In the case where Y takes values in a compact manifold with boundary M, Kendall (1990, 1991, 1992) has shown that uniqueness of the solution is equivalent to the statement that (M, Γ) has convex geometry.

All existence proofs so far have used specific discretization procedures, in which one constructs discrete time processes on M using conditional expectations in a sequence of nearby tangent spaces, and then uses tightness and weak convergence ideas to show that these processes converge to a limit which must be a Γ-martingale. This article presents a solution to this problem using the Pardoux–Peng construction without ever discretizing, taking conditional expectations or using any kind of weak convergence. The strategy of the proof can be simply described:

1. Solve a Pardoux–Peng backwards SDE (23) to construct a process which is close to being a Γ-martingale and which terminates at the desired random variable.
2. Use the solution to set up a forward SDE (66) for a Γ-martingale, whose terminal value is close to being the desired one.
3. Repeat steps 1 and 2 with a parameter tending to zero to create a sequence of Γ-martingales and show that their terminal values converge to the desired value (Lemma 6.4).
4. Show by an abstract method that this sequence has a limit, which must be a Γ-martingale with the desired terminal value (Theorem 7.1).

The assumptions used to obtain this solution are similar to the "convex geometry" used by Kendall (1990) and Picard (1994) (in fact, convexity with respect to both Γ and the Euclidean connection relative to the embedding is used in a number of places), but in certain respects the results presented here are more general. For example:

1. All proofs are carried out for nonlinear connections (which involves almost no extra difficulty) and extend immediately to the time-dependent case (Section 8.3).
2. No differentiability is assumed for the connection—only local Lipschitz properties (Condition 6.1).
3. It is not necessary to work within a compact manifold nor to assume that the terminal value is bounded (Section 8.2). For other results in this direction, see Darling (1994).
4. The method offers the possibility of extension to the case where W is replaced by a more general (possibly discontinuous) martingale, as in Antonelli (1993).

1. Basic concepts.

1.1. *Linear algebra.* Let V_W, V_X and V_Y be finite-dimensional Euclidean vector spaces, and let V_Z denote the vector space $L(V_W \to V_Y)$ (linear maps from V_W to V_Y) with the Hilbert–Schmidt norm

$$|z|^2 \equiv \mathrm{Tr}(z \cdot z), \qquad z \in V_Z,$$

where $z \cdot z$ is a square matrix whose (j, k) entry $(z \cdot z)^{jk}$ is the dot product $z^j \cdot z^k$ of the vectors in V_Y formed from the jth and kth columns of z. Alternatively, for z in V_Z we may identify $z \cdot z$ with the linear mapping $zz^* \in L(V_Y^* \to V_Y) \simeq V_Y \otimes V_Y$ and give it the norm

$$\|z \cdot z\| \equiv \left\{ \sum_{j,k} (z^j \cdot z^k)^2 \right\}^{1/2}. \tag{2}$$

Two applications of the Cauchy–Schwarz inequality suffice to prove that

$$\frac{1}{\sqrt{m}} |z|^2 \le \|z \cdot z\| \le |z|^2, \tag{3}$$

where $m \equiv \dim(V_Y)$.

1.2. *Connections.* For an elementary account of differential forms and linear connections, see Darling (1994). We shall deal here with nonlinear connections Γ on the tangent bundle of V_Y, and so the topology is Euclidean, and a single local trivialization of the tangent bundle will suffice. In this trivialization the connection is specified by a *local connector*, also denoted Γ, which is a map

$$\Gamma : V_Y \times (V_Y \otimes V_Y) \to V_Y. \tag{4}$$

For this paper it suffices that $\Gamma(y, \cdot)$ is defined on the set of elements of $V_Y \otimes V_Y$ of the form $\Sigma v_\alpha \otimes v_\alpha$ (finite sum). We do not deal with curvature in this paper, so Γ need not be differentiable. In the linear case where $\Gamma(y, \Sigma v_\alpha \otimes w_\alpha) = \Sigma \Gamma(y)(v_\alpha \otimes w_\alpha)$, the ith component of $\Gamma(y)(v \otimes w)$ is

$$\sum_{j,k} \Gamma^i_{jk}(y) v^j w^k,$$

using the Christoffel symbols Γ^i_{jk}. Whenever Γ is linear, it may be assumed to be torsion-free, that is, $\Gamma(y, v \otimes w) = \Gamma(y, w \otimes v)$, but need not be a metric connection. Indeed the use of the Euclidean metric on V_Y is just for analytic convenience. One should think of choosing a convenient coordinate system for studying Γ, and finally taking a metric whose metric tensor in this coordinate system is the identity.

For linear connections one uses the dual connection on the cotangent bundle to find the covariant derivative $\nabla d\phi$ of an exact 1-form $d\phi$ [i.e., the second covariant derivative of the function $\phi \in C^2(V_Y)$] by the formula

$$\nabla d\phi(D_j \otimes D_k) \equiv D_{jk}\phi - \sum_i (D_i \phi)\Gamma^i_{jk}, \tag{5}$$

where $D_j \equiv \partial/\partial y^j$ and $D_{jk} \equiv D_j D_k$. Since $\nabla d\phi$ is a $(0,2)$-tensor, we can think of $y \to \nabla d\phi(y)$ as a map from V_Y to $L(V_Y \otimes V_Y \to R)$. A coordinate-free form of (5) is

$$\nabla d\phi(y)(v \otimes w) = D^2\phi(y)(v \otimes w) - D\phi(y)\Gamma(y)(v \otimes w). \tag{6}$$

Imitating this expression in the case of a nonlinear connection gives the formula

$$\nabla d\phi(y)(z \cdot z) = D^2\phi(y)(z \cdot z) - D\phi(y)\Gamma(y, z \cdot z), \qquad z \in V_Z, \tag{7}$$

which is not necessarily linear in $z \cdot z$. Observe that it is only necessary to define the left side of (7) for elements of $V_Y \otimes V_Y$ of the form $\{z \cdot z\colon z \in V_Z\}$, not on the whole of $V_Y \otimes V_Y$.

1.3. *Convex functions.* We shall say that $\phi \in C^2(V_Y)$ is Γ-*convex* (resp. *strictly* Γ-*convex*) on $G \subseteq V_Y$ if for all $y \in G$ and $z \in V_Z$, $\nabla d\phi(y)(z \cdot z) \geq 0$ [resp. $\nabla d\phi(y)(z \cdot z) \geq \alpha |z|^2$ for some $\alpha > 0$]. Note that the second definition, but not the first, depends on the arbitrary choice of the metric, and that both depend on the dimension of V_W. A characterization of Γ-convexity of ϕ in the linear (but not the general) case is to say that $t \to \phi(\gamma(t))$ is convex for every geodesic γ, that is, for every solution of the geodesic equation $\ddot{\gamma} + \Gamma(\gamma)(\dot{\gamma} \otimes \dot{\gamma}) = 0$. Because of the nonintrinsic nature of certain constructions, we also need to introduce the expression

$$\mathrm{Hess}_+ \, \phi(y)(z \cdot z) \equiv \min\{\nabla d\phi(y)(z \cdot z), D^2\phi(y)(z \cdot z)\} \tag{8}$$

and we shall say that ϕ is *doubly convex* (resp. *strictly doubly convex*) on $G \subseteq V_Y$ if for all $y \in G$ and $z \in V_z$, $\mathrm{Hess}_+ \, \phi(y)(z \cdot z) \geq 0$ [resp. $\mathrm{Hess}_+ \, \phi(y)(z \cdot z) \geq \alpha |z|$ for some $\alpha > 0$]. Even for linear connections, these properties are nonintrinsic, that is, they depend on the choice of coordinate system.

1.4. *Stochastic differential equations with prescribed terminal value.* Let $\{W(t),\ 0 \le t < \infty\}$ be a V_W-valued Wiener process on a probability space $(\Omega, \mathfrak{T}, P)$, with the natural filtration $\mathfrak{T}_t = \sigma\{W(s),\ 0 \le s \le t\}$, augmented by the P-null sets and their complements; right-continuity is not assumed. Suppose that $0 < T < \infty$ and that we are given some random function

$$F\colon \Omega \times [0,T] \times V_Y \times V_Z \to V_Y,$$

where $F(t,\cdot,\cdot)$ is progressively measurable, and a $\mathfrak{T}_T$-measurable random variable U on V_Y, known as the *terminal value*. Following Pardoux and Peng (1990), the problem is to construct a pair (Y,Z) of adapted, progressively measurable processes on $V_Y \times Y_Z$ which satisfy the stochastic differential equation

$$dY(t) = Z(t)\,dW(t) - (1/2)F(t,Y(t)),Z(t)\,dt$$

such that $Y(T) = U$. Let us emphasize that $Y \equiv \{Y(t),\ 0 \le t \le T\}$ is a process with unknown initial value, but known terminal value. There will be no ambiguity if henceforth we omit the time variable t from the integrands in most stochastic integrals.

Subtracting the equation

$$Y(t) - Y(0) = \int_{(0,t]} Z\,dW - (1/2)\int_{(0,t]} F(s,Y,Z)\,ds$$

from its counterpart when $t = T$ gives the "Pardoux–Peng equation"

$$U = Y(t) + \int_{(t,T]} Z\,dW - (1/2)\int_{(t,T]} F(s,Y,Z)\,ds, \qquad 0 \le t \le T. \tag{9}$$

Conditions on U and F for existence and uniqueness of solutions to (9) are given in Pardoux and Peng (1990). We only quote a special case of Pardoux and Peng's (1990) result [for the non-Lipschitz case, see also Pardoux and Peng (1994)]. First here is some notation: For any Euclidean vector space V, $H^p(0,T;V)$ denotes the set of progressively measurable processes Z (with respect to $\{\mathfrak{T}_t\}$) such that

$$E\left[\left\{\int_{(0,T]} |Z|^2\,ds\right\}^{p/2}\right] < \infty. \tag{10}$$

THEOREM 1.5 (Pardoux–Peng theorem for the Lipschitz case). *Suppose $T < \infty$, $U \in L^2(\Omega, \mathfrak{T}_T, V_Y)$ and F: $\Omega \times [0,T] \times V_Y \times V_Z \to V_Y$ has the properties*:

(i) $F(\cdot,0,0) \in H^2(0,T;V_Y)$;

(ii) *There exists $c > 0$ such that $|F(\omega,t,y,z) - F(\omega,t,\tilde{y},\tilde{z})| \le c(|y - \tilde{y}| + |z - \tilde{z}|)$ for almost all t and ω.*

Then there exists a unique pair (Y,Z) of processes in $H^2(0,T;V_Y) \times H^2(0,T;V_Z)$ which satisfy (9).

2. The Pardoux–Peng system associated with a connection. Given a nonlinear connection Γ and a terminal value U, define the *Pardoux–Peng* *system* of stochastic differential equations associated with the triple (W, Γ, U) to be the Pardoux–Peng equation (9) with

$$F(t, y, z) \equiv \Gamma(y, z \cdot z).$$

Note that, even when Γ is linear, this F is usually not Lipschitz in y or z because the process Z is not known to be bounded. In the linear case, (9) can be written as

$$U^i = Y^i(t) + \int_{(t,T)} Z^i \, dW - (1/2) \sum_{j,k=1}^{m} \int_{(t,T]} \Gamma^i_{jk}(Y)(Z^j \cdot Z^k) \, ds, \qquad i = 1, 2, \ldots, m,.$$

In general, it becomes

$$U = Y(t) + \int_{(t,T]} Z \, dW - (1/2) \int_{(t,T]} \Gamma(Y, Z \cdot Z) \, ds. \tag{11}$$

Using (7) and Itô's formula, another statement of (11) is that, for all $f \in C^2(V_Y)$,

$$\begin{aligned} f(U) &= f(Y(t)) + \int_{(t,T]} df(Y)(Z \, dW) \\ &\quad + (1/2) \int_{(t,T]} \nabla df(Y, Z \cdot Z) \, ds. \end{aligned} \tag{12}$$

2.1. *Gamma-martingales.* A clear introduction to Γ-martingales in the case of a linear connection Γ is given in Emery and Meyer (1989). The case of nonlinear Γ is less well known, but is presented in Meyer (1981). We shall say that a continuous semimartingale X with values in V_Y is a Γ-*martingale*, with respect to the filtration $\{\mathfrak{T}_t, t \geq 0\}$ given above, if each component has a semimartingale decomposition $X^i = M^i + A^i$, where $d[M^i, M^j] = H^{ij} \, dt$ (absolute continuity of the joint quadratic variations with respect to Lebesgue measure holds for all continuous local martingales in this filtration) and

$$A(t) = -(1/2) \int_{(0,t]} \Gamma(X, H) \, ds. \tag{13}$$

Thus for the Euclidean connection, Γ-martingales are simply m-dimensional continuous local martingales. A more abstract way to give the definition is to say that X is a continuous semimartingale in V_Y such that

$$X + (1/2) \int \Gamma(X, (dX \otimes dX)/ds) \, ds \tag{14}$$

is a continuous local martingale in V_Y, where $(dX \otimes dX)^{ij}$ means $d[X^i, X^j]$. Its Itô representation [see Revuz and Yor (1991)] yields a progressively measurable process Z such that

$$\int_{(0,t]} Z \, dW \equiv X(t) - X(0) + (1/2) \int_{(0,t]} \Gamma(X, (dX \otimes dX)/ds) \, ds. \tag{15}$$

It is clear from (15) that $(dX \otimes dX)/ds = Z \cdot Z$, and so the solutions to (11) are precisely the Γ-martingales on $[0,T]$ with terminal value U. Notice also that, from Itô's formula, the real semimartingale

$$(16) \qquad f(X_t) - f(X_0) - (1/2)\int_{(0,t]} \nabla df(x,(dX \otimes dX)/ds)\, ds$$

belongs to M_{loc}^c, the space of continuous local martingales, for all $f \in C^2(M)$.

Given a Riemannian metric g on M (not necessarily related to Γ), we may associate with any M-valued continuous semimartingale Y a *Riemannian quadratic variation process* $\int\langle dY|dY\rangle$, given in local coordinates by

$$(17) \qquad \int_{(0,t]} \langle dY|dY\rangle = \int_{(0,t]} \sum g_{ij}(Y)\, d[Y^i, Y^j]$$

[see Emery and Meyer (1989)]. Suppose $0 < p < \infty$. A Γ-martingale X is called an H^p Γ-martingale on (M,g) or is said to belong to the *Hardy space* H^p [see Darling (1993)] if

$$(18) \qquad \int_{(0,\infty)} \langle dX|dX\rangle \in L^{p/2}.$$

In the case where g is the Euclidean metric, (17) becomes the Euclidean quadratic variation and (15) shows that $\int\langle dX|dX\rangle = \int|Z|^2\, ds$. We have now proved the following lemma.

LEMMA 2.2 (Gamma-martingales and backwards SDE's). *The solutions to* (11) [*with* $Z \in H^p(0,T;V_Z)$—*see* (10)] *are precisely the* $\{\mathfrak{F}_t\}$-*adapted* Γ-*martingales* (*in the class* H^p) *on* $[0,T]$ *with terminal value* U.

3. An approximation scheme with uniformly bounded variation. Equation (11), which we wish to solve, usually lies outside the scope of Pardoux–Peng Theorem 1.5. However, suppose $h\colon (0,1) \times V_Z \to [0,\infty)$ is a function with the following two properties: for each ε,

$$(19) \qquad |z|^2 \le 1/\varepsilon \quad \Rightarrow \quad h(\varepsilon, z) = 0,$$

$$(20) \qquad z \to z \cdot z/(1 + h(\varepsilon, z)) \quad \text{is bounded and Lipschitz.}$$

Note that the norm (2) is used for $z \cdot z$. An example of such a function h is

$$(21) \qquad h(\varepsilon, z) = (\varepsilon\|z \cdot z\| - 1)1_{\{\|z\cdot z\| < 1/\varepsilon\}}.$$

3.1. *Local Lipschitz property of the connection.* Assume that for each $r > 0$ there exists $c(r) > 0$ such that

$$(22) \qquad |\Gamma(y, z\cdot z) - \Gamma(\bar{y}, \bar{z}\cdot\bar{z})| \le c(r)\big(|y - \bar{y}| + \|z\cdot z - \bar{z}\cdot\bar{z}\|\big)$$

for all $y, \bar{y} \in V_Y$ and all $z, \bar{z} \in V_Z$ with $\|z \cdot z\| \le r$ and $\|\bar{z}\cdot\bar{z}\| \le r$.

REMARK 3.1.1. For example, in the linear case $\Gamma(y, z\cdot z) \equiv \Gamma(y)(z \cdot z)$, 3.1 holds when the Christoffel symbols are bounded and Lipschitz.

When Condition 3.1 holds, it follows that the map

$$(y, z) \to \Gamma\left(y, \frac{z \cdot z}{1 + h(\varepsilon, z)}\right)$$

is bounded and Lipschitz, for each ε, since the composition of two Lipschitz maps is Lipschitz, and $z \cdot z/(1 + h(\varepsilon, z))$ is bounded. By Theorem 1.5, for any $U \in L^2$ there is a unique progressively measurable solution $(Y_\varepsilon, Z_\varepsilon) \in H^2(0, T; V_Y \times V_Z)$ to the following approximation to (11):

$$U = Y_\varepsilon(t) + \int_{(t,T]} Z_\varepsilon \, dW - (1/2) \int_{(t,T]} \Gamma\left(Y_\varepsilon, \frac{Z_\varepsilon \cdot Z_\varepsilon}{1 + h(\varepsilon, Z_\varepsilon)}\right) ds. \tag{23}$$

Consider the following condition on Γ and on the coordinate system:

3.2. *A convexity condition.* (i) Assume there exists $\Phi \in C^2(V_Y)$ with bounded first and second derivatives and with $\nabla d\Phi(y, z \cdot z) \le c\|z\|$ for some $c > 0$, such that the set

$$G \equiv \{y \in V_Y : \Phi(y) \le 0\} \tag{24}$$

is compact and Φ is doubly convex on G^c [see (8)].

(ii) Moreover there exists $f \in C^2(V_Y)$ which is strictly doubly convex on G [see (8)].

REMARK 3.2.1. (i) The set G is Γ-convex in the sense that every geodesic segment γ with endpoints in G lies in G. Suppose $a < b$, $\gamma(a)$ and $\gamma(b)$ lie on the boundary of G and $\gamma(t) \in G^c$ for $a < t < b$. The Γ-convexity of Φ on G^c implies $\Phi \circ \gamma$ is convex, but

$$\Phi(\gamma(a)) = 0 = \Phi(\gamma(b)) \Rightarrow \Phi(\gamma(t)) \le 0 \qquad \forall\, t \in (a, b),$$

which is a contradiction to (24). The set G is convex in the Euclidean sense by an analogous argument. This implies in particular that G is connected.

(ii) Consider the case where $\Gamma(y, z \cdot z)$ is zero for y outside some ball of radius a. Condition 3.2(i) holds for any Φ such that $\Phi(y) \equiv |y| - a$ on $\{y : |y| \ge a\}$ and which is nonpositive on $G \equiv \{y : |y| \le a\}$. Provided the values of $\Gamma(y, z \cdot z)/\|z \cdot z\|$ are not too large, the function $f(y) \equiv |y|^2$ is strictly doubly convex.

(iii) The purpose of (i) will be to ensure that the processes $\{Y_\varepsilon\}$ are uniformly bounded (Lemma 3.3); (ii) will ensure that the $\{Z_\varepsilon\}$ are uniformly bounded in $H^2(0, T; V_Z)$ (Proposition 3.4).

(iv) These conditions refer to the choice of coordinate system, inasmuch as they refer to the Euclidean convexity of Φ and the strict Euclidean convexity of f.

Here is a weaker version of the condition, applicable to Γ-martingales.

3.2.2. *Convexity condition—weaker version.* This is the same as 3.2 except that Φ (resp. f) need only be Γ-convex (resp. strictly Γ-convex).

The following lemma is an easy extension of a result in Darling (1993).

LEMMA 3.3 (Boundedness of Pardoux–Peng solutions). *Assume that* Γ *satisfies Conditions* 3.1 *and* 3.2(*i*). *If the terminal value* U *of* (23) *lies in* G, *then* $Y_\varepsilon(t) \in G$ *for all* t, *with probability* 1.

PROOF. *Step* 1. Let us use the abbreviation

$$\bar{Z}_\varepsilon \equiv Z_\varepsilon/(1 + h(\varepsilon, Z_\varepsilon))^{1/2} \tag{25}$$

and so $dY_\varepsilon(t) = Z_\varepsilon\, dW - (1/2)\Gamma(Y_\varepsilon \bar{Z}_\varepsilon \cdot \bar{Z}_\varepsilon)\, dt$. Thus, for any C^2 f and any stopping times σ, τ with $0 \le \sigma \le \tau \le T$, we have

$$\begin{aligned} f(Y_\varepsilon(\tau)) - f(Y_\varepsilon(\sigma)) &= \int_{(\sigma,\tau]} df(Y_\varepsilon)\big(Z_\varepsilon\, dW - (1/2)\Gamma\big(Y_\varepsilon, \bar{Z}_\varepsilon \cdot \bar{Z}_\varepsilon\big)\, dt\big) \\ &\quad + (1/2)\int_{(\sigma,\tau)} D^2 f(Y_\varepsilon)(Z_\varepsilon \cdot Z_\varepsilon)\, dt, \\ &= \int_{(\sigma,\tau]} df(Y_\varepsilon)(Z_\varepsilon\, dW) \\ &\quad + \int_{(\sigma,\tau]} \big\{\nabla df(Y_\varepsilon)\big(\bar{Z}_\varepsilon \cdot \bar{Z}_\varepsilon\big) + hD^2 f(Y_\varepsilon)\big(\bar{Z}_\varepsilon \cdot \bar{Z}_\varepsilon\big)\big\}\frac{dt}{2}. \end{aligned} \tag{26}$$

Here we have used the fact that $Z_\varepsilon \cdot Z_\varepsilon = (1 + h)\bar{Z}_\varepsilon \bar{Z}_\varepsilon$, and so

$$D^2 f(Y_\varepsilon)(Z_\varepsilon \cdot Z_\varepsilon) = (1 + h) D^2 f(Y_\varepsilon)\big(\bar{Z}_\varepsilon \cdot \bar{Z}_\varepsilon\big).$$

Consider the case where f is chosen to be the function Φ appearing in Condition 3.2. The first integral in (26) has mean zero because Z_ε is in $H^2(0, T; V_Z)$ and $d\Phi(y)$ is bounded. By the double convexity of Φ, the second integral is nonnegative, and its expected value is finite because Z_ε is in $H^2(0,T; V_Z)$, and $D^2\Phi(y)$ and $\nabla d\Phi(y, z \cdot z)/\|z \cdot z\|$ are bounded. In other words,

$$E\big[\Phi(Y_\varepsilon(\tau)) | \mathfrak{T}_\sigma\big] \ge \Phi(Y_\varepsilon(\sigma))$$

and $\Phi \circ Y_\varepsilon$ is a submartingale.

Step 2. Fix $n \ge 1$ and a time $r \ge 0$, and define a stopping time σ by

$$\sigma = \inf\{t \ge r : \Phi(Y_\varepsilon(t)) \le 1/n\}.$$

Observe that $P(\sigma < \infty) = 1$ and $\Phi(Y_\varepsilon(\sigma)) \le 1/n$ because $\Phi(U) \le 0$ a.s. It follows from Step 1 that

$$\Phi(Y_\varepsilon(r)) \le E\big[\Phi(Y_\varepsilon(\sigma)) | \mathfrak{T}_r\big] \le 1/n. \quad \text{a.s.} \tag{27}$$

Since (27) holds for all n, we see that $P(Y_r \in G) = 1$, but r was arbitrary, so with probability 1, $Y_t \in G$ for every rational t, and hence for all t, by path continuity. $\square$

The next result, closely related to previous one, puts a uniform bound on the total quadratic variations of the martingales $\int Z_\varepsilon\, dW$.

PROPOSITION 3.4 (Uniformly bounded quadratic variation). *Suppose Conditions* 3.1 *and* 3.2 *hold for some* Φ *and* f, *with* $\mathrm{Hess}_+ f(y)(z \cdot z) \geq \alpha |z|^2$ *on* G. *Then*

$$\sup\left\{E\left[\int_{(0,T]} |Z_\varepsilon|^2\, dt\right] : 0 < \varepsilon < 1\right\} \tag{28}$$
$$\leq (2/\alpha)\sup\{f(x) - f(y) : (x,y) \in G \times G\}.$$

PROOF. For an f as in Condition 3.2, we see that, on the entire interval $[0,T]$,

$$\{\nabla df(Y_\varepsilon, \overline{Z}_\varepsilon \cdot \overline{Z}_\varepsilon) + hD^2 f(Y_\varepsilon)(\overline{Z}_\varepsilon \cdot \overline{Z}_\varepsilon)\} \geq \alpha(|\overline{Z}_\varepsilon|^2 + h|\overline{Z}_\varepsilon|^2) = \alpha|Z_\varepsilon|^2, \tag{29}$$

where (25) gives the last identity. Since the process Y_ε is bounded, by Lemma 3.2, it follows that $df(Y_\varepsilon)$ is bounded and so the first stochastic integral in (26) is a martingale. Now (26) and (29) combine to give

$$E\left[\int_{(0,T]} |Z_\varepsilon|^2\, dt\right] \leq (2/\alpha) E[f(Y_\varepsilon(T)) - f(Y_\varepsilon(0))], \tag{30}$$

and the result follows because, by Lemma 3.3, $Y_\varepsilon(0)$ as well as $U \equiv Y_\varepsilon(T)$ are in G. □

COROLLARY 3.5 (Similar result for gamma-martingales). *Assume Condition* 3.2.2. *Suppose* Y *is a bounded* $\{\mathfrak{F}_t\}$-*adapted* Γ-*martingale on* $[0,\infty)$ *with terminal value* U *and that* U *lies in* G. *Then* $Y_t \in G$ *for all* t *a.s. and* Y *is an* H^2 Γ-*martingale with Euclidean quadratic variation bounded by the right side of* (28).

PROOF. In view of Lemma 2.2 we consider a solution (Y, Z) to (11) instead of the solution to (23). Nowhere in the proof of Lemma 3.3 or Proposition 3.4 do we need the fact that T is finite or that ε is greater than zero. The only difficulty is that we do not know a priori that $Z \in H^2(0,\infty; V_Z)$, and therefore we cannot immediately take expectations in (26). To deal with this problem, define stopping times $\{\tau(\varepsilon),\ 0 < \varepsilon < 1\}$ as follows:

$$\tau(\varepsilon) \equiv \inf\left\{t \geq 0 : \int_{(0,t]} |Z|^2\, ds \geq 1/\varepsilon\right\}$$

or ∞ if there are no such t. Let $Y_\varepsilon(t) \equiv Y(t \wedge \tau(\varepsilon))$ and $Z_\varepsilon(t) \equiv Z(t)1_{\{t \leq \tau(\varepsilon)\}}$. Then Y_ε is an H^2 Γ-martingale, in which Z_ε plays the role of Z in (15). All the calculations of Proposition 3.4 go through for the pair $(Y_\varepsilon, Z_\varepsilon)$, taking $h = 0$, yielding the estimate (28), except that G is replaced by the bounded set in which Y spends its life. By Lebesgue's monotone convergence theorem,

$$E\left[\int_{[0,\infty)} |Z|^2\, ds\right] = \lim_{\varepsilon \to 0} E\left[\int_{[0,\infty)} |Z_\varepsilon|^2\, ds\right] < \infty.$$

By Lemma 2.2, Y is an H^2 Γ-martingale, and we may now use the submartingale argument of Lemma 3.3 to show that $Y_t \in G$ for all t a.s., and so (28) also follows. □

4. Coupling technique. In order to prove the main estimates of this paper, we need to study the stability of the $\mathfrak{T}_T$-measurable terminal value U with respect to variations in the path of the Wiener process W. In this construction [inspired by Picard (1994)] we shall assume temporarily that Ω is the canonical space $C([0,T]; V_W)$ of continuous paths starting at zero.

Let $\{W'(t), 0 \le t \le T\}$ be another V_W-valued Wiener process on a probability space $(\Omega', \mathfrak{T}', P')$, with Brownian filtration $\{\mathfrak{T}'_t\}$, all of which are copies of the original W, $(\Omega, \mathfrak{T}, P)$ and $\{\mathfrak{T}_t\}$, respectively. Construct a product probability space $(\overline{\Omega}, \overline{\mathfrak{T}}, \overline{P}) \equiv (\Omega \times \Omega', \mathfrak{T} \times \mathfrak{T}', P \otimes P')$, with filtration $\overline{\mathfrak{T}}_t \equiv \mathfrak{T}_t \times \mathfrak{T}'_t$, in which we identify $\mathfrak{T}_t$ and $\mathfrak{T}'_t$ as sub-sigma fields of $\overline{\mathfrak{T}}_t$.

Any $\{\mathfrak{T}_t\}$-progressively measurable subset H of $[0,T] \times \Omega$ may be identified with the $\{\overline{\mathfrak{T}}_t\}$-progressively measurable subset $H \times \Omega'$ of $[0,T] \times \overline{\Omega}$, and similarly for $\{\mathfrak{T}_t\}$-progressively measurable processes. For such an H, we may define an $\{\overline{\mathfrak{T}}_t\}$-adapted process W^H on $(\overline{\Omega}, \overline{\mathfrak{T}}, \overline{P})$ by

$$W_t^H \equiv \int_{(0,t]} 1_H \, dW' + \int_{(0,t]} 1_{H^c} \, dW. \tag{31}$$

For example, if $H =]\!]\sigma, \tau]\!]$ for some pair of $\{\mathfrak{T}_t\}$-stopping times σ, τ with $0 \le \sigma \le \tau \le T$, then:

1. On the random intervals $(0, \sigma]$ and $(\tau, T]$, the increments of W and W^H are identical.
2. On the random interval $(\sigma, \tau]$, the increments of W and W^H are independent.

The mapping $(W, W') \to (W, W^H)$ induces a new probability measure, denoted P_H (with corresponding expectation operator E_H), on $(\overline{\Omega}, \overline{\mathfrak{T}})$. Observe that the marginals of P_H are equal to P and P', respectively.

Given any $\mathfrak{T}_T$-measurable random variable $U \equiv U(W)$ on $(\Omega, \mathfrak{T}, P)$ with values in V_Y, and an identical copy $U' \equiv U'(W')$ on $(\Omega', \mathfrak{T}', P')$, the pair (U, U') has a joint law as a random variable on $(\overline{\Omega}, \overline{\mathfrak{T}}, \overline{P})$ with values in $V_Y \times V_Y$, induced by the composition map

$$(W, W') \to (W, W^H) \to (U(W), U'(W^H)).$$

For example, U and U' are independent if $H = [0,T] \times \Omega$ and identically equal if $H = \varnothing$.

In this and subsequent propositions, $\{\delta_1, \delta_2, \dots\}$ will refer to positive real numbers which we are free to choose as small as we like.

PROPOSITION 4.1 (Stability with respect to changes in the Wiener path). *Every random variable U in $L^2(\Omega, \mathfrak{T}_T, P; V)$, for any Euclidean vector space V, has the following property: for every $\delta_1 > 0$ there exists $\mu \equiv \mu(\delta_1) > 0$*

such that, for every $\{\mathfrak{T}_t\}$*-progressively measurable subset* H *of* $[0,T] \times \Omega$, *the joint law of* (U, U') *under* P_H *satisfies*

$$E_H\left[|U - U'|^2\right] \le \mu(\delta_1) E\left[\int_{(0,T]} 1_H \, dt\right] + \delta_1. \tag{32}$$

REMARKS. The idea is that U does not behave in an "unstable" fashion in response to a small change in the Wiener path. It is important to notice that the expectations on the right and left sides of the previous inequality are taken over different probability spaces. Since (32) refers only to the laws of the random variables involved, it remains true without the restriction that $\Omega \equiv C([0,T]; V_W)$.

PROOF. Let us give $\Omega \equiv C([0,T]; V_W)$ the topology of uniform convergence on compacts, under which it is a normal topological space. The Borel sets for this topology are identical with $\mathfrak{T}$, and Wiener measure P is a closed, regular measure [see Dudley (1989), Chapter 7, for the definitions]. Given $\delta > 0$, we first choose a compact set K in V such that

$$E_H\left[|U - U'|^2; A^c\right] \le \delta/3, \tag{33}$$

where $A \equiv \{(\omega, \omega'): (U(\omega), U'(\omega')) \in K \times K\}$. Let ρ denote

$$\rho \equiv \sup\left\{|x - y|^2: (x, y) \in K \times K\right\}.$$

By Lusin's theorem [Dudley (1989), page 190], there exists a compact set $F \subseteq \Omega$ with $P(F^c) < \delta/(6\rho)$ such that $U \equiv U(\omega)$ restricted to F is continuous. Let $F' \subseteq \Omega'$ be a copy of F. Now

$$\begin{aligned} P_H\left((F \times F')^c\right) &\le P_H(F^c \times \Omega') + P_H\left(\Omega \times (F')^c\right) \\ &\le 2P(F^c) \le \delta/(3\rho). \end{aligned} \tag{34}$$

Since $U - U'$ is continuous on the compact $F \times F'$, it is uniformly continuous, and so there exists an $\eta \equiv \eta(\delta) > 0$ such that for $(\omega, \omega') \in F \times F'$,

$$\sup\{|\omega(t) - \omega'(t)|: 0 \le t \le T\} \le \eta \quad \Rightarrow \quad |U(\omega) - U'(\omega')|^2 \le \delta/3. \tag{35}$$

Let us decompose the set A as follows:

$$A_0 \equiv A \cap (F \times F')^c,$$

$$A_1 \equiv A \cap \left\{(\omega, \omega') \in F \times F': \sup\{|\omega(t) - \omega'(t)|: 0 \le t \le T\} \le \eta\right\},$$

$$A_2 \equiv A \cap \left\{(\omega, \omega') \in F \times F': \sup\{|\omega(t) - \omega'(t)|: 0 \le t \le T\} > \eta\right\}.$$

Observe that by Doob's inequality and (31),

$$\begin{aligned} P_H(A_2) &\le P_H\left(\sup\left\{|W_t - W_t^H|^2: 0 \le t \le t\right\} > \eta^2\right) \\ &\le \eta^{-2} E\left[\sup\left\{|W_t - W_t^H|^2: 0 \le t \le T\right\}\right] \\ &\le 2\eta^{-2} E\left[|W_T - W_T^H|^2\right] = 2\eta^{-2} E\left[\int_{(0,T]} 1_H \, dt\right]. \end{aligned} \tag{36}$$

Finally we obtain from (33),

$$E_H\left[|U - U'|^2\right] \le \delta/3 + \sum_{0\le i\le 2} E_H\left[|U - U'|^2; A_i\right]$$
$$\le \delta + 2\rho\eta^{-2}E\left[\int_{(0,T]} 1_H\,dt\right],$$

using (34), (35) and (36). □

5. Stability of solutions with respect to the terminal value. Given a nonlinear connection Γ, the product connection $\Gamma^{(2)}$ on $V_Y \times V_Y$ is defined for our purposes by

$$\Gamma^{(2)}\left(\begin{bmatrix} y \\ y' \end{bmatrix}, \begin{bmatrix} z\cdot z & z\cdot z' \\ z'\cdot z & z'\cdot z' \end{bmatrix}\right) \equiv \Gamma^{(2)}\left(\begin{bmatrix} y \\ y' \end{bmatrix}, \begin{bmatrix} z\cdot z & 0 \\ 0 & z'\cdot z' \end{bmatrix}\right) \equiv \begin{bmatrix} \Gamma(y, z\cdot z) \\ \Gamma(y', z'\cdot z') \end{bmatrix}. \tag{37}$$

Note that we do not need to define it except when the second argument is of one of the two special forms above. As in (8), we shall say that a C^2 function Ψ: $V_Y \times V_Y \to R$ is doubly convex on a subset B of $V_Y \times V_Y$ if it is convex with respect to both the product connection $\Gamma^{(2)}$ and the Euclidean product connection; in other words,

$$\mathrm{Hess}_+ \Psi(y, y')((z \oplus z')\cdot(z \oplus z')) \ge 0 \tag{38}$$

for all $(y, y') \in B$ and all $z, z' \in V_Z$. $\nabla^{(2)} d\Psi$ will denote the second covariant derivative with respect to $\Gamma^{(2)}$, in the sense of (7). The following condition is based on Kendall (1990).

5.1. *Another convexity condition.* Let G be the compact set appearing in Condition 3.2. We shall say that G has *doubly convex geometry* if there exists a C^2 function Ψ: $V_Y \times V_Y \to [0, \infty)$ which is doubly convex on $B \times B$ for some open set $B \supset G$ and which vanishes precisely on the diagonal of $G \times G$, or to be specific,

$$\{(x, y) \in G \times G: \Psi(x, y) = 0\} = \{(x, x): x \in G\}. \tag{39}$$

REMARK 5.1.1. (i) For linear connections, such a Ψ always exists in a sufficiently small neighborhood of any point. The proof is a modification of (4.59) in Emery and Meyer (1989).

(ii) The condition is not intrinsic, since it refers to the Euclidean convexity of Ψ.

(iii) The purpose of the condition is to bound the distance between two approximate solutions in terms of the distance between their terminal values.

(iv) We know from Kendall (1990, 1991, 1992) that, in the case of linear connections, the weaker Condition 5.1.2 (with Ψ merely continuous, instead of C^2) is equivalent to uniqueness of a Γ-martingale Y with terminal value U in G, in the case where G is a compact manifold with boundary.

5.1.2. *Convexity condition—weaker version.* This is the same as Condition 5.1 except that Ψ need only be $\Gamma^{(2)}$-convex.

5.2. *Constructing a coupled pair of approximate solutions.* Let us change notation slightly from the previous section, so that (W, W') is the canonical random variable on $(\overline{\Omega}, \overline{\mathfrak{T}}, \overline{P}) \equiv (\Omega \times \Omega', \mathfrak{T} \times \mathfrak{T}', P_H)$, Ω being path space, H some $\{\mathfrak{T}_t\}$-progressively measurable set and P_H the probability measure described above such that the marginal distribution of both W and W' is that of a V_W-valued Wiener process. Given any $\mathfrak{T}_T$-measurable random variable $U_1 \equiv U_1(W)$ on $(\Omega, \mathfrak{T}, P)$ with values in V_Y, and a (possibly unrelated) $\mathfrak{T}'_T$-measurable random variable $U_2 \equiv U_2(W')$ on $(\Omega', \mathfrak{T}', P')$, we may solve both (23) for U_1 and the corresponding equation for U_2, namely,

$$U_2 = Y_\varepsilon'(t) + \int_{(t,T]} Z_\varepsilon' \, dW' - (1/2)\int_{(t,T]} \Gamma\left(Y_\varepsilon', \overline{Z}_\varepsilon' \cdot \overline{Z}_\varepsilon'\right) ds,$$

where h is as in (19) and (20) and

$$\overline{Z}_\varepsilon' \equiv Z_\varepsilon'/(1 + h(\varepsilon, Z_\varepsilon'))^{1/2}.$$

We now have a solution scheme analogous to (23), replacing Γ by the product connection $\Gamma^{(2)}$ as in (37), namely, a solution

$$((Y_\varepsilon, Z_\varepsilon), (Y_\varepsilon', Z_\varepsilon')) \in H^2(0, T; (V_Y \times V_Z) \times (V_Y \times V_Z))$$

to the backwards SDE:

$$\begin{aligned}\begin{bmatrix} U_1 \\ U_2 \end{bmatrix} &= \begin{bmatrix} Y_\varepsilon(t) \\ Y_\varepsilon'(t) \end{bmatrix} + \int_{(t,t]} \begin{bmatrix} Z_\varepsilon \, dW \\ Z_\varepsilon' \, dW' \end{bmatrix} \\ &\quad - \int_{(t,T]} \Gamma^{(2)}\left(\begin{bmatrix} Y_\varepsilon \\ Y_\varepsilon' \end{bmatrix}, \begin{bmatrix} \overline{Z}_\varepsilon \cdot \overline{Z}_\varepsilon & \overline{Z}_\varepsilon \cdot \overline{Z}_\varepsilon' \\ \overline{Z}_\varepsilon' \cdot \overline{Z}_\varepsilon & \overline{Z}_\varepsilon' \cdot \overline{Z}_\varepsilon' \end{bmatrix} \right) \frac{ds}{2}.\end{aligned} \tag{40}$$

Here $\overline{Z}_\varepsilon \equiv Z_\varepsilon/(1 + h(\varepsilon, Z_\varepsilon))^{1/2}$, and the top and bottom rows have a joint law induced by P_H. The meaning is the same if the off-diagonal entries of the last two-by-two matrix are set equal to zero, by (37). Note the technical point that the top row contains $\{\mathfrak{T}_t\}$-adapted processes and the bottom row, $\{\mathfrak{T}_t'\}$-adapted processes; thus everything is $\{\overline{\mathfrak{T}}_t\}$-adapted. Observe also that in expressions involving terms from both top and bottom, expectations must be denoted E_H, whereas for expressions involving the top only, E and E_H coincide.

PROPOSITION 5.3 (Bound on the distance between coupled solutions). *Suppose Γ satisfies Conditions* 3.1, 3.2 *and* 5.1. *Also assume that the range of the terminal values U_1 and U_2 is contained in the compact set G specified in Condition* 3.2 *a.s. Given $\delta_2 > 0$, there exists a constant $\lambda \equiv \lambda(\delta_2, G, \Gamma)$ such that, for every $\{\mathfrak{T}_t\}$-progressively measurable set H, every $\varepsilon > 0$, every $\{\overline{\mathfrak{T}}_t\}$-stopping time $v \leq T$ and every $A \in \overline{\mathfrak{T}}_v$,*

$$E_H\left[|Y_\varepsilon(v) - Y_\varepsilon'(v)|^2 1_A\right] \leq \lambda(\delta_2) E_H\left[|U_1 - U_2|^2 1_A\right] + \delta_2. \tag{41}$$

PROOF. Let the C^2 function Ψ: $V_Y \times V_Y \to [0, \infty)$ be as in Condition 5.1

and let $\tilde{Y}_\varepsilon = (Y_\varepsilon, Y'_\varepsilon)$. For any stochastic interval $]\!]\sigma, \tau]\!]$, we may write as in (26),

$$\begin{aligned}
&\Psi(\tilde{Y}_\varepsilon(\tau)) - \Psi(\tilde{Y}_\varepsilon(\sigma)) \\
&= \int_{(\sigma,\tau]} d\Psi(\tilde{Y}_\varepsilon)\begin{bmatrix} Z_\varepsilon\, dW \\ Z'_\varepsilon\, dW' \end{bmatrix} \\
&\quad + \int_{(\sigma,\tau]} 1_{H^c}\left\{\nabla^{(2)} d\Psi\left(\tilde{Y}_\varepsilon, \begin{bmatrix} \bar{Z}_\varepsilon \cdot \bar{Z}_\varepsilon & \bar{Z}_\varepsilon \cdot \bar{Z}'_\varepsilon \\ \bar{Z}'_\varepsilon \cdot \bar{Z}_\varepsilon & \bar{Z}'_\varepsilon \cdot \bar{Z}'_\varepsilon \end{bmatrix}\right)\right. \\
&\qquad\qquad \left. + D^2\Psi(\tilde{Y}_\varepsilon)\begin{bmatrix} \hat{Z}_\varepsilon \cdot \hat{Z}_\varepsilon & \hat{Z}_\varepsilon \cdot \hat{Z}'_\varepsilon \\ \hat{Z}'_\varepsilon \cdot \hat{Z}_\varepsilon & \hat{Z}'_\varepsilon \cdot \hat{Z}'_\varepsilon \end{bmatrix}\right\}\frac{dt}{2} \qquad (42) \\
&\quad + \int_{(\sigma,\tau]} 1_{H}\left\{\nabla^{(2)} d\Psi\left(\tilde{Y}_\varepsilon, \begin{bmatrix} \bar{Z}_\varepsilon \cdot \bar{Z}_\varepsilon & 0 \\ 0 & \bar{Z}'_\varepsilon \cdot \bar{Z}'_\varepsilon \end{bmatrix}\right)\right. \\
&\qquad\qquad \left. + D^2\Psi(\tilde{Y}_\varepsilon)\begin{bmatrix} \hat{Z}_\varepsilon \cdot \hat{Z}_\varepsilon & 0 \\ 0 & \hat{Z}'_\varepsilon \cdot \hat{Z}'_\varepsilon \end{bmatrix}\right\}\frac{dt}{2},
\end{aligned}$$

where $\hat{Z}_\varepsilon \equiv Z_\varepsilon\{h(\varepsilon, Z_\varepsilon)/(1 + h(\varepsilon, Z_\varepsilon))\}^{1/2}$, and similarly for $\hat{Z}'_\varepsilon$. This expression follows from (40) because the increments of W and W' are independent on H and identical otherwise. The boundedness of $(Y_\varepsilon, Y'_\varepsilon)$ (by Lemma 3.3) and the fact that Z_ε and Z'_ε are in $H^2(0, T; V_Z)$ ensure that all the stochastic integrals above have finite expectation. Double convexity makes the second and third stochastic integrals nonnegative, while the first is a martingale. Thus $\int 1_{]\!]\sigma,\tau]\!]}\, d(\Psi \circ \tilde{Y}_\varepsilon)$ is a nonnegative submartingale and, in particular,

$$E_H\big[\Psi(\tilde{Y}_\varepsilon(\tau)) - \Psi(\tilde{Y}_\varepsilon(\sigma))\big|\mathfrak{T}_\alpha\big] \geq 0, \qquad 0 \leq \sigma \leq \tau \leq T. \tag{43}$$

Let ρ denote $\sup\{|x - y|^2 : (x, y) \in G \times G\}$, and given $\delta > 0$, define

$$R_\delta \equiv \{(x, y) \in G \times G : |x - y|^2 \leq \delta/(2\rho)\}.$$

By virtue of (39), the following quantities are well-defined:

$$p(\delta) \equiv \sup\left\{\frac{|x - y|^2}{\Psi(x, y)} : (x, y) \in G \times G - R_\delta\right\},$$

$$q(\delta) \equiv \sup\left\{\frac{\Psi(x, y)}{|x - y|^2} : (x, y) \in G \times G - R_\delta\right\}.$$

Thus for every $\{\mathfrak{T}_t\}$-stopping time $v \leq t$ and every $A \in \mathfrak{T}_v$,

$$\begin{aligned}
E_H\big[|Y_\varepsilon(v) - Y'_\varepsilon(v)|^2 1_A\big] &\leq \delta/2 + E_H\big[|Y_\varepsilon(v) - Y'_\varepsilon(v)|^2 1_A, \tilde{Y}_\varepsilon(v) \notin R_\delta\big] \\
&\leq \delta/2 + p(\delta) E_H\big[\Psi(Y_\varepsilon(v), Y'_\varepsilon(v)) 1_A\big] \\
&\leq \delta/2 + p(\delta) E_H\big[\Psi(U, U') 1_A\big],
\end{aligned}$$

using (43). However, by a similar reasoning,

$$E_H\left[\Psi(U,U')1_A\right] \le \delta'/2 + q(\delta')E_H\left[|U-U'|^2 1_A\right].$$

Choosing δ' small enough so that $p(\delta)\delta' \le \delta$ now gives the result. □

COROLLARY 5.4 (A uniform bound in probability). *Given $a > 0$, let $v \equiv \inf\{t \ge 0\colon |Y_\varepsilon(t) - Y'_\varepsilon(t)| \ge a\} \wedge T$. Given $\delta_2 > 0$, there exists a constant $\lambda(\delta_2) \equiv \lambda(\delta_2, g, \Gamma)$ such that, for every $\{\mathfrak{T}_t\}$-progressively measurable set H and every $\varepsilon > 0$,*

$$P_H(v < T) \le \left\{\lambda(\delta_2)E_H\left[|U_1 - U_2|^2; v < t\right] + \delta_2\right\}/a^2. \tag{44}$$

PROOF. Clearly $E_H[|Y_\varepsilon(v) - Y'_\varepsilon(v)|^2;\ v < T] \ge a^2 P_H(v < T)$. Taking $A \equiv \{v < T\}$ in Proposition 5.3 gives the result. □

COROLLARY 5.5 (A result for gamma-martingales). *Assume Conditions 3.2.2 and 5.1.2. Formula (41) applies to bounded Γ-martingales Y and Y' with $\mathfrak{T}_\infty$-measurable terminal values U_1 and U_2 in G, respectively.*

PROOF. It follows from Corollary 3.5 that Y and Y' spend their whole lifetime in G and are H^2 Γ-martingales. The proof goes as before, taking $H \equiv \phi$, $h \equiv 0$ and $T \equiv \infty$. Many of the details are in the proof of the next corollary. □

COROLLARY 5.6 (Limit of a sequence of gamma-martingales). *Assume Conditions 3.2.2 and 5.1.2. Suppose $T \le \infty$ and for $0 < \varepsilon < 1$, $\{M_\varepsilon(t), 0 \le t \le T\}$ is a bounded Γ-martingale with values in the open set $B \supset G$ appearing in Condition 5.1, with the property*

$$E\left[|M_\varepsilon(T) - U|^2\right] \to 0 \quad \text{as } \varepsilon \to 0 \tag{45}$$

for some $\mathfrak{T}_T$-measurable U in G. Then there exists a Γ-martingale M with values in G such that $M(T) = U$ and $P(\sup\{|M_\varepsilon(t) - M(t)|\colon 0 \le t \le T\} \ge a) \to 0$ as $\varepsilon \to 0$ for each $a > 0$.

PROOF. Let us mimic the coupling construction (40), taking $H \equiv \varnothing$, so that in effect $W = W'$; thus the distinction between P and P_H need no longer be maintained. Let us use an η for the bottom row instead of an ε, for there is no need for the parameters in the top and bottom to coincide. We obtain an expression of the form

$$\begin{bmatrix} M_\varepsilon(T) \\ M_\eta(T) \end{bmatrix} = \begin{bmatrix} M_\varepsilon(t) \\ M_\eta(t) \end{bmatrix} + \int_{(t,T]} \begin{bmatrix} Z_\varepsilon\, dW \\ Z_\eta\, dW \end{bmatrix} - \int_{(t,T]} \begin{bmatrix} \Gamma(M_\varepsilon, Z_\varepsilon \cdot Z_\varepsilon) \\ \Gamma(M_\eta, Z_\eta \cdot Z_\eta) \end{bmatrix} \frac{ds}{2}.$$

All the calculations of Proposition 5.3 go through exactly as before, except that $h = 0$ and so we do not need any condition on $D^2\Psi$. Each Z_ε is in

$H^2(0, T; V_Z)$ by Corollary 3.5. Also we replace G by $\overline{B}$ from (43) onward. Given $a > 0$, let

$$v = \inf\{t \geq 0 : |M_\varepsilon(t) - M_\eta(t)| \geq a\} \wedge T.$$

Corollary 5.4 gives the following assertion: given $\delta_2 > 0$, there exists a constant $\lambda(\delta_2) \equiv \lambda(\delta_2, B, \Gamma)$ such that

$$(46) \qquad P(v < T) \leq \left\{\lambda(\delta_2) E\left[|M_\varepsilon(T) - M_\eta(T)|^2\right] + \delta_2\right\} \Big/ a^2.$$

However, it follows from (45) that

$$(47) \qquad E\left[|M_\varepsilon(T) - M_\eta(T)|^2\right] \to 0 \quad \text{as } \varepsilon, \eta \to 0.$$

Together (46) and (47) say that, for every $a > 0$,

$$P\left(\sup\{|M_\varepsilon(t) - M_\eta(t)| : 0 \leq t \leq T\} \geq a\right) \to 0 \quad \text{as } \varepsilon, \eta \to 0.$$

In other words, $\{M_\varepsilon, 0 < \varepsilon < 1\}$ forms a Cauchy family of continuous $\{\mathfrak{T}_t\}$-adapted processes with respect to the topology of uniform convergence in probability on $[0, T]$. There exists a subsequence which is Cauchy with respect to uniform convergence a.s., and which therefore has a unique continuous $\{\mathfrak{T}_t\}$-adapted process M on $[0, T]$ as its limit; all such subsequences have the same limit a.s. Necessarily we have, for all $a > 0$,

$$(48) \qquad \lim_{\varepsilon \to 0} P\left(\sup\{|M_\varepsilon(t) - M(t)| : 0 \leq t \leq T\} > a\right) = 0.$$

It follows from (48) that the terminal value $M(T)$ must be U. The fact that M is a Γ-martingale follows from Theorem 4.43 of Emery and Meyer (1989), which says that the limit under uniform convergence in probability of any sequence of Γ-martingales is a Γ-martingale. The fact that M takes values in G follows from Corollary 3.5. □

6. Estimates for approximating gamma-martingales.

6.1 *A stronger Lipschitz condition.* Assume that there exist constants $c, c' > 0$ such that, for all $z \in V_Z$,

$$(49) \qquad |\Gamma(y, z \cdot z) - \Gamma(y', z \cdot z)| \leq c|y - y'| \, \|z \cdot z\| \qquad \forall\, y, y' \in V_Y,$$

$$(50) \qquad |\Gamma(y, z \cdot z) - \Gamma(y, 0)| \leq c' \|z \cdot z\| \qquad \forall\, y \in G.$$

Also there exist constants $c(r) > 0$ for each $r > 0$ such that, for all $z, z' \in V_Z$ with $\|z \cdot z\| \leq r$ and $\|z' \cdot z'\| \leq r$,

$$(51) \qquad |\Gamma(y, z \cdot z) - \Gamma(y, z' \cdot z')| \leq c(r) \|z \cdot z - z' \cdot z'\| \qquad \forall\, y \in V_Y.$$

REMARK 6.1.1. (i) Condition 6.1 implies Condition 3.1.

(ii) Inequality (49) (taking $z \cdot z = 0$) implies that $\Gamma(y, 0)$ is the same for all y, and (50) implies that, for $y, y' \in G$,

$$(52) \qquad \begin{aligned} |\Gamma(y, z \cdot z) - \Gamma(y', z' \cdot z')| &\leq |\Gamma(y, z \cdot z) - \Gamma(y, 0)| \\ &\quad + |\Gamma(y', 0) - \Gamma(y', z' \cdot z')| \\ &\leq c'(\|z \cdot z\| + \|z' \cdot z'\|). \end{aligned}$$

(iii) Condition 6.1 is satisfied when Γ is a linear connection whose Christoffel symbols are bounded and Lipschitz.

The following proposition gives the crucial technical tool for controlling the solutions to (23) as $\varepsilon \to 0$.

PROPOSITION 6.2 (A uniform integrability property). *Suppose Γ satisfies Conditions 3.1, 3.2, 5.1 and (52), and that the range of U is contained in the compact set G specified in Condition 3.2, a.s. Then the solution $(Y_\varepsilon, Z_\varepsilon)$ to (23) satisfies*

$$\int_{(0,T]} 1_{\{|Z_\varepsilon|^2 \geq 1/\varepsilon\}} |Z_\varepsilon|^2 \, dt \to 0 \quad \textit{in probability as } \varepsilon \to 0.$$

To be precise, there exists a family of $\{\overline{\mathfrak{T}}_t\}$-stopping times $\{v(\varepsilon),\ 0 < \varepsilon < 1\}$ such that, for $H \equiv \{|Z_\varepsilon|^2 \geq 1/\varepsilon\}$, the following occur as $\varepsilon \to 0$:

$$P_H(v(\varepsilon) < T) \to 0; \tag{53}$$

$$E_H\left[\int_{(0,v(\varepsilon)]} 1_{\{|z_\varepsilon|^2 \geq 1/\varepsilon\}} |z_\varepsilon|^2 \, dt\right] \to 0. \tag{54}$$

PROOF. *Step* 1. Given any $\mathfrak{T}_T$-measurable random variable $U \equiv U(W)$ on $(\Omega, \mathfrak{T}, P)$ with values in G, let us take an identical copy $U \equiv U'(W')$ on $(\Omega', \mathfrak{T}', P')$, and consider the solution to the system (40) in the case where $(U_1, U_2) \equiv (U, U')$ and where the $\{\mathfrak{T}_t\}$-progressively measurable set H (which can also be regarded as $\{\overline{\mathfrak{T}}_t\}$-progressively measurable) is chosen to be

$$H \equiv \left\{|Z_\varepsilon|^2 \geq 1/\varepsilon\right\}.$$

It follows immediately from our previous estimate in Proposition 3.4 that

$$E\left[\int_{(0,T]} 1_H \, dt\right] \leq \varepsilon E\left[\int_{(0,T]} |Z_\varepsilon|^2 \, dt\right] \leq \varepsilon c_1,$$

where

$$c_1 \equiv c_1(\Gamma, G) \equiv \sup\left\{\left[\int_{(0,T]} |Z_\varepsilon|^2 \, dt\right] : 0 < \varepsilon < 1\right\}. \tag{55}$$

We see from (32) that for any $\delta_1 > 0$ there exists $\mu(\delta_1)$ such that

$$E_H\left[|U - U'|^2\right] \leq \mu(\delta_1)\varepsilon c_1 + \delta_1. \tag{56}$$

Step 2. Next observe that from (40) we obtain the forward SDE

$$dY_\varepsilon(t) - dY'_\varepsilon(t) = Z_\varepsilon \, dW - Z'_\varepsilon \, dw' - \left\{\Gamma\left(Y_\varepsilon, \overline{Z}_\varepsilon \cdot \overline{Z}_\varepsilon\right) - \Gamma\left(Y'_\varepsilon, Z'_\varepsilon \cdot \overline{Z}'_\varepsilon\right)\right\}\frac{dt}{2}.$$

Itô's formula now gives

$$d|Y_\varepsilon(t) - Y'_\varepsilon(t)|^2 = \left(|Z_\varepsilon|^2 + |Z'_\varepsilon|^2 - 2\,\mathrm{Tr}\{Z_\varepsilon \cdot Z'_\varepsilon\}1_{H^c}\right)dt + 2(Y_\varepsilon(t) - Y'_\varepsilon(t)) \cdot\left(Z_\varepsilon\, dW - Z'_\varepsilon\, dW' - \left\{\Gamma\left(Y_\varepsilon, \bar{Z}_\varepsilon \cdot \bar{Z}_\varepsilon\right) - \Gamma\left(Y'_\varepsilon, \bar{Z}'_\varepsilon \cdot \bar{Z}'_\varepsilon\right)\right\}\frac{dt}{2}\right). \tag{57}$$

For any $\delta_3 > 0$, we may define an $\{\mathfrak{F}_t\}$-progressively measurable set L by

$$L \equiv \{(t, \omega) : |Y_\varepsilon(t) - Y'_\varepsilon(t)| \le \delta_3/c'\},$$

where c' is the constant appearing in (52). It is evident from (52) that, on the set L,

$$\left|(Y_\varepsilon(t) - Y'_\varepsilon(t))\left\{\Gamma\left(Y_\varepsilon, \bar{Z}_\varepsilon \cdot \bar{Z}_\varepsilon\right) - \Gamma\left(Y'_\varepsilon, \bar{Z}'_\varepsilon \cdot \bar{Z}'_\varepsilon\right)\right\}\right| \le \delta_3\left(|Z_\varepsilon|^2 + |Z'_\varepsilon|^2\right). \tag{58}$$

The process $\int(Y_\varepsilon(t) - Y'_\varepsilon(t))(Z_\varepsilon\, dW - Z'_\varepsilon\, dW')$ is a martingale, since $(Y_\varepsilon, Y'_\varepsilon)$ spends its whole lifetime in $G \times G$. Thus (57) shows that, for $\delta_3 \le 1$,

$$\int 1_{H \cap L}\, d|Y_\varepsilon(t) - Y'_\varepsilon(t)|^2$$

is a submartingale. Moreover, for the $\{\mathfrak{F}_t\}$-stopping time

$$v \equiv \inf\{t \ge 0 : |Y_\varepsilon(t) - Y'_\varepsilon(t)| \ge \delta_3/c'\} \wedge T, \tag{59}$$

it is clear that stochastic interval $]\!]0, v]\!] \subseteq L$, and so for every $\delta_3 \le 1/2$,

$$E_H\left[\int_{(0,v]} 1_H\left(|Z_\varepsilon|^2 + |Z'_\varepsilon|^2\right)dt/2\right] \le E_H\left[\int_{(0,v]} 1_H\, d|Y_\varepsilon(t) - Y'_\varepsilon(t)|^2\right]. \tag{60}$$

Replacing H by H^c, we also obtain from (58) by a similar argument that

$$E_H\left[\int_{(0,v]} 1_{H^c} d|Y_\varepsilon(t) - Y'_\varepsilon(t)|^2\right] \ge -\delta_3 E_H\left[\int_{(0,v]} 1_{H^c}\left(|Z_\varepsilon|^2 + |Z'_\varepsilon|^2\right)dt\right]. \tag{61}$$

Step 3. To prove the proposition, with the stopping time v in (59) playing the role of $v(\varepsilon)$, we use the following estimate. From Proposition 5.3 we see that for any $\delta_2 > 0$ there is a $\lambda(\delta_2) \equiv \lambda(\delta_2, G, \Gamma)$, not depending on δ_3, such that

$$E_H\left[|Y_\varepsilon(v) - Y'_\varepsilon(\eta)|^2\right] \le \lambda(\delta_2)E_H\left[|U - U'|^2\right] + \delta_2. \tag{62}$$

By virtue of (56), we obtain

$$E_H\left[|Y_\varepsilon(v) - Y'_\varepsilon(v)|^2\right] \le \lambda(\delta_2)[\mu(\delta_1)\varepsilon c_1 + \delta_1] + \delta_2. \tag{63}$$

As in Corollary 5.4, we have

$$(\delta_3/c')^2 P_H(v < T) \le E_H\big[|Y_\varepsilon(v) - Y'_\varepsilon(v)|^2\big],$$

and so (63) gives

$$P_H(v < T) \le (c'/\delta_3)^2\big(\lambda(\delta_2)[\mu(\delta_1)\varepsilon c_1 + \delta_1] + \delta_2\big). \tag{64}$$

On the other hand, we have from (60) and (61) that, for $\delta_3 \le 1/2$,

$$\begin{aligned} E\left[\int_{(0,v]} 1_H |Z_\varepsilon|^2\, dt/2\right] &\le E_H\left[\int_{(0,v]} 1_H\, d|Y_\varepsilon(t) - Y'_\varepsilon(t)|^2\right] \\ &\le E_H\big[|Y_\varepsilon(v) - Y'_\varepsilon(v)|^2\big] \\ &\quad - E_H\left[\int_{(0,v]} 1_{H^c}\, d|Y_\varepsilon(t) - Y'_\varepsilon(t)|^2\right] \\ &\le E_H\big[|Y_\varepsilon(v) - Y'_\varepsilon(v)|^2\big] \\ &\quad + \delta_3 E_H\left[\int_{(0,v]} 1_{H^c}\big(|Z_\varepsilon|^2 + |Z'_\varepsilon|^2\big)\, dt\right]. \end{aligned}$$

Thus by formulas (55) (which of course applies to Z'_ε as well) and (63),

$$E\left[\int_{(0,v]} 1_H |Z_\varepsilon|^2\, dt\right] \le 2\{\lambda(\delta_2)[\mu(\delta_1)\varepsilon c_1 + \delta_1] + \delta_2\} + 4\delta_3 c_1. \tag{65}$$

By choosing δ_3, δ_2, δ_1 and ε (in that order) to be sufficiently small, formulas (64) and (65) show that the left sides of both (53) and (54) may be made arbitrarily small. □

6.3. *Approximating gamma-martingales.* Assume Condition 6.1. To each solution $(Y_\varepsilon, Z_\varepsilon)$ to (23) there corresponds a Γ-martingale $M_\varepsilon \equiv \{M_\varepsilon(t), \mathfrak{T}_t\}$ obtained by solving the following *forward* SDE:

$$M_\varepsilon(t) = Y_\varepsilon(0) + \int_{(0,t]} \bar{Z}_\varepsilon\, dW - \int_{(0,t]} \Gamma\big(M_\varepsilon, \bar{Z}_\varepsilon \cdot \bar{Z}_\varepsilon\big)\, ds/2, \qquad 0 \le t \le T, \tag{66}$$

where $\bar{Z}_\varepsilon \equiv Z_\varepsilon/(1 + h(\varepsilon, Z_\varepsilon))^{1/2}$ as in (40). Such a solution exists for all $0 \le t \le T$ because the first stochastic integral is an L^2 martingale and the integrand in the second integral can be expressed as $b\, ds + a(M_\varepsilon, s)\, dF_\varepsilon(s)$, where b is the constant value of $2\Gamma(y, 0)$ (see Remarks 6.1.1),

$$y \to a(y, s) \equiv \frac{\Gamma\big(y, \bar{Z}_\varepsilon \cdot \bar{Z}_\varepsilon\big) - \Gamma(y, 0)}{2\|\bar{Z}_\varepsilon \cdot \bar{Z}_\varepsilon\|} \tag{67}$$

is globally Lipschitz by Conditions 6.1 and

$$F_\varepsilon(t) \equiv \int_{(0,t]} \|\bar{Z}_\varepsilon \cdot \bar{Z}_\varepsilon\|\, ds, \tag{68}$$

which is a bounded, increasing process. Note carefully that M_ε has the same initial value as Y_ε but need not have the same terminal value, since the martingale parts of M_ε and Y_ε are different. Indeed, there is no guarantee that M_ε is confined to the compact set G. The strategy will be to show that the terminal value of M_ε converges in probability to U as $\varepsilon \to 0$, and that the family $\{M_\varepsilon,\ 0 < \varepsilon < 1\}$ converges uniformly in probability on $[0,T]$ to a limiting Γ-martingale M with terminal value U.

Although M_ε is adapted to the smaller filtration $\{\mathfrak{T}_t\}$, it will be convenient to work within the product probability space and the larger filtration $\{\overline{\mathfrak{T}}_t\}$ introduced in Section 4, taking $H \equiv \{|Z_\varepsilon|^2 \geq 1/\varepsilon\}$ as in Proposition 6.2.

LEMMA 6.4 (Stability of forward SDE). *Suppose Γ satisfies Conditions 3.2, 5.1 and 6.1, and that the range of U is contained in the compact set G specified in Condition 3.2, a.s. Then, for all $\eta > 0$,*

$$P\big(\sup\{|Y_\varepsilon(t) - M_\varepsilon(t)|: 0 \leq t \leq T\} > \eta\big) \to 0 \quad \text{as } \varepsilon \to 0.$$

PROOF. We are working with a pair of stochastic differential equations which may be written concisely in the notation of (67) and (68) as

$$Y_\varepsilon(t) = \zeta_\varepsilon(t) - bt - \int_{(0,t]} a(Y_\varepsilon, s)\, dF_\varepsilon(s), \tag{69}$$

$$M_\varepsilon(t) = \zeta_\varepsilon'(t) - bt - \int_{(0,t]} a(M_\varepsilon, s)\, dF_\varepsilon(s). \tag{70}$$

Observe that, by Condition 6.1, $|a(y,s) - a(y',s)| \leq c|y - y'|$ for all s, for c as in (49). The V_Y-valued martingales ζ_ε and ζ_ε' satisfy $\zeta_\varepsilon(0) = \zeta_\varepsilon'(0)$ and also

$$\zeta_\varepsilon(t) - \zeta_\varepsilon'(t) = \int_{(0,t]} \big(Z_\varepsilon - \overline{Z}_\varepsilon\big)\, dW$$
$$= \int_{(0,t]} Z_\varepsilon\Big\{1 - (1 + h(\varepsilon, Z_\varepsilon))^{-1/2}\Big\}\, dW.$$

Observe that, by (19),

$$0 \leq \Big\{1 - (1 + h(\varepsilon, Z_\varepsilon))^{-1/2}\Big\} \leq 1_{\{|Z_\varepsilon|^2 \geq 1/\varepsilon\}}.$$

Thus for the $\{\overline{\mathfrak{T}}_t\}$-stopping time $v \equiv v(\varepsilon)$ as in Proposition 6.2, we have

$$E_H\big[|\zeta_\varepsilon(v) - \zeta_\varepsilon'(v)|^2\big]$$
$$\leq E_H\bigg[\int_{(0,v]} 1_{\{|Z_\varepsilon|^2 \geq 1/\varepsilon\}} |Z_\varepsilon|^2\, dt\bigg] \to 0$$

as $\varepsilon \to 0$. By Doob's L^2 inequality,

$$\rho_\varepsilon \equiv E\big[\sup\{|\zeta_\varepsilon(t) - \zeta_\varepsilon'(t)|^2: 0 \leq t \leq v\}\big] \to 0 \quad \text{as } \varepsilon \to 0. \tag{71}$$

From this point on the proof follows the basic outline of Theorem 3.7.5 of Métivier and Pellaumail (1980), so we shall give the main points only. It follows from (3) that the process F_ε is "*-dominated" [in the sense of Métivier and Pellaumail (1980)] by the process

$$A_\varepsilon(t) \equiv \int_{(0,t]} |Z_\varepsilon|^2\,ds,$$

implying that for every Hilbert-valued bounded progressively measurable process ψ and every $\{\overline{\mathfrak{T}}_t\}$-stopping time $\tau \le T$ we have

$$E_H\left[\sup\left\{\left|\int_{(0,t]} \psi\|\overline{Z}_\varepsilon \cdot \overline{Z}_\varepsilon\|\,ds\right|^2 : 0 \le t < \tau\right\}\right] \le E_H\left[A_\varepsilon(\tau)\int_{(0,\tau]} |\psi|^2\,dA_\varepsilon\right]. \tag{72}$$

Introduce $\{\overline{\mathfrak{T}}_t\}$-stopping times $\{\nu(\varepsilon,n)\}$ by

$$\nu(\varepsilon,n) \equiv \inf\{t \ge 0 \colon A_\varepsilon(t) \ge n\} \wedge T. \tag{73}$$

It follows from Proposition 3.4 that $E[A_\varepsilon(T)]$ is bounded uniformly in ε and, therefore, $\sup\{P(\nu(\varepsilon,n) < T)\colon 0 < \varepsilon < 1\} \to 0$ as $n \to \infty$. Define $\{\overline{\mathfrak{T}}_t\}$-stopping times $\sigma(\varepsilon,n) \equiv v(\varepsilon) \wedge \nu(\varepsilon,n)$, with reference to (73) and Proposition 6.2. Thus

$$\sup\{P(\sigma(\varepsilon,n) < T) : 0 < \varepsilon < 1\} \to 0. \tag{74}$$

Let X_t^* denote $\sup\{|X_s|\colon 0 \le s < t\}$. Subtract (70) from (69), take the supremum up to an arbitrary $\{\overline{\mathfrak{T}}_t\}$-stopping time $\tau \le T$, take the second moment and apply (72) to obtain

$$E_H\left[\{|Y_\varepsilon - M_\varepsilon|_\tau^*\}^2\right] \le 2E_H\left[\{|\zeta_\varepsilon - \zeta_\varepsilon'|_\tau^*\}^2\right] + 2E_H\left[A_\varepsilon(\tau)\int_{(0,\tau]} |a(Y_\varepsilon,s) - a(M_\varepsilon,s)|^2\,dA_\varepsilon\right].$$

If $\varphi_s \equiv \sup\{|Y_\varepsilon(t) - M_\varepsilon(t)|^2\colon 0 \le t \le s\}$, it follows that, for $\tau \le \sigma(\varepsilon,n)$,

$$E_H[\varphi_\tau] \le 2\rho_\varepsilon + 2nc^2 E_H\left[\int_{[0,\tau]} \varphi_s\,dA_\varepsilon\right],$$

using the notation of (71). By Lemma 3.7.1 of Métivier and Pellaumail (1980), it follows that there exists a real number $\gamma \equiv \gamma(n,c)$ such that

$$E_H[\varphi_{\sigma(\varepsilon,n)}] \le \rho_\varepsilon \gamma(n,c).$$

So given $\eta > 0$, first choose n_0 so large that $P(\sigma(\varepsilon,n_0) < T) \le \eta/2$ for all ε, using (74), and then invoke (71) to choose ε_0 so small that $\rho_\varepsilon\gamma(n_0,c) \le \eta^3/2$ for $\varepsilon \le \varepsilon_0$. Now

$$P\big(\sup\{|Y_\varepsilon(t) - M_\varepsilon(t)|\colon 0 \le t \le T\} > \eta\big) \le \eta$$

for all $\varepsilon \le \varepsilon_0$. □

7. Existence and uniqueness theorem. The basic result of this paper is the following.

THEOREM 7.1 (Solution under convexity conditions). *Suppose* $T \le \infty$, *that* Γ *satisfies Conditions* 3.2, 5.1 *and* 6.1 *and that the range of* U *is contained in the compact set* G *specified in Condition* 3.2, *a.s. Then there exists a unique progressively measurable solution* (Y, Z) *to the Pardoux–Peng system* (11) *associated with* (W, Γ, U) *such that* $Z \in H^2(0, T; V_Z)$. *In particular,* Y *is an* H^2 Γ-*martingale with terminal value* U. *Moreover*:

(i) $Y(t) \in G$ *for all* t, *a.s.*
(ii) $E[\int_{(0,\infty)} |Z|^2\, ds]$ *satisfies the bound of Corollary* 3.5.
(iii) *The process* Y *depends continuously on the terminal value* U *in the sense of Corollary* 5.5.

PROOF. *Uniqueness*. The technique is standard from Emery (1985), but let us sketch it for the unfamiliar case of nonlinear connections. If there are two solutions (Y, Z) and (Y', Z), one writes out the Itô formula for $\Psi(Y, Y')$, giving an expression similar to (42) but without the $D^2\Psi$ terms. The convexity of Ψ (Condition 5.1) shows that

$$\Psi(Y(t), Y'(t)) - \int_{(0,t]} \{D_1\Psi(Y) Z\, dW + D_2\Psi(Y') Z'\, dW\}$$

is an integrable increasing process, and the stochastic integral above is a martingale. Thus $\Psi(Y, Y')$ is a bounded nonnegative submartingale converging to zero almost surely, and therefore identically zero. In view of the fact that Ψ vanishes only on the diagonal (Condition 5.1), this shows that Y and Y' are identical (and so are the associated Z and Z' by the reasoning below).

Existence. (*Step* 1). In the first step of the proof, take $T < \infty$. We wish to construct Γ-martingales on the time interval $[0, T]$ whose terminal values converge to U in L^2 as $\varepsilon \to 0$. Let $B \supset G$ be the open set appearing in Condition 5.1, which may be bounded if necessary so as to fit inside the Euclidean ball of radius r, for some $r > 0$. For each Γ-martingale M_ε introduced in Section 6.3, let

$$\tau'(\varepsilon) \equiv \inf\{t \ge 0: M_\varepsilon(t) \notin B\} \wedge T,$$

which is an $\{\mathfrak{F}_t\}$-stopping time, and let M'_ε denote M_ε stopped at $\tau'(\varepsilon)$; this is also a Γ-martingale. From the definition of r,

$$E\left[|M'_\varepsilon(T) - U|^2\right] \le 4r^2 P(\tau'(\varepsilon) < T) + E\left[|M_\varepsilon(T) - U|^2; \tau'(\varepsilon) = T\right].$$

Since $\sup\{|Y_\varepsilon(t) - M_\varepsilon(t)|: 0 \le t \le T\}$ goes to zero in probability by Lemma 6.4 and since Y_ε spends its whole life in G, it follows that

$$P(\tau'(\varepsilon) < T) \to 0 \quad \text{as } \varepsilon \to 0.$$

Lemma 6.4 also implies that $M_\varepsilon(T)$ converges to U in probability as $\varepsilon \to 0$. Therefore, given any $\delta > 0$, we can find ε_0 such that, for all $\varepsilon \le \varepsilon_0$,

$4r^2P(\tau'(\varepsilon) < T) \le \delta/4$ and $P(A) \le \delta/(16r^2)$, where $A \equiv \{|M_\varepsilon(T) - U|^2 > \delta/2\}$. Thus, for $\varepsilon \le \varepsilon_0$,

$$E\big[|M_\varepsilon(T) - U|^2; \tau'(\varepsilon) = T\big] \le \delta/2 + E\big[|M_\varepsilon(T) - U|^2; A \cap \{\tau'(\varepsilon) = T\}\big].$$

By the fact that U lies in G, the last expectation is bounded above by $\delta/4$, and we deduce that, for $\varepsilon \le \varepsilon_0$,

$$E\big[|M'_\varepsilon(T) - U|^2\big] \le \delta. \tag{75}$$

In other words, $\{M'_\varepsilon, 0 < \varepsilon < 1\}$ is a family of B-valued $\{\mathfrak{X}_t\}$-adapted Γ-martingales on the time interval $[0, T]$ whose terminal values converge to U in L^2 as $\varepsilon \to 0$. Now we apply Corollary 5.6 to the family $\{M'_\varepsilon, 0 < \varepsilon < 1\}$ to obtain a limiting G-valued Γ-martingale M with $M(T) = U$.

(*Step* 2). Finally we consider the case $T = \infty$. Recall from Remark 3.2.1 that G is convex in the Euclidean sense and, therefore, $E[U|\mathfrak{X}_n] \in G$ a.s. By the previous step, there exists a Γ-martingale $Y^{(n)}$ with terminal value $E[U|\mathfrak{X}_n]$ and which takes values in G. Since $E[U|\mathfrak{X}_n] \to U$ in L^2, we apply Corollary 5.6 again to obtain a limiting Γ-martingale Y with $Y(T) = U$. To construct the progressively measurable process Z as in (11), apply the Itô representation theorem to the V_Y-valued continuous local martingale $Y + (1/2)\int \Gamma(Y, (dY \otimes dY)/ds)\, ds$. Thus

$$\int_{(0,\infty)} Z\, dW \equiv U - Y(0) + (1/2)\int_{(0,\infty)} \Gamma(Y, (dY \otimes dY)/ds)\, ds.$$

It follows from Corollary 3.5 that Y is an H^2 Γ-martingale with values in G. This completes the construction of (Y, Z) satisfying (11), for the case where $T = \infty$. The assertions (ii) and (iii) follow from Corollaries 3.5 and 5.5. □

CONJECTURE 7.2. *The existence* (*but not uniquesness*) *assertion of Theorem* 7.1 *remains valid without Condition* 5.1 (*i.e., Conditions* 3.2 *and* 6.1 *suffice*). *Such a solution satisfies* (i) *and* (ii), *but possibly not* (iii).

8. Some generations and applications.

8.1. *Application to nonlinear elliptic PDE's.* Suppose (N, g) is a Riemannian manifold, and X is a Brownian motion with drift on (N, g), with generator L. Alternatively, think of X as a diffusion process on a Euclidean vector space V_X, which solves the SDE

$$dX(t) = b(X(t))\, dt + \sigma(X(t))\, dW(t), \qquad X(0) = x, \tag{76}$$

where $\sum \sigma_p^i \sigma_p^j = g^{ij}$, the inverse metric tensor. Let K be a compact submanifold with boundary of N, with boundary ∂K, and for $x \in K$, let ζ denote the first time that X hits the boundary; ζ is assumed finite a.s. Given a nonlinear connection Γ as in (4) and a continuous (or perhaps more regular) mapping

$$\bar{\varphi}: \partial K \to G \subset V_Y,$$

we wish to find a mapping φ: $K \to G$ which solves the following Dirichlet problem:

$$(77)\quad L\varphi(x) + \tfrac{1}{2}\Gamma\Big(\varphi(x), \sum_{i,j} D_i\varphi(x)g^{ij}(x)\, D_j\varphi(x)\Big) = 0, \qquad x \in K,$$

$$(78)\qquad \varphi(x) = \bar{\varphi}(x), \qquad x \in \partial K.$$

In the case where $L = (1/2)\Delta$ and Γ is a linear connection, (77) becomes the familiar condition that φ: $(K, g) \to (G, \Gamma)$ be a harmonic mapping [see Eells and Lemaire (1978, 1988)], namely,

$$(79)\qquad \Delta\varphi^{\alpha} + \sum \Gamma^{\alpha}_{\beta\gamma}(\varphi)\, D_i\varphi^{\beta} g^{ij} D_j\varphi^{\gamma} = 0, \qquad 1 \le \alpha \le m,$$

where Greek indices are used for coordinates in V_Y and $\varphi \equiv (\varphi^1, \dots, \varphi^m)$. The procedure suggested by Kendall (1990, 1994), and more generally Peng (1991), is as follows. Using the same Wiener process W with which we constructed the diffusion X in (76), whose value at time ζ (see above) for $X(0) = x$ is denoted $X_x(\zeta)$, solve the backwards SDE

$$(80)\quad \bar{\varphi}(X_x(\zeta)) = Y_x(t) + \int_{(t,\infty)} Z_x\, dW - (1/2)\int_{(t,\infty)} \Gamma(Y_x, Z_x \cdot Z_x)\, ds$$

and set $\varphi(x) \equiv Y_x(0)$. Note that, under the conditions of Theorem 7.1, $Y_x(0)$ is nonrandom (by adaptedness), unique and lies in G, and provided $x \to X_x(\zeta)$ is continuous in probability [see Métivier and Pellaumail (1980), Section 3.7], φ is continuous by Theorem 7.1, part (iii). The proof that this φ is "finely harmonic," in the sense that it sends L-diffusions to Γ-martingales, is relatively straightforward [Kendall (1990)]. For the case presented in (79), conditions under which φ is smooth are presented in Kendall (1994). In a different context, Pardoux and Peng (1992) use the Malliavin calculus to show that $Y_x(0)$ is a differentiable function of x, but unfortunately under Lipschitz conditions too restrictive for the present application.

8.2. *Terminal value with noncompact range.* To extend Theorem 7.1 to the case where the range of U is not relatively compact, one takes an increasing sequence of compact sets $G_1 \subset G_2 \subset \cdots$ whose union G contains the range of U, and C^2 functions Φ_n, f_n, such that $\{G_n, \Phi_n, f_n\}$ satisfies Condition 3.2 for each n. Assume that Ψ: $V_Y \times V_Y \to [0, \infty)$ is a function with bounded first and second derivatives which is doubly convex on $G \times G$. If U is in L^2, one can take a sequence $\{U_n\}$ of $\mathfrak{T}_\infty$-measurable random variables which converges to U in L^2 and such that the range of U_n is contained in G_n. Theorem 7.1 gives an H^2 Γ-martingale $Y^{(n)}$ with terminal value U_n and which takes values in G_n, for each n. Proposition 4.4 in Darling (1993) shows that there is a Γ-martingale Y, the uniform limit in probability of the $\{Y^{(n)}\}$, with $Y_\infty = U$, using the fact that $\Psi(U_n, U)$ converges to zero in L^1. We shall not write out the full details because a similar result has already been given in Theorem 5.2 of Darling (1993), together with some examples.

8.3. *The time-dependent case.* One can replace the nonlinear connection Γ as in (4) by a time-dependent function

$$\Lambda: [0,\infty) \times V_Y \times (V_Y \otimes V_Y) \to V_Y. \tag{81}$$

A function $\phi \in C^2(V_Y)$ is called Λ-convex on $G \subseteq V_Y$ if, for all $y \in G$,

$$D^2\phi(y)(z \cdot z) - D\phi(y)\Lambda(t, y, z \cdot z) \geq 0 \qquad \forall\, t, \forall\, z \in V_Z. \tag{82}$$

The definition of $\Lambda^{(2)}$-convexity for $\Psi \in C^2(V_Y \times V_Y)$ is analogous. Thus Conditions 3.2 and 5.1 can be interpreted in this case also.

THEOREM 8.4 (For the time-dependent case). *Suppose* Λ: $[0,\infty) \times V_Y \times (V_Y \otimes V_Y) \to V_Y$ *satisfies Conditions* 3.2 *and* 5.1 *in the sense described in* 8.3, *that* $\Lambda(t,\cdot,\cdot)$ *satisfies Condition* 6.1 *uniformly in t and that the range of U is contained in the compact set G specified in Condition* 3.2, *a.s. Then the conclusions of Theorem* 7.1 *apply to the backward SDE* (*for* $t \leq \infty$)

$$U = Y(t) + \int_{(t,T]} Z\, dW - (1/2)\int_{(t,T]} \Lambda(s, Y, Z \cdot Z)\, ds. \tag{83}$$

PROOF. Using Theorem 1.5, the construction of the approximations (23) presents no problem and the rest of the proof goes as before. □

8.5. *Application to nonlinear parabolic PDE's.* The following is based on Peng (1991). Let us proceed as in Application 8.1 except that now the generator L may be time-dependent, that is, of the form

$$L_t\varphi(x) \equiv (1/2)\sum_{i,j} a_{ij}(t,x)\, D_{ij}\varphi(x) + \sum_i b_i(t,x)\, D_i\varphi(x), \tag{84}$$

and likewise the coefficients of the SDE (76) now depend on t, with $\Sigma\sigma_p^{\,i}\sigma_p^{\,j} = a^{ij}$, the inverse of a (assumed to exist). Given $T < \infty$, a mapping Λ as in (81) and a mapping φ_T: $N \to G \subset V_Y$, we wish to find a mapping ψ: $[0,T] \times N \to G$ which solves the following quasilinear parabolic PDE:

$$\frac{\partial\psi}{\partial t} + L_t\psi + \frac{1}{2}\Lambda\big(t, \psi, (D\psi)a^{-1}(D\psi)^T\big) = 0. \tag{85}$$

$$\psi(T,x) = \varphi_T(x). \tag{86}$$

In the case where $L = (1/2)\Delta$ and G is a linear connection, (85) becomes the heat equation for harmonic mappings [see Eells and Lemaire (1978)], namely,

$$2\frac{\partial\psi}{\partial t} + \Delta\psi^\alpha + \sum\Gamma^\alpha_{\beta\gamma}(\psi)\, D_i\psi^\beta\, g^{ij}\, D_j\psi^\gamma = 0, \qquad 1 \leq \alpha \leq m, \tag{87}$$

in the notation of (79). Peng's (1991) proposal is to construct the diffusion X on a time interval $[t,T]$ using the forward SDE

$$X_x(s) = x + \int_{(t,s]} b(r, X_x(r))\, dr + \int_{(t,s]} \sigma(r, X_x(r))\, dW$$

and then to solve the backwards SDE on $[t, T]$:

$$\begin{aligned}\varphi_T(X_x(T)) = Y_x(s) &+ \int_{(s,T]} Z_x \, dW \\ &-(1/2)\int_{(s,T]} \Lambda(r, Y_x, Z_x \cdot Z_x)\, dr.\end{aligned} \tag{88}$$

Since $\{Y_x(s),\ t \le s \le T\}$ is adapted to the increments of W on $[t, T]$, it is clear that $Y_x(t)$ is nonrandom, and Peng (1991) shows that, under conditions more restrictive than ours,

$$\psi(t, x) = Y_x(t)$$

solves the system (85) and (86). The study of the heat equation for harmonic mappings using this method seems to be an open problem.

Acknowledgments. The author wishes to acknowledge the support of the M.R.E. (France) which made possible an extended stay at the Université de Provence where this research was conducted. The advice and encouragement of Etienne Pardoux played a major role in this work. Jean Picard made a valuable contribution by identifying significant errors in an earlier version of this paper, and by supplying a preprint.

REFERENCES

ANTONELLI, F. (1993). Backward–forward stochastic differential equations. *Ann. Appl. Probab.* **3** 777–793.

ARNAUDON, M. (1993). Sur l'existence sur une variété de martingales de valeur terminale donnée. Preprint, Univ. Louis Pasteur, Strasbourg.

DARLING, R. W. R. (1994). *Differential Forms and Connections*. Cambridge Univ. Press.

DARLING, R. W. R. (1995). Martingales on noncompact manifolds: maximal inequalities and prescribed limits. *Ann. Inst. H. Poincaré*. To appear.

DUDLEY, R. M. (1989). *Real Analysis and Probability*. Wadsworth, Belmont, CA.

DUFFIE, D. and EPSTEIN, L. G. (1992a). Stochastic differential utility. *Econometrica* **60** 353–394.

DUFFIE, D. and EPSTEIN, L. G. (1992b). Asset pricing with stochastic differential utility. *Rev. Financial Studies* **5** 411–436.

EELLS, J. and LEMAIRE, L. (1978). A report on harmonic maps. *Bull. London Math. Soc.* **10** 1–68.

EELLS, J. and LEMAIRE, L. (1988). Another report on harmonic maps. *Bull. London Math. Soc.* **20** 385–524.

EMERY, M. (1985). Convergence des martingales dans les variétés. Colloque en l' Hommage de Laurent Schwartz, 2. *Astérisque* **132** 47–63.

EMERY, M. and MEYER, P.-A. (1989). *Stochastic Calculus on Manifolds*. Springer, Berlin.

EMERY, M. and MOKOBODZKI, G. (1991). Sur le barycentre d'une probabilité dans une variété. *Séminaire de Probabilités XXV. Lecture Notes in Math.* **1485** 220–233. Springer, Berlin.

KENDALL, W. S. (1990). Probability, convexity and harmonic maps with small image I: uniqueness and fine existence. *Proc. London Math. Soc.* **61** 371–406.

KENDALL, W. S. (1991). Convex geometry and nonconfluent Γ-martingales I: Tightness and strict convexity. In *Stochastic Analysis: Proceedings of the Durham Symposium on Stochastic Analysis 1990* (M. T. Barlow and N. H. Bingham, eds.) 163–178. Cambridge Univ. Press.

KENDALL, W. S. (1992). The propeller: a counterexample to a conjectured criterion for the existence of certain convex functions. *J. London Math. Soc.* (2) **46** 364–374.

KENDALL, W. S. (1993). Brownian motion and partial differential equations: from the heat equation to harmonic maps. Preprint, Dept. Statistics, Univ. Warwick.

KENDALL, W. S. (1994). Probability, convexity, and harmonic maps II: smoothness via probabilisitic gradient inequalities. *J. Funct. Anal.* To appear.

MÉTIVIER, M. and PELLAUMAIL, J. (1980). *Stochastic Integration*. Academic Press, New York.

MEYER, P.-A. (1981). Géométrie stochastique sans larmes. *Séminaire de Probabilités XV. Lecture Notes in Math.* **850** 44–102. Springer, Berlin.

PARDOUX, E. and PENG, S. (1990). Adapted solution of a backward stochastic differential equation. *System Control Lett.* **14** 55–61.

PARDOUX, E. and PENG, S. (1992). Backward stochastic differential equations and quasilinear parabolic partial differential equations. *Stochastic Partial Differential Equations and Their Applications. Lecture Notes in Control and Inform. Sci.* **176**. Springer, Berlin.

PARDOUX, E. and PENG, S. (1994). Some backward stochastic differential equations with non-Lipschitz coefficients. Preprint 94-3, Université de Provence URA 225.

PENG, S. (1991). Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics* **37** 61–74.

PENG, S. (1993). Backward SDE and related quasi-expectation. Preprint, Université de Provence.

PICARD, J. (1989). Martingales sur le cercle. *Séminaire de Probabilités XXIII. Lecture Notes in Math.* **1372** 147–160. Springer, Berlin.

PICARD, J. (1991). Martingales on Riemannian manifolds with prescribed limit. *J. Funct. Anal.* **99** 223–261.

PICARD, J. (1994). Barycentres et martingales sur une variété. *Ann. Inst. H. Poincaré.* To appear.

REVUZ, D. and YOR, M. (1991). *Continuous Martingales and Brownian Motion*. Springer, Berlin.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTH FLORIDA
TAMPA, FLORIDA 33620-5700