Stability in $\mathbb{D}$ of martingales and backward equations under discretization of filtration

Francois Coquet, Vigirdas Mackevicius, Jean Mémin

IRMAR, Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France
Matematikos fakultetas, Vilnius universitetas, Naugarduko 24, 2600 Vilnius, France

Received 5 June 1997; received in revised form 2 February 1998; accepted 6 February 1998

Abstract

We consider a càdlàg process $Y$, $(\mathcal{F}_t)$ the filtration generated by $Y$ and $(\mathcal{F}_n^t)$ generated by step processes $Y_n$ defined from $Y$ by discretization in time. We study the stability in $\mathbb{D}$ (with Skorokhod topology) of $(\mathcal{F}_n^t)$-martingales and of $(\mathcal{F}_n^t)$-solutions of related backward equations, when $Y_n \to Y$. We get this stability (in law) when $Y$ is Markov and (in probability) under stronger assumptions on the coefficients of equations.

Keywords: Martingales; Backward equations; Skorokhod topology; Convergence in law

1. Introduction

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, P)$ be a filtered probability space, where the filtration $(\mathcal{F}_t) = (\mathcal{F}^t)$ is generated by a càdlàg (right continuous and admitting left limits) process $Y = (Y_t, t \in [0, T])$. Note that, in general, $(\mathcal{F}_t)$ is not right-continuous. Let

$$\pi_n = \{0 = t^n_0 < t^n_1 < \cdots < t^n_{k_n} = T\}, \quad n \in \mathbb{N},$$

be a sequence of refining partitions of an interval $[0, T]$ such that $|\pi_n| := \max |t^n_s - t^n_{s-1}| \to 0, \quad n \to \infty$. Denote $\mathcal{F}^n = \sigma(Y^n_s, \ s \leq t)$, where

$$Y^n_t := Y^n_{t^n_s} \quad \text{for} \quad t \in [t^n_s, t^n_{s+1}), \quad Y^n_T := Y^n_{t^n_{k_n}}.$$

At last, assume that each fixed point of discontinuity for $Y$ belongs to $\cup_n \pi_n$.

Given an integrable random variable $X$ and a sequence of random variables $X^n, \ n \in \mathbb{N}$, converging to $X$ in $L^1(P)$, consider the martingale $M = (M_t = E(X|\mathcal{F}_t), \ t \in [0, T])$, and the sequence of martingales $M^n = (M^n_t = E(X^n|\mathcal{F}^n_t), t \in [0, T]), \ n \in \mathbb{N}$, with respect to the perturbed filtrations $(\mathcal{F}^n_t)_{t \in [0, T]}$, $n \in \mathbb{N}$. Since $\mathcal{F}^n_t \uparrow \mathcal{F}_t, \ n \to \infty$, for each $t \in [0, T]$, we have that $M^n_t \to M_t$ in probability. In this paper we consider the problem of convergence $M^n \to M^+$ in the Skorokhod topology, where $M^+ := M(\cdot + 0)$.

Corresponding author.

0304-4149/98/$19.00 \quad \copyright 1998$ Published by Elsevier Science B.V. All rights reserved

PII: S0304-4149(98)00013-1
This problem was suggested by the paper of Antonelli and Kohatsu-Higa (1997), where a general problem of convergence of solutions of backward SDEs was considered. Under suitable conditions, they proved that, given the backward SDE
\[ V_t = E \left( \int_t^T g_s(V_s) \, dA_s + X|\mathcal{F}_t \right), \]  
the solutions \( V^n \) of perturbed equations
\[ V^n_t = E \left( \int_t^T g^n_s(V^n_s) \, dA^n_s + X^n|\mathcal{F}_t \right) \]  
converge to \( V \), the solution of (\( *) \), in law for the Meyer–Zheng topology. The problem of convergence for the Skorokhod topology appeared to be significantly more complicated even in the “degenerate” case \( g = g^n = 0 \) and \( X^n = X \).

In this paper we show that:

1. In general, the convergence \( M^n \to M^+ \) (i.e., the convergence of solutions \( V^n \to V \) in the “degenerate case” mentioned above) for the Skorokhod topology can fail; see the example in Section 2.
2. If \( Y \) is a Markov process (not necessarily continuous), then \( M^n \to M^+ \) in probability for the Skorokhod topology, (Theorem 1, Section 2).
3. If the hypotheses of Antonelli and Kohatsu-Higa (1997) are satisfied and if either the martingale part of \( V \) is continuous, or \( A \) is continuous and \( Y \) is a Markov process, then \( V^n \to V \) in law for the Skorokhod topology. Under additional assumptions (essentially, convergence in variation of \( A^n \) to \( A \), and \( A \) is nonrandom), \( V^n \to V \) in probability for the Skorokhod topology; these results are given in Theorem 2 and Proposition 3 of Section 3.

We further denote \( C = C[0,T], \mathbb{D} = D[0,T], \mathbb{D}^k = D((0,T], \mathbb{R}^k) \) (equipped with the Skorokhod topology).

2. Convergence of perturbed martingales and the “Markov case”

Example. Let \( B = (B_t, t \in [0,T]) \) be a standard Brownian motion, and \( X \) be a random variable independent of \( B \) and such that \( P\{X = 1\} = P\{X = 1/2\} = 1/2 \). Choose \( Y \) defined by

\[
Y_t = \begin{cases} 
0 & \text{for } t < \frac{1}{2}, \\
(B_t - B_{1/2})X & \text{for } t \geq \frac{1}{2}.
\end{cases}
\]

The main idea of the example is the following. The knowledge of the trajectory of \( Y_t = (B_t - B_{1/2})X \) on a whole, though arbitrary small, time interval \( [\frac{1}{2}, \frac{1}{2} + \varepsilon] \), gives a.s. knowledge of the true value of \( X \), while the knowledge of \( Y_t \) for a fixed time moment \( t > \frac{1}{2} \) gives, in contrast, little information on \( X \).

To be precise, note first that \( X \) is \( \mathcal{F}_t \)-measurable for each \( t > \frac{1}{2} \). To show this take a sequence \( \tilde{s}_n = \{ \frac{1}{2} = s^n_0 < s^n_1 < \cdots < s^n_n = t \}, \quad n \in \mathbb{N} \), of partitions of an interval \( [\frac{1}{2}, t] \)
such that \( |\tilde{\eta}_n| \to 0 \). Then

\[
\sum_i (Y_{s^n_i} - Y_{s^n_{i-1}})^2 = \sum_i (B_{s^n_i} - B_{s^n_{i-1}})^2 X^2 \to \left( t - \frac{1}{2} \right) X^2, \quad n \to \infty,
\]

in probability. Since all the summands \((Y_{s^n_i} - Y_{s^n_{i-1}})^2\) are \(\mathcal{F}_t\)-measurable, we have the \((\mathcal{F}_t)\)-measurability of \((t - \frac{1}{2})X^2\) and hence of \(X\). In particular, this implies that \(Y\) is a (continuous) \((\mathcal{F}_t)\)-martingale (and an \((\mathcal{F}_{t+})\)-martingale as well). We also have that

\[
M_t = E(X|\mathcal{F}_t) = \begin{cases} \frac{3}{4} & \text{for } t \leq \frac{1}{2}, \\ X & \text{for } t > \frac{1}{2}. \end{cases}
\]

Now consider \(M^n_t = E(X|\mathcal{F}^n_t)\). Fix \(n \in \mathbb{N} \) and denote \(i_\ast = \min\{i : t^n_i > \frac{1}{2}\}, \quad t_\ast = t_{i_\ast}, \quad \epsilon = t_\ast - \frac{1}{2} > 0, \quad Z_\ast = (B_{t_\ast} - B_{1/2})\). Let us calculate \(E(X|\mathcal{F}^n_{t_\ast})\). The distribution function of \(Y_{t_\ast} = X\) is

\[
F_{Z,X}(y) = P\{Z,X < y\} = \frac{1}{2}(P\{Z < y|X = 1\} + P\{Z < 2y|X = 1/2\})
\]

\[
= \frac{1}{2}(F_Z(y) + F_Z(2y)), \quad y \in \mathbb{R}.
\]

Hence, the density function of \(Y_{t_\ast}\) equals

\[
\varphi(y) = \varphi_{Z,X}(y) = \frac{1}{2}\varphi_{\ast}(y) + \varphi_{\ast}(2y), \quad y \in \mathbb{R},
\]

where \(\varphi_{\ast}(y) = \varphi_{Z}(y) = (2\pi\epsilon)^{-1/2}e^{-y^2/(2\epsilon)}, \quad y \in \mathbb{R}\), is the density of \(N(0,\epsilon)\). Therefore,

\[
P\{X = 1|Z,X = y\} = \frac{\varphi_{Z}(y)P\{X = 1\}}{\varphi_{Z,X}(y)} = \frac{\varphi_{\ast}(y)}{\varphi_{\ast}(y) + 2\varphi_{\ast}(2y)}
\]

and

\[
P\{X = 1/2|Z,X = y\} = \frac{2\varphi_{\ast}(2y)}{\varphi_{\ast}(y) + 2\varphi_{\ast}(2y)}, \quad y \in \mathbb{R}.
\]

Hence,

\[
E(X|\mathcal{F}^n_{t_\ast}) = E(X|Z,X) = \frac{\varphi_{\ast}(Z,X) + \varphi_{\ast}(2Z,X)}{\varphi_{\ast}(Z,X) + 2\varphi_{\ast}(2Z,X)} = \frac{1 + \exp\{-3(Z_\ast X)^2/2\epsilon\}}{1 + 2\exp\{-3(Z_\ast X)^2/2\epsilon\}}.
\]

Since \(Z_\ast/\sqrt{\epsilon}\) is a standard normal variable independent of \(X\), we have that the value of the first “step” \(M^n_t = M^n_{t^n_i}, \quad t \in [t_\ast,t_{i_\ast}]\) (after the time \(t = \frac{1}{2}\)) of the martingale \(M^n\) has a continuous distribution on the interval \((\frac{3}{4},1)\) independent of \(\epsilon\) and hence of \(n\). Now passing to the limit as \(n \to \infty\) we have that \(t^n_\ast = t_\ast \downarrow t\), but \(A\mathcal{M}_{t^n_\ast}\) does not converge.
(in probability or in law) to $\Delta M_{1/2}$ that has only two possible values $\pm \frac{1}{2}$. This makes impossible the convergence $M^n \to M^+$ in the Skorokhod sense.

In the sequel, $Y, Y^n, (\mathcal{F}_t), (\mathcal{F}_n^T), X, X^n, V, V^n, A$ and $A^n$ have the same meaning as in Introduction.

**Theorem 1.** Suppose $Y = (Y_t, \mathcal{F}_t = \mathcal{F}_T^T, t \in [0, T])$ is a Markov process. Let $X$ and $X^n$, $n \in \mathbb{N}$, be $\mathcal{F}_T$-measurable integrable random variables. Denote $M = E(X|\mathcal{F}_T)$, $M^n = E(X^n|\mathcal{F}_n^T)$, $t \in [0, T]$. Suppose that either $X^n = X$ for all $n \in \mathbb{N}$, or $X^n \to X$ in $L^p$ for some $p > 1$. Then $M^n \to M^+$ in probability for the Skorokhod topology.

**Proof.** In the case $X^n = X$ using the truncation argument we can also assume, without loss of generality, that $X$ is $p$-integrable for some $p > 1$. Note that, for each $t = t^n_i \in \pi_n$, the random variable $M^n_t = E(M^n_T|\mathcal{F}_n^T)$ can be written as

$$M^n_t = f(Y_{t_0}, Y_{t_1}, \ldots, Y_t)$$

with some measurable function $f : \mathbb{R}^{i+1} \to \mathbb{R}$. By the Markov property of $Y$ we have, for $s = t_j \in \pi_n$, $s \leq t$,

$$E(M^n_s|\mathcal{F}_s) = E(M^n_t|\mathcal{F}_s)$$

with some measurable function $E : \mathbb{R}^{i+1} \to \mathbb{R}$, i.e., $E(M^n_s|\mathcal{F}_s)$ is, in fact, $\mathcal{F}_s^n$-measurable, and thus

$$E(M^n_s|\mathcal{F}_s) = M^n_s$$

for $s \leq t$, $s, t \in \pi_n$, $n \in \mathbb{N}$.

This means that all the processes

$$(M^n_t, \mathcal{F}_t, t \in \pi_n), \ n \in \mathbb{N}$$

are martingales with respect to the filtrations $(\mathcal{F}_t)_{t \in \pi_n}$. Therefore, we can apply the martingale inequality for the martingales $(M^n_t - M_t, \mathcal{F}_t, t \in \pi_n)$ to obtain

$$E \left( \sup_{t \in \pi_n} |M^n_t - M_t|^p \right) \leq C_p E|M^n_T - M_T|^p \to 0, \ n \to \infty,$$

where the convergence $M^n_T \to M_T$ in $L^p$ is easily seen from

$$M^n_T - M_T = E(X^n|\mathcal{F}_n^T) - E(X|\mathcal{F}_T)$$

$$= E(X^n - X|\mathcal{F}_n^T) + [E(X|\mathcal{F}_n^T) - E(X|\mathcal{F}_T)].$$

Indeed, the first term converges to zero in $L^p$ by Jensen’s inequality, while the second one tends to zero in $L^p$ by the martingale convergence theorem (remind that $\mathcal{F}_n^T \uparrow \mathcal{F}_T$, $n \to \infty$).

Now denote $\tilde{M}^n_t = M^n_t$ for $t \in [t^n_i, t^n_{i+1})$, $\tilde{M}^n_T = M^n_{t^n_{i-1}}$. Then $\tilde{M}^n \to M^+$ in the Skorokhod topology a.s. One can check this using, for example, Lemma 6.2 from Kurtz–Protter [5]. On the other hand, since both $M^n$ and $\tilde{M}^n$ are constant on the partition intervals
by the preceding we have
\[
E \left( \sup_{t \in [0,T]} |M^n_t - \tilde{M}^n_t|^p \right) = E \left( \sup_{t \in \pi_n} |M^n_t - \tilde{M}^n_t|^p \right) \\
= E \left( \sup_{t \in \pi_n} |M^n_t - M_t|^p \right) \to 0, \quad n \to \infty,
\]
i.e., \( M^n - \tilde{M}^n \to 0 \) uniformly in probability. Together with the a.s. convergence \( \tilde{M}^n \to M^+ \) in the Skorokhod sense, this implies the convergence \( M^n \to M^+ \) in the Skorokhod sense (in probability), and the theorem is proved. \( \square \)

**Remark 1.** In view of our counterexample and Theorem 1, the Markov assumption for the “filtration generating” process \( Y \) seems rather natural, though one can indicate some situations when this assumption is not needed. We formulate a partial result in this direction when \( X \) has a chaos type representation in terms of \( Y \).

**Proposition 1.** Let \( Y \) be a square integrable martingale such that \( d \langle Y, Y \rangle \leq da_t \) for some bounded nonrandom increasing càdlàg function \( a \) (i.e., \((a - \langle Y, Y \rangle) \) is an increasing process). Suppose \( X \) admits the following representation: there exist \( k \in \mathbb{N} \) and functions \( u_p \in C[0,T], p = 1, 2, \ldots, k \), such that \( X = I_k(Y)_T \), where
\[
I_1(Y)_t = \int_0^t u_1(s) \, dY_s, \quad I_p(Y)_t = \int_0^t u_p(s) I_{p-1}(Y)_{s-} \, dY_s, \quad p = 2, \ldots, k.
\]
Then \( M^n \to M^+ \) in probability (for the Skorokhod topology).

**Sketch of the proof.** By assumptions made it is clear that \( M^n_0 = I_k(Y)_0 \). Denote \( \tilde{M}_t^n = I_k(Y^n)_t \). Then \( \tilde{M} \) is an \((\mathcal{F}_n^t)\)-martingale, and by induction it is easy to check that
\[
E(\tilde{M}_T^n)^2 \leq \int_0^T \cdots \int_0^{t_{k-1}} \int_0^{t_k} u_k^2(t_{k-1}) \cdots u_1^2(s) \, da_{k-1} \cdots da_1.
\]
Now, the convergence theorem for stochastic integrals (Jakubowski et al., 1989; Kurtz and Protter, 1991) easily yields that
\[
(Y^n, I_1(Y^n), \ldots, I_k(Y^n)) \xrightarrow{p} (Y, I_1(Y), \ldots, I_k(Y))
\]
for the Skorokhod topology in \( \mathbb{D}^{k+1} \), and hence \( \tilde{M}^n \xrightarrow{p} M^+ \) in \( \mathbb{D} \). Now using the uniform boundedness of \((I_k(Y^n)_T)\) in \( L^2 \) we get the convergence \( E(\tilde{M}_T^n - M_T)^2 \to 0 \). On the other hand, since \( M_T^n \to M_T \) a.s. and \( (M_T^n) \) is bounded in \( L^2 \), we also have \( E(M_T^n - M_T)^2 \to 0 \). Therefore,
\[
E \left[ \sup_{t \leq T} |M^n_t - \tilde{M}_T^n|^2 \right] \leq 4E(\tilde{M}_T^n - M_T)^2 \to 0,
\]
whence the result. \( \square \)

**Remark 2.** It is not difficult to generalize this proposition to random variables \( X \) given by sums of \( I_k(Y)_T \) with different \( k \) and then to \( X \) which are the sums of iterated
integrals of the form
\[ \int_{0<h_1<h_2<\cdots<h_{k-1}<t} u(s,t_1,\ldots,t_{k-1}) \, dY_s \, \cdots \, dY_{t_{k-1}} \]
with \( u \in L^2(S, \mathcal{F}_t, \mathbb{P}) \) and \( S_k = \{(t_1, t_2, \ldots, t_k) : 0 < t_1 < t_2 < \cdots < t_k < T\} \).

Without Markov assumption, in a more general setting we can get convergence of perturbed martingales in a weaker sense:

**Proposition 2.** Let \((\mathcal{G}_t)\) and \((\mathcal{G}_t^n)\) be filtrations such that, for each \( n \in \mathbb{N} \), \((\mathcal{G}_t^n)\) is right continuous and, for all \( t \leq T \), \( \mathcal{G}_t^n \triangleright \mathcal{G}_t \), \( n \to \infty \). Suppose that \( Z \) and \( Z^n \), \( n \in \mathbb{N} \), are \( \mathcal{G}_T \)-measurable random variables with \( Z^n \rightarrow Z \) in \( L^p \) for some \( p > 1 \). Then the sequence of processes \( (E(Z^n | \mathcal{G}_t^n)) \) converges to \( E(Z | \mathcal{G}_t) \) in probability for the Meyer–Zheng topology in \( \mathbb{D} \).

**Proof.** It suffices (see Lemma 1 of Meyer and Zheng, 1984) to prove that
\[ E \left( \int_0^T |E(Z^n | \mathcal{G}_t^n) - E(Z | \mathcal{G}_t)| \, dt \right) \rightarrow 0. \]

But
\[ E \left( \int_0^T |E(Z^n | \mathcal{G}_t^n) - E(Z | \mathcal{G}_t)| \, dt \right) = \int_0^T U_t^n \, dt, \]
where
\[ U_t^n := E|E(Z^n | \mathcal{G}_t^n) - E(Z | \mathcal{G}_t)| \rightarrow 0, \quad t \in [0, T]. \]

Since, by assumptions made, \( U_t^n \leq \sup_n E(|Z^n| + |Z|) < \infty \), \( t \in [0, T] \), applying the dominated convergence theorem we have the result.

\[ \square \]

3. **Convergence of solutions of backward equations**

Let us begin with an example based on a typical one of two converging in \( \mathbb{D} \) sequences of càdlàg functions having the sum diverging in \( \mathbb{D} \). Suppose that a martingale \( M_t^n = E(X | \mathcal{F}_t^n) \) is not continuous, that is, with a positive probability, it has at least one jump on the time interval \([0, T]\). Let then two numbers \( a > 0 \) and \( \epsilon > 0 \) be such that, for the stopping time \( \tau := \inf\{t \geq 0 : |\Delta M_t^n| > a\} \), the probability \( P\{\tau < T - \epsilon\} \) is positive. Suppose also that \( M_t^n \) has no jump equal to \( a \). Assuming that \( M^n \overset{D}{\rightarrow} M^+ \) a.s., denote \( \tau_n = \inf\{t \geq 0 : |\Delta M_t^n| > a\} = \min\{t \in \pi_n : |\Delta M_t^n| > a\} \), which is an \((\mathcal{G}_t^n)\)-stopping time. Then \( \tau_n \rightarrow \tau \) a.s. and, on the event \( \{\tau < T - \epsilon\} \), \( M_n^\tau \overset{D}{\rightarrow} M^+ \). Take \( A = \mathbf{1}_{\{\tau < T - \epsilon\}} \) and \( A^n = \mathbf{1}_{\{\tau < T - \epsilon + \delta_n\}} \), where the number sequence \((\delta_n)\) is taken so that \( \delta_n \rightarrow |\pi_n| \) and \( \delta_n \downarrow 0 \). Then \( A^n \overset{D}{\rightarrow} A \) and, even though \( M^n \overset{D}{\rightarrow} M^+ \), the sequence \( (M^n - A^n) \) does not converge in the Skorokhod topology to \( M^+ - A \) at least on the event \( \{\tau < T - \epsilon\} \). This implies that even for the coefficients \( g^n = g \equiv 1 \) we have no
convergence of $V^n = 1 + M^n - A^n$ to $V = 1 + M^+ - A$ in the Skorokhod topology for the solutions of the corresponding backward equations.

This example shows that to get the convergence for the Skorokhod topology for solutions of backward equations under perturbations of filtrations it is necessary to assume some additional hypotheses on $Y$ or $X$.

Now, we write the considered backward equations in the form

$$V_t = E \left( \int_0^T g_t(V_s) \, dA_s + X_t \big| \mathcal{F}_t \right) - \int_0^t g_t(V_s) \, dA_s,$$

$$V^n_t = E \left( \int_0^T g^n_t(V^n_s) \, dA^n_s + X^n_t \big| \mathcal{F}_t^n \right) - \int_0^t g^n_t(V^n_s) \, dA^n_s.$$ 

We assume the following hypotheses that are essentially taken from (Antonelli and Kohatsu-Higa, 1997):

(HX) $\exists \delta > 0 : X^n \to X$ in $L^{1+\delta}(P)$;

(HA) All $A^n$ are increasing ($\mathcal{F}_t^n$)-adapted processes, and $A^n \to A$; $A^n_t \leq \beta_n$; $\sup_n \beta_n = \beta_\infty < \infty$;

(Hg) $g, g^n : \Omega \times [0, T] \times \mathbb{R} \to \mathbb{R}$ are, respectively ($\mathcal{F}_t$) and $\mathcal{F}_t^n$-adapted, Lipschitz with constants $c, c_n$; $g$ is continuous in $s, g^n$ are càdlàg in $s$; $\sup_s |g_t(o)|$ is bounded;

$$\sup_n e^{c_n \beta_n(1+\delta)} \left\{ E \left[ \int_0^T |g^n_t(0)| \, dA^n_s \right]^{1+\delta} + E |X^n|^{1+\delta} \right\} < \infty,$$

$P$-a.s. $g^n \to g$ uniformly in $s$ and $x$ in compact sets.

In view of the first example of Theorem 1, we will also need the following hypothesis.

(H$\mathcal{F}$) The process $Y$ is a continuous Markov process with respect to the filtration ($\mathcal{F}_t$) it generates.

**Lemma 1** (Antonelli and Kohatsu-Higa, Theorem 2.5). Under hypotheses (HX), (HA), and (Hg), we have

1. $\sup_n E\left( \int_0^T |g^n_t(V^n_s)| \, dA^n_s + |X^n_s| \right)^{1+\delta} < \infty$,
2. $\sup_n E\left( \sup_{t \leq T} |V^n_t|^{1+\delta} \right) < \infty$.

**Remark.** Lemma 1 directly gives the uniform boundedness in probability and thus the tightness of the sequence $(V^n)$ for the Meyer–Zheng topology (see Stricker, 1985).

**Lemma 2.** Under hypotheses (HX), (HA), and (Hg), the sequence

$$\left( \int_0^T g^n_t(V^n_s) \, dA^n_s, A^n \right)$$

is tight in $\mathbb{D}^2$. 
Proof. Step 1: We begin with showing that

\[ \forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, \quad P \left[ \sup_{t \leq T} \left| \int_0^t (g_s(V^n_s) - g^n_s(x, \omega)) \, dA^n_s \right| > \varepsilon \right] \leq \frac{\varepsilon}{\beta_{\infty}}. \]

Let now \( K \) be such that \( P\{\sup_{t \leq T} |V^n_t| > K\} < \varepsilon/2 \) (such \( K \) exists by Lemmas 1 and 2), and let \( n_0 \) be such that

\[ P \left[ \sup_{|x| \leq K} \sup_{t \leq T} |g_t(x, \omega) - g^n_t(x, \omega)| > \frac{\varepsilon}{2\beta_{\infty}} \right] \leq \frac{\varepsilon}{2}, \quad n \geq n_0. \]

Then by the preceding we have

\[ P \left[ \sup_{t \leq T} \left| \int_0^t (g_s(V^n_s) - g^n_s(V^n_s)) \, dA^n_s \right| > \varepsilon \right] \]

\[ \leq P \left[ \sup_{t \leq T} \left| \int_0^t (g_s(V^n_s) - g^n_s(V^n_s)) \, dA^n_s \right| > \varepsilon, \sup_{t \leq T} |V^n_t| \leq K \right] + P[\sup_{t \leq T} |V^n_t| > K] \]

\[ \leq P \left[ \sup_{|x| \leq K} \left( \sup_{t \leq T} |g_t(x, \omega) - g^n_t(x, \omega)| \right) > \varepsilon \right] + \frac{\varepsilon}{2} \leq \frac{\varepsilon}{2}, \quad n \geq n_0. \]

Step 2: Now to obtain the result of Lemma it suffices, by (HA), to prove the tightness in \( \mathbb{D} \) of the sequence of the processes \( \left( \int_0^t g_s(V^n_s) \, dA^n_s \right) \), \( P \)-a.s. Denote \( g^n_s(V^n_s) = \max \{g_s(V^n_s), 0\} \) and \( g^n_s(\cdot) = \max \{-g_s(\cdot), 0\} \). Since

\[ \Delta \left( \int_0^t g^n_+(V^n_s) \, dA^n_s \right) = g^n_+(V^n_t) \Delta A^n_t \]

and

\[ \Delta \left( \int_0^t g^n_-(V^n_s) \, dA^n_s \right) = \Delta \left( \int_0^t g^n_+(V^n_s) \, dA^n_s \right) + \Delta \left( \int_0^t g^n_-(V^n_s) \, dA^n_s \right), \]

we have that the sequence \( \left( \int_0^t g^n_+(V^n_s) \, dA^n_s \right) \) is tight in \( \mathbb{D} \) provided both sequences \( \left( \int_0^t g^n_+(V^n_s) \, dA^n_s \right) \) are tight. Since

\[ \left( \int_0^t |g_s(V^n_s)| \, dA^n_s \right) \leq \int_0^t (|g_s(0)| + c|V^n_t|) \, dA^n_s \]

and the processes \( \int_0^t (|g_s(0)| + c|V^n_t| - g^n_+(V^n_s)) \, dA^n_s \) are increasing, we have that (see Jacod–Shiryaev, 1987, Ch. 6, Proposition 3.35) the sequence \( \left( \int_0^t g_s(V^n_s) \, dA^n_s \right) \) is tight provided the sequence \( \left( \int_0^t (|g_s(0)| + c|V^n_t|) \, dA^n_s \right) \) is tight. (Note that the increasing processes are well defined by hypothesis (Hg.).)

Step 3: Let \( K \) be such that \( P[\sup_{t \leq T} |V^n_t| > K] < \varepsilon/2 \). Then we have

\[ P \left[ \sup_{t \leq T} c \int_0^t |V^n_s| \mathbb{1}_{\{|V^n_t| > K\}} \, dA^n_s \right] > \varepsilon \]

\[ \leq P \left[ \sup_{t \leq T} |V^n_t| > K \right] < \frac{\varepsilon}{2}, \]

whence it suffices to show the tightness in \( \mathbb{D} \) of the sequence

\[ \left( \int_0^t (|g_s(0)| + c|V^n_t| \mathbb{1}_{\{|V^n_s| < K\}}) \, dA^n_s \right) \]
and hence of the sequence

\[ \left( \int_0^1 |g_s(0)| + cK \, dA_s^u \right). \]

Since \( A^u \overset{D}{\longrightarrow} A \), \(|A^u|\) are uniformly bounded, and \( g \) is continuous in \( s \), we have

\[ \int_0^1 |g_s(0)| + cK \, dA_s^u \overset{D}{\longrightarrow} \int_0^1 |g_s(0)| + cK \, dA_s, \]

and lemma is proved. \( \square \)

**Lemma 3.** Under hypotheses (HX), (HA), (Hg), and (H\( F \)), the sequence of processes

\[ \left( E \left[ \int_0^T g_\nu(V_s^u) \, dA_s + X^u \mid \mathcal{F}_n^u \right] \right) \]

is tight in \( \mathbb{D} \).

**Proof.** Since, by (H\( F \)), \( Y \) is a continuous process, we have by Lemma 2 that the sequence \( (X^u, Y^u, A^u, \int_0^1 g_s^u(V_s^u) \, dA_s^u) \) is tight in \( \mathbb{R} \times \mathbb{C} \times \mathbb{D}^2 \). By the Skorokhod representation theorem there exists a subsequence of \( \mathbb{N} \) (indexed by \( n' \)) and a probability space \( (\Omega, \mathcal{F}, \tilde{P}) \) with a sequence \( (\tilde{X}^u, \tilde{Y}^u, \tilde{A}^u, \tilde{\mathcal{F}}^u) \) defined on it such that

\[ \mathcal{L}(((\tilde{X}^u, \tilde{Y}^u, \tilde{A}^u, \tilde{\mathcal{F}}^u)) \overset{\tilde{P}}{\longrightarrow} (X, Y, A, \mathcal{F})) \quad \text{in} \quad \mathbb{R} \times \mathbb{C} \times \mathbb{D}^2. \]

Clearly,

\[ \mathcal{L}(((\tilde{X}, \tilde{Y}, \tilde{A})) \overset{\tilde{P}}{\longrightarrow} (X, Y, A) \quad \text{in} \quad \mathbb{R} \times \mathbb{C} \times \mathbb{D}^2. \]

Let \( \tilde{Y}^u \) be defined by

\[ \tilde{Y}^u = \sum \tilde{Y}_i \mathbb{1}_{[t_i, t_{i+1})}(t) \quad \text{when} \quad Y^u(t) = \sum \tilde{Y}_i \mathbb{1}_{[t_i, t_{i+1})}(t). \]

Denote \( \tilde{\mathcal{F}}^u = \mathcal{F}_t^{\tilde{Y}^u} \) and \( \tilde{\mathcal{F}} = \mathcal{F}_t^{\tilde{Y}}. \) Since

\[ \sup_n \tilde{E} \left\{ |\tilde{E}(\tilde{U}_T^u + \tilde{X}^u) |^{1+\delta} \right\} \leq \sup_n \tilde{E} \left\{ |\tilde{U}_T^u + \tilde{X}^u|^{1+\delta} \right\} \]

\[ = \sup_n \tilde{E} \left\{ \left| \int_0^T g_s^u(V_s^u) \, dA_s^u + X^u \right|^{1+\delta} \right\}, \]

we have that

\[ \tilde{E}(\tilde{U}_T^u + \tilde{X}^u \mid \tilde{\mathcal{F}}^u) \overset{L^{1+\delta}}{\longrightarrow} \tilde{E}(\tilde{U}_T^u + \tilde{X} \mid \tilde{\mathcal{F}}). \]
Therefore, applying Theorem 1 to the sequence of random variables (\( \hat{U}_t^n + \hat{X}_t^n \)) and filtrations (\( \mathcal{F}_t^n \), \( \mathcal{F}_t \)), we obtain the convergence in \( P \)-probability for the Skorokhod topology of (\( \hat{E}(\hat{U}_t^n + \hat{X}_t^n | \mathcal{F}_t^n) \)) to \( \hat{E}(\hat{U}_t + \hat{X} | \mathcal{F}_t) \). But

\[
\mathcal{L}(\hat{E}(\hat{U}_t^n + \hat{X}_t^n | \mathcal{F}_t^n) | P) = \mathcal{L}
\left( E \left( \int_0^T g^n_s(V^n_s) \, dA^n_s + X^n_s | \mathcal{F}_t^n \right) \right) | P,
\]

whence the assertion. \( \square \)

As the example at the beginning of this section shows, one cannot expect the relative compactness or tightness in \( D \) of the sequence of the solutions \( (V^n) \) without an additional hypothesis. On the other hand, under an adequate continuity hypothesis, we immediately have:

**Lemma 4.** Suppose hypotheses (HX), (HA), (Hg), and (H\( \mathcal{F} \)) are satisfied. In addition, assume the following hypothesis:

(Hco) Either \( A \) is a continuous process, or all \( (\mathcal{F}_t) \)-martingales are continuous. Then the sequence

\[
(X^n, A^n, V^n, E \left( \int_0^T g^n_s(V^n_s) \, dA^n_s + X^n_s | \mathcal{F}_t^n \right), \int_0^T g^n_s(V^n_s) \, dA^n_s, Y)
\]

is tight in \( \mathbb{R} \times D^4 \times C \).

**Proof.** From Lemmas 2 and 3, with (Hco) we immediately get the tightness in \( \mathbb{R} \times D^4 \times C \) of the sequence

\[
(X^n, A^n, V^n, E \left( \int_0^T g^n_s(V^n_s) \, dA^n_s + X^n_s | \mathcal{F}_t^n \right), \int_0^T g^n_s(V^n_s) \, dA^n_s, Y).
\]

As in the proof of Lemma 2, using the property 3) of Lemma 1 and the uniform convergence of \( g^n \to g \) one easily gets that

\[
E \left[ \int_0^T |g^n_s(V^n_s) - g_s(V^n_s)| \, dA^n_s | \mathcal{F}_t^n \right] \to 0 \quad \text{and} \quad \int_0^T |g^n_s(V^n_s) - g_s(V^n_s)| \, dA^n_s \to 0
\]

uniformly in probability. Therefore, in the latter sequence one can replace \( g^n_s(V^n_s) \) by \( g_s(V^n_s) \). \( \square \)

**Theorem 2.** Suppose hypotheses (HX), (HA), (Hg), (H\( \mathcal{F} \)), and (Hco) are satisfied, and \( Y \) is continuous. Then the sequence \( (V^n) \) of the solutions of equations \( (\ast^n) \) converge to the solution \( V \) of \( (\ast) \) in law for the Skorokhod topology.

**Proof.** As in the proof of Lemma 2, one can notice that, for all \( \varepsilon > 0 \),

\[
P \left[ \sup_{t \leq T} \left| V^n_t - E \left( \int_0^T g_s(V^n_s) \, dA^n_s + X^n_s | \mathcal{F}_t^n \right) - \int_0^T g_s(V^n_s) \, dA^n_s \right| > \varepsilon \right] \to 0, \quad n \to \infty.
\]

Hence, it suffices to show that

\[
E \left( \int_0^T g_s(V^n_s) \, dA^n_s + X^n_s | \mathcal{F}_t^n \right) - \int_0^T g_s(V^n_s) \, dA^n_s \overset{\mathcal{L}(D)}{\to} V.
\]
Now, let us remark that since the function \( g \) is \( \mathcal{B}_{[0, T]} \otimes \mathcal{F}_t \otimes \mathcal{B}_\mathbb{R} \)-measurable and \( \mathcal{F}_t = \mathcal{F}^Y_t \), there exists a measurable mapping \( \varphi : [0, T] \times C \times \mathbb{R} \) such that \( g(s, \omega, x) = \varphi(s, Y(\omega), x) \). Clearly, \( \varphi \) and \( g \) have the same properties in \( s \) and \( x \). Therefore, the sequence

\[
(X^n, A^n, V^n, M^n, U^n) = \left( X^n, A^n, V^n, E \left[ \int_0^T \varphi(s, Y, V^n_s) \, dA^n_s + X^n \big| \mathcal{F}^n \right], \int_0^T \varphi(s, Y, V^n_s) \, dA^n_s \right)
\]

is tight. As before, on another probability space \((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})\) take a subsequence

\[
(X''_n, A''_n, V''_n, M''_n, U''_n, \tilde{Y})
\]

with

\[
(\tilde{X}''_n, \tilde{A}''_n, \tilde{V}''_n, \tilde{Y}) \xrightarrow{P-a.s.}\ (\tilde{X}, \tilde{A}, \tilde{V}, \tilde{Y}) \quad \text{in} \quad \mathbb{R} \times \mathbb{D}^2 \times C
\]

and

\[
\mathcal{L}((\tilde{X}''_n, \tilde{A}''_n, \tilde{V}''_n, \tilde{M}''_n, \tilde{U}''_n, \tilde{Y})|\tilde{P}) = \mathcal{L}((X''_n, A''_n, V''_n, M''_n, U''_n, Y)|P),
\]

where \( \tilde{U}''_n \) and \( \tilde{M}''_n \) have the representations

\[
\tilde{U}'_n = \int_0^t \varphi(s, \tilde{Y}, \tilde{V}_s''_n) \, d\tilde{A}'_s'' \quad \text{and} \quad \tilde{M}'_n = \tilde{E} \left[ \int_0^T \varphi(s, \tilde{Y}, \tilde{V}_s''_n) \, d\tilde{A}'_s'' \big| \mathcal{F}'_t \right].
\]

Since

\[
\varphi(\cdot, \tilde{Y}, \tilde{V}'_n) \xrightarrow{P-a.s.} \varphi(\cdot, \tilde{Y}, \tilde{V}) \quad \text{in} \quad \mathbb{D},
\]

by the statements (1) and (2) of Lemma 1 we have the convergence

\[
\int_0^t \varphi(s, \tilde{Y}, \tilde{V}'_n) \, d\tilde{A}'_s'' \rightarrow \int_0^t \varphi(s, \tilde{Y}, \tilde{V}) \, d\tilde{A}_s \quad \text{in} \quad \mathbb{D}
\]

(see Jakubowski et al., 1989, Proposition 2.9, (c)). In particular, for \( 0 < \delta' < \delta \), we have

\[
\int_0^T \varphi(s, \tilde{Y}, \tilde{V}'_n) \, d\tilde{A}'_s'' \rightarrow \int_0^T \varphi(s, \tilde{Y}, \tilde{V}) \, d\tilde{A}_s \quad \text{in} \quad L^{1+\delta'}(\tilde{P})
\]

and thus \( \tilde{M}'_n \xrightarrow{D} \tilde{M} \) in probability, where \( \tilde{M}_t = \tilde{E} \left[ \int_0^T \varphi(s, \tilde{Y}, \tilde{V}_s) \, d\tilde{A}_s + \tilde{X}_s \big| \mathcal{F}_{\tau^+} \right] \). Thus, we get that \( \tilde{V} \) is a solution of the equation

\[
\tilde{V}_t = \tilde{E} \left( \int_0^T \varphi(s, \tilde{Y}, \tilde{V}_s) \, d\tilde{A}_s + \tilde{X}_s \big| \mathcal{F}_{\tau^+} \right) - \int_0^t \varphi(s, \tilde{Y}, \tilde{V}_s) \, d\tilde{A}_s,
\]

which has the same law as the unique solution of

\[
\tilde{V}_t = \tilde{E} \left( \int_0^T \varphi(s, Y, \tilde{V}_s) \, d\tilde{A}_s + \tilde{X}_s \big| \mathcal{F}_{\tau^+} \right) - \int_0^t \varphi(s, Y, \tilde{V}_s) \, dA_s,
\]

which, in turn, coincides with the solution of \((*)\). The theorem is proved.

\[\square\]

**Remark 3.** It remains an open problem, whether, actually, the convergence \( V^n \rightarrow V \) in probability for the Skorokhod topology holds. The main difficulty is to prove the
pointwise convergence $V^n_t \to V_t$ at least for $t$ from a countable dense subset of $[0, T]$. One might hope to apply, in this context, Gronwall-type lemmas (see, e.g., (Antonelli, 1996; Antonelli and Kohatsu-Higa, 1997)). This way gives a result in this direction under additional strong assumptions:

**Proposition 3.** Let, as before, the processes $A^n$, $g^n$ be $(\mathcal{F}^n_t)$-adapted, and $A$, $g$ be $(\mathcal{F}_t)$-adapted, $X^n$ be $\mathcal{F}^n_T$-measurable, and $X$ be $\mathcal{F}_T$-measurable. Moreover, assume the following hypotheses:

(HF) The filtration $(\mathcal{F}_t)$ is right continuous;
(HVA) $E[(\text{Var}(A^n - A)_T)^2] \to 0$, and the process $A$ is deterministic. (Var($) denotes the variation process of ($$));
(Hg) $|g^n(\omega, x) - g(\omega, x)| \leq k_n(1 + |x|)$ with non random $k_n$ such that $k_n \to 0$;
(HX) $\sup_x |g(0)|$ is in $L^2$; $g, g^n$ are Lipschitz with constants $c, c_n$; $\sup_n c_n < \infty$;
(HX2) $X^n \to X$ in $L^2$.

Then under (HF), (HVA), (Hg) with $\delta = 1$, (Hg2), and (HX2), we have

(1) For every $t \in [0, T]$, $V^n_t \to V_t$ in $L^1$.

(2) The sequence of processes $(V_n)$ converges to $V$ in probability for the Meyer–Zheng topology on $D$.

(3) If the process $Y$ is Markov, the convergence in probability holds for the Skorokhod topology.

**Proof.** (1) For every $t \in [0, T]$, we have

$$|V^n_t - V_t| \leq E \left( \int_0^T |g^n_s(V^n_n)| \text{dVar}(A^n - A)_s + \int_0^T |g^n_s(V^n_n) - g_s(V^n_n)| \text{d}A_s ight)$$

$$+ |X^n - X| + \int_0^T c|V^n - V_n| \text{d}A_s \left( \mathcal{F}^n_t \right)$$

$$+ |E(U_t|\mathcal{F}_t^n) - E(U_t|\mathcal{F}_t)|,$$

where $U_t := \int_0^T g_s(V^n_t) \text{d}A_s + X$.

Since the increasing process $A$ is deterministic and hence $(\mathcal{F}_t)$-adapted for all $n \in \mathbb{N}$, we can apply the Gronwall-type lemma (see Antonelli and Kohatsu-Higa, 1997, Theorem 2.1, Antonelli, 1996, Theorem 1.8) to obtain:

$$|V^n_t - V_t| \leq E \left( \int_0^T \delta(cA)_s \delta(cA)_s^{-1} |g^n_s(V^n_n) - g_s(V^n_n)| \text{d}A_s ight)$$

$$+ \int_0^T \delta(cA)_s \delta(cA)_s^{-1} c|E(U_t|\mathcal{F}^n_t) - E(U_t|\mathcal{F}_t)| \text{d}A_s$$

$$+ \int_0^T \delta(cA)_s \delta(cA)_s^{-1} |g^n_s(V^n_n)| \text{dVar}(A^n - A)_s$$

$$+ \delta(cA)_s \delta(cA)_s^{-1} |X^n - X| \left( \mathcal{F}^n_t \right)$$

$$+ |E(U_t|\mathcal{F}^n_t) - E(U_t|\mathcal{F}_t)|,$$
where $\delta(cA)$ denotes the Doléans exponential of $cA$. Then taking expectations yields

\[
E[V^n_t - V_t] \leq \int_0^T \delta(cA)_s \delta(cA)_s^{-1} E[g^n_s(V^n_s) - g_s(V^n_s)] \, dA_s + \int_t^T \delta(cA)_s \delta(cA)_s^{-1} E[E(U_s|\mathcal{F}_s^n) - E(U_s|\mathcal{F}_s)] \, dA_s
\]

\[
+ \delta(cA)_T \delta(cA)_T^{-1} E \left( \sup_s |g^n_s(V^n_s)| \text{Var}(A^n - A)_T \right)
\]

\[
+ \delta(cA)_T \delta(cA)_T^{-1} E|X^n - X| + E[E(U_t|\mathcal{F}_t^n) - E(U_t|\mathcal{F}_t)].
\]

Denote by (1)–(5) the five terms of the right side of the last inequality.

Since by the hypothesises $\sup_s E[(\sup_{s \leq T} |V^n_s|)^2] < \infty$, we easily get that (1), (3), and (4) tend to 0.

Since $U_t \in L^2$ and $\mathcal{F}_t^n \uparrow \mathcal{F}_t$, we have the convergence $E(U_t|\mathcal{F}_t^n) \to E(U_t|\mathcal{F}_t)$ in $L^2$ and hence the convergence of (5) to 0.

Since

\[
E[E(U_t|\mathcal{F}_t^n) - E(U_t|\mathcal{F}_t)] \leq 2E|U_0| < \infty, \quad s \in [0, T],
\]

term (2) also tends to 0 by the dominated convergence theorem, and we are finished with the first part of Proposition.

(2) The assertion is obtained easily, since by the dominated convergence theorem we get

\[
\int_0^T E[V^n_t - V_t] \, ds \to 0.
\]

(3) From the first part of the proposition we have the convergence $V^n_t \xrightarrow{p} V_t$ for all $t \in [0, T]$. By the dominated convergence theorem from the convergence of $A^n \to A$ in variation one obtains the uniform convergence in probability of the integrals

\[
\int_0^t g^n_s(V^n_s) \, dA^n_s \to \int_0^t g_s(V_s) \, dA_s.
\]

As in the proof of Theorem 1, we have the convergence of the martingale parts $(M^n_t, \mathcal{F}_t^n)$ of $V^n$ to that of $V$ in the following sense:

\[
\sup_{t \in \pi_n} |M^n_t - M_t| \xrightarrow{p} 0,
\]

and hence

\[
\sup_{t \in \pi_n} |V^n_t - V_t| \leq \sup_{t \in \pi_n} |M^n_t - M_t| + \sup_{t \leq T} \left| \int_0^t g^n_s(V^n_s) \, dA^n_s - \int_0^t g_s(V_s) \, dA_s \right| \xrightarrow{p} 0.
\]

Finishing as in the proof of Theorem 1, we obtain the convergence $V^n \to V$ in probability for the Skorokhod topology.

**Remark 4.** If $A$ is continuous, the right continuity of the filtration $(\mathcal{F}_t)$ is not needed for getting assertions (2) and (3) of Proposition 3, while assertion (1) should be changed...
by the following weaker one: 

(1') For all $t \in [0,T]$, $V^n_t - V_t - E(U_t|\mathcal{F}_t^+) + E(U_t|\mathcal{F}_t) \to 0$ in $L^1$.

Actually, in conditioning, $\mathcal{F}_t$ is changed by $\mathcal{F}_t^+$; after using Gronwall-type inequality, considering the process $(U_t)$ we can choose a càdlàg version of $E(U_t|\mathcal{F}_t)$ such that $E(U_t|\mathcal{F}_t) = E(U_t|\mathcal{F}_t^+)$, for all $t$, except $t$ from (at most) a countable set $D$. Then, for $t \in D^c$, $E(U_t|\mathcal{F}_t^+) \to E(U_t|\mathcal{F}_t^+)$. Since $A$ is continuous, by the dominated convergence theorem we still get the convergence of Stieltjes integrals with respect to $A$, and we are finished with (1').

The proofs of (2) and (3) need minimal changes as it suffices to notice that the Stieltjes integrals of $|E(U_t|\mathcal{F}_t^+) - E(U_t|\mathcal{F}_t)|$ with respect to $dA_t$ and $dt$ are zeros.

**Remark 5.** In assertions (1) and (2) of Proposition 3, for the filtrations we only need the hypotheses of Proposition 2; the process $Y$, actually, does not play any role.

For simplicity of the proof, the assumption of convergence $g^n \to g$ is expressed by technical hypothesis (Hg2) taken from Antonelli and Kohatsu-Higa (1997), which, clearly, is not the weakest possible one.

**References**


