Pathwise Construction of Stochastic Integrals

Marcel Nutz[*]

First version: August 14, 2011. This version: June 12, 2012.

Abstract

We propose a method to construct the stochastic integral simultaneously under a non-dominated family of probability measures. Path-by-path, and without referring to a probability measure, we construct a sequence of Lebesgue-Stieltjes integrals whose medial limit coincides with the usual stochastic integral under essentially any probability measure such that the integrator is a semimartingale. This method applies to any predictable integrand.

Keywords Pathwise stochastic integral, aggregation, non-dominated model, second order BSDE, G-expectation, medial limit

AMS 2000 Subject Classification 60H05

1 Introduction

The goal of this article is to construct the stochastic integral in a setting where a large family $\mathcal{P}$ of probability measures is considered simultaneously. More precisely, given a predictable integrand H and a process X which is a semimartingale under all $P \in \mathcal{P}$, we wish to construct a process which P-a.s. coincides with the P-Itô integral $^{(P)}\!\!\int H \, dX$ for all $P \in \mathcal{P}$; i.e., we seek to *aggregate* the family $\{^{(P)}\!\!\int H \, dX\}_{P \in \mathcal{P}}$ into a single process. This work is motivated by recent developments in probability theory and stochastic optimal control, where stochastic integrals under families of measures have arisen in the context of Denis and Martini's model of volatility uncertainty in financial markets [2], Peng's G-expectation [11] and in particular the representation of G-martingales of Soner et al. [15], the second order backward stochastic differential equations and target problems of Soner et al. [13, 14] and in the non-dominated optional decomposition of Nutz and Soner [10]. In all these examples, one considers a family $\mathcal{P}$ of (mutually singular) measures which cannot be dominated by a finite measure.

The key problem in stochastic integration is, of course, that the paths of the integrator X are not of finite variation. The classical Hilbert space construction depends strongly on the probability measure since it exploits the martingale properties of the integrator; in particular, it is far from being pathwise. A natural extension of the classical construction to a family $\mathcal{P}$, carried out in [2, 11, 8] under specific regularity assumptions, is to consider the upper expectation $\mathcal{E}_{\mathcal{P}}[\,\cdot\,] = \sup_{P \in \mathcal{P}} E^P[\,\cdot\,]$ instead of the usual expectation and define the stochastic integral along the lines of the usual closure operation from simple integrands, but under a norm induced by $\mathcal{E}_{\mathcal{P}}$. Since such a norm is rather strong, this leads to a space of integrands which is smaller than in the classical case.

A quite different approach is to construct the stochastic integral pathwise and without direct reference to a probability measure; in this case, the integral will be well defined under all $P \in \mathcal{P}$. The strongest previous result about pathwise integration is due to Bichteler [1]; it includes in particular the integrals which can be defined by Föllmer's approach [3]. A very readable account of that result and some applications also appear in Karandikar [6]. Bichteler's remarkable observation is that if $H = G_{-}$ is the left limit of a càdlàg process G, then one can obtain a very favorable discretization of $\int H\,dX$ by sampling H at a specific sequence of stopping times, given by the level-crossing times of H at a grid of mesh size 2^{-n}. Namely, the corresponding Riemann sums $\int H^n\,dX$ converge *pointwise* in ω (uniformly in time, outside a set which is negligible for all P), and this limit yields a pathwise definition of the integral. To appreciate this fact, recall that as soon as H is left-continuous, essentially any discretization will converge to the stochastic integral, but only *in measure*, so that it is not immediate to construct a single limiting process for all P: passing to a subsequence yields P-a.s. convergence for some P, but not for all P at once.

Finally, let us mention that the "skeleton approach" of Willinger and Taqqu [16, 17] is yet another pathwise stochastic integration theory (related to the problem of martingale representation); however, this construction cannot be used in our context since it depends strongly on the equivalence class of the probability measure. The main drawback of the previous results is that the restrictions on the admissible integrands can be too strong for applications. In particular, the integrals appearing in the main results of [13] and [10] could not be aggregated for that reason. In a specific setting where X is continuous, one possible solution, proposed by Soner et al. [12], is to impose a strong separability assumption on the set $\mathcal{P}$, which allows to glue together directly the processes $\{^{(P)}\!\int H\,dX\}_{P \in \mathcal{P}}$.

In the present paper, we propose a surprisingly simple, pathwise construction of the stochastic integral for arbitrary predictable integrands H (Theorem 2.2) and a very general set $\mathcal{P}$. In a first step, we average H in the time variable to obtain approximations H^n of finite variation, which allows to define the integral $\int H^n\,dX$ pathwise. This averaging requires a certain

2

domination assumption; however, by imposing the latter at the level of predictable characteristics, we achieve a condition which is satisfied in all cases of practical interest (Assumption 2.1). The second step is the passage to the limit, where we shall work with the (projective limit of the) convergence in measure and use a beautiful construction due to G. Mokobodzki, known as medial limit (cf. Meyer [9] and Section 2.2 below). As a Banach limit, this is not a limit in a proper sense, but it will allow us to define path-by-path a measurable process $\int H\,dX$ which coincides with the usual stochastic integral under every P; in fact, it seems that this technology may be useful in other aggregation problems as well. To be precise, this requires a suitable choice of the model of set theory: we shall work under the Zermelo–Fraenkel set theory with axiom of choice (ZFC) plus the Continuum Hypothesis.

2 Main Result

Let $(\Omega, \mathcal{F})$ be a measurable space equipped with a right-continuous filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in [0,1]}$ and let $\mathcal{P}$ be a family of probability measures on $(\Omega, \mathcal{F})$. In the sequel, we shall work in the $\mathcal{P}$-universally augmented filtration

$$\mathbb{F}^* = (\mathcal{F}_t^*)_{t \in [0,1]}, \quad \mathcal{F}_t^* := \bigcap_{P \in \mathcal{P}} \mathcal{F}_t \vee \mathcal{N}^P,$$

where $\mathcal{N}^P$ is the collection of $(\mathcal{F}, P)$-nullsets (but see also Remark 2.6(ii)). Moreover, let X be an adapted process with càdlàg paths such that X is a semimartingale under each $P \in \mathcal{P}$, and let H be a predictable process which is X-integrable under each $P \in \mathcal{P}$.

Since we shall average H in time, it is necessary to fix a measure on $[0,1]$, at least path-by-path. We shall work under the following condition; see Jacod and Shiryaev [5, Section II] for the notion of predictable characteristics.

Assumption 2.1. There exists a predictable càdlàg increasing process A such that

$$\mathrm{Var}(B^P) + \langle X^c \rangle^P + (x^2 \wedge 1) * \nu^P \ll A \quad P\text{-a.s.} \quad \text{for all} \quad P \in \mathcal{P};$$

where $(B^P, \langle X^c \rangle^P, \nu^P)$ is the triplet of predictable characteristics of X under P and $\mathrm{Var}(B^P)$ denotes the total variation of B^P.

With $^{(P)}\!\int H\,dX$ denoting the Itô integral under P, our main result can be stated as follows.

Theorem 2.2. *Under Assumption 2.1, there exists an $\mathbb{F}^*$-adapted càdlàg process, denoted by $\int H\,dX$, such that*

$$\int H\,dX = \,^{(P)}\!\!\int H\,dX \quad P\text{-a.s.} \quad \text{for all} \quad P \in \mathcal{P}.$$

Moreover, the construction of any path $(\int H\,dX)(\omega)$ involves only the paths $H(\omega)$ and $X(\omega)$.

3

Assumption 2.1 is quite weak and should not be confused with a domination property for $\mathcal{P}$ or the paths of X. In fact, most semimartingales of practical interest have characteristics absolutely continuous with respect to $A_t = t$ (diffusion processes, solutions of Lévy driven stochastic differential equations, etc.). The following example covers the applications from the introduction.

Example 2.3. Let X be a continuous local martingale under each $P \in \mathcal{P}$, then Assumption 2.1 is satisfied. Indeed, let

$$A := X^2 - X_0^2 - 2 \int X \, dX;$$

here the stochastic integral can be defined pathwise by Bichteler's construction [1, Theorem 7.14]. Then A is a continuous process and, by Itô's formula, $A = [X] - X_0^2 = \langle X^c \rangle^P$ P-a.s. for every $P \in \mathcal{P}$. Therefore, Assumption 2.1 is satisfied with equality.

The previous example should illustrate that Assumption 2.1 is much weaker than it may seem at first glance. For instance, let X be the canonical process on $\Omega = C([0, 1]; \mathbb{R})$ and let Λ be the set of *all* increasing continuous functions $f : [0, 1] \to \mathbb{R}_+$ with $f(0) = 0$. Using time-changed Brownian motions, construct a situation where for any $f \in \Lambda$ there exists $P \in \mathcal{P}$ under which f is the quadratic variation of X, P-a.s. Then we observe that, by the above, Assumption 2.1 is satisfied—even though it is clearly impossible to dominate the set Λ. The crucial point here is the flexibility to assign different values to A on the various supports of the measures P.

2.1 Approximating Sequence

Since our aim is to prove Theorem 2.2, we may assume without loss of generality that $X_0 = 0$. Moreover, we may assume that the jumps of X are bounded by one in magnitude,

$$|\Delta X| \leq 1.$$

Indeed, the process $\check{X} := \sum_{s \leq \cdot} \mathbf{1}_{\{|\Delta X_s| > 1\}} \Delta X_s$ is of finite variation and $\int H \, d\check{X}$ is easily defined since it is simply a sum. Decomposing

$$X = (X - \check{X}) + \check{X},$$

it suffices to construct the integral $\int H \, d(X - \check{X})$ whose integrator has jumps bounded by one; moreover, $X - \check{X}$ satisfies Assumption 2.1 if X does. (Of course, we cannot reduce further to the martingale case, since the semimartingale decomposition of X depends on P!)

In this section, we construct an approximating sequence of integrands H^n such that the integrals $\int H^n \, dX$ can be defined pathwise and tend to the

integral of H in measure. To this end, we shall assume that H is uniformly bounded by a constant,

$$|H| \leq c.$$

In fact, we can easily remove this condition later on, since

$$^{(P)}\!\!\int H \mathbf{1}_{\{|H| \leq n\}} \, dX \to \, ^{(P)}\!\!\int H \, dX \quad \text{in} \quad ucp(P) \quad \text{for all} \quad P \in \mathcal{P} \qquad (2.1)$$

by the definition of the usual stochastic integral. Here $ucp(P)$ stands for convergence in probability P, uniformly (on compacts) in time. We recall the process A from Assumption 2.1. We may assume that $A_t - A_s \geq t - s$ for all $0 \leq s \leq t \leq 1$ by replacing A_t with $A_t + t$ if necessary; moreover, to avoid complicated notation, we define $H_t = A_t = 0$ for $t < 0$.

Lemma 2.4. *For each $n \geq 1$, define the Lebesgue-Stieltjes integral*

$$H_t^n := \frac{1}{A_t - A_{t-1/n}} \int_{t-1/n}^t H_s \, dA_s, \quad t > 0 \qquad (2.2)$$

and $H_0^n := 0$. Then

$$Y^n := H^n X - \int X_- \, dH^n \qquad (2.3)$$

is well defined in the Lebesgue-Stieltjes sense and satisfies

$$Y^n = \, ^{(P)}\!\!\int H^n \, dX \to \, ^{(P)}\!\!\int H \, dX \quad in \quad ucp(P) \quad for \ all \quad P \in \mathcal{P}.$$

Proof. Recalling that $|H| \leq c$ and that A is a predictable increasing càdlàg process, we see that H^n is a predictable process satisfying $|H^n| \leq c$ identically and having càdlàg path of finite variation. In particular, we can use the Lebesgue-Stieltjes integral to define the process Y^n pathwise via (2.3). We deduce via integration by parts that Y^n coincides P-a.s. with the stochastic integral $^{(P)}\!\!\int H^n \, dX$ for each $P \in \mathcal{P}$.

By the standard theorem on approximate identities, we have

$$H^n(\omega) \to H(\omega) \quad \text{in} \quad L^1([0,1], dA(\omega)) \quad \text{for all} \quad \omega \in \Omega. \qquad (2.4)$$

For the remainder of the proof, we fix $P \in \mathcal{P}$ and show the convergence of Y^n to $^{(P)}\!\!\int H \, dX$ in $ucp(P)$. Since P is fixed, we may use the usual tools of stochastic analysis under P and write, as usual, E for the expectation operator under P, etc. (One can pass to the augmentation of $\mathbb{F}^*$ under P to have the "usual assumptions", although this is not important here.)

Recall that the jumps of X are bounded, so that there is a canonical decomposition $X = M + B$, where M is a local martingale and B is predictable of finite variation. Since the jumps of M and B are then also bounded, a

standard localization argument allows us to assume that $\mathrm{Var}(B)$ and the quadratic variation $[M]$ are uniformly bounded. We have

$$E\left[\sup_{t\leq 1}\left|\int_0^t H^n\,dX - \int_0^t H\,dX\right|^2\right]$$
$$\leq 2E\left[\sup_{t\leq 1}\left|\int_0^t (H^n - H)\,dM\right|^2\right] + 2E\left[\left|\int_0^1 |H^n - H|\,d\,\mathrm{Var}(B)\right|^2\right].$$

The second expectation on the right hand side converges to zero; indeed, recalling that $|H|, |H^n| \leq c$, we see that $\int_0^1 |H^n - H|\,d\,\mathrm{Var}(B)$ is uniformly bounded and converges to zero pointwise. The latter follows from (2.4) since $\mathrm{Var}(B)(\omega) \ll A(\omega)$ and

$$\left\{|H^n(\omega) - H(\omega)|\frac{d\,\mathrm{Var}(B)(\omega)}{dA(\omega)}\right\}_{n\geq 1} \subseteq L^1([0,1], dA(\omega))$$

is uniformly integrable.

It remains to show that the first expectation converges to zero. Let $\langle M \rangle$ be the predictable compensator of $[M]$, then the Burkholder-Davis-Gundy inequalities yield

$$E\left[\sup_{t\leq 1}\left|\int_0^t (H^n - H)\,dM\right|^2\right] \leq 4E\left[\int_0^1 |H^n - H|^2\,d[M]\right]$$
$$= 4E\left[\int_0^1 |H^n - H|^2\,d\langle M\rangle\right].$$

Recalling that $|\Delta X| \leq 1$, we have that

$$\langle M \rangle = \langle X^c \rangle + (x^2 \wedge 1) * \nu - \sum_{s\leq \cdot}(\Delta B_s)^2$$

and in particular that $\langle M \rangle \ll A$. Since $\langle M \rangle$ is bounded like $[M]$, we conclude exactly as above that $E[\int_0^1 |H^n - H|^2\,d\langle M\rangle]$ converges to zero. $\qquad\square$

2.2   Aggregation by Approximation in Measure

In this section, we shall find it very useful to employ Mokobodzki's medial limit, which yields a universal method (i.e., independent of the underlying probability) to identify the limit of a sequence which converges in probability. More precisely, "lim med" is a mapping on the set of real sequences with the following property (cf. [9, Theorems 3, 4]): If (Z_n) is a sequence of random variables on a measurable space $(\Omega', \mathcal{F}')$, then $Z(\omega) := \lim\mathrm{med}_n Z_n(\omega)$ is universally measurable and if P is a probability measure on $(\Omega', \mathcal{F}')$ such that Z_n converges to some random variable Z^P in probability P, then $Z = Z^P$ P-a.s. Here uniform measurability refers to the universal completion of $\mathcal{F}'$ under all probability measures on $(\Omega', \mathcal{F}')$.

Although developed in a different context and apparently not used before in ours, medial limits seem to be tailored to our task. Their construction is usually achieved through a transfinite induction that uses the Continuum Hypothesis (cf. [9]); in fact, it is known that medial limits exist under weaker hypotheses (Fremlin [4, 538S]), but not under ZFC alone (Larson [7]). We shall adopt a sufficient set of axioms; since the Continuum Hypothesis is independent of ZFC, we consider this a pragmatic choice of the model of set theory for our purposes.

We have the following result for càdlàg processes.

Lemma 2.5. *Let $(Y^n)_{n \geq 1}$ be a sequence of $\mathbb{F}^*$-adapted càdlàg processes. Assume that for each $P \in \mathcal{P}$ there exists a càdlàg process Y^P such that $Y_t^n \to Y_t^P$ in measure P for all $t \in [0,1]$. Then there exists an $\mathbb{F}^*$-adapted càdlàg process Y such that $Y = Y^P$ P-a.s. for all $P \in \mathcal{P}$.*

Proof. Let $r \in [0,1]$ be a rational number and define $\tilde{Y}_r := \lim \mathrm{med}_n Y_r^n$. Then $\tilde{Y}_r$ is measurable with respect to the universal completion of $\mathcal{F}_r$, which is contained in $\mathcal{F}_r^*$. Moreover,

$$\tilde{Y}_r = Y_r^P \quad P\text{-a.s.} \quad \text{for all} \quad P \in \mathcal{P}. \tag{2.5}$$

Given arbitrary $t \in [0,1)$, we define $Y_t := \limsup_{r \downarrow t} \tilde{Y}_r$, where r is rational (and $Y_1 := \tilde{Y}_1$). Fix $P \in \mathcal{P}$. Since Y^P is càdlàg, (2.5) entails that

$$Y_t = \limsup_{r \downarrow t} \tilde{Y}_r = \limsup_{r \downarrow t} Y_r^P = Y_t^P \quad P\text{-a.s.}$$

and that the $\limsup$ is actually a limit outside a P-nullset. As a consequence,

$$N = \{\omega \in \Omega : Y.(\omega) \text{ is not càdlàg}\}$$

is a P-nullset. Since $P \in \mathcal{P}$ was arbitrary, N is actually a nullset under every $P \in \mathcal{P}$ and therefore contained in $\mathcal{F}_0^*$. We redefine $Y \equiv 0$ on N, then Y is $\mathbb{F}^*$-adapted and all paths of Y are càdlàg. Since Y is also a P-modification of Y^P, we have $Y = Y^P$ P-a.s. $\qquad \square$

Our main result can then be proved as follows.

Proof of Theorem 2.2. We first assume that H is uniformly bounded. Then Lemma 2.4 yields a sequence $\int H^n \, dX$ of pathwise defined integrals which converge in $ucp(P)$ for all $P \in \mathcal{P}$. According to Lemma 2.5, there exists a process $\int H \, dX$ with the desired properties. For general H, we use the previous argument to define $\int H \mathbf{1}_{|H| \leq n} \, dX$ for $n \geq 1$. In view of (2.1), we may apply Lemma 2.5 once more to obtain $\int H \, dX$. $\qquad \square$

Remark 2.6. (i) If the integrand H has left-continuous paths, the assertion of Theorem 2.2 holds true without Assumption 2.1. Indeed, set $A_t := t$ and define H^n as in (2.2). Then, by the left-continuity, we have $H_t^n(\omega) \to H_t(\omega)$ for all t and ω, without exceptional set. The rest of the proof is as above. (Of course, there are other ways to define H^n in this case, such as discretization.)

(ii) Theorem 2.2 can be obtained in a filtration slightly smaller than $\mathbb{F}^*$. Indeed, the same proofs apply if $\mathbb{F}^*$ is replaced by the filtration obtained as follows: first, augment $\mathbb{F}$ by the collection $\cap_{P \in \mathcal{P}} \mathcal{N}^P$ of $\mathcal{P}$-polar sets, then, take the universal augmentation with respect to all probability measures (and not just those in $\mathcal{P}$). Our proofs also show that the random variable $\int_0^1 H \, dX$ is measurable with respect to the universal completion of $\mathcal{F}_1$, without adding the $\mathcal{P}$-polar sets.

(iii) Needless to say, our "construction" of the stochastic integral is not "constructive" in the proper sense; it merely yields an existence result.

References

[1] K. Bichteler. Stochastic integration and L^p-theory of semimartingales. *Ann. Probab.*, 9(1):49–89, 1981.

[2] L. Denis and C. Martini. A theoretical framework for the pricing of contingent claims in the presence of model uncertainty. *Ann. Appl. Probab.*, 16(2):827–852, 2006.

[3] H. Föllmer. Calcul d'Itô sans probabilités. In *Séminaire de Probabilités XV*, volume 850 of *Lecture Notes in Math.*, pages 143–150, Springer, Berlin, 1981.

[4] D. H. Fremlin. *Measure Theory*, volume 5. Torres Fremlin, Colchester, 2008.

[5] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer, Berlin, 2nd edition, 2003.

[6] R. L. Karandikar. On pathwise stochastic integration. *Stochastic Process. Appl.*, 57(1):11–18, 1995.

[7] P. B. Larson. The filter dichotomy and medial limits. *J. Math. Log.*, 9(2):159–165, 2009.

[8] X. Li and S. Peng. Stopping times and related Itô's calculus with G-Brownian motion. *Stochastic Process. Appl.*, 121(7):1492–1508, 2011.

[9] P. A. Meyer. Limites médiales, d'après Mokobodzki. In *Séminaire de Probabilités VII (1971/72)*, volume 321 of *Lecture Notes in Math.*, pages 198–204. Springer, Berlin, 1973.

[10] M. Nutz and H. M. Soner. Superhedging and dynamic risk measures under volatility uncertainty. To appear in *SIAM J. Control Optim.*

[11] S. Peng. Nonlinear expectations and stochastic calculus under uncertainty. *Preprint arXiv:1002.4546v1*, 2010.

[12] H. M. Soner, N. Touzi, and J. Zhang. Quasi-sure stochastic analysis through aggregation. *Electron. J. Probab.*, 16(2):1844–1879, 2011.

[13] H. M. Soner, N. Touzi, and J. Zhang. Wellposedness of second order backward SDEs. To appear in *Probab. Theory Related Fields.*

[14] H. M. Soner, N. Touzi, and J. Zhang. Dual formulation of second order target problems. To appear in *Ann. Appl. Probab.*

[15] H. M. Soner, N. Touzi, and J. Zhang. Martingale representation theorem for the G-expectation. *Stochastic Process. Appl.*, 121(2):265–287, 2011.

[16] W. Willinger and M. S. Taqqu. Pathwise approximations of processes based on the fine structure of their filtration. In *Séminaire de Probabilités XXII*, volume 1321 of *Lecture Notes in Math.*, pages 542–599. Springer, Berlin, 1988.

[17] W. Willinger and M. S. Taqqu. Pathwise stochastic integration and applications to the theory of continuous trading. *Stochastic Process. Appl.*, 32(2):253–280, 1989.