Weak Approximation of G-Expectations

Yan Dolinsky* Marcel Nutz† H. Mete Soner‡

March 2, 2011

Abstract

We introduce a notion of volatility uncertainty in discrete time and define the corresponding analogue of Peng's G-expectation. In the continuous-time limit, the resulting sublinear expectation converges weakly to the G-expectation. This can be seen as a Donsker-type result for the G-Brownian motion.

Keywords G-expectation, volatility uncertainty, weak limit theorem
AMS 2000 Subject Classifications 60F05, 60G44, 91B25, 91B30
JEL Classifications G13, G32

Acknowledgements Research supported by European Research Council Grant 228053-FiRM, Swiss National Science Foundation Grant PDFM2-120424/1, Swiss Finance Institute and ETH Foundation.

1 Introduction

The so-called G-expectation [12, 13, 14] is a nonlinear expectation advancing the notions of backward stochastic differential equations (BSDEs) [10] and g-expectations [11]; see also [2, 16] for a related theory of second order BSDEs. A G-expectation $\xi \mapsto \mathcal{E}^G(\xi)$ is a sublinear function which maps random variables ξ on the canonical space $\Omega = C([0,T];\mathbb{R})$ to the real numbers. The symbol G refers to a given function $G : \mathbb{R} \to \mathbb{R}$ of the form

$$G(\gamma) = \frac{1}{2}(R\gamma^+ - r\gamma^-) = \frac{1}{2} \sup_{a\in[r,R]} a\gamma,$$

where $0 \leq r \leq R < \infty$ are fixed numbers. More generally, the interval $[r, R]$ is replaced by a set $\mathbf{D}$ of nonnegative matrices in the multivariate case. The extension to a random set $\mathbf{D}$ is studied in [9].

The construction of $\mathcal{E}^G(\xi)$ runs as follows. When $\xi = f(B_T)$, where B_T is the canonical process at time T and f is a sufficiently regular function, then $\mathcal{E}^G(\xi)$ is defined to be the initial value $u(0, 0)$ of the solution of the nonlinear backward heat equation $-\partial_t u - G(u_{xx}) = 0$ with terminal condition

*ETH Zurich, Dept. of Mathematics, CH-8092 Zurich, yan.dolinsky@math.ethz.ch
†ETH Zurich, Dept. of Mathematics, CH-8092 Zurich, marcel.nutz@math.ethz.ch
‡ETH Zurich, Dept. of Mathematics, CH-8092 Zurich, and Swiss Finance Institute, mete.soner@math.ethz.ch

$u(\cdot,T)=f$. The mapping $\mathcal{E}^G$ can be extended to random variables of the form $\xi=f(B_{t_1},\dots,B_{t_n})$ by a stepwise evaluation of the PDE and then to the completion $\mathbb{L}^1_G$ of the space of all such random variables. The space $\mathbb{L}^1_G$ consists of so-called quasi-continuous functions and contains in particular all bounded continuous functions on Ω; however, not all bounded measurable functions are included (cf. [3]). While this setting is not based on a single probability measure, the so-called G-Brownian motion is given by the canonical process B "seen" under $\mathcal{E}^G$ (cf. [14]). It reduces to the standard Brownian motion if $r=R=1$ since $\mathcal{E}^G$ is then the (linear) expectation under the Wiener measure.

In this note we introduce a discrete-time analogue of the G-expectation and we prove a convergence result which resembles Donsker's theorem for the standard Brownian motion; the main purpose is to provide additional intuition for G-Brownian motion and volatility uncertainty. Our starting point is the dual view on G-expectation via volatility uncertainty [3, 4]: We consider the representation

$$\mathcal{E}^G(\xi)=\sup_{P\in\mathcal{P}}E^P[\xi], \tag{1.1}$$

where $\mathcal{P}$ is a set of probabilities on Ω such that under any $P\in\mathcal{P}$, the canonical process B is a martingale with volatility $d\langle B\rangle/dt$ taking values in $\mathbf{D}=[r,R]$, $P\times dt$-a.e. Therefore, $\mathbf{D}$ can be understood as the domain of (Knightian) volatility uncertainty and $\mathcal{E}^G$ as the corresponding worst-case expectation. In discrete-time, we translate this to uncertainty about the conditional variance of the increments. Thus we define a sublinear expectation $\mathcal{E}^n$ on the n-step canonical space in the spirit of (1.1), replacing $\mathcal{P}$ by a suitable set of martingale laws. A natural push-forward then yields a sublinear expectation on Ω, which we show to converge weakly to $\mathcal{E}^G$ as $n\to\infty$, if the domain $\mathbf{D}$ of uncertainty is scaled by $1/n$ (cf. Theorem 2.2). The proof relies on (linear) probability theory; in particular, it does not use the central limit theorem for sublinear expectations [14, 15]. The relation to the latter is nontrivial since our discrete-time models do not have independent increments. We remark that quite different approximations of the G-expectation (for the scalar case) can be found in discrete models for financial markets with transaction costs [8] or illiquidity [5].

The detailed setup and the main result are stated in Section 2, whereas the proofs and some ramifications are given in Section 3.

2 Main Result

We fix the dimension $d\in\mathbb{N}$ and denote by $|\cdot|$ the Euclidean norm on $\mathbb{R}^d$. Moreover, we denote by $\mathbb{S}^d$ the space of $d\times d$ symmetric matrices and by $\mathbb{S}^d_+$ its subset of nonnegative definite matrices. We fix a nonempty, convex and compact set $\mathbf{D} \subseteq \mathbb{S}^d_+$; the elements of $\mathbf{D}$ will be the possible values of our volatility processes.

Continuous-Time Formulation. Let $\Omega = C([0,T];\mathbb{R}^d)$ be the space of d-dimensional continuous paths $\omega = (\omega_t)_{0\le t\le T}$ with time horizon $T \in (0,\infty)$, endowed with the uniform norm $\|\omega\|_\infty = \sup_{0\le t\le T}|\omega_t|$. We denote by $B = (B_t)_{0\le t\le T}$ the canonical process $B_t(\omega) = \omega_t$ and by $\mathcal{F}_t := \sigma(B_s,\, 0 \le s \le t)$ the canonical filtration. A probability measure P on Ω is called a *martingale law* if B is a P-martingale and $B_0 = 0$ P-a.s. (All our martingales will start at the origin.) We set

$$\mathcal{P}_{\mathbf{D}} = \big\{P \text{ martingale law on } \Omega:\ d\langle B\rangle_t/dt \in \mathbf{D},\ P \times dt\text{-a.e.}\big\},$$

where $\langle B\rangle$ denotes the matrix-valued process of quadratic covariations. We can then define the sublinear expectation

$$\mathcal{E}_{\mathbf{D}}(\xi) := \sup_{P\in\mathcal{P}_{\mathbf{D}}} E^P[\xi] \quad \text{for any random variable } \xi : \Omega \to \mathbb{R}$$

such that ξ is $\mathcal{F}_T$-measurable and $E^P|\xi| < \infty$ for all $P \in \mathcal{P}_{\mathbf{D}}$. The mapping $\mathcal{E}_{\mathbf{D}}$ coincides with the G-expectation (on its domain $\mathbb{L}^1_G$) if $G : \mathbb{S}^d \to \mathbb{R}$ is (half) the support function of $\mathbf{D}$; i.e., $G(\Gamma) = \sup_{A\in\mathbf{D}} \operatorname{trace}(\Gamma A)/2$. Indeed, this follows from [3] with an additional density argument as detailed in Remark 3.6 below.

Discrete-Time Formulation. Given $n \in \mathbb{N}$, we consider $(\mathbb{R}^d)^{n+1}$ as the canonical space of d-dimensional paths in discrete time $k = 0, 1, \dots, n$. We denote by $X^n = (X^n_k)^n_{k=0}$ the canonical process defined by $X^n_k(x) = x_k$ for $x = (x_0, \dots, x_n) \in (\mathbb{R}^d)^{n+1}$. Moreover, $\mathcal{F}^n_k = \sigma(X^n_i,\, i = 0, \dots, k)$ defines the canonical filtration $(\mathcal{F}^n_k)^n_{k=0}$. We also introduce $0 \le r_{\mathbf{D}} \le R_{\mathbf{D}} < \infty$ such that $[r_{\mathbf{D}}, R_{\mathbf{D}}]$ is the spectrum of $\mathbf{D}$; i.e.,

$$r_{\mathbf{D}} = \inf_{\Gamma\in\mathbf{D}} \|\Gamma^{-1}\|^{-1} \quad \text{and} \quad R_{\mathbf{D}} = \sup_{\Gamma\in\mathbf{D}} \|\Gamma\|,$$

where $\|\cdot\|$ denotes the operator norm and we set $r_{\mathbf{D}} := 0$ if $\mathbf{D}$ has an element which is not invertible. We note that $[r_{\mathbf{D}}, R_{\mathbf{D}}] = \mathbf{D}$ if $d = 1$. Finally, a probability measure P on $(\mathbb{R}^d)^{n+1}$ is called a martingale law if X^n is a P-martingale and $X^n_0 = 0$ P-a.s. Denoting by $\Delta X^n_k = X^n_k - X^n_{k-1}$ the increments of X^n, we can now set

$$\mathcal{P}^n_{\mathbf{D}} = \left\{ \begin{array}{l} P \text{ martingale law on } (\mathbb{R}^d)^{n+1}: \ \text{ for } k = 1, \dots, n, \\ E^P[\Delta X^n_k (\Delta X^n_k)' | \mathcal{F}^n_{k-1}] \in \mathbf{D} \text{ and } d^2 r_{\mathbf{D}} \le |\Delta X^n_k|^2 \le d^2 R_{\mathbf{D}},\ P\text{-a.s.} \end{array} \right\},$$

where prime ($'$) denotes transposition. Note that ΔX^n_k is a column vector, so that $\Delta X^n (\Delta X^n)'$ takes values in $\mathbb{S}^d_+$. We introduce the sublinear expectation

$$\mathcal{E}^n_{\mathbf{D}}(\psi) := \sup_{P\in\mathcal{P}^n_{\mathbf{D}}} E^P[\psi] \quad \text{for any random variable } \psi : (\mathbb{R}^d)^{n+1} \to \mathbb{R}$$

such that ψ is $\mathcal{F}_n^n$-measurable and $E^P|\psi| < \infty$ for all $P \in \mathcal{P}_{\mathbf{D}}^n$, and we think of $\mathcal{E}_{\mathbf{D}}^n$ as a discrete-time analogue of the G-expectation.

Remark 2.1. The second condition in the definition of $\mathcal{P}_{\mathbf{D}}^n$ is motivated by the desire to generate the volatility uncertainty by a *small* set of scenarios; we remark that the main results remain true if, e.g., the lower bound $r_{\mathbf{D}}$ is omitted and the upper bound $R_{\mathbf{D}}$ replaced by any other condition yielding tightness. Our bounds are chosen so that

$$\mathcal{P}_{\mathbf{D}}^n = \big\{P \text{ martingale law on } (\mathbb{R}^d)^{n+1} : \Delta X^n (\Delta X^n)' \in \mathbf{D},\ P\text{-a.s.}\big\} \quad \text{if } d = 1.$$

Continuous-Time Limit. To compare our objects from the two formulations, we shall extend any discrete path $x \in (\mathbb{R}^d)^{n+1}$ to a continuous path $\widehat{x} \in \Omega$ by linear interpolation. More precisely, we define the interpolation operator

$$\begin{aligned} \widehat{\ }: \quad & (\mathbb{R}^d)^{n+1} \to \Omega, \quad x = (x_0, \dots, x_n) \mapsto \widehat{x} = (\widehat{x}_t)_{0 \le t \le T}, \quad \text{where} \\ & \widehat{x}_t := ([nt/T] + 1 - nt/T) x_{[nt/T]} + (nt/T - [nt/T]) x_{[nt/T]+1} \end{aligned}$$

and $[y] := \max\{m \in \mathbb{Z} : m \le y\}$ for $y \in \mathbb{R}$. In particular, if X^n is the canonical process on $(\mathbb{R}^d)^{n+1}$ and ξ is a random variable on Ω, then $\xi(\widehat{X^n})$ defines a random variable on $(\mathbb{R}^d)^{n+1}$. This allows us to define the following push-forward of $\mathcal{E}_{\mathbf{D}}^n$ to a continuous-time object,

$$\widehat{\mathcal{E}}_{\mathbf{D}}^n(\xi) := \mathcal{E}_{\mathbf{D}}^n(\xi(\widehat{X^n})) \quad \text{for} \quad \xi : \Omega \to \mathbb{R}$$

being suitably integrable.

Our main result states that this sublinear expectation with discrete-time volatility uncertainty converges to the G-expectation as the number n of periods tends to infinity, if the domain of volatility uncertainty is scaled as $\mathbf{D}/n := \{n^{-1}\Gamma : \Gamma \in \mathbf{D}\}$.

Theorem 2.2. *Let $\xi : \Omega \to \mathbb{R}$ be a continuous function satisfying $|\xi(\omega)| \le c(1 + \|\omega\|_\infty)^p$ for some constants $c, p > 0$. Then $\widehat{\mathcal{E}}_{\mathbf{D}/n}^n(\xi) \to \mathcal{E}_{\mathbf{D}}(\xi)$ as $n \to \infty$; that is,*

$$\sup_{P \in \mathcal{P}_{\mathbf{D}/n}^n} E^P[\xi(\widehat{X^n})] \to \sup_{P \in \mathcal{P}_{\mathbf{D}}} E^P[\xi]. \tag{2.1}$$

We shall see that all expressions in (2.1) are well defined and finite. Moreover, we will show in Theorem 3.8 that the result also holds true for a "strong" formulation of volatility uncertainty.

Remark 2.3. Theorem 2.2 cannot be extended to the case where ξ is merely in $\mathbb{L}_G^1$, which is defined as the completion of $C_b(\Omega; \mathbb{R})$ under the norm $\|\xi\|_{L_G^1} := \sup\{E^P|\xi|,\ P \in \mathcal{P}_{\mathbf{D}}\}$. This is because $\|\cdot\|_{L_G^1}$ "does not see" the discrete-time objects, as illustrated by the following example. Assume for simplicity that $0 \notin \mathbf{D}$ and let $A \subset \Omega$ be the set of paths with finite variation. Since $P(A) = 0$ for any $P \in \mathcal{P}_{\mathbf{D}}$, we have $\xi := 1 - \mathbf{1}_A = 1$ in $\mathbb{L}^1_G$ and the right hand side of (2.1) equals one. However, the trajectories of $\widehat{X^n}$ lie in A, so that $\xi(\widehat{X^n}) \equiv 0$ and the left hand side of (2.1) equals zero.

In view of the previous remark, we introduce a smaller space $\mathbb{L}^1_*$, defined as the completion of $C_b(\Omega; \mathbb{R})$ under the norm

$$\|\xi\|_* := \sup_{Q \in \mathcal{Q}} E^Q|\xi|, \quad \mathcal{Q} := \mathcal{P}_{\mathbf{D}} \cup \big\{P \circ (\widehat{X^n})^{-1} : P \in \mathcal{P}^n_{\mathbf{D}/n},\ n \in \mathbb{N}\big\}. \tag{2.2}$$

If ξ is as in Theorem 2.2, then $\xi \in \mathbb{L}^1_*$ by Lemma 3.4 below and so the following is a generalization of Theorem 2.2.

Corollary 2.4. *Let $\xi \in \mathbb{L}^1_*$. Then $\widehat{\mathcal{E}}^n_{\mathbf{D}/n}(\xi) \to \mathcal{E}_{\mathbf{D}}(\xi)$ as $n \to \infty$.*

Proof. This follows from Theorem 2.2 by approximation, using that $\|\xi\|_*$ and $\sup\{E^P|\xi| : P \in \mathcal{P}_{\mathbf{D}}\} + \sup\{E^P|\xi(\widehat{X^n})| : P \in \mathcal{P}^n_{\mathbf{D}/n},\ n \in \mathbb{N}\}$ are equivalent norms. □

3 Proofs and Ramifications

In the next two subsections, we prove separately two inequalities that jointly imply Theorem 2.2 and a slightly stronger result, reported in Theorem 3.8.

3.1 First Inequality

In this subsection we prove the first inequality of (2.1), namely that

$$\limsup_{n\to\infty} \sup_{P \in \mathcal{P}^n_{\mathbf{D}/n}} E^P[\xi(\widehat{X^n})] \le \sup_{P \in \mathcal{P}_{\mathbf{D}}} E^P[\xi]. \tag{3.1}$$

The essential step in this proof is a stability result for the volatility (see Lemma 3.3(ii) below); the necessary tightness follows from the compactness of $\mathbf{D}$; i.e., from $R_{\mathbf{D}} < \infty$. We shall denote $\lambda\mathbf{D} = \{\lambda\Gamma : \Gamma \in \mathbf{D}\}$ for $\lambda \in \mathbb{R}$.

Lemma 3.1. *Given $p \in [1, \infty)$, there exists a universal constant $K > 0$ such that for all $0 \le k \le l \le n$ and $P \in \mathcal{P}^n_{\mathbf{D}}$,*

(i) $E^P[\sup_{k=0,\dots,n} |X^n_k|^{2p}] \le K(nR_{\mathbf{D}})^p$,

(ii) $E^P|X^n_l - X^n_k|^4 \le KR^2_{\mathbf{D}}(l-k)^2$,

(iii) $E^P[(X^n_l - X^n_k)(X^n_l - X^n_k)'|\mathcal{F}^n_k] \in (l-k)\mathbf{D}$ *P-a.s.*

Proof. We set $X := X^n$ to ease the notation.

(i) Let $p \in [1, \infty)$. By the Burkholder-Davis-Gundy (BDG) inequalities there exists a universal constant $C = C(p, d)$ such that

$$E^P\left[\sup_{k=0,\dots,n} |X^n_k|^{2p}\right] \le CE^P\|[X]_n\|^p.$$

In view of $P \in \mathcal{P}^n_{\mathbf{D}}$, we have $\|[X]_n\| = \|\sum_{i=1}^n \Delta X_i(\Delta X_i)'\| \leq nd^2 R_{\mathbf{D}}$ P-a.s.

(ii) The BDG inequalities yield a universal constant C such that

$$E^P|X_l - X_k|^4 \leq CE^P\|[X]_l - [X]_k\|^2.$$

Similarly as in (i), $P \in \mathcal{P}^n_{\mathbf{D}}$ implies that $\|[X]_l - [X]_k\| \leq (l-k)d^2 R_{\mathbf{D}}$ P-a.s.

(iii) The orthogonality of the martingale increments yields that

$$E^P[(X_l - X_k)(X_l - X_k)'|\mathcal{F}^n_k] = \sum_{i=k+1}^{l} E^P[\Delta X_i(\Delta X_i)'|\mathcal{F}^n_k].$$

Since $E^P[\Delta X_i(\Delta X_i)'|\mathcal{F}^n_{i-1}] \in \mathbf{D}$ P-a.s. and since $\mathbf{D}$ is convex,

$$E^P[\Delta X_i(\Delta X_i)'|\mathcal{F}^n_k] = E^P\big[E^P[\Delta X_i(\Delta X_i)'|\mathcal{F}^n_{i-1}]\big|\mathcal{F}^n_k\big]$$

again takes values in $\mathbf{D}$. It remains to observe that if $\Gamma_1, \dots, \Gamma_m \in \mathbf{D}$, then $\Gamma_1 + \dots + \Gamma_m \in m\mathbf{D}$ by convexity. □

The following lemma shows in particular that all expressions in Theorem 2.2 are well defined and finite.

Lemma 3.2. *Let $\xi : \Omega \to \mathbb{R}$ be as in Theorem 2.2. Then $\|\xi\|_* < \infty$; that is,*

$$\sup_{n\in\mathbb{N}} \sup_{P\in\mathcal{P}^n_{\mathbf{D}/n}} E^P|\xi(\widehat{X^n})| < \infty \quad \text{and} \quad \sup_{P\in\mathcal{P}_{\mathbf{D}}} E^P|\xi| < \infty. \tag{3.2}$$

Proof. Let $n \in \mathbb{N}$ and $P \in \mathcal{P}^n_{\mathbf{D}/n}$. By the assumption on ξ, there exist constants $c, p > 0$ such that

$$E^P|\xi(\widehat{X^n})| \leq c + cE^P\bigg[\sup_{0\leq t\leq T} |\widehat{X^n_t}|^p\bigg] \leq c + cE^P\bigg[\sup_{k=0,\dots,n} |X^n_k|^p\bigg].$$

Hence Lemma 3.1(i) and the observation that $R_{\mathbf{D}/n} = R_{\mathbf{D}}/n$ yield that $E^P|\xi(\widehat{X^n})| \leq KR_{\mathbf{D}}^{p/2}$ and the first claim follows. The second claim similarly follows from the estimate that $E^P[\sup_{0\leq t\leq T}|B_t|^p] \leq C_p$ for all $P \in \mathcal{P}_{\mathbf{D}}$, which is obtained from the BDG inequalities by using that $\mathbf{D}$ is bounded. □

We can now prove the key result of this subsection.

Lemma 3.3. *For each $n \in \mathbb{N}$, let $\{M^n = (M^n_k)_{k=0}^n, \tilde{P}^n\}$ be a martingale with law $P^n \in \mathcal{P}^n_{\mathbf{D}/n}$ on $(\mathbb{R}^d)^{n+1}$ and let Q^n be the law of $\widehat{M^n}$ on Ω. Then*

(i) the sequence (Q^n) is tight on Ω,

(ii) any cluster point of (Q^n) is an element of $\mathcal{P}_{\mathbf{D}}$.

Proof. (i) Let $0 \leq s \leq t \leq T$. As $R_{\mathbf{D}/n} = R_{\mathbf{D}}/n$, Lemma 3.1(ii) implies that

$$E^{Q^n}|B_t - B_s|^4 = E^{\tilde{P}^n}|\widehat{M^n_t} - \widehat{M^n_s}|^4 \leq C|t-s|^2$$

for a constant $C > 0$. Hence (Q^n) is tight by the moment criterion.

(ii) Let Q be a cluster point, then B is a Q-martingale as a consequence of the uniform integrability implied by Lemma 3.1(i) and it remains to show that $d\langle B\rangle_t/dt \in \mathbf{D}$ holds $Q \times dt$-a.e. It will be useful to characterize $\mathbf{D}$ by scalar inequalities: given $\Gamma \in \mathbb{S}^d$, the separating hyperplane theorem implies that

$$\Gamma \in \mathbf{D} \quad \text{if and only if} \quad \ell(\Gamma) \leq C^{\ell}_{\mathbf{D}} := \sup_{A \in \mathbf{D}} \ell(A) \quad \text{for all} \quad \ell \in (\mathbb{S}^d)^*, \tag{3.3}$$

where $(\mathbb{S}^d)^*$ is the set of all linear functionals $\ell : \mathbb{S}^d \to \mathbb{R}$.

Let $H : [0,T] \times \Omega \to [0,1]$ be a continuous and adapted function and let $\ell \in (\mathbb{S}^d)^*$. We fix $0 \leq s < t \leq T$ and denote $\Delta_{s,t}Y := Y_t - Y_s$ for a process $Y = (Y_u)_{0 \leq u \leq T}$. Let $\varepsilon > 0$ and let $\tilde{\mathbf{D}}$ be any neighborhood of $\mathbf{D}$, then for n sufficiently large,

$$E^{\tilde{P}^n}\Big[(\Delta_{s,t}\widehat{M^n})(\Delta_{s,t}\widehat{M^n})'\Big|\sigma\big(\widehat{M^n_u}, 0 \leq u \leq s-\varepsilon\big)\Big] \in (t-s)\widetilde{\mathbf{D}} \quad \tilde{P}^n\text{-a.s.}$$

as a consequence of Lemma 3.1(iii). Since $\tilde{\mathbf{D}}$ was arbitrary, it follows by (3.3) that

$$\begin{aligned}
&\limsup_{n\to\infty} E^{Q^n}\big[H(s-\varepsilon, B)\,\big\{\ell\big((\Delta_{s,t}B)(\Delta_{s,t}B)'\big) - C^{\ell}_{\mathbf{D}}(t-s)\big\}\big] \\
&= \limsup_{n\to\infty} E^{\tilde{P}^n}\big[H(s-\varepsilon, \widehat{M^n})\,\big\{\ell\big((\Delta_{s,t}\widehat{M^n})(\Delta_{s,t}\widehat{M^n})'\big) - C^{\ell}_{\mathbf{D}}(t-s)\big\}\big] \leq 0.
\end{aligned}$$

Using (3.2) with $\xi(\omega) = \|\omega\|^2_\infty$, we may pass to the limit and conclude that

$$E^Q\big[H(s-\varepsilon, B)\,\ell\big((\Delta_{s,t}B)(\Delta_{s,t}B)'\big)\big] \leq E^Q\big[H(s-\varepsilon, B)\,C^{\ell}_{\mathbf{D}}(t-s)\big]. \tag{3.4}$$

Since $H(s-\varepsilon, B)$ is $\mathcal{F}_s$-measurable and

$$E^Q[(\Delta_{s,t}B)(\Delta_{s,t}B)'|\mathcal{F}_s] = E^Q[B_tB_t' - B_sB_s'|\mathcal{F}_s] = E^Q[\langle B\rangle_t - \langle B\rangle_s|\mathcal{F}_s]$$

as B is a square-integrable Q-martingale, (3.4) is equivalent to

$$E^Q\big[H(s-\varepsilon, B)\,\ell\big(\langle B\rangle_t - \langle B\rangle_s\big)\big] \leq E^Q\big[H(s-\varepsilon, B)\,C^{\ell}_{\mathbf{D}}(t-s)\big].$$

Using the continuity of H and dominated convergence as $\varepsilon \to 0$, we obtain

$$E^Q\big[H(s, B)\,\ell\big(\langle B\rangle_t - \langle B\rangle_s\big)\big] \leq E^Q\big[H(s, B)\,C^{\ell}_{\mathbf{D}}(t-s)\big]$$

and then it follows that

$$E^Q\bigg[\int_0^T H(t, B)\,\ell(d\langle B\rangle_t)\bigg] \leq E^Q\bigg[\int_0^T H(t, B) C^{\ell}_{\mathbf{D}}\,dt\bigg].$$

By an approximation argument, this inequality extends to functions H which are measurable instead of continuous. It follows that $\ell(d\langle B\rangle_t/dt) \leq C^\ell_{\mathbf{D}}$ holds $Q\times dt$-a.e., and since $\ell \in (\mathbb{S}^d)^*$ was arbitrary, (3.3) shows that $d\langle B\rangle_t/dt \in \mathbf{D}$ holds $Q \times dt$-a.e. □

We can now deduce the first inequality of Theorem 2.2 as follows.

Proof of (3.1). Let ξ be as in Theorem 2.2 and let $\varepsilon > 0$. For each $n \in \mathbb{N}$ there exists an ε-optimizer $P^n \in \mathcal{P}^n_{\mathbf{D}/n}$; i.e., if Q^n denotes the law of $\widehat{X^n}$ on Ω under P_n, then

$$E^{Q^n}[\xi] = E^{P^n}[\xi(\widehat{X^n})] \geq \sup_{P\in\mathcal{P}^n_{\mathbf{D}/n}} E^P[\xi(\widehat{X^n})] - \varepsilon.$$

By Lemma 3.3, the sequence (Q^n) is tight and any cluster point belongs to $\mathcal{P}_{\mathbf{D}}$. Since ξ is continuous and (3.2) implies $\sup_n E^{Q_n}|\xi| < \infty$, tightness yields that $\limsup_n E^{Q^n}[\xi] \leq \sup_{P\in\mathcal{P}_{\mathbf{D}}} E^P[\xi]$. Therefore,

$$\limsup_{n\to\infty} \sup_{P\in\mathcal{P}^n_{\mathbf{D}/n}} E^P[\xi(\widehat{X^n})] \leq \sup_{P\in\mathcal{P}_{\mathbf{D}}} E^P[\xi] + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, it follows that (3.1) holds. □

Finally, we also prove the statement preceding Corollary 2.4.

Lemma 3.4. *Let $\xi : \Omega \to \mathbb{R}$ be as in Theorem 2.2. Then $\xi \in \mathbb{L}^1_*$.*

Proof. We show that $\xi^m := (\xi \wedge m) \vee m$ converges to ξ in the norm $\|\cdot\|_*$ as $m \to \infty$, or equivalently, that the upper expectation $\sup\{E^Q[\,\cdot\,] : Q \in \mathcal{Q}\}$ is continuous along the decreasing sequence $|\xi - \xi^m|$, where $\mathcal{Q}$ is as in (2.2). Indeed, $\mathcal{Q}$ is tight by (the proof of) Lemma 3.3. Using that $\|\xi\|_* < \infty$ by Lemma 3.2, we can then argue as in the proof of [3, Theorem 12] to obtain the claim. □

3.2 Second Inequality

The main purpose of this subsection is to show the second inequality "$\geq$" of (2.1). Our proof will yield a more precise version of Theorem 2.2. Namely, we will include "strong" formulations of volatility uncertainty both in discrete and in continuous time; i.e., consider laws generated by integrals with respect to a fixed random walk (resp. Brownian motion). In the financial interpretation, this means that the uncertainty can be generated by *complete* market models.

Strong Formulation in Continuous Time. Here we shall consider *Brownian* martingales: with P_0 denoting the Wiener measure, we define

$$\mathcal{Q}_{\mathbf{D}} = \left\{ P_0 \circ \Big(\int f(t,B)\, dB_t \Big)^{-1} : \ f \in C\big([0,T] \times \Omega; \sqrt{\mathbf{D}}\big) \text{ adapted} \right\},$$

where $\sqrt{\mathbf{D}} = \{\sqrt{\Gamma} : \Gamma \in \mathbf{D}\}$. (For $\Gamma \in \mathbb{S}^d_+$, $\sqrt{\Gamma}$ denotes the unique square-root in $\mathbb{S}^d_+$.) We note that $\mathcal{Q}_{\mathbf{D}}$ is a (typically strict) subset of $\mathcal{P}_{\mathbf{D}}$. The elements of $\mathcal{Q}_{\mathbf{D}}$ with nondegenerate f have the predictable representation property; i.e., they correspond to a complete market in the terminology of mathematical finance. We have the following density result; the proof is deferred to the end of the section.

Proposition 3.5. *The convex hull of $\mathcal{Q}_D$ is a weakly dense subset of $\mathcal{P}_{\mathbf{D}}$.*

We can now deduce the connection between $\mathcal{E}_{\mathbf{D}}$ and the G-expectation associated with $\mathbf{D}$.

Remark 3.6. (i) Proposition 3.5 implies that

$$\sup_{P \in \mathcal{Q}_{\mathbf{D}}} E^P[\xi] = \sup_{P \in \mathcal{P}_{\mathbf{D}}} E^P[\xi], \quad \xi \in C_b(\Omega; \mathbb{R}). \tag{3.5}$$

In [3, Section 3] it is shown that the G-expectation as introduced in [12, 13] coincides with the mapping $\xi \mapsto \sup_{P \in \mathcal{Q}^*_{\mathbf{D}}} E^P[\xi]$ for a certain set $\mathcal{Q}^*_{\mathbf{D}}$ satisfying $\mathcal{Q}_{\mathbf{D}} \subseteq \mathcal{Q}^*_{\mathbf{D}} \subseteq \mathcal{P}_{\mathbf{D}}$. In particular, we deduce that the right hand side of (3.5) is indeed equal to the G-expectation, as claimed in Section 2.

(ii) A result similar to Proposition 3.5 can also be deduced from [17, Proposition 3.4.], which relies on a PDE-based verification argument of stochastic control. We include a (possibly more enlightening) probabilistic proof at the end of the section.

Strong Formulation in Discrete Time. For fixed $n \in \mathbb{N}$, we consider

$$\Omega_n := \big\{\omega = (\omega_1, \dots, \omega_n) : \omega_i \in \{1, \dots, d+1\},\ i = 1, \dots, n\big\}$$

equipped with its power set and let $P_n := \{(d+1)^{-1}, \dots, (d+1)^{-1}\}^n$ be the product probability associated with the uniform distribution. Moreover, let $\xi_1, \dots, \xi_n$ be an i.i.d. sequence of $\mathbb{R}^d$-valued random variables on Ω_n such that $|\xi_k| = d$ and such that the components of ξ_k are orthonormal in $L^2(P_n)$, for each $k = 1, \dots, n$. Let $Z_k = \sum_{l=1}^k \xi_l$ be the associated random walk. Then, we consider martingales M^f which are discrete-time integrals of Z of the form

$$M^f_k = \sum_{l=1}^k f(l-1, Z) \Delta Z_l,$$

where f is measurable and adapted with respect to the filtration generated by Z; i.e., $f(l, Z)$ depends only on $Z|_{\{0,\dots,l\}}$. We define

$$\mathcal{Q}^n_{\mathbf{D}} = \Big\{ P_n \circ (M^f)^{-1};\ f : \{0, \dots, n-1\} \times (\mathbb{R}^d)^{n+1} \to \sqrt{\mathbf{D}} \text{ measurable, adapted} \Big\}.$$

To see that $\mathcal{Q}^n_{\mathbf{D}} \subseteq \mathcal{P}^n_{\mathbf{D}}$, we note that $\Delta_k M^f = f(k-1, Z)\xi_k$ and the orthonormality property of ξ_k yield

$$E^{P_n}\big[\Delta_k M^f \big(\Delta_k M^f\big)' \big| \sigma(Z_1, \dots, Z_{k-1})\big] = f(k-1, Z)^2 \in \mathbf{D} \quad P_n\text{-a.s.},$$

while $|\xi_k| = d$ and $f^2 \in \mathbf{D}$ imply that

$$\big\|\Delta_k M^f \big(\Delta_k M^f\big)'\big\| = |f(k-1, Z)\xi_k|^2 \in \big[d^2 r_{\mathbf{D}}, d^2 R_{\mathbf{D}}\big] \quad P_n\text{-a.s.}$$

Remark 3.7. We recall from [7] that such $\xi_1, \dots, \xi_n$ can be constructed as follows. Let A be an orthogonal $(d+1) \times (d+1)$ matrix whose last row is $((d+1)^{-1/2}, \dots, (d+1)^{-1/2})$ and let $v_l \in \mathbb{R}^d$ be column vectors such that $[v_1, \dots, v_{d+1}]$ is the matrix obtained from A by deleting the last row. Setting $\xi_k(\omega) := (d+1)^{1/2} v_{\omega_k}$ for $\omega = (\omega_1, \dots, \omega_n)$ and $k = 1, \dots, n$, the above requirements are satisfied.

We can now formulate a result which includes Theorem 2.2.

Theorem 3.8. *Let $\xi : \Omega \to \mathbb{R}$ be as in Theorem 2.2. Then*

$$\begin{aligned}
\lim_{n\to\infty} \sup_{P \in \mathcal{Q}^n_{\mathbf{D}/n}} E^P[\xi(\widehat{X^n})] &= \lim_{n\to\infty} \sup_{P \in \mathcal{P}^n_{\mathbf{D}/n}} E^P[\xi(\widehat{X^n})] \\
&= \sup_{P \in \mathcal{Q}_{\mathbf{D}}} E^P[\xi] \\
&= \sup_{P \in \mathcal{P}_{\mathbf{D}}} E^P[\xi]. \qquad (3.6)
\end{aligned}$$

Proof. Since $\mathcal{Q}^n_{\mathbf{D}/n} \subseteq \mathcal{P}^n_{\mathbf{D}/n}$ for each $n \geq 1$, the inequality (3.1) yields that

$$\limsup_{n\to\infty} \sup_{P \in \mathcal{Q}^n_{\mathbf{D}/n}} E^P[\xi(\widehat{X^n})] \leq \sup_{P \in \mathcal{P}_{\mathbf{D}}} E^P[\xi].$$

As the equality in (3.6) follows from Proposition 3.5, it remains to show that

$$\liminf_{n\to\infty} \sup_{P \in \mathcal{Q}^n_{\mathbf{D}/n}} E^P[\xi(\widehat{X^n})] \geq \sup_{P \in \mathcal{Q}_{\mathbf{D}}} E^P[\xi].$$

To this end, let $P \in \mathcal{Q}_{\mathbf{D}}$; i.e., P is the law of a martingale of the form

$$M = \int f(t, W)\, dW_t,$$

where W is a Brownian motion and $f \in C\big([0,T]\times\Omega;\sqrt{\mathbf{D}}\big)$ is an adapted function. We shall construct martingales $M^{(n)}$ whose laws are in $\mathcal{Q}^n_{\mathbf{D}/n}$ and tend to P.

For $n \geq 1$, let $Z^{(n)}_k = \sum_{l=1}^k \xi_l$ be the random walk on (Ω_n, P_n) as introduced before Remark 3.7. Let

$$W^{(n)}_t := n^{-1/2} \sum_{k=1}^{[nt/T]} \xi_k, \quad 0 \leq t \leq T$$

be the piecewise constant càdlàg version of the scaled random walk and let $\hat{W}^{(n)} := n^{-1/2}\widehat{Z^{(n)}}$ be its continuous counterpart obtained by linear interpolation. It follows from the central limit theorem that

$$\big(W^{(n)}, \hat{W}^{(n)}\big) \Rightarrow (W, W) \quad \text{on} \quad D([0,T];\mathbb{R}^{2d}),$$

the space of càdlàg paths equipped with the Skorohod topology. Moreover, since f is continuous, we also have that

$$\big(W^{(n)}, f\big([nt/T]T/n, \hat{W}^{(n)}\big)\big) \Rightarrow (W, f(t,W)) \quad \text{on} \quad D([0,T];\mathbb{R}^{d+d^2}).$$

Thus, if we introduce the discrete-time integral

$$M^{(n)}_k := \sum_{l=1}^k f\big((l-1)T/n, \hat{W}^{(n)}\big)\Big(\hat{W}^{(n)}_{lT/n} - \hat{W}^{(n)}_{(l-1)T/n}\Big),$$

it follows from the stability of stochastic integrals (see [6, Theorem 4.3 and Definition 4.1]) that

$$\Big(M^{(n)}_{[nt/T]}\Big)_{0\leq t\leq T} \Rightarrow M \quad \text{on} \quad D([0,T];\mathbb{R}^d).$$

Moreover, since the increments of $M^{(n)}$ uniformly tend to 0 as $n \to \infty$, it also follows that

$$\widehat{M}^{(n)} \Rightarrow M \quad \text{on} \quad \Omega.$$

As f^2/n takes values in $\mathbf{D}/n$, the law of $M^{(n)}$ is contained in $\mathcal{Q}^n_{\mathbf{D}/n}$ and the proof is complete. □

It remains to give the proof of Proposition 3.5, which we will obtain by a randomization technique. Since similar arguments, at least for the scalar case, can be found elsewhere (e.g., [8, Section 5]), we shall be brief.

Proof of Proposition 3.5. We may assume without loss of generality that

$$\text{there exists an invertible element } \Gamma_* \in \mathbf{D}. \tag{3.7}$$

Indeed, using that $\mathbf{D}$ is a convex subset of $\mathbb{S}^d_+$, we observe that (3.7) is equivalent to $K = \{0\}$ for $K := \bigcap_{\Gamma\in\mathbf{D}} \ker \Gamma$. If $k = \dim K > 0$, a change of coordinates bring us to the situation where K corresponds to the last k coordinates of $\mathbb{R}^d$. We can then reduce all considerations to $\mathbb{R}^{d-k}$ and thereby recover the situation of (3.7).

1. Regularization. We first observe that the set

$$\big\{P \in \mathcal{P}_{\mathbf{D}} : d\langle B\rangle_t/dt \geq \varepsilon \mathbb{1}_d \ P \times dt\text{-a.e. for some } \varepsilon > 0\big\} \tag{3.8}$$

is weakly dense in $\mathcal{P}_{\mathbf{D}}$. (Here $\mathbb{1}_d$ denotes the unit matrix.) Indeed, let M be a martingale whose law is in $\mathcal{P}_{\mathbf{D}}$. Recall (3.7) and let N be an independent continuous Gaussian martingale with $d\langle N\rangle_t/dt = \Gamma_*$. For $\lambda \uparrow 1$, the law of $\lambda M + (1-\lambda)N$ tends to the law of M and is contained in the set (3.8), since $\mathbf{D}$ is convex.

2. Discretization. Next, we reduce to martingales with piecewise constant volatility. Let M be a martingale whose law belongs to (3.8). We have

$$M = \int \sigma_t \, dW_t \quad \text{for} \quad \sigma_t := \sqrt{d\langle M\rangle/dt} \quad \text{and} \quad W := \int \sigma_t^{-1} \, dM_t,$$

where W is a Brownian motion by Lévy's theorem. For $n \geq 1$, we introduce $M^{(n)} = \int \sigma_t^{(n)} \, dW_t$, where $\sigma^{(n)}$ is an $\mathbb{S}^d_+$-valued piecewise constant process satisfying

$$\big(\sigma_t^{(n)}\big)^2 = \Pi_{\mathbf{D}}\left[\left(\frac{n}{T}\int_{(k-1)T/n}^{kT/n} \sigma_s \, ds\right)^2\right], \quad t \in \big(kT/n, (k+1)T/n\big]$$

for $k = 1, \dots, n-1$, where $\Pi_{\mathbf{D}} : \mathbb{S}^d \to \mathbf{D}$ is the Euclidean projection. On $[0, T/n]$ one can take, e.g., $\sigma^{(n)} := \sqrt{\Gamma_*}$. We then have

$$E\big\|\langle M - M^{(n)}\rangle_T\big\| = E\int_0^T \big\|\sigma_t - \sigma_t^{(n)}\big\|^2 \, dt \to 0$$

and in particular $M^{(n)}$ converges weakly to M.

3. Randomization. Consider a martingale of the form $M = \int \sigma_t \, dW_t$, where W is a Brownian motion on some given filtered probability space and σ is an adapted $\sqrt{\mathbf{D}}$-valued process which is piecewise constant; i.e.,

$$\sigma = \sum_{k=0}^{n-1} \mathbf{1}_{[t_k, t_{k+1})}\sigma(k) \quad \text{for some} \quad 0 = t_0 < t_1 < \cdots < t_n = T$$

and some $n \geq 1$. Consider also a second probability space carrying a Brownian motion $\tilde{W}$ and a sequence $U^1, \dots, U^n$ of $\mathbb{R}^{d\times d}$-valued random variables such that the components $\{U^k_{ij} : 1 \leq i, j \leq d;\ 1 \leq k \leq n\}$ are i.i.d. uniformly distributed on $(0,1)$ and independent of $\tilde{W}$.

Using the existence of regular conditional probability distributions, we can construct functions $\Theta_k : C([0,t_k]; \mathbb{R}^d) \times (0,1)^{d^2} \times \cdots \times (0,1)^{d^2} \to \sqrt{\mathbf{D}}$ such that the random variables $\tilde{\sigma}(k) := \Theta_k(\tilde{W}|_{[0,t_k]}, U^1, \dots, U^k)$ satisfy

$$\big\{\tilde{W}, \tilde{\sigma}(0), \dots, \tilde{\sigma}(n-1)\big\} = \big\{W, \sigma(0), \dots, \sigma(n-1)\big\} \quad \text{in law.} \tag{3.9}$$

We can then consider the volatility corresponding to a fixed realization of $U^1, \dots, U^n$. Indeed, for $u = (u^1, \dots, u^n) \in (0,1)^{nd^2}$, let

$$\tilde{\sigma}(k; u) := \Theta_k\big(\tilde{W}|_{[0,t_k]}, u^1, \dots, u^k\big)$$

and consider $\tilde{M}^u = \int \tilde{\sigma}^u_t \, d\tilde{W}_t$, where $\tilde{\sigma}^u := \sum_{k=0}^{n-1} \mathbf{1}_{[t_k, t_{k+1})} \tilde{\sigma}(k; u)$. For any $F \in C_b(\Omega; \mathbb{R})$, the equality (3.9) and Fubini's theorem yield that

$$\begin{aligned} E[F(M)] = E\big[F\big(\tilde{M}^{(U^1,\dots,U^n)}\big)\big] &= \int_{(0,1)^{nd^2}} E[F(\tilde{M}^u)] \, du \\ &\leq \sup_{u \in (0,1)^{nd^2}} E[F(\tilde{M}^u)]. \end{aligned}$$

Hence, by the Hahn-Banach theorem, the law of M is contained in the weak closure of the convex hull of the laws of $\{\tilde{M}^u : u \in (0,1)^{nd^2}\}$. We note that $\tilde{M}^u$ is of the form $\tilde{M}^u = \int g(t, \tilde{W}) \, d\tilde{W}_t$ with a measurable, adapted, $\sqrt{\mathbf{D}}$-valued function g, for each fixed u.

4. Smoothing. As $\mathcal{Q}_D$ is defined through continuous functions, it remains to approximate g by a continuous function f. Let $g : [0,T] \times \Omega \to \sqrt{\mathbf{D}}$ be a measurable adapted function and $\delta > 0$. By standard density arguments there exists $\tilde{f} \in C\big([0,T] \times \Omega; \mathbb{S}^d\big)$ such that

$$E \int_0^T \|\tilde{f}(t, \tilde{W}) - g(t, \tilde{W})\|^2 \, dt \leq \delta.$$

Let $f(t,x) := \sqrt{\Pi_{\mathbf{D}}\big(\tilde{f}(t,x)^2\big)}$. Then $f \in C\big([0,T] \times \Omega; \sqrt{\mathbf{D}}\big)$ and

$$\|f - g\|^2 \leq \|f^2 - g^2\| \leq \|\tilde{f}^2 - g^2\| \leq (\|\tilde{f}\| + \|g\|)\|\tilde{f} - g\| \leq 2\sqrt{R_{\mathbf{D}}} \, \|\tilde{f} - g\|$$

(see [1, Theorem X.1.1] for the first inequality). By Jensen's inequality we conclude that $E \int_0^T \|f(t, \tilde{W}) - g(t, \tilde{W})\|^2 \, dt \leq 2\sqrt{T R_{\mathbf{D}} \delta}$, which, in view of the above steps, completes the proof. $\square$

References

[1] R. Bhatia. *Matrix Analysis.* Springer, New York, 1997.

[2] P. Cheridito, H. M. Soner, N. Touzi, and N. Victoir. Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs. *Comm. Pure Appl. Math.*, 60(7):1081–1110, 2007.

[3] L. Denis, M. Hu, and S. Peng. Function spaces and capacity related to a sublinear expectation: application to G-Brownian motion paths. *Potential Anal.*, 34(2):139–161, 2011.

[4] 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.

[5] Y. Dolinsky and H. M. Soner. Limit theorems for binomial markets with friction. *Preprint*, 2011.

[6] D. Duffie and P. Protter. From discrete- to continuous-time finance: weak convergence of the financial gain process. *Math. Finance*, 2(1):1–15, 1992.

[7] H. He. Convergence from discrete- to continuous-time contingent claims prices. *Rev. Financ. Stud.*, 3(4):523–546, 1990.

[8] S. Kusuoka. Limit theorem on option replication cost with transaction costs. *Ann. Appl. Probab.*, 5(1):198–221, 1995.

[9] M. Nutz. Random G-expectations. *Preprint arXiv:1009.2168v1*, 2010.

[10] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems Control Lett.*, 14(1):55–61, 1990.

[11] S. Peng. Backward SDE and related g-expectation. In *Backward stochastic differential equations*, volume 364 of *Pitman Res. Notes Math. Ser.*, pages 141–159. Longman, 1997.

[12] S. Peng. G-expectation, G-Brownian motion and related stochastic calculus of Itô type. In *Stochastic Analysis and Applications*, volume 2 of *Abel Symp.*, pages 541–567, Springer, Berlin, 2007.

[13] S. Peng. Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. *Stochastic Process. Appl.*, 118(12):2223–2253, 2008.

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

[15] S. Peng. Tightness, weak compactness of nonlinear expectations and application to CLT. *Preprint arXiv:1006.2541v1*, 2010.

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

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