Electronic Journal of Probability

Electron. J. Probab. **17** (2012), no. 98, 1–15.
ISSN: 1083-6489 DOI: 10.1214/EJP.v17-2284

Numerical schemes for G–Expectations

Yan Dolinsky*

Abstract

We consider a discrete time analog of G–expectations and we prove that in the case where the time step goes to zero the corresponding values converge to the original G–expectation. Furthermore we provide error estimates for the convergence rate. This paper is continuation of Dolinsky, Nutz, and Soner (2012). Our main tool is a strong approximation theorem which we derive for general discrete time martingales.

Keywords: G-expectations; volatility uncertainty; strong approximation theorems.
AMS MSC 2010: 60F15; 60G44; 91B24.
Submitted to EJP on October 8, 2011, final version accepted on August 25, 2012.

1 Introduction

In this paper we study numerical schemes for G–expectations, which were introduced recently by Peng (see [7] and [8]). A G–expectation is a sublinear function which maps random variables on the canonical space $\Omega := C([0,T];\mathbb{R}^d)$ to the real numbers. The motivation to study G–expectations comes from mathematical finance, and in particular from risk measures (see [6] and [9]) and pricing under volatility uncertainty (see [2] ,[6] and [12]).

Our starting point is the dual view on G–expectation via volatility uncertainty (see [1]), which yields the representation $\xi \to \sup_{P\in\mathcal{P}} E_P[\xi]$ where $\mathcal{P}$ is the set of probabilities on $C([0,T];\mathbb{R}^d)$ 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 a compact convex subset $\mathbf{D}\subset\mathbb{S}^d_+$ of positive definite matrices. Thus the set $\mathbf{D}$ can be understood as the domain of (Knightian) volatility uncertainty and the functional above represents the European option (with reward ξ) super–hedging price. For details see [2] and [6].

In the current work we assume that ξ is of the form $F(B,\langle B\rangle)$ where F is a path–dependent functional which satisfies some regularity conditions. In particular, ξ can represent an award of a path dependent European contingent claim. In this case the reward is a functional of the stock price, which is equal to the Doolean exponential of the canonical process, and so quadratic variation appears naturally.

In [4] the authors introduced a volatility uncertainty in discrete time and an analog of the Peng G–expectation. They proved that the discrete time values converge to

*ETH Zurich, Switzerland; and Hebrew University of Jerusalem, Israel.
E-mail: yan.dolinsky@math.ethz.ch; and E-mail: yan.yolinsky@mail.huji.ac.il

the continuous time G–expectation. The main tools that were used there are the weak convergence machinery together with a randomization technique. The main disadvantage of the weak convergence approach is that it cannot provide error estimates. In order to obtain error estimates we should consider all the market models on the same probability space, and so methods based on strong approximation theorems come into picture.

In this paper we consider a different (from the one in [4]) discrete time analog of G–expectation and prove that in a case where the time step goes to zero the corresponding values converge to the original G–expectation. In the current scheme, the discrete time martingales that we consider have an explicit representation, and so we can write a dynamical programming for the discrete time analog of the G–expectation and for the corresponding optimal control. For the scheme that was introduced in [4] dynamical programming is not available. Furthermore, by deriving a strong invariance principle for general discrete time martingales, we are able to provide error estimates for the convergence rate of the current scheme.

The paper is organized as following. In the next section we introduce the setup and formulate the main results. In Section 3 we present the main machinery which we use, namely we obtain a strong approximation theorem for general martingales. In Section 4 we derive auxiliary lemmas that are used in the proof of the main results. In Section 5 we complete the proof of Theorems 2.3 and 2.5, and remark why our estimates are also valid for the approximations which were considered in [4].

2 Preliminaries and main results

We fix the dimension $d \in \mathbb{N}$ and denote by $||\cdot||$ the sup 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. Consider the space $\mathbb{S}^d$ with the operator norm $||A|| := \sup_{||v||=1} ||A(v)||$. We fix a nonempty, convex and compact set $\mathbf{D} \subset \mathbb{S}^d_+$; the elements of $\mathbf{D}$ will be the possible values of our volatility process. Denote by $\Omega = C([0,T];\mathbb{R}^d)$ and $\Gamma = C([0,T];\mathbb{S}^d)$, the spaces of continuous functions with values in $\mathbb{R}^d$ and $\mathbb{S}^d$, respectively. We consider these spaces with the sup norm $||x|| := \sup_{0 \le t \le T} ||x_t||$. Let $F : \Omega \times \Gamma \to \mathbb{R}$ be a function which satisfies the following assumption. There exist constants $H_1, H_2 > 0$ such that

$$|F(u_1,v_1) - F(u_2,v_2)| \le H_1 \exp\left(H_2(||u_1|| + ||u_2|| + ||v_1|| + ||v_2||)\right) \times \\ (||u_1 - u_2|| + ||v_1 - v_2||), \quad u_1,u_2 \in \Omega, \quad v_1,v_2 \in \Gamma. \tag{2.1}$$

Without loss of generality we assume that the maturity date $T = 1$. We denote by $B = (B_t)_{0 \le t \le 1}$ the canonical process (on the space Ω) $B_t(\omega) := \omega_t$, $\omega \in \Omega$ 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 (with respect to the filtration $\mathcal{F}$) and $B_0 = 0$ P-a.s. (all our martingales start at the origin). We set

$$\mathcal{P}_{\mathbf{D}} := \{P \text{ martingale law on } \Omega : \ d\langle B \rangle / dt \in \mathbf{D}, \ P \times dt \text{ a.s.}\}, \tag{2.2}$$

observe that under any measure $P \in \mathcal{P}_D$ the stochastic processes B and $\langle B \rangle$, are random elements in Ω and Γ, respectively. Consider the G–expectation

$$V := \sup_{P \in \mathcal{P}_{\mathbf{D}}} E_P F(B, \langle B \rangle) \tag{2.3}$$

where E_P denotes the expectation with respect to P. A measure $P \in \mathcal{P}_{\mathbf{D}}$ will be called ϵ–optimal if

$$V < \epsilon + E_P F(B, \langle B \rangle). \tag{2.4}$$

Our goal is to find discrete time approximations for V. The advantage of discrete time approximations is that the corresponding values can be calculated by dynamical programming. Furthermore, we will apply these approximations in order to find ϵ–optimal measures in the continuous time setting.

Remark 2.1. *Let $S = \{(S^1_t, ..., S^d_t)\}_{t=0}^1$ be the Doolean's exponential $\mathcal{E}(B)$ of the canonical process B, namely $S^i_t := S^i_0 \exp\left(B^i_t - \langle B^i\rangle_t\right)$, $i \le d$, $t \in [0,1]$. The stochastic process S represents the stock prices in a financial model with volatility uncertainty. Clearly any random variable of the from $g(S)$ where $g : C([0,T];\mathbb{R}^d) \to \mathbb{R}_+$ is a Lipschitz continuous function, can be written in the form $g(S) = F(B, \langle B\rangle)$ for a suitable F which satisfies (2.1). Thus we see that our setup includes payoffs which correspond to path dependent European options.*

Remark 2.2. *In general it can be shown (see [1]) that there is no loss of generality in assuming that D is convex. Namely, any G–expectation can be represented by the right–hand side of (2.3) for a compact convex set $\boldsymbol{D}$. The proof in [1] relied on PDE technique, however let us briefly explain the probabilistic intuition behind the result. Assume that we start with a compact (not necessarily convex) set $\boldsymbol{D}$. Let $\hat{\mathbf{D}}$ be the convex hull of $\boldsymbol{D}$. Clearly, $\hat{\mathbf{D}}$ is a compact convex set. It can be shown that the set of probability measures $\mathcal{P}_{\hat{\mathbf{D}}}$ is the closure (with respect to weak convergence) of the convex hull of $\mathcal{P}_{\boldsymbol{D}}$. This together with the regularity of $F(B, \langle B\rangle)$ yields*

$$\sup_{P \in \mathcal{P}_{\boldsymbol{D}}} E_P F(B, \langle B\rangle) = \sup_{P \in \mathcal{P}_{\hat{\mathbf{D}}}} E_P F(B, \langle B\rangle).$$

Let us emphasis that in the proof of our main results we will use the fact that the set $\boldsymbol{D}$ is convex and compact.

Next, we formulate the main approximation results. Let ν be a distribution on $\mathbb{R}^d$ which satisfies the following

$$\int_{\mathbb{R}^d} x d\nu(x) = 0 \ \text{ and } \ \int_{\mathbb{R}^d} x^i x^j d\nu(x) = \delta_{ij}, \ \ 1 \le i,j \le d \tag{2.5}$$

where δ_{ij} is the Kronecker–Delta. Furthermore, we assume that the moment generating function $\psi_\nu(y) := \int_{x\in\mathbb{R}^d} \exp(\sum_{i=1}^d x^i y^i) d\nu(x) < \infty$ exists for any $y \in \mathbb{R}^d$, and for any compact set $K \subset \mathbb{R}^d$

$$\sup_{n\in\mathbb{N}} \sup_{y\in K} \psi^n_\nu\left(\frac{y}{\sqrt{n}}\right) < \infty. \tag{2.6}$$

Observe that the standard d–dimensional normal distribution $\nu = N(0, I)$ is satisfying the assumptions (2.5)–(2.6).

Let $n \in \mathbb{N}$ and $Y_1, ..., Y_n$ be a sequence of i.i.d. random vectors with $\mathcal{L}(Y_1) = \nu$, i.e., the distribution of the random vectors is ν. We denote by $\mathcal{A}^\nu_n$ the set of all d–dimensional stochastic process $M = (M_0, ..., M_n)$ of the form, $M_0 := 0$ and

$$M_i := \sum_{j=1}^{i} \frac{1}{\sqrt{n}} \phi_j(Y_1, ..., Y_{j-1}) Y_j, \ \ 1 \le i \le n \tag{2.7}$$

where $\phi_j : (\mathbb{R}^d)^{j-1} \to \sqrt{\mathbf{D}} := \{\sqrt{A} : A \in \mathbf{D}\}$ are measurable functions (which can be chosen arbitrary) and $Y_1, ..., Y_n$ are column vectors. As usual for a matrix $A \in \mathbb{S}^d_+$ we denote by $\sqrt{A}$ the unique square root in $\mathbb{S}^d_+$. Observe that M is a martingale under the filtration which is generated by $Y_1, ..., Y_n$. Let $\langle M\rangle$ be the ($\mathbb{S}^d_+$ valued) predictable variation of M. In view of (2.5) we get

$$\langle M\rangle_k = \frac{1}{n} \sum_{j=1}^{k} \phi^2_j(Y_1, ..., Y_{j-1}), \ \ 1 \le k \le n \tag{2.8}$$

and we set $\langle M\rangle_0 = 0$. Let $\mathcal{W}_n : (\mathbb{R}^d)^{n+1} \times (\mathbb{S}^d)^{n+1} \to \Omega \times \Gamma$ be the linear interpolation operator given by

$$\mathcal{W}_n(u,v)(t) := ([nt]+1-nt)\left(u_{[nt]}, v_{[nt]}\right) + (nt-[nt])\left(u_{[nt]+1}, v_{[nt]+1}\right),\ t \in [0,1]$$

where $u = (u_0, u_1, ..., u_n)$, $v = (v_0, v_1, ..., v_n)$ and $[z]$ denotes the integer part of z. Namely, we consider a continuous time polygonal curve which is defined by the vectors u, v. This operator gives a natural way to map discrete time processes such as M and $\langle M\rangle$, to continuous time processes, and this is essential since the function F is defined on $\Omega \times \Gamma$. Set

$$V_n^\nu := \sup_{M \in \mathcal{A}_n^\nu} \mathbb{E} F\left(\mathcal{W}_n(M, \langle M\rangle)\right), \tag{2.9}$$

we denote by $\mathbb{E}$ the expectation with respect to the underlying probability measure.

The following theorem which will be proved in Section 5 is the main result of the paper.

Theorem 2.3. *For any $\epsilon > 0$ there exists a constant $\mathcal{C}_\epsilon = \mathcal{C}_\epsilon(\nu)$ which depends only on the distribution ν such that*

$$|V_n^\nu - V| \le \mathcal{C}_\epsilon n^{\epsilon - 1/8},\ \ \forall n \in \mathbb{N}. \tag{2.10}$$

Furthermore, if the function F is bounded, then there exists a constant $\mathcal{C} = \mathcal{C}(\nu)$ for which

$$|V_n^\nu - V| \le \mathcal{C} n^{-1/8},\ \ \forall n \in \mathbb{N}. \tag{2.11}$$

Next, we describe a dynamical programming algorithm for V_n^ν and for the corresponding optimal control, which in general should not be unique. For the later we will need the following definition.

Definition 2.4. *Let $(\mathcal{X}, \mathcal{B})$ be a measurable space, $m \in \mathbb{N}$ and let $K \subset \mathbb{R}^m$ be a compact set. Consider a measurable map $\mathbb{H} : \mathcal{X} \times K \to \mathbb{R}$ such that $\mathbb{H}(x, \cdot) : K \to \mathbb{R}$ is a continuous function for any $x \in \mathcal{X}$. Define $arg\max_{y \in K} \mathbb{H}(x,y) : \mathcal{X} \to K$ by*

$$arg\max_{y \in K} \mathbb{H}(x,y) := \max\{z \in K | \mathbb{H}(x,z) = \hat{\mathbb{H}}(x)\}$$

where $\hat{\mathbb{H}}(x) := \sup_{u \in K} \mathbb{H}(x,u)$ and the maximum in the above right–hand side is taken with respect to the lexicographical order $\succ$ on $\mathbb{R}^m$. Namely, for any $u, v \in \mathbb{R}^m$ we define $u \succ v$ if there exists $k > 0$ such that $u_i = v_i$ for $i < k$ and $u_k > v_k$. From the continuity of $\mathbb{H}$ in the second variable we get that $\hat{\mathbb{H}} : \mathcal{X} \to \mathbb{R}$ is a measurable function and the set $\{z \in K | \mathbb{H}(x,z) = \hat{\mathbb{H}}(x)\}$ is a non empty compact set. Thus, the above map is well defined. Next, let us briefly verify that the defined map is measurable. Let $arg\max_{y \in K} \mathbb{H}(x,y) := (\tilde{\mathbb{H}}_1(x), ..., \tilde{\mathbb{H}}_m(x))$. By applying the continuity of $\mathbb{H}$ in the second variable it follows that

$$\{x \in \mathcal{X} | \tilde{\mathbb{H}}_1(x) \ge a\} = \left\{x \in \mathcal{X} | \max_{z \in K \cap [a,\infty) \times \mathbb{R}^{m-1}} \mathbb{H}(x,z) = \hat{\mathbb{H}}(x)\right\},$$

since $a \in \mathbb{R}$ was arbitrary, we conclude that $\tilde{\mathbb{H}}_1$ is a measurable function. Finally, to complete the argument, we assume by induction that $\tilde{\mathbb{H}}_1, ..., \tilde{\mathbb{H}}_k$ are measurable functions, and we show that $\tilde{\mathbb{H}}_{k+1}$ is measurable. Let $\mathcal{J} \subset K$ be a countable set that is dense in K. Choose $a \in \mathbb{R}$. For any $n \in \mathbb{N}$ and $u = (u_1, ..., u_m) \in \mathcal{J}$ define the function $\mathbb{H}_{u,n} : \mathcal{X} \to \mathbb{R}$ by

$$\begin{aligned}\mathbb{H}_{u,n}(x) := \mathbb{H}(x,u)\mathbb{I}_{\sum_{i=1}^k |u_i - \tilde{\mathbb{H}}_i(x)| + \max(0, a - u_{i+1}) < 1/n} + \\ (\hat{\mathbb{H}}(x) - 1)\mathbb{I}_{\sum_{i=1}^k |u_i - \tilde{\mathbb{H}}_i(x)| + \max(0, a - u_{i+1}) \ge 1/n},\ \ x \in \mathcal{X}\end{aligned}$$

where for any event $\mathbb{A}$ we set $\mathbb{I}_{\mathbb{A}} = 1$ if an event $\mathbb{A}$ occurs and $\mathbb{I}_{\mathbb{A}} = 0$ if not. From the induction assumption it follows that for any $n \in \mathbb{N}$ and $u \in J$, the function $\mathbb{H}_{u,n} : \mathcal{X} \to \mathbb{R}$ is measurable. From the continuity of $\mathbb{H}$ in the second variable we get

$$\{x \in \mathcal{X} | \tilde{\mathbb{H}}_{k+1}(x) \geq a\} = \{x \in \mathcal{X} | \inf_{n \in \mathbb{N}} \sup_{u \in \mathcal{J}} \mathbb{H}_{u,n}(x) = \hat{\mathbb{H}}(x)\}.$$

Thus $\{x \in \mathcal{X} | \tilde{\mathbb{H}}_{k+1}(x) \geq a\}$ is a measurable set, and the argument is completed.

Now we are ready to introduce the dynamical programming for V_n^ν and for the corresponding optimal control. Fix $n \in \mathbb{N}$ and define a sequence of functions $J_k^{\nu,n} : (\mathbb{R}^d)^{k+1} \times (\mathbb{S}^d)^{k+1} \to \mathbb{R}$, $k = 0, 1, ..., n$ by the backward recursion

$$\begin{gathered} J_n^{\nu,n}(u_0, u_1, ..., u_n, v_0, v_1, ..., v_n) := F\left(\mathcal{W}_n(u, v)\right) \quad \text{and} \\ J_k^{\nu,n}(u_0, u_1, ..., u_k, v_0, v_1, ..., v_k) := \\ \sup_{\gamma \in \sqrt{\mathbf{D}}} \mathbb{E}\left(J_{k+1}^{\nu,n}\left(u_0, u_1, ..., u_k, u_k + \tfrac{\gamma Y_{k+1}}{\sqrt{n}}, v_0, v_1, ..., v_k, v_k + \tfrac{\gamma^2}{n}\right)\right) = \\ \sup_{\gamma \in \sqrt{\mathbf{D}}} \int_{\mathbb{R}^d} J_{k+1}^{\nu,n}\left(u_0, u_1, ..., u_k, u_k + \tfrac{\gamma x}{\sqrt{n}}, v_0, v_1, ..., v_k, v_k + \tfrac{\gamma^2}{n}\right) d\nu(x) \\ \text{for} \quad k = 0, 1, ..., n-1. \end{gathered} \tag{2.12}$$

From (2.1) and (2.6) it follows that there exists a constant $\hat{H}$ such that

$$J_k^{\nu,n}\left(u_0, ..., u_k, v_0, ..., v_k\right) \leq \hat{H} \exp\left((H_2 + 1) \sum_{i=0}^{k} (||u_i|| + ||v_i||)\right), \quad \forall k, u_0, ..., u_k, v_0, ..., v_k.$$

Fix k. By applying (2.6) again we conclude that for any compact sets $K_1 \subset \mathbb{R}^d$ and $K_2 \subset \mathbb{S}_+^d$, the family of random variables

$$\begin{gathered} J_{k+1}^{\nu,n}\left(u_0, ..., u_k, u_k + \tfrac{\gamma Y_{k+1}}{\sqrt{n}}, v_0, ..., v_k, v_k + \tfrac{\gamma^2}{n}\right), \\ \gamma \in \sqrt{\mathbf{D}}, \quad u_0, ..., u_k \in K_1, \quad v_0, ..., v_k \in K_2 \end{gathered}$$

is uniformly integrable. This together with the fact that the set $\mathbf{D}$ is compact gives (by backward induction) that for any k, the function $J_k^{\nu,n}$ is continuous. Thus (following Definition 2.4) we introduce the measurable functions $h_k^{\nu,n} : (\mathbb{R}^d)^{k+1} \times (\mathbb{S}^d)^{k+1} \to \sqrt{\mathbf{D}}$, $k = 0, 1, ..., n-1$ by

$$\begin{gathered} h_k^{\nu,n}(u_0, ..., u_k, v_0, ..., v_k) := \\ arg \max_{\gamma \in \sqrt{\mathbf{D}}} \int_{\mathbb{R}^d} J_{k+1}^{\nu,n}\left(u_0, u_1, ..., u_k, u_k + \tfrac{\gamma x}{\sqrt{n}}, v_0, v_1, ..., v_k, v_k + \tfrac{\gamma^2}{n}\right) d\nu(x). \end{gathered} \tag{2.13}$$

Finally, define by induction the stochastic processes $\{M_k^{\nu,n}\}_{k=0}^n$ and $\{N_k^{\nu,n}\}_{k=0}^n$, with values in $\mathbb{R}^d$ and $\mathbb{S}^d$, respectively by $M_0^{\nu,n} := 0$, $N_0^{\nu,n} := 0$ and for $k < n$

$$\begin{gathered} N_{k+1}^{\nu,n} := N_k^{\nu,n} + \tfrac{1}{n}\left(h_k^{\nu,n}(M_0^{\nu,n}, ..., M_k^{\nu,n}, N_0^{\nu,n}, ..., N_k^{\nu,n})\right)^2 \\ \text{and} \quad M_{k+1}^{\nu,n} := M_k^{\nu,n} + \tfrac{1}{\sqrt{n}} h_k^{\nu,n}(M_0^{\nu,n}, ..., M_k^{\nu,n}, N_0^{\nu,n}, ..., N_k^{\nu,n}) Y_{k+1}. \end{gathered} \tag{2.14}$$

Observe that $M^{\nu,n} \in \mathcal{A}_n^\nu$ and $N^{\nu,n} = \langle M^{\nu,n} \rangle$. From the dynamical programming principle it follows that

$$V_n^\nu = J_0^{\nu,n}(0,0) = \mathbb{E} F\left(\mathcal{W}_n(M^{\nu,n}, \langle M^{\nu,n} \rangle)\right). \tag{2.15}$$

In the following theorem (which will be proved in Section 5) we provide an explicit construction of ϵ–optimal measures for the G–expectation which is defined in (2.3).

Theorem 2.5. *Let $(\Omega^W, \mathcal{F}^W, \mathbb{P}^W)$ be a complete probability space together with a standard d-dimensional Brownian motion $\{W_t\}_{t\in[0,1]}$ and its natural filtration $\mathcal{F}^W_t := \sigma\{W(s)|s \le t\}$. Consider the standard normal distribution $\nu_g = \mathcal{N}(0, I)$. For any $n \in \mathbb{N}$, let $f_n : (\mathbb{R}^d)^n \to \mathbb{R}$ be a function which is satisfying $f_n(Y^g_1, ..., Y^g_n) = M^{\nu_g,n}_n$, where $Y^g_1, ..., Y^g_n$ are i.i.d. and $\mathcal{L}(Y^g_1) = \nu_g$. Observe that f_n can be calculated from (2.12)–(2.14). Define the stochastic process $\{M^n_t\}^1_{t=0}$ by*

$$M^n_t := \mathbb{E}^W \left(f_n \left(\sqrt{n} W_{\frac{1}{n}}, \sqrt{n}(W_{\frac{2}{n}} - W_{\frac{1}{n}}), ..., \sqrt{n}(W_1 - W_{\frac{n-1}{n}}) \right) |\mathcal{F}^W_t \right), \quad t \in [0, 1] \tag{2.16}$$

where $\mathbb{E}^W$ denotes the expectation with respect to $\mathbb{P}^W$. Notice that since M^n is a martingale with respect to Brownian filtration, its a continuous stochastic process. Thus, let P_n be the distribution of M^n on the canonical space Ω. Then $P_n \in \mathcal{P}_D$, and for any $\epsilon > 0$ there exists a constant $\tilde{\mathcal{C}}_\epsilon$ such that

$$V < E_n F(B, \langle B\rangle) + \tilde{\mathcal{C}}_\epsilon n^{\epsilon - 1/8}, \quad \forall n \tag{2.17}$$

where E_n denotes the expectation with respect to P_n. If the function F is bounded then there exists a constant $\tilde{\mathcal{C}}$ for which

$$V < E_n F(B, \langle B\rangle) + \tilde{\mathcal{C}} n^{-1/8}, \quad \forall n. \tag{2.18}$$

Remark 2.6. *One question which remains open is whether the estimates of Theorem 2.3 are sharp. Another unanswered question is which distribution ν provides the best estimates in Theorem 2.3. In this paper we do not study the dependence of the constants which appear in (2.10)–(2.11), as functions of the distribution ν. The Gaussian distribution $\nu_g = \mathcal{N}(0, I)$ is the most natural for establishing Theorem 2.5. In fact, we will use the Gaussian distribution and the martingale representation theorem on the Wiener space in order to establish Theorem 2.5 and one side of the estimates in (2.10)–(2.11). Another interesting choice of the distribution ν is a purely atomic distribution which is supported by $d+1$ points (for the exact construction see Remark 5.1). It seems that the later choice of ν is the most efficient for the implementation of the dynamical programming which is give by (2.12). Another application of this choice is explained in Remark 5.1.*

3 The main tool

In this section we derive a strong approximation theorem (Lemma 3.2) which is the main tool in the proof of Theorems 2.3 and 2.5. This theorem is an extension of the main result in [11].

For any two distributions ν_1, ν_2 on the same measurable space $(\mathcal{X}, \mathcal{B})$ we define the distance in variation

$$\rho(\nu_1, \nu_2) := \sup_{B\in\mathcal{B}} |\nu_1(B) - \nu_2(B)|. \tag{3.1}$$

First we state some results (without a proof) from [11] (Lemmas 4.5 and 7.2 in [11]) that will be used in the proof of Lemma 3.2.

Lemma 3.1.
i. There exists a distribution μ on $\mathbb{R}^d$ which is supported on the set $(-1/2, 1/2)^d$ and has the following property. There exists a constant $\mathcal{C}_1 > 0$ such that for any distributions ν_1, ν_2 on $\mathbb{R}^d$ which satisfy

$$\begin{gathered} \int_{\mathbb{R}^d} x d\nu_1(x) = \int_{\mathbb{R}^d} x d\nu_2(x) \ \text{ and for } \ 1 \le i, j \le d \\ \int_{\mathbb{R}^d} x^i x^j d\nu_1(x) = \int_{\mathbb{R}^d} x^i x^j d\nu_2(x) \end{gathered} \tag{3.2}$$

we have

$$\rho(\nu_1 * \mu, \nu_2 * \mu) \leq \mathcal{C}_1 \left(\int_{\mathbb{R}^d} ||x||^3 d\nu_1(x) + \int_{\mathbb{R}^d} ||x||^3 d\nu_2(x) \right) \tag{3.3}$$

where $\nu * \mu$ denotes the convolution of the measures ν and μ.
ii. Let $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ be a probability space together with a d–dimensional random vector Y, a m–dimensional random vector Z (m is some natural number), and a random variable α which is distributed uniformly on the interval $[0,1]$ and independent of Y and Z. Let ν be a distribution on $\mathbb{R}^d$ and let $\hat{\nu}$ be a distribution on $\mathbb{R}^m \times \mathbb{R}^d$ such that $\hat{\nu}(A \times \mathbb{R}^d) = \tilde{P}(Z \in A)$ for any $A \in \mathcal{B}(\mathbb{R}^m)$, i.e. the marginal distribution of $\hat{\nu}$ on $\mathbb{R}^m$ equals to $\mathcal{L}(Z)$. There exists a measurable function $\Psi = \Psi_{\nu,\hat{\nu},\mathcal{L}(Z,Y)} : \mathbb{R}^m \times \mathbb{R}^d \times [0,1] \to \mathbb{R}^d \times \mathbb{R}^d$ such that for the vector

$$(U, X) := \Psi(Z, Y, \alpha) \tag{3.4}$$

we have the following: $\mathcal{L}(U) = \nu$, $\mathcal{L}(Z, X) = \hat{\nu}$, U is independent of X, Z and

$$\tilde{P}(U + X \neq Y | Z) = \rho(\mathcal{L}(U) * \mathcal{L}(X|Z), \mathcal{L}(Y|Z)). \tag{3.5}$$

Now we are ready to prove the main result of this section. For any stochastic process $Z = \{Z_k\}_{k=0}^n$ we denote $\Delta Z_k := Z_k - Z_{k-1}$ for $k \geq 0$, where we set $Z_{-1} = 0$. Fix $n \in \mathbb{N}$ and consider a d–dimensional martingale $\{M_k\}_{k=0}^n$ with respect to its natural filtration, which satisfies $M_0 = 0$. For any $1 \leq k \leq n$, there exists a measurable function $\hat{\phi}_k : (\mathbb{R}^d)^k \to \mathbb{S}^d$ such that

$$\begin{aligned} \sqrt{\Delta\langle M \rangle_k} = \sqrt{\mathbb{E}\left(\Delta M_k \Delta M_k' \middle| \sigma\{M_0, M_1, ..., M_{k-1}\}\right)} := \\ \hat{\phi}_k(\Delta M_0, \Delta M_1, ..., \Delta M_{k-1}), \end{aligned} \tag{3.6}$$

where $\{\langle M \rangle_k\}_{k=0}^n$ is the predictable variation ($\mathbb{S}_+^d$ valued) of M and the symbol $\cdot'$ denotes transposition. We assume that there exists a constant H for which

$$\mathbb{E}\left(||\Delta M_k||^3 \middle| \sigma\{M_0, ..., M_{k-1}\}\right) + ||\sqrt{\Delta\langle M \rangle_k}||^3 \leq H, \quad a.s. \ \forall k. \tag{3.7}$$

Lemma 3.2. *Let ν a distribution on $\mathbb{R}^d$ such that*

$$\begin{gathered} \textstyle\int_{\mathbb{R}^d} x d\nu(x) = 0, \ \ \int_{\mathbb{R}^d} x^i x^j d\nu(x) = \delta_{ij} \ \ \forall i, j \leq d \\ \text{and} \ \ \textstyle\int_{\mathbb{R}^d} ||x||^3 d\nu(x) < \infty. \end{gathered} \tag{3.8}$$

For any $\Theta > 0$ its possible to reconstruct the martingale $\{M_k\}_{k=0}^n$ on some probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ (namely we construct a martingale which has the same distribution as the original martingale M, and for simplicity we denote the new martingale also by M) together with a sequence of i.i.d. random vectors $Y_1, ..., Y_n$ with the following properties:
i. $\mathcal{L}(Y_1) = \nu$.
ii. For any k, the random vectors $M_1, ..., M_{k-1}$ are independent of Y_k.
iii. There exists a constant $\mathcal{C}_2 = \mathcal{C}_2(\nu)$ which depends only on the distribution ν such that

$$\tilde{P}\left(\max_{1 \leq k \leq n} ||M_k - \sum_{j=1}^k \sqrt{\Delta\langle M \rangle_j} Y_j|| > \Theta \right) \leq \frac{\mathcal{C}_2 H n}{\Theta^3}. \tag{3.9}$$

Proof. Fix $\Theta > 0$. For any k let ν_k be the distribution of the random vector $\frac{1}{\Theta}(\Delta M_0, ..., \Delta M_k)$. Let $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ be a probability space which contains a sequence of i.i.d. random vectors $Y_1, ..., Y_n$ such that $\mathcal{L}(Y_1) = \nu$, a sequence of i.i.d. random variables $\alpha_1, ..., \alpha_n$ which are distributed uniformly on the interval $[0,1]$ and independent of $Y_1, ..., Y_n$, and a random vector U_0 which is independent of $Y_1, ..., Y_n, \alpha_1, ..., \alpha_n$ and satisfies $\mathcal{L}(U_0) = \mu$, where the distribution μ is defined in the first part of Lemma 3.1. Define the sequences $\{X_i\}_{i=0}^n$ and $\{U_i\}_{i=1}^n$ by the following recursive relations, $X_0 = 0$ and

$$(U_k, X_k) := \Psi_{\mu,\nu_k,\hat{\nu}_k}(X_0, ..., X_{k-1}, U_{k-1} + \frac{1}{\Theta}\hat{\phi}_k(\Theta X_0, ..., \Theta X_{k-1})Y_k, \alpha_k), \ \ 1 \leq k \leq n \quad (3.10)$$

where $\hat{\nu}_k$ is the distribution of $(X_0, ..., X_{k-1}, U_{k-1} + \frac{1}{\Theta}\hat{\phi}_k(\Theta X_0, ..., \Theta X_{k-1})Y_k)$. Recall, that the function Ψ was introduced before (3.4) and the functions $\hat{\phi}_k$, $k < n$ were introduced before (3.6). From the definition of the map Ψ it follows (by induction) that $\mathcal{L}(\Theta X_0, ..., \Theta X_n) = \mathcal{L}(\Delta M_0, ..., \Delta M_n)$. We conclude that the stochastic process $\Theta \sum_{i=0}^k X_i$, $0 \leq k \leq n$ is distributed as $\{M_k\}_{k=0}^n$, and so we set (on the current probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$),

$$M_k := \Theta \sum_{i=0}^k X_i, \ \ 0 \leq k \leq n. \quad (3.11)$$

Let $1 \leq k \leq n$. From (3.10)–(3.11) and the fact that Y_k is independent of $Y_1, ..., Y_{k-1}$, $\alpha_1, ..., \alpha_{k-1}$ it follows that Y_k is independent of $M_0, ..., M_{k-1}$. Thus in order to complete the proof, it remains to establish (3.9). Set

$$\begin{aligned} &\delta_k := U_k + X_k - U_{k-1} - \tfrac{1}{\Theta}\hat{\phi}_k(\Theta X_0, ..., \Theta X_{k-1})Y_k, \text{ and } \rho_k(x_0, ..., x_{k-1}) \quad (3.12) \\ &:= \tilde{P}(\delta_k \neq 0 | X_0 = x_0, ..., X_{k-1} = x_{k-1}), \ \ x_0, ..., x_{k-1} \in \mathbb{R}^d \ \ 1 \leq k \leq n. \end{aligned}$$

From the properties of the map Ψ it follows that for any k, U_k is independent of $X_0, ..., X_k$ and $\mathcal{L}(U_k) = \mu$. This together with (3.5) and (3.10) gives

$$\begin{aligned} &\rho_k(x_0, ..., x_{k-1}) = \rho\big(\mathcal{L}(X_k | X_0 = x_0, ..., X_{k-1} = x_{k-1}) * \mu, \quad (3.13) \\ &\mathcal{L}(\tfrac{1}{\Theta}\phi(\Theta x_0, ..., \Theta x_{k-1})Y_k) * \mu\big) \ \ x_0, ..., x_{k-1} \in \mathbb{R}^d, \ \ 1 \leq k \leq n. \end{aligned}$$

From (3.3), (3.6)–(3.7), (3.11) and (3.13)

$$\rho_k(x_0, ..., x_{k-1}) \leq \frac{\mathcal{C}_2 H}{\Theta^3}, \ \ x_0, ..., x_{k-1} \in \mathbb{R}^d, \ \ 1 \leq k \leq n \quad (3.14)$$

for some constant $\mathcal{C}_2 = \mathcal{C}_2(\nu)$ which depends only on the distribution ν. From (3.11)–(3.12), (3.14) and the fact that $\max_{0 \leq k \leq n} ||U_k|| < \frac{1}{2}$ a.s. we obtain

$$\begin{aligned} &\tilde{P}\left(\max_{1 \leq k \leq n} ||M_k - \textstyle\sum_{j=1}^k \sqrt{\Delta \langle M \rangle_j} Y_j|| > \Theta\right) = \\ &\tilde{P}\left(\max_{1 \leq k \leq n} ||M_k - \textstyle\sum_{j=0}^{k-1} \hat{\phi}_j(\Delta M_0, ..., \Delta M_j) Y_{j+1}|| > \Theta\right) = \\ &\tilde{P}\left(\max_{1 \leq k \leq n} \Theta || \textstyle\sum_{i=1}^k \delta_i + U_0 - U_k|| > \Theta\right) \leq \textstyle\sum_{i=1}^n \tilde{P}(\delta_i \neq 0) \leq \frac{\mathcal{C}_2 H n}{\Theta^3} \end{aligned}$$

and we conclude the proof. □

4 Auxiliary lemmas

In this section we derive several estimates which are essential for the proof of Theorem 2.3 and 2.5. We start with the following general result.

Lemma 4.1. *Let $\{M_t\}_{t=0}^1$ be a one dimensional continuous martingale which satisfies $\frac{d\langle M \rangle_t}{dt} \leq \mathcal{H}$ a.s. for some constant $\mathcal{H}$. Consider the discrete time martingale $N_k := M_{k/n}$, $0 \leq k \leq n$ together with its predictable variation process $\{\langle N \rangle_k\}_{k=0}^n$ which is given by $\langle N \rangle_0 := 0$ and*

$$\langle N \rangle_k := \sum_{i=1}^k \mathbb{E}((\Delta N_i)^2 | \sigma\{N_0, ..., N_{i-1}\}), \ \ 1 \leq k \leq n.$$

There exists constants $\mathcal{C}_3, \mathcal{C}_4$ (which depend only on $\mathcal{H}$) such that

$$\mathbb{E}\left(\max_{0\le k\le n-1}\max_{k/n\le t\le (k+1)/n}|M_t-N_k|^4\right)\le \frac{\mathcal{C}_3}{n} \tag{4.1}$$

and

$$\mathbb{E}\left(\max_{0\le k\le n-1}\max_{k/n\le t\le (k+1)/n}|\langle M\rangle_t-\langle N\rangle_k|^2\right)\le \frac{\mathcal{C}_4}{\sqrt{n}}. \tag{4.2}$$

Proof. From the Burkholder–Davis–Gundy inequality it follows that there exists a constant c_1 such that

$$\begin{gathered}\mathbb{E}\left(\max_{0\le k\le n-1}\max_{k/n\le t\le (k+1)/n}|M_t-N_k|^4\right)\le \\ \sum_{k=0}^{n-1}\mathbb{E}\left(\max_{k/n\le t\le (k+1)/n}|M_t-M_{k/n}|^4\right)\le \\ c_1\sum_{k=0}^{n-1}\mathbb{E}\left(|\langle M\rangle_{(k+1)/n}-\langle M\rangle_{k/n}|^2\right)\le c_1 n\frac{\mathcal{H}^2}{n^2}=\frac{c_1\mathcal{H}^2}{n}\end{gathered} \tag{4.3}$$

this completes the proof of (4.1). Next, we prove (4.2). Define the optional variation of the martingale $\{N_k\}_{k=0}^n$ by $[N]_0:=0$ and

$$[N]_k:=\sum_{i=1}^k(\Delta N_i)^2,\ \ 1\le k\le n. \tag{4.4}$$

From the relation $\mathbb{E}(\Delta[N]_k|\sigma\{N_0,...,N_{k-1}\})=\Delta\langle N\rangle_k$ and the Doob–Kolmogorov inequality we obtain

$$\begin{gathered}\mathbb{E}\left(\max_{0\le k\le n}|[N]_k-\langle N\rangle_k|^2\right)\le 4\mathbb{E}\left(|[N]_n-\langle N\rangle_n|^2\right)= \\ 4\mathbb{E}\left(|\sum_{i=1}^n\Delta[N]_i-\Delta\langle N\rangle_i|^2\right)=4\sum_{i=1}^n\mathbb{E}\left(|\Delta[N]_i-\Delta\langle N\rangle_i|^2\right)\le \\ 4\sum_{i=1}^n\mathbb{E}((\Delta[N]_i)^2)=4\sum_{i=1}^n\mathbb{E}\left(|M_{i/n}-M_{(i-1)/n}|^4\right)\le\frac{4c_1\mathcal{H}^2}{n}\end{gathered} \tag{4.5}$$

where the last inequality follows from the Burkholder–Davis–Gundy inequality. Next, observe that

$$[N]_k=N_k^2-2\sum_{i=1}^{k-1}N_i(N_{i+1}-N_i)=N_k^2-2\int_0^{k/n}N_{[nt]}dM_t,\ \ 1\le k\le n. \tag{4.6}$$

From the Doob–Kolmogorov inequality and Ito's Isometry we get

$$\begin{gathered}\mathbb{E}\left(\sup_{0\le u\le 1}\left|\int_0^u(M_t-N_{[nt]})dM_t\right|^2\right)\le \\ 4\mathbb{E}\left(|\int_0^1(M_t-N_{[nt]})dM_t|^2\right)=4\mathbb{E}\left(\int_0^1(M_t-N_{[nt]})^2d\langle M\rangle_t\right)\le \\ 4\mathcal{H}\mathbb{E}\left(\max_{0\le k\le n-1}\max_{k/n\le t\le (k+1)/n}|M_t-N_k|^2\right)\le\frac{4\mathcal{H}\sqrt{\mathcal{C}_3}}{\sqrt{n}},\end{gathered} \tag{4.7}$$

the last inequality follows from (4.1) and Jensen's inequality. From (4.6)–(4.7) and the equality $2\int_0^{k/n}M_tdM_t=N_k^2-\langle M\rangle_{k/n}$ it follows that

$$\mathbb{E}\left(\max_{1\le k\le n}|[N]_k-\langle M\rangle_{k/n}|^2\right)\le\frac{16\mathcal{H}\sqrt{\mathcal{C}_3}}{\sqrt{n}}.$$

This together with (4.5) and the inequality $(a+b+c)^2\le 4(a^2+b^2+c^2)$ yields

$$\begin{gathered}\mathbb{E}\left(\max_{0\le k\le n-1}\max_{k/n\le t\le (k+1)/n}|\langle M\rangle_t-\langle N\rangle_k|^2\right)\le \\ \frac{4\mathcal{H}^2}{n^2}+4\mathbb{E}\left(\max_{1\le k\le n}|[N]_k-\langle M\rangle_{k/n}|^2\right)+4\mathbb{E}\left(\max_{1\le k\le n}|[N]_k-\langle N\rangle_k|^2\right) \\ \le\frac{4\mathcal{H}^2}{n^2}+\frac{64\mathcal{H}\sqrt{\mathcal{C}_3}}{\sqrt{n}}+\frac{16c_1\mathcal{H}^2}{n}\end{gathered} \tag{4.8}$$

and the proof is completed. □

Next, we apply the above lemma in order to derive some estimates in our setup.

Lemma 4.2. *Let $n \in \mathbb{N}$ and $P \in \mathcal{P}_{\mathbf{D}}$. Consider the d–dimensional martingale $N_k := B_{k/n}$, $0 \leq k \leq n$ together with its predictable variation $\{\langle N \rangle_k\}_{k=0}^n$, under the measure P. There exists a constant $\mathcal{C}_5$ (which is independent of n and P) such that*

$$E_P \left(||\mathcal{W}_n(N) - B||^2 \right) \leq \frac{\mathcal{C}_5}{\sqrt{n}} \tag{4.9}$$

and

$$E_P \left(||\mathcal{W}_n(\langle N \rangle) - \langle B \rangle||^2 \right) \leq \frac{\mathcal{C}_5}{\sqrt{n}}. \tag{4.10}$$

In the equations (4.9) and (4.10), $\mathcal{W}_n$ is the linear interpolation operator which is defined on the spaces $(\mathbb{R}^d)^{n+1}$ and $(\mathbb{S}^d)^{n+1}$, respectively.

Proof. Inequality (4.9) follows immediately from (4.1) and the relation

$$||\mathcal{W}_n(N) - B|| \leq 2 \sum_{i=1}^{d} \max_{1 \leq k \leq n} \max_{k/n \leq t \leq (k+1)/n} |N_k^i - B_t^i|.$$

Next, we prove (4.10). For any $1 \leq i, j \leq d$ denote by $\langle N \rangle_k^{i,j}$ and $\langle B \rangle_t^{i,j}$, the $i-$th row and the $j-$th column of the matrices $\langle N \rangle_k$ and $\langle B \rangle_t$, respectively. Notice that $\langle B \rangle_t^{i,j} = \frac{1}{2}(\langle B^i + B^j \rangle_t - \langle B^i \rangle_t - \langle B^j \rangle_t)$ and $\langle N \rangle_k^{i,j} = \frac{1}{2}(\langle N^i + N^j \rangle_k - \langle N^i \rangle_k - \langle N^j \rangle_k)$. Thus (4.10) follows from (4.2) and the inequality

$$||\mathcal{W}_n(\langle N \rangle) - \langle B \rangle|| \leq 2 \sum_{i=1}^{d} \sum_{j=1}^{d} \max_{0 \leq k \leq n-1} \max_{k/n \leq t \leq (k+1)/n} |\langle N \rangle_k^{i,j} - \langle B \rangle_t^{i,j}|.$$

□

We conclude this section with the following technical lemma.

Lemma 4.3. *Let $A > 0$. Then we have:*
i.

$$\sup_{P \in \mathcal{P}_{\mathbf{D}}} E_P \exp(A \sup_{0 \leq t \leq 1} ||B_t||) < \infty. \tag{4.11}$$

ii. Let $n \in \mathbb{N}$ and ν be a distribution which satisfies (2.5)–(2.6). Consider a filtered probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_k\}_{k=0}^n, \tilde{P})$ together with a sequence of i.i.d. random vectors $Y_1, ..., Y_n$ which satisfy $\mathcal{L}(Y_1) = \nu$. Assume that for any i, Y_i is $\tilde{\mathcal{F}}_i$ measurable and independent of $\tilde{\mathcal{F}}_{i-1}$. Let $\{M_k\}_{k=0}^n$ be a d–dimensional stochastic process of the following form: $M_0 := 0$ and

$$M_k := \sqrt{\frac{1}{n}} \sum_{i=1}^{k} \gamma_i Y_i, \quad 1 \leq k \leq n \tag{4.12}$$

where for any i, γ_i is $\tilde{\mathcal{F}}_{i-1}$ measurable random matrix, which takes values in $\sqrt{\mathbf{D}}$. There exists a constant $\mathcal{C}_6$ (which may depend on A and ν) such that

$$\exp\left(A \max_{0 \leq k \leq n} ||M_k|| \right) < \mathcal{C}_6. \tag{4.13}$$

Proof. i. Let $P \in \mathcal{P}_{\mathbf{D}}$. From the Novikov condition it follows that for any $1 \leq i \leq d$ and $a \in \mathbb{R}$, $E_P \exp\left(aB_1^i - \frac{a^2}{2}\langle B^i \rangle_1\right) = 1$. Thus

$$E_P \left(\exp(a|B_1^i|) \right) \leq E_P(\exp(aB_1^i)) + E_P(\exp(-aB_1^i)) \leq 2 \exp\left(\frac{a^2}{2} ||\mathbf{D}|| \right)$$

where $||\mathbf{D}|| = \sup_{\mathcal{D}\in\mathbf{D}} ||\mathcal{D}||$. This together with the CauchyŰ-Schwarz inequality and the Doob–Kolomogorov inequality for the sub–martingale $\exp(A||B_t||/2)$, $t \in [0,1]$ completes the proof of (4.11).
ii. Consider the compact set $K := \{x \in \mathbb{R}^d : ||x|| \le ||\sqrt{\mathbf{D}}||\}$. Clearly, the rows of the matrices γ_j, $1 \le j \le n$ are in K. Fix $1 \le i \le d$ and consider the $i-$th component of the process M, namely we consider the process $(M^i_0, ..., M^i_n)$. From (4.12) we get that for any $a \in \mathbb{R}$

$$\mathbb{E}\left(\exp(a(M^i_k - M^i_{k-1}))|\tilde{\mathcal{F}}_{k-1}\right) \le \sup_{y\in K} \psi_\nu\left(\frac{ay}{\sqrt{n}}\right)$$

where ψ_ν is the function which is defined below (2.5). This together with (2.6) gives

$$\mathbb{E}\left(\exp(aM^i_n)\right) \le \sup_{n\in\mathbb{N}} \sup_{y\in K} \psi^n_\nu\left(\frac{ay}{\sqrt{n}}\right) < \infty. \tag{4.14}$$

From the inequality $\mathbb{E}\exp(|aM^i_n|) \le \mathbb{E}\left(\exp(aM^i_n)\right) + \mathbb{E}\left(\exp(-aM^i_n)\right)$ and the Cauchy–Schwartz inequality it follows that there exists a constant c_2 (which may depend on A and ν) such that

$$\mathbb{E}(\exp(A||M_n||)) < c_2. \tag{4.15}$$

Finally, since for any i the process M^i_k, $k \le n$ is a martingale with respect to the filtration $\{\tilde{\mathcal{F}}_k\}^n_{k=0}$ we conclude that the stochastic process $\{\exp(A||M_k||/2)\}^n_{k=0}$ is a sub–martingale and so, from (4.15) and the Doob–Kolomogorv inequality $\mathbb{E}\exp(A\max_{0\le k\le n}||M_k||) \le 4c_2$ and the proof is completed. □

5 Proof of the main results

In this section we complete the proof of Theorems 2.3 and 2.5. Let ν be a distribution which satisfies (2.5)–(2.6). Fix $\epsilon > 0$. We start with proving the following statements

$$V^\nu_n > V - \mathcal{C}_\epsilon n^{\epsilon-1/8}, \;\; \forall n \in \mathbb{N} \tag{5.1}$$

and for a bounded F

$$V^\nu_n > V - \mathcal{C}n^{-1/8}, \;\; \forall n \in \mathbb{N}. \tag{5.2}$$

Choose $n \in \mathbb{N}$ and $\delta > 0$. There exists a measure $Q \in \mathcal{P}_\mathbf{D}$ for which

$$V < \delta + E_Q F(B, \langle B\rangle). \tag{5.3}$$

Consider the stochastic process $N_k := B_{k/n}$, $0 \le k \le n$ together with its predictable variation $\{\langle N\rangle_k\}^n_{k=0}$. From (2.2) and the fact that $\mathbf{D}$ is a convex compact set (notice that convexity is essential here) we obtain that there exists a sequence of measurable functions $\tilde{\phi}_j : (\mathbb{R}^d)^j \to \sqrt{\mathbf{D}}$, $1 \le j \le n$ such that

$$\sqrt{\Delta\langle N\rangle_k} = \sqrt{\mathbb{E}\left(\Delta N_k \Delta N'_k \middle| \sigma\{N_0, N_1, ..., N_{k-1}\}\right)} := \frac{1}{\sqrt{n}}\tilde{\phi}_k(N_0, ..., N_{k-1}), \;\; \forall k \;\text{ a.s.} \tag{5.4}$$

From the Burkholder–Davis–Gundy inequality it follows that there exists a constant c_3 for which

$$E_Q\left(||\Delta N_k||^3 \middle| \sigma\{N_0, ..., N_{k-1}\}\right) \le c_3 n^{-3/2}, \;\; \forall k \;\text{ a.s.} \tag{5.5}$$

By applying (2.1), Lemmas 4.2–4.3 and Cauchy–Schwartz inequality we get

$$E_Q|F(B, \langle B\rangle) - F(\mathcal{W}_n(N), \mathcal{W}_n(\langle N\rangle))| \le c_4 n^{-1/4} \tag{5.6}$$

for some constant c_4 (which depends only on the distribution ν). From (5.5) and Lemma 3.2 we obtain that there exists a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ which contains the martingale N, a sequence of i.i.d. random vectors $Y_1, ..., Y_n$ and satisfies, $\mathcal{L}(Y_1) = \nu$, for any k the random vectors $N_1, ..., N_{k-1}$ are independent of Y_k, and

$$\tilde{P}\left(\max_{1\leq k\leq n} ||N_k - \sum_{j=1}^{k} \sqrt{\Delta\langle N\rangle_j} Y_j|| > n^{-1/8}\right) < \frac{c_5 n^{-3/2} n}{n^{-3/8}} = c_5 n^{-1/8} \tag{5.7}$$

for some constant c_5 which depends only on the distribution ν. Denote $M_k := \sum_{j=1}^{k} \sqrt{\Delta\langle N\rangle_j} Y_j$, $1 \leq k \leq n$ and $\mathcal{A} := \{\max_{1\leq k\leq n} ||N_k - M_k|| > n^{-1/8}\}$. From (2.5) and the fact that $N_1, ..., N_{k-1}$ are independent of Y_k we obtain that M is a martingale, and $\langle M\rangle = \langle N\rangle$. Next, from Lemma 4.3, the Cauchy–Schwartz inequality and the simple inequality $\exp(Ax)x^q < \exp((A+q)x)$, $A, q, x > 0$ we get that for any $A, q \geq 0$

$$\tilde{E}\Bigg(\exp\left(A(\max_{1\leq k\leq n} ||M||_k + \max_{1\leq k\leq n} ||N||_k)\right)(||\mathcal{W}_n(N)|| + ||\mathcal{W}_n(M)||)^q\Bigg) < \infty$$

where $\tilde{E}$ denotes the expectation with respect to $\tilde{P}$. This together with (5.7), the Markov inequality and the Holder inequality (for $p = \frac{1}{1-8\epsilon}$ and $q = \frac{1}{8\epsilon}$) yields that there exists constants c_6, c_7 which depend on ϵ and ν such that

$$\begin{aligned} &\tilde{E}|F(\mathcal{W}_n(N), \mathcal{W}_n(\langle N\rangle)) - F(\mathcal{W}_n(M), \mathcal{W}_n(\langle M\rangle))| \leq \\ &H_1\tilde{E}\Bigg(\exp\left(H_2(\max_{1\leq k\leq n} ||M||_k + \max_{1\leq k\leq n} ||N||_k + 2||\mathbf{D}||)\right) \times \\ &\left(n^{-1/8} + \mathbb{I}_{\mathcal{A}}(||\mathcal{W}_n(N)|| + ||\mathcal{W}_n(M)||)\right)\Bigg) \leq \\ &\leq c_6(n^{-1/8} + \tilde{P}(A)^{\frac{1}{1-8\epsilon}}) \leq c_7 n^{\epsilon - 1/8}. \end{aligned} \tag{5.8}$$

If the function F is bounded, say $F \leq R$, then we have

$$\begin{aligned} &\tilde{E}|F(\mathcal{W}_n(N), \mathcal{W}_n(\langle N\rangle)) - F(\mathcal{W}_n(M), \mathcal{W}_n(\langle M\rangle))| \leq R\tilde{P}(A) + H_1 n^{-1/8} \\ &\times\tilde{E}\left(\exp(H_2(\max_{1\leq k\leq n} ||M||_k + \max_{1\leq k\leq n} ||N||_k + 2||\mathbf{D}||))\right) \leq c_8 n^{-1/8} \end{aligned} \tag{5.9}$$

for some constant c_8 which depends only on ν. Since $\delta > 0$ was arbitrary, then in view of (5.3), (5.6) and (5.8)–(5.9) we conclude that in order to prove (5.1)–(5.2) it remains to establish the following inequality

$$V_n^\nu \geq \tilde{E}F(\mathcal{W}_n(M), \mathcal{W}_n(\langle M\rangle)). \tag{5.10}$$

Define a sequence of functions $L_k : (\mathbb{R}^d)^{k+1} \times (\mathbb{S}^d)^{k+1} \to \mathbb{R}$, $k = 0, 1, ..., n$ by the backward recursion

$$\begin{aligned} &L_n(u_0, ..., u_n, v_0, ..., v_n) := F\left(\mathcal{W}_n(u, v)\right) \quad \text{and} \\ &L_k(u_0, ..., u_k, v_0, ..., v_k) := \\ &\tilde{E}L_{k+1}\big(u_0, ..., u_k, u_k + \tfrac{1}{\sqrt{n}}\tilde{\phi}_{k+1}(u_0, ..., u_k)Y_{k+1}, v_0, ..., v_k, \\ &v_k + \tfrac{1}{n}\tilde{\phi}^2_{k+1}(u_0, ..., u_k)\big) \;\text{ for }\; k = 0, 1, ..., n-1. \end{aligned} \tag{5.11}$$

From the fact that Y_{k+1} is independent of $Y_1, ..., Y_k, N_1, ..., N_{k-1}$ it follows (by backward induction) that for any k,

$$\begin{aligned} &\tilde{E}\left(F(\mathcal{W}_n(M), \mathcal{W}_n(\langle M\rangle))|\sigma\{N_1, ..., N_{k-1}, Y_1, ..., Y_k\}\right) = \\ &L_k\left(M_0, ..., M_k, \langle N\rangle_0, ..., \langle N\rangle_k\right). \end{aligned} \tag{5.12}$$

Finally, from (2.12), (5.11)–(5.12) and the fact that $\tilde{\phi}_k$ takes values in $\sqrt{\mathbf{D}}$ for any k, we obtain (by backward induction) that $L_k \leq J_k^{\nu,n}$, $k \leq n$, and in particular

$$V_n^{\nu} = J_0^{\nu,n}(0,0) \geq L_0(0,0) = \tilde{E}F(\mathcal{W}_n(M), \mathcal{W}_n(\langle M \rangle)). \tag{5.13}$$

This completes the proof of (5.1)–(5.2). Next, fix $n \in \mathbb{N}$, a distribution ν which satisfies (2.5)–(2.6) and consider the optimal control $M^{\nu,n}$ which is given by (2.12)–(2.14). By applying Lemma 3.2 for the standard normal distribution ν_g it follows that there exists a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ which contains the martingale $M^{\nu,n}$, a sequence of i.i.d. standard Gaussian random vectors (d–dimensional) $Y_1^g, ..., Y_n^g$ such that for any k the random vectors $M_1^{\nu,n}, ..., M_{k-1}^{\nu,n}$ are independent of Y_k^g, and

$$\tilde{P}\left(\max_{1\leq k\leq n} ||M_k^{\nu,n} - \sum_{j=1}^{k} \sqrt{\Delta\langle M^{\nu,n}\rangle_j} Y_j^g|| > n^{-1/8}\right) < c_9 n^{-1/8} \tag{5.14}$$

for some constant c_9. Denote $\hat{M}_k := \sum_{j=1}^{k} \sqrt{\Delta\langle M^{\nu,n}\rangle_j} Y_j^g$, $1 \leq k \leq n$. Observe that $\langle \hat{M} \rangle = \langle M^{\nu,n} \rangle$. Thus by using similar argument to those as in (5.8)–(5.9) we obtain that there exist constants c_{10}, c_{11} such that

$$\begin{gathered}|\tilde{E}F(\mathcal{W}_n(\hat{M}), \mathcal{W}_n(\langle \hat{M} \rangle)) - V_n^{\nu}| \leq \\ \tilde{E}|F(\mathcal{W}_n(\hat{M}), \mathcal{W}_n(\langle \hat{M} \rangle)) - F(\mathcal{W}_n(M^{\nu,n}), \mathcal{W}_n(\langle M^{\nu,n} \rangle))| \leq c_{10} n^{\epsilon-1/8}\end{gathered} \tag{5.15}$$

and if the function F is bounded,

$$\begin{gathered}|\tilde{E}F(\mathcal{W}_n(\hat{M}), \mathcal{W}_n(\langle \hat{M} \rangle)) - V_n^{\nu}| \leq \\ \tilde{E}|F(\mathcal{W}_n(\hat{M}), \mathcal{W}_n(\langle \hat{M} \rangle)) - F(\mathcal{W}_n(M^{\nu,n}), \mathcal{W}_n(\langle M^{\nu,n} \rangle))| \leq c_{11} n^{-1/8}.\end{gathered} \tag{5.16}$$

By applying similar arguments to those as in (5.11)–(5.13) we conclude that

$$V_n^{\nu_g} = J_0^{\nu_g,n}(0,0) \geq \tilde{E}F(\mathcal{W}_n(\hat{M}), \mathcal{W}_n(\langle \hat{M} \rangle)). \tag{5.17}$$

Next, since the functions $h_k^{\nu_g,n}$, $1 \leq k \leq n-1$ are measurable there exists a measurable functions $z_k : (\mathbb{R}^d)^k \to \sqrt{\mathbf{D}}$, $1 \leq k \leq n-1$ such that for any $1 \leq k \leq n-1$, $z_k(Y_1^g, ..., Y_k^g) = h_k^{\nu_g,n}(M_0^{\nu_g,n}, ..., M_k^{\nu_g,n}, N_0^{\nu_g,n}, ..., N_k^{\nu_g,n})$, where the terms $M^{\nu_g,n}, N^{\nu_g,n}$ are given by (2.12)–(2.14). From the martingale representation theorem if follows that the martingale M_n which is defined by (2.16) equals to

$$\begin{gathered}M_t^n = h_0^{\nu_g,n}(0,0) W_t + \mathbb{I}_{t>1/n} \times \\ \int_{1/n}^t z_{[nu]}(\sqrt{n} W_{1/n}, \sqrt{n}(W_{2/n} - W_{1/n}), ..., \sqrt{n}(W_{[nu]} - W_{[nu]-1})) dW_u, \ \ t \in [0,1]\end{gathered}$$

and so we obtain that $P_n \in \mathcal{P}_{\mathbf{D}}$. As in (5.6) we have

$$E_n|F(B, \langle B \rangle) - F(\mathcal{W}_n(N), \mathcal{W}_n(\langle N \rangle))| \leq c_4 n^{-1/4} \tag{5.18}$$

where, recall, $N_k = B_{k/n}$, $0 \leq k \leq n$. Finally, observe that the distribution of N under P_n equals to the distribution of the martingale $M^{\nu_g,n}$. Thus from (2.15) and (5.18) we conclude that

$$V \geq E_{P_n} F(B, \langle B \rangle) \geq V_n^{\nu_g} - c_4 n^{-1/4}.$$

This together with (5.1)–(5.2) and (5.15)–(5.17) completes the proof of Theorems 2.3–2.5. □

Remark 5.1. *In [4], the following analogue of the G–expectation was presented. Given $n \in \mathbb{N}$ we consider $(\mathbb{R}^d)^{n+1}$ as the canonical space of d–dimensional paths in discrete time $k = 0, 1, ..., n$. We denote by $Z^n = \{Z^n_k\}_{k=0}^n$ the canonical process defined by $Z^n_k(z) := z_k$ for $z = (z_0, z_1, ..., z_n) \in (\mathbb{R}^d)^{n+1}$. Denote by $\mathcal{P}^n_{\boldsymbol{D}}$ the set of all measures P on the space $(\mathbb{R}^d)^{n+1}$ which satisfy the following:*
i. The canonical process Z is a martingale under P and $Z_0 = 0$ P a.s.
ii. For any k,

$$nE\left((\Delta Z^n_k)(\Delta Z^n_k)'|\sigma\{Z^n_1, ..., Z^n_{k-1}\}\right) \in \boldsymbol{D}, \quad P \text{ a.s.}$$

and

$$d^2 \inf_{\Upsilon \in \boldsymbol{D}} ||\Upsilon^{-1}||^{-1} \leq n|\Delta Z^n_k|^2 \leq d^2 \sup_{\Upsilon \in \boldsymbol{D}} ||\Upsilon||, \quad P \text{ a.s.}$$

where we set $\inf_{\Upsilon \in \boldsymbol{D}} ||\Upsilon^{-1}||^{-1} = 0$ if $\boldsymbol{D}$ has an element which is not invertible. The n–step discrete time version of the G–expectation is defined by

$$V_n := \sup_{P \in \mathcal{P}^n_{\boldsymbol{D}}} E_P F\left(\mathcal{W}_n(Z, \langle Z \rangle)\right). \tag{5.19}$$

We want to establish

$$|V_n - V| \leq \mathcal{C}_\epsilon n^{\epsilon - 1/8} \tag{5.20}$$

and if the function F is bounded

$$|V_n - V| \leq \mathcal{C} n^{-1/8}. \tag{5.21}$$

By using Lemma 3.2 (in a similar way to above) we can prove that for any distribution ν which satisfies (2.5)–(2.6),

$$V^\nu_n \geq V_n - \mathcal{C}_\epsilon n^{\epsilon - 1/8},$$

and if the function F is bounded, then

$$V^\nu_n \geq V_n - \mathcal{C} n^{-1/8}.$$

Thus, in view of Theorem 2.3, in order to establish (5.20)–(5.21) it is sufficient to find a distribution ν which is satisfying (2.5)–(2.6) and the inequality $V^\nu_n \leq V_n$. Let A be an orthogonal $(d+1) \times (d+1)$ matrix whose last raw is equals to $(\frac{1}{\sqrt{d+1}}, ..., \frac{1}{\sqrt{d+1}})$ and let $v_i \in \mathbb{R}^d$ be column vectors such that $[v_1, ..., v_{d+1}]$ is the matrix obtained from A by deleting the last row. Consider a random vector Y which values are $\sqrt{d+1}v_1, ..., \sqrt{d+1}v_{d+1}$, and the probability of each value is equal to $\frac{1}{d+1}$. It is straightforward to check that the the distribution $\nu := \mathcal{L}(Y)$ is satisfying (2.5)–(2.6) and for any $M \in \mathcal{A}^\nu_n$ the distribution of M (on the space $(\mathbb{R}^d)^{n+1}$) belongs to $\mathcal{P}^n_{\mathbf{D}}$, and so $V^\nu_n \leq V_n$.

Acknowledgments. I am very grateful to M. Nutz, H.M. Soner and O. Zeitouni for valuable discussions.

References

[1] L. Denis, M. Hu, and S. Peng. *Function spaces and capacity related to a sublinear expectations: applications to G–Brownian motion paths,* Potential Anal., **2** (2011), 139–161. MR-2754968

[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** (2006), 827–852. MR-2244434

[3] S. Deparis and C. Martini, *Superhedging Strategies and Balayage in Discrete Time,* Proceedings of the 4th Ascona Conferecne on Stochastic Analysis, Random Fields and Applications. **58** (2004), 205–219, Birkhäuser. MR-2096290

[4] Y. Dolinsky, M. Nutz and H.M. Soner, *Weak Approximations of G–Expectations,* Stochastic Processes and their Applications. **2** (2012), 664–675 MR-2868935

[5] Y. Hu and S. Peng, *Some Estimates for Martingale Representation under G–Expectation,* preprint.

[6] M. Nutz and H.M. Soner, *Superhedging and Dynamic Risk Measures under Volatility Uncertainty,* to appear in SIAM Journal on Control and Optimization.

[7] S. Peng, *G–expectation, G–Brownian motion and related stochastic calculus of Itô type,* Stochastic Analysis and Applications, volume 2 of Abel Symp., (2007), 541–567, Springer Berlin. MR-2397805

[8] S. Peng, *Multi–dimensional G–Brownian motion and related stochastic calculus under G–expectation.,* Stochastic. Processes. Appl., **12** (2008), 2223–2253. MR-2474349

[9] S. Peng, *Nonlinear expectations and stochastic calculus under uncertainty,* preprint.

[10] S. Peng *Tightness, weak compactness of nonlinear expectations and application to CLT,* preprint.

[11] A.I Sakhanenko, *A New Way to Obtain Estimates in the Invariance Principle,* High Dimensional Probability II, (2000), 221–243.

[12] H.M. Soner, N. Touzi, and J. Zhang, *Martingale representation theorem for the G–expectation,* Stochastic. Processes. Appl., **2** (2011), 265–287. MR-2746175