A Central Limit Theorem for Sets of Probability Measures*

Zengjing Chen Larry G. Epstein

June 25, 2020

Abstract

We prove a central limit theorem for a sequence of random variables whose means are ambiguous and vary in an unstructured way. Their joint distribution is described by a set of measures. The limit is (not the normal distribution and is) defined by a backward stochastic differential equation that can be interpreted as modeling an ambiguous continuous-time random walk.

1 Introduction

We present a Central Limit Theorem (CLT) for situations where random events (or experiments) are describable by nonsingleton sets of probability measures. Such sets arise in economics and finance as the subjective prior beliefs of an agent within a model who does not have sufficient information to justify reliance on a single probability measure (e.g. [20, 19, 15]), in mathematical statistics and econometrics where, for example, they represent the predictions of the theory being tested or estimated empirically and where predictions are multivalued because the theory is incomplete (e.g. [21, 37, 36, 11]). We refer to such situations as featuring ambiguity. Our focus in this paper is on a sequential or temporal context, where experiments are ordered. The set of probability measures can be taken to be objective (the set of logically possible probability laws) or subjective (representing an individual's beliefs about future experiments).

*Chen is at School of Mathematics, Shandong University, zjchen@sdu.edu.cn, and Epstein is at Department of Economics, Boston University, lepstein@bu.edu. We thank Juan Li and Shige Peng for helpful comments. Chen gratefully acknowledges the support of the National Key R&D Program of China (grant No. 2018YFA0703900).

Our first main result can be outlined roughly as follows. Let $(\Omega, \mathcal{G})$ be a measurable space and let (X_i) be a sequence of (real-valued) random variables, where X_i describes the outcome of experiment i. Let $\mathcal{P}$ be a set of probability measures on $(\Omega, \mathcal{G})$. Information is represented by the filtration $\{\mathcal{G}_i\}$, $(\mathcal{G}_0 = \{\emptyset, \Omega\})$, such that (X_i) is adapted to $\{\mathcal{G}_i\}$ and $\mathcal{G} = \sigma(\cup_1^\infty \mathcal{G}_i)$. Assume that the upper and lower conditional means of the X_is satisfy:

$$ess \sup_{Q\in\mathcal{P}} E_Q[X_i|\mathcal{G}_{i-1}] = \overline{\mu} \text{ and } ess \inf_{Q\in\mathcal{P}} E_Q[X_i|\mathcal{G}_{i-1}] = \underline{\mu}, \text{ for all } i \geq 1. \quad (1.1)$$

Ambiguity about means is indicated if $\overline{\mu} > \underline{\mu}$. Conditional variances are taken to be unambiguous and common to all X_is:

$$E_Q\left[(X_i - E_Q[X_i|\mathcal{G}_{i-1}])^2|\mathcal{G}_{i-1}\right] = \sigma^2 > 0 \text{ for all } Q \in \mathcal{P} \text{ and all } i. \quad (1.2)$$

Then, under suitable additional assumptions, we show that for every $\varphi \in C([-\infty, \infty])$, the class of all bounded continuous functions with finite limits at $\pm\infty$,

$$\lim_{n\to\infty} \sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{1}{n}\sum_{i=1}^n X_i + \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{1}{\sigma}(X_i - E_Q[X_i|\mathcal{G}_{i-1}])\right)\right] = \mathbb{E}_{[\underline{\mu},\overline{\mu}]}\left[\varphi(B_1)\right], \quad (1.3)$$

where the right side of this equation is defined to be Y_0, given that (Y_t, Z_t) is the solution of the backward stochastic differential equation (BSDE)

$$Y_t = \varphi(B_1) + \int_t^1 \max_{\underline{\mu}\leq\mu\leq\overline{\mu}}(\mu Z_s)ds - \int_t^1 Z_s dB_s,\ 0 \leq t \leq 1, \quad (1.4)$$

and (B_t) is a standard Brownian motion on a probability space $(\Omega^*, \mathcal{F}^*, P^*)$.

The result highlights the connection between CLTs and BSDEs. If $\overline{\mu} = \underline{\mu} = \mu$, and given (1.1) and any fixed measure in $\mathcal{P}$, then $(X_i - \mu)$ is a martingale difference and the limit result reduces to a form of the classical martingale CLT (applying the strong Law of Large Numbers (LLN) for martingales which gives a.s. convergence of $\frac{1}{n}\sum_{i=1}^n X_i$ to μ). In addition, the right side reduces to a linear BSDE that, through its solution, yields the expectation of $\varphi(B_1)$ under the normal distribution $\boldsymbol{N}(\mu, 1)$. More generally, in our CLT accommodating ambiguity about means, the associated BSDE is nonlinear. Rather it corresponds to a model in which a Brownian motion is augmented by a drift that can vary stochastically thru time subject only to remaining in the interval $\left[\underline{\mu}, \overline{\mu}\right]$. For example, when φ is the indicator function $I_{[a,b]}$, [3] shows that $sgn(Z_s) = -sgn\left(B_s - \frac{a+b-\left(\overline{\mu}+\underline{\mu}\right)(1-s)}{2}\right)$, and hence

$$\arg\max_{\underline{\mu}\leq\mu\leq\overline{\mu}}(\mu Z_s) = \begin{cases} \underline{\mu} & \text{if } B_s \geq \frac{a+b-\left(\overline{\mu}+\underline{\mu}\right)(1-s)}{2}, \\ \overline{\mu} & \text{if } B_s < \frac{a+b-\left(\overline{\mu}+\underline{\mu}\right)(1-s)}{2}. \end{cases}$$

This stochastic variability of the maximizing mean μ leads to a non-normal limiting distribution.

Two important points regarding tractability should be noted. First, from [3] and also Lemma 6.11 below, the indicated BSDE can be solved in closed-form for some specifications of φ. For example, when φ is the indicator for the interval $[a,b]$, then the right side of (1.3) is given by

$$\mathbb{E}_{[\underline{\mu},\overline{\mu}]}[I_{[a,b]}(B_1)] = \begin{cases} \Phi_{-\overline{\mu}}(-a) - e^{-\frac{(\overline{\mu}-\underline{\mu})(b-a)}{2}}\Phi_{-\overline{\mu}}(-b) & \text{if } a+b \geq d, \\ \Phi_{\underline{\mu}}(b) - e^{-\frac{(\overline{\mu}-\underline{\mu})(b-a)}{2}}\Phi_{\underline{\mu}}(a) & \text{if } a+b < d, \end{cases} \tag{1.5}$$

where $d \equiv \overline{\mu} + \underline{\mu}$ and Φ_μ is the normal cdf with mean μ and unit variance. The second point concerns the left side of (1.3) which is nonstandard in that the argument of φ, whose distribution is at issue, includes the measures Q in $\mathcal{P}$ and hence is not a function only of past realizations of the X_is. However, our second principal result (Theorem 4.3) is that for a class of functions φ, including indicators and quadratics, both of which are prominent in statistical theory and methods, (1.3) is valid also when each conditional expectation $E_Q[X_i|\mathcal{G}_{i-1}]$ is replaced by a suitable (and explicit) function of $(X_1, ..., X_{i-1})$ alone. Potential usefulness of this result is illustrated by an application to hypothesis testing.

The key additional assumption underlying both theorems is that the set $\mathcal{P}$ is "rectangular", or closed with respect to the pasting of alien marginals and conditionals. (Rectangularity was introduced in [12] in the context of recursive utility theory, where an axiomatic analysis demonstrated its central role in modeling dynamic behavior. It has been studied and applied also in robust stochastic dynamic optimization [34], in the literature on dynamic risk measures [33, 8, 1], and in continuous-time modeling in finance [6].) It can be understood as endowing $\mathcal{P}$ with a recursive structure that yields a form of the law of iterated expectations. If $\mathcal{P} = \{P\}$, which implies (and, for our purposes, is essentially equivalent to) $\overline{\mu} = \underline{\mu}$, then the law of iterated expectations is a consequence of updating by Bayes rule and rectangularity is vacuously satisfied. (Sections 2 and 3 provide a precise definition of rectangularity and some motivating informal interpretation.)

Some connections to the literature conclude this introduction. In the classical probability framework, there are numerous CLTs with non-normal limiting distributions (with stable laws, for example) [24, 9], all of which have much different motivation and limits than our result. There exist alternative generalizations of the classical theorem that are motivated by robustness to ambiguity. In [11] (see also the generalization in [35]), experiments are not ordered and the analysis is intended for a cross-sectional context. In addition, $\mathcal{P}$ is assumed to be the core of a convex (that is, supermodular) capacity, which renders it inconsistent with a recursive structure [5]. Finally, the limiting distribution in their result is the normal, in contrast to our novel BSDE-based limit.

Closer to this paper is the CLT due to Peng [31, 32] who also assumes that experiments are ordered. Peng's focus is on ambiguity about variance (or at least about the second moment), while our focus is on ambiguity about the mean. A more recent paper [16] provides a CLT (Theorem 3.2) with ambiguity about both mean and variance. (Their theorem also considers rates of convergence, which are ignored here.) To compare it with this paper, consider the special case where there is ambiguity about means only. Then their CLT is related primarily to our Theorem 5.1, rather than to our central results Theorems 4.1 and 4.3. In particular, only in the latter are limits defined by a BSDE rather than by a normal distribution (as in [16]). See section 5 for elaboration. Another difference is that our approach is more probability-theoretic: Peng and coauthors take a nonlinear expectation as the core primitive and adopt the PDE approach, while our primitive is a set of probability measures and conditionals are central only in our analysis.

The next section describes the model's primitives and key assumptions formally. These are illustrated in section 3 via a canonical example that can be understood as generalizing the classical random walk to accommodate ambiguity. The two main CLT results (Theorems 4.1 and 4.3) are presented in section 4. Section 5 provides perspective on our main results by relating them to an alternative CLT (Theorem 5.1) and a weak LLN for our setting (Corollary 5.2). Proofs of Theorems 4.1 and 4.3 are presented in section 6. An appendix contains other proofs and supplementary material.

2 Primitives and assumptions

Let $(\Pi_1^\infty \Omega_i, \{\mathcal{G}_n\}_{n=1}^\infty)$ be a filtered space modeling a sequence of experiments. The set of possible outcomes for the i^{th} experiment is Ω_i. For each n, $\mathcal{G}_n$ is a σ-algebra on $\Pi_1^n \Omega_i$ representing the observable events regarding experiments $1, ..., n$. (Accordingly, we assume that $\mathcal{G}_n$ is increasing with n and we take $\mathcal{G}_0$ to be the trivial σ-algebra.) The observable events for the collection of all experiments are given by $\mathcal{G}$,

$$\mathcal{G} = \sigma(\cup_1^\infty \mathcal{G}_n),$$

a σ-algebra on Ω, where

$$\Omega = \Pi_1^\infty \Omega_i.$$

(Here and in the sequel, we identify each $\mathcal{G}_n$ in the obvious way with a σ-algebra on Ω.) The ex ante probabilities of observable events are not known precisely and are represented by a set $\mathcal{P}$ of probability measures,[1]

$$\mathcal{P} \subset \Delta(\Omega, \mathcal{G}).$$

We limit ambiguity about which events are possible and assume that all measures in $\mathcal{P}$ are equivalent on each $\mathcal{G}_n$.

Below we assume that for each measure P in $\mathcal{P}$ and each n, there exists a regular conditional measure $P(\cdot \mid \mathcal{G}_n)$. For example, a well-known [28, Theorem 7.1] sufficient condition for regular $\mathcal{G}_n$-conditionals to exist for every P in $\Delta(\Omega, \mathcal{G})$ is that $(\Omega, \mathcal{G})$ is a separable standard Borel space (a special case is where Ω is a complete separable metric space and $\mathcal{G}$ is its Borel σ-algebra).

Finally, we consider a sequence (X_i) of real-valued random variables (r.v.), $X_i : \Pi_1^\infty \Omega_j \longrightarrow \mathbb{R}$, such that X_i is $\mathcal{G}_i$-measurable (using the Borel σ-algebra on $\mathbb{R}$). Think of X_i as a scalar measure of the outcome of experiment i or of the value (or utility) of that outcome. In general, X_i can depend also on the outcomes of earlier experiments.

Remark 2.1. *We presume a particular ordering of experiments, which may be arbitrary in cross-sectional contexts. Thus we view the analysis and the resulting CLT as more relevant to sequential or time-series contexts where an ordering is given.*

In the rest of this section, we describe our assumptions on the above primitives. We use the following notation. $\mathcal{H}$ denotes the set of all r.v. X on $(\Omega, \mathcal{G})$ satisfying $\sup_{Q \in \mathcal{P}} E_Q[|X|] < \infty$; $E_Q[\cdot]$ is the expectation under the probability measure Q. For any X in $\mathcal{H}$, its *upper and lower expectations* are defined respectively by

$$\mathbb{E}[X] \equiv \sup_{Q \in \mathcal{P}} E_Q[X], \qquad \mathcal{E}[X] \equiv \inf_{Q \in \mathcal{P}} E_Q[X] = -\mathbb{E}[-X],$$

and its *conditional upper and lower expectations* are defined respectively by

$$\mathbb{E}[X \mid \mathcal{G}_n] \equiv ess \sup_{Q \in \mathcal{P}} E_Q[X \mid \mathcal{G}_n], \quad \mathcal{E}[X \mid \mathcal{G}_n] \equiv ess \inf_{Q \in \mathcal{P}} E_Q[X \mid \mathcal{G}_n].$$

Obviously, the conditional expectations are well-defined due to equivalence of all measures in $\mathcal{P}$ on each $\mathcal{G}_n$. (See section 6 for key properties of these expectations.) Rewritten with this notation, (1.1) takes the form

$$\mathbb{E}[X_i \mid \mathcal{G}_{i-1}] = \overline{\mu} \text{ and } \mathcal{E}[X_i \mid \mathcal{G}_{i-1}] = \underline{\mu} \text{ for all } i.$$

[1]For any measurable space $(Y, \mathcal{F})$, the corresponding set of probability measures is denoted $\Delta(Y, \mathcal{F})$.

Say that (X_i) has an *unambiguous conditional variance* σ^2 if (1.2) is satisfied. Say that (X_i) satisfies the *Lindeberg condition* if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\left[|X_i|^2 \, I_{\{|X_i| > \sqrt{n}\epsilon\}} \right] = 0, \quad \forall \epsilon > 0. \tag{2.1}$$

To formulate the remaining assumption requires additional notation and terminology. Write

$$\begin{aligned}
\omega_{(n)} &= (\omega_n, ...)\,, \; \omega^{(n)} = (\omega_1, ..., \omega_n)\,, \\
\mathcal{P}_{0,n} &= \left\{ P_{|\mathcal{G}_n} : P \in \mathcal{P} \right\} \text{ and} \\
\mathcal{G}_{(n+1)} &= \left\{ A \subset \Pi_{n+1}^{\infty} \Omega_i : \Pi_1^n \Omega_i \times A \in \mathcal{G} \right\}.
\end{aligned}$$

A probability kernel from $(\Pi_1^n \Omega_i, \mathcal{G}_n)$ to $\left(\Pi_{n+1}^{\infty} \Omega_i, \mathcal{G}_{(n+1)}\right)$ is a function $\lambda : \Pi_1^n \Omega_i \times \mathcal{G}_{(n+1)} \longrightarrow [0,1]$ satisfying:

Kernel1 $\forall \omega^{(n)} \in \Pi_1^n \Omega_i$, $\lambda\left(\omega^{(n)}, \cdot\right)$ is a probability measure on $\left(\Pi_{n+1}^{\infty} \Omega_i, \mathcal{G}_{(n+1)}\right)$,

Kernel2 $\forall A \in \mathcal{G}_{(n+1)}$, $\lambda(\cdot, A)$ is a $\mathcal{G}_n$-measurable function on $\Pi_1^n \Omega_i$.

Any pair (p_n, λ) consisting of a probability measure p_n on $(\Pi_1^n \Omega_i, \mathcal{G}_n)$ and a probability kernel λ as above, induces a unique probability measure P on $(\Pi_1^{\infty} \Omega_i, \mathcal{G})$ that coincides with p_n on $\mathcal{G}_n$. It is given by, $\forall A \in \mathcal{G}$,

$$P(A) = \int_{\Pi_1^n \Omega_i} \int_{\Pi_{n+1}^{\infty} \Omega_i} I_A\left(\omega^{(n)}, \omega_{(n+1)}\right) \lambda\left(\omega^{(n)}, d\omega_{(n+1)}\right) p_n\left(d\omega^{(n)}\right). \tag{2.2}$$

For $Q \in \mathcal{P}$, let $Q(\cdot \mid \mathcal{G}_n)$ denote its regular conditional. Then it defines a probability kernel λ by: $\forall \omega^{(n)} \in \Pi_1^n \Omega_i$,

$$\lambda\left(\omega^{(n)}, A\right) = Q\left(\Pi_1^n \Omega_i \times A \mid \mathcal{G}_n\right)\left(\omega^{(n)}\right), \forall A \in \mathcal{G}_{(n+1)}. \tag{2.3}$$

A feature of such a kernel is that the single measure Q is used to define the conditional at every $\omega^{(n)}$. We are interested in kernels for which the measure to be conditioned can vary with $\omega^{(n)}$. Say that the probability kernel λ is a $\mathcal{P}$*-kernel* if: $\forall \omega^{(n)} \in \Pi_1^n \Omega_i$, $\exists Q \in \mathcal{P}$ satisfying (2.3).

Finally, say that $\mathcal{P}$ is *rectangular* (with respect to the filtration $\{\mathcal{G}_n\}$) if: $\forall n$, $\forall p_n \in \mathcal{P}_{0,n}$ and for every $\mathcal{P}$*-kernel* λ, if P is defined as in (2.2), then $P \in \mathcal{P}$. (Note that a measure $P \in \Delta(\Omega, \mathcal{G})$ is well-defined by (2.2), for any $p_n \in \mathcal{P}_{0,n}$ and $\mathcal{P}$-*kernel* λ, because of the assumption that all measures in $\mathcal{P}$ are equivalent on $\mathcal{G}_n$). When $\mathcal{P}$ is the singleton $\{P\}$, rectangularity is trivially implied by Bayesian updating, specifically by the fact that after decomposing P into a marginal and conditional, these can be pasted together to recover P. More generally, rectangularity requires that the set $\mathcal{P}$ is closed also with respect to pasting together conditionals and marginals that are *alien*, that is, induced by possibly different measures in $\mathcal{P}$. In this sense, $\mathcal{P}$ does not restrict the pattern of heterogeneity across experiments (see the next section for elaboration).

The significance of rectangularity is illuminated by the following lemma. (See Appendix A.1 for a partial proof. The complement of any $A \subset \Omega$ is denoted A^c.)

Lemma 2.2. *$\mathcal{P}$ rectangular implies the following (for any $0 \leq m \leq n \in \mathbb{N}$).*
(i) ***Stability by composition:*** *For any $Q, R \in \mathcal{P}$, $\exists P \in \mathcal{P}$ such that, for any $X \in \mathcal{H}$,*

$$E_P[X|\mathcal{G}_m] = E_Q[E_R[X|\mathcal{G}_n]|\mathcal{G}_m].$$

(ii) ***Stability by bifurcation:*** *For any $Q, R \in \mathcal{P}$, and any $A_n \in \mathcal{G}_n$, $\exists P \in \mathcal{P}$ such that, for any $X \in \mathcal{H}$,*

$$E_P[X|\mathcal{G}_n] = I_{A_n} E_Q[X|\mathcal{G}_n] + I_{A_n^c} E_R[X|\mathcal{G}_n].$$

(iii) ***Law of iterated upper expectations****: For any $X \in \mathcal{H}$,*

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}_n]|\mathcal{G}_m] = \mathbb{E}[X|\mathcal{G}_m]. \tag{2.4}$$

(iv) Let $\{X_i\}$ be a sequence in $\mathcal{H}$. Set $S_{n-1} = \sum_{i=1}^{n-1} X_i$ and, for any $Q \in \mathcal{P}$, $S_{n-1}^Q = S_{n-1} - \sum_{i=1}^{n-1} E_Q[X_i|\mathcal{G}_{i-1}]$. Then, for any continuous bounded functions f, h:

$$\sup_{Q \in \mathcal{P}} E_Q \left[f\left(\frac{S_{n-1}}{n} + \frac{S_{n-1}^Q}{\sqrt{n}} \right) + h\left(\frac{S_{n-1}}{n} + \frac{S_{n-1}^Q}{\sqrt{n}} \right) X_n \right]$$
$$= \sup_{Q \in \mathcal{P}} E_Q \left[ess \sup_{R \in \mathcal{P}} E_R \left[f\left(\frac{S_{n-1}}{n} + \frac{S_{n-1}^Q}{\sqrt{n}} \right) + h\left(\frac{S_{n-1}}{n} + \frac{S_{n-1}^Q}{\sqrt{n}} \right) X_n | \mathcal{G}_{n-1} \right] \right].$$

(v) If $\{X_i\}$ is a sequence in $\mathcal{H}$ satisfying (1.1), then

$$\mathbb{E}\left[X_n \mid \mathcal{G}_{n-1}\right] = \mathbb{E}\left[X_n\right] \text{ and } \mathcal{E}\left[X_n \mid \mathcal{G}_{n-1}\right] = \mathcal{E}\left[X_n\right]. \tag{2.5}$$

(i) and (ii) make explicit two senses in which rectangularity of $\mathcal{P}$ implies that combinations of distinct measures ($Q \neq R$) from $\mathcal{P}$ leave one within $\mathcal{P}$. Together they lead to (iii). The latter is built into the classical model but must be adopted explicitly, via rectangularity, for upper (or lower) expectations. For a general set $\mathcal{P}$, one would expect the supremum on the left in (2.4) to be (weakly) larger because it permits the choices of measures conditional on each history $\omega^{(n-1)}$ and the ex ante measure on $\mathcal{G}_{n-1}$ to be alien. However, rectangularity implies that any such combination of measures yields a measure in $\mathcal{P}$, and thus the single-stage supremum on the right is no smaller. The proof of our CLT employs a similar recursive relation also in instances when the r.v. itself depends on $Q \in \mathcal{P}$ as in (iv), the intuition for which is similar to that for (iii). (v) states that conditional upper and lower expectations do not vary with the outcomes of previous experiments. It is an immediate consequence of (1.1) and (iii); for example,

$$\mathbb{E}\left[X_n\right] = \mathbb{E}\left[\mathbb{E}\left[X_n \mid \mathcal{G}_{n-1}\right]\right] = \mathbb{E}\left[\overline{\mu}\right] = \overline{\mu} = \mathbb{E}\left[X_n \mid \mathcal{G}_{n-1}\right].$$

3 Example: IID

Our canonical example (adapted from [13]) is as follows. Specialize the above framework by assuming that there exists a measurable space $(\overline{\Omega}, \overline{\mathcal{F}})$ such that, for all $1 \leq i \leq n$,

$$(\Omega_i, \mathcal{F}_i) = (\overline{\Omega}, \overline{\mathcal{F}}) \text{ and } \mathcal{G}_n = \Pi_1^n \mathcal{F}_i.$$

That is, experiments have a common set of possible outcomes $\overline{\Omega}$ and an associated common σ-algebra $\overline{\mathcal{F}}$. In addition, suppose that, for all i,

$$X_i = \overline{X} : (\Omega_i, \mathcal{F}_i) \to \mathbb{R}.$$

One-step-ahead conditionals are central. Thus, for each P in $\mathcal{P}$, and each n, let $P_{n,n+1}\left(\omega^{(n)}\right)$ denote the restriction to $\mathcal{G}_{n+1}$ of $P\left(\cdot \mid \mathcal{G}_n\right)\left(\omega^{(n)}\right)$.

Fix a subset $\mathcal{L}$ of $\Delta(\overline{\Omega}, \overline{\mathcal{F}})$, all of whose measures are equivalent. Then the IID model is defined via the set $\mathcal{P}^{IID}$,

$$\mathcal{P}^{IID} = \left\{P \in \Delta\left(\Omega, \mathcal{G}\right) : P_{n,n+1}\left(\omega^{(n)}\right) \in \mathcal{L},\ \forall n, \omega^{(n)} \in \mathcal{G}_n\right\}. \tag{3.1}$$

The set consists of all measures whose one-step-ahead conditionals, at every history, lie in $\mathcal{L}$. Thus, $\mathcal{L}$ is the set of plausible probability laws for each experiment, independent of history, modeling partial ignorance about each experiment separately. There remains the question of the perception of, or information about, the sequence of experiments, that is, how experiments are related to one another. In spite of $\mathcal{L}$ being common to all i, in this model experiments are not necessarily identical. (Accordingly, we refer to experiments as being *indistinguishable rather than identical* and take IID to mean "*indistinguishably* and independently distributed".) Indeed, any measure in $\mathcal{L}$ is plausible as the law describing the i^{th} experiment in conjunction with any possibly different measure in $\mathcal{L}$ being the law describing the j^{th} experiment. Indeed, $\mathcal{P}^{IID}$ imposes no restrictions on joint distributions thus capturing *agnosticism about the pattern of heterogeneity across experiments.* As demonstrated below, this feature is closely related to rectangularity.

In the special case where $\mathcal{L} = \{P\}$, $\mathcal{P}^{IID}$ consists of the single i.i.d. product of P, as in a random walk. One might think of $\mathcal{P}^{IID}$ as modeling an "ambiguous random walk".

The following lemma gives some readily verified properties of $\mathcal{P}^{IID}$ (see Appendix A.2 for some proof details).

Lemma 3.1. *The set $\mathcal{P}^{IID}$ satisfies (for every $n \in \mathbb{N}$):*
(i) $\mathcal{P}^{IID}$ is rectangular.
(ii) Measures in $\mathcal{P}^{IID}$ are mutually equivalent on each $\mathcal{G}_n$.
(iii) For any $\varphi \in C(\mathbb{R})$, with $\varphi(X_n) \in \mathcal{H}$,

$$\mathbb{E}\left[\varphi\left(X_n\right) \mid \mathcal{G}_{n-1}\right] = \sup_{q \in \mathcal{L}} E_q\left[\varphi\left(\overline{X}\right)\right] = \mathbb{E}\left[\varphi(X_n)\right] = \mathbb{E}\left[\varphi(X_1)\right].$$

(iv) Conditional variances satisfy:

$$\begin{aligned}\sup_{Q \in \mathcal{P}^{IID}} E_Q\left[(X_n - E_Q[X_n|\mathcal{G}_{n-1}])^2|\mathcal{G}_{n-1}\right] &= \sup_{q \in \mathcal{L}} E_q\left[(\overline{X} - E_q[\overline{X}])^2\right], \\ \inf_{Q \in \mathcal{P}^{IID}} E_Q\left[(X_n - E_Q[X_n|\mathcal{G}_{n-1}])^2|\mathcal{G}_{n-1}\right] &= \inf_{q \in \mathcal{L}} E_q\left[(\overline{X} - E_q[\overline{X}])^2\right].\end{aligned}$$

The key property of $\mathcal{P}^{IID}$ is rectangularity. Because of its centrality, we verify rectangularity here: Let p_n, λ and P be as in (2.3). Then, for the given $\omega^{(n)}$,

$$\begin{aligned}P\left(\Pi_1^n\Omega_i \times \cdot \mid \mathcal{G}_n\right)\left(\omega^{(n)}\right) &= \lambda\left(\omega^{(n)}, \cdot\right) \\ &= Q\left(\Pi_1^n\Omega_i \times \cdot \mid \mathcal{G}_n\right)\left(\omega^{(n)}\right),\end{aligned}$$

for some $Q \in \mathcal{P}^{IID}$. Therefore, the one-step-ahead conditional of P at history $\omega^{(n)}$ equals that of Q and hence lies in $\mathcal{L}$. Therefore, $P \in \mathcal{P}^{IID}$.

The lemma implies that $\mathcal{P}^{IID}$ readily accommodates also the other assumptions in the CLT below. For example, (1.1) is implied by (iii) and conditional variances are common and unambiguous if and only if[2]

$$var_q\left(\overline{X}\right) \equiv E_q\left[(\overline{X} - E_q[\overline{X}])^2\right] = \sigma^2, \text{ for all } q \in \mathcal{L}. \tag{3.2}$$

For perspective, consider also the set $\mathcal{P}^{prod}$, consisting of all (nonidentical) product measures that can be constructed from $\mathcal{L}$ - refer to this as the *product model*. The set $\mathcal{P}^{prod}$ also implies a degree of agnosticism about heterogeneity – after all, it consists of product measures $\Pi_{i=1}^{\infty}\ell_i$, where $\ell_i \neq \ell_j$ in general, and these measures are restricted only by the requirement that they lie in $\mathcal{L}$. However, the two models differ in a significant way in that $\mathcal{P}^{prod}$ violates rectangularity, and hence also (2.4), for example. This is because $\mathcal{P}^{prod}$ is "too small" in the sense of not being closed with respect to the pasting of alien marginals and conditionals (note that $\mathcal{P}^{prod}$ is a strict subset of $\mathcal{P}^{IID}$). Our interpretation of $\mathcal{P}^{prod}$ is that it models *certainty* that the probability law for experiment i does not vary with the outcomes of preceding experiments (note that invariance to these outcomes is exhibited by each individual measure in $\mathcal{P}^{prod}$). In contrast, in $\mathcal{P}^{IID}$ one-step-ahead conditionals can vary arbitrarily across different histories subject only to lying in $\mathcal{L}$. Thus $\mathcal{P}^{IID}$ permits heterogeneity across experiments to vary stochastically and thereby models greater agnosticism regarding heterogeneity.

A simple concrete example illustrates both models and the difference between them. Each experiment can produce one of three outcomes: success (s), failure (f) and the neutral outcome (n). Thus $\overline{\Omega} = \{s, f, n\}$ and $\overline{\mathcal{F}}$ is the power set. Outcomes are valued by $\overline{X}$ according to

$$\overline{X}(s) = 1, \overline{X}(f) = -1, \overline{X}(n) = 0.$$

Outcomes are uncertain but their probabilities are not known precisely. Let

$$0 < q < p, \;\; p + q \leq 1.$$

It is known that, for each experiment, and regardless of the outcomes in preceding experiments, the outcomes s, f and n (in that order) are given

[2] In decision theory (in [20], for example), it is often innocuous and a convenient normalization to take sets of measures to be convex. But because variances are not linear in the measure q, convexity of $\mathcal{L}$ precludes (3.2) except in the degenerate case where means are also unambiguous. Thus we do not assume that $\mathcal{L}$ is convex.

either by the favorable distribution $(p, q, 1-p-q)$ *or* by the unfavorable distribution $(q, p, 1-p-q)$, that is,

$$\mathcal{L} = \{(p, q, 1-p-q), (q, p, 1-p-q)\}.$$

There is no additional information provided that would justify, for example, assigning weights (or probabilities) to these two distributions and then using the average as the Bayesian model would require - there is complete ignorance about which distribution applies for any given experiment. Consequently, conditional on any history, the implied upper and lower means of each X_i equal $\overline{\mu} = p - q$ and $\underline{\mu} = -(p-q)$ respectively, and the implied conditional variance σ^2 of each X_i is unambiguous and equals $p + q - (p-q)^2$. Thus $p+q$ and $p-q$ parametrize risk (measured by σ^2) and ambiguity (measured by $\frac{\overline{\mu}-\underline{\mu}}{2}$) respectively in the sense that a change in $p+q$ alone changes only risk and a change in $p-q$ alone changes only ambiguity.

The final issue is the relation between experiments. Arguably, ignorance about which probability law applies to any given experiment, logically implies (or at least suggests) ignorance about how experiments are related. Accordingly, $\mathcal{P}^{IID}$ does not restrict measures on the entire sequence of experiments beyond requiring that each one-step-ahead conditional lie in $\mathcal{L}$. In contrast, $\mathcal{P}^{prod}$ admits only measures for which the conditional law for the i^{th} experiment, though it can be either favorable or unfavorable, is necessarily the same for all histories of outcomes. Thus, for example, $\mathcal{P}^{prod}$ excludes measures that specify both (1) the favorable law for experiment i after a successful outcome in $i-1$, and (2) the unfavorable law for experiment i after a failure in $i-1$.

4 The main results

4.1 Two theorems

We extend (a version of) the classical martingale CLT to admit ambiguity about means while maintaining the assumption of unambiguous variances. Though the theorems deal with real-valued random variables, multidimensional versions can be proven in a similar fashion and will be reported elsewhere.

Theorem 4.1. *Let the sequence (X_i) be such that $X_i \in \mathcal{H}$ for each i, and where (X_i) satisfies (1.1) and (1.2), with conditional upper and lower means $\overline{\mu}$ and $\underline{\mu}$, and unambiguous conditional variance $\sigma^2 > 0$. Assume also the Lindeberg condition (2.1) and that $\mathcal{P}$ is rectangular. Then, for any $\varphi \in C([-\infty, \infty])$,*

$$\lim_{n\to\infty} \sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{1}{n}\sum_{i=1}^{n} X_i + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sigma}(X_i - E_Q[X_i|\mathcal{G}_{i-1}])\right)\right] = \mathbb{E}_{[\underline{\mu},\overline{\mu}]}[\varphi(B_1)], \tag{4.1}$$

or equivalently,

$$\lim_{n\to\infty} \inf_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{1}{n}\sum_{i=1}^{n} X_i + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sigma}(X_i - E_Q[X_i|\mathcal{G}_{i-1}])\right)\right] = \mathcal{E}_{[\underline{\mu},\overline{\mu}]}[\varphi(B_1)], \tag{4.2}$$

where $\mathbb{E}_{[\underline{\mu},\overline{\mu}]}[\varphi(B_1)] \equiv Y_0$ ***is called g-expectation by Peng in [30],*** *given that (Y_t, Z_t) is the solution of the BSDE*

$$Y_t = \varphi(B_1) + \int_t^1 \max_{\underline{\mu}\le\mu\le\overline{\mu}}(\mu Z_s)ds - \int_t^1 Z_s dB_s,\ 0 \le t \le 1,$$

and $\mathcal{E}_{[\underline{\mu},\overline{\mu}]}[\varphi(B_1)] \equiv y_0$, given that (y_t, z_t) is the solution of the BSDE

$$y_t = \varphi(B_1) + \int_t^1 \min_{\underline{\mu}\le\mu\le\overline{\mu}}(\mu z_s)ds - \int_t^1 z_s dB_s,\ 0 \le t \le 1. \tag{4.3}$$

Here (B_t) is a standard Brownian motion on a probability space $(\Omega^*, \mathcal{F}^*, P^*)$.

Remark 4.2. *By standard limiting arguments, (4.1) can be extended to indicator functions for intervals. Such indicators are sufficient in the classical CLT, because of the additivity of a single probability measure. But when dealing with sets of measures, (4.1) is strictly stronger. Another remark is that while in (4.1) the second term inside $\varphi(\cdot)$ is normalized by the standard deviation σ, a change of variables delivers a CLT without that normalization. (Set $\alpha = \sigma$ and $\beta = 1$ in the statement of Theorem A.2 in the appendix.)*

Three differences from classical results stand out. First, the limiting distribution is not normal but rather is given by the BSDE (1.4). Another notable difference is that the r.v. on the left in (4.1) combines the sample average, typical of LLNs, with a term that is more typical of CLTs. Both of these features will be discussed in section 5 below.

Here we consider the fact that the argument of φ above, whose distribution is the focus, includes measures Q from $\mathcal{P}$, which might raise concerns about tractability. To partially alleviate such concerns, we show that (4.1) takes on a more tractable form when restricted to "symmetric" functions φ. Say that $\varphi : \mathbb{R} \to \mathbb{R}$ is *symmetric with center* $c \in \mathbb{R}$ if $\varphi(c-x) = \varphi(c+x)$ for all $x \in \mathbb{R}$. Examples include indicator(s) $\varphi(t) = \pm I_{[a,b]}(t)$ with $c = \frac{a+b}{2}$, and quadratic functions $\varphi(t) = \pm(t-c)^2$, both of which are prominent in statistical methods. It is important to emphasize also that for both of these classes of functions [3] provides closed-form expressions for the BSDE-based limits appearing on the right sides of (4.1) and (4.2) above, and (4.8) and (4.9) below; recall (1.5), for example. Section 4.2 exploits these closed-forms in an application to hypothesis testing.

The next theorem is the second major result of the paper. (Throughout sums of the form $\Sigma_n^0 x_i$, $n \geq 1$,are taken to equal 0, and increasing/decreasing are intended in the weak sense.)

Theorem 4.3. *Adopt the assumptions in Theorem 4.1 and let the function $\varphi \in C([-\infty, \infty])$ be symmetric with center $c \in \mathbb{R}$. For $n \geq 1$ and $0 \leq m \leq n$, define*

$$M_{m,n} = \frac{1}{n}\sum_{i=1}^{m} X_i + \frac{1}{\sqrt{n}}\sum_{i=1}^{m}\frac{1}{\sigma}\left(X_i - \mu_i^n\right), \quad M_{0,n} \equiv 0, \tag{4.4}$$

$$\widetilde{M}_{m,n} = \frac{1}{n}\sum_{i=1}^{m} X_i + \frac{1}{\sqrt{n}}\sum_{i=1}^{m}\frac{1}{\sigma}\left(X_i - \widetilde{\mu}_i^n\right), \quad \widetilde{M}_{0,n} \equiv 0, \tag{4.5}$$

where

$$\begin{aligned} \mu_m^n &= \overline{\mu} I_{A_{m-1,n}} + \underline{\mu} I_{A_{m-1,n}^c}, \\ A_{m-1,n} &= \left\{ M_{m-1,n} \leq -\tfrac{\overline{\mu}+\underline{\mu}}{2}\left(1 - \tfrac{m-1}{n}\right) + c \right\}, \end{aligned} \tag{4.6}$$

and

$$\begin{aligned} \widetilde{\mu}_m^n &= \overline{\mu} I_{\widetilde{A}_{m-1,n}} + \underline{\mu} I_{\widetilde{A}_{m-1,n}^c}, \\ \widetilde{A}_{m-1,n} &= \left\{ \widetilde{M}_{m-1,n} \geq -\tfrac{\overline{\mu}+\underline{\mu}}{2}\left(1 - \tfrac{m-1}{n}\right) + c \right\}. \end{aligned} \tag{4.7}$$

(1) Assume that φ is decreasing on (c, ∞). Then

$$\lim_{n\to\infty} \sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(M_{n,n}\right)\right] = \mathbb{E}_{[\underline{\mu},\overline{\mu}]}[\varphi\left(B_1\right)]. \tag{4.8}$$

(2) Assume that φ is increasing on (c,∞). Then

$$\liminf_{n\to\infty}\sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\widetilde{M}_{n,n}\right)\right] \geq \mathbb{E}_{[\underline{\mu},\overline{\mu}]}[\varphi(B_1)]. \tag{4.9}$$

Furthermore, assume also that

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\frac{1}{n}\sum_{m=1}^{n}\sup_{Q\in\mathcal{P}} E_Q\left[\left|E_Q[X_m|\mathcal{G}_{m-1}]-\widetilde{\mu}_m^n\right| I_{\widetilde{A}^{\delta}_{m-1,n}}\right]=0, \tag{4.10}$$

where

$$\widetilde{A}^{\delta}_{m-1,n}=\left\{\left|\widetilde{M}_{m-1,n}+\tfrac{\overline{\mu}+\underline{\mu}}{2}\left(1-\tfrac{m-1}{n}\right)-c\right|\leq\delta\right\},\quad \delta>0.$$

Then

$$\lim_{n\to\infty}\sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\widetilde{M}_{n,n}\right)\right] = \mathbb{E}_{[\underline{\mu},\overline{\mu}]}[\varphi(B_1)]. \tag{4.11}$$

Consider (1). Given $n\geq 1$, $\{\mu_m^n : m\leq n\}$ are defined recursively with μ_m^n being a function of $(X_1,...,X_{m-1})$. The definition is clearer in the special case where

$$c=0 \text{ and } \underline{\mu}+\overline{\mu}=0. \tag{4.12}$$

Then

$$\mu_m^n=\begin{cases}\overline{\mu} & \text{if } \frac{1}{n}\sum\limits_{i=1}^{m-1}X_i+\frac{1}{\sqrt{n}}\sum\limits_{i=1}^{m-1}\frac{1}{\sigma}\left(X_i-\mu_i^n\right)\leq 0,\\ \underline{\mu} & \text{if } \frac{1}{n}\sum\limits_{i=1}^{m-1}X_i+\frac{1}{\sqrt{n}}\sum\limits_{i=1}^{m-1}\frac{1}{\sigma}\left(X_i-\mu_i^n\right)> 0.\end{cases} \tag{4.13}$$

That is, μ_m^n is set as large (small) as possible when $M_{m-1,n}\leq(>)0$, hence lying in the region where φ is increasing (decreasing).

Conclude that the theorem delivers the statistic $M_n=M_{n,n}$ defined in (4.4), and, through the upper expectation of $\varphi(M_n)$ for the indicated set of functions φ, (4.8) gives information about its asymptotic distribution. Moreover, in combination with (1.5), this information can be expressed in closed-form when φ is the indicator for an interval (for a simpler proof than in [3] see Lemma 6.11 below). In particular, we have: For any $a<b\in\mathbb{R}$,

$$\begin{aligned}&\lim_{n\to\infty}\sup_{Q\in\mathcal{P}} Q\left(a\leq M_n\leq b\right)\\ &=\begin{cases}\Phi_{-\overline{\mu}}(-a)-e^{-\frac{(\overline{\mu}-\underline{\mu})(b-a)}{2}}\Phi_{-\overline{\mu}}(-b) & \text{if } a+b\geq\overline{\mu}+\underline{\mu},\\ \Phi_{\underline{\mu}}(b)-e^{-\frac{(\overline{\mu}-\underline{\mu})(b-a)}{2}}\Phi_{\underline{\mu}}(a) & \text{if } a+b<\overline{\mu}+\underline{\mu}.\end{cases}\end{aligned} \tag{4.14}$$

Similarly, part (2) produces the statistic $\widetilde{M}_n = \widetilde{M}_{n,n}$ defined in (4.5), that plays a corresponding role. A difference is that only the inequality (4.9) is proven in general, though equality obtains under the condition (4.10). In that case one obtains (as above) that: For any $a < b \in \mathbb{R}$,

$$\lim_{n\to\infty} \inf_{Q\in\mathcal{P}} Q\left(a \leq \widetilde{M}_n \leq b\right)$$
$$= \begin{cases} \Phi_{-\underline{\mu}}\left(-a\right) - e^{\frac{(\overline{\mu}-\underline{\mu})(b-a)}{2}} \Phi_{-\underline{\mu}}\left(-b\right) & \text{if } a+b \geq \overline{\mu}+\underline{\mu}, \\ \Phi_{\overline{\mu}}\left(b\right) - e^{\frac{(\overline{\mu}-\underline{\mu})(b-a)}{2}} \Phi_{\overline{\mu}}\left(a\right) & \text{if } a+b < \overline{\mu}+\underline{\mu}. \end{cases}$$

Finally, we note that (4.10) is easily verified when $\overline{\mu} = \underline{\mu} = \mu$, because then $E_Q[X_m|\mathcal{G}_{m-1}] = \mu = \widetilde{\mu}_m^n$, for any $Q \in \mathcal{P}$ and $1 \leq m \leq n$. More generally, (4.10) is satisfied if

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} Q(\widetilde{A}_{m-1,n}^{\delta}) = 0.$$

When $c = \pm\infty$, the assumptions in the theorem imply global monotonicity conditions for φ, and lead to the fixed means $\underline{\mu}$ and $\overline{\mu}$ replacing the stochastic means appearing in (4.1), (4.8) and (4.9) respectively, and to the normal as the limiting distribution. These features apply, in particular, to one-sided indicators $I_{(-\infty,b]}$ and $I_{[a,\infty)}$, and stand in contrast to the implications described above for two-sided indicators $I_{[a,b]}$.

Corollary 4.4. *Adopt the assumptions in Theorem 4.1 and assume that* $\varphi \in C([-\infty,\infty])$.

(1) If φ *is decreasing on* $\mathbb{R}$*, then*

$$\lim_{n\to\infty} \sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{1}{n}\sum_{i=1}^{n} X_i + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sigma}(X_i - \underline{\mu})\right)\right] = \int \varphi\left(t\right) d\Phi_{\underline{\mu}}\left(t\right). \tag{4.15}$$

(2) If φ *is increasing on* $\mathbb{R}$*, then*

$$\lim_{n\to\infty} \sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{1}{n}\sum_{i=1}^{n} X_i + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sigma}(X_i - \overline{\mu})\right)\right] = \int \varphi\left(t\right) d\Phi_{\overline{\mu}}\left(t\right). \tag{4.16}$$

4.2 An application to hypothesis testing

We give an illustrative application of Theorem 4.3 to hypothesis testing that demonstrates tractability; a more comprehensive study of statistical applications is beyond the scope of this paper. Here we exploit also explicit solutions to BSDEs established in [3], an example of which is provided in (1.5).

Consider the model

$$X_i = \theta + Y_i,\ i = 1, 2, ...,$$

where $\theta \in \mathbb{R}$ is a parameter of interest, (X_i) describes observable data, and (Y_i) is an unobservable error process. The usual assumption on errors is that they are i.i.d. with zero mean. Since errors are unobservable, a weaker a priori specification is natural. Thus, for example, assume the IID model $\mathcal{P}^{IID}$, and for simplicity, that errors have means that lie in the interval $[-\kappa, \kappa]$. Both the variance σ and κ, which measures ambiguity, are assumed known. In the special case $\kappa = 0$, θ is the unknown mean of each X_i and one can test hypotheses about its value by exploiting the classical CLT. Here we generalize that test procedure to cover $\kappa > 0$.

Let $\varphi = I_{[a,b]}$, which is symmetric with center $c = \frac{a+b}{2}$, and define the statistic $M_n = M_{n,n}$ by (4.4). It follows from Theorem 4.3(1) and (1.5) that, for any θ, (see Appendix A.4),

$$\begin{aligned}
&\lim_{n\to\infty} \sup_{Q\in\mathcal{P}^{IID}} Q\left(\{M_n - b \leq \theta \leq M_n - a\}\right) \\
&= \lim_{n\to\infty} \sup_{Q\in\mathcal{P}^{IID}} Q\left(\{a \leq M_n - \theta \leq b\}\right) = \mathbb{E}_{[-\kappa,\kappa]}[I_{[a,b]}\left(B_1\right)] \qquad (4.17) \\
&= \begin{cases} \Phi_{-\kappa}\left(-a\right) - e^{-\kappa(b-a)}\Phi_{-\kappa}\left(-b\right) & \text{if } a+b \geq 0, \\ \Phi_{-\kappa}\left(b\right) - e^{-\kappa(b-a)}\Phi_{-\kappa}\left(a\right) & \text{if } a+b < 0. \end{cases}
\end{aligned}$$

The null hypothesis is $H_0 : \theta \in \Theta$ and the alternative is H_1: $\theta \notin \Theta$, for some $\Theta \subset \mathbb{R}$. A nonstandard feature is that there are several probability laws that conceivably describe the data even given a specific θ. One test procedure is to accept H_0 if and only if the realized statistic M_n is "sufficiently consistent" with some $\theta \in \Theta$ and some probability law in $\mathcal{P}^{IID}$. Precisely, choose $[a, b]$ so that $\mathbb{E}_{[-\kappa,\kappa]}[I_{[a,b]}\left(B_1\right)] = 1 - \alpha$, for a suitable α, and accept H_0 if and only if $\mathcal{C}_n \cap \Theta \neq \varnothing$, where the random interval $\mathcal{C}_n$ is given by

$$\mathcal{C}_n = \left[M_n - b, M_n - a\right].$$

Then, if H_0 is true, in the limit for large samples the (upper) probability of acceptance is approximately $1-\alpha$. The upper probability of wrongly rejecting

H_0 is typically greater than α because of the multiplicity of measures in $\mathcal{P}^{IID}$:

$$\begin{aligned}\sup_{Q\in\mathcal{P}^{IID}} Q\left(\{\mathcal{C}_n \cap \Theta = \varnothing\}\right) &= 1 - \inf_{Q\in\mathcal{P}^{IID}} Q\left(\{\mathcal{C}_n \cap \Theta \neq \varnothing\}\right) \\ &\geq 1 - \sup_{Q\in\mathcal{P}^{IID}} Q\left(\{\mathcal{C}_n \cap \Theta \neq \varnothing\}\right).\end{aligned}$$

Let $\Theta = \{\theta_0\}$ and suppose that the truth is $\theta = \theta_1 \equiv \theta_0 + \xi$, $\xi \neq 0$. Then the limiting upper probability of wrongly accepting θ_0 is given by

$$\begin{aligned}&\lim_{n\to\infty} \sup_{Q\in\mathcal{P}^{IID}} Q\left(\{a \leq M_n - \theta_0 \leq b\}\right) \\ &= \lim_{n\to\infty} \sup_{Q\in\mathcal{P}^{IID}} Q\left(\{a \leq M_n - \theta_1 + \xi \leq b\}\right) \\ &= \mathbb{E}_{[-\kappa+\xi,\kappa+\xi]}[I_{[a,b]}\left(B_1\right)] = \mathbb{E}_{[-\kappa,\kappa]}[I_{[a-\xi,b-\xi]}\left(B_1\right)]\end{aligned}$$

We emphasize that, given a and b, $\mathbb{E}_{[-\kappa,\kappa]}[I_{[a-\xi,b-\xi]}\left(B_1\right)]$ can be expressed in closed-form (using (1.5)); and a and b might be chosen by solving

$$\min_{a\leq b} \mathbb{E}_{[-\kappa,\kappa]}[I_{[a-\xi,b-\xi]}\left(B_1\right)] \text{ s.t. } \mathbb{E}_{[-\kappa,\kappa]}[I_{[a,b]}\left(B_1\right)] \geq 1-\alpha. \tag{4.18}$$

5 Further discussion

We turn attention to two nonstandard features of the CLT Theorem 4.1 mentioned only briefly above. One novel feature is that the limit is defined by the BSDE (1.4). It is shown in [6, Theorem 2.2], using the Girsanov Theorem, that $\mathbb{E}_{[\underline{\mu},\overline{\mu}]}[\cdot]$ is also an upper expectation for a set of probability measures, where these are defined on $C([0,1])$, the space of continuous trajectories. Moreover, measures in this set define differing models of the underlying stochastically varying (instantaneous) drift. Stochastic variability of the drift is suggested by (1.4), according to which it varies between $\underline{\mu}$ and $\overline{\mu}$ depending on the sign of Z_s. When the mean is unambiguous ($\underline{\mu} = \overline{\mu} = \mu$), then the drift is constant and $\mathbb{E}_{[\underline{\mu},\overline{\mu}]}[\cdot]$ reduces to a linear expectation with normal distribution. However, in general, *the limit is given by a two parameter ($\underline{\mu}$ and $\overline{\mu}$) family of upper expectations* that model stochastically varying drift in a continuous-time context. This limiting family is common to a large class of models (for example, to all IID models in section 3), thus endowing the BSDE with special significance for asymptotic approximations in a sequential context with considerable unstructured heterogeneity in means.

The other notable feature is that the r.v. on the left in (4.1) combines the sample average, typical of LLNs, with a term that is more typical of CLTs. In the classical i.i.d. or martingale model, including the empirical average $\frac{1}{n}\sum_{i=1}^{n} X_i$ is of little consequence for the CLT because the LLN permits replacing it by the common mean of the X_is, thereby merely shifting the mean of the limiting normal distribution. This supports the common view that, in large samples, sample average reveals location of the population distribution while the ($\sqrt{n}$-scaled) average deviation from the mean reflects the distribution about that location. But this separation of roles is not true in our framework because empirical averages need not converge given ambiguity (see related LLNs in [13, 26, 31, 4], for example). Next we show that both a LLN and a "more standard-looking" CLT can be obtained from Theorem 4.1 - the former as a corollary and the latter by adapting the proof of our CLT. However, our CLT is more than the "sum of these parts"; for example, a BSDE-based limit as in (4.1) is not present or at all evident from inspection of the two derivative results.

Theorem 5.1. *Let the sequence (X_i) be such that $X_i \in \mathcal{H}$ for each i, and where (X_i) satisfies (1.1) and (1.2), with conditional upper and lower means $\overline{\mu}$ and $\underline{\mu}$, and unambiguous conditional variance $\sigma^2 > 0$. Suppose also that (X_i) satisfies the Lindeberg condition (2.1). Then, for any $\varphi \in C([-\infty, \infty])$,*

$$\lim_{n\to\infty} \sup_{Q\in\mathcal{P}} E_Q \left[\varphi \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sigma} (X_i - E_Q[X_i|\mathcal{G}_{i-1}] \right) \right] = \int \varphi(t)\, d\Phi_0(t). \quad (5.1)$$

A proof can be constructed along the lines of that of Theorem 4.1 as indicated in Remarks 6.5 and 6.12 and in Appendix A.5.

In comparison with Theorem 4.1, the above theorem drops rectangularity and yields a limit given by the normal distribution as in the classical martingale CLT. This is intuitive since, as argued earlier, the non-normal limit in Theorem 4.1 reflects agnosticism about the *stochastic* variation in means, which is implicit in rectangularity. The difference between the two theorems can be seen clearly through their canonical examples, the IID model $\mathcal{P}^{IID}$ for Theorem 4.1 and, we would argue, the product model $\mathcal{P}^{prod}$ for the second theorem. The noted agnosticism motivates $\mathcal{P}^{IID}$ but is excluded by $\mathcal{P}^{prod}$ (section 3).

Another point of comparison is that while Theorem 5.1 adopts weaker assumptions, there is a sense in which it also produces a weaker result. For example, it does not discriminate between the IID and product models - the limit is the same for both. In contrast, it can be shown that Theorem 4.1, where the sample average term is included, is not valid for the product model.

Theorem 5.1 also clarifies the relation (outlined in the introduction) between this paper and CLTs by Peng and coauthors. In particular, in common with (5.1) and unlike (4.1), [16, Theorem 3.2] excludes the sample average term and delivers a normal distribution in the limit.

Finally, we show that if Theorem 4.1 is modified so as to include only the sample average term, then one obtains the following LLN. (The idea in the proof, found in Appendix A.6, is first to note the appropriate form of (4.1) when the deviation term is weighted by $\alpha > 0$, and then to let $\alpha \to 0$.)

Corollary 5.2. *Adopt the assumptions in Theorem 4.1. Then, for any* $\varphi \in C([-\infty, \infty])$,

$$\lim_{n\to\infty} \sup_{Q\in\mathcal{P}} E_Q \left[\varphi\left(\frac{1}{n} \sum_{i=1}^{n} X_i \right) \right] = \sup_{\underline{\mu}\leq\mu\leq\overline{\mu}} \varphi(\mu). \tag{5.2}$$

For example, if $\varphi = I_{[a,b]}$, then (5.2) takes the form

$$\lim_{n\to\infty} \sup_{Q\in\mathcal{P}} Q\left(a \leq \frac{1}{n} \sum_{i=1}^{n} X_i \leq b \right) = \begin{cases} 1 & \text{if } [a,b] \cap \left[\underline{\mu}, \overline{\mu}\right] \neq \varnothing \\ 0 & \text{otherwise.} \end{cases}$$

6 Main proofs

This section proves Theorems 4.1 and 4.3. Throughout we use the following well-known properties of (conditional) upper expectations, understood to hold for all X and Y in $\mathcal{H}$, and all $n \geq 0$.

1. Monotonicity: $X \geq Y$ implies $\mathbb{E}[X \mid \mathcal{G}_n] \geq \mathbb{E}[Y \mid \mathcal{G}_n]$.

2. Sub-additivity: $\mathbb{E}[X + Y \mid \mathcal{G}_n] \leq \mathbb{E}[X \mid \mathcal{G}_n] + \mathbb{E}[Y \mid \mathcal{G}_n]$.

3. Homogeneity: If Z is $\mathcal{G}_n$ measurable,
$$\mathbb{E}[ZX \mid \mathcal{G}_n] = Z^+ \mathbb{E}[X \mid \mathcal{G}_n] - Z^- \mathcal{E}[X \mid \mathcal{G}_n].$$

4. Translation homogeneity: If Z is $\mathcal{G}_n$ measurable,
$$\mathbb{E}[Z + X \mid \mathcal{G}_n] = Z + \mathbb{E}[X \mid \mathcal{G}_n].$$

The assumptions in Theorem 4.1 are adopted throughout. As indicated following (4.3), (B_t) is a standard Brownian motion on a filtered probability space $(\Omega^*, \mathcal{F}^*, \{\mathcal{F}_t\}, P^*)$; $\{\mathcal{F}_t\}$ is the natural filtration generated by (B_t).

For both theorems, we prove them first for the special case where

$$-\underline{\mu} = \overline{\mu} = \kappa \geq 0, \tag{6.1}$$

that is,

$$\mathbb{E}[X_i \mid \mathcal{G}_{i-1}] = \kappa, \quad \mathcal{E}[X_i \mid \mathcal{G}_{i-1}] = -\kappa.$$

Then the results asserted for general $\underline{\mu}$ and $\overline{\mu}$ are established by applying the preceding special case to (Y_i), where $Y_i = X_i - \frac{\overline{\mu}+\underline{\mu}}{2}$, and thus

$$\mathbb{E}[Y_i \mid \mathcal{G}_{i-1}] = \frac{\overline{\mu}-\underline{\mu}}{2}, \quad \mathcal{E}[Y_i \mid \mathcal{G}_{i-1}] = -\frac{\overline{\mu}-\underline{\mu}}{2}.$$

6.1 Lemmas

The following lemmas prepare the groundwork for proofs of both Theorems 4.1 and 4.3. The special case (6.1) is assumed throughout unless specified otherwise.

For any fixed $\epsilon > 0$, define $g_\epsilon : \mathbb{R} \to \mathbb{R}$ by

$$g_\epsilon(z) = \kappa \left(\sqrt{z^2+\epsilon^2} - \epsilon\right). \tag{6.2}$$

Obviously, $g_\epsilon(0) = 0$ and g_ϵ is symmetric with center $c = 0$. For any suitably integrable random variable $\xi \in \mathcal{F}_1$, define g-expectation by $\mathbb{E}_{g_\epsilon}[\xi] = Y_0^\epsilon$, where $(Y_t^\epsilon, Z_t^\epsilon)$ is the unique solution to the BSDE

$$Y_t^\epsilon = \xi + \int_t^1 g_\epsilon\left(Z_s^\epsilon\right) ds - \int_t^1 Z_s^\epsilon dB_s,\ 0 \leq t \leq 1. \tag{6.3}$$

(Existence of a unique solution follows from [27].) Moreover, by [10, Proposition 2.1], for any suitably integrable $\xi \in \mathcal{F}_1$,

$$\mathbb{E}_{g_\epsilon}[\xi] \to \mathbb{E}_{g_0}[\xi], \text{ as } \epsilon \to 0,$$

where $\mathbb{E}_{g_0}[\xi] = Y_0^0$, and (Y_t^0, Z_t^0) is the unique solution to the BSDE (6.3) for the extreme case corresponding to $\epsilon = 0$, where

$$g_0\left(z\right) = \kappa |z|, \tag{6.4}$$

and $\mathbb{E}_{g_0}$ is alternative notation for $\mathbb{E}_{[-\kappa,\kappa]}$. We consider g_ϵ for $\epsilon > 0$ in order to overcome the nondifferentiability of g_0 at $z = 0$. (The relevant smoothness is exploited in Lemma 6.1.)

We introduce a sequence of functions generated by g-expectation $\mathbb{E}_{g_\epsilon}$. Some properties of g-expectations can be found in [30], we need to prove the following properties.

Given $\varphi \in C_b^3(\mathbb{R})$, let $\xi = \varphi\left(x + B_1 - B_{\frac{m}{n}}\right)$ in (6.3) and define the functions $\{H_{m,n}\}_{m=0}^n$ by

$$H_{m,n}(x) \equiv \mathbb{E}_{g_\epsilon}\left[\varphi\left(x + B_1 - B_{\frac{m}{n}}\right)\right],\ m = 0, \cdots, n. \tag{6.5}$$

($\epsilon > 0$ is fixed and dependence on ϵ is suppressed notationally.) Obviously,

$$H_{n,n}(x) = \varphi(x),\ H_{0,n}(x) = \mathbb{E}_{g_\epsilon}\left[\varphi\left(x + B_1\right)\right].$$

The following lemma shows that the functions $\{H_{m,n}\}_{m=0}^n$ are suitably differentiable given that $\varphi \in C_b^3(\mathbb{R})$ and $\epsilon > 0$.

Lemma 6.1. *The functions $\{H_{m,n}\}_{m=0}^n$ satisfy:*

(1) $H_{m,n} \in C_b^2(\mathbb{R})$, for $n \geq 1,\ m = 0, 1, \cdots n$.

(2) The second derivatives of $H_{m,n}$ are uniformly bounded and Lipschitz continuous with uniform Lipschitz constant for $\{(m,n) : 0 \leq m \leq n\}$.

(3) Dynamic programming principle: for $n \geq 1,\ m = 1, ..., n$,

$$H_{m-1,n}(x) = \mathbb{E}_{g_\epsilon}\left[H_{m,n}\left(x + B_{\frac{m}{n}} - B_{\frac{m-1}{n}}\right)\right],\ x \in \mathbb{R}.$$

(4) Identically distributed: for $n \geq 1,\ m = 1, ..., n$,

$$\mathbb{E}_{g_\epsilon}\left[H_{m,n}\left(x + B_{\frac{m}{n}} - B_{\frac{m-1}{n}}\right)\right] = \mathbb{E}_{g_\epsilon}\left[H_{m,n}\left(x + B_{\frac{1}{n}}\right)\right],\ x \in \mathbb{R}.$$

Proof: (1) and (2): From the nonlinear Feynman-Kac Formula [10, Proposition 4.3], we have $H_{m,n}(x) = u(\frac{m}{n}, x)$ and u is the solution of the PDE

$$\begin{cases} \partial_t u + \frac{1}{2}\partial_{xx}^2 u + \kappa\left(\sqrt{|\partial_x u|^2 + \epsilon^2} - \epsilon\right) = 0, \\ u(1,x) = \varphi(x). \end{cases} \tag{6.6}$$

Next we prove that for any $t \in [0,1]$, $u(t,\cdot) \in C_b^2(\mathbb{R})$; $u(t,\cdot)$, $\partial_x u(t,\cdot)$, $\partial_{xx}^2 u(t,\cdot)$ are bounded uniformly in $t \in [0,1]$; and for any $x, x' \in \mathbb{R}$, $\exists C > 0$ such that $|\partial_{xx}^2 u(t,x) - \partial_{xx}^2 u(t,x')| \leq C|x - x'|,\ \forall t \in [0,1]$.

By the definition of g_ϵ,

$$g'_\epsilon(z) = \kappa\frac{z}{\sqrt{\epsilon^2+z^2}} \Rightarrow |g'_\epsilon(z)| \le \kappa,$$

$$g''_\epsilon(z) = \kappa\frac{\epsilon^2}{(\epsilon^2+z^2)^{3/2}} \Rightarrow |g''_\epsilon(z)| \le \frac{\kappa}{\epsilon},$$

$$g'''_\epsilon(z) = -\kappa\frac{3z\epsilon^2}{(\epsilon^2+z^2)^{5/2}} \Rightarrow |g'''_\epsilon(z)| \le \frac{3\kappa}{\epsilon^2}.$$

Consider the following BSDE,

$$Y_s^{t,x} = \varphi(x+B_1-B_t) + \int_s^1 g_\epsilon(Z_r^{t,x})dr - \int_s^1 Z_r^{t,x}dB_r,\ s\in[t,1]. \tag{6.7}$$

Then $u(t,x) = Y_t^{t,x}$ is the classical unique solution of PDE (6.6), and

$$\partial_x Y_s^{t,x} = \varphi'(x+B_1-B_t) + \int_s^1 g'_\epsilon(Z_r^{t,x})\partial_x Z_r^{t,x}dr - \int_s^1 \partial_x Z_r^{t,x}dB_r,\ s\in[t,1]. \tag{6.8}$$

$$\begin{aligned}\partial_x^2 Y_s^{t,x} =& \varphi''(x+B_1-B_t) + \int_s^1 g''_\epsilon(Z_r^{t,x})|\partial_x Z_r^{t,x}|^2dr + \int_s^1 g'_\epsilon(Z_r^{t,x})\partial_x^2 Z_r^{t,x}dr\\ &- \int_s^1 \partial_x^2 Z_r^{t,x}dB_r,\ s\in[t,1].\end{aligned} \tag{6.9}$$

From standard estimates of BSDEs ([10]), we have, $\forall p\ge 2$, $\forall x\in\mathbb{R}$,

$$\begin{aligned}
&E_{P^*}\left[\sup_{s\in[t,1]}|Y_s^{t,x}|^p|\mathcal{F}_t\right] + E_{P^*}\left[(\int_t^1 |Z_s^{t,x}|^2ds)^{\frac{p}{2}}|\mathcal{F}_t\right]\\
\le& C_p^0 E_{P^*}\left[|\varphi(x+B_1-B_t)|^p + (\int_t^1 |g_\epsilon(0)|dr)^p|\mathcal{F}_t\right] \le C_p^0\|\varphi\|^p;\\
&E_{P^*}\left[\sup_{s\in[t,1]}|\partial_x Y_s^{t,x}|^p|\mathcal{F}_t\right] + E_{P^*}\left[(\int_t^1 |\partial_x Z_s^{t,x}|^2ds)^{\frac{p}{2}}|\mathcal{F}_t\right]\\
\le& C_p^1 E_{P^*}\left[|\varphi'(x+B_1-B_t)|^p|\mathcal{F}_t\right] \le C_p^1\|\varphi'\|^p;\\
&E_{P^*}\left[\sup_{s\in[t,1]}|\partial_x^2 Y_s^{t,x}|^p|\mathcal{F}_t\right] + E_{P^*}\left[(\int_t^1 |\partial_x^2 Z_s^{t,x}|^2ds)^{\frac{p}{2}}|\mathcal{F}_t\right]\\
\le& C_p^2 E_{P^*}\left[|\varphi''(x+B_1-B_t)|^p + (\int_t^1 |g''_\epsilon(Z_s^{t,x})||\partial_x Z_s^{t,x}|^2ds)^p|\mathcal{F}_t\right]\\
\le& C_p^2(\|\varphi''\|^p + (\tfrac{\kappa}{\epsilon})^p C_{2p}^1\|\varphi'\|^{2p}),
\end{aligned}$$

where $C_p^0, C_p^1, C_{2p}^1, C_p^2$ are constants independent of t and $\|f\| = \sup_{x\in\mathbb{R}} f(x)$ denote the sup norm of function f. Then, for any $t \in [0,1]$, $u(t,\cdot) \in C_b^2(\mathbb{R})$ and $u(t,\cdot), \partial_x u(t,\cdot), \partial_{xx}^2 u(t,\cdot)$ are bounded uniformly in $t \in [0,1]$.

From (6.7), the Malliavin derivative satisfies, for $u \in [t,s)$,

$$D_u Y_s^{t,x} = \varphi'(x + B_1 - B_t) + \int_s^1 g'_\epsilon(Z_r^{t,x}) D_u Z_r^{t,x} dr - \int_s^1 D_u Z_r^{t,x} dB_r, \ s \in [t,1].$$

From standard estimates for BSDEs, for $s \in [t,1]$, we have

$$E_{P^*}\left[\left(\int_s^1 |D_u Z_r^{t,x}|^2 dr\right)^{\frac{p}{2}} |\mathcal{F}_s\right] \leq C_p^1 E_{P^*}\left[|\varphi'(s + B_1 - B_t)|^p |\mathcal{F}_s\right] \leq C_p^1 \|\varphi'\|^p,$$

and from (6.8), we have

$$\begin{aligned} D_u\left[\partial_x Y_s^{t,x}\right] =& \varphi''(x + B_1 - B_t) + \int_s^1 g''_\epsilon(Z_r^{t,x}) D_u[Z_r^{t,x}] \partial_x Z_r^{t,x} dr \\ &+ \int_s^1 g'_\epsilon(Z_r^{t,x}) D_u\left[\partial_x Z_r^{t,x}\right] dr - \int_s^1 D_u\left[\partial_x Z_r^{t,x}\right] dB_r, \ s \in [t,1]. \end{aligned}$$

Let $d\tilde{B}_s = dB_s - g'_\epsilon(Z_s^{t,x})ds$, $\rho_s = \exp\left\{\int_0^s g'_\epsilon(Z_r^{t,x}) dB_r - \frac{1}{2}\int_0^s |g'_\epsilon(Z_r^{t,x})|^2 dr\right\}$, and $E_{P^*}[\frac{d\tilde{P}}{dP^*}|_{\mathcal{F}_s}] = \rho_s$. Then,

$$\begin{aligned} &\left|D_u\left[\partial_x Y_s^{t,x}\right]\right| \\ =& \left|E_{\tilde{P}}\left[\varphi''(x + B_1 - B_t) + \int_s^1 g''_\epsilon(Z_r^{t,x}) D_u[Z_r^{t,x}] \partial_x Z_r^{t,x} dr |\mathcal{F}_s\right]\right| \\ \leq& \|\varphi''\| + \frac{\kappa}{\epsilon} E_{P^*}\left[\rho_1(\rho_s)^{-1} \cdot \int_s^1 |D_u[Z_r^{t,x}]| \cdot |\partial_x Z_r^{t,x}| dr |\mathcal{F}_s\right] \\ \leq& \|\varphi''\| + \frac{\kappa}{\epsilon} M_s \left(E_{P^*}\left[\left(\int_s^1 |D_u[Z_r^{t,x}]|^2 dr\right)^2 |\mathcal{F}_s\right]\right)^{\frac{1}{4}} \left(E_{P^*}\left[\left(\int_s^1 |\partial_x Z_r^{t,x}|^2 dr\right)^2 |\mathcal{F}_s\right]\right)^{\frac{1}{4}} \\ \leq& K, \end{aligned}$$

where $M_s \equiv \left(E_{P^*}\left[\rho_1^2(\rho_s)^{-2}|\mathcal{F}_s\right]\right)^{\frac{1}{2}}$, and satisfies

$$\begin{aligned} M_s^2 =& E_{P^*}\left[e^{\int_s^1 2g'_\epsilon(Z_r^{t,x})dB_r - \int_s^1 |g'_\epsilon(Z_r^{t,x})|^2 dr} |\mathcal{F}_s\right] \\ =& E_{P^*}\left[e^{\int_s^1 2g'_\epsilon(Z_r^{t,x})dB_r - \frac{1}{2}\int_s^1 |2g'_\epsilon(Z_r^{t,x})|^2 dr} e^{\int_s^1 |g'_\epsilon(Z_r^{t,x})|^2 dr} |\mathcal{F}_s\right] \leq e^{\kappa^2}. \end{aligned}$$

Here K is a constant that depends on $\kappa, \epsilon, p, \|\varphi'\|, \|\varphi''\|$. With $u \in [t,s)$, from (6.8), the Malliavin derivative satisfies, $\forall s \in [t,1]$,

$$D_u\left[\partial_x Y_s^{t,x}\right] = -\int_u^s g''_\epsilon(Z_r^{t,x}) D_u\left[Z_r^{t,x}\right] \partial_x Z_r^{t,x} dr - \int_u^s g'_\epsilon(Z_r^{t,x}) D_u\left[\partial_x Z_r^{t,x}\right] dr$$

$$+\int_u^s D_u\left[\partial_x Z_r^{t,x}\right] dB_r + \partial_x Z_u^{t,x}, \text{ and}$$

$$\lim_{s\downarrow u} D_u\left[\partial_x Y_s^{t,x}\right] = \partial_x Z_u^{t,x} \quad P^*\text{-a.s.}$$

We have,

$$|\partial_x Z_u^{t,x}| \le K \quad du \times dP^*\text{-}a.s.$$

Thus, from (6.9), by standard estimates for BSDEs again, $\forall p \ge 2$, $\forall x, x' \in \mathbb{R}$,

$$\begin{aligned}
&E_{P^*}\left[\sup_{s\in[t,1]} |\partial_x^2 Y_s^{t,x} - \partial_x^2 Y_s^{t,x'}|^p |\mathcal{F}_t\right] + E_{P^*}\left[\left(\int_t^1 |\partial_x^2 Z_r^{t,x} - \partial_x^2 Z_r^{t,x'}|^2 dr\right)^{\frac{p}{2}} |\mathcal{F}_t\right]\\
\le& C_p E_{P^*}\left[|\varphi''(x+B_1-B_t) - \varphi''(x'+B_1-B_t)|^p|\mathcal{F}_t\right]\\
&+ C_p E_{P^*}\left[\left(\int_t^1 |g_\epsilon''(Z_r^{t,x})(\partial_x Z_r^{t,x})^2 - g_\epsilon''(Z_r^{t,x'})(\partial_x Z_r^{t,x'})^2| dr\right)^p |\mathcal{F}_t\right]\\
&+ C_p E_{P^*}\left[\left(\int_t^1 |g_\epsilon'(Z_r^{t,x}) - g_\epsilon'(Z_r^{t,x'})||\partial_x^2 Z_r^{t,x}| dr\right)^p |\mathcal{F}_t\right]\\
\equiv& I_1 + I_2 + I_3,
\end{aligned}$$

where C_p is a constant independent of t, and I_1, I_2, I_3 satisfied

$$\begin{aligned}
I_1 =& C_p E_{P^*}\left[|\varphi''(x+B_1-B_t) - \varphi''(x'+B_1-B_t)|^p|\mathcal{F}_t\right]\\
\le& C_p\|\varphi'''\|^p|x-x'|^p = C_{1,p}|x-x'|^p, \quad (\text{where } C_{1,p} = C_p\|\varphi'''\|^p)\\
I_2 =& C_p E_{P^*}\Bigg[\Bigg(\int_t^1 \Big| g_\epsilon''(Z_r^{t,x})(\partial_x Z_r^{t,x})(\partial_x Z_r^{t,x} - \partial_x Z_r^{t,x'})\\
&\qquad + (g_\epsilon''(Z_r^{t,x}) - g_\epsilon''(Z_r^{t,x'}))(\partial_x Z_r^{t,x})(\partial_x Z_r^{t,x'})\\
&\qquad + g_\epsilon''(Z_r^{t,x'})(\partial_x Z_r^{t,x'})(\partial_x Z_r^{t,x} - \partial_x Z_r^{t,x'})\Big| dr\Bigg)^p |\mathcal{F}_t\Bigg]\\
\le& 2\cdot 3^{p-1} C_p (\tfrac{\kappa}{\epsilon})^p K^p E_{P^*}\left[\left(\int_t^1 |\partial_x Z_r^{t,x} - \partial_x Z_r^{t,x'}| dr\right)^p |\mathcal{F}_t\right]\\
&+ 3^{p-1} C_p K^{2p} (\tfrac{3\kappa}{\epsilon^2})^p E_{P^*}\left[\left(\int_t^1 |Z_r^{t,x} - Z_r^{t,x'}| dr\right)^p |\mathcal{F}_t\right]\\
\le& C_{2,p}|x-x'|^p,
\end{aligned}$$

where $C_{2,p}$ is a constant depend on $\kappa, \epsilon, p, \|\varphi'\|$ and $\|\varphi''\|$,

$$\begin{aligned}
I_3 =& C_p E_{P^*}\left[\left(\int_t^1 |g_\epsilon'(Z_r^{t,x}) - g_\epsilon'(Z_r^{t,x'})||\partial_x^2 Z_r^{t,x}| dr\right)^p |\mathcal{F}_t\right]\\
\le& C_p(\tfrac{\kappa}{\epsilon})^p E_{P^*}\left[\left(\int_t^1 |Z_r^{t,x} - Z_r^{t,x'}||\partial_x^2 Z_r^{t,x}| dr\right)^p |\mathcal{F}_t\right]\\
\le& C_p(\tfrac{\kappa}{\epsilon})^p \Big(E_{P^*}\Big[(\int_t^1 |Z_r^{t,x} - Z_r^{t,x'}|^2 |dr)^p|\mathcal{F}_t\Big]\Big)^{\frac{1}{2}} \Big(E_{P^*}\Big[(\int_t^1 |\partial_x^2 Z_r^{t,x}|^2 dr)^p|\mathcal{F}_t\Big]\Big)^{\frac{1}{2}}
\end{aligned}$$

$$\leq C_{3,p}|x-x'|^p,$$

where $C_{3,p}$ is a constant depend on $\kappa, \epsilon, p, \|\varphi'\|$ and $\|\varphi''\|$. Therefore,

$$E_{P^*}\left[\sup_{s\in[t,1]}|\partial_x^2 Y_s^{t,x}-\partial_x^2 Y_s^{t,x'}|^p|\mathcal{F}_t\right]+E_{P^*}\left[\left(\int_t^1|\partial_x^2 Z_r^{t,x}-\partial_x^2 Z_r^{t,x'}|^2dr\right)^{\frac{p}{2}}|\mathcal{F}_t\right]$$
$$\leq (C_{1,p}+C_{2,p}+C_{3,p})|x-x'|^p$$

Thus we obtain Claims (1) and (2).
(3): Follows from Peng's dynamic programming principle [29, Theorem 3.2].
(4): It is a direct consequence of [6, Theorem 3.1]. ■

The following lemma is adapted from [2, Proposition 2.3].

Lemma 6.2. *Suppose that (b_t) and (σ_t) are two continuous, bounded $\mathcal{F}_t$-adapted processes and that (X_t) is of the form*

$$X_t = x+\int_0^t b_s ds+\int_0^t \sigma_s dB_s, \quad x\in\mathbb{R}.$$

Then

$$\lim_{n\to\infty} n\sup_{x\in\mathbb{R}}\left|\mathbb{E}_{g_\epsilon}\left[X_{\frac{1}{n}}\right]-x-\frac{1}{n}g_\epsilon(\sigma_0)-\frac{1}{n}b_0\right|=0.$$

The next lemma is an immediate consequence.

Lemma 6.3. *For any $\varphi\in C_b^2(\mathbb{R})$,*

$$\lim_{n\to\infty} n\sup_{x\in\mathbb{R}}\left|\mathbb{E}_{g_\epsilon}\left[\varphi\left(x+B_{\frac{1}{n}}\right)\right]-\varphi(x)-\frac{1}{n}g_\epsilon(\varphi'(x))-\frac{1}{2n}\varphi''(x)\right|=0. \quad (6.10)$$

Proof: Let $X_s\equiv\varphi(x+B_s)$. By Ito's formula,

$$X_t=\varphi(x)+\frac{1}{2}\int_0^t\varphi''(x+B_s)ds+\int_0^t\varphi'(x+B_s)dB_s.$$

Apply Lemma 6.2 to complete the proof. ■

Lemma 6.4. *Let g_0 be defined by (6.4). For any $\varphi\in C_b^3(\mathbb{R})$, let $\{H_{m,n}\}_{m=0}^n$ be the functions defined in (6.5). Define functions $\{L_{m,n}\}_{m=0}^n$ by*

$$L_{m,n}(x)=H_{m,n}(x)+\frac{1}{n}g_0(H'_{m,n}(x))+\frac{1}{2n}H''_{m,n}(x). \quad (6.11)$$

Let $\{T_{m,n}\}_{m,n\geq 0}$ be an array of r.v.s satisfying

$$T_{0,n}=0, \text{ and } T_{m,n}\in\mathcal{H} \text{ is } \mathcal{G}_m\text{-measurable for all } m\geq 1,\ n\geq 1,$$

and, for any $Q \in \mathcal{P}$, set $Y_m^Q = \frac{1}{\sigma}(X_m - E_Q[X_m|\mathcal{G}_{m-1}])$. Then

$$\lim_{n\to\infty}\sum_{m=1}^{n}\left|\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{Y_m^Q}{\sqrt{n}}\right)\right]-\sup_{Q\in\mathcal{P}} E_Q\left[L_{m,n}(T_{m-1,n})\right]\right|=0. \tag{6.12}$$

Remark 6.5. *In Theorem 5.1, the sample average term is absent, and accordingly its proof involves a counterpart of this lemma where the term $\frac{X_m}{n}$ is deleted above. Then the proof of (6.12), so modified, simplifies, in particular, rectangularity is no longer needed and the generators g_ϵ in (6.3) and g_0 in (6.11) can be set equal to* 0*. (Appendix A.5 provides some details.)*

Proof: We proceed in two steps.

Step 1: We first give a remainder estimate that will also be used later in the proof of Lemma 6.8. Let $\{\theta_m\}_{m\geq 1}$ be a sequence of $\mathcal{G}_{m-1}$-measurable random variables satisfying

$$|\theta_m| \leq \kappa, \text{ for } m \geq 1.$$

We prove that

$$\lim_{n\to\infty}\sum_{m=1}^{n}\left|\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}(T_{m-1,n}+\frac{X_m}{n}+\frac{X_m-\theta_m}{\sigma\sqrt{n}})\right]\right.$$
$$\left. -\sup_{Q\in\mathcal{P}} E_Q\left[F(\theta_m,m,n)\right]\right|=0, \tag{6.13}$$

where $F(\theta_m, m, n) \equiv$

$$H_{m,n}(T_{m-1,n})+H'_{m,n}(T_{m-1,n})\left(\frac{X_m}{n}+\frac{X_m-\theta_m}{\sigma\sqrt{n}}\right)+\frac{1}{2}H''_{m,n}(T_{m-1,n})\left(\frac{X_m-\theta_m}{\sigma\sqrt{n}}\right)^2. \tag{6.14}$$

By Lemma 6.1, $\exists C > 0$ such that (for all m and n),

$$\sup_{m\leq n}\sup_{x\in\mathbb{R}}|H''_{m,n}(x)| \leq C \quad \text{and} \quad \sup_{m\leq n}\sup_{x,y\in\mathbb{R},x\neq y}\frac{|H''_{m,n}(x)-H''_{m,n}(y)|}{|x-y|}\leq C.$$

By the Taylor expansion of $H_{m,n} \in C_b^2(\mathbb{R})$, $\forall\bar{\epsilon} > 0$, $\exists\delta > 0$ (δ depends only on C and $\bar{\epsilon}$), such that $\forall x, y \in \mathbb{R}$, and all $n \geq m \geq 1$,

$$\left|H_{m,n}(x+y)-H_{m,n}(x)-H'_{m,n}(x)y-\frac{1}{2}H''_{m,n}(x)y^2\right| \leq \bar{\epsilon}|y|^2 I_{\{|y|<\delta\}}+C|y|^2 I_{\{|y|\geq\delta\}}. \tag{6.15}$$

Let $x = T_{m-1,n}$ and $y = \frac{X_m}{n} + \frac{X_m-\theta_m}{\sigma\sqrt{n}}$ in (6.15), and obtain

$$\sum_{m=1}^{n}\left|\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{X_m-\theta_m}{\sigma\sqrt{n}}\right)\right]-\sup_{Q\in\mathcal{P}} E_Q\left[F(\theta_m,m,n)\right]\right|$$

$$
\begin{aligned}
&\le \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[\left| H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{X_m-\theta_m}{\sigma\sqrt{n}} \right) - F(\theta_m, m, n) \right| \right] \\
&\le R_1(\overline{\epsilon}, n) + R_2(C,n) + R_3(C,n), \quad \text{where}
\end{aligned}
$$

$$
\begin{aligned}
R_1(\overline{\epsilon}, n) &:= \overline{\epsilon} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[\left| \tfrac{X_m}{n} + \tfrac{X_m-\theta_m}{\sigma\sqrt{n}} \right|^2 I_{\left\{ \left| \frac{X_m}{n} + \frac{X_m-\theta_m}{\sigma\sqrt{n}} \right| < \delta \right\}} \right], \\
R_2(C, n) &:= C \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[\left| \tfrac{X_m}{n} + \tfrac{X_m-\theta_m}{\sigma\sqrt{n}} \right|^2 I_{\left\{ \left| \frac{X_m}{n} + \frac{X_m-\theta_m}{\sigma\sqrt{n}} \right| \ge \delta \right\}} \right], \\
R_3(C, n) &:= \frac{C}{2} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[\left| \tfrac{X_m}{n} \right|^2 + 2 \left| \tfrac{X_m}{n} \right| \left| \tfrac{X_m-\theta_m}{\sigma\sqrt{n}} \right| \right].
\end{aligned}
$$

It is readily proven that, for sufficiently large n,

$$
\begin{aligned}
R_1(\overline{\epsilon}, n) \le& \frac{2\overline{\epsilon}}{n^2} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[|X_m|^2 \right] + \frac{2\overline{\epsilon}}{n} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[\left| \tfrac{X_m-\theta_m}{\sigma} \right|^2 \right] \\
\le& \frac{4\overline{\epsilon}}{n} \left(\sigma^2 + \kappa^2 \right) + \frac{4\overline{\epsilon}}{\sigma^2} \left(\sigma^2 + 4\kappa^2 \right), \\
R_2(C, n) \le& 2C \left(\frac{1}{n} + \frac{1}{\sigma\sqrt{n}} \right)^2 \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[|X_m|^2 I_{\left\{ \left| \frac{X_m}{n} + \frac{X_m-\theta_m}{\sigma\sqrt{n}} \right| \ge \delta \right\}} \right] \\
&+ \frac{2C}{\sigma^2 n} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[|\theta_m|^2 I_{\left\{ \left| \frac{X_m}{n} + \frac{X_m-\theta_m}{\sigma\sqrt{n}} \right| \ge \delta \right\}} \right] \\
\le& \frac{2C}{\sigma^2} \frac{(\sigma+\sqrt{n})^2}{n^2} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[|X_m|^2 I_{\left\{ |X_m| > \frac{\sigma n}{\sigma+\sqrt{n}}\delta - \kappa \right\}} \right] \\
&+ \frac{2C}{\sigma^2 n} \frac{\kappa^2}{\delta^2} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[\left| \tfrac{X_m}{n} + \tfrac{X_m-\theta_m}{\sigma\sqrt{n}} \right|^2 \right], \\
R_3(C, n) \le& \frac{C}{2n^2} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[|X_m|^2 \right] + \frac{C}{n^{3/2}\sigma} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[|X_m| \, |X_m - \theta_m| \right] \\
\le& \left(\frac{C}{n} + \frac{2C}{\sqrt{n}\sigma} \right) \left(\sigma^2 + \kappa^2 \right) + \frac{2C\kappa}{\sqrt{n}\sigma} \sqrt{\sigma^2 + \kappa^2}.
\end{aligned}
$$

By the finiteness of κ, σ and the Lindeberg condition (2.1),

$$
\lim_{\overline{\epsilon}\to 0} \lim_{n\to\infty} \left(R_1(\overline{\epsilon}, n) + R_2(C,n) + R_3(C,n) \right) = 0,
$$

which proves (6.13).

Step 2: To prove (6.12), it suffices to prove that if we take $\theta_m = E_Q[X_m|\mathcal{G}_{m-1}]$ in (6.14), then

$$\sup_{Q\in\mathcal{P}} E_Q\left[F(\theta_m, m, n)\right] = \sup_{Q\in\mathcal{P}} E_Q\left[L_{m,n}(T_{m-1,n})\right], \ \forall n \geq m \geq 1.$$

In fact, if $\theta_m = E_Q[X_m|\mathcal{G}_{m-1}]$, then by a generalization of Lemma 2.2(iii) (in the proof of Theorem 4.1, we shall take $T_{m-1,n} = \frac{S_{m-1}}{n} + \frac{S^Q_{m-1}}{\sqrt{n}}$, in which case part (iv) of Lemma 2.2 suffices),

$$\begin{aligned}
&\sup_{Q\in\mathcal{P}} E_Q\left[F(\theta_m, m, n)\right] \\
=& \sup_{Q\in\mathcal{P}} E_Q\Bigg[H_{m,n}(T_{m-1,n}) + H'_{m,n}(T_{m-1,n})\left(\frac{X_m}{n} + \frac{Y^Q_m}{\sqrt{n}}\right) \\
&\qquad\qquad + \frac{1}{2}H''_{m,n}(T_{m-1,n})\left(\frac{Y^Q_m}{\sqrt{n}}\right)^2\Bigg] \\
=& \sup_{Q\in\mathcal{P}} E_Q\Bigg[H_{m,n}(T_{m-1,n}) + H'_{m,n}(T_{m-1,n})E_Q\left[\left(\frac{X_m}{n} + \frac{Y^Q_m}{\sqrt{n}}\right)|\mathcal{G}_{m-1}\right] \\
&\qquad\qquad + \frac{1}{2n}H''_{m,n}(T_{m-1,n})E_Q\left[\left(Y^Q_m\right)^2|\mathcal{G}_{m-1}\right]\Bigg] \\
=& \sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}(T_{m-1,n}) + \frac{1}{n}H'_{m,n}(T_{m-1,n})X_m + \frac{1}{2n}H''_{m,n}(T_{m-1,n})\right] \\
=& \sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}(T_{m-1,n}) + \frac{1}{n}\,\mathbb{E}[H'_{m,n}(T_{m-1,n})X_m|\mathcal{G}_{m-1}] + \frac{1}{2n}H''_{m,n}(T_{m-1,n})\right] \\
=& \sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}(T_{m-1,n}) + \frac{1}{n}g_0\left(H'_{m,n}(T_{m-1,n})\right) + \frac{1}{2n}H''_{m,n}(T_{m-1,n})\right] \\
=& \sup_{Q\in\mathcal{P}} E_Q\left[L_{m,n}(T_{m-1,n})\right].
\end{aligned}$$

Combine with (6.13) to complete the proof. ■

The next three lemmas consider the special implications of symmetry and thus relate to the proof of Theorem 4.3.

Lemma 6.6 ([3]). *Let $\varphi \in C^3_b(\mathbb{R})$ be symmetric with center $c \in \mathbb{R}$, and $v(t,x)$ be the unique solution of Cauchy's problem for the parabolic equation*

$$\begin{cases} \partial_t v(t,x) &= \frac{1}{2}\partial^2_{xx}v(t,x) + g_\epsilon(\partial_x v(t,x)) \\ v(0,x) &= \varphi(x). \end{cases} \tag{6.16}$$

(1) *For any* $t \geq 0$, $v(t, \cdot)$ *is symmetric with center* c.

(2) *If* $sgn(\varphi'(x)) = -sgn(x - c)$, *then, for any* $t \geq 0$,

$$sgn\left(\partial_x v(t, x)\right) = -sgn(x - c).$$

(3) *If* $sgn(\varphi'(x)) = sgn(x - c)$, *then, for any* $t \geq 0$,

$$sgn\left(\partial_x v(t, x)\right) = sgn(x - c).$$

Lemma 6.7. *Let* $\varphi \in C_b^3(\mathbb{R})$ *be symmetric with center* $c \in \mathbb{R}$. *Then the functions* $\{H_{m,n}\}_{m=0}^n$ *defined in (6.5) satisfy, for any* n *and* $m = 0, \cdots, n$:

(1) $H_{m,n}$ *is symmetric with center* c.

(2) *If* $sgn(\varphi'(x)) = -sgn(x - c)$, *then*

$$sgn(H'_{m,n}(x)) = -sgn(x - c), \quad \text{and} \quad H''_{m,n}(c) \leq 0.$$

(3) *If* $sgn(\varphi'(x)) = sgn(x - c)$, *then*

$$sgn(H'_{m,n}(x)) = sgn(x - c), \quad \text{and} \quad H''_{m,n}(c) \geq 0.$$

Proof: By the definition of $H_{m,n}(x)$ via (6.5) and the nonlinear Feynman-Kac formula, we know that $H_{m,n}(x) = v(1 - \frac{m}{n}, x)$, where $v(t, x)$ is the solution of equation (6.16). Then (1)-(3) follows from Lemma 6.6. ■

Lemma 6.8. *Adopt the assumptions and notation in (4.6), (4.7) and Lemma 6.4, and let* $\varphi \in C_b^3(\mathbb{R})$ *be symmetric with center* $c \in \mathbb{R}$.

(1) *If* $sgn(\varphi'(x)) = -sgn(x - c)$, *and if* Y_m^Q *in (6.12) is replaced by* Z_m^n, *where*

$$Z_m^n = \frac{1}{\sigma}(X_m - \mu_m^n), \ \mu_m^n = \kappa I_{A_{m-1,n}} - \kappa I_{A_{m-1,n}^c}, \ A_{m-1,n} = \{T_{m-1,n} \leq c\},$$

then

$$\lim_{n\to\infty} \sum_{m=1}^n \left| \sup_{Q\in\mathcal{P}} E_Q \left[H_{m,n} \left(T_{m-1,n} + \frac{X_m}{n} + \frac{Z_m^n}{\sqrt{n}} \right) \right] - \sup_{Q\in\mathcal{P}} E_Q \left[L_{m,n} \left(T_{m-1,n} \right) \right] \right| = 0. \quad (6.17)$$

(2) If $sgn(\varphi'(x)) = sgn(x-c)$, and if Y_m^Q in (6.12) is replaced by $\widetilde{Z}_m^n$, where

$$\widetilde{Z}_m^n = \frac{1}{\sigma}(X_m - \widetilde{\mu}_m^n),\ \widetilde{\mu}_m^n = \kappa I_{\widetilde{A}_{m-1,n}} - \kappa I_{\widetilde{A}_{m-1,n}^c},\ \widetilde{A}_{m-1,n} = \{T_{m-1,n} \geq c\},$$

then

$$\liminf_{n\to\infty} \sum_{m=1}^n \left\{ \sup_{Q\in\mathcal{P}} E_Q \left[H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{\widetilde{Z}_m^n}{\sqrt{n}} \right) \right] - \sup_{Q\in\mathcal{P}} E_Q \left[L_{m,n}\left(T_{m-1,n}\right) \right] \right\} \geq 0. \quad (6.18)$$

Furthermore, if

$$\lim_{\bar{\delta}\to 0} \limsup_{n\to\infty} \frac{1}{n} \sum_{m=1}^n \sup_{Q\in\mathcal{P}} E_Q \left[\left| E_Q[X_m|\mathcal{G}_{m-1}] - \widetilde{\mu}_m^n \right| I_{\{|T_{m-1,n}-c|\leq\bar{\delta}\}} \right] = 0, \quad (6.19)$$

then

$$\lim_{n\to\infty} \sum_{m=1}^n \left| \sup_{Q\in\mathcal{P}} E_Q \left[H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{\widetilde{Z}_m^n}{\sqrt{n}} \right) \right] - \sup_{Q\in\mathcal{P}} E_Q \left[L_{m,n}\left(T_{m-1,n}\right) \right] \right| = 0. \quad (6.20)$$

Remark 6.9. *The lemma is valid also if $c = \pm\infty$. Taking $c = +\infty$ in (1) means that φ is increasing on $\mathbb{R}$. Then $A_{m-1,n} = \Omega$ and $\mu_m^n = \kappa$ for any $1 \leq m \leq n$. If $c = -\infty$ in (1), then φ is decreasing on $\mathbb{R}$, $A_{m-1,n} = \varnothing$ and $\mu_m^n = -\kappa$. Similarly for (2).*

Proof of (1): We proceed in two steps.
Step 1: Firstly, we prove

$$\limsup_{n\to\infty} \sum_{m=1}^n \left\{ \sup_{Q\in\mathcal{P}} E_Q \left[H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Z_m^n}{\sqrt{n}} \right) \right] - \sup_{Q\in\mathcal{P}} E_Q \left[L_{m,n}\left(T_{m-1,n}\right) \right] \right\} \leq 0. \quad (6.21)$$

By Lemma 6.4, we only need to prove the non-positivity of

$$\limsup_{n\to\infty} \sum_{m=1}^n \left\{ \sup_{Q\in\mathcal{P}} E_Q \left[H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Z_m^n}{\sqrt{n}} \right) \right] \right.$$

$$-\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{Y_m^Q}{\sqrt{n}}\right)\right]\right\}.$$

For any $\overline{\delta}>0$, we set

$$D^{\overline{\delta},1}_{m-1,n}=\{T_{m-1,n}>c+\overline{\delta}\},\ D^{\overline{\delta},2}_{m-1,n}=\{T_{m-1,n}<c-\overline{\delta}\},$$

$$D^{\overline{\delta},3}_{m-1,n}=\{|T_{m-1,n}-c|\le\overline{\delta}\},\ N^{\overline{\delta}}_{m,n}=\left\{|X_m|\le\tfrac{\sigma n}{\sigma+\sqrt{n}}\overline{\delta}-\kappa\right\}.$$

For any $\omega\in D^{\overline{\delta},1}_{m-1,n}\cap N^{\overline{\delta}}_{m,n}$, we have

$$\begin{aligned}c&<\left(T_{m-1,n}+\frac{X_m}{n}+\frac{X_m-E_Q[X_m|\mathcal{G}_{m-1}]}{\sqrt{n}\sigma}\right)(\omega)\\&\le\left(T_{m-1,n}+\frac{X_m}{n}+\frac{X_m+\kappa}{\sqrt{n}\sigma}\right)(\omega).\end{aligned}$$

By Lemma 6.7, $H_{m,n}$ is decreasing on $(c,+\infty)$. Thus

$$\begin{aligned}&I_{D^{\overline{\delta},1}_{m-1,n}\cap N^{\overline{\delta}}_{m,n}}H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{Y_m^Q}{\sqrt{n}}\right)\\\ge&I_{D^{\overline{\delta},1}_{m-1,n}\cap N^{\overline{\delta}}_{m,n}}H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{Z_m^n}{\sqrt{n}}\right).\end{aligned}$$

Also, for any $\omega\in D^{\overline{\delta},2}_{m-1,n}\cap N^{\overline{\delta}}_{m,n}$,

$$\begin{aligned}c&>\left(T_{m-1,n}+\frac{X_m}{n}+\frac{X_m-E_Q[X_m|\mathcal{G}_{m-1}]}{\sqrt{n}\sigma}\right)(\omega)\\&\ge\left(T_{m-1,n}+\frac{X_m}{n}+\frac{X_m-\kappa}{\sqrt{n}\sigma}\right)(\omega).\end{aligned}$$

By Lemma 6.7, $H_{m,n}$ is increasing on $(-\infty,c)$. Thus

$$\begin{aligned}&I_{D^{\overline{\delta},2}_{m-1,n}\cap N^{\overline{\delta}}_{m,n}}H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{Y_m^Q}{\sqrt{n}}\right)\\\ge&I_{D^{\overline{\delta},2}_{m-1,n}\cap N^{\overline{\delta}}_{m,n}}H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{Z_m^n}{\sqrt{n}}\right).\end{aligned}$$

For $F(\theta_m,m,n)$ defined in (6.14), we have

$$H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{Z_m^n}{\sqrt{n}}\right)-H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{Y_m^Q}{\sqrt{n}}\right)$$

$$
\begin{aligned}
=&F(\mu_m^n, m, n) - F\left(E_Q[X_m|\mathcal{G}_{m-1}], m, n\right) \\
&+ H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Z_m^n}{\sqrt{n}}\right) - F(\mu_m^n, m, n) \\
&- \left(H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Y_m^Q}{\sqrt{n}}\right) - F\left(E_Q[X_m|\mathcal{G}_{m-1}], m, n\right)\right).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\sum_{m=1}^{n}\left\{\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Z_m^n}{\sqrt{n}}\right)\right]\right. \\
&\qquad\qquad\qquad \left. - \sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Y_m^Q}{\sqrt{n}}\right)\right]\right\} \\
\leq &\sum_{m=1}^{n}\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Z_m^n}{\sqrt{n}}\right) - H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Y_m^Q}{\sqrt{n}}\right)\right] \\
\leq &\sum_{m=1}^{n}\sup_{Q\in\mathcal{P}} E_Q\left[I_{D_{m-1,n}^{\bar{\delta},3}}\left[H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Z_m^n}{\sqrt{n}}\right)\right.\right. \\
&\qquad\qquad\qquad \left.\left. -H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Y_m^Q}{\sqrt{n}}\right)\right]\right] \\
&+ \sum_{m=1}^{n}\sup_{Q\in\mathcal{P}} E_Q\left[I_{\{|X_m| > \frac{\sigma n}{\sigma+\sqrt{n}}\bar{\delta}-\kappa\}}\left|H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Z_m^n}{\sqrt{n}}\right)\right.\right. \\
&\qquad\qquad\qquad \left.\left. -H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Y_m^Q}{\sqrt{n}}\right)\right|\right] \\
\leq &I_n^1 + I_n^2 + I_n^3,
\end{aligned}
$$

where I_n^1, I_n^2, I_n^3 are defined by

$$
\begin{aligned}
I_n^1 &\equiv \sum_{m=1}^{n}\sup_{Q\in\mathcal{P}} E_Q\left[I_{D_{m-1,n}^{\bar{\delta},3}}\left(F(\mu_m^n, m, n) - F\left(E_Q[X_m|\mathcal{G}_{m-1}], m, n\right)\right)\right], \\
I_n^2 &\equiv \sum_{m=1}^{n}\sup_{Q\in\mathcal{P}} E_Q\left[2\|\varphi\| I_{\{|X_m| > \frac{\sigma n}{\sigma+\sqrt{n}}\bar{\delta}-\kappa\}}\right] \qquad (\text{ where } \|\varphi\| = \sup_{x\in\mathbb{R}}\varphi(x)), \\
I_n^3 &\equiv \sum_{m=1}^{n}\sup_{Q\in\mathcal{P}} E_Q\left[\left|H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Z_m^n}{\sqrt{n}}\right) - F(\mu_m^n, m, n)\right|\right] \\
&\quad + \sum_{m=1}^{n}\sup_{Q\in\mathcal{P}} E_Q\left[\left|H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Y_m^Q}{\sqrt{n}}\right) - F\left(E_Q[X_m|\mathcal{G}_{m-1}], m, n\right)\right|\right].
\end{aligned}
$$

By the Lindeberg condition (2.1), $\lim_{n\to\infty} I_n^2 = 0$, and by the remainder estimate in the proof of Lemma 6.4, $\lim_{n\to\infty} I_n^3 = 0$. Thus it suffices to show that $\lim_{n\to\infty} I_n^1 = 0$, which is proven as follows. From Lemma 6.7(2), $H''_{m,n}(c) \leq 0$. Therefore,

$$
\begin{aligned}
I_n^1 &= \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[I_{D^{\bar{\delta},3}_{m-1,n}} \left(F(\mu_m^n, m, n) - F\left(E_Q[X_m|\mathcal{G}_{m-1}], m, n\right)\right)\right] \\
&\leq \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[I_{D^{\bar{\delta},3}_{m-1,n}} H''_{m,n}(T_{m-1,n}) \frac{(E_Q[X_m|\mathcal{G}_{m-1}] - \mu_m^n)^2}{2n\sigma^2}\right] \\
&\leq \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[I_{D^{\bar{\delta},3}_{m-1,n}} \left(H''_{m,n}(T_{m-1,n}) - H''_{m,n}(c)\right) \frac{(E_Q[X_m|\mathcal{G}_{m-1}] - \mu_m^n)^2}{2n\sigma^2}\right] \\
&\quad + \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[I_{D^{\bar{\delta},3}_{m-1,n}} H''_{m,n}(c) \frac{(E_Q[X_m|\mathcal{G}_{m-1}] - \mu_m^n)^2}{2n\sigma^2}\right] \\
&\leq \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[I_{D^{\bar{\delta},3}_{m-1,n}} L|T_{m-1,n} - c| \frac{(E_Q[X_m|\mathcal{G}_{m-1}] - \mu_m^n)^2}{2n\sigma^2}\right] \leq \frac{L2\kappa^2}{\sigma^2}\bar{\delta},
\end{aligned}
$$

where L is the uniform Lipschitz constant for $H''_{m,n}$ given in Lemma 6.1.

Step 2: Next we prove

$$
\begin{aligned}
\liminf_{n\to\infty} \sum_{m=1}^{n} \Bigg\{ \sup_{Q\in\mathcal{P}} E_Q & \left[H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Z_m^n}{\sqrt{n}}\right)\right] \\
& - \sup_{Q\in\mathcal{P}} E_Q \left[L_{m,n}\left(T_{m-1,n}\right)\right] \Bigg\} \geq 0. \qquad (6.22)
\end{aligned}
$$

By the remainder estimate in the proof of Lemma 6.4, it suffices to take $\theta_m = \mu_m^n$ in (6.14) and to show that

$$
\sup_{Q\in\mathcal{P}} E_Q\left[F(\theta_m, m, n)\right] \geq \sup_{Q\in\mathcal{P}} E_Q\left[L_{m,n}\left(T_{m-1,n}\right)\right], \quad n \geq m \geq 1.
$$

By Lemma A.1, there exist $\{Q_k\}_{k\geq 1}, \{\overline{P}_j^m\}_{j\geq 1}, \{\underline{P}_j^m\}_{j\geq 1} \subset \mathcal{P}$ such that

$$
\lim_{j\to\infty} E_{\overline{P}_j^m}[X_m|\mathcal{G}_{m-1}] = \kappa, \ \lim_{j\to\infty} E_{\underline{P}_j^m}[X_m|\mathcal{G}_{m-1}] = -\kappa,
$$

and

$$
\begin{aligned}
&\sup_{Q\in\mathcal{P}} E_Q \left[H_{m,n}(T_{m-1,n}) + \frac{1}{n} g_0\left(H'_{m,n}(T_{m-1,n})\right) + \frac{1}{2n} H''_{m,n}(T_{m-1,n})\right] \\
&= \lim_{k\to\infty} E_{Q_k} \left[H_{m,n}(T_{m-1,n}) + \frac{1}{n} g_0\left(H'_{m,n}(T_{m-1,n})\right) + \frac{1}{2n} H''_{m,n}(T_{m-1,n})\right].
\end{aligned}
$$

By Lemma 2.2(ii), there exist $\{R_j^m\}_{j\geq 1}\subset\mathcal{P}$ satisfying

$$E_{R_j^m}[X_m|\mathcal{G}_{m-1}] = I_{A_{m-1,n}} E_{\overline{P}_j^m}[X_m|\mathcal{G}_{m-1}] + I_{A_{m-1,n}^c} E_{\underline{P}_j^m}[X_m|\mathcal{G}_{m-1}].$$

By Lemma 2.2(iii) and the dominated convergence theorem,

$$\begin{aligned}
&\sup_{Q\in\mathcal{P}} E_Q\left[L_{m,n}(T_{m-1,n})\right] \\
=&\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}(T_{m-1,n}) + \frac{1}{n} g_0\left(H'_{m,n}(T_{m-1,n})\right) + \frac{1}{2n} H''_{m,n}(T_{m-1,n})\right] \\
=&\lim_{k\to\infty} E_{Q_k}\left[H_{m,n}(T_{m-1,n}) + \frac{1}{n} g_0\left(H'_{m,n}(T_{m-1,n})\right) + \frac{1}{2n} H''_{m,n}(T_{m-1,n})\right] \\
=&\lim_{k\to\infty} E_{Q_k}\Bigg[H_{m,n}(T_{m-1,n}) + I_{A_{m-1,n}} H'_{m,n}(T_{m-1,n})\frac{\kappa}{n} \\
&\qquad\qquad + I_{A_{m-1,n}^c} H'_{m,n}(T_{m-1,n})\frac{-\kappa}{n} + \frac{1}{2n} H''_{m,n}(T_{m-1,n})\Bigg] \\
=&\lim_{k\to\infty} E_{Q_k}\Bigg[H_{m,n}(T_{m-1,n}) \\
&\qquad + \lim_{j\to\infty} E_{\overline{P}_j^m}\left[I_{A_{m-1,n}} H'_{m,n}(T_{m-1,n})\left(\frac{X_m}{n} + \frac{X_m-\kappa}{\sigma\sqrt{n}}\right)|\mathcal{G}_{m-1}\right] \\
&\qquad + \lim_{j\to\infty} E_{\overline{P}_j^m}\left[I_{A_{m-1,n}} H''_{m,n}(T_{m-1,n})\frac{(X_m-\kappa)^2}{2n\sigma^2}|\mathcal{G}_{m-1}\right] \\
&\qquad + \lim_{j\to\infty} E_{\underline{P}_j^m}\left[I_{A_{m-1,n}^c} H'_{m,n}(T_{m-1,n})\left(\frac{X_m}{n} + \frac{X_m+\kappa}{\sigma\sqrt{n}}\right)|\mathcal{G}_{m-1}\right] \\
&\qquad + \lim_{j\to\infty} E_{\underline{P}_j^m}\left[I_{A_{m-1,n}^c} H''_{m,n}(T_{m-1,n})\frac{(X_m+\kappa)^2}{2n\sigma^2}|\mathcal{G}_{m-1}\right]\Bigg] \\
=&\lim_{k\to\infty}\lim_{j\to\infty} E_{Q_k}\Bigg[H_{m,n}(T_{m-1,n}) + E_{R_j^m}\Bigg[H'_{m,n}(T_{m-1,n})\left(\frac{X_m}{n} + \frac{Z_m^n}{\sqrt{n}}\right) \\
&\qquad\qquad + \frac{1}{2n} H''_{m,n}(T_{m-1,n})(Z_m^n)^2|\mathcal{G}_{m-1}\Bigg]\Bigg] \\
\leq&\sup_{Q\in\mathcal{P}} E_Q\Bigg[H_{m,n}(T_{m-1,n}) + ess\sup_{R\in\mathcal{P}} E_R\Bigg[H'_{m,n}(T_{m-1,n})\left(\frac{X_m}{n} + \frac{Z_m^n}{\sqrt{n}}\right) \\
&\qquad\qquad + \frac{1}{2n} H''_{m,n}(T_{m-1,n})(Z_m^n)^2|\mathcal{G}_{m-1}\Bigg]\Bigg] \\
=&\sup_{Q\in\mathcal{P}} E_Q\Bigg[H_{m,n}(T_{m-1,n}) + H'_{m,n}(T_{m-1,n})\left(\frac{X_m}{n} + \frac{Z_m^n}{\sqrt{n}}\right) \\
&\qquad\qquad + \frac{1}{2n} H''_{m,n}(T_{m-1,n})(Z_m^n)^2\Bigg].
\end{aligned}$$

Combined with (6.13), this implies (6.22), thus completing the proof of (1).

Proof of (2): Proof of inequality (6.18) is similar to that of (6.22).

To prove (6.20), assuming (6.19), we need only prove

$$\limsup_{n\to\infty}\sum_{m=1}^{n}\left\{\sup_{Q\in\mathcal{P}}E_Q\left[H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{\widetilde{Z}_m^n}{\sqrt{n}}\right)\right]\right.$$
$$\left.-\sup_{Q\in\mathcal{P}}E_Q\left[L_{m,n}\left(T_{m-1,n}\right)\right]\right\}\le 0. \quad (6.23)$$

By assumption (6.19), $\forall\varepsilon>0$, $\exists\overline{\delta}_\varepsilon>0$ such that

$$\limsup_{n\to\infty}\frac{1}{n}\sum_{m=1}^{n}\sup_{Q\in\mathcal{P}}E_Q\left[\left|E_Q[X_m|\mathcal{G}_{m-1}]-\widetilde{\mu}_m^n\right|I_{\{|T_{m-1,n}-c|\le\overline{\delta}_\varepsilon\}}\right]\le\varepsilon.$$

By Lemma 6.4, we only need to prove the non-positivity of

$$\limsup_{n\to\infty}\sum_{m=1}^{n}\left\{\sup_{Q\in\mathcal{P}}E_Q\left[H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{\widetilde{Z}_m^n}{\sqrt{n}}\right)\right]\right.$$
$$\left.-\sup_{Q\in\mathcal{P}}E_Q\left[H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{Y_m^Q}{\sqrt{n}}\right)\right]\right\}.$$

Define

$$D_{m-1,n}^{\overline{\delta}_\varepsilon,1}=\{T_{m-1,n}>c+\overline{\delta}_\varepsilon\},\ D_{m-1,n}^{\overline{\delta}_\varepsilon,2}=\{T_{m-1,n}<c-\overline{\delta}_\varepsilon\},$$

$$D_{m-1,n}^{\overline{\delta}_\varepsilon,3}=\{|T_{m-1,n}-c|\le\overline{\delta}_\varepsilon\},\ N_{m,n}^{\overline{\delta}_\varepsilon}=\left\{|X_m|\le\tfrac{\sigma n}{\sigma+\sqrt{n}}\overline{\delta}_\varepsilon-\kappa\right\}.$$

For any $\omega\in D_{m-1,n}^{\overline{\delta}_\varepsilon,1}\cap N_{m,n}^{\overline{\delta}_\varepsilon}$,

$$\begin{aligned}c&<\left(T_{m-1,n}+\frac{X_m}{n}+\frac{X_m-\kappa}{\sqrt{n}\sigma}\right)(\omega)\\&\le\left(T_{m-1,n}+\frac{X_m}{n}+\frac{X_m-E_Q[X_m|\mathcal{G}_{m-1}]}{\sqrt{n}\sigma}\right)(\omega).\end{aligned}$$

By Lemma 6.7, $H_{m,n}$ is increasing on $(c,+\infty)$. Thus

$$I_{D_{m-1,n}^{\overline{\delta}_\varepsilon,1}\cap N_{m,n}^{\overline{\delta}_\varepsilon}}H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{Y_m^Q}{\sqrt{n}}\right)$$

$$\geq I_{D^{\bar{\delta}_\varepsilon,1}_{m-1,n}\cap N^{\bar{\delta}_\varepsilon}_{m,n}} H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{\widetilde{Z}^n_m}{\sqrt{n}}\right).$$

Also, for any $\omega \in D^{\bar{\delta}_\varepsilon,2}_{m-1,n}\cap N^{\bar{\delta}_\varepsilon}_{m,n}$,

$$\begin{aligned} c &> \left(T_{m-1,n}+\frac{X_m}{n}+\frac{X_m+\kappa}{\sqrt{n}\sigma}\right)(\omega) \\ &\geq \left(T_{m-1,n}+\frac{X_m}{n}+\frac{X_m-E_Q[X_m|\mathcal{G}_{m-1}]}{\sqrt{n}\sigma}\right)(\omega). \end{aligned}$$

By Lemma 6.7, $H_{m,n}$ is decreasing on $(-\infty, c)$. Thus

$$\begin{aligned} &I_{D^{\bar{\delta}_\varepsilon,2}_{m-1,n}\cap N^{\bar{\delta}_\varepsilon}_{m,n}} H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{Y^Q_m}{\sqrt{n}}\right) \\ \geq &I_{D^{\bar{\delta}_\varepsilon,2}_{m-1,n}\cap N^{\bar{\delta}_\varepsilon}_{m,n}} H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{\widetilde{Z}^n_m}{\sqrt{n}}\right). \end{aligned}$$

Therefore,

$$\begin{aligned} &\sum_{m=1}^n \left\{\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{\widetilde{Z}^n_m}{\sqrt{n}}\right)\right]\right. \\ &\qquad\qquad \left. -\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{Y^Q_m}{\sqrt{n}}\right)\right]\right\} \\ \leq &\sum_{m=1}^n \sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{\widetilde{Z}^n_m}{\sqrt{n}}\right)\right. \\ &\qquad\qquad \left. -H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{Y^Q_m}{\sqrt{n}}\right)\right] \\ \leq &\sum_{m=1}^n \sup_{Q\in\mathcal{P}} E_Q\left[I_{D^{\bar{\delta}_\varepsilon,3}_{m-1,n}}\left[H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{\widetilde{Z}^n_m}{\sqrt{n}}\right)\right.\right. \\ &\qquad\qquad \left.\left. -H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{Y^Q_m}{\sqrt{n}}\right)\right]\right] \\ &+\sum_{m=1}^n \sup_{Q\in\mathcal{P}} E_Q\left[I_{\{|X_m|>\frac{\sigma n}{\sigma+\sqrt{n}}\bar{\delta}_\varepsilon-\kappa\}}\left|H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{\widetilde{Z}^n_m}{\sqrt{n}}\right)\right.\right. \\ &\qquad\qquad \left.\left. -H_{m,n}\left(T_{m-1,n}+\frac{X_m}{n}+\frac{Y^Q_m}{\sqrt{n}}\right)\right|\right] \end{aligned}$$

$$\le \widetilde{I}_n^1 + \widetilde{I}_n^2 + \widetilde{I}_n^3,$$

where $\widetilde{I}_n^1, \widetilde{I}_n^2, \widetilde{I}_n^3$ are defined by

$$\widetilde{I}_n^1 \equiv \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[I_{D_{m-1,n}^{\overline{\delta}_\varepsilon,3}} \left(F(\widetilde{\mu}_m^n, m, n) - F\left(E_Q[X_m|\mathcal{G}_{m-1}], m, n\right)\right)\right],$$

$$\widetilde{I}_n^2 \equiv \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[2\|\varphi\| I_{\{|X_m|>\frac{\sigma n}{\sigma+\sqrt{n}}\overline{\delta}_\varepsilon-\kappa\}} \right] \qquad (\text{ where } \|\varphi\| = \sup_{x\in\mathbb{R}} \varphi(x)),$$

$$\widetilde{I}_n^3 \equiv \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[\left| H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{\widetilde{Z}_m^n}{\sqrt{n}}\right) - F(\widetilde{\mu}_m^n, m, n)\right|\right]$$

$$+ \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[\left| H_{m,n}\left(T_{m-1,n} + \frac{X_m}{n} + \frac{Y_m^Q}{\sqrt{n}}\right) - F\left(E_Q[X_m|\mathcal{G}_{m-1}], m, n\right)\right|\right].$$

By the Lindeberg condition (2.1), $\lim_{n\to\infty} \widetilde{I}_n^2 = 0$, and by the remainder estimate in the proof of Lemma 6.4, $\lim_{n\to\infty} \widetilde{I}_n^3 = 0$. Finally, we prove that $\lim_{n\to\infty} \widetilde{I}_n^3 = 0$:

$$\begin{aligned}
\widetilde{I}_n^1 =& \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[I_{D_{m-1,n}^{\overline{\delta}_\varepsilon,3}} \left(F(\widetilde{\mu}_m^n, m, n) - F\left(E_Q[X_m|\mathcal{G}_{m-1}], m, n\right)\right)\right] \\
\le& \frac{1}{2n\sigma^2} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[H''_{m,n}(T_{m-1,n})(E_Q[X_m|\mathcal{G}_{m-1}] - \widetilde{\mu}_m^n)^2 I_{\{|T_{m-1,n}-c|\le\overline{\delta}_\varepsilon\}}\right] \\
\le& \frac{C\kappa}{n\sigma^2} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[|E_Q[X_m|\mathcal{G}_{m-1}] - \widetilde{\mu}_m^n|\, I_{\{|T_{m-1,n}-c|\le\overline{\delta}_\varepsilon\}}\right],
\end{aligned}$$

where C is the uniform bound given in Lemma 6.1 (2). Thus $\limsup_{n\to\infty} \widetilde{I}_n^1 \le C\kappa/\varepsilon$, where ε is arbitrary. This proves (6.23) and completes the proof of part (2). ■

The next lemma is used in extending the two theorems from the special case (6.1) to general $\underline{\mu}$ and $\overline{\mu}$.

Lemma 6.10. *For any $\kappa > 0$ and $c \in \mathbb{R}$,*

$$\mathbb{E}_{[-\kappa,\kappa]}\left[\varphi\left(c + B_1\right)\right] = \mathbb{E}_{[-\kappa+c,\,\kappa+c]}\left[\varphi\left(B_1\right)\right],$$

where $\mathbb{E}_{[-\kappa+c,\,\kappa+c]}\left[\varphi\left(B_1\right)\right]$ is defined in (1.4).

Proof: $\mathbb{E}_{[-\kappa+c,\kappa+c]}[\varphi(B_1)] = Y_0^c$, where (Y_t^c, Z_t^c) solves

$$Y_t^c = \varphi(B_1) + \int_t^1 \max_{-\kappa+c\le\mu\le\kappa+c} (\mu Z_s^c) ds - \int_t^1 Z_s^c dB_s$$

$$=\varphi(B_1)+\int_t^1\left(\max_{-\kappa\leq\mu\leq\kappa}(\mu Z_s^c)+cZ_s^c\right)ds-\int_t^1 Z_s^c dB_s,\ 0\leq t\leq 1.$$

where the last equality is due to

$$\max_{-\kappa+c\leq\mu\leq\kappa+c}(\mu z)=(\kappa+c)z^+-(-\kappa+c)z^-=\max_{-\kappa\leq\mu\leq\kappa}(\mu z)+cz.$$

Let Q^c be the probability measure satisfying

$$E_{P^*}\left[\frac{dQ^c}{dP^*}|\mathcal{F}_t\right]=\exp\left\{-\frac{c^2t}{2}+cB_t\right\},\quad t\geq 0.$$

Then $W_t=B_t-ct$ is a Brownian motion under Q^c and (Y_t^c,Z_t^c) solves

$$Y_t^c=\varphi(c+W_1)+\int_t^1\max_{-\kappa\leq\mu\leq\kappa}(\mu Z_s^c)ds-\int_t^1 Z_s^c dW_s,\ 0\leq t\leq 1.$$

Hence $Y_0^c=\mathbb{E}_{[-\kappa,\kappa]}\left[\varphi\left(c+B_1\right)\right]$. ■

Chen et al [3] derive closed-form solutions for a class of BSDEs by using the properties of BSDEs and related PDEs. The next lemma provides a simpler derivation for the special case where the terminal value of the BSDE is a suitably defined indicator function.

Lemma 6.11. *For any $a<b\in\mathbb{R}$ and $\kappa>0$,*

$$\mathbb{E}_{[-\kappa,\kappa]}[I_{[a,b]}(B_1)]=\begin{cases}\Phi_{-\kappa}\left(-a\right)-e^{-\kappa(b-a)}\Phi_{-\kappa}\left(-b\right) & \text{if } a+b\geq 0\\ \Phi_{-\kappa}\left(b\right)-e^{-\kappa(b-a)}\Phi_{-\kappa}\left(a\right) & \text{if } a+b<0.\end{cases}$$

and

$$\mathcal{E}_{[-\kappa,\kappa]}[I_{[a,b]}(B_1)]=\begin{cases}\Phi_{\kappa}\left(-a\right)-e^{\kappa(b-a)}\ \Phi_{\kappa}\left(-b\right) & \text{if } a+b\geq 0\\ \Phi_{\kappa}\left(b\right)-e^{\kappa(b-a)}\ \Phi_{\kappa}\left(a\right) & \text{if } a+b<0.\end{cases}$$

Proof: For $\kappa>0$, let

$$\mathcal{P}\equiv\left\{Q^v: E_{P^*}[\frac{dQ^v}{dP^*}|\mathcal{F}_1]=e^{-\frac{1}{2}\int_0^1 v_s^2ds+\int_0^1 v_s dB_s},(v_t)\text{ is }\mathcal{F}_t\text{-adapted and }\sup_{s\in[0,1]}|v_s|\leq\kappa\right\}.$$

Let $\varphi=I_{[a,b]}$, then by [6, Theorem 2.2] or [7, Lemma 3],

$$\mathbb{E}_{[-\kappa,\kappa]}[\varphi\left(B_1\right)]=\sup_{Q\in\mathcal{P}}E_Q\left[\varphi(B_1)\right]=\sup_{|v|\leq\kappa}E_{Q^v}\left[\varphi\left(B_1^v+\int_0^1 v_s ds\right)\right],$$

where $B_t^v\equiv B_t-\int_0^t v_s ds$ is the Brownian motion under Q^v.

Let (v_s) be any $\mathcal{F}_t$-adapted process valued in $[-\kappa, \kappa]$, and consider the following BSDEs:

$$\begin{aligned}
Y_t =&\varphi(B_1) + \int_t^1 \max_{-\kappa\le\mu\le\kappa}(\mu Z_s)ds - \int_t^1 Z_s dB_s \\
Y_t^v =&\varphi(\overline{B}_1^v) + \int_t^1 v_s Z_s^v ds - \int_t^1 Z_s^v d\overline{B}_s^v \\
=&\varphi(\overline{B}_1^v) - \int_t^1 Z_s^v dB_s \implies \\
Y_t' =&\varphi(\overline{B}_1^v) + \int_t^1 \max_{-\kappa\le\mu\le\kappa}(\mu Z_s')ds - \int_t^1 Z_s' d\overline{B}_s^v,
\end{aligned}$$

where $\overline{B}_t^v \equiv B_t + \int_0^t v_s ds$. Clearly, $Y_0 = Y_0' \geq Y_0^v$, and thus

$$\sup_{|v|\le\kappa} \mathbb{E}_{Q^v}[\varphi(B_1^v + \int_0^1 v_s ds)] = Y_0 \geq \sup_{|v|\le\kappa} Y_0^v = \sup_{|v|\le\kappa} E_{P^*}[\varphi(B_1 + \int_0^1 v_s ds)]$$

Let $(X_t^{v,x})$ and $(X_t^{*,x})$ be the solutions respectively of

$$\begin{aligned}
X_t^{v,x} &= x + \int_0^t v_s ds + B_t, \quad 0 \le t \le 1, \text{ and} \\
X_t^{*,x} &= x - \kappa \int_0^t sgn\,(X_s^{*,x})\, ds + B_t, \quad 0 \le t \le 1.
\end{aligned}$$

By the comparison theorem for stochastic differential equations [22, Thm. 2.1],

$$\begin{aligned}
&\sup_{|v|\le\kappa} P^* \left(a \le B_1 + \int_0^1 v_s ds \le b\right) \\
=&\sup_{|v|\le\kappa} P^* \left(|X_1^{v,c}| \le \frac{b-a}{2}\right) = P^* \left(|X_1^{*,c}| \le \frac{b-a}{2}\right),
\end{aligned}$$

where $c = -\frac{a+b}{2}$.

On the other hand, let (α_s) be any $\mathcal{F}_t$-adapted process valued in $[-\kappa, \kappa]$, and $(\overline{X}_t^{\alpha,x})$ be the solution of

$$\overline{X}_t^{\alpha,x} = x + \int_0^t \alpha_s ds + B_t^v, \quad 0 \le t \le 1$$

and let $(\overline{X}_t^{*,x})$ be the solution of

$$\overline{X}_t^{*,x} = x - \kappa \int_0^t sgn\left(\overline{X}_s^{*,x}\right) ds + B_t^v, \quad 0 \le t \le 1.$$

Then

$$
\begin{aligned}
&Q^v\left(a \le B_1^v + \int_0^1 \alpha_s ds \le b\right)\\
=&Q^v\left(|\overline{X}_1^{\alpha,c}| \le \frac{b-a}{2}\right) \le Q^v\left(|\overline{X}_1^{*,c}| \le \frac{b-a}{2}\right) = P^*\left(|X_1^{*,c}| \le \frac{b-a}{2}\right),
\end{aligned}
$$

and

$$
\begin{aligned}
&\sup_{|v|\le\kappa} Q^v\left(a \le B_1^v + \int_0^1 v_s ds \le b\right)\\
\le &\sup_{|\alpha|\le\kappa} Q^v\left(a \le B_1^v + \int_0^1 \alpha_s ds \le b\right) \le P^*\left(|X_1^{*,c}| \le \frac{b-a}{2}\right).
\end{aligned}
$$

That is,

$$
\sup_{|v|\le\kappa} Q^v\left(a \le B_1^v + \int_0^1 v_s ds \le b\right) = P^*\left(|X_1^{*,c}| \le \frac{b-a}{2}\right)
$$

By [23, Proposition 5.1], the transition probability density of $(X_t^{*,x})$ is given by (for all $t \in (0,1]$, $z \in \mathbb{R}$),

$$
q_x(t,z) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-z)^2+2\kappa t(|z|-|x|)+\kappa^2 t^2}{2t}} + \kappa e^{-2\kappa|z|}\int_{|x|+|z|-\kappa t}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{u^2}{2t}} du.
$$

Thus

$$
\begin{aligned}
P^*\left(|X_1^{*,c}| \le \frac{b-a}{2}\right) &= \int_{\frac{a-b}{2}}^{\frac{b-a}{2}} q_c(1,z)dz\\
&= \begin{cases} \Phi_{-\kappa}(-a) - e^{-\kappa(b-a)}\Phi_{-\kappa}(-b) & \text{if } a+b \ge 0\\ \Phi_{-\kappa}(b) - e^{-\kappa(b-a)}\Phi_{-\kappa}(a) & \text{if } a+b < 0. \end{cases}
\end{aligned}
$$

The rest can be proven in the same way. ■

6.2 Proof of Theorem 4.1

It is enough to prove (4.1). We prove it for $\varphi \in C_b^\infty(\mathbb{R})$. This suffices because any $\varphi \in C([-\infty,\infty])$ can be approximated uniformly by a sequence of functions in $C_b^\infty(\mathbb{R})$ (see Approximation Lemma in [17, Ch. VIII]).

Let

$$
S_0 \equiv 0,\ S_n \equiv \sum_{i=1}^n X_i,\ S_n^Q \equiv \sum_{i=1}^n Y_i^Q,\ Y_i^Q \equiv \frac{1}{\sigma}\left(X_i - E_Q[X_i|\mathcal{G}_{i-1}]\right).
$$

First we prove that

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\left|\sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{S_n}{n}+\frac{S_n^Q}{\sqrt{n}}\right)\right]-\mathbb{E}_{g_\epsilon}\left[\varphi\left(B_1\right)\right]\right|=0. \tag{6.24}$$

By the definition of $\{H_{m,n}\}_{m=0}^n$,

$$\begin{aligned}
&\sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{S_n}{n}+\frac{S_n^Q}{\sqrt{n}}\right)\right]-\mathbb{E}_{g_\epsilon}\left[\varphi\left(B_1\right)\right]\\
=&\sup_{Q\in\mathcal{P}} E_Q\left[H_{n,n}\left(\frac{S_n}{n}+\frac{S_n^Q}{\sqrt{n}}\right)\right]-H_{0,n}(0)\\
=&\sup_{Q\in\mathcal{P}} E_Q\left[H_{n,n}\left(\frac{S_n}{n}+\frac{S_n^Q}{\sqrt{n}}\right)\right]-\sup_{Q\in\mathcal{P}} E_Q\left[H_{n-1,n}\left(\frac{S_{n-1}}{n}+\frac{S_{n-1}^Q}{\sqrt{n}}\right)\right]\\
&+\sup_{Q\in\mathcal{P}} E_Q\left[H_{n-1,n}\left(\frac{S_{n-1}}{n}+\frac{S_{n-1}^Q}{\sqrt{n}}\right)\right]-\sup_{Q\in\mathcal{P}} E_Q\left[H_{n-2,n}\left(\frac{S_{n-2}}{n}+\frac{S_{n-2}^Q}{\sqrt{n}}\right)\right]\\
&+\dots\\
&+\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(\frac{S_m}{n}+\frac{S_m^Q}{\sqrt{n}}\right)\right]-\sup_{Q\in\mathcal{P}} E_Q\left[H_{m-1,n}\left(\frac{S_{m-1}}{n}+\frac{S_{m-1}^Q}{\sqrt{n}}\right)\right]\\
&+\dots+\sup_{Q\in\mathcal{P}} E_Q\left[H_{1,n}\left(\frac{S_1}{n}+\frac{S_1^Q}{\sqrt{n}}\right)\right]-H_{0,n}(0)\\
=&\sum_{m=1}^n\left\{\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(\frac{S_m}{n}+\frac{S_m^Q}{\sqrt{n}}\right)\right]-\sup_{Q\in\mathcal{P}} E_Q\left[H_{m-1,n}\left(\frac{S_{m-1}}{n}+\frac{S_{m-1}^Q}{\sqrt{n}}\right)\right]\right\}\\
=&\sum_{m=1}^n\left\{\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(\frac{S_m}{n}+\frac{S_m^Q}{\sqrt{n}}\right)\right]-\sup_{Q\in\mathcal{P}} E_Q\left[L_{m,n}\left(\frac{S_{m-1}}{n}+\frac{S_{m-1}^Q}{\sqrt{n}}\right)\right]\right\}\\
&+\sum_{m=1}^n\left\{\sup_{Q\in\mathcal{P}} E_Q\left[L_{m,n}\left(\frac{S_{m-1}}{n}+\frac{S_{m-1}^Q}{\sqrt{n}}\right)\right]\right.\\
&\qquad\left.-\sup_{Q\in\mathcal{P}} E_Q\left[H_{m-1,n}\left(\frac{S_{m-1}}{n}+\frac{S_{m-1}^Q}{\sqrt{n}}\right)\right]\right\}\\
\equiv& I_{1n}+I_{2n},
\end{aligned}$$

where $L_{m,n}(x)=H_{m,n}(x)+\frac{1}{n}g_0\left(H'_{m,n}(x)\right)+\frac{1}{2n}H''_{m,n}(x)$.

By Lemma 6.4, if $T_{m,n}=\frac{S_m}{n}+\frac{S_m^Q}{\sqrt{n}}$, then

$$|I_{1n}|\le\sum_{m=1}^n\left|\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(\frac{S_m}{n}+\frac{S_m^Q}{\sqrt{n}}\right)\right]-\sup_{Q\in\mathcal{P}} E_Q\left[L_{m,n}\left(\frac{S_{m-1}}{n}+\frac{S_{m-1}^Q}{\sqrt{n}}\right)\right]\right|$$

$$\to 0 \quad \text{as } n \to \infty.$$

Furthermore, by Lemmas 6.1 and 6.3, as $n \to \infty$,

$$
\begin{aligned}
|I_{2n}| &\le \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[\left| L_{m,n}\left(\frac{S_{m-1}}{n} + \frac{S^Q_{m-1}}{\sqrt{n}}\right) - H_{m-1,n}\left(\frac{S_{m-1}}{n} + \frac{S^Q_{m-1}}{\sqrt{n}}\right) \right| \right] \\
&\le \sum_{m=1}^{n} \sup_{x\in\mathbb{R}} |L_{m,n}(x) - H_{m-1,n}(x)| \\
&= \sum_{m=1}^{n} \sup_{x\in\mathbb{R}} \left| H_{m,n}(x) + \frac{1}{n} g_0\left(H'_{m,n}(x)\right) + \frac{1}{2n} H''_{m,n}(x) \right. \\
&\qquad\qquad \left. - \mathbb{E}_{g_\epsilon}\left[H_{m,n}\left(x + B_{\frac{m}{n}} - B_{\frac{m-1}{n}}\right)\right] \right| \\
&\le \sum_{m=1}^{n} \sup_{x\in\mathbb{R}} \left| H_{m,n}(x) + \frac{1}{n} g_\epsilon\left(H'_{m,n}(x)\right) + \frac{1}{2n} H''_{m,n}(x) - \mathbb{E}_{g_\epsilon}\left[H_{m,n}\left(x + B_{\frac{1}{n}}\right)\right] \right| \\
&\quad + \frac{1}{n} \sum_{m=1}^{n} \sup_{x\in\mathbb{R}} \left| g_\epsilon\left(H'_{m,n}(x)\right) - g_0\left(H'_{m,n}(x)\right) \right| \\
&\le \sum_{m=1}^{n} \sup_{x\in\mathbb{R}} \left| \mathbb{E}_{g_\epsilon}\left[H_{m,n}\left(x + B_{\frac{1}{n}}\right)\right] - H_{m,n}(x) - \frac{1}{n} g_\epsilon\left(H'_{m,n}(x)\right) - \frac{1}{2n} H''_{m,n}(x) \right| \\
&\quad + 2\kappa\epsilon,
\end{aligned}
$$

which sum converges to $2\kappa\epsilon$. This proves (6.24).

From the standard estimates for BSDEs [10, Proposition 2.1],

$$
\begin{aligned}
\left|\mathbb{E}_{g_\epsilon}\left[\varphi\left(B_1\right)\right] - \mathbb{E}_{g_0}\left[\varphi\left(B_1\right)\right]\right|^2 &\le \widehat{C} E_{P^*}\left[\left(\int_0^1 |g_\epsilon(Z^\epsilon_s) - g_0(Z^\epsilon_s)|\, ds \right)^2 \right] \\
&< \widehat{C} 4\kappa^2\epsilon^2,
\end{aligned}
$$

where $\widehat{C} > 0$ is a constant. Combine with (6.24) to obtain

$$
\begin{aligned}
&\lim_{n\to\infty} \left| \sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{S_n}{n} + \frac{S^Q_n}{n}\right)\right] - \mathbb{E}_{g_0}\left[\varphi\left(B_1\right)\right] \right| \\
\le &\lim_{\epsilon\to 0} \lim_{n\to\infty} \left| \sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{S_n}{n} + \frac{S^Q_n}{n}\right)\right] - \mathbb{E}_{g_\epsilon}\left[\varphi\left(B_1\right)\right] \right| \\
&+ \lim_{\epsilon\to 0} \left|\mathbb{E}_{g_\epsilon}\left[\varphi\left(B_1\right)\right] - \mathbb{E}_{g_0}\left[\varphi\left(B_1\right)\right]\right|.
\end{aligned}
$$

The latter sum equals 0, thus completing the proof under condition (6.1).

Finally, we describe the proof for general $\underline{\mu}$ and $\overline{\mu}$. Let $Y_i = X_i - \frac{\overline{\mu}+\underline{\mu}}{2}$ and $\kappa = \frac{\overline{\mu}-\underline{\mu}}{2}$. Then

$$\mathbb{E}[Y_i \mid \mathcal{G}_{i-1}] = \frac{\overline{\mu}-\underline{\mu}}{2} = \kappa, \quad \mathcal{E}[Y_i \mid \mathcal{G}_{i-1}] = -\frac{\overline{\mu}-\underline{\mu}}{2} = -\kappa.$$

Apply the above result to (Y_i) and $\widehat{\varphi}$, $\widehat{\varphi}(x) = \varphi\left(x + \frac{\overline{\mu}+\underline{\mu}}{2}\right)$, to obtain

$$\begin{aligned}
&\lim_{n\to\infty} \sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{S_n}{n} + \frac{S_n^Q}{n}\right)\right] \\
&= \lim_{n\to\infty} \sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{1}{n}\sum_{i=1}^n X_i + \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{1}{\sigma}\left(X_i - E_Q[X_i|\mathcal{G}_{i-1}]\right)\right)\right] \\
&= \lim_{n\to\infty} \sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{1}{n}\sum_{i=1}^n Y_i + \frac{\overline{\mu}+\underline{\mu}}{2} + \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{1}{\sigma}\left(Y_i - E_Q[Y_i|\mathcal{G}_{i-1}]\right)\right)\right] \\
&= \lim_{n\to\infty} \sup_{Q\in\mathcal{P}} E_Q\left[\widehat{\varphi}\left(\frac{1}{n}\sum_{i=1}^n Y_i + \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{1}{\sigma}\left(Y_i - E_Q[Y_i|\mathcal{G}_{i-1}]\right)\right)\right] \\
&= \mathbb{E}_{\left[\frac{\underline{\mu}-\overline{\mu}}{2}, \frac{\overline{\mu}-\underline{\mu}}{2}\right]}\left[\widehat{\varphi}\left(B_1\right)\right] \\
&= \mathbb{E}_{\left[\frac{\underline{\mu}-\overline{\mu}}{2}, \frac{\overline{\mu}-\underline{\mu}}{2}\right]}\left[\varphi\left(\frac{\overline{\mu}+\underline{\mu}}{2} + B_1\right)\right] = \mathbb{E}_{[\underline{\mu},\overline{\mu}]}\left[\varphi\left(B_1\right)\right],
\end{aligned}$$

where the last equality is due to Lemma 6.10. This completes the proof. ■

Remark 6.12. *Straightforward modifications of the preceding arguments deliver a proof of Theorem 5.1. The key is modification of Lemma 6.4 (see Remark 6.5 and Appendix A.5). The remaining arguments are similar to those given above and are omitted.*

6.3 Proof of Theorem 4.3

Proof of (1): Let $\varphi \in C([-\infty,\infty])$ be symmetric with center $c \in \mathbb{R}$ and decreasing on (c,∞). The result is clear if φ is globally constant. Thus we assume that φ is not a globally constant function. Then φ can be approximated uniformly by φ_h defined by

$$\varphi_h(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \varphi(x+hy) e^{-\frac{y^2}{2}} dy, \tag{6.25}$$

and (see Appendix A.3), φ_h is symmetric with center c, and satisfies

$$sgn(\varphi_h'(x)) = -sgn(x-c), \;\; \forall h > 0.$$

Consider the special case (6.1). Let $\{H_{m,n}\}_{m=0}^{n}$ be defined via (6.5) using φ, where, without loss of generality we assume $\varphi \in C_b^3(\mathbb{R})$; otherwise, we can use φ_h defined in (6.25).

Let

$$S_0 \equiv 0,\ S_n \equiv \sum_{i=1}^{n} X_i,\ \overline{S}_n \equiv \sum_{i=1}^{n} Z_i^n,\ \ Z_i^n \equiv \frac{1}{\sigma}\left(X_i - \mu_i^n\right).$$

First, prove that

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\left|\sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{S_n}{n}+\frac{\overline{S}_n}{\sqrt{n}}\right)\right]-\mathbb{E}_{g_\epsilon}\left[\varphi\left(B_1\right)\right]\right| = 0. \tag{6.26}$$

Argue as follows:

$$\begin{aligned}
&\sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{S_n}{n}+\frac{\overline{S}_n}{\sqrt{n}}\right)\right]-\mathbb{E}_{g_\epsilon}\left[\varphi\left(B_1\right)\right]\\
=&\sup_{Q\in\mathcal{P}} E_Q\left[H_{n,n}\left(\frac{S_n}{n}+\frac{\overline{S}_n}{\sqrt{n}}\right)\right]-H_{0,n}(0)\\
=&\sup_{Q\in\mathcal{P}} E_Q\left[H_{n,n}\left(\frac{S_n}{n}+\frac{\overline{S}_n}{\sqrt{n}}\right)\right]-\sup_{Q\in\mathcal{P}} E_Q\left[H_{n-1,n}\left(\frac{S_{n-1}}{n}+\frac{\overline{S}_{n-1}}{\sqrt{n}}\right)\right]\\
&+\sup_{Q\in\mathcal{P}} E_Q\left[H_{n-1,n}\left(\frac{S_{n-1}}{n}+\frac{\overline{S}_{n-1}}{\sqrt{n}}\right)\right]-\sup_{Q\in\mathcal{P}} E_Q\left[H_{n-2,n}\left(\frac{S_{n-2}}{n}+\frac{\overline{S}_{n-2}}{\sqrt{n}}\right)\right]\\
&+\dots\\
&+\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(\frac{S_m}{n}+\frac{\overline{S}_m}{\sqrt{n}}\right)\right]-\sup_{Q\in\mathcal{P}} E_Q\left[H_{m-1,n}\left(\frac{S_{m-1}}{n}+\frac{\overline{S}_{m-1}}{\sqrt{n}}\right)\right]\\
&+\dots\\
&+\sup_{Q\in\mathcal{P}} E_Q\left[H_{1,n}\left(\frac{S_1}{n}+\frac{\overline{S}_1}{\sqrt{n}}\right)\right]-H_{0,n}(0)\\
=&\sum_{m=1}^{n}\left\{\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(\frac{S_m}{n}+\frac{\overline{S}_m}{\sqrt{n}}\right)\right]-\sup_{Q\in\mathcal{P}} E_Q\left[L_{m,n}\left(\frac{S_{m-1}}{n}+\frac{\overline{S}_{m-1}}{\sqrt{n}}\right)\right]\right\}\\
&+\sum_{m=1}^{n}\left\{\sup_{Q\in\mathcal{P}} E_Q\left[L_{m,n}\left(\frac{S_{m-1}}{n}+\frac{\overline{S}_{m-1}}{\sqrt{n}}\right)\right]\right.\\
&\qquad\left.-\sup_{Q\in\mathcal{P}} E_Q\left[H_{m-1,n}\left(\frac{S_{m-1}}{n}+\frac{\overline{S}_{m-1}}{\sqrt{n}}\right)\right]\right\}\\
\equiv&J_{1n}+J_{2n},
\end{aligned}$$

where $L_{m,n}(x) = H_{m,n}(x) + \frac{1}{n} g_0\left(H'_{m,n}(x)\right) + \frac{1}{2n} H''_{m,n}(x)$.

By Lemma 6.8(1), with $T_{m,n} = \frac{S_m}{n} + \frac{\overline{S}_m}{\sqrt{n}}$, we have

$$
\begin{aligned}
|J_{1n}| \leq & \sum_{m=1}^{n} \left| \sup_{Q \in \mathcal{P}} E_Q \left[H_{m,n}\left(\frac{S_m}{n} + \frac{\overline{S}_m}{\sqrt{n}} \right) \right] - \sup_{Q \in \mathcal{P}} E_Q \left[L_{m,n}\left(\frac{S_{m-1}}{n} + \frac{\overline{S}_{m-1}}{\sqrt{n}} \right) \right] \right| \\
& \to 0 \quad \text{as } n \to \infty.
\end{aligned}
$$

As in the proof of Theorem 4.1, we have $|J_{2n}| \to 0,$ as $n \to \infty$. Hence, (6.26) holds. Combine it with standard estimate for BSDEs to complete the proof under condition (6.1).

For the case of general $\underline{\mu}$ and $\overline{\mu}$, let $Y_i = X_i - \frac{\overline{\mu}+\underline{\mu}}{2}$. Then

$$
\mathbb{E}[Y_i \mid \mathcal{G}_{i-1}] = \frac{\overline{\mu} - \underline{\mu}}{2}, \quad \mathcal{E}[Y_i \mid \mathcal{G}_{i-1}] = -\frac{\overline{\mu} - \underline{\mu}}{2}.
$$

Apply the above result for (Y_i) to $\widehat{\varphi}$, $\widehat{\varphi}(x) = \varphi\left(x + \frac{\overline{\mu}+\underline{\mu}}{2}\right)$, to obtain

$$
\begin{aligned}
& \lim_{n \to \infty} \sup_{Q \in \mathcal{P}} E_Q \left[\varphi \left(\frac{1}{n} \sum_{i=1}^{n} X_i + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sigma} (X_i - \mu_i^n) \right) \right] \\
= & \lim_{n \to \infty} \sup_{Q \in \mathcal{P}} E_Q \left[\varphi \left(\frac{1}{n} \sum_{i=1}^{n} Y_i + \frac{\overline{\mu}+\underline{\mu}}{2} + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sigma} \left(Y_i - \left(\mu_i^n - \frac{\overline{\mu}+\underline{\mu}}{2} \right) \right) \right) \right] \\
= & \lim_{n \to \infty} \sup_{Q \in \mathcal{P}} E_Q \left[\widehat{\varphi} \left(\frac{1}{n} \sum_{i=1}^{n} Y_i + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sigma} (Y_i - \gamma_i^n) \right) \right] \\
= & \mathbb{E}_{\left[\frac{\underline{\mu}-\overline{\mu}}{2}, \frac{\overline{\mu}-\underline{\mu}}{2}\right]} \left[\widehat{\varphi}(B_1) \right] \\
= & \mathbb{E}_{\left[\frac{\underline{\mu}-\overline{\mu}}{2}, \frac{\overline{\mu}-\underline{\mu}}{2}\right]} \left[\varphi \left(\frac{\overline{\mu}+\underline{\mu}}{2} + B_1 \right) \right] \\
= & \mathbb{E}_{[\underline{\mu}, \overline{\mu}]} \left[\varphi(B_1) \right],
\end{aligned}
$$

where the last equality is due to Lemma 6.10. Also,

$$
\gamma_m^n = \tfrac{\overline{\mu}-\underline{\mu}}{2} I_{\widehat{A}_{m-1,n}} + \tfrac{\underline{\mu}-\overline{\mu}}{2} I_{\widehat{A}^c_{m-1,n}}, \quad \text{and}
$$

$$
\begin{aligned}
\widehat{A}_{m-1,n} = & \left\{ \tfrac{1}{n} \sum_{i=1}^{m-1} Y_i + \tfrac{1}{\sqrt{n}} \sum_{i=1}^{m-1} \tfrac{1}{\sigma} (Y_i - \gamma_i^n) \leq -\tfrac{\overline{\mu}+\underline{\mu}}{2} + c \right\} \\
= & \left\{ \tfrac{1}{n} \sum_{i=1}^{m-1} X_i + \tfrac{1}{\sqrt{n}} \sum_{i=1}^{m-1} \tfrac{1}{\sigma} \left(X_i - \gamma_i^n - \tfrac{\overline{\mu}+\underline{\mu}}{2} \right) \leq -\tfrac{\overline{\mu}+\underline{\mu}}{2} \left(1 - \tfrac{m-1}{n}\right) + c \right\}.
\end{aligned}
$$

Thus $\widehat{A}_{0,n} = A_{0,n}$, and

$$\gamma_1^n + \tfrac{\overline{\mu}+\underline{\mu}}{2} = \overline{\mu} I_{\widehat{A}_{0,n}} + \underline{\mu} I_{\widehat{A}_{0,n}^c} = \overline{\mu} I_{A_{0,n}} + \underline{\mu} I_{A_{0,n}^c} = \mu_1^n.$$

By induction, $A_{m-1,n} = \widehat{A}_{m-1,n}$, for $m \geq 1$, and

$$\gamma_m^n + \tfrac{\overline{\mu}+\underline{\mu}}{2} = \overline{\mu} I_{\widehat{A}_{m-1,n}} + \underline{\mu} I_{\widehat{A}_{m-1,n}^c} = \overline{\mu} I_{A_{m-1,n}} + \underline{\mu} I_{A_{m-1,n}^c} = \mu_m^n.$$

This completes the proof of (4.8).

By standard limiting arguments, (4.8) can be extended to indicator functions for intervals. Then (4.14) follows from Lemmas 6.10 and 6.11.

Proof of (2): In light of Lemma 6.8(2), the proof of part (2) is similar to that of (1) and is omitted. ∎

References

[1] Acciaio, B. and Penner, I. (2011). Dynamic convex risk measures. In G. Di Nunno and B. Øksendal eds. *Advanced Mathematical Methods for Finance.* Springer, Berlin Heidelberg, 1-34.

[2] Briand, P., Coquet, F., Hu, Y., Memin, J. and Peng, S. (2000). A converse comparison theorem for BSDEs and related properties of g-expectation. *Electron. Comm. Probab.* **5** 101-117.

[3] Chen, Z. , Liu, S., Qian, Z. and Xu, X. (2020). Explicit solutions for a class of nonlinear BSDEs and their nodal sets. *arXiv:2006.00222 [math.PR].*

[4] Chen, Z. (2016). Strong laws of large numbers for sub-linear expectations. *Science China Math.* **59** 945-954.

[5] Chen, Z., Chen, T. and Davison, M. (2005). Choquet expectation and Peng's g-expectation. *Ann. Probab.* **33** 1179-1199.

[6] Chen, Z. and Epstein, L.G. (2002). Ambiguity, risk and asset returns in continuous time. *Econometrica* **70** 1403-1443.

[7] Chen, Z. and Peng, S. (2000). A general downcrossing inequality for g-martingales. *Statist. Probab. Letters.* **46** 169-175.

[8] Cheridito, P., Delbaen, F. and Kupper, M. (2006). Dynamic monetary risk measures for bounded discrete-time processes. *Electron. J. Probab.* **11** 57-106.

[9] Dasgupta, A. (2008). *Asymptotic Theory of Statistics and Probability.* Springer.

[10] El Karoui, N., Peng, S. and Quenez, M. (1997). Backward stochastic differential equations in finance. *Math. Finan.* **7** 1-71.

[11] Epstein, L.G., Kaido, H. and Seo, K. (2016). Robust confidence regions for incomplete models. *Econometrica* **84** 1799-1838.

[12] Epstein, L.G. and Schneider, M. (2003). Recursive multiple-priors. *J. Econom. Theory* **113** 1-31.

[13] Epstein, L.G. and Schneider, M. (2003). IID: independently and indistinguishably distributed. *J. Econom. Theory* **113** 32-50.

[14] Epstein, L.G. and Schneider, M. (2008). Ambiguity, information quality and asset pricing. *J. Finan.* **63** 197-228.

[15] Epstein, L.G. and Schneider, M. (2010). Ambiguity and asset markets. *Annual Rev. Finan. Econom.* **2** 315-34.

[16] Fang, X., Peng, S., Shao, Q., and Song, Y. (2019). Limit theorems with rate of convergence under sublinear expectations. *Bernoulli* **25** 2564-2596.

[17] Feller, W. (1958). *An Introduction to Probability Theory and its Applications-Vol II.* Cambridge U. Press.

[18] Follmer, H., and Schied, A. (2002). *Stochastic Finance: An Introduction in Discrete Time.* De Gruyter Studies in Mathematics 27.

[19] Gilboa, I. (2009). *Theory of Decision under Uncertainty.* Cambridge U. Press.

[20] Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique priors. *J. Math. Econom.* **18** 141-153.

[21] Huber, P. J. (1981). *Robust Statistics.* John Wiley & Sons.

[22] Ikeda, N., and Watanabe, S. (1977). A comparison theorem for solutions of stochastic differential equations and its applications. *Osaka J. Math.* **14** 619-633.

[23] Karatzas, I., and Shreve, S. E. (1984). Trivariate density of Brownian motion, its local and occupation times, with application to stochastic control. *Ann. Probab.* **12**, 819-828.

[24] Kolmogorov, A.N. and Gnedenko, B.V. (1954). *Limit Distributions for Sums of Independent Random Variables.* Translated by K.L. Chung. Addison-Wesley.

[25] Ma, J.,and Zhang, J. (2002). Representation theorems for backward stochastic differential equations. *Ann. Appl. Probab.* **12** 1390-1418.

[26] Maccheroni, M. and Marinacci, M. (2005). A strong law of large numbers for capacities. *Ann. Probab.* **33** 1171-1178.

[27] Pardoux, E. and Peng, S. (1990). Adapted solution of a backward stochastic differential equation. *Systems Control Letters.* **14** 55-61.

[28] Parthasarathy, K.R. (1967). *Probability Measures on Metric Spaces.* AMS Chelsea Publishing.

[29] Peng, S. (1992). A generalized dynamic programming principle and Hamilton-Jacobi-Bellman equation. *Stochastics.* **38** 119-134.

[30] Peng, S. *(1997).* Backward SDE and related g-expectation. In N. El Karoui and L. Mazliak eds. *Backward Stochastic Differential Equations.* Pitman Research Notes in Math. Series ***364*** 141-159.

[31] Peng, S. (2019). Nonlinear expectations and stochastic calculus under uncertainty. *Springer*, 23–44.

[32] Peng, S. (2009). Survey on normal distributions, central limit theorem, Brownian motion and the related stochastic calculus under sublinear expectations. *Sci. China Ser. A* **52** 1391–1411.

[33] Riedel, F. (2004). Dynamic coherent risk measures. *Stoch. Proc. and Appl.* **112** 185-200.

[34] Shapiro, A. (2016). Rectangular sets of probability measures. *Oper. Res.* **64** 528-541.

[35] Shi, X. (2018). Central limit theorems for bounded random variables under belief measures. *J. Math. Anal. Appl.* **460** 546-560.

[36] Tamer, E. (2010). Partial identification in econometrics. *Annual Rev. Econom.* **2** 167-195.

[37] Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities.* Chapman and Hall.

A Supplementary Appendix

A.1 Rectangularity

Let $\mathcal{P} \subset \Delta(\Omega, \mathcal{G})$ be rectangular. All measures in $\mathcal{P}$ are equivalent on each $\mathcal{G}_n$ and relations between $\mathcal{G}_n$-measurable r.v.s should be understood to hold P_0-a.s. for some fixed measure P_0 in $\mathcal{P}$. $\mathcal{H}$ denotes the set of r.v.s X on $(\Omega, \mathcal{G})$ satisfying $\sup_{Q\in\mathcal{P}} E_Q[|X|] < \infty$.

Lemma A.1. *For any $X \in \mathcal{H}$ and any n, there is a sequence $E_{P_i}[X|\mathcal{G}_n]$ in $\{E_P[X|\mathcal{G}_n] : P \in \mathcal{P}\}$ such that $ess\sup_{P\in\mathcal{P}} E_P[X|\mathcal{G}_n]$ is the increasing limit of $E_{P_i}[X|\mathcal{G}_n]$.*

Proof: We prove that $\{E_P[X|\mathcal{G}_n] : P \in \mathcal{P}\}$ is an upward-directed set. Then the result follows from [18, Theorem A.32].
Let $Q_1, Q_2 \in \mathcal{P}$, and $\forall B \in \mathcal{G}_n$, $\forall \omega = (\omega^{(n)}, \omega_{(n+1)}) \in \Omega$, $\forall D \in \mathcal{G}_{(n+1)}$, define

$$\lambda(\omega^{(n)}, D) = \begin{cases} Q_1(\prod_1^n \Omega_i \times D|\mathcal{G}_n)(\omega^{(n)}) & \text{if } \omega \in B, \\ Q_2(\prod_1^n \Omega_i \times D|\mathcal{G}_n)(\omega^{(n)}) & \text{if } \omega \in B^c, \end{cases} \qquad \text{and } p_n = Q_1|_{\mathcal{G}_n}.$$

Then, $p_n \in \mathcal{P}_{0,n}$ and λ is a $\mathcal{P}$-kernel. By rectangularity, $P \in \mathcal{P}$, where, for $A \in \mathcal{G}$,

$$\begin{aligned} P(A) &= \int_{\prod_1^n \Omega_i} \int_{\prod_{n+1}^\infty \Omega_i} I_A(\omega^{(n)}, \omega_{(n+1)}) \lambda(\omega^{(n)}, d\omega_{(n+1)}) p_n(d\omega^{(n)}) \\ &= \int_{\prod_1^n \Omega_i} \int_{\prod_{n+1}^\infty \Omega_i} I_{A\cap B}(\omega^{(n)}, \omega_{(n+1)}) Q_1(\prod_{i=1}^n \Omega_i \times d\omega_{(n+1)}|\mathcal{G}_n)(\omega^{(n)}) p_n(d\omega^{(n)}) \\ &+ \int_{\prod_1^n \Omega_i} \int_{\prod_{n+1}^\infty \Omega_i} I_{A\cap B^c}(\omega^{(n)}, \omega_{(n+1)}) Q_2(\prod_{i=1}^n \Omega_i \times d\omega_{(n+1)}|\mathcal{G}_n)(\omega^{(n)}) p_n(d\omega^{(n)}). \end{aligned}$$

Consider the probability measure $\widetilde{P}$ with Radon-Nikodym derivative

$$\frac{d\widetilde{P}}{dP_0} = \frac{dQ_1}{dP_0} I_B + \frac{(\frac{dQ_1}{dP_0})_n}{(\frac{dQ_2}{dP_0})_n} \frac{dQ_2}{dP_0} I_{B^c},$$

where $(\frac{dQ_1}{dP_0})_n$ means $E_{P_0}[\frac{dQ_1}{dP_0}|\mathcal{G}_n]$. We claim that $P(A) = \widetilde{P}(A)$, $\forall A \in \mathcal{G}$. Indeed, by the definitions, for all n, $P(A) = \widetilde{P}(A)$, $\forall A \in \mathcal{C}_n$, where

$$\mathcal{C}_n = \{A^{(n)} \times A_{(n+1)} : A^{(n)} \in \mathcal{G}_n,\ A_{(n+1)} \in \mathcal{G}_{(n+1)}\}.$$

Since $\mathcal{C}_n$ is a π class, and satisfies $\sigma(\mathcal{C}_n) = \mathcal{G}$, P and $\widetilde{P}$ are identical on $\mathcal{G}$.

Note that $(\frac{d\widetilde{P}}{dP_0})_n = (\frac{dQ_1}{dP_0})_n$, and, by Bayes rule,

$$\begin{aligned} E_P[X|\mathcal{G}_n] &= E_{\widetilde{P}}[X|\mathcal{G}_n] \\ &= E_{P_0}\left[X\frac{d\widetilde{P}}{dP_0}|\mathcal{G}_n\right]\left((\tfrac{dQ_1}{dP_0})_n\right)^{-1} \\ &= E_{P_0}\left[X\left(\frac{dQ_1}{dP_0}I_B + \frac{(\frac{dQ_1}{dP_0})_n}{(\frac{dQ_2}{dP_0})_n}\frac{dQ_2}{dP_0}I_{B^c}\right)|\mathcal{G}_n\right][(\tfrac{dQ_1}{dP_0})_n]^{-1} \\ &= I_B E_{Q_1}\left[X|\mathcal{G}_n\right] + I_{B^c}E_{Q_2}\left[X|\mathcal{G}_n\right]. \end{aligned} \tag{A.1}$$

If $B = \{\omega \in \Omega : E_{Q_1}[X|\mathcal{G}_n](\omega) > E_{Q_2}[X|\mathcal{G}_n](\omega)\}$, then

$$E_P[X|\mathcal{G}_n] = ess\sup\{E_{Q_1}\left[X|\mathcal{G}_n\right], E_{Q_2}\left[X|\mathcal{G}_n\right]\}. \quad \blacksquare$$

Proof of Lemma 2.2: (i) For any $\omega^{(n)} \in \prod_1^n \Omega_i$ and $B \in \mathcal{G}_{(n+1)}$, define

$$\lambda(\omega^{(n)}, B) = Q(\prod_{i=1}^n \Omega_i \times B|\mathcal{G}_n)(\omega^{(n)}) \text{ and } p_n = R|_{\mathcal{G}_n}.$$

Then $p_n \in \mathcal{P}_{0,n}$ and λ is a $\mathcal{P}$-kernel. By rectangularity, $P \in \mathcal{P}$, where, for $A \in \mathcal{G}$,

$$P(A) = \int_{\prod_1^n \Omega_i}\int_{\prod_{n+1}^\infty \Omega_i} I_A(\omega^{(n)}, \omega_{(n+1)})Q(\prod_{i=1}^n \Omega_i \times d\omega_{(n+1)}|\mathcal{G}_n)(\omega^{(n)})p_n(d\omega^{(n)}).$$

Consider the probability measure $\widetilde{P}$ with Radon-Nikodym derivative

$$\frac{d\widetilde{P}}{dP_0} = \frac{(\frac{dR}{dP_0})_n}{(\frac{dQ}{dP_0})_n}\frac{dQ}{dP_0},$$

where $(\frac{dR}{dP_0})_n$ means $E_{P_0}[\frac{dR}{dP_0}|_{\mathcal{G}_n}]$. Argue as in the proof of Lemma A.1, to show that P and $\widetilde{P}$ are identical on $\mathcal{G}$. Note that $(\frac{d\widetilde{P}}{dP_0})_n = (\frac{dR}{dP_0})_n$, and, by Bayes rule, for any $m < n$ and $X \in \mathcal{H}$,

$$\begin{aligned} E_P[X|\mathcal{G}_m] = E_{\widetilde{P}}[X|\mathcal{G}_m] &= E_{\widetilde{P}}[E_{\widetilde{P}}[X|\mathcal{G}_n]|\mathcal{G}_m] \\ &= E_{\widetilde{P}}\left[E_{P_0}\left[X\frac{d\widetilde{P}}{dP_0}|\mathcal{G}_n\right]\left((\tfrac{dR}{dP_0})_n\right)^{-1}|\mathcal{G}_m\right] \end{aligned}$$

$$
\begin{aligned}
&= E_{\widetilde{P}}\left[E_{P_0}\left[X\frac{(\frac{dR}{dP_0})_n}{(\frac{dQ}{dP_0})_n}\frac{dQ}{dP_0}|\mathcal{G}_n\right]\left((\tfrac{dR}{dP_0})_n\right)^{-1}|\mathcal{G}_m\right] \\
&= E_{\widetilde{P}}\left[E_Q\left[X|\mathcal{G}_n\right]|\mathcal{G}_m\right] = E_R\left[E_Q\left[X|\mathcal{G}_n\right]|\mathcal{G}_m\right].
\end{aligned}
$$

(ii) can be proven using (A.1).
(iii) By Lemma A.1, there exist increasing sequences $\{E_{Q_i}[\phi(X)|\mathcal{G}_n]\}$ and $\{E_{P_j}[\mathbb{E}[X|\mathcal{G}_n]|\mathcal{G}_m]|\}$, with $Q_i, P_j \in \mathcal{P}$ for all i and j, such that

$$
\begin{aligned}
\mathbb{E}[X|\mathcal{G}_n] &= \lim_{i\to\infty} E_{Q_i}[X|\mathcal{G}_n], \text{ and} \\
\mathbb{E}[\mathbb{E}[X|\mathcal{G}_n]|\mathcal{G}_m] &= \lim_{j\to\infty} E_{P_j}[\mathbb{E}[X|\mathcal{G}_n]|\mathcal{G}_m].
\end{aligned}
$$

By the monotone convergence theorem and (i),

$$
\begin{aligned}
\mathbb{E}[\mathbb{E}[X|\mathcal{G}_n]|\mathcal{G}_m] &= \lim_{j\to\infty}\lim_{i\to\infty} E_{P_j}[E_{Q_i}[X|\mathcal{G}_n]|\mathcal{G}_m] \\
&\leq ess\sup_{P\in\mathcal{P}} E_P[X|\mathcal{G}_m] \\
&= \mathbb{E}[X|\mathcal{G}_m].
\end{aligned}
$$

For the reverse inequality, we have

$$
\begin{aligned}
\mathbb{E}[X|\mathcal{G}_m] &= ess\sup_{P\in\mathcal{P}} E_P[E_P[X|\mathcal{G}_n]|\mathcal{G}_m] \\
&\leq ess\sup_{P\in\mathcal{P}} E_P[ess\sup_{Q\in\mathcal{P}} E_Q[X|\mathcal{G}_n]|\mathcal{G}_m] \\
&= \mathbb{E}[\mathbb{E}[X|\mathcal{G}_n]|\mathcal{G}_m].
\end{aligned}
$$

(iv) can be proven using (iii). ■

A.2 IID model: Lemma 3.1

Part (i) was proven in the text. (iii) follows from (i) and Lemma 2.2.
For (iv), use (i) and (iii) to argue that, for example,

$$
\begin{aligned}
&\sup_{Q\in\mathcal{P}^{IID}} E_Q\left[(X_n - E_Q[X_n|\mathcal{G}_{n-1}])^2|\mathcal{G}_{n-1}\right] \\
&= \sup_{Q\in\mathcal{P}^{IID}}\left\{E_Q\left[X_n^2|\mathcal{G}_{n-1}\right] - (E_Q[X_n|\mathcal{G}_{n-1}])^2\right\} \\
&= \sup_{q\in\mathcal{L}}\left\{E_q[X_n^2] - (E_q[X_n])^2\right\} = \sup_{q\in\mathcal{L}} E_q\left[(\overline{X} - E_q\overline{X})^2\right].
\end{aligned}
$$

The equivalence on each $\mathcal{G}_n$ stated in (ii) is proven by induction. Let $P, Q \in \mathcal{P}^{IID}$. Equivalence on $\mathcal{G}_1$ is due to the equivalence of measures in $\mathcal{L}$. Suppose P and Q are equivalent on $\mathcal{G}_{n-1}$, and prove equivalence on $\mathcal{G}_n$.
Let $A \in \mathcal{G}_n$, $P(A) = 0$. Then

$$E_P\left[E_P\left[I_A|\mathcal{G}_{n-1}\right]\right] = 0 \Leftrightarrow \ P\left(\{E_P[I_A|\mathcal{G}_{n-1}] > 0\}\right) = 0.$$

By the equivalence of measures in $\mathcal{L}$, $\forall \omega^{(n-1)} \in \prod_1^{n-1} \Omega_i$,

$$\left\{\omega^{(n-1)} : \ E_P[I_A|\mathcal{G}_{n-1}](\omega^{(n-1)}) > 0\right\} = \left\{\omega^{(n-1)} : \ E_Q[I_A|\mathcal{G}_{n-1}](\omega^{(n-1)}) > 0\right\}.$$

Given also equivalence of P and Q on $\mathcal{G}_{n-1}$, conclude that

$$\begin{aligned} Q\left(\{E_Q[I_A|\mathcal{G}_{n-1}] > 0\}\right) &= P\left(\{E_Q[I_A|\mathcal{G}_{n-1}] > 0\}\right) \\ &= P\left(\{E_P[I_A|\mathcal{G}_{n-1}] > 0\}\right) = 0, \end{aligned}$$

and hence $Q(A) = 0$. ■

A.3 Some details for proof of Theorem 4.3

Let $\varphi \in C([-\infty, \infty])$ be symmetric with center $c \in \mathbb{R}$ and decreasing on (c, ∞). Define φ_h, for $h > 0$, by (6.25). Here we prove that:
(i) φ_h is symmetric with center c; and (ii) $sgn(\varphi_h'(x)) = -sgn(x - c)$.
Proof: (i) By the definition of φ_h,

$$\begin{aligned} \varphi_h(x + c) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \varphi(x + c + hy) e^{-\frac{y^2}{2}} dy \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \varphi(-x + c - hy) e^{-\frac{y^2}{2}} dy \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \varphi(-x + c + hy) e^{-\frac{y^2}{2}} dy \\ &= \varphi_h(-x + c) \end{aligned}$$

(ii) Compute that

$$\varphi_h'(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi} h^3} \varphi(x + y) y e^{-\frac{y^2}{2h^2}} dy.$$

Since φ_h is symmetric with c, we have for any $x > c$,

$$\varphi_h'(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi} h^3} \varphi(x + y) y e^{-\frac{y^2}{2h^2}} dy$$

$$
\begin{aligned}
&= \int_0^{\infty} \frac{1}{\sqrt{2\pi h^3}} \varphi(c+y+x-c) y e^{-\frac{y^2}{2h^2}} dy \\
&\quad + \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi h^3}} \varphi(c+y+x-c) y e^{-\frac{y^2}{2h^2}} dy \\
&= \int_0^{\infty} \frac{1}{\sqrt{2\pi h^3}} \varphi(c+y+x-c) y e^{-\frac{y^2}{2h^2}} dy \\
&\quad - \int_0^{\infty} \frac{1}{\sqrt{2\pi h^3}} \varphi(c+y+c-x) y e^{-\frac{y^2}{2h^2}} dy \\
&= \int_0^{\infty} \frac{1}{\sqrt{2\pi h^3}} \left(\varphi(c+y+x-c) - \varphi(c+y+c-x)\right) y e^{-\frac{y^2}{2h^2}} dy \\
&< 0
\end{aligned}
$$

Thus $sgn(\varphi_h'(x)) = -sgn(x-c)$. ■

A.4 Some details for hypothesis testing

Both (X_i) and (Y_i) described in section 4.2 conform to the IID model, with the common variance σ^2 and mean intervals $\left[\underline{\mu}, \overline{\mu}\right]$ and $\left[\underline{\mu}-\theta, \overline{\mu}-\theta\right]$ respectively. (The text considers the special case $\left[\underline{\mu}-\theta, \overline{\mu}-\theta\right] = [-\kappa, \kappa]$.) Let μ_m^n be defined by the form of (4.6) appropriate for (X_i) and denote by γ_m^n the corresponding variables appropriate for (Y_i). Here we prove (4.17), for which it suffices to show that

$$
\mu_m^n = \theta + \gamma_m^n. \tag{A.2}
$$

By Theorem 4.3(1), if φ is decreasing on $(c, +\infty)$, then

$$
\lim_{n\to\infty} \sup_{Q\in\mathcal{P}} E_Q \left[\varphi\left(\frac{1}{n}\sum_{i=1}^{n} X_i + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sigma}\left(X_i - \mu_i^n\right)\right)\right] = \mathbb{E}_{\left[\underline{\mu}, \overline{\mu}\right]}[\varphi(B_1)],
$$

where, by (4.6), $\mu_m^n = \overline{\mu} I_{A_{m-1,n}} + \underline{\mu} I_{A^c_{m-1,n}}$ and

$$
A_{m-1,n} = \left\{ \frac{1}{n}\sum_{i=1}^{m-1} X_i + \frac{1}{\sqrt{n}}\sum_{i=1}^{m-1}\frac{1}{\sigma}\left(X_i - \mu_i^n\right) \le -\frac{\overline{\mu}+\underline{\mu}}{2}\left(1 - \frac{m-1}{n}\right) + c \right\}.
$$

Let $\phi(x) = \varphi(x+\theta)$. Then ϕ is symmetric with center $\widehat{c} = c - \theta$. Theorem 4.3(1) applied to (Y_i) yields

$$
\lim_{n\to\infty} \sup_{Q\in\mathcal{P}} E_Q \left[\phi\left(\frac{1}{n}\sum_{i=1}^{n} Y_i + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sigma}\left(Y_i - \gamma_i^n\right)\right)\right] = \mathbb{E}_{\left[\underline{\mu}-\theta, \overline{\mu}-\theta\right]}[\phi(B_1)]
$$

where $\gamma_m^n = (\overline{\mu}-\theta)I_{\widehat{A}_{m-1,n}} + (\underline{\mu}-\theta)I_{\widehat{A}^c_{m-1,n}}$ for $m = 1, \cdots, n$, and, for $m \geq 1$,

$$\widehat{A}_{m-1,n} = \left\{ \tfrac{1}{n}\sum_{i=1}^{m-1} Y_i + \tfrac{1}{\sqrt{n}}\sum_{i=1}^{m-1} \tfrac{1}{\sigma}\left(Y_i - \gamma_i^n\right) \leq -\left(\tfrac{\overline{\mu}+\underline{\mu}}{2} - \theta\right)\left(1 - \tfrac{m-1}{n}\right) + \widehat{c} \right\}.$$

Replace Y_i by $X_i - \theta$ to obtain

$$\begin{aligned}\widehat{A}_{m-1,n} &= \left\{ \tfrac{1}{n}\sum_{i=1}^{m-1} Y_i + \tfrac{1}{\sqrt{n}}\sum_{i=1}^{m-1} \tfrac{1}{\sigma}\left(Y_i - \gamma_i^n\right) \leq -\left(\tfrac{\overline{\mu}+\underline{\mu}}{2} - \theta\right)\left(1 - \tfrac{m-1}{n}\right) + \widehat{c} \right\} \\ &= \left\{ \tfrac{1}{n}\sum_{i=1}^{m-1} X_i + \tfrac{1}{\sqrt{n}}\sum_{i=1}^{m-1} \tfrac{1}{\sigma}\left(X_i - \gamma_i^n - \theta\right) \leq -\tfrac{\overline{\mu}+\underline{\mu}}{2}\left(1 - \tfrac{m-1}{n}\right) + c \right\}\end{aligned}$$

Thus $A_{0,n} = \widehat{A}_{0,n}$, and

$$\gamma_1^n + \theta = \overline{\mu}I_{\widehat{A}_{0,n}} + \underline{\mu}I_{\widehat{A}^c_{0,n}} = \overline{\mu}I_{A_{0,n}} + \underline{\mu}I_{A^c_{0,n}} = \mu_1^n$$

By induction, $A_{m-1,n} = \widehat{A}_{m-1,n}$, for $m \geq 1$, and

$$\gamma_m^n + \theta = \overline{\mu}I_{\widehat{A}_{m-1,n}} + \underline{\mu}I_{\widehat{A}^c_{m-1,n}} = \overline{\mu}I_{A_{m-1,n}} + \underline{\mu}I_{A^c_{m-1,n}} = \mu_m^n.$$

A.5 Proof of Theorem 5.1

As noted previously (Remarks 6.5 and 6.12), a suitably modified version of Lemma 6.4 is the key to proof of Theorem 5.1. Here we outline a proof of the modified lemma. We prove it in two steps.

Step 1: For every $m \geq 1$, let θ_m be a $\mathcal{G}_{m-1}$-measurable r.v. satisfying

$$|\theta_m| \leq \kappa.$$

We prove that

$$\lim_{n\to\infty}\sum_{m=1}^{n}\left|\sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}\left(T_{m-1,n} + \frac{X_m - \theta_m}{\sigma\sqrt{n}}\right)\right] - \sup_{Q\in\mathcal{P}} E_Q\left[f(\theta_m, m, n)\right]\right| = 0, \tag{A.3}$$

where $f(\theta_m, m, n)$ is given by: $f(\theta_m, m, n) =$

$$H_{m,n}(T_{m-1,n}) + H'_{m,n}(T_{m-1,n})\left(\frac{X_m - \theta_m}{\sigma\sqrt{n}}\right) + \frac{1}{2}H''_{m,n}(T_{m-1,n})\left(\frac{X_m - \theta_m}{\sigma\sqrt{n}}\right)^2. \tag{A.4}$$

Let $x = T_{m-1,n}$ and $y = \frac{X_m - \theta_m}{\sigma\sqrt{n}}$ in inequality (6.15), and obtain

$$\begin{aligned}
&\sum_{m=1}^{n} \left| \sup_{Q\in\mathcal{P}} E_Q \left[H_{m,n}\left(T_{m-1,n} + \frac{X_m - \theta_m}{\sigma\sqrt{n}} \right) \right] - \sup_{Q\in\mathcal{P}} E_Q \left[f(\theta_m, m, n) \right] \right| \\
\leq & r_1(\overline{\epsilon}, n) + r_2(C, n),
\end{aligned}$$

where

$$\begin{aligned}
r_1(\overline{\epsilon}, n) &:= \overline{\epsilon} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[\left| \frac{X_m - \theta_m}{\sigma\sqrt{n}} \right|^2 I_{\left\{ \left| \frac{X_m - \theta_m}{\sigma\sqrt{n}} \right| < \delta \right\}} \right] \\
r_2(C, n) &:= C \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[\left| \frac{X_m - \theta_m}{\sigma\sqrt{n}} \right|^2 I_{\left\{ \left| \frac{X_m - \theta_m}{\sigma\sqrt{n}} \right| \geq \delta \right\}} \right].
\end{aligned}$$

It is readily proven that, for sufficiently large n,

$$\begin{aligned}
r_1(\overline{\epsilon}, n) \leq & \frac{\overline{\epsilon}}{\sigma^2} \left(\sigma^2 + 4\kappa^2 \right) \\
r_2(C, n) \leq & \frac{2C}{n\sigma^2} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[|X_m|^2 I_{\left\{ \left| \frac{X_m - \theta_m}{\sigma\sqrt{n}} \right| \geq \delta \right\}} \right] \\
& + \frac{2C}{\sigma^2 n} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[|\theta_m|^2 I_{\left\{ \left| \frac{X_m - \theta_m}{\sigma\sqrt{n}} \right| \geq \delta \right\}} \right] \\
\leq & \frac{2C}{n\sigma^2} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[|X_m|^2 I_{\left\{ |X_m| > \sigma\sqrt{n}\delta - \kappa \right\}} \right] \\
& + \frac{2C}{\sigma^2 n} \frac{\kappa^2}{\delta^2} \sum_{m=1}^{n} \sup_{Q\in\mathcal{P}} E_Q \left[\left| \frac{X_m - \theta_m}{\sigma\sqrt{n}} \right|^2 \right].
\end{aligned}$$

By the finiteness of κ, σ and the Lindeberg condition (2.1),

$$\lim_{\overline{\epsilon}\to 0} \lim_{n\to\infty} \left(r_1(\overline{\epsilon}, n) + r_2(C, n) \right) = 0,$$

which proves (A.3).

Step 2: We take $\theta_m = E_Q[X_m|\mathcal{G}_{m-1}]$ in (A.4). Then for all $n \geq m \geq 1$,

$$\sup_{Q\in\mathcal{P}} E_Q \left[f(\theta_m, m, n) \right] = \sup_{Q\in\mathcal{P}} E_Q \left[H_{m,n}(T_{m-1,n}) + \tfrac{1}{2n} H''_{m,n}(T_{m-1,n}) \right].$$

In fact,

$$\sup_{Q\in\mathcal{P}} E_Q \left[f(\theta_m, m, n) \right]$$

$$
\begin{aligned}
&= \sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}(T_{m-1,n}) + H'_{m,n}(T_{m-1,n})\left(\frac{X_m - E_Q[X_m|\mathcal{G}_{m-1}]}{\sigma\sqrt{n}}\right)\right.\\
&\qquad\qquad\qquad\qquad \left.+\frac{1}{2}H''_{m,n}(T_{m-1,n})\left(\frac{X_m - E_Q[X_m|\mathcal{G}_{m-1}]}{\sigma\sqrt{n}}\right)^2\right]\\
&= \sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}(T_{m-1,n}) + H'_{m,n}(T_{m-1,n})E_Q\left[\left(\frac{X_m - E_Q[X_m|\mathcal{G}_{m-1}]}{\sigma\sqrt{n}}\right)|\mathcal{G}_{m-1}\right]\right.\\
&\qquad\qquad\qquad \left.+\frac{1}{2n}H''_{m,n}(T_{m-1,n})E_Q\left[\left(\frac{X_m - E_Q[X_m|\mathcal{G}_{m-1}]}{\sigma\sqrt{n}}\right)^2|\mathcal{G}_{m-1}\right]\right]\\
&= \sup_{Q\in\mathcal{P}} E_Q\left[H_{m,n}(T_{m-1,n}) + \frac{1}{2n}H''_{m,n}(T_{m-1,n})\right]\\
&= \sup_{Q\in\mathcal{P}} E_Q\left[L_{m,n}(T_{m-1,n})\right].
\end{aligned}
$$

The last equality follows when g_0 in (6.11) equals 0.

A.6 Proof of LLN: Corollary 5.2

Here we prove Corollary 5.2, showing how it can be derived from our main result Theorem 4.1, or more precisely, from the following slight generalization.

Theorem A.2. *Adopt the assumptions of Theorem 4.1. Then, for any $\varphi \in C([-\infty,\infty])$, $\beta \geq 0$ and $\alpha > 0$,*

$$
\lim_{n\to\infty}\sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{\beta}{n}\sum_{i=1}^{n}X_i + \frac{\alpha}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sigma}(X_i - E_Q[X_i|\mathcal{G}_{i-1}])\right)\right] = \mathbb{E}_g[\varphi(\alpha B_1)], \tag{A.5}
$$

where the right side of this equation is defined to be Y_0, given that (Y_t, Z_t) is the solution of the BSDE

$$
Y_t = \varphi(\alpha B_1) + \int_t^1 g(Z_s)ds - \int_t^1 Z_s dB_s,\ 0\leq t\leq 1, \tag{A.6}
$$

Here $g(z) := \frac{\beta}{\alpha}\max_{\underline{\mu}\leq\mu\leq\overline{\mu}}(\mu z)$, and (B_t) is a standard Brownian motion.

Proof: Change variables to $\tilde{X}_i = \frac{\beta}{\alpha}X_i$ and let $\tilde{\varphi}(x) = \varphi(\alpha x)$. Then $\mathbb{E}[\tilde{X}_i] = \frac{\overline{\mu}\beta}{\alpha}$, $\mathcal{E}[\tilde{X}_i] = \frac{\underline{\mu}\beta}{\alpha}$ and their variance is $\left(\frac{\beta\sigma}{\alpha}\right)^2$. Apply Theorem 4.1 to obtain

$$
\begin{aligned}
&\sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{\beta}{n}\sum_{i=1}^{n} X_i + \frac{\alpha}{\sqrt{n}\sigma}\sum_{i=1}^{n}(X_i - E_Q[X_i|\mathcal{G}_{i-1}])\right)\right] \\
&= \lim_{n\to\infty}\sup_{Q\in\mathcal{P}} E_Q\left[\tilde{\varphi}\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{X}_i + \frac{\alpha}{\sqrt{n}\beta\sigma}\sum_{i=1}^{n}(\tilde{X}_i - E_Q[\tilde{X}_i|\mathcal{G}_{i-1}])\right)\right] \\
&= \mathbb{E}_g[\tilde{\varphi}\,(B_1)] = \mathbb{E}_g[\varphi\,(\alpha B_1)].
\end{aligned}
$$

■

Proof of Corollary 5.2: It suffices to take $\varphi \in C_b^\infty(\mathbb{R})$. Let

$$
S_0 = 0,\ S_n = \sum_{i=1}^{n} X_i,\ S_n^Q = \sum_{i=1}^{n} Y_i^Q,\ Y_i^Q = \frac{1}{\sigma}\left(X_i - E_Q[X_i|\mathcal{G}_{i-1}]\right).
$$

Then, for any $Q \in \mathcal{P}$,

$$
E_Q[Y_i^Q|\mathcal{G}_{i-1}] = 0 \text{ and } E_Q[(Y_i^Q)^2|\mathcal{G}_{i-1}] = 1 \text{ for all } i.
$$

Step 1: Prove that

$$
\lim_{\alpha\to 0}\lim_{n\to\infty}\sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{S_n}{n} + \alpha\frac{S_n^Q}{\sqrt{n}}\right)\right] = \lim_{n\to\infty}\sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{S_n}{n}\right)\right].
$$

Since $\varphi \in C_b^\infty(\mathbb{R})$, φ is uniformly Lipschitz continuous, i.e. $\exists C > 0$ such that $|\varphi(x+y) - \varphi(x)| \le C|x-y|$ for $x, y \in\mathbb{R}$. Thus

$$
\begin{aligned}
&\left|\sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{S_n}{n} + \alpha\frac{S_n^Q}{\sqrt{n}}\right)\right] - \sup_{Q\in\mathcal{P}} E_Q\left[\varphi\left(\frac{S_n}{n}\right)\right]\right| \\
\le&\frac{\alpha C}{\sqrt{n}}\sup_{Q\in\mathcal{P}} E_Q\left[\left|S_n^Q\right|\right] \\
\le&\frac{\alpha C}{\sqrt{n}}\left(\sup_{Q\in\mathcal{P}} E_Q\left[\left(S_{n-1}^Q + Y_n^Q\right)^2\right]\right)^{\frac{1}{2}} \\
=&\frac{\alpha C}{\sqrt{n}}\left(\sup_{Q\in\mathcal{P}} E_Q\left[(S_{n-1}^Q)^2 + 2S_{n-1}^Q Y_n^Q + (Y_n^Q)^2\right]\right)^{\frac{1}{2}} \\
=&\frac{\alpha C}{\sqrt{n}}\left(\sup_{Q\in\mathcal{P}} E_Q\left[(S_{n-1}^Q)^2 + 2S_{n-1}^Q E_Q[Y_n^Q|\mathcal{G}_{n-1}] + E_Q[(Y_n^Q)^2|\mathcal{G}_{n-1}]\right]\right)^{\frac{1}{2}}
\end{aligned}
$$

$$
\begin{aligned}
=&\frac{\alpha C}{\sqrt{n}}\left(\sup_{Q\in\mathcal{P}} E_Q\left[(S_{n-1}^Q)^2\right]+1\right)^{\frac{1}{2}}=\cdots=\frac{\alpha C}{\sqrt{n}}\,(n)^{\frac{1}{2}}\\
=&\alpha C\to 0 \text{ as } \alpha\to 0.
\end{aligned}
$$

Step 2: Prove that if $\beta=1$ and $g(z)=\frac{1}{\alpha}\max\limits_{\underline{\mu}\le x\le\overline{\mu}}(xz)$, then

$$
\lim_{\alpha\to 0}\mathbb{E}_g[\varphi(\alpha B_1)]=\sup_{\underline{\mu}\le x\le\overline{\mu}}\varphi(x)\,. \tag{A.7}
$$

Let

$$
\mathcal{P}\equiv
\left\{Q^v: E_{P^*}[\frac{dQ^v}{dP^*}|\mathcal{F}_1]=e^{-\frac{1}{2}\int_0^1\frac{v_s^2}{\alpha^2}ds+\int_0^1\frac{v_s}{\alpha}dB_s},\,(v_t)\text{ is }\mathcal{F}_t\text{-adapted and }v\in[\underline{\mu},\overline{\mu}]\right\}
$$

where $v\in[\underline{\mu},\overline{\mu}]$ is in the sense of $\inf_{0\le s\le 1}v_s\ge\underline{\mu}$ and $\sup_{0\le s\le 1}v_s\le\overline{\mu}$ a.s..

By [6, Theorem 2.2] or [7, Lemma 3],

$$
\begin{aligned}
\mathbb{E}_g[\varphi(\alpha B_1)]&=\sup_{Q\in\mathcal{P}}E_Q\left[\varphi(\alpha B_1)\right]\\
&=\sup_{\underline{\mu}\le v\le\overline{\mu}}E_{Q^v}\left[\varphi(\alpha B_1)\right]\\
&=\sup_{\underline{\mu}\le v\le\overline{\mu}}E_{Q^v}\left[\varphi\left(\alpha\left(B_1-\int_0^1\frac{v_s}{\alpha}ds+\int_0^1\frac{v_s}{\alpha}ds\right)\right)\right]\\
&=\sup_{\underline{\mu}\le v\le\overline{\mu}}E_{Q^v}\left[\varphi\left(\alpha B_1^v+\int_0^1 v_s ds\right)\right],
\end{aligned}\tag{A.8}
$$

where $B_t^v\equiv B_t-\int_0^t\frac{v_s}{\alpha}ds$ is the Brownian motion under Q^v.

We now prove that

$$
\lim_{\alpha\to 0}\left|\sup_{\underline{\mu}\le v\le\overline{\mu}}E_{Q^v}\left[\varphi\left(\alpha B_1^v+\int_0^1 v_s ds\right)\right]-\sup_{\underline{\mu}\le v\le\overline{\mu}}E_{Q^v}\left[\varphi\left(\int_0^1 v_s ds\right)\right]\right|=0. \tag{A.9}
$$

Because φ has Lipschitz constant C>0,

$$
\begin{aligned}
&\left|\sup_{\underline{\mu}\le v\le\overline{\mu}}E_{Q^v}\left[\varphi\left(\alpha B_1^v+\int_0^1 v_s ds\right)\right]-\sup_{\underline{\mu}\le v\le\overline{\mu}}E_{Q^v}\left[\varphi\left(\int_0^1 v_s ds\right)\right]\right|\\
\le&\sup_{\underline{\mu}\le v\le\overline{\mu}}E_{Q^v}\left[\left|\varphi\left(\alpha B_1^v+\int_0^1 v_s ds\right)-\varphi\left(\int_0^1 v_s ds\right)\right|\right]
\end{aligned}
$$

$$\leq C \sup_{\underline{\mu}\leq v\leq\overline{\mu}} E_{Q^v}\left[|\alpha B_1^v|\right] \leq \alpha C \to 0 \quad \text{as } \alpha \to 0,$$

because (B_t^v) is Q^v-Brownian motion and $E_{Q^v}\left[|B_1^v|\right] \leq 1$. This proves (A.9).

We now prove (A.7): For any $x \in [\underline{\mu}, \overline{\mu}]$, let $v_s = x$, $s \in [0,1]$. Then

$$\sup_{\underline{\mu}\leq v\leq\overline{\mu}} E_{Q^v}\left[\varphi\left(\int_0^1 v_s ds\right)\right] \geq \sup_{\underline{\mu}\leq x\leq\overline{\mu}} E_{Q^v}\left[\varphi\left(x\right)\right] = \sup_{\underline{\mu}\leq x\leq\overline{\mu}} \varphi\left(x\right). \tag{A.10}$$

In addition, since $\underline{\mu} \leq \inf_{s\in[0,1]} v_s \leq \sup_{s\in[0,1]} v_s \leq \overline{\mu}$ a.s., we have

$$\sup_{\underline{\mu}\leq v\leq\overline{\mu}} \varphi\left(\int_0^1 v_s(\omega) ds\right) \leq \sup_{\underline{\mu}\leq x\leq\overline{\mu}} \varphi\left(x\right), \text{ a.s. } .$$

Therefore,

$$\begin{aligned}\sup_{\underline{\mu}\leq v\leq\overline{\mu}} E_{Q^v}\left[\varphi\left(\int_0^1 v_s ds\right)\right] &\leq \sup_{\underline{\mu}\leq v\leq\overline{\mu}} E_{Q^v}\left[\sup_{\underline{\mu}\leq v\leq\overline{\mu}} \varphi\left(\int_0^1 v_s ds\right)\right] \\ &\leq \sup_{\underline{\mu}\leq x\leq\overline{\mu}} \varphi\left(x\right),\end{aligned}$$

which implies, given (A.10), that

$$\sup_{\underline{\mu}\leq v\leq\overline{\mu}} E_{Q^v}\left[\varphi\left(\int_0^1 v_s ds\right)\right] = \sup_{\underline{\mu}\leq x\leq\overline{\mu}} \varphi\left(x\right).$$

From (A.8) and (A.9), we have

$$\begin{aligned}&\lim_{\alpha\to 0} \mathbb{E}_g[\varphi\left(\alpha B_1\right)] = \lim_{\alpha\to 0} \sup_{\underline{\mu}\leq v\leq\overline{\mu}} E_{Q^v}\left[\varphi\left(\alpha B_1^v + \int_0^1 v_s ds\right)\right] \\ &= \lim_{\alpha\to 0}\left\{\sup_{\underline{\mu}\leq v\leq\overline{\mu}} E_{Q^v}\left[\varphi\left(\alpha B_1^v + \int_0^1 v_s ds\right)\right] - \sup_{\underline{\mu}\leq v\leq\overline{\mu}} E_{Q^v}\left[\varphi\left(\int_0^1 v_s ds\right)\right]\right\} \\ &\quad + \sup_{\underline{\mu}\leq x\leq\overline{\mu}} \varphi\left(x\right) \\ &= \sup_{\underline{\mu}\leq x\leq\overline{\mu}} \varphi\left(x\right).\end{aligned}$$

The proof of Step 2 is complete.

Finally, let $\alpha \to 0$ on both sides of (A.5) and apply Steps 1 and 2. ■