Central limit theorem for discretization errors based on stopping time sampling

Emmanuel Gobet, Nicolas Landon, Uladzislau Stazhynski

Central limit theorem for discretization errors based on stopping time sampling

Emmanuel Gobet* Nicolas Landon† Uladzislau Stazhynski‡

Abstract

We study the convergence in distribution of the renormalized error arising from the discretization of a Brownian semimartingale sampled at stopping times. Our mild assumptions on the form of stopping times allow the time grid to be a combination of hitting times of stochastic domains and of Poisson-like random times. Remarkably, a Functional Central Limit Theorem holds under great generality on the semimartingale and on the form of stopping times. Furthermore, the asymptotic characteristics are quite explicit. Along the derivation of such results, we also establish some key estimates related to approximations and sensitivities of hitting time/position with respect to model and domain perturbations.

Keywords: discretization of semimartingales, functional central limit theorem, stopping times, exit time from a domain.

MSC2010: 60F05, 60H05, 60G40.

Contents

1 Introduction **2**

2 Stochastic model, random grids, main result **6**

2.1 Probabilistic model 6

2.2 Class of random discretization grids 7

2.3 Main result: functional Central Limit Theorem 12

2.4 Examples 14

3 Proof of the main result (Theorem 2.4) **16**

3.1 A more general CLT 17

3.2 Properties of exit times from domain 19

3.3 Completion of the proof of Theorem 2.4 21

*CMAP, Ecole Polytechnique and CNRS, Université Paris Saclay, Route de Saclay, 91128 Palaiseau cedex, France. Email: emmanuel.gobet@polytechnique.edu

†CMAP, Ecole Polytechnique and CNRS, Université Paris Saclay, Route de Saclay, 91128 Palaiseau cedex, France. Email: nico.landon@gmail.com

‡CMAP, Ecole Polytechnique and CNRS, Université Paris Saclay, Route de Saclay, 91128 Palaiseau cedex, France. Email: uladzislau.stazhynski@polytechnique.edu

4 Proof of the general CLT (Theorem 3.1) 24
4.1 Part I: Preliminary almost sure convergence results . . . 24
4.2 Part II: Conclusion of the proof . . . 29

5 Proofs of domain exit time properties (Lemma 3.3, Propositions 3.4 and 3.5) 32
5.1 Proof of Lemma 3.3 . . . 32
5.2 Preparing the proof of Propositions 3.4 and 3.5 . . . 33
5.3 Proofs of Propositions 3.4 and 3.5 . . . 38

A Technical proofs 39
A.1 Proof of Lemma 3.6 . . . 39
A.2 Proof of Proposition 4.2 . . . 41

B Supplementary material 44
B.1 Decomposition of symmetric matrix into non-negative and non-positive parts . 44
B.2 Fundamental lemma on the almost sure convergence of processes . . . 44
B.3 Almost sure uniform convergence of stochastic integrals w.r.t. a Brownian semi-martingale . . . 45

1 Introduction

Statement of the problem and motivation. Let S be a $\mathbb{R}^d$-valued Itô semimartingale driven by a d-dimensional Brownian motion B and let us consider the discretization of S at random stopping times $\tau_0^n = 0 < \tau_1^n < \cdots < \tau_{N_T^n}^n = T$. The number of discretization times N_T^n may be random as well. Our goal is to establish a functional Central Limit Theorem (CLT) for the renormalized discretization error process $(\sqrt{N_t^n}\mathcal{E}_t^n)_{0\le t\le T}$, where $\mathcal{E}_t^n$ is $\mathbb{R}^m$-valued and has the form $\mathcal{E}_t^n := \mathcal{E}_t^{n,1} + \mathcal{E}_t^{n,2}$ with

$$\mathcal{E}_t^{n,1} := \sum_{\tau_{i-1}^n < t} \int_{\tau_{i-1}^n}^{\tau_i^n \wedge t} \mathcal{M}_{\tau_{i-1}^n}(S_s - S_{\tau_{i-1}^n})\mathrm{d}s, \quad \mathcal{E}_t^{n,2} := \sum_{\tau_{i-1}^n < t} \int_{\tau_{i-1}^n}^{\tau_i^n \wedge t} (S_s - S_{\tau_{i-1}^n})^{\mathsf{T}} \mathcal{A}_{\tau_{i-1}^n} \mathrm{d}B_s. \quad (1.1)$$

Here, $\mathcal{M}$ and $\mathcal{A}$ are arbitrary adapted continuous processes with values in $\mathrm{Mat}_{m,d}$ and $\mathrm{Mat}_{d,d} \otimes \mathbb{R}^m$ respectively (so that $\mathcal{A}_t$ maps bilinearly $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ to $x^{\mathsf{T}}\mathcal{A}_t y \in \mathbb{R}^m$; see the notation at the end of this section). We consider quite general sequences of stopping times, combining exit times by S of random domains and Poisson-like random times, as for instance

$$\tau_i^n := \inf\{t > \tau_{i-1}^n : (S_t - S_{\tau_{i-1}^n}) \notin \varepsilon_n D_{\tau_{i-1}^n}^n\} \wedge (\tau_{i-1}^n + \varepsilon_n^2 G_{\tau_{i-1}^n}(U_{n,i}) + \Delta_{n,i}) \wedge T, \quad (1.2)$$

for some parameter $\varepsilon_n \to 0$, some stochastic domains $\mathrm{D}_\cdot^n$ indexed by time, some independent random variables $(U_{n,i})_{i,n}$, some negligible error terms $\Delta_{n,i}$. More general forms are even allowed in Section 3.

Actually, the representation (1.1) of the error term covers important applications such as those presented below, where a discretization error process can be typically decomposed into

a linear part like (1.1) and the rest, that gives negligible contribution. To illustrate this, set $\Delta S_t := S_t - S_{\tau_i^n}$ where τ_i^n is the largest discretization time before t.

1. *Integrated variance estimation.* Here the goal is to estimate $\int_0^t \mathrm{Tr}(\sigma_s \sigma_s^\mathsf{T})\mathrm{d}s$ using observations at random times (see, e.g., [RR12, LZZ13, LMR+14]). Using the Itô formula, the estimation error writes

$$\sum_{\tau_{i-1}^n < t} |\Delta S_{\tau_i^n \wedge t}|^2 - \int_0^t \mathrm{Tr}(\sigma_s \sigma_s^\mathsf{T})\mathrm{d}s = 2\int_0^t \Delta S_s^\mathsf{T} \sigma_s \mathrm{d}B_s + 2\int_0^t b_s^\mathsf{T} \Delta S_s \mathrm{d}s.$$

2. *Optimal tracking strategies.* This is related to the minimization of the tracking error of a continuous-times strategy, which, for some function $v : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$, may be written in the form

$$\int_0^t v(s, S_s)\mathrm{d}S_s - \sum_{\tau_{i-1}^n < t} v(\tau_{i-1}^n, S_{\tau_{i-1}^n})\Delta S_{\tau_i^n \wedge s} \approx \sum_{\tau_{i-1}^n < t} \int_{\tau_{i-1}^n}^{\tau_i^n \wedge t} \nabla_S v(\tau_{i-1}^n, S_{\tau_{i-1}^n})\Delta S_s \mathrm{d}S_s,$$

   which is a particular case of (1.1). See, e.g., [Fuk11b, Fuk11a, GL14, GS18b].

3. *Parametric estimation for processes.* Regarding the parametric inference of a diffusion model based on discrete time observations, the study of the asymptotic statistical fluctuations of minimum contrast estimators boils down to investigate the CLT for estimation errors of the form (1.1). See [GJ93] in the case of deterministic observation times, and [GS18c] for random observation times, where furthermore optimal observation times are derived.

Besides, the randomness of observation times is a quite common feature in real-life applications: in [GW02] the authors bring empirical evidence about the connection between volatility and inter-transaction duration in finance; in [Fuk10] a relation between the bid/ask quotation data and tick time sampling is highlighted.

Our contributions and comparison with background results. To the best of our knowledge, this is the first attempt to study the convergence in distribution of discretization errors for a general class of Itô processes and random discretization grids given by stopping times of the general form (1.2). In particular, our models for the process and the discretization times are specified directly, in simple terms and without abstract assumptions, so that verification for a specific example is quite straightforward. In addition, we provide explicitly the limit distribution (the asymptotic bias and covariance matrix) in a tractable form in terms of the underlying model. We consider both multidimensional process and multidimensional error term, which covers simultaneously most of the applications of interest. Our class of random discretization grids (1.2) includes, in particular, hitting times of general random multidimensional domains (under quite mild assumptions), but it also allows a combination of endogenous (e.g. given by hitting times) and exogenous noise (given by independent random variables, e.g. Poisson-like random times), while a majority of previous works is restricted to only one of these cases. Note that we do not impose any Markovian assumptions either on the process or on discretization times.

As a comparison, let us mention [Lan13, Chapter 7], where the second author investigates the case where S is a Markovian Stochastic Differential Equation, $D_\cdot^n = D_\cdot$ is an ellipsoid

and where there is no Poisson-like random times (i.e. $G_\cdot(.) = +\infty$); in this reference, the approach strongly uses the Markov structure of the problem and related Partial Differential Equations, thus it is quite different from the current work which offers much more flexibility on the setting.

Another situation, where a functional CLT can be derived, corresponds to one-dimensional Itô process S, see [Fuk10, FR12, RR10, RR12]: when the time step is small, this situation is locally close to a case of scaled Brownian motion which hits ± 1, for which the distribution of hitting time/location are known. Therefore, the computations of the asymptotic characteristics are easy to perform. Here, as a difference, we consider multidimensional S and general domains D^n.

Certain works (such as e.g. [AM03, AM04, LR13, ZS16]) consider the case of random but, so called, strongly predictable discretization times, possibly up to conditioning on some independent noise. This implies that conditionally to the current time, the next increment of Brownian stochastic integral can be well approximated by a Gaussian variable, and therefore all the conditional moments are quite explicit up to some negligible errors. Then a functional CLT can be derived, using the general machinery of [JS02], and it usually leads to a mixture of Gaussian variables having zero bias and zero correlation with the ambient Brownian motion B. Though important, this case is more basic compared to general stopping times.

In [Fuk11b], the author handles multidimensional S and derives CLT-like results for errors of the form (1.1). However, the asymptotic characteristics of the CLT depend on moment conditions about the increments of the driving martingale along stopping times, see [Fuk11b, Condition 2.3]. On the one hand, these conditions are natural extensions of those observed in the one-dimensional case, but on the other hand, checking these conditions in multidimensional case is really though, not to say impossible except in simple situations. Consequently, it is not clear from [Fuk11b, Condition 2.3] which sequences of stopping times are compatible with a CLT. As a comparison, in our setting, we show that the explicit and general family of stopping times as defined in (1.2) leads to a functional CLT for $(\sqrt{N_t}\mathcal{E}_t^n)_{0\leq t\leq T}$; we do not try to check [Fuk11b, Condition 2.3] and we tackle the problem directly. More general forms of stopping times are even allowed in Section 3. In our CLT results, the asymptotic Gaussian distribution may exhibit non-zero bias and non-zero correlation with the ambient Brownian motion.

To achieve this high level of generality and to derive the above CLT for general grids, we have proved several important results about approximations of exit times/positions of Brownian semimartingales from bounded domains, on sensitivities of these quantities with respect to perturbations of model and domain. All these results are of their own interest and may be useful in other problems.

Organization of the paper. In Section 2 we introduce the stochastic model for the semimartingale S and describe the class of random discretization grids under study. Further we state the main theorem of this work and provide various examples and applications of our result. Section 3 is devoted to the proof of the main theorem, which contains two important blocks: a general abstract CLT for discretization errors based on random grids (Section 3.1) and certain important properties of the semimartingale exit times from general domains (Section 3.2). The completion of the proof is given in Section 3.3. In Section 4 we continue with the proof of the general abstract CLT, while Section 5 is devoted to the proof of the semimartingale exit time properties. Supplementary material and technical results are given in Appendix.

Notation used throughout this work.

- $v \cdot w$ denotes the scalar product in $\mathbb{R}^d$.
- $\mathrm{Mat}_{m,d}$ denotes the set of $m \times d$ real matrices. $\mathrm{Tr}(.)$ and $^\intercal$ stand respectively for the trace and transpose operators.
- We write $(M)_{ij}$ for the components of a matrix M, $M_{i:}$ (resp. $M_{:i}$) its i-th row (resp. i-th column), and a^k for the components of the vector a.
- $\mathcal{S}_d$, $\mathcal{S}_d^+$ and $\mathcal{S}_d^{++}$ denote respectively the set of symmetric, positive semidefinite symmetric and positive definite symmetric real $d \times d$ matrices.
- For $M \in \mathrm{Mat}_{m,d}$ we denote by $\|M\| := \sqrt{\mathrm{Tr}(MM^\intercal)}$ its Frobenius norm. For $M \in \mathrm{Mat}_{d,d}$, we recall the easy inequality $|\mathrm{Tr}(M)| \leq \sqrt{d}\|M\|$.
- For $M \in \mathcal{S}_d$ we denote $\lambda_{\min}(M)$ and $\lambda_{\max}(M)$ the smallest and the largest eigenvalue of M.
- We denote by: $\underset{n\to+\infty}{\overset{u.c.a.s.}{\longrightarrow}}$ - a.s. convergence uniform on $[0,T]$, $\underset{n\to+\infty}{\overset{u.c.p.}{\longrightarrow}}$ - convergence in probability uniform on $[0,T]$, $\underset{[0,T]}{\overset{d}{\Longrightarrow}}$ - convergence in distribution on $[0,T]$ in the sense of processes w.r.t. the uniform topology.
- $B_d(x_0, R)$ denotes a d-dimensional closed ball with radius R and center x_0.
- $\mathcal{U}(0,1)$ stands for the distribution of a uniform random variable on $[0,1]$.
- $\mathcal{C}_{\sup}([0,T])$ denotes the normed vector space of continuous processes on $[0,T]$ with the sup-norm.
- If $f : \mathbb{R}^d \mapsto \mathbb{R}$ is a smooth function, then ∇f (resp. $\nabla^2 f$) stands for the gradient (resp. the Hessian) of f, as a row vector (resp. as a square matrix).
- A $f : \mathbb{R}^d \mapsto \mathbb{R}$ is an α-homogeneous function (for some $\alpha \in \mathbb{N}$) if $f(cx) = c^\alpha f(x)$ for all $c \geq 0, x \in \mathbb{R}^d$.
- All the further asymptotic convergences are stated through a positive deterministic sequence $(\varepsilon_n)_{n\geq 0}$ with $\varepsilon_n \to 0$. Without loss of generality and for the sake of simplicity, from now on we assume $\varepsilon_n \leq 1$ for any n.
- For any subinterval $I \subset [0,T]$ denote $N^n(I) := \#\{\tau_i^n \in I\}$ for the number of grid times in I. Let $|I|$ denote the length of I.
- In what follows, we may consider the conditional expectation of scalar random variables $\mathcal{X}$ that are non necessarily integrable. We adopt the following convention. When $\mathcal{X}$ is non-negative, $\mathbb{E}_t(\mathcal{X})$ can be properly defined as a random variable valued in $\mathbb{R}_+ \cup \{+\infty\}$. In the case of $\mathbb{E}_t(|\mathcal{X}|) < +\infty$ a.s. we define $\mathbb{E}_t(\mathcal{X}) := \mathbb{E}_t(\mathcal{X}_+) - \mathbb{E}_t(\mathcal{X}_-)$ where $\mathcal{X}_+$ and $\mathcal{X}_-$ are the positive and the negative parts of $\mathcal{X}$.

2 Stochastic model, random grids, main result

2.1 Probabilistic model

Let $T > 0$ and let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, \mathbb{P})$ be a filtered probability space supporting a d-dimensional Brownian motion $(B_t)_{0\le t\le T}$. We assume that the filtration $(\mathcal{F}_t)_{0\le t\le T}$ satisfies the usual assumptions of being right-continuous and $\mathbb{P}$-complete. Let $(S_t)_{0\le t\le T}$ be a d-dimensional continuous $\mathcal{F}$-adapted semimartingale.

Our first CLT (Theorem 2.4) and the computation of explicit limits in Section 2.4 will be derived under the following assumptions and for stopping times of the form (2.6). A slightly more general version of CLT is established in Section 3.1, for abstract stopping times satisfying some structure conditions ($\mathbf{H}_R$)-($\mathbf{H}_\mathcal{B}$).

($\mathbf{H}_S$): The process S is of the form

$$S_t = S_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}B_s, \quad t \in [0,T], \tag{2.1}$$

where

- the starting point S_0 is a $\mathcal{F}_0$-measurable random variable;
- $(b_t)_{0\le t\le T}$ is a $\mathcal{F}$-adapted d-dimensional stochastic process;
- $(\sigma_t)_{0\le t\le T}$ is a continuous $\mathcal{F}$-adapted $\mathrm{Mat}_{d,d}$-valued process, such that σ_t is invertible a.s. for all $t \in [0,T]$ and σ_0, σ_0^{-1} are bounded;
- for some a.s. finite random variable $C_\sigma > 0$ satisfying $\mathbb{E}\left(C_\sigma^4 | \mathcal{F}_0\right) < +\infty$ and a parameter $\eta_\sigma \in (0,1]$, we have

$$|\sigma_t - \sigma_s| \le C_\sigma |t-s|^{\eta_\sigma/2} \qquad \forall s,t \in [0,T] \quad \text{a.s.}$$

We remark that the boundedness of σ_0 and σ_0^{-1} above is needed mainly to guarantee that certain processes are integrable at 0 in the proof of Proposition 4.2 in Section A.2, which is an important step of our main proof. Later similar boundedness condition is assumed for some other processes for the same reason.

($\mathbf{H}_\Delta$): There exist positive $\mathcal{F}$-adapted processes $(v_t)_{0\le t\le T}$ and $(\delta_t)_{0\le t\le T}$, such that v_t is a.s. bounded and δ_t is a.s. continuous, and for which we have a.s. for all $t \in [0,T]$

$$v_t^{-1} \le \inf_{t\le s\le \psi(t)} \lambda_{\min}(\sigma_s \sigma_s^\mathsf{T}) \le \sup_{t\le s\le \psi(t)} \|\sigma_s \sigma_s^\mathsf{T}\| \le v_t, \qquad \sup_{t\le s\le \psi(t)} |b_s| \le v_t,$$

where

$$\psi(t) := \inf\{s \ge t : |S_s - S_t| \ge \delta_t\} \wedge T, \quad t \in [0,T].$$

The role of ($\mathbf{H}_\Delta$) is to ensure ω-ise uniform controls on the coefficients of S, while the process stays in a local neighborhood. This is a technical condition for the proofs, which is easily satisfied as exemplified below. In ($\mathbf{H}_\Delta$) the key assumption is that v_t is $\mathcal{F}$-adapted, so that it allows $\mathcal{F}_t$-measurable control on $[t, \psi(t)]$ for $t \in [0,T]$.

Example 1. *On $(\Omega, \mathcal{F}, \mathbb{P})$ consider a Brownian motion $(B_t)_{0\le t\le T}$ and a continuous-time Markov chain $(P_t)_{0\le t\le T}$ taking values in $\mathbb{N}_R := \{1, \dots, R\}$, that is aimed at modeling a regime-switching behavior (see [Nor98, Chapter 2]). The label $r \in \mathbb{N}_R$ stands for indexing the different regimes. The transition from state r to state r' in two successive times is given by a Frobenius matrix M_F and the distributions of time interval between two jumps are exponential distributions, with a parameter depending on M_F. Define the $\mathbb{P}$-augmented right-continuous extension $(\mathcal{F}_t)_{0\le t\le T}$ of the filtration generated by (B, P). Consider the processes*

$$\sigma_t = \sigma\left(t, (S_{s\wedge t})_{0\le s\le T}\right), \qquad b_t = b\left(P_t, t, (S_{s\wedge t})_{0\le s\le T}\right)$$

for functions $\sigma : [0,T] \times \mathcal{C}_{\sup}([0,T]) \to \mathrm{Mat}_{d,d}$ such that σ_t^{-1} exists for all $t \in [0,T]$ a.s. and $b : \mathbb{N}_R \times [0,T] \times \mathcal{C}_{\sup}([0,T]) \to \mathbb{R}^d$. Suppose that $\sigma(\cdot,\cdot)$ is continuous and that $b(r,\cdot,\cdot)$ is continuous for all $r \in \mathbb{N}_R$. Thus for a given continuous positive process v_t, since σ_t is invertible, we may choose δ_t (continuous in t) small enough, such that if the trajectory $(S_{s\wedge\psi(t)})_{0\le s\le T}$ is at distance at most δ_t from $(S_{s\wedge t})_{0\le s\le T}$ we may upper and lower bound the eigenvalues of $\sigma\left(u, (S_{s\wedge u})_{0\le s\le T}\right), u \in [t, \psi(t)]$, using v_t. Similar reasoning yields the condition on b_t in ($\mathbf{H}_\Delta$). We remark that this model is path-dependent (thus non-Markovian) and non-only driven by Brownian motions (which justifies the use of general filtration). It also includes the diffusion model $\sigma_t = \sigma(t, S_t)$ as a particular case.

2.2 Class of random discretization grids

In this section we discuss the class of random discretization grids for which we study the discretization error, in particular, for which we establish the functional CLT with explicit limit characterization.

- This class is quite large and includes the hitting times of general random domains. Notably, it allows almost arbitrary random domain processes under some mild regularity assumptions. We claim that this is the most general concrete framework (i.e. with explicit description and without any abstract assumption) for endogenously generated discretization schemes for multidimensional processes considered in the literature.
- In addition we allow to incorporate additional independent noise of quite general form while constructing the discretization times.

In particular, examples include random grids given by a combination of the hitting times of random domains with the times generated by a Poisson process having general random path-dependent intensity and independent source of randomness.

We recall that $(\varepsilon_n)_{n\ge 0}$ is a deterministic sequence with $\varepsilon_n \in (0,1]$ and $\varepsilon_n \to 0$.

2.2.1 A set of regular bounded domains

We recall that a domain is a non-empty open connected set, see [GT83, p.10]. Let $\widetilde{\mathcal{D}}$ be the set of bounded domains D in $\mathbb{R}^d$ which contains 0, and let $\mathcal{D}$ be the subset of $\widetilde{\mathcal{D}}$ which element D has a boundary ∂D of class $\mathcal{C}^2$. For any $D \in \widetilde{\mathcal{D}}$, define the signed distance $\delta_{\partial D} : \mathbb{R}^d \to \mathbb{R}$ to its boundary by

$$\delta_{\partial D}(x) := (\mathbb{1}_{x\in D} - \mathbb{1}_{x\notin D}) \inf\{|x-y| : y \in \partial D\}. \tag{2.2}$$

We recall that without any regularity on ∂D, $\delta_{\partial D}$ is a Lipschitz function with Lipschitz constant smaller than 1 (see [GT83, Section 14.6, p. 354]). For any $D^1, D^2 \in \widetilde{\mathcal{D}}$ define

$$\mu(D^1, D^2) := \sup_{x \in \partial D^1} |\delta_{\partial D^2}(x)| + \sup_{x \in \partial D^2} |\delta_{\partial D^1}(x)| .$$

The above definition is not exactly related to the usual Hausdorff distance, as described in [HP18, Chapter 2], it is slightly more adapted to our setting.

Lemma 2.1. *$\mu(.,.)$ is a distance on the set $\widetilde{\mathcal{D}}$ of domains of $\mathbb{R}^d$ containing 0.*

Proof. It is obviously non-negative and symmetric.
Assume that $\mu(D^1, D^2) = 0$ for $D^1, D^2 \in \widetilde{\mathcal{D}}$ and let us show that $D^1 = D^2$. We have $0 = \sup_{x \in \partial D^1} |\delta_{\partial D^2}(x)| = \sup_{x \in \partial D^2} |\delta_{\partial D^1}(x)|$, which means that the Hausdorff distance between the closed compact sets ∂D^1 and ∂D^2 is zero, therefore $\partial D^1 = \partial D^2$, see [HP18, Section 2.2.3]. But since D^1 and D^2 are open connected sets containing 0, we must have $D^1 = D^2$.
It remains to prove that μ satisfies to the triangular inequality: this is an easy verification that we leave to the reader. The proof is complete. □

To allow greater generality and deal with intersection of J smooth domains (to encompass domains with corners like polyhedrons) we introduce appropriate notations. For any integer $J > 0$, let

$$\mathcal{D}^J := \{(D_1, \dots, D_J) : D_j \in \mathcal{D}\}, \qquad \mathcal{D}^J_\cap := \Big\{ \bigcap_{j=1}^J D_j : D_j \in \mathcal{D} \Big\}. \tag{2.3}$$

An element of $\mathcal{D}^J$ is a sequence of J domains, while an element of $\mathcal{D}^J_\cap$ is a domain of $\mathbb{R}^d$. We generalize $\mu(\cdot,\cdot)$ to $\mu^J(\cdot,\cdot)$ on $\mathcal{D}^J$ (resp. $\mathcal{D}^J_\cap$) by setting, for any D^1, D^2 in $\mathcal{D}^J$ (resp. $\mathcal{D}^J_\cap$),

$$\mu^J(D^1, D^2) := \sum_{j=1}^J \mu(D^1_j, D^2_j),$$

with obvious definitions of D^i_j. Since μ is a distance on $\widetilde{\mathcal{D}}$, μ^J defines also a distance on $\mathcal{D}^J$ (resp. $\mathcal{D}^J_\cap$). In what follows the continuity for a $\mathcal{D}^J$ or $\mathcal{D}^J_\cap$-valued process is meant with respect to $\mu^J(\cdot,\cdot)$.

For a domain $D \in \mathcal{D}^J_\cap$, the notation εD stands naturally as $\varepsilon D := \{y \in \mathbb{R}^d : y/\varepsilon \in D\}$ and similarly for $D \in \mathcal{D}^J$.

2.2.2 Class of random discretization grids

Fix some integer $J > 0$. We consider a $\mathcal{D}^J_\cap$-valued continuous $\mathcal{F}$-adapted process $(D_t)_{0 \le t \le T}$ and a sequence of $\mathcal{D}^J_\cap$-valued continuous $\mathcal{F}$-adapted processes $\{(D^n_t)_{0 \le t \le T} : n \ge 0\}$. All these domains of $\mathcal{D}^J_\cap$ are under the form

$$D^n_t := \bigcap_{j=1}^J D^n_{j,t}, \qquad D_t := \bigcap_{j=1}^J D_{j,t}.$$

Suppose that for some positive constants $r_0, \bar{r}_0$ the initial domain D_0 verifies

$$B_d(0, r_0) \subset D_0 \subset B_d(0, \bar{r}_0) \quad \text{a.s.} \tag{2.4}$$

We will assume the following approximation and continuity properties.

($\mathbf{H}_D^1$): There exists a constant $\eta_{\mathcal{D}} > 0$ such that

$$\sup_{n\geq 0}\left(\varepsilon_n^{-\eta_{\mathcal{D}}}\sup_{0\leq t\leq T}\mu^J(D_t^n, D_t)\right) < +\infty. \tag{2.5}$$

($\mathbf{H}_D^2$): There exists a continuous $\mathcal{F}$-adapted positive process $(L_t)_{0\leq t\leq T}$ such that L_0^{-1} is a bounded random variable and the following holds a.s. for all $t \in [0,T]$ and any $D \in \{D_{j,t}^n, D_{j,t},\ n \geq 0, j = 1,\dots,J\}$

1. the signed distance $\delta_{\partial D}(\cdot)$ is $\mathcal{C}^2$ on the set $\{x \in \mathbb{R}^d : |\delta_{\partial D}(x)| \leq L_t\}$;
2. we have $\sup_{x\in D}|x| \leq L_t^{-1}$ and

$$\inf_{x:|\delta_{\partial D}(x)|\leq L_t}|\nabla\delta_{\partial D}(x)| \geq \frac{1}{2}, \qquad \sup_{x:|\delta_{\partial D}(x)|\leq L_t}(|\nabla\delta_{\partial D}(x)| + \|\nabla^2\delta_{\partial D}(x)\|) \leq L_t^{-1}.$$

Assumption ($\mathbf{H}_D^2$) ensures in a way that the main geometric characteristics of the domain (diameter, distance function, curvature) remain ω-ise locally uniformly controlled, this is a technical condition for the subsequent proofs.

Remark 1. *Actually Assumption ($\mathbf{H}_D^2$) is quite mild. Indeed, following [GT83, Lemma 14.16] for any $D \in \mathcal{D}$ there exists $L_D > 0$ such that the distance function (2.2) is $\mathcal{C}^2$ on the set $\{x \in \mathbb{R}^d : |\delta_{\partial D}(x)| \leq L_D\}$. Further, using that $\nabla\delta_{\partial D}(\cdot)$ restricted to ∂D is the inward unit vector at the boundary, the boundedness of D and ∂D, we get the existence of $L_D > 0$ such that, in addition, $\sup_{x\in D}|x| \leq L_D^{-1}$ and*

$$\inf_{x:|\delta_{\partial D}(x)|\leq L_D}|\nabla\delta_{\partial D}(x)| \geq \frac{1}{2}, \qquad \sup_{x:|\delta_{\partial D}(x)|\leq L_D}(|\nabla\delta_{\partial D}(x)| + \|\nabla^2\delta_{\partial D}(x)\|) \leq L_D^{-1}.$$

Therefore ($\mathbf{H}_D^2$) only requires some continuity and uniformity properties of L_D for the random domain-valued processes $D_{j,t}^n, D_{j,t},\ n \geq 0, j = 1,\dots,J$.

Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ supports an i.i.d. family of random variables $U := \{U_{n,i} : i, n \in \mathbb{N}\}$ with $U_{n,i} \sim \mathcal{U}(0,1)$, that are independent of $\mathcal{F}_T$. Define the filtration $\mathcal{F}_t^U := \mathcal{F}_t \vee \sigma(U)$. Let $G : (t,\omega,u) \in [0,T]\times\Omega\times[0,1] \mapsto \mathbb{R}^+ \cup \{+\infty\}$ be a $\mathcal{P}\otimes\mathcal{B}([0,1])$-measurable mapping, where $\mathcal{P}$ denotes the σ-field of predictable sets of $[0,T]\times\Omega$. In what follows, we will simply write $G_t(u)$.

Now we present the class of random discretization grids that constitutes the principal object of our analysis. Define a sequence of discretization grids $\mathcal{T} := \{\mathcal{T}^n : n \geq 0\}$ with $\mathcal{T}^n = \{\tau_i^n, i = 0,\dots,N_T^n\}$ given by

$$\begin{cases}\tau_0^n := 0,\\ \tau_i^n := \inf\{t > \tau_{i-1}^n : (S_t - S_{\tau_{i-1}^n}) \notin \varepsilon_n D_{\tau_{i-1}^n}^n\} \wedge (\tau_{i-1}^n + \varepsilon_n^2 G_{\tau_{i-1}^n}(U_{n,i}) + \Delta_{n,i}) \wedge T,\end{cases} \tag{2.6}$$

where $(\Delta_{n,i})_{n,i\in\mathbb{N}}$ is a family of random variables such that τ_i^n's are $\mathcal{F}^U$-stopping times and $\Delta_{n,i}$ is independent of $U_{m,j}$ for $m \neq n$ or $j > i$. The variables $\Delta_{n,i}$ play the role of error terms, we make an additional assumption on it later.

Remark 2. *Note that $G_t(\cdot)$ may take the value of $+\infty$. However τ_i^n is always well defined since we take the minimum with the exit time in (2.6). In particular, if $G_t(\cdot) = +\infty$ for all $t \in [0,T]$ we simply get a sequence of random grids given by exit times without exogenous source of randomness.*

We consider the counting process $N_t^n := \#\{i \geq 1 : \tau_i^n \leq t\}$ for any $t \in [0,T]$, this is a càdlàg $\mathcal{F}^U$-adapted process. Define the normed vector space

$$\mathcal{H} := \left\{ u = (u_n, n \in \mathbb{N}) : u_n \in \mathbb{R},\ \|u\|_{\mathcal{H}} := \sum_{n\in\mathbb{N}} \frac{|u_n|}{2^n} < +\infty \right\},$$

and consider the $\mathcal{H}$-valued $\mathcal{F}^U$-adapted càdlàg process $Z_t := (Z_{n,t}, n \in \mathbb{N})$ on $[0,T]$ defined by

$$Z_{n,t} := \frac{N_t^n}{N_t^n + 1}, \quad n \in \mathbb{N}.$$

Let $(\bar{\mathcal{F}}_t)_{0\leq t\leq T}$ be the right-continuous extension of the filtration $(\mathcal{F}_t \vee \sigma(Z_r, r \leq t))_{0\leq t\leq T}$. Since Z_t is $\mathcal{F}^U$-adapted and $\mathcal{F}^U$ is right-continuous, we naturally have

$$\mathcal{F}_t \subset \bar{\mathcal{F}}_t \subset \mathcal{F}_t^U. \tag{2.7}$$

Thus the filtration $\bar{\mathcal{F}}$ verifies the usual conditions. We also remark that the definition of Z_t implies that the $\mathcal{F}^U$-stopping times τ_i^n given by (2.6) are $\bar{\mathcal{F}}$-stopping times.

Suppose the following condition:

($\mathbf{H}_G$): 1. With probability 1, for all $u \in [0,1]$ the process $(G_t(u))_{0\leq t\leq T}$ is continuous on $\mathbb{R}^+ \cup \{+\infty\}$. Moreover there exists an $\mathcal{F}_T \otimes \mathcal{B}([0,1])$-measurable mapping $G_* : \Omega \times [0,1] \to \mathbb{R}^+$ not a.e. equal to zero , such that a.s. for all $n \geq 0$ and $1 \leq i \leq N_T^n$ we have

$$G_{\tau_{i-1}^n}(U_{n,i}) + \varepsilon_n^{-2}\Delta_{n,i} \geq G_*(U_{n,i}).$$

2. For some constant $\eta > 0$ and an $\bar{\mathcal{F}}$-adapted bounded process $(p_t)_{0\leq t\leq T}$ we have a.s. for all $n \geq 0$ and $1 \leq i \leq N_T^n$

$$\mathbb{E}\left(|\Delta_{n,i}||\bar{\mathcal{F}}_{\tau_{i-1}^n}\right) \leq p_{\tau_{i-1}^n}\varepsilon_n^{2+\eta}. \tag{2.8}$$

The following lemma states certain important properties of the filtration $\bar{\mathcal{F}}$.

Lemma 2.2. *The following properties hold.*

(i) The $\mathcal{F}$-Brownian motion $(B_t)_{0\leq t\leq T}$ is also a $\bar{\mathcal{F}}$-Brownian motion. Moreover any $\mathcal{F}$-adapted continuous semimartingale has the same characteristics (finite variation part, local martingale part and quadratic variation) w.r.t. $\bar{\mathcal{F}}$.

(ii) For any $\bar{\mathcal{F}}_{\tau_{i-1}^n} \otimes \mathcal{B}([0,1])$-measurable mapping $f : \Omega \times [0,1] \to \mathbb{R}^+$ we have

$$\mathbb{E}(f(\omega,U_{n,i})|\bar{\mathcal{F}}_{\tau_{i-1}^n}) = \int_0^1 f(\omega,x)\mathrm{d}x.$$

Proof. *Item* (i). Observe that [Pro04, Theorem 2, Chap. VI] ensures that any $\mathcal{F}$-semimartingale remains a $\mathcal{F}^U$-semimartingale with the same characteristics. Now we extend this property to the filtration $\bar{\mathcal{F}}$. For this, consider a square-integrable continuous $\mathcal{F}$-martingale M: using that it is a $\mathcal{F}^U$-martingale as recalled before, M is also a $\bar{\mathcal{F}}$-martingale in view of (2.7) and of the equality

$$\mathbb{E}(M_t|\bar{\mathcal{F}}_s) = \mathbb{E}(\mathbb{E}(M_t|\mathcal{F}^U_s)|\bar{\mathcal{F}}_s) = \mathbb{E}(M_s|\bar{\mathcal{F}}_s) = M_s.$$

In addition, M has the same quadratic variation $\langle M\rangle$ w.r.t. $\bar{\mathcal{F}}$ since it is characterized by the fact that $M^2 - \langle M\rangle$ is a martingale. The same conclusion can be extended to the case of local martingales since the localization times may be chosen as $\nu^k = \inf\{t \in [0,T] : \langle M\rangle_t \geq k\}$, which are $\bar{\mathcal{F}}$-stopping times, and thus by the previous argument each process $M_{\cdot\wedge\nu^k}$ is a $\bar{\mathcal{F}}$-martingale. Finally the property of having finite variation is independent of the filtration.

Item (ii). It is sufficient to show that $U_{n,i}$ is independent of $\bar{\mathcal{F}}_{\tau^n_{i-1}}$. Indeed, $U_{n,i}$ is independent of $\mathcal{F}_T$ and of $(Z_{m,t})_{0\leq t\leq T}$ for $m \neq n$. Moreover, $N_{n,\cdot}$ is a counting process, thus its natural filtration (or equivalently that of $Z_{n,\cdot}$) is right-continuous (see [Pro04, Theorem 25, Chap. I]). So, it is enough to show that $U_{n,i}$ is independent of $Z_{n,\tau^n_{i-1}}$. This follows from the construction (2.6) of the times τ^n_i and the properties of $\Delta_{n,i}$, in particular, since $U_{n,i}$ is completely unused up to the time τ^n_{i-1}, and no information about it is available at τ^n_{i-1}. □

In what follows by adapted process we mean $\bar{\mathcal{F}}$-adapted, for $\mathcal{F}$-adapted processes we will specify it explicitly if this property is needed. We also denote $\mathbb{E}_t(\cdot) := \mathbb{E}(\cdot|\bar{\mathcal{F}}_t)$.

2.2.3 Example: combination of hitting times and Poisson point process with general stochastic intensity

In this section we present the example of Poisson random times having general random path-dependent intensity and based on independent source of randomness (see [Str10] for an introduction to Poisson point processes), for which ($\mathbf{H}_G$) holds.

Let $(\lambda_t)_{0\leq t\leq T}$ be a strictly positive $\mathcal{F}$-adapted continuous stochastic process, playing the role of a stochastic intensity, and suppose that the following assumption holds.

($\mathbf{H}_\lambda$): For some constant $\eta_\lambda \in (0,1]$ we have

$$|\lambda_t - \lambda_s| \leq C_\lambda |t-s|^{\eta_\lambda}, \quad 0 \leq s \leq t \leq T, \quad \text{a.s.}$$

and, in addition, $\mathbb{E}(C_\lambda \lambda_*^{-(2+\eta_\lambda)}) < +\infty$ where $\lambda_* := \inf_{0\leq t\leq T} \lambda_t$.

For a given trajectory of $(\lambda_t)_{0\leq t\leq T}$ define a sequence of independent Poisson point processes $(\mathcal{P}^n)_{n\geq 0}$, where for each $n \geq 0$ the process $\mathcal{P}^n$ has the intensity $\{\varepsilon_n^{-2}\lambda_t, t \in [0,T]\}$ and is based on the random noise $(U_{n,i})_{i\in\mathbb{N}}$ (see (2.11) below for a precise definition). Define a sequence of random discretization grids $\mathcal{T} := \{\mathcal{T}^n : n \geq 0\}$ with $\mathcal{T}^n = \{\tau^n_i, i = 0, \ldots, N^n_T\}$ as follows

$$\begin{cases} \tau^n_0 := 0, \\ \tau^n_i := \inf\{t > \tau^n_{i-1} : (S_t - S_{\tau^n_{i-1}}) \notin \varepsilon_n D^n_{\tau^n_{i-1}} \text{ or } t \in \mathcal{P}^n\} \wedge T. \end{cases} \tag{2.9}$$

Then our claim is that $\mathcal{T}$ belongs to the class of grids described in Section 2.2.2, of the form (2.6), and it satisfies to ($\mathbf{H}_G$). Indeed, let

$$G_t(u) := -\frac{1}{\lambda_t}\log(1-u), \tag{2.10}$$

which is the inverse c.d.f. of the exponential distribution with parameter λ_t. The next Poisson time $\tilde{\tau}_i^n$ after τ_{i-1}^n is defined by the equation

$$\varepsilon_n^{-2} \int_{\tau_{i-1}^n}^{\tilde{\tau}_i^n} \lambda_s \mathrm{d}s = -\log(1-U_{n,i}), \tag{2.11}$$

so that $\Delta_{n,i}$ is such that (in view of (2.6))

$$\tilde{\tau}_i^n = \tau_{i-1}^n + \varepsilon_n^2 G_{\tau_{i-1}^n}(U_{n,i}) + \Delta_{n,i}. \tag{2.12}$$

It readily follows that

$$G_{\tau_{i-1}^n}(U_{n,i}) + \varepsilon_n^{-2}\Delta_{n,i} = \varepsilon_n^{-2}(\tilde{\tau}_i^n - \tau_{i-1}^n) \geq (\sup_{0\leq t\leq T} \lambda_t)^{-1} \varepsilon_n^{-2} \int_{\tau_{i-1}^n}^{\tilde{\tau}_i^n} \lambda_s \mathrm{d}s = (\sup_{0\leq t\leq T} \lambda_t)^{-1} |\log(1-U_{n,i})|.$$

We have completed the proof of ($\mathbf{H}_G$)-1.
Now, let us establish (2.8). Combining (2.10)-(2.11)-(2.12) and invoking Assumption ($\mathbf{H}_\lambda$), we obtain

$$|\Delta_{n,i}| = \left| \tilde{\tau}_i^n - \tau_{i-1}^n - \lambda_{\tau_{i-1}^n}^{-1} \int_{\tau_{i-1}^n}^{\tilde{\tau}_i^n} \lambda_s \mathrm{d}s \right| \leq \lambda_{\tau_{i-1}^n}^{-1} \left| \int_{\tau_{i-1}^n}^{\tilde{\tau}_i^n} |\lambda_s - \lambda_{\tau_{i-1}^n}| \mathrm{d}s \right| \leq \lambda_*^{-1} C_\lambda (\tilde{\tau}_i^n - \tau_{i-1}^n)^{1+\eta_\lambda}.$$

Further (2.11) yields

$$\tilde{\tau}_i^n - \tau_{i-1}^n \leq \lambda_*^{-1} \int_{\tau_{i-1}^n}^{\tilde{\tau}_i^n} \lambda_s \mathrm{d}s = \lambda_*^{-1} |\log(1-U_{n,i})| \varepsilon_n^2,$$

which finally implies

$$|\Delta_{n,i}| \leq C_\lambda \lambda_*^{-(2+\eta_\lambda)} |\log(1-U_{n,i})|^{1+\eta_\lambda} \varepsilon_n^{2+2\eta_\lambda}.$$

Using Lemma 2.2-(ii), we deduce that

$$\mathbb{E}_{\tau_{i-1}^n}(|\Delta_{n,i}|) \leq \left(\int_0^1 |\log(1-x)|^{1+\eta_\lambda} \mathrm{d}x \right) \mathbb{E}_{\tau_{i-1}^n} \left(C_\lambda \lambda_*^{-(2+\eta_\lambda)} \right) \varepsilon_n^{2+2\eta_\lambda}.$$

The process $\mathbb{E}_t \left(C_\lambda \lambda_*^{-(2+\eta_\lambda)} \right) < +\infty$ is a martingale due to ($\mathbf{H}_\lambda$) and thus has a cádlág version, hence it is a.s. bounded. We have proved ($\mathbf{H}_G$)-2. All in all, ($\mathbf{H}_G$) holds in this general framework of Poisson point process with stochastic intensity. □

2.3 Main result: functional Central Limit Theorem

We are now in a position to state a functional CLT for a general multidimensional discretization error in the setting presented in the previous subsections. The CLT limit is defined in terms of the solution to the following matrix-valued quadratic equation.

Lemma 2.3 ([GL14, Lemma 3.1]). *Let c be a $d\times d$-matrix symmetric non-negative real matrix. Then the equation*

$$2\,\mathrm{Tr}(x)x + 4x^2 = c \tag{2.13}$$

admits exactly one solution $x(c) \in \mathcal{S}_d^+$. Moreover, the mapping $c \mapsto x(c)$ is continuous.

Proof. We remark that in [GL14, Lemma 3.1], the input matrix on the right-hand side of (2.13) is $\tilde{c}^2$ instead of c here. Of course, it does not modify the existence and uniqueness properties in the form we state them here. Only the continuity property is questionable: in [GL14, Lemma 3.1] the continuity of $\tilde{c} \mapsto x(\tilde{c}^2) = x(c)$ is proved. However one may easily deduce the continuity of $c \mapsto x(c)$ from their proof as well: indeed, this is a direct consequence of the representation [GL14, eq. (A.7)] and of the fact that y_λ is continuous in $(\lambda_i^2)_{i=1}^d$ (in the notation of [GL14, Section A.4]). □

Fix a random grid sequence $\mathcal{T} := \{\mathcal{T}^n : n \geq 0\}$ of the form (2.6). Define

$$\begin{aligned} &\varphi(t) := \max\{\tau \in \mathcal{T}^n : \tau \leq t\}, \qquad \bar{\varphi}(t) := \min\{\tau \in \mathcal{T}^n : \tau > t\},\ \bar{\varphi}(T) := T, \\ &\Delta X_t := X_t - X_{\varphi(t)}, \end{aligned} \tag{2.14}$$

where the dependence on n is omitted for the sake of simplicity.

Let $(\mathcal{M}_t)_{0\leq t\leq T}$ and $(\mathcal{A}_t)_{0\leq t\leq T}$ be adapted continuous processes with values in $\mathrm{Mat}_{m,d}$ and $\mathrm{Mat}_{d,d}\otimes\mathbb{R}^m$ respectively (recall that an element $\mathcal{A}_t \in \mathrm{Mat}_{d,d}\otimes\mathbb{R}^m$ is given by m real $d\times d$ matrices as $[\mathcal{A}_{1,t},\dots,\mathcal{A}_{m,t}]^\mathsf{T}$ for which we write $x^\mathsf{T}\mathcal{A}_t y := [x^\mathsf{T}\mathcal{A}_{1,t}y,\dots,x^\mathsf{T}\mathcal{A}_{m,t}y]^\mathsf{T} \in \mathbb{R}^m$). Consider an $\mathbb{R}^m$-valued discretization error process given by

$$\mathcal{E}_t^n := \mathcal{E}_t^{n,1} + \mathcal{E}_t^{n,2}, \quad t \in [0,T],$$

with $\mathcal{E}_t^{n,1}$ and $\mathcal{E}_t^{n,2}$ of the form

$$\mathcal{E}_t^{n,1} := \sum_{\tau_{i-1}^n < t} \int_{\tau_{i-1}^n}^{\tau_i^n \wedge t} \mathcal{M}_{\tau_{i-1}^n} \Delta S_s \mathrm{d}s, \qquad \mathcal{E}_t^{n,2} := \sum_{\tau_{i-1}^n < t} \int_{\tau_{i-1}^n}^{\tau_i^n \wedge t} \Delta S_s^\mathsf{T} \mathcal{A}_{\tau_{i-1}^n} \mathrm{d}B_s. \tag{2.15}$$

Note that this is the most general form of an error term which is linear (or bi-linear) in terms of ΔS_s and $\mathrm{d}B_s$.

Now we introduce some processes that are involved in the explicit characterization of the limit distribution. Let W be a standard Brownian motion with $W_0 = 0$ and $U \sim \mathcal{U}(0,1)$ be independent of W, both independent of $\bar{\mathcal{F}}_T$. Set

$$\tau(t) := \inf\{s \geq 0 : \sigma_t W_s \notin D_t\} \wedge G_t(U),\ t \in [0,T].$$

In addition, for any measurable $f : \mathbb{R}^d \to \mathbb{R}$ define

$$\mathcal{B}_t[f(\cdot)] := \mathbb{E}_t\left(f(\sigma_t W_{\tau(t)})\right),\ t \in [0,T], \tag{2.16}$$

and

$$m_t := \mathbb{E}_t(\tau(t)),\ t \in [0,T]. \tag{2.17}$$

Define an $\mathbb{R}^d$-valued adapted continuous process $(Q_t)_{0\leq t\leq T}$ by

$$Q_t := \frac{1}{3} m_t^{-1} \begin{pmatrix} (\sigma_t\sigma_t^\mathsf{T})_{11}^{-1}\mathcal{B}_t[f(x) := (x^1)^3] \\ \vdots \\ (\sigma_t\sigma_t^\mathsf{T})_{dd}^{-1}\mathcal{B}_t[f(x) := (x^d)^3] \end{pmatrix}. \tag{2.18}$$

Denote $\mathcal{A}_t^{\mathsf{T}} := [\mathcal{A}_{1,t}^{\mathsf{T}}, \dots, \mathcal{A}_{m,t}^{\mathsf{T}}]^{\mathsf{T}}$ and $\mathcal{A}_t^{ij} := \frac{1}{2}(\mathcal{A}_{i,t}\mathcal{A}_{j,t}^{\mathsf{T}} + \mathcal{A}_{i,t}^{\mathsf{T}}\mathcal{A}_{j,t})$. Since $\mathcal{A}_t^{ij}$ is symmetric, by Lemma B.1 we may write $\mathcal{A}_t^{ij} = \mathcal{A}_t^{ij+} - \mathcal{A}_t^{ij-}$, where $\mathcal{A}_t^{ij+}$ and $\mathcal{A}_t^{ij-}$ are continuous symmetric non-negative definite matrices. Define a $\mathrm{Mat}_{m,m}$-valued process $(\mathcal{K}_t)_{0\le t\le T}$ by

$$\mathcal{K}_t^{ij} := m_t^{-1}\mathcal{B}_t\left[f(x) := ((\sigma_t^{-1}x)^{\mathsf{T}}X_t^{ij+}(\sigma_t^{-1}x))^2 - ((\sigma_t^{-1}x)^{\mathsf{T}}X_t^{ij-}(\sigma_t^{-1}x))^2\right] - Q_t^{\mathsf{T}}\mathcal{A}_t^{ij}Q_t, \tag{2.19}$$

for all $1 \le i, j \le m$, where X_t^{ij+} (resp. X_t^{ij-}) is the solution of the matrix equation (2.13) for $c = \sigma_t^{\mathsf{T}}\mathcal{A}_t^{ij+}\sigma_t$ (resp. $\sigma_t^{\mathsf{T}}\mathcal{A}_t^{ij-}\sigma_t$).

Here is the main result of this paper which provides the $\bar{\mathcal{F}}$-stable functional convergence of $(\sqrt{N_t^n}\mathcal{E}_t^n)_{0\le t\le T}$ in distribution as $n \to \infty$. For stable convergence, see [JS02, p. 512]-[JP12, Section 2.2.1.] for definition and properties.

Theorem 2.4. *Assume that S satisfies $(\mathbf{H}_S)$, $(\mathbf{H}_\Delta)$and $\mathcal{T}$ is given by (2.6) and satisfies $(\mathbf{H}_D^1)$, $(\mathbf{H}_D^2)$ and $(\mathbf{H}_G)$. Assume that $\mathcal{M}_0$ and $\mathcal{A}_0$ are bounded random variables. Then the processes Q and $\mathcal{K}$ are adapted continuous and $\mathcal{K}_t \in \mathcal{S}_m^+$ a.s. for all $t \in [0,T]$. Denote $\mathcal{K}_t^{1/2}$ the matrix principal square root of $\mathcal{K}_t$. Then there exists an m-dimensional Brownian motion W defined on an extended probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and independent of B such that the following functional $\bar{\mathcal{F}}$-stable convergence in distribution holds:*

$$\sqrt{N_t^n}\mathcal{E}_t^n \underset{[0,T]}{\overset{d}{\Longrightarrow}} \sqrt{\int_0^t m_s^{-1}\mathrm{d}s}\left(\int_0^t \mathcal{M}_s Q_s \mathrm{d}s + \int_0^t Q_s^{\mathsf{T}}\mathcal{A}_s \mathrm{d}B_s + \int_0^t \mathcal{K}_s^{1/2}\mathrm{d}W_s\right). \tag{2.20}$$

2.4 Examples

Below we discuss several examples where the characteristics m,Q,$\mathcal{K}$ of the limit distribution (2.20) may be explicit or easily computable using only some basic numerical calculations. We consider a general process $(S_t)_{0\le t\le T}$ verifying $(\mathbf{H}_S)$, $(\mathbf{H}_\Delta)$ and sequence of domain-valued processes $(D_t^n)_{0\le t\le T}, n \ge 0$ verifying $(\mathbf{H}_D^1)$, $(\mathbf{H}_D^2)$, while we only specify explicitly the process $(D_t)_{0\le t\le T}$.

Case $d = 1$, hitting times of stochastic time-dependent barriers. First consider the case $d = 1$, $G_t(\cdot) \equiv +\infty$ and the domain-valued process $D_t := (-\alpha_t, \beta_t) \subset \mathbb{R}$ for some adapted continuous a.s. positive processes $(\alpha_t)_{0\le t\le T}$ and $(\beta_t)_{0\le t\le T}$. Recall that

$$\tau(t) := \inf\{r > 0 : \sigma_t W_r \notin (-\alpha_t, \beta_t)\}, \qquad \mathcal{B}_t[f(\cdot)] := \mathbb{E}_t\left(f(\sigma_t W_{\tau(t)})\right).$$

In this case the distribution of $\sigma_t W_{\tau(t)}$ is explicitly known: $\mathbb{P}_t(\sigma_t W_{\tau(t)} = -\alpha_t) = \frac{\beta_t}{\alpha_t+\beta_t}$ and $\mathbb{P}_t(\sigma_t W_{\tau(t)} = \beta_t) = \frac{\alpha_t}{\alpha_t+\beta_t}$, so that $\mathcal{B}_t[f(x) := x^k] = \frac{\alpha_t\beta_t^k + (-1)^k\beta_t\alpha_t^k}{\alpha_t+\beta_t}$. In particular, an easy calculation from (2.16) and (2.17) yields

$$m_t = \mathbb{E}_t(\tau(t)) = \mathbb{E}_t((W_{\tau(t)})^2) = \alpha_t\beta_t\sigma_t^{-2}, \qquad Q_t = \frac{1}{3}m_t^{-1}\sigma_t^{-2}\mathcal{B}_t[f(x) := x^3] = \frac{1}{3}(\beta_t - \alpha_t).$$

To calculate $\mathcal{K}_t$ we remark that $\mathcal{A}_t^{11+} = (\mathcal{A}_t)^2$, $\mathcal{A}_t^{11-} = 0$ and thus $(X_t^{11+})^2 = \frac{1}{6}\sigma_t^2(\mathcal{A}_t)^2$. This further implies

$$\mathcal{K}_t = m_t^{-1}\frac{1}{6}\sigma_t^2(\mathcal{A}_t)^2\sigma_t^{-4}\mathcal{B}_t[f(x) := x^4] - Q_t^2(\mathcal{A}_t)^2 = \frac{(\mathcal{A}_t)^2}{18}(\alpha_t^2 + \beta_t^2 + \alpha_t\beta_t).$$

So finally we get

$$\sqrt{N_t^n}\mathcal{E}_t^n \xRightarrow[[0,T]]{d} \frac{1}{3}\sqrt{\int_0^t \frac{\sigma_s^2}{\alpha_s\beta_s}\mathrm{d}s}\Big(\int_0^t \mathcal{M}_s(\beta_s-\alpha_s)\mathrm{d}s + \int_0^t (\beta_s-\alpha_s)\mathcal{A}_s\mathrm{d}B_s + \frac{1}{\sqrt{2}}\int_0^t \mathcal{A}_s\sqrt{\alpha_s^2+\beta_s^2+\alpha_s\beta_s}\mathrm{d}W_s\Big). \tag{2.21}$$

From (2.21) we can easily deduce the result of [Fuk10, Theorem 3.1] (for $\varphi(x) = x$; the general case may be easily deduce by applying $\varphi^{-1}(\cdot)$ to S_t) which studies a particular case of $\alpha_t = \beta_t = 1$ and considers the estimation of integrated variance (see Section 1), so that $\mathcal{A}_t = 2\sigma_t$. In this case, invoking Theorem 3.1 yields

$$\varepsilon_n^{-1}\mathcal{E}_t^n \xRightarrow[[0,T]]{d} \int_0^t \mathcal{K}_s^{1/2}\mathrm{d}W_s$$

where $\mathcal{K}_t = \frac{2\sigma_t^2}{3}$, and Theorem 4.4 justifies that

$$\varepsilon_n^{-2}\sum_{\tau_{i-1}^n<T} |\Delta S_{\tau_i^n}|^4 \xrightarrow[n\to+\infty]{\mathbb{P}} \int_0^T \sigma_t^2\mathrm{d}t,$$

which, all in all, coincide with the results in [Fuk10, Theorem 3.1]. Theorem 2.4 uses the normalization $\sqrt{N_t^n}$, which is somewhat more natural for a CLT, and it writes

$$\sqrt{N_t^n}\mathcal{E}_t^n \xRightarrow[[0,T]]{d} \sqrt{\frac{2}{3}\int_0^t \sigma_s^2\mathrm{d}s}\int_0^t \sigma_s\mathrm{d}W_s.$$

Note that our work provides tractable limit distribution characterization in a more general setting than [Fuk10] in terms of the discretization times, the shape of the error terms; furthermore it covers the multidimensional case.

Now suppose that $G_t(\cdot)$ is not always $+\infty$. Let T_0 be deterministic and τ be the first exit time of σW from an interval $[-\alpha,\beta]$. Thus the distribution of $W_{\tau\wedge T_0}$ is equal to

$$\mathbb{P}(\tau\le T_0, \sigma W_\tau = -\alpha)\delta_{-\alpha}(\mathrm{d}x) + k(x)\mathbb{1}_{[-\alpha,\beta]}(x)\mathrm{d}x + \mathbb{P}(\tau\le T_0, \sigma W_\tau = \beta)\delta_\beta(\mathrm{d}x),$$

where, following [RY99, p.111, Exercise 3.15], $k(x)$ equals

$$\frac{1}{(2\pi T_0\sigma^2)^{1/2}}\sum_{k=-\infty}^{+\infty}\left\{\exp\left(-\frac{1}{2T_0\sigma^2}(x+2k(\alpha+\beta))^2\right) - \exp\left(-\frac{1}{2T_0\sigma^2}(x-2\beta+2k(\alpha+\beta))^2\right)\right\},$$

and, from [BS02, p.212, formulas 3.0.6],

$$\mathbb{P}(\tau\le T_0, \sigma W_\tau = -\alpha) = \int_0^{\sigma^2T_0}\mathrm{ss}_s(\beta,\alpha+\beta)\mathrm{d}s, \quad \mathbb{P}(\tau\le T_0, \sigma W_\tau = \beta) = \int_0^{\sigma^2T_0}\mathrm{ss}_s(\alpha,\alpha+\beta)\mathrm{d}s$$

for $\mathrm{ss}_t(\cdot,\cdot)$ given under an explicit form in [BS02, p.641].

Let $\mathcal{N}(\alpha,\beta,\mu,\sigma^2,p) := \int_{-\alpha}^{\beta} x^p p_{\mu,\sigma}(x)\mathrm{d}x$, where $p_{\mu,\sigma}(x) := (2\pi\sigma^2)^{-1/2}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$. Note that the explicit value of $\mathcal{N}(\alpha,\beta,\mu,\sigma^2,p)$ in terms of the standard Gaussian c.d.f. maybe easily deduced (recursively in p) via integration by parts. Further define

$$M_p(\alpha,\beta,\sigma,T_0) := \sum_{k=-\infty}^{+\infty}\Big\{\mathcal{N}(\alpha,\beta,-2k(\alpha+\beta),T_0\sigma^2,p) - \mathcal{N}(\alpha,\beta,2\beta-2k(\alpha+\beta),T_0\sigma^2,p)\Big\}.$$

Note that in practice $M_p(\alpha,\beta,\sigma,T_0)$ is well approximated by a finite sum due to the fast decay of e^{-x^2}. Now a simple calculation yields that $\mathcal{B}_t[f(x):=x^p]$ equals

$$\int_0^1 \sigma_t^p \left(M_p(\alpha_t,\beta_t,\sigma_t,G_t(u)) + \int_0^{\sigma_t^2 G_t(u)} ((-\alpha_t)^p \,\mathrm{ss}_s(\beta_t,\alpha_t+\beta_t) + \beta_t^p \,\mathrm{ss}_s(\alpha_t,\alpha_t+\beta_t))\mathrm{d}s \right) \mathrm{d}u,$$

which allows to easily deduce the explicit form of the limit distribution in (2.20) through the computations of $m, Q, \mathcal{K}$ (at least, using a numerical integration routine).

Case $d>1$, hitting times of symmetric domains, ellipsoid based grids. Suppose that for all $t\in[0,T]$ the domain D_t is symmetric (i.e. $D_t=-D_t$), denote $\tau(t)=\inf\{r>0 : \sigma_t W_r \notin D_t\}\wedge G_t(U)$. Let us prove that $Q_t=0$. Indeed, in view of (2.18), this follows from

$$\mathbb{E}_t((W^i_{\tau(D_t)\wedge T})^3) = \mathbb{E}_t((-W^i_{\tau(-D_t)\wedge T})^3) = \mathbb{E}_t((-W^i_{\tau(D_t)\wedge T})^3) = -\mathbb{E}_t((W^i_{\tau(D_t)\wedge T})^3),$$

where we denote $\tau(D)$ the first exist time of $\sigma_t W$ from a domain D, and $T>0$ is fixed.

We suppose again that $G_t(\cdot)\equiv+\infty$. Consider the case $d>1$. For an ${\mathcal{S}_d}^{++}$-valued process $(\Sigma_t)_{0\le t\le T}$ we take $D_t=\{x\in\mathbb{R}^d : x^{\mathsf{T}}\Sigma_t x\le 1\}$. Hence

$$\tau(t)=\inf\{r>0 : W_r^{\mathsf{T}}(\sigma_t^{\mathsf{T}}\Sigma_t\sigma_t)W_r\ge 1\}.$$

Let $\sigma_t^{\mathsf{T}}\Sigma_t\sigma_t = U_t^{\mathsf{T}}\Lambda_t U_t$ where U_t is orthogonal and Λ_t is diagonal. Then $\tau(t)$ is equal in distribution to $\inf\{r>0 : W_r^{\mathsf{T}}\Lambda_t W_r\ge 1\}$. To characterize explicitly the limit distribution (conditionally on σ_t) in (2.20), it is enough to calculate $\mathcal{K}_t$ (since $Q_t=0$), which requires only the calculation of $\mathbb{E}_t(\tau(t))$ and $\mathbb{E}_t\left(\prod_{i=1}^d (W^i_{\tau(t)})^{k_i}\right)$ for $k_1+\cdots+k_d=4, k_i\ge 0$.

In the case $d=2$ we need only to calculate numerically the following 3 functions

$$f_1(\lambda):=\mathbb{E}((W^1_{\tau(\lambda)})^4),\quad f_2(\lambda):=\mathbb{E}((W^1_{\tau(\lambda)}W^2_{\tau(\lambda)})^2),\quad f_3(\lambda):=\mathbb{E}((W^1_{\tau(\lambda)})^3 W^2_{\tau(\lambda)}),$$

where $\tau(\lambda):=\inf\{r>0 : (W^1_r)^2+\lambda(W^2_r)^2\ge 1\}$ for $\lambda>0$ (other calculations follow from setting $\lambda\mapsto\frac{1}{\lambda}$ and using basic scaling properties). To treat the case with general $G_t(\cdot)$ it is enough to numerically calculate the following 3 functions in 2 parameters

$$\begin{aligned} &f_1(\lambda,T_0):=\mathbb{E}((W^1_{\tau(\lambda)\wedge T_0})^4), \qquad f_2(\lambda,T_0):=\mathbb{E}((W^1_{\tau(\lambda)\wedge T_0}W^2_{\tau(\lambda)\wedge T_0})^2),\\ &f_3(\lambda,T_0):=\mathbb{E}((W^1_{\tau(\lambda)\wedge T_0})^3 W^2_{\tau(\lambda)\wedge T_0}). \end{aligned}$$

To the best of our knowledge, explicit formulas for these functions are not available and we have to resort to numerical methods like Monte Carlo methods. For related efficient schemes, see the boundary shifting scheme of [GM10], the walk on moving spheres algorithm of [DH13].

3 Proof of the main result (Theorem 2.4)

This is based on two general results: first, a CLT (Section 3.1) for discretization errors in an abstract setting; second, general properties of exit times from intersection of regular domains (Section 3.2). The proof of Theorem 2.4 is then completed in Section 3.3.

3.1 A more general CLT

The result of this section is the key ingredient of the proof of Theorem 2.4 and constitutes itself a stand-alone contribution. In particular, it generalizes the result of [Fuk11b] in our framework of multidimensional process and general multidimensional error term, with explicit limit coefficients (as opposed to the non-explicit Condition 2.3 of [Fuk11b]).

Within Section 3.1 (and Section 4 for the proofs) we are working in a slightly more abstract framework regarding S than in Section 2. Let $(\Omega, \mathcal{F}, (\bar{\mathcal{F}}_t)_{0\leq t\leq T}, \mathbb{P})$ be a filtered probability space (with $(\bar{\mathcal{F}}_t)_{0\leq t\leq T}$ satisfying the usual conditions) and consider a more general semimartingale S satisfying the following extended assumption.

($\mathbf{H}_S^{gen.}$): The process S on $[0,T]$ is given by

$$S_t = A_t + \int_0^t \sigma_s \mathrm{d}B_s, \quad t \in [0,T],$$

where

- the process A is continuous, adapted and of finite variation, and satisfies
$$|A_t - A_s| \leq C_A |t-s|^{\eta_A} \qquad \forall s,t \in [0,T] \quad \text{a.s.}, \tag{3.1}$$
for a random variable C_A, a.s. finite, and a parameter $\eta_A \in (1/2, 1]$;
- $(\sigma_t)_{0\leq t\leq T}$ is a continuous adapted $\mathrm{Mat}_{d,d}$-valued process, such that σ_t is invertible a.s. for all $t \in [0,T]$ and σ_0, σ_0^{-1} are bounded random variables;
- for some a.s. finite random variable $C_\sigma > 0$ and a parameter $\eta_\sigma \in (0,1]$, we have
$$|\sigma_t - \sigma_s| \leq C_\sigma |t-s|^{\eta_\sigma/2} \qquad \forall s,t \in [0,T] \quad \text{a.s.}$$

Let $\mathcal{T} = \{\mathcal{T}^n : n \geq 0\}$ be a sequence of discretization grids made of stopping times, where $\mathcal{T}^n = \{\tau_i^n, i = 0, \ldots, N_T^n\}$. We introduce two assumptions, whose formulation depends on the choice of a particular sequence $(\varepsilon_n)_{n\geq 0}$. For the subsequent CLT, we consider $\varepsilon_n \to 0$; with loss of generality, we assume $\varepsilon_n \leq 1$ for any n.

($\mathbf{H}_R$):

1. There exists an adapted continuous non-decreasing process $(C_t^{(3.2)})_{0\leq t\leq T}$ with bounded $C_0^{(3.2)}$, such that for $\alpha \in \{2,3,4\}$ and for all $n \geq 0$ and $1 \leq i \leq N_T^n$
$$\sup_{\tau_{i-1}^n < t \leq T} \left(\mathbb{E}_t(|S_{\tau_i^n} - S_{\tau_{i-1}^n}|^\alpha) + |S_{t\wedge\tau_i^n} - S_{\tau_{i-1}^n}|^\alpha \right) \leq C_{\tau_{i-1}^n}^{(3.2)} \varepsilon_n^\alpha \tag{3.2}$$
where $\mathbb{E}_t(.) := \mathbb{E}(. \mid \bar{\mathcal{F}}_t)$.
2. The following non-negative random variable is a.s. finite:
$$C_{(3.3)} := \sup_{n\geq 0} \left(\varepsilon_n^2 N_T^n \right) < +\infty. \tag{3.3}$$

Observe that it is enough to verify (3.2) with $\alpha = 4$, by invoking the non-expansion property of (conditional) L^p-norms.

For $\alpha \in \mathbb{N}$ we denote by $\mathcal{P}^\alpha$ the vector space spanned by α-homogeneous polynomial functions $f : \mathbb{R}^d \to \mathbb{R}$. The next set of assumptions is related to the mapping $\mathcal{B}_t[\cdot]$ arising in (2.16) in our applications. Since we deal here with a more general setting, we state a more general assumption.

($\mathbf{H}_{\mathcal{B}}$):

1. There is a linear operator $\mathcal{B}_\cdot[.]$ from the vector space spanned by $\mathcal{P}^\alpha, \alpha = 2, 3, 4$, into scalar adapted continuous process $(\mathcal{B}_t[f(\cdot)])_{0\le t\le T}$, such that the random variable $\mathcal{B}_0[f]$ is bounded for any such f.

2. The $\mathbb{R}$-valued process $m_t := \dfrac{\mathcal{B}_t[f(x) := |x|^2]}{\mathrm{Tr}(\sigma_t\sigma_t^\intercal)}$ is strictly positive and such that m_0^{-1} is bounded.

3. There exists a function $g : [0,1] \to \mathbb{R}_+$ with $\lim_{\varepsilon\to 0}(g(\varepsilon) + \varepsilon^{2(1-\rho)}g(\varepsilon)^{-1}) = 0$ for some $\rho \in (0,1)$, such that for any $f \in \mathcal{P}^\alpha$ with $\alpha \in \{2,3,4\}$ we have, for some a.s. finite random variable $C_{(3.4)}$ and a parameter $\eta \in (0,1]$, that
$$\sup_{\tau^n_{i-1}<(T-g(\varepsilon_n))_+} \left|\varepsilon_n^{-\alpha}\mathbb{E}_{\tau^n_{i-1}}(f(S_{\tau^n_i} - S_{\tau^n_{i-1}})) - \mathcal{B}_{\tau^n_{i-1}}[f(\cdot)]\right| \le C_{(3.4)}\varepsilon_n^\eta \tag{3.4}$$
for all $n \ge 0$ a.s.

4. We have $\varepsilon_n^{-2}\#\{\tau_i^n : (T - g(\varepsilon_n))_+ \le \tau_i^n \le T\} \xrightarrow[n\to+\infty]{\text{a.s.}} 0$.

The assumption ($\mathbf{H}_{\mathcal{B}}$) imposes consistency on the distribution of the discretization grids for various n and specifies a "scaling" property for the grid sequence as $n \to +\infty$. At first sight it looks like similar to [Fuk11b, Condition 2.3], but as we see in Section 3.3, it is quite tractable. Moreover, we remark that [Fuk11b, Condition 2.3] involves higher moments (up to 12, as opposed to 4 in our work) and is stated for moment ratios which makes the generalization to the multidimensional case and the practical verification of this condition much harder.

We adopt some of the notations from Section 2.3 but with the general notion of $\mathcal{B}_t[f(\cdot)]$ and m_t in ($\mathbf{H}_{\mathcal{B}}$) instead of (2.16) and (2.17), and for a general sequence of discretization grids $\mathcal{T}$. In particular, we similarly denote $\varphi(t), \bar\varphi(t)$ and ΔX_t (for any process X_t) as in (2.14).

We consider an $\mathbb{R}^m$-valued discretization error process $\mathcal{E}_t^n := \mathcal{E}_t^{n,1} + \mathcal{E}_t^{n,2}$ with $\mathcal{E}_t^{n,1}$ and $\mathcal{E}_t^{n,2}$ given by (2.15). The processes $(Q_t)_{0\le t\le T}$ and $(\mathcal{K}_t)_{0\le t\le T}$ are derived from m_t and $\mathcal{B}_t[f(\cdot)]$ in the same way as in (2.18) and (2.19). Here is a general result which provides the $\bar{\mathcal{F}}$-stable functional convergence of $(\sqrt{N_t^n}\mathcal{E}_t^n)_{0\le t\le T}$ in distribution.

Theorem 3.1. *Assume that S satisfies ($\mathbf{H}_S^{gen.}$) and consider a sequence of discretization grids $\mathcal{T} := \{\mathcal{T}^n : n \ge 0\}$ with $\mathcal{T}^n = \{\tau_i^n, i = 0, \dots, N_T^n\}$. Assume that S and $\mathcal{T}$ are such that, there is a positive sequence ε_n with $\varepsilon_n \to 0$, such that for any subsequence $(\varepsilon_{\iota(n)})_{n\ge0}$ there exists another subsequence $(\varepsilon_{\iota'\circ\iota(n)})_{n\ge0}$ for which ($\mathbf{H}_R$) and ($\mathbf{H}_{\mathcal{B}}$) hold (for this sub-subsequence). Suppose that $\mathcal{M}_0$ and $\mathcal{A}_0$ are bounded random variables.*
Then there exists an m-dimensional Brownian motion W defined on an extended probability space $(\tilde\Omega, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and independent of $\bar{\mathcal{F}}_T$ such that the following convergences hold:

1. *the functional $\bar{\mathcal{F}}$-stable convergence in distribution*
$$\varepsilon_n^{-1}\mathcal{E}_t^n \xrightarrow[[0,T]]{d} \left(\int_0^t \mathcal{M}_s Q_s \mathrm{d}s + \int_0^t Q_s^\intercal \mathcal{A}_s \mathrm{d}B_s + \int_0^t \mathcal{K}_s^{1/2}\mathrm{d}W_s\right);$$

2. *the uniform convergence in probability*
$$\varepsilon_n^2 N_t^n \xrightarrow[n\to+\infty]{u.c.p.} \int_0^t m_s^{-1}\mathrm{d}s. \tag{3.5}$$

As a consequence, this justifies the convergence in distribution for $(\sqrt{N_t^n}\mathcal{E}_t^n : 0 \le t \le T)$ in the functional sense (see [JP12, p.45]). The proof will be given in Section 4.

3.2 Properties of exit times from domain

Let B be a d-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\bar{\mathcal{F}}_t)_{t\geq 0}, \mathbb{P})$, which filtration satisfies the usual assumptions of being right-continuous and $\mathbb{P}$-complete. In this section we present some general properties of domain exit times for d-dimensional continuous Itô semimartingales $(S_t)_{0\leq t\leq T}$ and $(\bar{S}_t)_{0\leq t\leq T}$ of the form

$$S_t = \int_0^t b_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}B_s, \quad \bar{S}_t = \sigma_0 B_t, \quad t \geq 0, \tag{3.6}$$

where $(b_t)_{t\geq 0}$ and $(\sigma_t)_{t\geq 0}$ are respectively $\mathbb{R}^d$-valued and $\mathrm{Mat}_{d,d}$-valued $\bar{\mathcal{F}}$-adapted stochastic processes, satisfying some assumptions presented below. Here the starting point is $S_0 = 0$, for the sake of simplicity; actually, this is enough for our analysis, since the stopping times under study are essentially defined regarding the increments of S, extensions to $S_0 \neq 0$ would be straightforward. The subsequent results (Lemma 3.2, Propositions 3.4 and 3.5) play a key role in the proof of the CLT (Theorem 2.4, which proof is provided in Section 3.3).

($\mathbf{H}_{loc}^{D,\sigma}$): The following assumptions hold.

i) Let $J \geq 1$ and $D \in \mathcal{D}_\cap^J$ (i.e. $D = \cap_{j=1}^J D_j$ for some $D_j \in \mathcal{D}$). Define the functions $\delta_{\partial D_j} : \mathbb{R}^d \to \mathbb{R}$ which are the signed distances to ∂D_j (defined in (2.2)). Set $L_D > 0$ such that for all j we have $\delta_{\partial D_j}(\cdot) \in \mathcal{C}^2$ on $\{x : |\delta_{\partial D_j}(x)| \leq L_D\}$ and

$$\begin{aligned} &\sup_{x\in D_j} |x| \leq L_D^{-1}, \qquad \inf_{x:|\delta_{\partial D_j}(x)|\leq L_D} |\nabla \delta_{\partial D_j}(x)| \geq \frac{1}{2}, \\ &\sup_{x:|\delta_{\partial D_j}(x)|\leq L_D} (|\nabla \delta_{\partial D_j}(x)| + \|\nabla^2 \delta_{\partial D_j}(x)\|) \leq L_D^{-1}. \end{aligned} \tag{3.7}$$

ii) The $\mathrm{Mat}_{d,d}$-valued process $(\sigma_t)_{0\leq t\leq T}$ is adapted continuous, such that for all $t \geq 0$ the matrix σ_t is invertible and

$$|\sigma_t - \sigma_0| \leq C_\sigma t^{\eta_\sigma/2}, \quad \forall t \in [0, T] \quad \text{a.s.}$$

for some $\eta_\sigma > 0$ and some random variable $C_\sigma > 0$ satisfying $m_\sigma := \mathbb{E}\left(C_\sigma^4\right) < +\infty$. In addition, there exist strictly positive and finite constants $\Lambda_{\min}^\sigma, \Lambda_{\max}^\sigma, b_{\max}$ such that

$$\Lambda_{\min}^\sigma \leq \inf_{t\in[0,\tau_0]} \lambda_{\min}(\sigma_t \sigma_t^\mathsf{T}) \leq \sup_{t\in[0,\tau_0]} \|\sigma_t \sigma_t^\mathsf{T}\| \leq \Lambda_{\max}^\sigma, \qquad \sup_{t\in[0,\tau_0]} |b_t| \leq b_{\max}, \tag{3.8}$$

where we denote $\tau_0 := \inf\{t \geq 0 : S_t \notin D\}$.

Let $f \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ be an α-homogeneous function for some $\alpha \geq 2$. It is easy to check that for some constant C_f we have for all $x \in \mathbb{R}^d$

$$|f(x)| \leq C_f |x|^\alpha, \qquad |\nabla f(x)| \leq C_f |x|^{\alpha-1}, \qquad \|\nabla^2 f(x)\| \leq C_f |x|^{\alpha-2}. \tag{3.9}$$

In what follows, we fix the parameters $L_D, \eta_\sigma, m_\sigma, \Lambda_{\min}^\sigma, \Lambda_{\max}^\sigma, b_{\max}, C_f$ that are specified by the model. The following notation is quite convenient for the subsequent analysis, it will be repeatedly used.

Notation 1. *Let $\mathfrak{S}$ be a set of variables. We denote by $\mathcal{C}(\mathfrak{S})$ the set of strictly positive and continuous functions of the variables of $\mathfrak{S}$.*

Remark that such a set $\mathcal{C}(\mathfrak{S})$ is closed under addition, multiplication and all usual operations we may perform in the following analysis.

Let us fix $\mathfrak{S} := \{L_D, \eta_\sigma, m_\sigma, \Lambda^\sigma_{\min}, \Lambda^\sigma_{\max}, b_{\max}, C_f\}$. For the elements of $\mathcal{C}(\mathfrak{S})$ we will omit the dependence on the arguments, the value of a function in $\mathcal{C}(\mathfrak{S})$ is by default assumed to be equal to the value on the parameters fixed above.

Now we state the main results of this section (proofs postponed to Section 5). The next lemma is a simple technical result.

Lemma 3.2. *Assume* $(\mathbf{H}^{D,\sigma}_{loc})$. *For any* $\varepsilon \in (0,1]$ *any stopping times* $\nu_1, \nu_2 \in [0, \tau]$, *with* $\tau := \inf\{t \geq 0 : S_t \notin \varepsilon D\}$, *we have*

$$|\mathbb{E}(f(S_{\nu_1}) - f(S_{\nu_2}))| \leq C_f(b_{\max} L_D^{-(\alpha-1)} + \frac{1}{2}\sqrt{d}\Lambda^\sigma_{\max} L_D^{-(\alpha-2)})\varepsilon^{\alpha-2}\mathbb{E}\left(|\nu_1 - \nu_2|\right).$$

Proof. Using the Itô formula, the inequality $|\operatorname{Tr}(M)| \leq \sqrt{d}\|M\|$ for any $M \in \mathrm{Mat}_{m,d}$, the sub-multiplicativity of the Frobenius norm, and since $\varepsilon \leq 1$, we obtain

$$\begin{aligned}|\mathbb{E}\left(f(S_{\nu_1}) - f(S_{\nu_2})\right)| &\leq \left|\mathbb{E}\left(\int_{\nu_1}^{\nu_2} \left[\nabla f(S_t) b_t + \frac{1}{2}\operatorname{Tr}(\sigma_t^\mathsf{T} \nabla^2 f(S_t)\sigma_t)\right]\mathrm{d}t\right)\right| \\ &\leq C_f(b_{\max} L_D^{-(\alpha-1)} + \frac{1}{2}\sqrt{d}\Lambda^\sigma_{\max} L_D^{-(\alpha-2)})\varepsilon^{\alpha-2}\mathbb{E}\left(|\nu_1 - \nu_2|\right).\end{aligned}$$

□

The next results state some important properties of domain exit times, their proofs are postponed to Section 5.2. These results are interesting on their own.

Lemma 3.3. *Assume* $(\mathbf{H}^{D,\sigma}_{loc})$ *with* $D \in \mathcal{D}$ $(J = 1)$. *There exists* $R_D \in \mathcal{C}(\mathfrak{S})$ *such that, for any* $\varepsilon \in (0,1]$, $\tau = \inf\{t \geq 0 : S_t \notin \varepsilon D\}$ *and any stopping time* ν, *the following holds:*

i) for any $p \in \mathbb{N}^*$, a.s. *on the event* $\{\nu \leq \tau\}$ *we have* $\mathbb{E}_\nu\left((\tau - \nu)^p\right) \leq p!\ (R_D\varepsilon^2)^p$;

ii) for any $c \geq 0$, a.s. *on the event* $\{\nu \leq \tau\}$ *we have* a.s. $\mathbb{P}_\nu(\tau - \nu \geq \varepsilon^2 c) \leq 2e^{-\frac{c}{2R_D}}$.

The next proposition estimates the weak error between the exit values for S and $\bar{S}$.

Proposition 3.4. *Assume* $(\mathbf{H}^{D,\sigma}_{loc})$ *and let* $f \in \mathcal{C}(\mathbb{R}^d, \mathbb{R})$ *be an* α*-homogeneous function with* $\alpha \in \{2, 3, 4\}$. *There exists* $K \in \mathcal{C}(\mathfrak{S})$ *such that for any* $\varepsilon \in (0,1]$, *the stopping times*

$$\tau = \inf\{t \geq 0 : S_t \notin \varepsilon D\} \quad \text{and} \quad \bar{\tau} = \inf\{t \geq 0 : \bar{S}_t \notin \varepsilon D\}$$

satisfy, for any $T > 0$,

$$\varepsilon^{-\alpha}\left|\mathbb{E}(f(S_{\tau\wedge T}) - f(\bar{S}_{\bar{\tau}\wedge T}))\right| \leq K\varepsilon^{\eta_\sigma}. \tag{3.10}$$

The next result gives the estimation of the weak error between the exit values of S from two domains that are close to each other.

Proposition 3.5. *Assume* $(\mathbf{H}^{D,\sigma}_{loc})$ *and let* $f \in \mathcal{C}(\mathbb{R}^d, \mathbb{R})$ *be an* α*-homogeneous function with* $\alpha \in \{2, 3, 4\}$. *There exists* $K \in \mathcal{C}(\mathfrak{S} \cup \{K'\})$ *such that for any* $\varepsilon \in (0,1]$, *any strictly positive constants* K', η' *and any* $D' \in \mathcal{D}^J_\cap$ *such that* $\mu^J(D, D') \leq K'\varepsilon^{\eta'}$, *and for which* (3.7) *and* $(\mathbf{H}^{D,\sigma}_{loc})$ *hold for* D' *instead of* D *with the same constants* $L_D, \Lambda^\sigma_{\min}, \Lambda^\sigma_{\max}, b_{\max}$, *we have*

$$\varepsilon^{-\alpha}\left|\mathbb{E}(f(S_{\tau\wedge T}) - f(S_{\tau'\wedge T}))\right| \leq K\varepsilon^{\eta'},$$

for all $T > 0$, *where*

$$\tau = \inf\{t \geq 0 : S_t \notin \varepsilon D\}, \quad \tau' = \inf\{t \geq 0 : S_t \notin \varepsilon D'\}.$$

3.3 Completion of the proof of Theorem 2.4

We come back to the setting of Section 2.3. Our strategy is to apply the general CLT stated in Theorem 3.1. In particular, we aim at checking ($\mathbf{H}_R$) and ($\mathbf{H}_\mathcal{B}$) for the $\mathcal{B}_t[\cdot]$ given by (2.16) for any ε_n satisfying $\sum_{n\geq 0}\varepsilon_n^2 < +\infty$. For a general sequence $\varepsilon_n \to 0$ the result will follow in view of the subsequence formulation of Theorem 3.1: it is enough to verify the assumptions for some subsequence $\varepsilon_{\iota'\circ\iota(n)}$ (that may be chosen square summable) of arbitrary subsequence $\varepsilon_{\iota(n)}$ of ε_n.

Let us prove ($\mathbf{H}_R$)-1. Recall that we denote $\mathbb{E}_t(\cdot) := \mathbb{E}(\cdot|\bar{\mathcal{F}}_t)$. From the definition of $\mathcal{T}$ in (2.6), we have by ($\mathbf{H}_D^2$) that for all $n \geq 0$ and $1 \leq i \leq N_T^n$

$$\sup_{\tau_{i-1}^n < t \leq T} \left(\mathbb{E}_t(|S_{\tau_i^n} - S_{\tau_{i-1}^n}|^\alpha) + |S_{t\wedge\tau_i^n} - S_{\tau_{i-1}^n}|^\alpha \right) \leq 2 \left(\sup_{0\leq s\leq \tau_{i-1}^n, \alpha\in\{2,3,4\}} L_s^{-\alpha} \right) \varepsilon_n^\alpha,$$

which shows ($\mathbf{H}_R$)-1 with $C_t^{(3.2)} := 2\left(\sup_{0\leq s\leq t, \alpha\in\{2,3,4\}} L_s^{-\alpha} \right)$, so that by ($\mathbf{H}_D^2$) the process $C^{(3.2)}$ is continuous and $C_0^{(3.2)}$ is bounded.

The verification of the assumptions ($\mathbf{H}_R$)-2 and ($\mathbf{H}_\mathcal{B}$)-4 is technical, and it relies on the next Lemma, which is proved in Appendix A.1. The result below gives a quantitative comparison between the empirical measure related to the grid times and the Lebesgue measure.

Lemma 3.6. *Assume the conditions of Theorem 2.4 and $\sum_{n\geq 0}\varepsilon_n^2 < +\infty$. Then, for any sequence of non-empty deterministic intervals $I_n \subset [0,T]$, such that for some $\rho \in (0,1)$*

$$\varepsilon_n^{-(2-2\rho)}|I_n| \to +\infty, \tag{3.11}$$

there exists an a.s. finite random variable C such that

$$N^n(I_n) \leq C\varepsilon_n^{-2}|I_n|, \quad \forall n \geq 0, \text{ a.s.} \tag{3.12}$$

The condition ($\mathbf{H}_R$)-2 follows from Lemma 3.6 (with $I_n = [0,T]$ and any $\rho \in (0,1)$), while the condition ($\mathbf{H}_\mathcal{B}$)-4 follows from Lemma 3.6 with $I_n := [(T - g(\varepsilon_n))_+, T]$ and the choice $g(\varepsilon) = \varepsilon$, $\rho = 1/3$.

We now prove that the statements 1-2-3 of ($\mathbf{H}_\mathcal{B}$) hold with $\mathcal{B}[f]$ and m defined in (2.16)-(2.17). For a Brownian motion W starting at 0 and $U \sim \mathcal{U}(0,1)$ independent of W (both independent of $\bar{\mathcal{F}}_T$) let

$$\tau(t) := \inf\{s \geq 0 : \sigma_t W_s \notin D_t\} \wedge G_t(U), \tag{3.13}$$

$$\tau^n(t) := \inf\{s \geq 0 : \sigma_t W_s \notin D_t\} \wedge G_t(U) \wedge \varepsilon_n^{-2}(T-t). \tag{3.14}$$

Since a.s. D_t is a bounded domain and σ_t is invertible, $\tau(t)$ and $\tau^n(t)$ are a.s. finite random variables. Moreover $(W_{s\wedge\tau(t)} : s \geq 0)$ is a bounded martingale (with a $\mathcal{F}_t$-measurable bound depending on $\sigma_t, \sigma_t^{-1}, L_t^{-1}$), thus

$$\mathcal{B}_t[f(\cdot)] := \mathbb{E}_t\left(f(\sigma_t W_{\tau(t)})\right)$$

(given in (2.16)) is well defined for any function $f \in \mathcal{P}^\alpha, \alpha \in \{2,3,4\}$. It obviously defines a linear operator from the vector space spanned by $\mathcal{P}^\alpha, \alpha = 2,3,4$, into scalar adapted processes.

Note that $\mathcal{B}_0[f]$ is bounded owing to the boundedness of $\sigma_0, \sigma_0^{-1}, L_0^{-1}$.
The aforementioned boundedness on $W_{\cdot\wedge\tau(t)}$ implies $\mathbb{E}_t(W^i_{\tau(t)}W^j_{\tau(t)}) = 0$ for $0 \le i < j \le d$ and $\mathbb{E}_t((W^i_{\tau(t)})^2) = \mathbb{E}_t(\tau(t))$: to see these, apply the optional sampling theorem at the stopping time $\tau(t) \wedge k$ and take the limit as $k \uparrow +\infty$, each right-hand side converges using the dominated convergence theorem, each left-hand side using the monotone convergence theorem. As a consequence and using easy manipulations, we obtain the identity

$$\frac{\mathcal{B}_t[f(x) := |x|^2]}{\mathrm{Tr}(\sigma_t\sigma_t^{\mathsf{T}})} = \mathbb{E}_t(\tau(t)) \underset{(2.17)}{=} m_t.$$

Since D_t contains $0 \in \mathbb{R}^d$, $\tau(t) > 0$ a.s. and therefore $m_t > 0$ a.s.; in addition from (2.4), we get the boundedness of m_0^{-1} and $\mathcal{B}_0[f(\cdot)]$. We are done with the proof of ($\mathbf{H}_{\mathcal{B}}$)-2.

Observe that to get ($\mathbf{H}_{\mathcal{B}}$)-1, it remains only to justify the continuity of $\mathcal{B}_t[f(\cdot)]$. Using that $\cup_{0\le t\le T} D_t$ is a.s. bounded and the local Lipschitz condition of f, we have for some a.s. finite C_T and all $0 \le s \le t \le T$ that

$$|\mathcal{B}_t[f(\cdot)] - \mathcal{B}_s[f(\cdot)]| = \left|\mathbb{E}_t(f(\sigma_t W_{\tau(t)})) - \mathbb{E}_s(f(\sigma_s W_{\tau(s)}))\right| \le C_T\left(|\sigma_t - \sigma_s| + \mathbb{E}_T(|W_{\tau(t)} - W_{\tau(s)}|)\right).$$

The first term on the right-hand side is clearly continuous under our assumptions on σ. For the second, write

$$\mathbb{E}_T(|W_{\tau(t)} - W_{\tau(s)}|) \le \mathbb{E}_T(|W_{\tau(t)} - W_{\tau(s)}|^2)^{1/2} = \mathbb{E}_T(|\tau(t) - \tau(s)|)^{1/2}.$$

Let us fix t, assume $s \to t$ and let us prove that $\mathbb{E}_T(|\tau(t) - \tau(s)|) \to 0$. Define the domains $\tilde{D}_t := \sigma_t^{-1}D_t, \tilde{D}_s := \sigma_s^{-1}D_s$ (where $\sigma^{-1}D = \{\sigma^{-1}x : x \in D\}$), and set

$$\tilde{\tau}(s,t) := \inf\{r \ge 0 : W_r \notin \tilde{D}_s\} \wedge G_t(U),$$

so that

$$\mathbb{E}_T(|\tau(t) - \tau(s)|) \le \mathbb{E}_T(|\tau(t) - \tilde{\tau}(s,t)|) + \mathbb{E}_T(|\tilde{\tau}(s,t) - \tau(s)|). \tag{3.15}$$

From the continuity of σ_t and D_t (w.r.t. $\mu^J(\cdot,\cdot)$) one may check that $\mu^J(\tilde{D}_s, \tilde{D}_t) \to 0$: thus, the convergence to 0 of the first term in (3.15) readily follows by invoking Corollary 5.5 with D and D' equal to the components of $\tilde{D}_s$ and $\tilde{D}_t$ respectively (see (2.3)), with $S = W$, and making $K' \to 0$ (in the notation of Corollary 5.5).
The second term in (3.15) is bounded by $\mathbb{E}_T(|\tilde{\tau}\wedge G_t(U) - \tilde{\tau}\wedge G_s(U)|)$ (where $\tilde{\tau}$ denotes the first exit time of W from $\cup_t \tilde{D}_t$), which converges to zero by the dominated convergence theorem in view of ($\mathbf{H}_G$)-1. The proof of ($\mathbf{H}_{\mathcal{B}}$)-1 is now complete.

It remains to show the condition ($\mathbf{H}_{\mathcal{B}}$)-3 with the choice $g(\varepsilon) = \varepsilon$ made at the beginning. Fix n and i, let $f : \mathbb{R}^d \to \mathbb{R}$ be any α-homogeneous polynomial function of degree $\alpha = 2, 3, 4$. Let

$$\tilde{\tau}_i^n := \inf\{t > \tau_{i-1}^n : S_t - S_{\tau_{i-1}^n} \notin \varepsilon_n D^n_{\tau_{i-1}^n}\} \wedge (\tau_{i-1}^n + \varepsilon_n^2 G_{\tau_{i-1}^n}(U_{n,i})) \wedge T,$$

$$\hat{\tau}_i^n := \inf\{t > \tau_{i-1}^n : S_t - S_{\tau_{i-1}^n} \notin \varepsilon_n D_{\tau_{i-1}^n}\} \wedge (\tau_{i-1}^n + \varepsilon_n^2 G_{\tau_{i-1}^n}(U_{n,i})) \wedge T$$

($\hat{\tau}_i^n$ differs from $\tilde{\tau}_i^n$ by the use of $D_{\tau_{i-1}^n}$ instead of $D^n_{\tau_{i-1}^n}$ in the definition, and τ_i^n differs from $\tilde{\tau}_i^n$ by the use of $\Delta_{n,i}$ in (2.6)). Recall that by ($\mathbf{H}^2_D$) $\sup_{n\ge0}\sup_{x\in D_t\cup D_t^n}|x| \le L_t^{-1}$. Define a sequence of events $\Omega_n := \{\varepsilon_n L_t^{-1} \le \delta_t \ \forall t \in [0,T]\}, n \ge 0$, where δ_t is given by ($\mathbf{H}_\Delta$). For any

$\tau^n_{i-1} < (T-\varepsilon_n)_+$ (since we consider $g(\varepsilon)=\varepsilon$) and in view of (2.16), write

$$\begin{aligned}
&\mathbb{1}_{\Omega_n}\left|\varepsilon_n^{-\alpha}\mathbb{E}_{\tau^n_{i-1}}(f(S_{\tau^n_i}-S_{\tau^n_{i-1}}))-\mathcal{B}_{\tau^n_{i-1}}[f(\cdot)]\right| \\
&\qquad\le \mathbb{1}_{\Omega_n}\varepsilon_n^{-\alpha}\left|\mathbb{E}_{\tau^n_{i-1}}(f(S_{\tau^n_i}-S_{\tau^n_{i-1}}))-\mathbb{E}_{\tau^n_{i-1}}(f(S_{\tilde\tau^n_i}-S_{\tau^n_{i-1}}))\right| \\
&\qquad+ \mathbb{1}_{\Omega_n}\varepsilon_n^{-\alpha}\left|\mathbb{E}_{\tau^n_{i-1}}(f(S_{\tilde\tau^n_i}-S_{\tau^n_{i-1}}))-\mathbb{E}_{\tau^n_{i-1}}(f(S_{\hat\tau^n_i}-S_{\tau^n_{i-1}}))\right| \\
&\qquad+ \mathbb{1}_{\Omega_n}\varepsilon_n^{-\alpha}\left|\mathbb{E}_{\tau^n_{i-1}}(f(S_{\hat\tau^n_i}-S_{\tau^n_{i-1}}))-\mathbb{E}_{\tau^n_{i-1}}(f(\varepsilon_n\sigma_{\tau^n_{i-1}}W_{\tau^n(\tau^n_{i-1})}))\right| \\
&\qquad+ \mathbb{1}_{\Omega_n}\varepsilon_n^{-\alpha}\left|\mathbb{E}_{\tau^n_{i-1}}(f(\varepsilon_n\sigma_{\tau^n_{i-1}}W_{\tau^n(\tau^n_{i-1})}))-\mathbb{E}_{\tau^n_{i-1}}(f(\varepsilon_n\sigma_{\tau^n_{i-1}}W_{\tau(\tau^n_{i-1})}))\right|.
\end{aligned} \tag{3.16}$$

Remark that the assumption $(\mathbf{H}^{D,\sigma}_{loc})$ is verified on Ω_n for $D_{\tau^n_{i-1}}$ and $D^n_{\tau^n_{i-1}}$ due to $(\mathbf{H}_\Delta)$, η_σ given by $(\mathbf{H}_S)$, $m_\sigma=\mathbb{E}_{\tau^n_{i-1}}(C^4_\sigma)$, $\Lambda^\sigma_{\min}=v^{-1}_{\tau^n_{i-1}}$ and $\Lambda^\sigma_{\max}=b_{\max}=v_{\tau^n_{i-1}}$. In addition we may take $L_D=L_{\tau^n_{i-1}}$.
For the first term of the right-hand side of (3.16), by applying Lemma 3.2 and using that $|\tilde\tau^n_i-\tau^n_i|\le|\Delta_{n,i}|$ together with $(\mathbf{H}_G)$-2 we have for some $\bar{\mathcal{F}}_{\tau^n_{i-1}}$-measurable K and for some constant $\eta>0$

$$\mathbb{1}_{\Omega_n}\varepsilon_n^{-\alpha}\left|\mathbb{E}_{\tau^n_{i-1}}(f(S_{\tau^n_i}-S_{\tau^n_{i-1}}))-\mathbb{E}_{\tau^n_{i-1}}(f(S_{\tilde\tau^n_i}-S_{\tau^n_{i-1}}))\right|\le p_{\tau^n_{i-1}}K\varepsilon_n^{\eta}.$$

For the second term we apply Propositions 3.5 with $D=D_{\tau^n_{i-1}}$ and $D=D^n_{\tau^n_{i-1}}$ conditionally on $U_{n,i}$ and taking $T:=\varepsilon_n^2G_{\tau^n_{i-1}}(U_{n,i})\wedge(T-\tau^n_{i-1})$. Note that the necessary conditions are verified due to $(\mathbf{H}^1_D)$. Since in Proposition 3.5 the variable K is independent of T, we may further take (in view of Lemma 2.2-(ii)) expectation w.r.t. $U_{n,i}$. Thus we get for some $\bar{\mathcal{F}}_{\tau^n_{i-1}}$-measurable K and the constant $\eta_{\mathcal{D}}>0$

$$\mathbb{1}_{\Omega_n}\varepsilon_n^{-\alpha}\left|\mathbb{E}_{\tau^n_{i-1}}(f(S_{\tilde\tau^n_i}-S_{\tau^n_{i-1}}))-\mathbb{E}_{\tau^n_{i-1}}(f(S_{\hat\tau^n_i}-S_{\tau^n_{i-1}}))\right|\le K\varepsilon_n^{\eta_{\mathcal{D}}}.$$

For the third term we similarly apply Propositions 3.4 with $D=D_{\tau^n_{i-1}}$ and $D=D^n_{\tau^n_{i-1}}$ conditionally on the coupling $U_{n,i}=U$ and taking $T:=\varepsilon_n^2G_{\tau^n_{i-1}}(U)\wedge(T-\tau^n_{i-1})$. Again for some $\bar{\mathcal{F}}_{\tau^n_{i-1}}$-measurable $K>0$ (integrating with respect to $U_{n,i}=U\sim\mathcal{U}(0,1)$ since K is independent of T in Proposition 3.4, and in view of Lemma 2.2-(ii)) we get

$$\mathbb{1}_{\Omega_n}\varepsilon_n^{-\alpha}\left|\mathbb{E}_{\tau^n_{i-1}}(f(S_{\hat\tau^n_i}-S_{\tau^n_{i-1}}))-\mathbb{E}_{\tau^n_{i-1}}(f(\varepsilon_n\sigma_{\tau^n_{i-1}}W_{\tau^n(\tau^n_{i-1})}))\right|\le K\varepsilon_n^{\eta_\sigma}.$$

Finally for the last term we write using Lemma 3.3, $\tau^n_{i-1}<(T-\varepsilon_n)_+$ and (3.9), that

$$\begin{aligned}
&\mathbb{1}_{\Omega_n}\varepsilon_n^{-\alpha}\left|\mathbb{E}_{\tau^n_{i-1}}(f(\varepsilon_n\sigma_{\tau^n_{i-1}}W_{\tau^n(\tau^n_{i-1})}))-\mathbb{E}_{\tau^n_{i-1}}(f(\varepsilon_n\sigma_{\tau^n_{i-1}}W_{\tau(\tau^n_{i-1})}))\right| \\
&\qquad\le 2\mathbb{1}_{\Omega_n}C_fL^{-\alpha}_{\tau^n_{i-1}}\mathbb{P}_{\tau^n_{i-1}}(\tau(\tau^n_{i-1})>(T-\tau^n_{i-1})\varepsilon_n^{-2}) \\
&\qquad\le KC\exp(-C\varepsilon_n^{-1})\le K\varepsilon_n\sup_{x\ge0}\left(xe^{-x}\right)
\end{aligned}$$

for some a.s. finite K (independent of T and τ^n_{i-1}) and an $\bar{\mathcal{F}}_{\tau^n_{i-1}}$-measurable C.
In addition from $(\mathbf{H}_G)$-2, Lemmas 3.2, 3.3 and Propositions 3.4, 3.5 we also deduce that $\bar{\mathcal{F}}_{\tau^n_{i-1}}$-measurable K in the four latter bounds may be expressed as continuous positive simple expressions of $\eta_\sigma,\mathbb{E}_{\tau^n_{i-1}}(C^4_\sigma),v_{\tau^n_{i-1}}$ and $L_{\tau^n_{i-1}}$. This implies that, due to boundedness of the

processes v_t, L_t, $\mathbb{E}_t(C_\sigma^4)$ (since it is a martingale and thus has a càdlàg version) and also p_t, we may choose $K > 0$ uniformly in $n \geq 0$ and $i = 1, \ldots, N_T^n$ so that for all $n \geq 0$,

$$\mathbb{1}_{\Omega_n} \sup_{\tau_{i-1}^n < (T-\varepsilon_n)_+} \left| \varepsilon_n^{-\alpha} \mathbb{E}_{\tau_{i-1}^n}(f(S_{\tau_i^n} - S_{\tau_{i-1}^n})) - \mathcal{B}_{\tau_{i-1}^n}[f(\cdot)] \right| \leq K \varepsilon_n^{\eta \wedge \eta_{\mathcal{D}} \wedge \eta_\sigma \wedge 1}.$$

Finally, $\mathbb{1}_{\Omega_n} = 1$ except for a finite number of n a.s., hence we easily derive the inequality (3.4). Thus, ($\mathbf{H}_{\mathcal{B}}$)-3 is verified. The proof of Theorem 2.4 is finished. □

4 Proof of the general CLT (Theorem 3.1)

We adopt the framework of Section 3.1. The overall strategy of proof is standard and consists in proving that the drift and the quadratic variation/covariation of the error $\mathcal{E}_n$ converge in probability to some limits (see details in Subsection 4.2). The trick is to switch from convergence in probability to a.s. convergence by using a subsequence principle.

Lemma 4.1 ([Bil95, Theorem 20.5]). *Consider real-valued random variables. $\mathcal{X}_n \overset{\mathbb{P}}{\underset{n\to+\infty}{\to}} \mathcal{X}$ if, and only if, for any subsequence $(\mathcal{X}_{\iota(n)})_{n\geq 0}$ of $(\mathcal{X}_n)_{n\geq 0}$, we can extract another subsequence $(\mathcal{X}_{\iota\circ\iota'(n)})_{n\geq 0}$ such that $\mathcal{X}_{\iota\circ\iota'(n)} \overset{a.s.}{\underset{n\to+\infty}{\to}} \mathcal{X}$.*

In our framework, the flexibility in choosing another subsequence ι' is that it can be made to guarantee $\sum_{n\geq 0} \varepsilon^2_{\iota\circ\iota'(n)} < +\infty$ and to make ($\mathbf{H}_R$)-($\mathbf{H}_{\mathcal{B}}$) valid along this sequence $\tilde{\varepsilon}_n = \varepsilon_{\iota\circ\iota'(n)}$. In doing so, we define a new sequence of discretization grids $\tilde{\mathcal{T}} := \{\mathcal{T}^{\iota\circ\iota'(n)} : n \geq 0\}$. Because the new sequence $(\tilde{\varepsilon}_n : n \geq 0)$ is square summable and ($\mathbf{H}_R$)-($\mathbf{H}_{\mathcal{B}}$) hold for $(\tilde{\varepsilon}_n : n \geq 0)$, we are back to the framework of admissible sequences of discretization grids studied in [GL14, GS18b, GS18a] with a parameter $\rho_N = 1$. This latter framework is quite interesting since some a.s. results for discretization errors are already available.

The careful reader will have observed that the above references study these convergence results for admissible grid sequences in the context of a Brownian filtration $\mathcal{F}^B$ (this choice of filtration was motivated by the application at hand). However, the reader can check easily that the results of [GL14, GS18b, GS18a] hold true even if the filtration satisfies the usual assumptions of being only right continuous and $\mathbb{P}$-complete, as for $\bar{\mathcal{F}}$ in particular, because the proofs of the above references mostly use the Itô formula for the continuous semimartingale S of the form ($\mathbf{H}_S^{gen.}$) and the BDG inequalities for the Brownian integral (as in the decomposition of S), both being available when the filtration satisfies the usual assumptions.

4.1 Part I: Preliminary almost sure convergence results

We now provide some auxiliary almost sure convergence results that are necessary for the proof of Theorem 3.1. These results are, however, of their own interest and hence we put them in a separate section. In view of the above subsequence principle, these results will have to be established for a sub-subsequence $(\tilde{\varepsilon}_n : n \geq 0)$ instead of $(\varepsilon_n : n \geq 0)$. But to maintain simple notation, we keep writing ε_n (instead of $\tilde{\varepsilon}_n$), and therefore, we will have to assume that $(\varepsilon_n : n \geq 0)$ is square summable and ($\mathbf{H}_R$)-($\mathbf{H}_{\mathcal{B}}$) hold for $(\varepsilon_n : n \geq 0)$.

The next lemma allows to replace locally the values of homogeneous functions of the process increments by their conditional expectations.

Proposition 4.2. *Assume the hypotheses* ($\mathbf{H}_S^{gen.}$) *and* ($\mathbf{H}_R$) *for the sequence* $(\varepsilon_n)_{n\geq 0}$ *with* $\sum_{n\geq 0}\varepsilon_n^2 < +\infty$. *Let* $\alpha \in \{2,3,4\}$. *For any adapted continuous* $\mathcal{P}^\alpha$*-valued process* $(f_t)_{0\leq t\leq T}$ *with bounded* f_0 *(i.e. given by* $f_t = \sum_{\text{finitely many } k} f_t^k P_k$ *where* P_k *are monomials of degree* α *and* f_t^k *are adapted continuous scalar process with bounded random variables* f_0^k*), and for any adapted continuous scalar process* $(H_t)_{0\leq t\leq T}$ *with bounded* H_0, *we have*

$$\varepsilon_n^{2-\alpha}\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}\left(f_{\tau_{i-1}^n}(\Delta S_{\tau_i^n\wedge t}) - \mathbb{E}_{\tau_{i-1}^n}(f_{\tau_{i-1}^n}(\Delta S_{\tau_i^n}))\right) \overset{u.c.a.s.}{\underset{n\to+\infty}{\longrightarrow}} 0.$$

Similar convergence-in-probability results are typically deduced using the Lenglart inequality (see e.g. [Fuk11b, Proof of Lemma A.2]). However, here, since we need a.s. results to leverage the setting of admissible grid sequences, and due to lack of suitable references we provide our own proof in Section A.2.

Next, we reformulate the above convergence in a form ready to be used in combination with ($\mathbf{H}_\mathcal{B}$).

Proposition 4.3. *Assume* ($\mathbf{H}_S^{gen.}$), ($\mathbf{H}_R$) *and* ($\mathbf{H}_\mathcal{B}$) *for the sequence* $(\varepsilon_n)_{n\geq 0}$ *with* $\sum_{n\geq 0}\varepsilon_n^2 < +\infty$. *Let* $(f_t)_{0\leq t\leq T}$ *be adapted continuous* $\mathcal{P}^\alpha$*-valued process for* $\alpha \in \{2,3,4\}$ *with bounded* f_0 *(see the definition in Proposition 4.2). Then*

(i) the process $(\mathcal{B}_t[f_t(\cdot)])_{0\leq t\leq T}$ *is adapted continuous;*

(ii) for some random variable $C_{(4.1)}$ *a.s. finite and independent of* n, *we have a.s. for all* $n \geq 0$

$$\sup_{\tau_{i-1}^n<(T-g(\varepsilon_n))_+}\left|\varepsilon_n^{-\alpha}\mathbb{E}_{\tau_{i-1}^n}(f_{\tau_{i-1}^n}(S_{\tau_i^n}-S_{\tau_{i-1}^n})) - \mathcal{B}_{\tau_{i-1}^n}[f_{\tau_{i-1}^n}(\cdot)]\right| \leq C_{(4.1)}\varepsilon_n^\eta; \tag{4.1}$$

(iii) for any adapted continuous scalar process $(H_t)_{0\leq t\leq T}$ *we have*

$$\varepsilon_n^2\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}\left(\varepsilon_n^{-\alpha}\mathbb{E}_{\tau_{i-1}^n}(f_{\tau_{i-1}^n}(S_{\tau_i^n}-S_{\tau_{i-1}^n})) - \mathcal{B}_{\tau_{i-1}^n}[f_{\tau_{i-1}^n}(\cdot)]\right) \overset{u.c.a.s.}{\underset{n\to+\infty}{\longrightarrow}} 0. \tag{4.2}$$

Proof. Statements (i) and (ii) are obvious to check from ($\mathbf{H}_\mathcal{B}$)-1 and ($\mathbf{H}_\mathcal{B}$)-3.

Let us now prove (iii). Decomposing the sum in (4.2) into the contributions of the intervals $[0, t\wedge(T-g(\varepsilon_n))_+)$ and $[t\wedge(T-g(\varepsilon_n))_+, t]$, we write using (4.1)

$$\begin{aligned}
&\varepsilon_n^2\left|\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}\left(\varepsilon_n^{-\alpha}\mathbb{E}_{\tau_{i-1}^n}(f_{\tau_{i-1}^n}(S_{\tau_i^n}-S_{\tau_{i-1}^n})) - \mathcal{B}_{\tau_{i-1}^n}[f_{\tau_{i-1}^n}(\cdot)]\right)\right| \\
&\leq C_{(4.1)}\varepsilon_n^2 N_t^n \sup_{0\leq s\leq t}|H_s|\varepsilon_n^\eta \\
&+ \left(\sum_{(T-g(\varepsilon_n))_+\leq\tau_{i-1}^n<T}\varepsilon_n^2\right)\sup_{0\leq s\leq t}\left(|H_{\varphi(s)}|\left(|\varepsilon_n^{-\alpha}\mathbb{E}_{\varphi(s)}(f_{\varphi(s)}(S_{\bar\varphi(s)}-S_{\varphi(s)}))| + |\mathcal{B}_{\varphi(s)}[f_{\varphi(s)}(\cdot)]|\right)\right) \\
&\overset{u.c.a.s.}{\underset{n\to+\infty}{\longrightarrow}} 0
\end{aligned}$$

where for the first term we used that $\varepsilon_n^2 N_t^n$ is a.s. bounded owing to ($\mathbf{H}_R$)-2, and for the second term the convergence is proved by (i), ($\mathbf{H}_R$)-1 and using that $\sum_{(T-g(\varepsilon_n))_+\leq\tau_{i-1}^n<T}\varepsilon_n^2 \overset{\text{a.s.}}{\underset{n\to+\infty}{\longrightarrow}} 0$ by ($\mathbf{H}_\mathcal{B}$)-4. □

The next theorem states the convergence of the renormalized sum of process values at the discretization grid points.

Theorem 4.4. *Assume* ($\mathbf{H}_S^{gen.}$), ($\mathbf{H}_R$) *and* ($\mathbf{H}_\mathcal{B}$) *for the sequence* $(\varepsilon_n)_{n\geq 0}$ *with* $\sum_{n\geq 0}\varepsilon_n^2 < +\infty$. *Let* $(m_t)_{0\leq t\leq T}$ *be given by* ($\mathbf{H}_\mathcal{B}$)-2. *Let* $(H_t)_{0\leq t\leq T}$ *be an adapted continuous scalar process with bounded* H_0. *Let* $\alpha\in\{2,3,4\}$ *and* $(f_t)_{0\leq t\leq T}$ *be an adapted continuous* $\mathcal{P}^\alpha$*-valued process with bounded* f_0. *Then the following uniform convergences hold on* $[0,T]$*:*

$$\varepsilon_n^2\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n} \overset{u.c.a.s.}{\underset{n\to+\infty}{\longrightarrow}} \int_0^t H_s m_s^{-1}\mathrm{d}s, \tag{4.3}$$

$$\varepsilon_n^{2-\alpha}\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n} f_{\tau_{i-1}^n}(S_{\tau_i^n\wedge t}-S_{\tau_{i-1}^n}) \overset{u.c.a.s.}{\underset{n\to+\infty}{\longrightarrow}} \int_0^t H_s m_s^{-1}\mathcal{B}_s[f_s(\cdot)]\mathrm{d}s. \tag{4.4}$$

Proof. Let us first prove (4.3). The assumption ($\mathbf{H}_\mathcal{B}$)-2 reads $\mathcal{B}_t[f(x):=|x|^2]=m_t\,\mathrm{Tr}(\sigma_t\sigma_t^\mathsf{T})$, where the above right-hand side is positive continuous. Let

$$\xi_t := m_t^{-1}\,\mathrm{Tr}(\sigma_t\sigma_t^\mathsf{T})^{-1},\quad t\in[0,T]; \tag{4.5}$$

note that ξ is adapted continuous, ξ_0 is bounded in view of ($\mathbf{H}_\mathcal{B}$)-2 and ($\mathbf{H}_S^{gen.}$), and we have

$$\xi_t\mathcal{B}_t[f(x):=|x|^2]=1,\quad t\in[0,T]. \tag{4.6}$$

Now leverage the above equality to write

$$\begin{aligned}
\varepsilon_n^2\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n} &= \varepsilon_n^2\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}\xi_{\tau_{i-1}^n}\mathcal{B}_{\tau_{i-1}^n}[f(x):=|x|^2]\\
&= \sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}\xi_{\tau_{i-1}^n}|\Delta S_{\tau_i^n\wedge t}|^2\\
&\quad+\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}\xi_{\tau_{i-1}^n}\left(\mathbb{E}_{\tau_{i-1}^n}(|\Delta S_{\tau_i^n}|^2)-|\Delta S_{\tau_i^n\wedge t}|^2\right)\\
&\quad+\varepsilon_n^2\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}\xi_{\tau_{i-1}^n}\left(\mathcal{B}_{\tau_{i-1}^n}[f(x)=|x|^2]-\varepsilon_n^{-2}\mathbb{E}_{\tau_{i-1}^n}|\Delta S_{\tau_i^n}|^2\right).
\end{aligned}$$

Applying (4.2) from Proposition 4.3 with $f_t(x)=|x|^2$, $\alpha=2$, we justify that the third term above converges uniformly a.s. to 0. Further using Proposition 4.2 with $f_t(x)=|x|^2$, $\alpha=2$, the second term above also converges uniformly a.s. to 0. Finally by [GS18b, Proposition 3.8] (it easy to check in the proof that the convergence there holds in the sup-norm) we obtain

$$\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}\xi_{\tau_{i-1}^n}|\Delta S_{\tau_i^n\wedge t}|^2 \overset{u.c.a.s.}{\underset{n\to+\infty}{\longrightarrow}} \int_0^t H_s\xi_s\,\mathrm{Tr}(\sigma_s\sigma_s^\mathsf{T})\mathrm{d}s = \int_0^t H_s m_s^{-1}\mathrm{d}s,$$

where for the last equality we recast the definition of ξ. The proof of (4.3) is finished.

Regarding (4.4), write

$$\begin{aligned}
\varepsilon_n^{2-\alpha}\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n} f_{\tau_{i-1}^n}(S_{\tau_i^n\wedge t}-S_{\tau_{i-1}^n}) &= \varepsilon_n^2\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}\mathcal{B}_{\tau_{i-1}^n}[f_{\tau_{i-1}^n}(\cdot)]\\
&\quad+\varepsilon_n^2\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}\left(\varepsilon_n^{-\alpha}\mathbb{E}_{\tau_{i-1}^n}\left(f_{\tau_{i-1}^n}(S_{\tau_i^n}-S_{\tau_{i-1}^n})\right)-\mathcal{B}_{\tau_{i-1}^n}[f_{\tau_{i-1}^n}(\cdot)]\right)\\
&\quad+\varepsilon_n^{2-\alpha}\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}\left(f_{\tau_{i-1}^n}(S_{\tau_i^n\wedge t}-S_{\tau_{i-1}^n})-\mathbb{E}_{\tau_{i-1}^n}\left(f_{\tau_{i-1}^n}(S_{\tau_i^n}-S_{\tau_{i-1}^n})\right)\right).
\end{aligned} \tag{4.7}$$

Proposition 4.3 and Proposition 4.2 imply respectively that the second and the third terms in the above right-hand side converge uniformly a.s. to 0. Last, apply (4.3) to the process $(H_t\mathcal{B}_t[f_t(\cdot)])_{0\le t\le T}$ (which is adapted continuous by Proposition 4.3): this shows that the first term of the right-hand side of (4.7) converges uniformly a.s. to $\int_0^t H_s m_s^{-1}\mathcal{B}_s[f_s(\cdot)]\mathrm{d}s$. We are done. □

The next lemma gives the limit of integral of weighted increments of S.

Lemma 4.5. *Assume* $(\mathbf{H}_S^{gen.})$, $(\mathbf{H}_R)$ *and* $(\mathbf{H}_\mathcal{B})$ *for the sequence* $(\varepsilon_n)_{n\ge0}$ *with* $\sum_{n\ge0}\varepsilon_n^2<+\infty$. *Let* $(\mathcal{M}_t)_{0\le t\le T}$ *be a* $\mathrm{Mat}_{m,d}$*-valued adapted continuous process with bounded* $\mathcal{M}_0$, *and recall the definition* (2.18) *of the* $\mathbb{R}^d$*-valued adapted continuous process* $(Q_t)_{0\le t\le T}$*:*

$$Q_t := \frac{1}{3}m_t^{-1}\begin{pmatrix}(\sigma_t\sigma_t^\mathsf{T})^{-1}_{11}\mathcal{B}_t[f(x):=(x^1)^3]\\ \vdots\\ (\sigma_t\sigma_t^\mathsf{T})^{-1}_{dd}\mathcal{B}_t[f(x):=(x^d)^3]\end{pmatrix}. \tag{4.8}$$

Then

$$\varepsilon_n^{-1}\int_0^t \mathcal{M}_{\varphi(s)}\Delta S_s\mathrm{d}s \xrightarrow[n\to+\infty]{u.c.a.s.} \int_0^t \mathcal{M}_s Q_s\mathrm{d}s.$$

Proof. For any adapted continuous scalar process $(H_t)_{0\le t\le T}$ with bounded H_0 and any coordinate $k\in\{1,\dots,d\}$, the Itô formula yields that

$$\begin{aligned}\varepsilon_n^{-1}\int_0^t H_{\varphi(s)}\Delta S_s^k\mathrm{d}s = \varepsilon_n^{-1}\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}(\sigma_{\tau_{i-1}^n}\sigma_{\tau_{i-1}^n}^\mathsf{T})^{-1}_{kk}\times\\ \times\left(\frac{1}{3}(\Delta S^k_{\tau_i^n\wedge t})^3-\int_{\tau_{i-1}^n}^{\tau_i^n\wedge t}(\Delta S_s^k)^2\mathrm{d}S_s^k-\int_{\tau_{i-1}^n}^{\tau_i^n\wedge t}\Delta S_s^k\Delta(\sigma_s\sigma_s^\mathsf{T})_{kk}\mathrm{d}s\right).\end{aligned}$$

First, by Theorem 4.4 applied with $f_t(x)=(x^k)^3$ we obtain

$$\varepsilon_n^{-1}\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}(\sigma_{\tau_{i-1}^n}\sigma_{\tau_{i-1}^n}^\mathsf{T})^{-1}_{kk}(\Delta S^k_{\tau_i^n\wedge t})^3 \xrightarrow[n\to+\infty]{u.c.a.s.}\int_0^t H_s m_s^{-1}(\sigma_s\sigma_s^\mathsf{T})^{-1}_{kk}\mathcal{B}_s[f(x):=(x^k)^3]\mathrm{d}s.$$

Second, apply Lemma B.3 with $\alpha=2$ to get

$$\varepsilon_n^{-1}\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}(\sigma_{\tau_{i-1}^n}\sigma_{\tau_{i-1}^n}^\mathsf{T})^{-1}_{kk}\int_{\tau_{i-1}^n}^{\tau_i^n\wedge t}(\Delta S_s^k)^2\mathrm{d}S_s^k \xrightarrow[n\to+\infty]{u.c.a.s.} 0.$$

Finally, in view of (3.2) in $(\mathbf{H}_R)$ and using the Hölder continuity of σ in $(\mathbf{H}_S^{gen.})$, it readily follows that

$$\begin{aligned}&\left|\varepsilon_n^{-1}\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}(\sigma_{\tau_{i-1}^n}\sigma_{\tau_{i-1}^n}^\mathsf{T})^{-1}_{kk}\int_{\tau_{i-1}^n}^{\tau_i^n\wedge t}\Delta S_s^k\Delta(\sigma_s\sigma_s^\mathsf{T})_{kk}\mathrm{d}s\right|\\ &\le \varepsilon_n^{-1}\sup_{0\le s\le t}|H_s(\sigma_s\sigma_s^\mathsf{T})^{-1}_{kk}|\sup_{0\le s\le t}|\Delta S_s^k|\sup_{0\le s\le t}|\Delta(\sigma_s\sigma_s^\mathsf{T})_{kk}|\,t\le C\Big(\sup_{1\le i\le N_T^n}\Delta\tau_i^n\Big)^{\eta_\sigma/2}\end{aligned}$$

for some finite random variable C. The above time step goes almost surely to 0, this is a consequence of ($\mathbf{H}_S^{gen.}$)-($\mathbf{H}_R$), see [GS18b, Theorem 3.4 and Lemma 3.2]. All in all, this implies

$$\varepsilon_n^{-1}\int_0^t H_{\varphi(s)}\Delta S_s^k \mathrm{d}s \xrightarrow[n\to+\infty]{u.c.a.s.} \frac{1}{3}\int_0^t H_s m_s^{-1}(\sigma_s\sigma_s^\mathsf{T})^{-1}_{kk}\mathcal{B}_s[f(x):=(x^k)^3]. \tag{4.9}$$

Now, apply the above for each component $\int_0^t \mathcal{M}^{lk}_{\varphi(s)}\Delta S_s^k \mathrm{d}s$ arising in the product matrix-vector $\int_0^t \mathcal{M}_{\varphi(s)}\Delta S_s \mathrm{d}s$, we get the announced convergence. □

The next lemma handles the convergence of integral of weighted squared increments of S.

Lemma 4.6. *Assume ($\mathbf{H}_S^{gen.}$), ($\mathbf{H}_R$) and ($\mathbf{H}_\mathcal{B}$) for the sequence $(\varepsilon_n)_{n\geq 0}$ with $\sum_{n\geq 0}\varepsilon_n^2 < +\infty$. Let $(H_t)_{0\leq t\leq T}$ be an adapted continuous $\mathcal{S}_d^+$-valued process with bounded H_0. Then*

$$\varepsilon_n^{-2}\int_0^t \Delta S_s^\mathsf{T} H_{\varphi(s)}\Delta S_s \mathrm{d}s \xrightarrow[n\to+\infty]{u.c.a.s.} \int_0^t m_s^{-1}\mathcal{B}_s[f(x):=((\sigma_s^{-1}x)^\mathsf{T}X_s(\sigma_s^{-1}x))^2]\mathrm{d}s,$$

where X_s the solution of the matrix equation (2.13) for $c=\sigma_s^\mathsf{T}H_s\sigma_s$ (remark that $\sigma_s^\mathsf{T}H_s\sigma_s$ is in $\mathcal{S}_d^+$).

Proof. Set $\Lambda_s := (\sigma_s^{-1})^\mathsf{T}X_s\sigma_s^{-1}$. First observe that, owing to the properties of Lemma 2.3, X and Λ are adapted continuous processes. Moreover, multiply (2.13) (with $c=\sigma_s^\mathsf{T}H_s\sigma_s$) by $(\sigma_s^{-1})^\top$ on the left and σ_s^{-1} on the right: this gives the identity

$$2\Lambda_s\,\mathrm{Tr}(\sigma_s\sigma_s^\top\Lambda_s) + 4\Lambda_s\sigma_s\sigma_s^\top\Lambda_s = H_s. \tag{4.10}$$

Besides, for $\tau_{i-1}^n < t$, the Itô formula gives

$$\begin{aligned}(\Delta S^\mathsf{T}_{\tau_i^n\wedge t}\Lambda_{\tau_{i-1}^n}\Delta S_{\tau_i^n\wedge t})^2 &= 4\int_{\tau_{i-1}^n}^{\tau_i^n\wedge t}\Delta S_s^\top\Lambda_{\varphi(s)}\Delta S_s\Delta S_s^\top\Lambda_{\varphi(s)}\mathrm{d}S_s\\ &\quad+\int_{\tau_{i-1}^n}^{\tau_i^n\wedge t}\Delta S_s^\top\left[2\Lambda_{\varphi(s)}\,\mathrm{Tr}(\sigma_s\sigma_s^\top\Lambda_{\varphi(s)})+4\Lambda_{\varphi(s)}\sigma_s\sigma_s^\top\Lambda_{\varphi(s)}^\top\right]\Delta S_s\mathrm{d}s.\end{aligned}$$

Therefore, summing over i for $\tau_{i-1}^n < t$ and using the idendity (4.10), we get

$$\begin{aligned}\varepsilon_n^{-2}\sum_{\tau_{i-1}^n<t}(\Delta S^\mathsf{T}_{\tau_i^n\wedge t}\Lambda_{\tau_{i-1}^n}\Delta S_{\tau_i^n\wedge t})^2 &= 4\varepsilon_n^{-2}\int_0^t\Delta S_s^\top\Lambda_{\varphi(s)}\Delta S_s\Delta S_s^\top\Lambda_{\varphi(s)}\mathrm{d}S_s\\ &\quad+\varepsilon_n^{-2}\int_0^t\Delta S_s^\top\left[2\Lambda_{\varphi(s)}\,\mathrm{Tr}(\Delta(\sigma_s\sigma_s^\top)\Lambda_{\varphi(s)})+4\Lambda_{\varphi(s)}\Delta(\sigma_s\sigma_s^\top)\Lambda_{\varphi(s)}^\top\right]\Delta S_s\mathrm{d}s\\ &\quad+\varepsilon_n^{-2}\int_0^t\Delta S_s^\top H_{\varphi(s)}\Delta S_s\mathrm{d}s.\end{aligned} \tag{4.11}$$

Lemma B.3 with $\alpha=3$ implies that $\varepsilon_n^{-2}\int_0^t\Delta S_s^\top\Lambda_{\varphi(s)}\Delta S_s\Delta S_s^\top\Lambda_{\varphi(s)}\mathrm{d}S_s \xrightarrow[n\to+\infty]{u.c.a.s.} 0$. Moreover, the Hölder continuity of σ in ($\mathbf{H}_S^{gen.}$) and the bound (3.2) of ($\mathbf{H}_R$) ensure the existence of a a.s. finite random variable $C>0$ such that

$$\sup_{t\leq T}\left|\varepsilon_n^{-2}\int_0^t\Delta S_s^\top\left[2\Lambda_{\varphi(s)}\,\mathrm{Tr}(\Delta(\sigma_s\sigma_s^\top)\Lambda_{\varphi(s)})+4\Lambda_{\varphi(s)}\Delta(\sigma_s\sigma_s^\top)\Lambda_{\varphi(s)}^\top\right]\Delta S_s\mathrm{d}s\right| \leq C\Big(\sup_{1\leq i\leq N_T^n}\Delta\tau_i^n\Big)^{\eta_\sigma/2}.$$

The latter bound converges to 0, see the arguments in the proof of Lemma 4.5. Therefore, from (4.11), we obtain

$$\varepsilon_n^{-2}\int_0^t \Delta S_s^{\mathsf{T}} H_{\varphi(s)} \Delta S_s \mathrm{d}s - \varepsilon_n^{-2} \sum_{\tau_{i-1}^n < t} (\Delta S_{\tau_i^n \wedge t}^{\mathsf{T}} \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n \wedge t})^2 \xrightarrow[n\to+\infty]{u.c.a.s.} 0.$$

Observe that due to the boundedness of σ_0, σ_0^{-1} and H_0, and the properties of the solution of (2.13), the coefficients of X_0 and Λ_0 are bounded random variables. Thus, we can apply Theorem 4.4 with $\alpha = 4$ and $f_s(x) := (x^{\mathsf{T}}\Lambda_s x)^2$, to obtain

$$\begin{aligned}\varepsilon_n^{-2} \sum_{\tau_{i-1}^n < t} (\Delta S_{\tau_i^n}^{\mathsf{T}} \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n})^2 &\xrightarrow[n\to+\infty]{u.c.a.s.} \int_0^t m_s^{-1} \mathcal{B}_s[f(x) := (x^{\mathsf{T}}\Lambda_s x)^2]\mathrm{d}s \\ &= \int_0^t m_s^{-1} \mathcal{B}_s[f(x) := ((\sigma_s^{-1}x)^{\mathsf{T}} X_s (\sigma_s^{-1}x))^2]\mathrm{d}s.\end{aligned}$$

The proof is complete. □

4.2 Part II: Conclusion of the proof

Now we are in a position to finish the proof of Theorem 3.1. It boils down to combine previous preliminary results with the application of an abstract CLT for semimartingale sequences. The reference result on this subject is [JS02, Chapter IX, Theorem 7.3]. Next we state a theorem that essentially follows from a simplified version of this general result given in [Fuk11b, Theorem A.1]. For notions of stable convergence in distribution, [JS02, p. 512]-[JP12, Section 2.2.1.].

Theorem 4.7. *Let $(\Omega, \mathcal{F}, (\bar{\mathcal{F}}_t)_{0\le t\le T}, \mathbb{P})$ be a filtered probability space supporting a $\bar{\mathcal{F}}$-adapted d-dimensional Brownian motion $(B_t)_{0\le t\le T}$. Let $(S^n)_{n\ge 0}$ be a sequence of adapted continuous semimartingales of the form*

$$S^n = A^n + M^n,$$

where M^n are $\mathbb{R}^m$-valued $\bar{\mathcal{F}}$-local martingales of the form $M^n = \int_0^{\cdot} \alpha_s^n \mathrm{d}B_s$, and A^n are $\mathbb{R}^m$-valued adapted continuous processes with finite variation (note that m and d are not necessarily equal). Suppose that:

a) $\langle M^n\rangle_t \xrightarrow[n\to+\infty]{\mathbb{P}} \int_0^t \mathcal{K}_s \mathrm{d}s$ for all $t \in [0,T]$ and $(\mathcal{K}_t)_{0\le t\le T}$ is a $\mathcal{S}_m^+$-valued adapted process;

b) $\langle M^n, B\rangle_t \xrightarrow[n\to+\infty]{\mathbb{P}} 0$ for all $t \in [0,T]$;

c) there exists an adapted continuous $\mathbb{R}^m$-valued process A such that $\sup_{0\le t\le T} |A_t^n - A_t| \xrightarrow[n\to+\infty]{\mathbb{P}} 0$.

We denote by $\mathcal{K}_t^{1/2}$ the principal square root of the symmetric non-negative definite matrix $\mathcal{K}_t$. Let W be a m-dimensional Brownian motion independent of $\bar{\mathcal{F}}_T$ defined on an extended probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$. Then, we have the following functional $\bar{\mathcal{F}}$-stable convergence in distribution

$$S_t^n \xrightarrow[[0,T]]{d} A_t + \int_0^t \mathcal{K}_s^{1/2} \mathrm{d}W_s.$$

Proof. First we apply [Fuk11b, Theorem A.1] to the martingale sequence M^n. The conditions of [Fuk11b, Theorem A.1] follow from (a)-(b) and the fact that $M^n = \int_0^\cdot \alpha_s^n \mathrm{d}B_s$ is orthogonal to all martingales that are orthogonal to B. Note that this result in [Fuk11b] can be easily extended to our multidimensional setting using the standard Cramér-Wold argument. Finally the convergence of S^n follows from (c) and the $\bar{\mathcal{F}}$-stability in [Fuk11b, Theorem A.1]. □

We now proceed to the proof of Theorem 3.1. We come back to the setting of Theorem 3.1 with general sequence $\varepsilon_n \to 0$. Take any subsequence $(\varepsilon_{\iota(n)})_{n\geq 0}$. Then there exists another subsequence $(\varepsilon_{\iota'\circ\iota(n)})_{n\geq 0}$ which is square summable and for which the assumptions $(\mathbf{H}_R)$ and $(\mathbf{H}_{\mathcal{B}})$ are verified. To simplify the notation we write simply ε_n instead of $\varepsilon_{\iota'\circ\iota(n)}$ until the final part of the proof.

Recall (see definitions (2.15)) that

$$\mathcal{E}_t^n = \mathcal{E}_t^{n,1} + \mathcal{E}_t^{n,2},$$

with $\mathcal{E}_t^{n,1}$ and $\mathcal{E}_t^{n,2}$ given by

$$\mathcal{E}_t^{n,1} = \int_0^t \mathcal{M}_{\varphi(s)} \Delta S_s \mathrm{d}s, \qquad \mathcal{E}_t^{n,2} = \int_0^t \Delta S_s^\mathsf{T} \mathcal{A}_{\varphi(s)} \mathrm{d}B_s.$$

For two continuous semimartingales $(a_t)_{0\leq t\leq T}$ and $(b_t)_{0\leq t\leq T}$ with values in $\mathbb{R}^m$ and $\mathbb{R}^d$ respectively we denote by $(\langle a; b\rangle_t)_{0\leq t\leq T}$ their $\mathrm{Mat}_{m,d}$-valued quadratic covariation process. Recall that $\mathcal{A}_t = (\mathcal{A}_{1,t}, \ldots, \mathcal{A}_{m,t})^\mathsf{T}$ and set

$$\mathcal{A}_t^{ij} := \frac{1}{2}(\mathcal{A}_{i,t}\mathcal{A}_{j,t}^\mathsf{T} + \mathcal{A}_{i,t}^\mathsf{T}\mathcal{A}_{j,t}).$$

Using Lemma 4.5 we obtain for any $l = 1, \ldots, m$ and $(Q_t)_{0\leq t\leq T}$ given by (4.8)

$$\left\langle \varepsilon_n^{-1} \int_0^\cdot \Delta S_s^\mathsf{T} \mathcal{A}_{l,\varphi(s)} \mathrm{d}B_s; B \right\rangle_t = \varepsilon_n^{-1} \int_0^t \Delta S_s^\mathsf{T} \mathcal{A}_{l,\varphi(s)} \mathrm{d}s \overset{u.c.a.s.}{\underset{n\to+\infty}{\longrightarrow}} \int_0^t Q_s^\mathsf{T} \mathcal{A}_{l,s} \mathrm{d}s. \tag{4.12}$$

Hence

$$\left\langle \varepsilon_n^{-1} \mathcal{E}^{n,2}; B \right\rangle_t = \left\langle \varepsilon_n^{-1} \int_0^\cdot \Delta S_s^\mathsf{T} \mathcal{A}_{\varphi(s)} \mathrm{d}B_s; B \right\rangle_t \overset{u.c.a.s.}{\underset{n\to+\infty}{\longrightarrow}} \int_0^t Q_s^\mathsf{T} \mathcal{A}_s \mathrm{d}s. \tag{4.13}$$

Further we have

$$\left\langle \int_0^t Q_s^\mathsf{T} \mathcal{A}_s \mathrm{d}B_s; B \right\rangle_t = \int_0^t Q_s^\mathsf{T} \mathcal{A}_s \mathrm{d}s, \tag{4.14}$$

which in view of (4.13) yields

$$\left\langle \varepsilon_n^{-1} \mathcal{E}^{n,2} - \int_0^\cdot Q_s^\mathsf{T} \mathcal{A}_s \mathrm{d}B_s; B \right\rangle_t \overset{u.c.a.s.}{\underset{n\to+\infty}{\longrightarrow}} 0. \tag{4.15}$$

We decompose the quadratic covariation matrix of $\varepsilon_n^{-1}\mathcal{E}^{n,2} - \int_0^\cdot Q_s^\mathsf{T} \mathcal{A}_s \mathrm{d}B_s$ at time t as follows: for any $1 \leq i, j \leq m$, we have

$$\begin{aligned}\left\langle \varepsilon_n^{-1} \mathcal{E}^{n,2} - \int_0^\cdot Q_s^\mathsf{T} \mathcal{A}_s \mathrm{d}B_s \right\rangle_t^{ij} &= \varepsilon_n^{-2} \int_0^t \Delta S_s^\mathsf{T} \mathcal{A}_{i,\varphi(s)} \mathcal{A}_{j,\varphi(s)}^\mathsf{T} \Delta S_s \mathrm{d}s + \int_0^t Q_s^\mathsf{T} \mathcal{A}_{i,s} \mathcal{A}_{j,s}^\mathsf{T} Q_s \mathrm{d}s \\ &\quad - \varepsilon_n^{-1} \int_0^t Q_s^\mathsf{T} (\mathcal{A}_{j,s} \mathcal{A}_{i,\varphi(s)}^\mathsf{T} + \mathcal{A}_{i,s} \mathcal{A}_{j,\varphi(s)}^\mathsf{T}) \Delta S_s \mathrm{d}s.\end{aligned}$$

By symmetry of the matrix $\left\langle \varepsilon_n^{-1}\mathcal{E}^{n,2} - \int_0^\cdot Q_s^\mathsf{T}\mathcal{A}_s \mathrm{d}B_s \right\rangle_t$, we deduce

$$\left\langle \varepsilon_n^{-1}\mathcal{E}^{n,2} - \int_0^\cdot Q_s^\mathsf{T}\mathcal{A}_s \mathrm{d}B_s \right\rangle_t^{ij} = \varepsilon_n^{-2}\int_0^t \Delta S_s^\mathsf{T}\mathcal{A}^{ij}_{\varphi(s)}\Delta S_s \mathrm{d}s + \int_0^t Q_s^\mathsf{T}\mathcal{A}_s^{ij}Q_s \mathrm{d}s - \varepsilon_n^{-1}\int_0^t Q_s^\mathsf{T}(\mathcal{A}_{j,s}\mathcal{A}^\mathsf{T}_{i,\varphi(s)} + \mathcal{A}_{i,s}\mathcal{A}^\mathsf{T}_{j,\varphi(s)})\Delta S_s \mathrm{d}s.$$

First, apply the dominated convergence theorem by invoking the a.s. continuity of $\mathcal{A}_\cdot$ and $Q_\cdot$ on $[0,T]$, ($\mathbf{H}_R$) and the convergence to 0 of the mesh size of $\mathcal{T}^n$ (see the proof of Lemma 4.5), it gives

$$\varepsilon_n^{-1}\int_0^t \left| 2Q^\mathsf{T}_{\varphi(s)}\mathcal{A}^{ij}_{\varphi(s)} - Q_s^\mathsf{T}(\mathcal{A}_{j,s}\mathcal{A}^\mathsf{T}_{i,\varphi(s)} + \mathcal{A}_{i,s}\mathcal{A}^\mathsf{T}_{j,\varphi(s)}) \right| \Delta S_s \mathrm{d}s \xrightarrow[n\to+\infty]{u.c.a.s.} 0. \tag{4.16}$$

Second, from Lemma 4.5 we obtain

$$\varepsilon_n^{-1}\int_0^t Q^\mathsf{T}_{\varphi(s)}\mathcal{A}^{ij}_{\varphi(s)}\Delta S_s \mathrm{d}s \xrightarrow[n\to+\infty]{u.c.a.s.} \int_0^t Q_s^\mathsf{T}\mathcal{A}_s^{ij}Q_s \mathrm{d}s. \tag{4.17}$$

Last, we write $\mathcal{A}_s^{ij} = \mathcal{A}_s^{ij+} - \mathcal{A}_s^{ij-}$ (see Lemma B.1), where $\mathcal{A}_s^{ij+}$ and $\mathcal{A}_s^{ij-}$ are adapted continuous symmetric non-negative definite matrices. Owing to Lemma 4.6 we get

$$\varepsilon_n^{-2}\int_0^t \Delta S_s^\mathsf{T}(\mathcal{A}^{ij}_{\varphi(s)})^+\Delta S_s \mathrm{d}s \xrightarrow[n\to+\infty]{u.c.a.s.} \int_0^t m_s^{-1}\mathcal{B}_s[f(x) := ((\sigma_s^{-1}x)^\mathsf{T}X_s^{ij+}(\sigma_s^{-1}x))^2]\mathrm{d}s,$$
$$\varepsilon_n^{-2}\int_0^t \Delta S_s^\mathsf{T}(\mathcal{A}^{ij}_{\varphi(s)})^-\Delta S_s \mathrm{d}s \xrightarrow[n\to+\infty]{u.c.a.s.} \int_0^t m_s^{-1}\mathcal{B}_s[f(x) := ((\sigma_s^{-1}x)^\mathsf{T}X_s^{ij-}(\sigma_s^{-1}x))^2]\mathrm{d}s,$$

where X_s^{ij+} (resp. X_s^{ij-}) is the solution of the matrix equation (2.13) for $c = \sigma_s^\mathsf{T}\mathcal{A}_s^{ij+}\sigma_s$ (resp. $\sigma_s^\mathsf{T}\mathcal{A}_s^{ij-}\sigma_s$). Hence, using that $\mathcal{B}_s[\cdot]$ is linear, we obtain

$$\varepsilon_n^{-2}\int_0^t \Delta S_s^\mathsf{T}\mathcal{A}^{ij}_{\varphi(s)}\Delta S_s \mathrm{d}s \xrightarrow[n\to+\infty]{u.c.a.s.} \int_0^t m_s^{-1}\mathcal{B}_s[f(x) := ((\sigma_s^{-1}x)^\mathsf{T}X_s^{ij+}(\sigma_s^{-1}x))^2 - ((\sigma_s^{-1}x)^\mathsf{T}X_s^{ij-}(\sigma_s^{-1}x))^2]\mathrm{d}s. \tag{4.18}$$

Recall the definition (2.19), i.e.

$$\mathcal{K}_t^{ij} = m_t^{-1}\mathcal{B}_t\left[f(x) := ((\sigma_t^{-1}x)^\mathsf{T}X_t^{ij+}(\sigma_t^{-1}x))^2 - ((\sigma_t^{-1}x)^\mathsf{T}X_t^{ij-}(\sigma_t^{-1}x))^2)\right] - Q_t^\mathsf{T}\mathcal{A}_t^{ij}Q_t.$$

Thus from (4.16), (4.17) and (4.18) we get the convergence

$$\left\langle \varepsilon_n^{-1}\mathcal{E}^{n,2} - \int_0^\cdot Q_s^\mathsf{T}\mathcal{A}_s \mathrm{d}B_s \right\rangle_t \xrightarrow[n\to+\infty]{u.c.a.s.} \int_0^t \mathcal{K}_s \mathrm{d}s. \tag{4.19}$$

Note that $\mathcal{K}_s$ is a symmetric non-negative definite matrix since it is the a.s. limit of covariation matrices.

Further we compute the limit for the finite variation part $\mathcal{E}_t^{n,1}$. Owing to Lemma 4.5 we directly have

$$\varepsilon_n^{-1}\mathcal{E}_t^{n,1} = \varepsilon_n^{-1}\int_0^t \mathcal{M}_{\varphi(s)}\Delta S_s \mathrm{d}s \xrightarrow[n\to+\infty]{u.c.a.s.} \int_0^t \mathcal{M}_s Q_s \mathrm{d}s. \tag{4.20}$$

For the convergence of $\varepsilon_n^2 N_t^n$ we take advantage of Theorem 4.4 to write

$$\varepsilon_n^2 N_t^n = \sum_{\tau_{i-1}^n < t} \varepsilon_n^2 \overset{u.c.a.s.}{\underset{n\to+\infty}{\longrightarrow}} \int_0^t m_s^{-1} \mathrm{d}s. \tag{4.21}$$

Now we come back to the initial notation $\varepsilon_{\iota'\circ\iota}$ for the subsequence. Having proved the a.s. convergences (4.15), (4.19), (4.20) and (4.21) for $\varepsilon_n = \varepsilon_{\iota'\circ\iota}$, we use the arbitrary choice of $\iota(n)$ and the subsequence principle from Lemma 4.1 to get the same convergences in probability along the initial sequence $(\varepsilon_n : n \geq 0)$. So, in particular, we can apply Theorem 4.7 with

$$M_t^n = \varepsilon_n^{-1}\mathcal{E}_t^{n,2} - \int_0^t Q_s^\mathsf{T} \mathcal{A}_s \mathrm{d}B_s \quad \text{and} \quad A_t^n = \varepsilon_n^{-1}\mathcal{E}_t^{n,1}$$

and after easy manipulations, we obtain the following functional $\bar{\mathcal{F}}$-stable convergence in distribution:

$$\varepsilon_n^{-1}\mathcal{E}_t^n \underset{[0,T]}{\overset{d}{\Longrightarrow}} \left(\int_0^t \mathcal{M}_s Q_s \mathrm{d}s + \int_0^t Q_s^\mathsf{T} \mathcal{A}_s \mathrm{d}B_s + \int_0^t \mathcal{K}_s^{1/2} \mathrm{d}W_s \right).$$

The uniform convergence in probability (3.5) follows similarly from the $\mathbb{P}$-version of the convergence (4.21). The proof of Theorem 3.1 is now complete. □

5 Proofs of domain exit time properties (Lemma 3.3, Propositions 3.4 and 3.5)

We assume the notation of Section 3.2. In particular L_D denotes the constant given by (3.7).

5.1 Proof of Lemma 3.3

We begin by justifying *i)* with $p = 1$. For this we assume without loss of generality that the process S has its coefficients such that

$$\Lambda_{\min}^\sigma \leq \inf_{t\geq 0} \lambda_{\min}(\sigma_t \sigma_t^\mathsf{T}) \leq \sup_{t\geq 0} \|\sigma_t \sigma_t^\mathsf{T}\| \leq \Lambda_{\max}^\sigma, \qquad \sup_{t\geq 0} |b_t| \leq b_{\max}. \tag{5.1}$$

Indeed, we can still define new $\bar{\mathcal{F}}$-adapted coefficients $\tilde{b}_t = b_t \mathbf{1}_{\tau<t}$ and $\tilde{\sigma}_t = \sigma_t \mathbf{1}_{\tau<t} + \mathbf{1}_{\tau\geq t}\sqrt{\Lambda_{\max}^\sigma / d}$: they satisfy to the above bounds, they coincide with those of S before τ, and therefore the process with new coefficients has the same exit time τ. For the proof of the above lemma, this is enough to consider such a modified process instead of the initial S, or equivalently to assume (5.1) for S.

Now, we invoke the rough bound $\tau \leq \bar{\tau} = \inf\{t \geq 0 : |S_{1,t}| \geq \varepsilon L_D^{-1}\}$ which holds since D is included in a ball centered at 0 with radius L_D^{-1}. We now derive two bounds, one for any $\varepsilon \leq \varepsilon_0 \leq 1$, the other for small ε.

1. Take λ as the unique positive solution to $-\lambda b_{\max} + \frac{1}{2}\lambda^2 \Lambda_{\min}^\sigma = 1$: clearly $\lambda \in \mathcal{C}(\mathfrak{S})$; then apply the Itô formula in expectation to get

$$\begin{aligned} e^{\lambda L_D^{-1}} \geq \mathbb{E}_\nu\left(e^{\lambda S_{1,\bar{\tau}}}\right) &= e^{\lambda S_{1,\nu}} + \mathbb{E}_\nu\left(\int_\nu^{\bar{\tau}} e^{\lambda S_{1,s}} \big(\lambda b_s^1 + \frac{1}{2}\lambda^2 |\sigma_{1:,s}|^2\big) \mathrm{d}s\right) \\ &\geq \mathbb{E}_\nu\left(\int_\nu^{\bar{\tau}} e^{\lambda S_{1,s}} \mathrm{d}s\right) \geq e^{-\lambda L_D^{-1}} \mathbb{E}_\nu(\bar{\tau} - \nu). \end{aligned}$$

   This holds for any $\varepsilon \leq 1$.

2. Now, for $\varepsilon \le \min(1, \Lambda^\sigma_{\min} L_D/(4b_{\max})) := \bar\varepsilon_0 \in \mathcal{C}(\mathfrak{S})$ so that $-2\varepsilon L_D^{-1} b_{\max} + \Lambda^\sigma_{\min} \ge \Lambda^\sigma_{\min}/2$, we have with similar arguments

$$\varepsilon^2 L_D^{-2} \ge \mathbb{E}_\nu\left(S^2_{1,\bar\tau}\right) = S^2_{1,\nu} + \mathbb{E}_\nu\left(\int_\nu^{\bar\tau} \left(2S_{1,s}b_{1,s} + |\sigma_{1:,s}|^2\right)\mathrm{d}s\right) \ge \mathbb{E}_\nu(\bar\tau - \nu)\ \Lambda^\sigma_{\min}/2.$$

To summarize, we have justified that for any stopping time ν, a.s. on $\{\nu \le \tau\}$ we have

$$\mathbb{E}_\nu\left(\tau - \nu\right) \le \mathbb{E}_\nu\left(\bar\tau - \nu\right) \le e^{2\lambda/L_D}\mathbf{1}_{\varepsilon > \bar\varepsilon_0} + 2\varepsilon^2/(L_D^2\Lambda^\sigma_{\min})\mathbf{1}_{\varepsilon\le\bar\varepsilon_0} \le \max\left(e^{2\lambda/L_D}/\bar\varepsilon_0^2, 2/(L_D^2\Lambda^\sigma_{\min})\right)\varepsilon^2 =: R_D\varepsilon^2$$

with $R_D \in \mathcal{C}(\mathfrak{S})$.

We now establish *i)* for $p \ge 2$ by induction. Assume that *i)* holds for some $p \ge 1$ and for any stopping time ν: then, on $\{\nu \le \tau\}$,

$$\begin{aligned}
\mathbb{E}_\nu\Big((\tau-\nu)^{p+1}\Big) &= \int_0^\infty (p+1)\mathbb{E}_\nu\Big((\tau-\nu-t)^p\mathbf{1}_{\tau-\nu\ge t}\Big)\mathrm{d}t \\
&= \int_0^\infty (p+1)\mathbb{E}_\nu\Big(\mathbb{E}_{\nu+t}\big((\tau-\nu-t)^p\big)\mathbf{1}_{\tau-\nu\ge t}\Big)\mathrm{d}t \\
&\le \int_0^\infty (p+1)\mathbb{E}_\nu\Big(p!(R_D\varepsilon^2)^p\mathbf{1}_{\tau-\nu\ge t}\Big)\mathrm{d}t \\
&= (p+1)!(R_D\varepsilon^2)^p\mathbb{E}_\nu(\tau-\nu) \le (p+1)!(R_D\varepsilon^2)^{p+1}
\end{aligned}$$

using twice the induction assumption (first for the stopping time $\nu+t$ on the event $\{\nu+t\le\tau\}$, second for ν on the event $\{\nu\le\tau\}$).

Last we derive *ii)*. On $\{\nu \le \tau\}$, use the exponential Markov inequality and the estimates *i)* to get

$$\mathbb{P}_\nu(\tau-\nu \ge \varepsilon^2 c) \le \mathbb{E}_\nu\left(e^{\frac{1}{2R_D\varepsilon^2}(\tau-\nu-\varepsilon^2 c)}\right) \le e^{-\frac{c}{2R_D}}\sum_{p\ge0}\frac{1}{p!2^p}\mathbb{E}_\nu\left(\left(\frac{\tau-\nu}{R_D\varepsilon^2}\right)^p\right) \le 2e^{-\frac{c}{2R_D}}.$$

We are done. □

5.2 Preparing the proof of Propositions 3.4 and 3.5

This section is devoted to some preliminary results. Only within this section we assume that

$$\boxed{D \in \mathcal{D}}$$

(we pass to the general case $D \in \mathcal{D}^J_\cap$ in Section 5.3). For simplicity we write $\delta(\cdot)$ instead of $\delta_{\partial D}(\cdot)$ since $D \in \mathcal{D}$ is fixed and no confusion is possible. For $\varepsilon > 0$ denote $\delta_\varepsilon(x) := \varepsilon\delta(\varepsilon^{-1}x)$.

Lemma 5.1. *Assume* $(\mathbf{H}^{D,\sigma}_{loc})$ *with* $D \in \mathcal{D}$ *(*$J = 1$*). Let* $\tau_0 := \inf\{t \ge 0 : S_t \notin D)\}$. *There exists* $L_\sigma \in \mathcal{C}(\mathfrak{S})$ *such that* $L_\sigma \le L_D$ *and for any* $t \in [0, \tau_0]$ *we have a.s.*

$$\inf_{0\le\delta(x)\le L_\sigma} \mathrm{Tr}(\sigma_t^\mathsf{T}(\nabla\delta^\mathsf{T}\nabla\delta + \delta\nabla^2\delta)(x)\sigma_t) \ge \frac{1}{8}\Lambda^\sigma_{\min}.$$

Proof. Remind of the convention on $\nabla\delta$ as a row vector. By (3.7) on the set $|\delta(x)| \le L_D$ the function $\delta(\cdot)$ is $\mathcal{C}^2$ and $\inf_{0\le\delta(x)\le L_D} |\nabla\delta(x)|^2 \ge \frac{1}{4}$. For any $x \in D$ such that $|\delta(x)| \le L_D$ we have

$$\mathrm{Tr}(\sigma_t^\mathsf{T}(\nabla\delta^\mathsf{T}\nabla\delta)(x)\sigma_t) = \nabla\delta(x)^\mathsf{T} \cdot \sigma_t\sigma_t^\mathsf{T}\nabla\delta(x)^\mathsf{T} \ge \frac{1}{4}\Lambda^\sigma_{\min}. \tag{5.2}$$

Further using $|\mathrm{Tr}(M)|\sqrt{d}\|M\|$ for $M \in \mathrm{Mat}_{d,d}(\mathbb{R})$ and the sub-multiplicative property of the Frobenius norm, for any $0 \le L \le L_D$ and $x \in D$ with $|\delta(x)| \le L$, we have

$$|\mathrm{Tr}(\sigma_t^\mathsf{T}(\delta\nabla^2\delta)(x)\sigma_t)| \le \sqrt{d}\|\sigma_t\sigma_t^\mathsf{T}(\delta\nabla^2\delta)(x)\| \le \sqrt{d}\|\sigma_t\sigma_t^\mathsf{T}\| \times \|(\delta\nabla^2\delta)(x)\| \le \sqrt{d}LL_D^{-1}\Lambda^\sigma_{\max}. \tag{5.3}$$

We set $L_\sigma := L_D \min\left(1, \frac{\Lambda^\sigma_{\min}}{8\sqrt{d}\Lambda^\sigma_{\max}}\right)$, which is a continuous function of L_D, $\Lambda^\sigma_{\min}$ and $\Lambda^\sigma_{\max}$, so that $\frac{1}{4}\Lambda^\sigma_{\min} - \sqrt{d}L_\sigma L_D^{-1}\Lambda^\sigma_{\max} \ge \frac{1}{8}\Lambda^\sigma_{\min}$, which together with (5.2) and (5.3) implies the announced result. □

Lemma 5.2. *Assume* $(\mathbf{H}^{D,\sigma}_{loc})$ *with* $D \in \mathcal{D}$ *($J = 1$). There exists* $K \in \mathcal{C}(\mathfrak{S})$ *such that for any* $\varepsilon \in (0, 1]$ *and the stopping time*

$$\tau = \inf\{t \ge 0 : S_t \notin \varepsilon D\}$$

and any stopping time ν *such that* $\nu \le \tau$ *a.s. we have*

$$\mathbb{E}(\tau - \nu) \le K\varepsilon^2\mathbb{E}(\delta(\varepsilon^{-1}S_\nu)). \tag{5.4}$$

Proof. Take $\varepsilon \in (0, 1]$. Let $L_\sigma \in (0, L_D]$ be given by Lemma 5.1 (L_D is defined in (3.7)), $l \in (0, L_\sigma]$. We have

$$\mathbb{E}(\tau - \nu) = \mathbb{E}((\tau - \nu)\mathbb{1}_{\delta(\varepsilon^{-1}S_\nu)>l}) + \mathbb{E}((\tau - \nu)\mathbb{1}_{\delta(\varepsilon^{-1}S_\nu)\le l}). \tag{5.5}$$

Using Lemma 3.3 we get

$$\mathbb{E}((\tau - \nu)\mathbb{1}_{\delta(\varepsilon^{-1}S_\nu)>l}) = \mathbb{E}(\mathbb{1}_{\delta(\varepsilon^{-1}S_\nu)>l}\mathbb{E}_\nu(\tau - \nu)) \le R_D\varepsilon^2\mathbb{P}(\delta(\varepsilon^{-1}S_\nu) > l). \tag{5.6}$$

The rest of the proof consists in estimating $\mathbb{1}_{\delta(\varepsilon^{-1}S_\nu)\le l}\mathbb{E}_\nu(\tau - \nu)$. For simplicity we omit the indicator in the calculations, so that *we are working on the event* $\{\delta(\varepsilon^{-1}S_\nu) \le l\}$. Denote $\tau_l := \inf\{t > \nu : \delta_\varepsilon(S_t) \ge l\varepsilon\}$. Note that $\delta(\cdot)$ is $\mathcal{C}^2$ on the set $|\delta(x)| \le l$ since $l \le L_\sigma \le L_D$. Let us write the Itô formula for $\delta^2_\varepsilon(S_t)$ on $[\nu, \tau \wedge \tau_l]$:

$$\begin{aligned}\delta^2_\varepsilon(S_{\tau\wedge\tau_l}) = \delta^2_\varepsilon(S_\nu) &+ 2\int_\nu^{\tau\wedge\tau_l} (\delta_\varepsilon\nabla\delta_\varepsilon)(S_s)\,\mathrm{d}S_s \\ &+ \int_\nu^{\tau\wedge\tau_l} \mathrm{Tr}(\sigma_s^\mathsf{T}(\nabla\delta_\varepsilon^\mathsf{T}\nabla\delta_\varepsilon + \delta_\varepsilon\nabla^2\delta_\varepsilon)(S_s)\sigma_s)\mathrm{d}s.\end{aligned} \tag{5.7}$$

Note that by Lemma 5.1, $s \le \tau \le \tau_0$, $0 \le \delta(\varepsilon^{-1}S_s) \le l \le L_\sigma$, $(\nabla\delta_\varepsilon^\mathsf{T}\nabla\delta_\varepsilon + \delta_\varepsilon\nabla^2\delta_\varepsilon)(x) = (\nabla\delta^\mathsf{T}\nabla\delta + \delta\nabla^2\delta)(\varepsilon^{-1}x)$ we have for all $s \in [\nu, \tau \wedge \tau_l]$ a.s.

$$\mathrm{Tr}(\sigma_s^\mathsf{T}(\nabla\delta_\varepsilon^\mathsf{T}\nabla\delta_\varepsilon + \delta_\varepsilon\nabla^2\delta_\varepsilon)(S_s)\sigma_s) \ge \frac{1}{8}\Lambda^\sigma_{\min}.$$

So we obtain

$$\int_\nu^{\tau\wedge\tau_l} \mathrm{Tr}(\sigma_s^\mathsf{T}(\nabla\delta_\varepsilon^\mathsf{T}\nabla\delta_\varepsilon + \delta_\varepsilon\nabla^2\delta_\varepsilon)(S_s)\sigma_s)\mathrm{d}s \ge \frac{1}{8}\Lambda^\sigma_{\min}(\tau \wedge \tau_l - \nu). \tag{5.8}$$

Further

$$\left|\mathbb{E}_\nu\left(\int_\nu^{\tau\wedge\tau_l}(\delta_\varepsilon\nabla\delta_\varepsilon)(S_s)\ \mathrm{d}S_s\right)\right| = \left|\mathbb{E}_\nu\left(\int_\nu^{\tau\wedge\tau_l}(\delta_\varepsilon\nabla\delta_\varepsilon)(S_s)\ b_s\mathrm{d}s\right)\right| \leq lL_D^{-1}b_{\max}\mathbb{E}_\nu(\tau\wedge\tau_l-\nu). \tag{5.9}$$

Thus from (5.7), applying $\mathbb{E}_\nu(\cdot)$, using (5.8), (5.9) and simply that $\delta_\varepsilon^2(S_\nu)\geq 0$ we get

$$C_1\mathbb{E}_\nu(\tau\wedge\tau_l-\nu)\leq \mathbb{E}_\nu(\delta_\varepsilon^2(S_{\tau\wedge\tau_l})),$$

where $C_1=\frac{1}{16}\Lambda^\sigma_{\min}\in\mathcal{C}(\mathfrak{S})$, for any l satisfying

$$0<l\leq L_\sigma\wedge(L_D\Lambda^\sigma_{\min}b^{-1}_{\max}/16). \tag{5.10}$$

We continue with l satisfying (5.10). Now using that $\delta_\varepsilon(S_\tau)=0$ and from the definition of τ_l we get $\mathbb{E}_\nu(\delta_\varepsilon^2(S_{\tau\wedge\tau_l}))=\mathbb{E}_\nu(\delta_\varepsilon^2(S_{\tau_l})\mathbb{1}_{\tau>\tau_l})=l^2\varepsilon^2\mathbb{P}_\nu(\tau>\tau_l)$, and consequently

$$C_1\mathbb{E}_\nu(\tau\wedge\tau_l-\nu)\leq l^2\varepsilon^2\mathbb{P}_\nu(\tau>\tau_l). \tag{5.11}$$

Further we write

$$\mathbb{E}_\nu(\tau\wedge\tau_l-\nu)=\mathbb{E}_\nu((\tau-\nu)\mathbb{1}_{\tau<\tau_l})+\mathbb{E}_\nu((\tau_l-\nu)\mathbb{1}_{\tau>\tau_l})=\mathbb{E}_\nu(\tau-\nu)-\mathbb{E}_\nu((\tau-\tau_l)\mathbb{1}_{\tau>\tau_l}). \tag{5.12}$$

Using Lemma 3.3 (with $R_D\in\mathcal{C}(\mathfrak{S})$) we obtain

$$\mathbb{E}_\nu((\tau-\tau_l)\mathbb{1}_{\tau>\tau_l})=\mathbb{E}_\nu(\mathbb{1}_{\tau>\tau_l}\mathbb{E}_{\tau_l}(\tau-\tau_l))\leq R_D\varepsilon^2\mathbb{P}_\nu(\tau>\tau_l). \tag{5.13}$$

Hence plugging (5.11) and (5.13) into (5.12) yields

$$\mathbb{E}_\nu(\tau-\nu)\leq(R_D+C_1^{-1}l^2)\varepsilon^2\mathbb{P}_\nu(\tau>\tau_l). \tag{5.14}$$

Now, we aim at upper bounding the above probability. By taking the conditional expectation $\mathbb{E}_\nu(\cdot)$ of the Itô formula for $\delta_\varepsilon(S_t)$ on $[\nu,\tau\wedge\tau_l]$, we get

$$l\varepsilon\mathbb{P}_\nu(\tau>\tau_l)=\mathbb{E}_\nu(\delta_\varepsilon(S_{\tau\wedge\tau_l}))=\delta_\varepsilon(S_\nu)+\mathbb{E}_\nu\left(\int_\nu^{\tau\wedge\tau_l}\nabla\delta_\varepsilon(S_s)\ b_s\mathrm{d}s\right) + \frac{1}{2}\mathbb{E}_\nu\left(\int_\nu^{\tau\wedge\tau_l}\mathrm{Tr}(\sigma_s^\intercal\nabla^2\delta_\varepsilon(S_s)\sigma_s)\mathrm{d}s\right). \tag{5.15}$$

The first expectation in the right-hand side of (5.15) is bounded by $L_D^{-1}b_{\max}\mathbb{E}_\nu(\tau\wedge\tau_l-\nu)$, while the second expectation, in view of (3.7) and ($\mathbf{H}^{D,\sigma}_{loc}$), is bounded by $\varepsilon^{-1}\sqrt{d}L_D^{-1}\Lambda^\sigma_{\max}\mathbb{E}_\nu(\tau\wedge\tau_l-\nu)$. Therefore, plugging the above into (5.15) and using then (5.11), we readily obtain

$$\begin{aligned}l\varepsilon^2\mathbb{P}_\nu(\tau>\tau_l)&\leq\varepsilon\delta_\varepsilon(S_\nu)+(\frac{1}{2}\sqrt{d}L_D^{-1}\Lambda^\sigma_{\max}+L_D^{-1}b_{\max})\mathbb{E}_\nu(\tau\wedge\tau_l-\nu)\\&\leq\varepsilon\delta_\varepsilon(S_\nu)+\varepsilon^2C_2l^2\mathbb{P}_\nu(\tau>\tau_l),\end{aligned}$$

where $C_2:=(\frac{1}{2}\sqrt{d}L_D^{-1}\Lambda^\sigma_{\max}+L_D^{-1}b_{\max})C_1^{-1}$, so that $C_2\in\mathcal{C}(\mathfrak{S})$. Note that all the previous analysis is valid for any l verifying (5.10) and the elements of $\mathcal{C}(\mathfrak{S})$ do not depend on l, so we may now fix $l=l_0:=\min(C_2^{-1}/2,L_\sigma,(L_D\Lambda^\sigma_{\min}b^{-1}_{\max})/16)$ which implies $C_3:=l_0-C_2l_0^2\geq\frac{l_0}{2}>0$. Observe that $l_0,C_3\in\mathcal{C}(\mathfrak{S})$. Hence we obtain

$$\mathbb{P}_\nu(\tau>\tau_{l_0})\leq C_3^{-1}\delta(\varepsilon^{-1}S_\nu). \tag{5.16}$$

Combining (5.14) and (5.16) and setting $K := (R_D + C_1^{-1} l_0^2) C_3^{-1} \in \mathcal{C}(\mathfrak{S})$, we get

$$\mathbb{E}_\nu(\tau - \nu) \leq K\varepsilon^2 \delta(\varepsilon^{-1} S_\nu). \tag{5.17}$$

Remember that this result is obtained on the event $\{\delta(\varepsilon^{-1}S_\nu) \leq l_0\}$. Going back to the general notation we have $\mathbb{1}_{\delta(\varepsilon^{-1}S_\nu)\leq l_0}\mathbb{E}_\nu(\tau-\nu) \leq K\varepsilon^2 \mathbb{1}_{\delta(\varepsilon^{-1}S_\nu)\leq l_0}\delta(\varepsilon^{-1}S_\nu)$, and then by taking expectation and combining this with (5.6) and (5.5), we finally obtain

$$\mathbb{E}(\tau - \nu) \leq K\varepsilon^2 \mathbb{E}(\delta(\varepsilon^{-1}S_\nu)) + R_D \varepsilon^2 \mathbb{P}(\delta(\varepsilon^{-1}S_\nu) > l_0) \leq (K + R_D l_0^{-1})\varepsilon^2 \mathbb{E}(\delta(\varepsilon^{-1}S_\nu))$$

where we have applied the Markov inequality at the last inequality. We are done. □

Lemma 5.3. *Assume* $(\mathbf{H}^{D,\sigma}_{loc})$ *with* $D \in \mathcal{D}$ $(J = 1)$*, and let* $f \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ *be an* α*-homogeneous function with* $\alpha \in \{2, 3, 4\}$*. There exists* $K \in \mathcal{C}(\mathfrak{S})$ *such that for any* $\varepsilon \in (0, 1]$*, for the stopping times*

$$\tau = \inf\{t \geq 0 : S_t \notin \varepsilon D\}, \quad \bar{\tau} = \inf\{t \geq 0 : \bar{S}_t \notin \varepsilon D\}$$

and any stopping time ν *such that* $\nu \leq \tau \wedge \bar{\tau}$ a.s.*, we have*

$$\varepsilon^{-2}\mathbb{E}(|S_\nu - \bar{S}_\nu|^2) + \varepsilon^{-\alpha}|\mathbb{E}(f(S_\nu) - f(\bar{S}_\nu))| \leq K\varepsilon^{\eta_\sigma}. \tag{5.18}$$

Proof. We start with a bound on $\mathbb{E}(|S_\nu - \bar{S}_\nu|^2)$:

$$\begin{aligned}\mathbb{E}\left(|S_\nu - \bar{S}_\nu|^2\right) &\leq \mathbb{E}\left(\sum_{k=1}^{+\infty} \mathbb{1}_{\nu/\varepsilon^2 \in [k-1,k)} \sup_{t\leq k\varepsilon^2} |\bar{S}_t - S_t|^2\right) \\ &\leq \sum_{k=1}^{+\infty} \mathbb{P}(\nu/\varepsilon^2 \in [k-1,k))^{1/2} \left[\mathbb{E}\left(\sup_{t\leq k\varepsilon^2} |\bar{S}_t - S_t|^4\right)\right]^{1/2}.\end{aligned}$$

a) Estimate for $\mathbb{E}\left(\sup_{t\leq k\varepsilon^2} |\bar{S}_t - S_t|^4\right)$*:* Denote

$$\bar{M}_t := \int_0^t (\sigma_s - \sigma_0) \mathrm{d}W_s$$

so that $S_t - \bar{S}_t = \int_0^t b_s \mathrm{d}s + \bar{M}_t$. Using the BDG inequalities and $(\mathbf{H}^{D,\sigma}_{loc})$ we obtain

$$\begin{aligned}\mathbb{E}\left(\sup_{t\leq k\varepsilon^2} |S_t - \bar{S}_t|^4\right) &\leq 8\left(b_{\max}^4 (k\varepsilon^2)^4 + \mathbb{E}\left(\sup_{t\leq k\varepsilon^2} |\bar{M}_t|^4\right)\right) \\ &\leq C\left(b_{\max}^4 (k\varepsilon^2)^4 + \mathbb{E}\left(\langle \bar{M}\rangle_{k\varepsilon^2}^2\right)\right),\end{aligned}$$

where C is some universal constant. For the quadratic variation part we get

$$\mathbb{E}\left(\langle \bar{M}\rangle_{k\varepsilon^2}^2\right) = \mathbb{E}\left(\left(\int_0^{k\varepsilon^2} |\sigma_t - \sigma_0|^2 \mathrm{d}t\right)^2\right) \leq \mathbb{E}\left(C_\sigma^4\right)\left(\int_0^{k\varepsilon^2} t^{\eta_\sigma} \mathrm{d}t\right)^2 = C_0 (k\varepsilon^2)^{2(\eta_\sigma+1)},$$

with $C_0 := \frac{m_\sigma}{(\eta_\sigma+1)^2}$. So we conclude, using that $k \geq 1, \varepsilon \leq 1$,

$$\mathbb{E}\left(\sup_{t\leq k\varepsilon^2} |S_t - \bar{S}_t|^4\right) \leq C_1 k^4 \varepsilon^{2(2+2\eta_\sigma)}, \tag{5.19}$$

where $C_1 := C(b_{\max}^4 + C_0) \in \mathcal{C}(\mathfrak{S})$.

b) Estimate for $\mathbb{P}(\nu/\varepsilon^2 \in [k-1,k))^{1/2}$: Lemma 3.3-ii) directly yields

$$\mathbb{P}(\nu/\varepsilon^2 \in [k-1,k)) \le \mathbb{P}(\nu \ge \varepsilon^2(k-1)) \le R_D e^{-R_D(k-1)}$$

for some $R_D \in \mathcal{C}(\mathfrak{S})$ Hence combining this with (5.19) we get

$$\sum_{k=1}^{+\infty} \mathbb{P}(\nu/\varepsilon^2 \in [k-1,k))^{1/2} \left[\mathbb{E}\left(\sup_{t\le k\varepsilon^2} |S_t - \bar{S}_t|^4\right)\right]^{1/2} \le \sqrt{R_D C_1}\left(\sum_{k=1}^{+\infty} e^{-R_D(k-1)/2} k^2\right)\varepsilon^{2+2\eta_\sigma}.$$

Thus for $K = \sqrt{R_D C_1}\left(\sum_{k=1}^{+\infty} e^{-R_D(k-1)/2}k^2\right)$ (so that $K \in \mathcal{C}(\mathfrak{S})$), we get

$$\mathbb{E}_\nu(|S_\nu - \bar{S}_\nu|^2) \le K\varepsilon^{2+2\eta_\sigma}. \tag{5.20}$$

Now we proceed with the proof of (5.18) regarding f. Recall that the function f verifies (3.9). We have

$$\begin{aligned}|\mathbb{E}\left(f(S_\nu) - f(\bar{S}_\nu)\right)| &\le \mathbb{E}\left(|S_\nu - \bar{S}_\nu| \int_0^1 |\nabla f(\lambda S_\nu + (1-\lambda)\bar{S}_\nu)| \mathrm{d}\lambda\right) \\ &\le \left[\mathbb{E}\left(\left(\int_0^1 |\nabla f(\lambda S_\nu + (1-\lambda)\bar{S}_\nu)|\mathrm{d}\lambda\right)^2\right)\right]^{1/2} \left[\mathbb{E}\left(|S_\nu - \bar{S}_\nu|^2\right)\right]^{1/2}.\end{aligned}$$

Using that $\nu \le \tau \wedge \bar{\tau}$ we obtain $|S_\nu| \le \varepsilon L_D^{-1}$ and $|\bar{S}_\nu| \le \varepsilon L_D^{-1}$ so that

$$\left[\mathbb{E}\left(\left(\int_0^1 |\nabla f(\lambda S_\nu + (1-\lambda)\bar{S}_\nu)|\mathrm{d}\lambda\right)^2\right)\right]^{1/2} \le C_f\ L_D^{-(\alpha-1)}\varepsilon^{\alpha-1}. \tag{5.21}$$

Now combine (5.20) and (5.21) to get (up to changing $K \in \mathcal{C}(\mathfrak{S})$) the announced estimate. □

Corollary 5.4. *Assume $(\mathbf{H}_{loc}^{D,\sigma})$ with $D \in \mathcal{D}$ $(J=1)$. There exists $K \in \mathcal{C}(\mathfrak{S})$ such that for any $\varepsilon \in (0,1]$, the stopping times*

$$\tau = \inf\{t \ge 0 : S_t \notin \varepsilon D\}, \quad \bar{\tau} = \inf\{t \ge 0 : \bar{S}_t \notin \varepsilon D\}$$

satisfy

$$\mathbb{E}(|\tau - \bar{\tau}|) \le K\varepsilon^{2+\eta_\sigma}, \tag{5.22}$$

Proof. Let $\nu := \tau \wedge \bar{\tau}$. Applying Lemma 5.2, we get for some $K \in \mathcal{C}(\mathfrak{S})$

$$\mathbb{E}\left(\tau - \nu\right) \le K\varepsilon^2 \mathbb{E}\left(\delta(\varepsilon^{-1}S_\nu)\right). \tag{5.23}$$

Using that $\mathbb{1}_{\nu<\tau}\delta(\varepsilon^{-1}\bar{S}_\nu) = 0$ and $\mathbb{1}_{\nu=\tau}\delta(\varepsilon^{-1}S_\nu) = 0$ we write

$$\mathbb{E}\left(\delta(\varepsilon^{-1}S_\nu)\right) = \mathbb{E}\left(\mathbb{1}_{\nu<\tau}(\delta(\varepsilon^{-1}S_\nu) - \delta(\varepsilon^{-1}\bar{S}_\nu))\right) \le L_D^{-1}\varepsilon^{-1}\mathbb{E}(|S_\nu - \bar{S}_\nu|^2)^{1/2}.$$

Using (5.18) from Lemma 5.3 we get $\varepsilon^2\mathbb{E}(\delta(\varepsilon^{-1}S_\nu)) \le L_D^{-1}K^{1/2}\varepsilon^{2+\eta_\sigma/2}$. In view of (5.23), we have proved (up to redefining $K \in \mathcal{C}(\mathfrak{S})$)

$$\mathbb{E}(\mathbb{1}_{\tau>\bar{\tau}}(\tau - \bar{\tau})) = \mathbb{E}(\tau - \nu) \le K\varepsilon^{2+\eta_\sigma}.$$

A similar bound holds for $\mathbb{E}(\mathbb{1}_{\tau<\bar{\tau}}(\bar{\tau} - \tau))$: this is justified in the same way, applying Lemma 5.2 to $\bar{S}$ and Lemma 5.3. Consequently, the proof of the bound for $\mathbb{E}(|\tau - \bar{\tau}|)$ is complete. □

Corollary 5.5. *Assume* ($\mathbf{H}^{D,\sigma}_{loc}$) *with* $D \in \mathcal{D}$ *(*$J = 1$*). There exists* $K \in \mathcal{C}(\mathfrak{S} \cup \{K'\})$ *such that for any* $\varepsilon \in (0,1]$*, any strictly positive constants* K', η' *and for* $D' \in \mathcal{D}$ *such that* $\mu(D, D') \leq K'\varepsilon^{\eta'}$*, and for which* (3.7) *and* ($\mathbf{H}^{D,\sigma}_{loc}$) *hold for* D' *instead of* D *with the same constants* $L_D, \Lambda^\sigma_{\min}, \Lambda^\sigma_{\max}, b_{\max}$*, we have*

$$\mathbb{E}(|\tau - \tau'|) \leq K\varepsilon^{2+\eta'} \tag{5.24}$$

where

$$\tau = \inf\{t \geq 0 : S_t \notin \varepsilon D\}, \quad \tau' = \inf\{t \geq 0 : S_t \notin \varepsilon D'\}.$$

In particular, K *is a multiple of* K'*, so that* $K \to 0$ *as* $K' \to 0$*.*

Proof. Let $\nu := \tau \wedge \tau'$ and denote by $\delta(\cdot)$ the distance $\delta_{\partial D}(\cdot)$. Using Lemma 5.2, we obtain for some $K \in \mathcal{C}(\mathfrak{S})$

$$\mathbb{E}(\tau - \nu) \leq K\varepsilon^2 \mathbb{E}(\delta(\varepsilon^{-1} S_\nu)).$$

Observe that $\delta(\varepsilon^{-1} S_\nu) \leq \mu(D, D') \leq K'\varepsilon^{\eta'}$, which gives

$$\mathbb{E}(\mathbb{1}_{\tau \geq \tau'}(\tau - \tau')) = \mathbb{E}(\tau - \nu) \leq KK'\varepsilon^{2+\eta'}.$$

A similar bound on $\mathbb{E}(\mathbb{1}_{\tau' \geq \tau}(\tau' - \tau))$ follows from the symmetry between D and D'. □

5.3 Proofs of Propositions 3.4 and 3.5

Now we pass to the general case of $D \in \mathcal{D}^J_\cap$, i.e. of the form $D = \cap_{j=1}^J D_j$. Note that the results of Section 5.2 are valid for each $D_j, j = 1, \ldots, J$.

Proof of Proposition 3.4. Let $\nu := \tau \wedge \bar{\tau}$. Denote for $j = 1, \ldots, J$

$$\tau_j = \inf\{t \geq 0 : S_t \notin \varepsilon D_j\}, \quad \bar{\tau}_j = \inf\{t \geq 0 : \bar{S}_t \notin \varepsilon D_j\},$$

so that $\tau = \min(\tau_1, \ldots, \tau_J)$ and $\bar{\tau} = \min(\bar{\tau}_1, \ldots, \bar{\tau}_J)$. Write

$$\begin{aligned} |\mathbb{E}(f(S_{\tau \wedge T}) - f(\bar{S}_{\bar{\tau} \wedge T}))| \leq\ & |\mathbb{E}(f(S_{\nu \wedge T}) - f(\bar{S}_{\nu \wedge T}))| + |\mathbb{E}(f(S_{\tau \wedge T}) - f(S_{\nu \wedge T}))| \\ & + |\mathbb{E}(f(\bar{S}_{\bar{\tau} \wedge T}) - f(\bar{S}_{\nu \wedge T}))|. \end{aligned}$$

By Lemma 5.3 (applied for any j to the domain D_j and the stopping time $\nu \wedge T \leq \tau_j \wedge \bar{\tau}_j$) we have for some $K \in \mathcal{C}(\mathfrak{S})$

$$\varepsilon^{-\alpha} |\mathbb{E}(f(S_{\nu \wedge T}) - f(\bar{S}_{\nu \wedge T}))| \leq K\varepsilon^{\eta_\sigma}. \tag{5.25}$$

For the next term we have (using that $\varepsilon \leq 1$)

$$\begin{aligned} |\mathbb{E}(f(S_{\tau \wedge T}) - f(S_{\nu \wedge T}))| &\leq \mathbb{E}\left(\left| \int_{\nu \wedge T}^{\tau \wedge T} \left[\nabla f(S_t) b_t + \frac{1}{2} \operatorname{Tr}(\sigma_t^\mathsf{T} \nabla^2 f(S_t) \sigma_t) \right] \mathrm{d}t \right| \right) \\ &\leq C_f (b_{\max} L_D^{-(\alpha-1)} + \frac{1}{2} \sqrt{d} \Lambda^\sigma_{\max} L_D^{-(\alpha-2)}) \varepsilon^{\alpha-2} \mathbb{E}(|\tau \wedge T - \nu \wedge T|) \\ &\leq C_f (b_{\max} L_D^{-(\alpha-1)} + \frac{1}{2} \sqrt{d} \Lambda^\sigma_{\max} L_D^{-(\alpha-2)}) \varepsilon^{\alpha-2} \sum_{j=1}^J \mathbb{E}(|\tau_j - \bar{\tau}_j|) \end{aligned}$$

(since the min function is Lipschitz)

$$\leq C_f(b_{\max}L_D^{-(\alpha-1)} + \frac{1}{2}\sqrt{d}\Lambda_{\max}^{\sigma}L_D^{-(\alpha-2)})\varepsilon^{\alpha-2}J\ K\ \varepsilon^{2+\eta_\sigma}$$

where we have applied Corollary 5.4 at the last inequality. We can show a similar bound for $\bar{S}$ and at the end, we obtain the advertised inequality (3.10). □

Proof of Proposition 3.5. The proof is quite similar to that of Proposition 3.4, at the end we invoke Corollary 5.5 instead of Corollary 5.4. □

Appendix

A Technical proofs

A.1 Proof of Lemma 3.6

We start with some preliminary analysis. Let $(U_i, i \geq 0)$ be i.i.d. random variables uniformly distributed on $[0,1]$ and independent of $\bar{\mathcal{F}}_T$. We keep the same notation for the extended probability space supporting these extra random variables and we simply write $\mathbb{P}_T(\cdot)$ (resp. $\mathbb{E}_T(\cdot)$) for the probability (resp. expectation) conditionally on $\bar{\mathcal{F}}_T$.
Set $V_j = G_*(U_j)$, $j \geq 0$ where $G_*(\cdot)$ is given by ($\mathbf{H}_G$): conditionally on $\bar{\mathcal{F}}_T$, these random variables are i.i.d. Let Y be the random variable given by

$$Y := \inf\{m \geq 1 : \sum_{j=1}^{m} V_j \geq T\}.$$

In view of ($\mathbf{H}_G$) there exists an a.s. finite $\bar{\mathcal{F}}_T$-measurable random integer m_0 such that a.s. we have

$$\gamma := \mathbb{P}_T\left((V_1 + \cdots + V_{m_0}) < T\right) < 1.$$

Our goal is to show that Y has finite (conditional) moments. We write for all $p > 0$

$$\begin{aligned}\mathbb{E}_T(|Y|^p) &\leq \sum_{k\geq 1} k^p \mathbb{P}_T\left(\sum_{j=1}^{k-1} V_j < T\right) \leq \sum_{k\geq 1} k^p \mathbb{P}_T\left((V_1 + \cdots + V_{m_0}) < T\right)^{\lfloor (k-1)/m_0 \rfloor} \\ &= \sum_{k\geq 1} k^p \gamma^{\lfloor (k-1)/m_0 \rfloor} < +\infty.\end{aligned} \tag{A.1}$$

We now come back to the main point about proving (3.12). For any $n \geq 0$ the grid $\mathcal{T}^n$ may be represented as a union $\mathcal{T}^{n,1} \cup \mathcal{T}^{n,2}$ (possibly non-disjoint), where $\mathcal{T}^{n,1}$ is the grid points with $\tau_i^n = \tau_{i-1}^n + \varepsilon_n^2 G_{\tau_{i-1}^n}(U_{n,i}) + \Delta_{n,i}$ and $\mathcal{T}^{n,2}$ contains the points where exit times occur first (see (2.6)). We have $N^n(I_n) \leq N^{n,1}(I_n) + N^{n,2}(I_n)$ with respect to the decomposition $\mathcal{T}^n = \mathcal{T}^{n,1} \cup \mathcal{T}^{n,2}$.

▷ *Upper bound on* $N^{n,1}(I_n)$. Note that from (A.1) we get $\mathbb{E}_T(Y) < +\infty$ a.s. Set $C := 1 + 2\mathbb{E}_T(Y)$, let $(Y_i)_{i\geq 0}$ be i.i.d. copies of Y conditionally on $\bar{\mathcal{F}}_T$, and put $m_n := \lceil \varepsilon_n^{-2}|I_n|/T \rceil \to$

$+\infty$. Let $\preceq$ denote the relation of first-order stochastic domination (conditionally on $\bar{\mathcal{F}}_T$). Then using ($\mathbf{H}_G$)-1 and the subadditivity property of counting processes we obtain

$$N^{n,1}(I_n) \preceq \inf\{m \geq 0 : \sum_{j=1}^{m} V_j \geq \varepsilon_n^{-2}|I_n|\} \preceq \sum_{i=1}^{m_n} Y_i.$$

Remark that the latter relation of domination turns into equality in distribution in the particular case of V_j having an exponential distribution due to the additivity of Poisson variables. Let $p := 2/\rho \geq 2$ for ρ in (3.11). Note that from (3.11) we have $\varepsilon_n^2/|I_n| \leq C_0\varepsilon_n^{2\rho}$ so that $\sum_{n\geq 0}(\varepsilon_n^2/|I_n|)^{p/2} < +\infty$. Applying the Markov inequality, the Burkholder inequality (see e.g. [HH80, Theorem 2.10]) and the Minkowsky inequality we obtain (for n large enough so that $m_n \geq 2$)

$$\begin{aligned}
&\mathbb{P}_T\Big(T|I_n|^{-1}\varepsilon_n^2 N^{n,1}(I_n) \geq C\Big) \\
&\quad \leq \mathbb{P}_T\left(\frac{\sum_{i=1}^{m_n} Y_i}{m_n - 1} \geq C\right) \leq \mathbb{P}_T\left(\frac{\sum_{i=1}^{m_n}(Y_i - \mathbb{E}_T(Y))}{m_n - 1} \geq 1\right) \\
&\quad \leq \mathbb{E}_T\left(\left|\frac{\sum_{i=1}^{m_n}(Y_i - \mathbb{E}_T(Y))}{m_n - 1}\right|^p\right) \leq C_{\text{Burk.}}\, m_n^{-p}\mathbb{E}_T\left(\left|\sum_{i=1}^{m_n}(Y_i - \mathbb{E}_T(Y))^2\right|^{p/2}\right) \\
&\quad \leq C_{\text{Burk.}} m_n^{-p}\left(\sum_{i=1}^{m_n}\mathbb{E}_T(|Y_i - \mathbb{E}_T(Y)|^p)^{2/p}\right)^{p/2} = C_{\text{Burk.}} m_n^{-p/2}\mathbb{E}_T(|Y - \mathbb{E}_T(Y)|^p)^{2/p} \\
&\quad \leq C_{\text{Burk.}}\mathbb{E}_T(|Y - \mathbb{E}_T(Y)|^p)^{2/p}\left(\frac{T\varepsilon_n^2}{|I_n|}\right)^{p/2}.
\end{aligned}$$

So we get

$$\sum_{n\geq 0}\mathbb{P}_T(T|I_n|^{-1}\varepsilon_n^2 N^{n,1}(I_n) \geq C) < +\infty \text{ a.s.}$$

and thus, by the Borel-Cantelli lemma, the event $\{T|I_n|^{-1}\varepsilon_n^2 N^{n,1}(I_n) \geq C\}$ occurs finitely many times conditionally on $\bar{\mathcal{F}}_T$ a.s. This proves $\sup_{n\geq 0}\varepsilon_n^2|I_n|^{-1}N^{n,1}(I_n) \leq C_1$ a.s. for some a.s. finite C_1.

$\triangleright$ *Upper bound on* $N^{n,2}(I_n)$. Denote $r_* := \inf_{0\leq t\leq T}\sup\{r \geq 0 : B_d(0,r) \subset \cap_{n\geq 0} D_t^n\}$. Let us show that $r_* > 0$ a.s. Indeed for any $n \geq 0$, we have $\inf_{0\leq t\leq T}\sup\{r \geq 0 : B_d(0,r) \subset D_t^n\} > 0$ since each D_t^n contains 0 and in view of the time-continuity of D_t^n w.r.t. the distance $\mu^J(\cdot,\cdot)$. The same holds for $(D_t)_{0\leq t\leq T}$. Now the positivity of r_* follows from the convergence of D_t^n to D_t w.r.t. $\mu^J(\cdot,\cdot)$ uniformly in $t \in [0,T]$ by ($\mathbf{H}_D^1$). For $N^{n,2}(I_n)$, we write

$$\begin{aligned}
N^{n,2}(I_n)\varepsilon_n^2 &\leq \varepsilon_n^2 + r_*^{-2}\sum_{\tau_i^n \in \mathcal{T}^{n,2}\cap I_n, \tau_{i-1}^n \in I_n}\left|S_{\tau_i^n} - S_{\tau_{i-1}^n}\right|^2 \\
&\leq \varepsilon_n^2 + r_*^{-2}\sum_{\tau_i^n \in \mathcal{T}^{n}\cap I_n, \tau_{i-1}^n \in I_n}\left|S_{\tau_i^n} - S_{\tau_{i-1}^n}\right|^2.
\end{aligned}$$

We have

$$\sum_{\tau_i^n \in \mathcal{T}^n\cap I_n, \tau_{i-1}^n \in I_n}|\Delta S_{\tau_i^n}|^2 - \int_{I_n}\mathrm{Tr}(\sigma_t\sigma_t^\mathsf{T})\mathrm{d}t = 2\int_{I_n}\Delta S_t^\mathsf{T}\sigma_t\mathrm{d}B_t + 2\int_{I_n}\Delta S_t^\mathsf{T} b_t\mathrm{d}t.$$

Further, using ($\mathbf{H}_R$)-1, we obtain that there exists an a.s. finite random variable C such that

$$\left|\int_{I_n} \Delta S_t^\mathsf{T} b_t \mathrm{d}t\right| \leq C\varepsilon_n |I_n| \quad \text{and} \quad \left|\int_{I_n} \Delta S_t^\mathsf{T} \sigma_t \mathrm{d}B_t\right| \leq C\varepsilon_n^{1-\rho}\sqrt{|I_n|},$$

where for the last inequality we apply [GL14, Corollary 2.1] for the sequence of martingales

$$M_t^n := \frac{\varepsilon_n^{\rho-1}}{\sqrt{|I_n|}} \int_0^t \mathbb{1}_{I_n}(s) \Delta S_s^\mathsf{T} \sigma_s \mathrm{d}B_s,$$

for the parameter $p := 2/\rho$ with ρ given by (3.11), in view of the quadratic variation bound

$$\langle M^n \rangle_T = \frac{\varepsilon_n^{2\rho-2}}{|I_n|} \left|\int_{I_n} \Delta S_t^\mathsf{T} \sigma_t \sigma_t^\mathsf{T} \Delta S_t \mathrm{d}t\right| \leq C\varepsilon_n^{4/p}, \qquad \sum_{n\geq 0} \langle M^n \rangle_T^{p/2} < +\infty \quad \text{a.s..}$$

Using that $\varepsilon_n^{1-\rho}/\sqrt{|I_n|} \to 0$ by (3.11), this finally implies

$$N^{n,2}(I_n) \leq 1 + r_*^{-2}\varepsilon_n^{-2} \left(|I_n| \sup_{0\leq t\leq T} \mathrm{Tr}(\sigma_t \sigma_t^\mathsf{T}) + o_n^{\text{a.s.}}(|I_n|) \right),$$

which finishes the proof. □

A.2 Proof of Proposition 4.2

First let us prove the statement for $f_t = f$, for any $t \in [0,T]$, where $f : \mathbb{R}^d \to \mathbb{R}$ is a continuous α-homogeneous deterministic function. Let $C_f := \sup_{|x|=1} |f(x)|$ and $C^{(3.2)}$ be given by ($\mathbf{H}_R$). First note that from ($\mathbf{H}_R$) and the homogeneity of f we have for all $n \geq 0$ and for all $t \in [0,T]$ a.s.

$$|f(\Delta S_t)| + |\mathbb{E}_t(f(\Delta S_{\bar{\varphi}(t)}))| \leq C_f C^{(3.2)}_{\varphi(t)} \varepsilon_n^\alpha. \tag{A.2}$$

Fix $n \geq 0$. Consider the adapted process

$$Z_t^n := \sum_{\tau_{i-1}^n < t} H_{\tau_{i-1}^n} \left(\mathbb{E}_t(f(\Delta S_{\tau_i^n})) - \mathbb{E}_{\tau_{i-1}^n}(f(\Delta S_{\tau_i^n})) \right)$$

(note that the conditional expectations are well defined, see our conventions at the end of the introduction). Define the process

$$\bar{V}_t := C_t^{(3.2)} + \sup_{0\leq s\leq t} |H_s| + \left(2C_f C_t^{(3.2)} \sup_{0\leq s\leq t} |H_s| + 1 \right)^2 \left(1 + \sum_{n\geq 0} \varepsilon_n^4 N_t^n \right).$$

Note that $\bar{V}_t$ takes finite values due to ($\mathbf{H}_R$)-2 and is adapted càdlàg and non-decreasing. Define

$$\nu_k := \inf\{t \geq 0 : \bar{V}_t \geq k\} \tag{A.3}$$

(with the convention $\nu_k = +\infty$ a.s. if $k > \bar{V}_T$). Due to boundedness of H_0 and $C_0^{(3.2)}$ we have that

$$\bar{V}_0 = C_0^{(3.2)} + |H_0| + \left(2C_f C_0^{(3.2)} |H_0| + 1 \right)^2 \leq C_{\bar{V}_0}$$

for some deterministic constant $C_{\bar{V}_0}$. Now observe that (since the jumps of N^n_t are of size 1)

$$\bar{V}_{\nu_k} \leq \left(1 + \sum_{n\geq 0} \varepsilon_n^4\right)(k \vee C_{\bar{V}_0}) =: q(k). \tag{A.4}$$

In order to justify the manipulations with the conditional expectations below we remark the following properties

$$\mathbb{1}_{\tau^n_{i-1}\leq \nu^k}|\Delta S_{\tau^n_i}|^\alpha \leq C^{(3.2)}_{\nu^k}\varepsilon_n^\alpha \leq q(k)\varepsilon_n^\alpha, \qquad \mathbb{1}_{\tau^n_{i-1}\leq \nu^k}|H_{\tau^n_{i-1}}| \leq q(k) \qquad \text{a.s. .} \tag{A.5}$$

It implies that for any stopping time θ and any continuous function Φ we have the equality

$$\mathbb{1}_{\tau^n_{i-1}\leq \nu^k\wedge\theta}\mathbb{E}_{\nu^k\wedge\theta}\Big(\mathbb{1}_{\tau^n_{i-1}\leq \nu^k\wedge\theta}\Phi(\Delta S_{\tau^n_i})\Big) = \mathbb{1}_{\tau^n_{i-1}\leq \nu^k\wedge\theta}\mathbb{E}_{\nu^k\wedge\theta}\Big(\Phi(\Delta S_{\tau^n_i})\Big). \tag{A.6}$$

Owing to (A.5), the random variable inside the conditional expectation on the left-hand side is bounded, and therefore its conditional expectation is well-defined (and in any L^p). The random variable inside the conditional expectation on the right-hand side is not necessarily integrable (essentially controlled thanks to ($\mathbf{H}_R$)), but actually, in the next computations, it will be still localised on a set of the form $\{\tau^n_{i-1} \leq \nu^k \wedge \theta\}$, on which we have the equality (A.6). Therefore, in what follows, writing $\mathbb{1}_{\tau^n_{i-1}\leq \nu^k\wedge\theta}\mathbb{E}_{\nu^k\wedge\theta}\Big(\Phi(\Delta S_{\tau^n_i})\Big)$ or $\mathbb{1}_{\tau^n_{i-1}\leq \nu^k\wedge\theta}\mathbb{E}_{\nu^k\wedge\theta}\Big(\mathbb{1}_{\tau^n_{i-1}\leq \nu^k\wedge\theta}\Phi(\Delta S_{\tau^n_i})\Big)$ is the same and gives random variables that are bounded: for the sake of brevity, we use the notation on the left-hand side of (A.6).
For $\tau^n_{i-1} < t \wedge \nu_k$ we obtain

$$|H_{\tau^n_{i-1}}|\Big(|\mathbb{E}_{t\wedge\nu_k}(f(\Delta S_{\tau^n_i}))| + |\mathbb{E}_{\tau^n_{i-1}}(f(\Delta S_{\tau^n_i}))|\Big) \leq \left(\sup_{0\leq s\leq \nu^k}|H_s|\right) 2C_f C^{(3.2)}_{\nu^k}\varepsilon_n^\alpha \leq \sqrt{q(k)}\varepsilon_n^\alpha. \tag{A.7}$$

Using in addition that $\varepsilon_n^4 N^n_{t\wedge\nu_k} \leq q(k)$, we obtain a.s.

$$\begin{aligned}|Z^n_{t\wedge\nu_k}| &\leq \sum_{\tau^n_{i-1}<t\wedge\nu_k}|H_{\tau^n_{i-1}}|\Big(|\mathbb{E}_{t\wedge\nu_k}(f(\Delta S_{\tau^n_i}))| + |\mathbb{E}_{\tau^n_{i-1}}(f(\Delta S_{\tau^n_i}))|\Big) \leq N^n_{t\wedge\nu_k}\sqrt{q(k)}\varepsilon_n^\alpha \\ &\leq q(k)^{3/2}\varepsilon_n^{\alpha-4}.\end{aligned} \tag{A.8}$$

Hence, we get that $\mathbb{E}(|Z^n_{t\wedge\nu_k}|^p) < +\infty$ a.s. for all $p \geq 1$ with an L^p-norm bound independent of $t \in [0,T]$. Using (A.5)-(A.6) to deal with the conditional expectations and (A.8) to be able to interchange the sum and the conditional expectation below, we verify that for any $0 \leq s \leq t \leq T$ we have a.s.

$$\begin{aligned}\mathbb{E}_s(Z^n_{t\wedge\nu^k}) = &\sum_{\tau^n_{i-1}<s\wedge\nu^k} H_{\tau^n_{i-1}}\Big(\mathbb{E}_{s\wedge\nu^k}(f(\Delta S_{\tau^n_i})) - \mathbb{E}_{\tau^n_{i-1}}(f(\Delta S_{\tau^n_i}))\Big) \\ &+ \mathbb{E}_s\left(\sum_{s\wedge\nu^k\leq\tau^n_{i-1}<t\wedge\nu^k} H_{\tau^n_{i-1}}\Big(\mathbb{E}_{t\wedge\nu^k}(f(\Delta S_{\tau^n_i})) - \mathbb{E}_{\tau^n_{i-1}}(f(\Delta S_{\tau^n_i}))\Big)\right) = Z^n_{s\wedge\nu^k}.\end{aligned}$$

Hence the process $(Z^n_{t\wedge\nu^k})_{0\leq t\leq T}$ is a martingale, and, in particular, it has a càdlàg modification. Using that $\nu^k = +\infty$ for $k > \bar{V}_T$ we deduce that the process $(Z^n_t)_{0\leq t\leq T}$ is cádlág.

In view of (A.2), the final result will follow from the convergence $\varepsilon_n^{2-\alpha} Z_t^n \overset{u.c.a.s.}{\underset{n\to+\infty}{\longrightarrow}} 0$. We prove it by leveraging Lemma B.2. Define

$$U_t^n := \varepsilon_n^{4-2\alpha} \sup_{0\le s\le t} |Z_s^n|^2, \qquad V_t^n := \varepsilon_n^4 N_t^n \left(2C_f C_t^{(3.2)} \sup_{0\le s\le t} |H_s|\right)^2.$$

Since N^n and Z^n are càdlag, it readily follows that U^n and V^n are càdlàg adapted processes, non-decreasing, vanishing at 0. Note that

$$\sum_{n\ge 0} V_t^n \le \left(2C_f C_t^{(3.2)} \sup_{0\le s\le t} |H_s|\right)^2 \sum_{n\ge 0} \varepsilon_n^4 N_t^n \le \bar{V}_t. \tag{A.9}$$

Let us check the hypotheses (i)-(ii)-(iii) of Lemma B.2. The assumptions (i)-(ii) follow from (A.9). We have already proved (iv) in (A.4). Now, we check the relation of domination (iii). We need to show that for some (deterministic) constant $C_0 > 0$ we have, uniformly in k and n,

$$\mathbb{E}(U_{t\wedge\nu_k}^n) \le C_0 \mathbb{E}(V_{t\wedge\nu_k}^n). \tag{A.10}$$

We proceed with the following estimate of $\mathbb{E}(|Z_{t\wedge\nu_k}^n|^2)$ using Fubini's theorem

$$\begin{aligned}
\mathbb{E}\left(|Z_{t\wedge\nu_k}^n|^2\right) &= \mathbb{E}\left(\sum_{\tau_{i-1}^n < t\wedge\nu_k} H_{\tau_{i-1}^n}^2 \left(\mathbb{E}_{t\wedge\nu_k}(f(\Delta S_{\tau_i^n})) - \mathbb{E}_{\tau_{i-1}^n}(f(\Delta S_{\tau_i^n}))\right)^2\right) \\
&\quad + 2 \sum_{1\le i<j<+\infty} \mathbb{E}\Big(1_{\tau_{j-1}^n < t\wedge\nu_k} H_{\tau_{i-1}^n} \left(\mathbb{E}_{t\wedge\nu_k}(f(\Delta S_{\tau_i^n})) - \mathbb{E}_{\tau_{i-1}^n}(f(\Delta S_{\tau_i^n}))\right) \\
&\qquad\qquad \times H_{\tau_{j-1}^n} \mathbb{E}_{\tau_{j-1}^n}\left(\mathbb{E}_{t\wedge\nu_k}(f(\Delta S_{\tau_j^n})) - \mathbb{E}_{\tau_{j-1}^n}(f(\Delta S_{\tau_j^n}))\right)\Big) \\
&\le \mathbb{E}\left(\left(2C_f C_{t\wedge\nu^k}^{(3.2)} \sup_{0\le s\le t\wedge\nu^k} |H_s|\right)^2 N_{t\wedge\nu^k}^n \varepsilon_n^{2\alpha}\right) = \varepsilon_n^{2\alpha-4} \mathbb{E}\left(V_{t\wedge\nu^k}^n\right),
\end{aligned}$$

where we used (A.8) to interchange the sum and the expectation, and (A.6)-(A.7) to justify that the expectations of the cross-products are well defined and equal 0. In particular, since the process in the right-hand side of the last inequality is non-decreasing, we obtain

$$\varepsilon_n^{4-2\alpha} \sup_{0\le s\le t} \mathbb{E}(|Z_{s\wedge\nu_k}^n|^2) \le \mathbb{E}(V_{t\wedge\nu_k}^n). \tag{A.11}$$

Applying Doob's L^2-inequality ([RY99, Theorem II.1.7]) to the càdlàg martingale $(Z_{t\wedge\nu^k})_{0\le t\le T}$, we obtain

$$\mathbb{E}\left(\sup_{0\le s\le t} |Z_{s\wedge\nu_k}^n|^2\right) \le 4 \sup_{0\le s\le t} \mathbb{E}\left(|Z_{s\wedge\nu_k}^n|^2\right).$$

Combining this estimate with (A.11) and from the definition of U_t^n we get

$$\mathbb{E}(U_{t\wedge\nu_k}^n) = \varepsilon_n^{4-2\alpha} \mathbb{E}\left(\sup_{0\le s\le t} |Z_{s\wedge\nu_k}^n|^2\right) \le 4\mathbb{E}(V_{t\wedge\nu_k}^n).$$

The convergence $\varepsilon_n^{2-\alpha} Z_t^n \overset{u.c.a.s.}{\underset{n\to+\infty}{\longrightarrow}} 0$ now follows from Lemma B.2.

To complete the proof in the general case $f_t = \sum_{\text{finitely many } k} f_t^k P_k$, simply apply the above result to $H_t f_t^k$ and P_k for each k. □

B Supplementary material

B.1 Decomposition of symmetric matrix into non-negative and non-positive parts

Lemma B.1. *Let $(M_t)_{0\le t\le T}$ be an $\mathcal{S}_d$-valued continuous adapted process on some filtered probability space. Then we can decompose $M_t = M_t^+ - M_t^-$ where M_t^+ and M_t^- are $\mathcal{S}_d^+$-valued continuous adapted processes; this decomposition, however, is not unique.*

Proof. Let $\lambda_t := \max(\lambda_{\max}(M_t), 0)$. By Hoffman and Wielandt's theorem [HJ90, p. 368], $(\lambda_t)_{0\le t\le T}$ is continuous and we may take $M_t^+ := \lambda_t \,\mathrm{Id}, \quad M_t^- := \lambda_t \,\mathrm{Id} - M_t$. □

B.2 Fundamental lemma on the almost sure convergence of processes

The following lemma is inspired from [GL14, Lemma 2.1], but its assumptions better fit our setting.

Lemma B.2. *Let $(U^n)_{n\ge0}$ and $(V^n)_{n\ge0}$ be two sequences of non-negative measurable processes. Assume that:*

(i) the series $\sum_{n\ge0} V_t^n$ converges for all $t\in[0,T]$ a.s.;

(ii) the above limit is upper bounded by a non-decreasing adapted càdlàg process $\bar V$;

(iii) there is a constant $c_{(B.1)} \ge 0$ such that, for every $n\in\mathbb{N}$, $k\in\mathbb{N}$ and $t\in[0,T]$, we have

$$\mathbb{E}[U^n_{t\wedge\nu^k}] \le c_{(B.1)}\ \mathbb{E}[V^n_{t\wedge\nu^k}] \tag{B.1}$$

with the stopping time $\nu^k := \inf\{s\in[0,T] : \bar V_s \ge k\}$ (with the usual convention that $\inf\emptyset = +\infty$);

(iv) there is a deterministic function $q:\mathbb{N}\to\mathbb{R}^+$ such that $q(k)\ge k$ and $\bar V_{\nu^k}\le q(k)$ for any k a.s.

Then for any $t\in[0,T]$, the series $\sum_{n\ge0} U_t^n$ converges almost surely. As a consequence, $U_t^n \overset{\text{a.s.}}{\to} 0$.

Proof. Let $t\in[0,T]$ be fixed. Denote by $\mathcal{N}_V$ the subset of Ω on which the series $(\sum_{n\ge0} V_t^n)_{0\le t\le T}$ do not converge, on which $\bar V$ and then $(\nu^k)_{k\ge0}$ are not defined and on which the inequalities of (iv) are not fulfilled; note that $\mathcal{N}_V$ is built as a countable union of negligible sets, thus it is $\mathbb{P}$-negligible.
For $\omega\notin\mathcal{N}_V$, we have $\bar V_{t\wedge\nu^k}(\omega)\le q(k)$ for any $k\in\mathbb{N}$. Set $\bar V^p := \sum_{n=0}^p V^n$: we have $\bar V^p \le \bar V$ on $\mathcal{N}_V^c$; thus, the localization of $\bar V$ entails that of $\bar V^p$ and we have $\bar V^p_{t\wedge\nu^k}\le q(k)$ for any k,p (on $\mathcal{N}_V^c$). Furthermore the relation of domination (iii) writes

$$\mathbb{E}\left[\sum_{n=0}^p U^n_{t\wedge\nu^k}\right] \le c_{(B.1)}\ \mathbb{E}\left[\sum_{n=0}^p V^n_{t\wedge\nu^k}\right] = c_{(B.1)}\ \mathbb{E}\big[\bar V^p_{t\wedge\nu^k}\big] \le c_{(B.1)}\ q(k) \tag{B.2}$$

for any k,p (on $\mathcal{N}_V^c$). From Fatou's lemma we get $\mathbb{E}[\sum_{n\ge0} U^n_{t\wedge\nu^k}] < +\infty$ for any k, therefore the series $\sum_{n\ge0} U^n_{t\wedge\nu^k}(\omega)$ converges for all ω outside of a $\mathbb{P}$-negligible set $\mathcal{N}_{k,t}$. The set $\mathcal{N}_t :=$

$\bigcup_{k\in\mathbb{N}}\mathcal{N}_{k,t}\bigcup\mathcal{N}_V$ is $\mathbb{P}$-negligible, and it follows that for $\omega\notin\mathcal{N}_t$, the series $\sum_{n\geq 0}U^n_{t\wedge\nu^k}(\omega)$ converges for all $k\in\mathbb{N}$. For $\omega\notin\mathcal{N}_t$, we have $\nu^k(\omega)=+\infty$ as soon as $k>\bar{V}_T(\omega)$; thus by taking such k, we complete the convergence of $\sum_{n\geq 0}U^n_t$ on $\mathcal{N}_t^c$. □

B.3 Almost sure uniform convergence of stochastic integrals w.r.t. a Brownian semimartingale

Lemma B.3. *Assume that a process S and a sequence of discretization grids $\mathcal{T}$ verify ($\mathbf{H}_S^{gen.}$) and ($\mathbf{H}_R$)-1 with a sequence $(\varepsilon_n : n\geq 0)$ such that $\sum_{n\geq 0}\varepsilon_n^2<+\infty$. Let $(H_t)_{0\leq t\leq T}$ be an adapted continuous scalar process and let $f:\mathbb{R}^d\to\mathbb{R}$ be a α-homogeneous function with $\alpha>0$. Then for any $k=1,\dots,d$ we have*

$$\varepsilon_n^{-\alpha+1}\sum_{\tau^n_{i-1}<t}H_{\tau^n_{i-1}}\int_{\tau^n_{i-1}}^{\tau^n_i\wedge t}f(\Delta S_s)\mathrm{d}S^k_s\xrightarrow[n\to+\infty]{u.c.a.s.}0.$$

Proof. Using the decomposition $S=A+M$, we write

$$\sum_{\tau^n_{i-1}<t}H_{\tau^n_{i-1}}\int_{\tau^n_{i-1}}^{\tau^n_i\wedge t}f(\Delta S_s)\mathrm{d}S^k_s=\int_0^t H_{\varphi(s)}f(\Delta S_s)\mathrm{d}A^k_s+\int_0^t H_{\varphi(s)}f(\Delta S_s)\mathrm{d}M^k_s.$$

First, the assumption ($\mathbf{H}_R$)-1 and the inequality $|f(x)|\leq C_f|x|^\alpha$ yield

$$\varepsilon_n^{-\alpha+1}\left|\int_0^t H_{\varphi(s)}f(\Delta S_s)\mathrm{d}A^k_s\right|\leq\varepsilon_n C_f\sup_{0\leq t\leq T}|H_t|\int_0^t(\varepsilon_n^{-1}|\Delta S_s|)^\alpha\mathrm{d}|A^k|_s\xrightarrow[n\to+\infty]{u.c.a.s.}0.$$

Second, the quadratic variation of the Brownian stochastic integral is

$$\left\langle\varepsilon_n^{-\alpha+1}\int_0^\cdot H_{\varphi(s)}f(\Delta S_s)\mathrm{d}M^k_s\right\rangle_T\leq C_f^2\sup_{0\leq t\leq T}|H_t|^2\sup_{0\leq t\leq T}|(\sigma_t\sigma_t^\mathsf{T})_{kk}|\varepsilon_n^{-2\alpha+2}\int_0^T|\Delta S_s|^{2\alpha}\mathrm{d}s\leq C\varepsilon_n^2$$

for some a.s. finite random variable $C>0$ (using again ($\mathbf{H}_R$)-1). Thus using that $\sum_{n\geq 0}\varepsilon_n^2<+\infty$ and applying [GL14, Corollary 2.1] we get

$$\varepsilon_n^{-\alpha+1}\int_0^t H_{\varphi(s)}f(\Delta S_s)\mathrm{d}M^k_s\xrightarrow[n\to+\infty]{u.c.a.s.}0,$$

which implies the result. □

References

[AM03] Y. Ait-Sahalia and P.A. Mykland. The effects of random and discrete sampling when estimating continuous-time diffusions. *Econometrica*, 71(2):483–549, 2003.

[AM04] Y. Ait-Sahalia and P.A. Mykland. Estimators of diffusions with randomly spaced discrete observations: A general theory. *Annals of Statistics*, 32:2186–2222, 2004.

[Bil95] P. Billingsley. *Probability and measure*. Wiley Series in Probability and Statistics. Wiley-Interscience, third edition, 1995.

[BS02] A.N. Borodin and P. Salminen. *Handbook of Brownian motion: facts and formulae*. Springer Basel AG, second edition, 2002.

[DH13] M. Deaconu and S. Herrmann. Hitting time for Bessel processes - walk on moving spheres algorithm. *Annals of Applied Probability*, 23(6):2259–2289, 2013.

[FR12] M. Fukasawa and M. Rosenbaum. Central limit theorems for realized volatility under hitting times of an irregular grid. *Stochastic Processes and their Applications*, 122(12):3901–3920, 2012.

[Fuk10] M. Fukasawa. Central limit theorem for the realized volatility based on tick time sampling. *Finance Stoch.*, 14:209–233, 2010.

[Fuk11a] M. Fukasawa. Asymptotically efficient discrete hedging. Stochastic analysis with financial applications. *Progress in Probability*, 65:331–346, 2011.

[Fuk11b] M. Fukasawa. Discretization error of stochastic integrals. *Annals of Applied Probability*, 21(4):1436–1465, 2011.

[GJ93] V. Genon-Catalot and J. Jacod. On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Annales de l'I.H.P. Probabilités et statistiques*, 29(1):119–151, 1993.

[GL14] E. Gobet and N. Landon. Almost sure optimal hedging strategy. *Annals of Applied Probability*, 24(4):1652–1690, 2014.

[GM10] E. Gobet and S. Menozzi. Stopped diffusion processes: boundary corrections and overshoot. *Stochastic Processes and Their Applications*, 120:130–162, 2010.

[GS18a] E. Gobet and U. Stazhynski. Model-adaptive optimal discretization of stochastic integrals. *In revision for Stochastics*, 2018.

[GS18b] E. Gobet and U. Stazhynski. Optimal discretization of stochastic integrals driven by general Brownian semimartingale. *Annales de l'Institut Henri Poincaré*, 54(3):1556–1582, 2018.

[GS18c] E. Gobet and U. Stazhynski. Parametric inference for diffusions observed at stopping times. *Preprint*, 2018.

[GT83] D. Gilbarg and N.S. Trudinger. *Elliptic partial differential equations of second order*. Springer Verlag, second edition, 1983.

[GW02] J. Grammig and M. Wellner. Modeling the interdependence of volatility and intertransaction duration processes. *Journal of Econometrics*, 106:369–400, 2002.

[HH80] P. Hall and C. C. Heyde. *Martingale limit theory and its application*. Academic Press, Inc., 1980.

[HJ90] R.A. Horn and C.R. Johnson. *Matrix analysis*. Cambridge University Press, 1990. Reprinted with corrections.

[HP18] A. Henrot and M. Pierre. *Shape Variation and Optimization, a geometrical analysis*. Tracts in Mathematics 28. European Mathematical Society, 2018.

[JP12] J. Jacod and P. Protter. *Discretization of Processes*, volume 67 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, 2012.

[JS02] J. Jacod and A.N. Shiryaev. *Limit theorems for stochastic processes*. Springer Verlag, Berlin, second edition, 2002.

[Lan13] N. Landon. *Almost sure optimal stopping times : theory and applications*. PhD thesis, Ecole Polytechnique, http://pastel.archives-ouvertes.fr/pastel-00788067, 2013.

[LMR+14] Y. Li, P.A. Mykland, E. Renault, L. Zhang, and X. Zheng. Realized volatility when sampling times are possibly endogenous. *Econometric Theory*, 30(3):580–605, 2014.

[LR13] C. Lindberg and H. Rootzén. Error distributions for random grid approximations of multidimensional stochastic integrals. *Ann. Appl. Probab.*, 23:834–857, 2013.

[LZZ13] Y. Li, Z. Zhang, and X. Zheng. Volatility inference in the presence of both endogenous time and microstructure noise. *Stochastic Processes and their Applications*, 123(7):2696–2727, 2013.

[Nor98] J.R. Norris. *Markov chains*. Cambridge university press, 1998.

[Pro04] P. Protter. *Stochastic integration and differential equations*. Springer Verlag, second edition, 2004.

[RR10] C.Y. Robert and M. Rosenbaum. On the microstructural hedging error. *SIAM J. Financial Math.*, 1:427–453, 2010.

[RR12] C.Y. Robert and M. Rosenbaum. Volatility and covariation estimation when microstructure noise and trading times are endogenous. *Mathematical Finance*, 22(1):133–164, 2012.

[RY99] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*. Berlin: Springer, third edition, 1999.

[Str10] R.L. Streit. *Poisson Point Processes. Imaging, Tracking, and Sensing*. Springer Science+Business Media, LLC. Springer, 2010.

[ZS16] L. Zhou and Z. Su. Discretization error of irregular sampling approximations of stochastic integrals. *Applied Mathematics-A Journal of Chinese Universities*, 31:296–306, 2016.