

Non-asymptotic Gaussian estimates for the recursive approximation of the invariant distribution of a diffusion

I. Honoré^a, S. Menozzi^{a,b} and G. Pagès^c

^aLaboratoire de Mathématiques et Modélisation d'Évry (LaMME), Université d'Évry Val d'Essonne, Paris Saclay, 23 Boulevard de France, 91037, Evry, France. E-mail: igor.honore@univ-evry.fr; stephane.menozzi@univ-evry.fr

^bLaboratory of Stochastic Analysis, Higher School of Economics, Pokrovsky Boulevard, 11, Moscow, Russian Federation

^cLaboratoire de Probabilités, Statistique & Modélisation. Campus Pierre et Marie Curie, Sorbonne Université, case courrier 158, 4 pl. Jussieu 75252 Paris Cedex 5, France. E-mail: gilles.pages@sorbonne-universite.fr

Received 10 July 2017; revised 12 March 2019; accepted 25 March 2019

Abstract. We obtain non-asymptotic Gaussian concentration bounds for the difference between the invariant distribution ν of an ergodic Brownian diffusion process and the empirical distribution of an approximating scheme with decreasing time step along a suitable class of (smooth enough) test functions f such that $f - \nu(f)$ is a coboundary of the infinitesimal generator. We show that these bounds can still be improved when some suitable squared-norms of the diffusion coefficient also belong to this class. We apply these estimates to design computable non-asymptotic confidence intervals for the approximating scheme. As a theoretical application, we finally derive non-asymptotic deviation bounds for the almost sure Central Limit Theorem.

Résumé. Nous obtenons des estimées de concentration gaussienne non asymptotiques pour la différence entre la mesure invariante ν d'une diffusion brownienne ergodique et la mesure empirique d'un schéma d'approximation à pas décroissants évaluée le long d'une classe admissible de fonctions tests f telles que $f - \nu(f)$ soit un co-bord du générateur infinitésimal. Nous montrons que ces bornes peuvent être améliorées lorsque le carré de certaines normes du coefficient de diffusion appartient également à cette classe. Nous déduisons de ces estimées des intervalles de confiance non asymptotiques explicitement calculables pour le schéma d'approximation. Nous obtenons également, en terme d'application théorique, des estimées de déviations non asymptotiques pour le théorème de la limite centrale presque sûr.

MSC: Primary 60H35; 60E15; secondary 35J15

Keywords: Invariant distribution; Diffusion processes; Inhomogeneous Markov chains; Non-asymptotic Gaussian concentration; Almost sure Central Limit Theorem

1. Introduction

1.1. Setting

The aim of this article is to approach the invariant distribution of the solution of the diffusion equation:

$$dY_t = b(Y_t) dt + \sigma(Y_t) dW_t, \quad (1.1)$$

where $(W_t)_{t \geq 0}$ is a Wiener process of dimension r on a given filtered probability space $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0}, \mathbb{P})$, $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$, and $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^r$ are assumed to be Lipschitz continuous functions and to satisfy a mean-reverting assumption in the following sense. If \mathcal{A} denotes the infinitesimal generator of the diffusion (1.1), there exists a twice continuously differentiable Lyapunov function $V : \mathbb{R}^d \rightarrow (0, +\infty)$ such that $\lim_{|x| \rightarrow +\infty} V(x) = +\infty$ and $\mathcal{A}V \leq \beta - \alpha V$ where $\beta \in \mathbb{R}$ and $\alpha > 0$. Such a condition ensures the existence of an invariant distribution, see [16]. We will also assume uniqueness of the invariant distribution, denoted from now by ν . For in-depth discussions on the conditions yielding such existence and uniqueness results, we refer to the monograph by Khasminskii [23] (see also its augmented second edition [22]) and the survey paper [33]. Let us also mention Villani [45] for the convergence to the equilibrium distribution of the solution of a particular Fokker–Planck equation which ensures the existence of an invariant distribution.

We introduce an approximation algorithm based on an Euler like discretization with decreasing time step, which may use more general innovations than the Brownian increments. Namely, for the step sequence $(\gamma_k)_{k \geq 1}$ and $n \geq 0$, we define:

$$X_{n+1} = X_n + \gamma_{n+1} b(X_n) + \sqrt{\gamma_{n+1}} \sigma(X_n) U_{n+1}, \quad (\text{S})$$

where $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P})$ and $(U_n)_{n \geq 1}$ is an i.i.d. sequence of centered random variables matching the moments of the Gaussian law on \mathbb{R}^r up to order three, independent of X_0 .

We define the empirical (random) occupation measures of the scheme in the following way. For any $A \in \mathcal{B}(\mathbb{R}^d)$ (where $\mathcal{B}(\mathbb{R}^d)$ denotes the Borel σ -field on \mathbb{R}^d):

$$\nu_n(A) := \nu_n(\omega, A) := \frac{\sum_{k=1}^n \gamma_k \delta_{X_{k-1}(\omega)}(A)}{\sum_{k=1}^n \gamma_k}, \quad n \geq 1. \quad (1.2)$$

The measure ν_n is here defined accordingly to the *intrinsic* time scale of the scheme. Namely, $\Gamma_n = \sum_{k=1}^n \gamma_k$ represents the current time associated with the Euler scheme (S) after n iterations. Since we are interested in long time approximation, we consider steps $(\gamma_k)_{k \geq 1}$ such that $\Gamma_n := \sum_{k=1}^n \gamma_k \xrightarrow{n} +\infty$. We also assume $\gamma_k \downarrow_k 0$. Observe that, for a bounded ν -a.s. continuous function f , it is proved in [25] (see e.g. Theorem 1), that:

$$\nu_n(f) = \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k f(X_{k-1}) \xrightarrow[n]{a.s.} \nu(f) = \int_{\mathbb{R}^d} f(x) \nu(dx), \quad (1.3)$$

or equivalently that $\nu_n(\omega, \cdot) \xrightarrow[n]{w} \nu$, $\mathbb{P}(d\omega)$ -a.s. The above result can be seen as an inhomogeneous counterpart of *stability results* discussed for homogeneous Markov chains in Duflo [15]. Intuitively, the decreasing steps make the approximation more and more accurate in long time and, therefore, the ergodic empirical mean of the scheme converges to the quantity of interest. Put it differently, there is no bias. This is a significant advantage w.r.t. a more *naive* discretization method that would rely on a constant step scheme. Indeed, even if this latter approach gains in simplicity, taking $\gamma_k = h > 0$ in (S) would lead to replace the r.h.s. of (1.3) by the quantity $\nu^h(f) := \int_{\mathbb{R}^d} f(x) \nu^h(dx)$, with ν^h denoting the invariant distribution of the Euler scheme with step h . In such a case, for the analysis to be complete, one needs to investigate the difference $\nu - \nu^h$ through the corresponding continuous and discrete Poisson problems. We refer to Talay *et al.* [43,44] for a precise presentation of this approach.

Now, once (1.3) is available, the next question naturally concerns the rate of that convergence. This was originally investigated by Lamberton and Pagès [25] for functions f of the form $f - \nu(f) = \mathcal{A}\varphi$, i.e. $f - \nu(f)$ is a *coboundary* for \mathcal{A} . The specific reason for focusing on such a class of functions is that an invariant distribution ν is characterized as a solution in the distribution sense of the stationary Fokker–Planck equation $\mathcal{A}^* \nu = 0$ (where \mathcal{A}^* stands for the adjoint of \mathcal{A}). Thus, for smooth enough functions φ (at least $\mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$), one has $\nu(\mathcal{A}\varphi) = \int_{\mathbb{R}^d} \mathcal{A}\varphi(x) \nu(dx) = 0$. The authors then investigate the weak convergence of $\nu_n(f) - \nu(f)$ once suitably renormalized. However, in these results, the assumptions are made on the function φ itself rather than on f . To overcome this limitation and exploit directly some assumptions on the function f requires to solve the Poisson equation $\mathcal{A}\varphi = f - \nu(f)$. This is precisely for this step that some structure conditions are needed, namely (hypo)ellipticity or confluence conditions. We refer for instance to the work of Pardoux and Veretennikov [37], Rothschild and Stein [41] or Villani [45] who discuss the solvability of the Poisson and of the Fokker–Planck problems under some ellipticity or hypoellipticity assumptions. We also mention the work of Pagès and Panloup [35] who exploit some confluence conditions allowing to handle for instance the case of an Ornstein–Uhlenbeck process with degenerate covariance matrix. We refer to Sections 2.2 and 2.3 for precise assumptions giving the uniqueness of the invariant distribution of (1.1) and the expected smoothness properties for the associated Poisson problem.

In the current paper, our goal is to establish for this recursive procedure a non-asymptotic Gaussian control for the deviations of the quantity $\nu_n(f) - \nu(f)$ for *possibly unbounded* Lipschitz continuous functions f . Such non-asymptotic bounds are crucial in many applicative fields. Indeed, for specific practical simulations, it is not always possible to run ergodic means for very large values of n . It will be direct to derive, as a by-product of our deviations estimates, some *computable* non-asymptotic confidence intervals. A specific feature of such non-asymptotic deviation inequalities is that their accuracy depends again on the status of the diffusion coefficient σ with respect to the Poisson equation. Thus, if $\|\sigma\|^2 - \nu(\|\sigma\|^2) = \mathcal{A}\vartheta$ is a coboundary (where $\|\cdot\|$ denotes a matrix norm), we manage to improve our analysis, to derive better concentration bounds in a certain deviation range as well as some additional deviation regimes. Also, this additional study seems rather efficient to capture the numerical behavior of the empirical deviations. We refer to Sections 4 and 6.2 for details about these points. Eventually, our main deviation results allow to provide deviation inequalities for plain Lipschitz continuous sources f in the ergodic approximation, by using a suitable regularization procedure, as established in Theorem 7. As expected, dealing with this general class of functions requires more stringent constraints on

the time steps, that must be *small* enough, and prevents from obtaining the fastest convergence rates (see again Theorem 7 and Section 5.3).

The main feature of the sequence (1.2) of weighted empirical measures is that it targets the *true* invariant distribution ν of the continuous time diffusion. The price to pay is the use of an Euler scheme with decreasing step which is a non-homogeneous Markov chain. This induces new difficulties compared to the extensive literature on deviation inequalities for ergodic homogeneous Markov chains. In particular, our approximation procedure produces some remainder terms that need to be controlled accurately enough in a non-asymptotic way to produce tractable deviation inequalities asymptotically close to their counterparts for the diffusion itself. This is a major difficulty compared to a *CLT* where these remainder terms are simply requested to go to 0 fast enough.

As mentioned above and like for the *CLT* (see [6] for the diffusion or [25] for the weighted empirical measures ν_n), these deviation inequalities are naturally established for coboundaries $f - \nu(f) = \mathcal{A}(\varphi)$, the assumptions being made on φ . Our second objective in this paper is to state our results so that all assumptions could be read on the source function f itself. This first requires to solve the Poisson equation in that spirit, that means deriving pointwise regularity on φ from that supposed on f . Again, for Lipschitz sources, this step will require an appropriate regularization procedure.

In particular, we will not rely on the Sobolev regularity (see e.g. Pardoux and Veretennikov [37]) but rather on some Schauder estimates in line with the works by Krylov and Priola [24], which allow to benefit from the elliptic regularity for operators with unbounded coefficients. For more details, we refer to the introduction of Section 5.

1.2. Assumptions and related asymptotic results

From now on, we will extensively use the following notations.

For a given step sequence $(\gamma_n)_{n \geq 1}$, we denote:

$$\forall \ell \in \mathbb{R}, \quad \Gamma_n^{(\ell)} := \sum_{k=1}^n \gamma_k^\ell, \quad \Gamma_n := \sum_{k=1}^n \gamma_k = \Gamma_n^{(1)}.$$

In practice, we will consider time step sequences: $\gamma_n \asymp \frac{1}{n^\theta}$ with $\theta \in (0, 1]$, where for two sequences $(u_n)_{n \in \mathbb{N}}, (v_n)_{n \in \mathbb{N}}$ the notation $u_n \asymp v_n$ means that $\exists n_0 \in \mathbb{N}, \exists C \geq 1$ such that $\forall n \geq n_0, C^{-1}v_n \leq u_n \leq Cv_n$.

For a vector $v \in \mathbb{R}^k, k \in \{d, r\}$, we denote by $|v| := (\sum_{j=1}^k v_j^2)^{\frac{1}{2}}$ its (canonical) Euclidean norm. Also, for a function $\psi : \mathbb{R}^q \rightarrow \mathbb{R}^d$, we set $\|\psi\|_\infty := \sup_{x \in \mathbb{R}^q} |\psi(x)|$.

Hypotheses

(C1) The random variable X_0 is supposed to be sub-Gaussian, *i.e.* its square is exponentially integrable up to some threshold. Namely, there exists $\lambda_0 \in \mathbb{R}_+^*$ such that:

$$\forall \lambda < \lambda_0, \quad \mathbb{E}[\exp(\lambda|X_0|^2)] < +\infty.$$

(GC) The μ -distributed i.i.d. innovation sequence $(U_n)_{n \geq 1}$ is such that $\mathbb{E}[U_1] = 0$ and for each $(i, j, k) \in \{1, \dots, r\}^3$, $\mathbb{E}[U_1^i U_1^j] = \delta_{ij}, \mathbb{E}[U_1^i U_1^j U_1^k] = 0$. Also, $(U_n)_{n \geq 1}$ and X_0 are independent. Eventually, U_1 satisfies the following Gaussian concentration property, *i.e.* for every 1-Lipschitz continuous function $g : \mathbb{R}^r \rightarrow \mathbb{R}$ and every $\lambda > 0$:

$$\mathbb{E}[\exp(\lambda g(U_1))] \leq \exp\left(\lambda \mathbb{E}[g(U_1)] + \frac{\lambda^2}{2}\right).$$

Observe that if $U_1 \stackrel{\text{(law)}}{=} \mathcal{N}(0, I_r)$ or $U_1 \stackrel{\text{(law)}}{=} (\frac{1}{2}(\delta_1 + \delta_{-1}))^{\otimes r}$, *i.e.* for Gaussian or symmetric Bernoulli increments which are the most commonly used sequences for the innovations, the above identity holds. On the other hand, what follows can be adapted almost straightforwardly for a wider class of sub-Gaussian distributions satisfying that for some $\varpi > 0$ and for any $\lambda > 0$:

$$\mathbb{E}[\exp(\lambda g(U_1))] \leq \exp\left(\lambda \mathbb{E}[g(U_1)] + \frac{\varpi \lambda^2}{4}\right), \tag{1.4}$$

which yields that for any $r > 0, \mathbb{P}[|U_1| \geq r] \leq 2 \exp(-\frac{r^2}{\varpi})$ (sub-Gaussian concentration of the innovation). The case $\varpi = 2$ corresponds to the standard Gaussian concentration. This is also the constant in the logarithmic Sobolev inequality fulfilled by the standard Gaussian measure, see [4].

(C2) There exists a positive constant κ such that,

$$\sup_{x \in \mathbb{R}^d} \|\sigma(x)\|^2 \leq \kappa,$$

where $\|\sigma(x)\|$ stands for the operator norm of $\sigma(x)$, i.e. $\|\sigma(x)\| = \sup_{z \in \mathbb{R}^r, |z| \leq 1} |\sigma(x)z|$ (keep in mind that $\|\sigma(x)\| = \|\sigma^*(x)\| = \|\sigma\sigma^*(x)\|^{\frac{1}{2}}$). We then set $\|\sigma\|_\infty := \sup_{x \in \mathbb{R}^d} \|\sigma(x)\|$.

(L ν) There exists a Lyapunov function $V : \mathbb{R}^d \rightarrow [v^*, +\infty)$, with $v^* > 0$, satisfying the following conditions:

- (i) Regularity–Coercivity. V is a \mathcal{C}^2 function, $\|D^2V\|_\infty < +\infty$ and $\lim_{|x| \rightarrow \infty} V(x) = +\infty$.
- (ii) Growth control. There exists $C_V \in (0, +\infty)$ such that for any $x \in \mathbb{R}^d$:

$$|\nabla V(x)|^2 + |b(x)|^2 \leq C_V V(x).$$

- (iii) Stability. Let \mathcal{A} be the infinitesimal generator associated with the diffusion equation (1.1), defined for all $\varphi \in C_0^2(\mathbb{R}^d, \mathbb{R})$ and $x \in \mathbb{R}^d$ by:

$$\mathcal{A}\varphi(x) = \langle b(x), \nabla\varphi(x) \rangle + \frac{1}{2} \text{Tr}(\Sigma(x)D^2\varphi(x)), \quad \Sigma(x) := \sigma\sigma^*(x),$$

where, for two vectors $v_1, v_2 \in \mathbb{R}^d$, the symbol $\langle v_1, v_2 \rangle$ stands for the canonical inner product of v_1 and v_2 and for $M \in \mathbb{R}^d \otimes \mathbb{R}^d$, $\text{Tr}(M)$ denotes the trace of the matrix M .

There exist $\alpha_V > 0, \beta_V \in \mathbb{R}^+$ such that for all $x \in \mathbb{R}^d$,

$$\mathcal{A}V(x) \leq -\alpha_V V(x) + \beta_V.$$

As a consequence of (L ν)(i), there exist constants K and \bar{c} such that for $|x| \geq K$, $|V(x)| \leq \bar{c}|x|^2$, which in turn implies, from (L ν)(ii), that $|b(x)| \leq \sqrt{C_V \bar{c}}|x|$.

(U) There exists a unique invariant distribution, denoted from now on by ν , for equation (1.1).

(S) For a Lyapunov function V satisfying (L ν), we assume that the step sequence $(\gamma_k)_{k \geq 1}$ satisfies for each $k \geq 1$, $\gamma_k \leq \frac{1}{2} \min(\frac{1}{\sqrt{C_V \bar{c}}}, \frac{\alpha_V}{C_V \|D^2V\|_\infty})$.

Condition (S) means that we assume that the time steps are sufficiently *small* w.r.t. the upper bounds of the coefficients and the Lyapunov function.

Remark 1. We have assumed (U) without imposing some specific non-degeneracy conditions. Observe that (L ν) yields existence of an invariant distribution (see e.g. Chapter 4.9 in [16]). Additional structure conditions on the coefficients (hypo-)ellipticity [22,37–39], or confluence [35]) then yield uniqueness.

Assumption (S) is a technical condition which is exploited in order to derive the non-asymptotic controls of Theorem 2 (see especially the proof of Proposition 1 below).

Observe that, as soon as conditions (C2), (L ν), (U) are satisfied and $\mathbb{E}[U_1] = 0, \mathbb{E}[U_1^{\otimes 3}] = 0$, the following Central Limit Theorem (CLT) holds (see Theorems 9, 10 in [25]).

Theorem 1 (CLT). *Under (C2), (L ν), (U), if $\mathbb{E}[U_1] = 0, \mathbb{E}[U_1^{\otimes 3}] = 0$ and $\mathbb{E}[V(X_0)] < +\infty$, we have the following results.*

(a) *Fast decreasing step. If $\lim_n \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} = 0$ and $\mathbb{E}[|U_1|^6] < +\infty$, then, for any Lipschitz continuous function φ in $\mathcal{C}^3(\mathbb{R}^d, \mathbb{R})$ with $D^2\varphi$ bounded and $D^3\varphi$ bounded and Lipschitz, one has (with (\mathcal{L}) denoting weak convergence)*

$$\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \xrightarrow{(\mathcal{L})} \mathcal{N}\left(0, \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi|^2 d\nu\right).$$

(b) *Critical and slowly decreasing step. If $\lim_n \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} = \tilde{\gamma} \in (0, +\infty)$ and if $\mathbb{E}[|U_1|^8] < +\infty$, then for every Lipschitz continuous function $\varphi \in \mathcal{C}^4(\mathbb{R}^d, \mathbb{R})$ with $(D^i \varphi)_{i \in \{2,3,4\}}$ bounded:*

$$\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \xrightarrow{(\mathcal{L})} \mathcal{N}\left(\tilde{\gamma} m, \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi|^2 d\nu\right) \quad \text{if } \tilde{\gamma} < +\infty \text{ (critical decreasing step),}$$

$$\frac{\Gamma_n}{\Gamma_n^{(2)}} \nu_n(\mathcal{A}\varphi) \xrightarrow{\mathbb{P}} m \quad \text{if } \tilde{\gamma} = +\infty \text{ (slowly decreasing step),}$$

with the notations of [25]

$$m := - \int_{\mathbb{R}^d} \left(\frac{1}{2} D^2 \varphi(x) b(x)^{\otimes 2} + \Phi_4(x) \right) \nu(dx),$$

where

$$\Phi_4(x) := \int_{\mathbb{R}^r} \left(\frac{1}{2} (D^3 \varphi(x) b(x), (\sigma(x)u)^{\otimes 2}) + \frac{1}{24} D^4 \varphi(x) (\sigma(x)u)^{\otimes 4} \right) \mu(du)$$

and μ denotes the distribution of the innovations $(U_k)_{k \geq 1}$. In the above definition of Φ_4 , the term $D^3 \varphi$ stands for the order 3 tensor $(\partial_{x_i, x_j, x_k}^3 \varphi)_{(i, j, k) \in \llbracket 1, d \rrbracket^3}$ and we denote, for any $x \in \mathbb{R}^d$, by $D^3 \varphi(x) b(x)$ the $\mathbb{R}^d \otimes \mathbb{R}^d$ matrix with entries $(D^3 \varphi(x) b(x))_{ij} = \sum_{k=1}^d (D^3 \varphi(x))_{ijk} b_k(x)$, $(i, j) \in \llbracket 1, d \rrbracket^2$.

Remark 2. Let us specify that for a step sequence $(\gamma_n)_{n \in \mathbb{N}}$ such that $\gamma_n \asymp n^{-\theta}$, $\theta \in (0, 1]$, it is easily checked that case (a) occurs for $\theta \in (\frac{1}{3}, 1]$ for which $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow 0$. In case (b), that is for $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow \tilde{\gamma} > 0$, we have $\theta \in (0, \frac{1}{3}]$. If $\tilde{\gamma} < +\infty$ then $\theta = \frac{1}{3}$ and if $\tilde{\gamma} = +\infty$ then $\theta \in (0, \frac{1}{3})$.

Let us mention that, when $\sup_{n \geq 0} \frac{\Gamma_n^{(3/2)}}{\sqrt{\Gamma_n}} < +\infty$, i.e. for $\gamma_n \asymp n^{-\theta}$, $\theta \in (1/2, 1]$, the CLT of point (a) holds without the condition $\mathbb{E}[U_1^{\otimes 3}] = 0$ provided $\mathbb{E}[|U_1|^4] < +\infty$ (see Theorem 9 in [25]). Moreover, the boundedness condition (C2) can be relaxed to derive the CLT, which holds provided $\lim_{|x| \rightarrow +\infty} \frac{|\sigma^* \nabla \varphi(x)|^2}{V(x)} = 0$ (strictly sublinear diffusion) in case (a) and $\sup_{x \in \mathbb{R}^d} \frac{|\sigma^* \nabla \varphi(x)|^2}{V(x)} < +\infty$ (sublinear diffusion) in case (b). We refer again to Theorems 9 and 10 in [25] for further considerations.

Remark 3. The reader should have in mind that an ergodic result similar to the one stated in the fast decreasing step setting holds for the diffusion itself under the same structure assumptions, i.e. (C2), (L ν), (U) (see Bhattacharya [6]). In fact (C2) can be partially relaxed as well, as mentioned above. Precisely, for $(Y_s)_{s \geq 0}$ as in (1.1),

$$\frac{1}{\sqrt{t}} \int_0^t \mathcal{A}\varphi(Y_s) ds \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi|^2 dv\right) \text{ as } t \rightarrow +\infty.$$

Note that the asymptotic variance corresponds to the usual integral of the “carré du champ” w.r.t. to the invariant distribution (see again Bhattacharya [6] or the monograph by Bakry *et al.* [4]), i.e.:

$$\int_{\mathbb{R}^d} |\sigma^* \nabla \varphi(x)|^2 \nu(dx) = -2 \int_{\mathbb{R}^d} \langle \mathcal{A}\varphi, \varphi \rangle(x) \nu(dx).$$

In both settings, the normalization is the same: \sqrt{t} for the diffusion and $\sqrt{\Gamma_n}$ for the scheme. Except that, as emphasized by Theorem 1, for slowly decreasing step – when $\theta < 1/3$ – the time discretization effect becomes prominent and “hides” the CLT so that $\theta = 1/3$ (critical value between “fast” and “slow” settings) yields the fastest rate with a biased CLT.

Remark 4. We would like to mention that, in the biased case (b), for steps of the form $\gamma_k = \gamma_0 k^{-1/3}$, $k \geq 1$, it is important for a practical implementation to choose γ_0 in an appropriate way, namely by minimizing the function $\gamma_0 \mapsto c_1 \gamma_0 + c_2 \gamma_0^{-1/2}$, $c_1 = \lim_n \frac{\sum_{k=1}^n k^{-2/3}}{(\sum_{k=1}^n k^{-1/3})^{1/2}}$, $c_2 = \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi|^2 dv$, which corresponds to the mean-variance contribution deriving from the biased limit Theorem. Of course, c_2 is usually unknown, and the concrete optimization has to be performed replacing c_2 by a computable estimate, like for instance upper bounds, $c_2 \leq \|\sigma\|_\infty \|\nabla \varphi\|_\infty$, see Theorem 4 and Section 5 below.

The purpose of this work is to obtain non-asymptotic deviation results which match with the above CLT. In the current ergodic framework, the very first non-asymptotic results were established for the Euler scheme with constant time step by Malrieu and Talay in [30] when the diffusion coefficient σ in (1.1) is constant. The key tool in their approach consists in establishing a Log Sobolev inequality, which implies Gaussian concentration, for the Euler scheme. This approach allows to easily control the invariant distribution associated with the diffusion process (1.1), see e.g. Ledoux [27] or Bakry *et al.* [4] in a general framework. However Log Sobolev, and even Poincaré, inequalities turn out to be rather *rigid* tools and are not very well adapted for discretization schemes like (S) with or without decreasing steps.

Our approach relies on martingale techniques, which were already a crucial tool to establish the asymptotic results of [25] and have been successfully used in Frikha and Menozzi [17] as well to establish non-asymptotic bounds for the regular Monte Carlo error associated with the Euler discretization of a diffusion over a finite time interval $[0, T]$ and a class of stochastic algorithms of Robbins–Monro type. Roughly speaking, for a given n , we decompose the quantity $\sqrt{\Gamma_n} v_n(\mathcal{A}\varphi)$ as $M_n + R_n$ where $(M_k)_{k \geq 0}$ is a martingale which has Gaussian concentration and R_n is a remainder term to be controlled in a non-asymptotic way.

We can as well refer to the recent work by Dedecker and Gouëzel [14] who also use a martingale approach to derive non-asymptotic deviation bounds for separately bounded functionals of geometrically ergodic Markov chains on a general state space.

Let us also mention that many non-asymptotic results have been obtained based on functional inequalities. Bolley, Guillin and Villani [10] derived non-asymptotic controls for the deviations of the Wasserstein distance between a reference measure and its empirical counterpart, establishing a non-asymptotic version of the Sanov theorem. Deviation estimates for sums of weakly dependent random variables (with sub exponential mixing rates) have been considered in Merlevède *et al.* [31]. From a more dynamical viewpoint, let us mention the work of Joulin and Ollivier [21], who introduced for rather general homogeneous Markov chains a kind of *curvature* condition to derive a spectral gap for the chain, and therefore an exponential convergence of the marginal laws towards the stationary distribution. We also mention a work of Blower and Bolley [7], who obtain Gaussian concentration properties for deviations of functional of the path for metric space valued homogeneous Markov chains or Boissard [9] who established non-asymptotic deviation bounds for the Wasserstein distance between the marginal distributions and the stationary law, still in the homogeneous case. The common idea of these works is to prove some contraction properties of the transition kernel of the Markov chain in Wasserstein metric. However, this usually requires to have some continuity in Wasserstein metric for the involved transition law, see *e.g.* condition (ii) in Theorems 1.2 and 2.1 of [7]. Checking such continuity conditions figures out to be difficult in practice. Sufficient conditions, which require absolute continuity and smoothness of the transition laws are given in Proposition 2.2 of [7].

Though potentially less sharp for the derivation of constants, the adopted martingale-based approach in this work turns out to be rather simple, robust and can be very naturally adapted to both discrete innovations and inhomogeneous time steps dynamics like the one we currently consider.

It should as well allow to control deviations for functionals of the path, in the spirit of those considered in [34,35]. Also, the approach could possibly extend to diffusions with less stringent Lyapunov conditions, like the weakly mean reverting drifts considered in [26], or even to more general ergodic Markov processes, see *e.g.* Pagès and Rey [36]. These aspects will concern further research.

As an application of our non-asymptotic concentration results, we will discuss two important topics:

- The first one is of numerical interest and deals with non-asymptotic confidence intervals associated with the estimation of the ergodic mean. Such results can be very useful in practice when the computational resources are constrained (by time, by the model itself, ...). If we assume that $\varphi \in \mathcal{C}^3(\mathbb{R}^d, \mathbb{R})$, Lipschitz continuous with $(D^i \varphi)_{i \in \{2,3\}}$ bounded, such that the mapping $\mathbb{R}^d \ni x \mapsto \langle b(x), \nabla \varphi(x) \rangle$ and $D^3 \varphi$ are Lipschitz continuous, we then establish that there are explicit monotonic sequences $c_n \leq 1 \leq C_n$ converging to 1 such that for all $n \in \mathbb{N}$, $a > 0$ and $\gamma_k \asymp k^{-\theta}$, $\theta \in (\frac{1}{3}, 1]$,

$$\mathbb{P}\left[\sqrt{\Gamma_n} v_n(\mathcal{A}\varphi) \geq a\right] \leq C_n \exp\left(-c_n \frac{a^2}{2\|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2}\right). \quad (1.5)$$

When the diffusion coefficient σ is such that $\|\sigma\|^2 - \nu(\|\sigma\|^2)$ is itself a coboundary (or its counterpart for any other norm dominating $\|\cdot\|$), the previous bound can be improved in a certain deviation range for a . Namely, we are able to replace $\|\sigma\|_\infty^2$ by $\nu(\|\sigma\|^2)$ in (1.5), going thus closer to the theoretical limit variance involving the “*carré du champ*”. Moreover, a mixed regime appears in the non-asymptotic deviation bounds which dramatically improves, from the numerical viewpoint, the general case for a certain deviation range. In particular, the corresponding variance is closer to the asymptotic one given by the “*carré du champ*” (see Theorem 8 below). In accordance with the limit results of Theorem 1, the drifts associated with the fastest convergence rates can be handled as well. We obtain in full generality, results of type (1.5) under slightly weaker smoothness assumptions, considering *e.g.* $D^3 \varphi$ being $\beta \in (0, 1]$ -Hölder continuous. Eventually, under suitable ellipticity conditions on σ , we are able to give non-asymptotic deviation bounds for a Lipschitz source f as well as explicit gradient bounds for the solution φ of the corresponding Poisson problem.

- The second one is mainly theoretical and concerns non-asymptotic deviation bounds for the *celebrated* almost-sure CLT first established by Brosamler and Schatte (see [11] and [42]) and revisited through the ergodic discretization schemes viewpoint in [25].

Both applications require a careful investigation of the corresponding Poisson equation $\mathcal{A}\varphi = f - \nu(f)$. We will in particular prove that some pointwise regularity properties can be transferred from f to φ .

The paper is organized as follows. We conclude this section by introducing some notations. Our main results are presented in Section 2. We first state therein the specific concentration results for functions f writing $f = \mathcal{A}\varphi + \nu(f)$ (see Section 2.1). We then proceed with some suitable controls on the Poisson problem associated with \mathcal{A} and f in a confluent framework under the two main cases considered: namely a possibly degenerate setting, which requires a strong confluence condition and *smooth* source and coefficients, and a non degenerate setting, which allows to weaken the confluence condition as well as the smoothness assumptions on the source and the coefficients since in that case we manage to benefit from an elliptic bootstrap property (see Section 2.2). We eventually give in Section 2.3 some practical and tractable deviation bounds and non-asymptotic confidence intervals, including a Slutsky like result, for a given specific source f under the afore mentioned conditions on the coefficients of (1.1).

We prove our main concentration result in Section 3. Section 4 is devoted to the case where $\|\sigma\|^2 - \nu(\|\sigma\|^2)$ is a coboundary. We then prove in Section 5 the required controls on the Poisson equation for our deviation result to hold as well as the practical controls of Section 2.3. Section 6.1 is dedicated to the non-asymptotic deviation bounds for the almost-sure CLT and Section 6.2 to the numerical illustration of our non-asymptotic confidence intervals.

1.3. Notations

In the following, we will denote by C a real constant that may change from line to line and depend, uniformly in time, on known parameters appearing in (C1), (GC), (C2), (L_V), (S). Other possible dependencies will be explicitly specified. We will also denote by \mathcal{R}_n and e_n deterministic remainder terms that respectively converge to 1 and 0 with n . The explicit dependencies of those sequences again appear in the proofs.

For a function $f \in C^\beta(\mathbb{R}^d, \mathbb{R})$, $\beta \in (0, 1]$, we denote

$$[f]_\beta := \sup_{x \neq x'} \frac{|f(x) - f(x')|}{|x - x'|^\beta} < +\infty$$

its Hölder modulus of continuity. Observe carefully that, when f is also bounded, we have that for all $0 < \beta' < \beta$:

$$[f]_{\beta'} \leq [f]_\beta^{\frac{\beta}{\beta'}} (2\|f\|_\infty)^{1 - \frac{\beta}{\beta'}}. \tag{1.6}$$

Additionally, for $f \in C^p(\mathbb{R}^d, \mathbb{R})$, $p \in \mathbb{N}$, we set for $\beta \in (0, 1]$:

$$[f^{(p)}]_\beta := \sup_{x \neq x', |\alpha|=p} \frac{|D^\alpha f(x) - D^\alpha f(x')|}{|x - x'|^\beta} \leq +\infty,$$

where α (viewed as an element of $\mathbb{N}_0^d \setminus \{0\}$ with $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$) is a multi-index of length p , i.e. $|\alpha| := \sum_{i=1}^d \alpha_i = p$.

For notational convenience, we also introduce for $k \in \mathbb{N}_0$, $\beta \in (0, 1]$ and $m \in \{1, d, d \times r\}$ the Hölder space

$$C^{k,\beta}(\mathbb{R}^d, \mathbb{R}^m) := \left\{ f \in C^k(\mathbb{R}^d, \mathbb{R}^m) : \forall \alpha, |\alpha| \in [1, k], \sup_{x \in \mathbb{R}^d} |D^\alpha f(x)| < +\infty, [f^{(k)}]_\beta < +\infty \right\}. \tag{1.7}$$

We also denote by $C_b^{k,\beta}$ the subset of $C^{k,\beta}$ for which the functions themselves are bounded. In particular, $C^{0,1}(\mathbb{R}^d, \mathbb{R}^m)$ is the space of Lipschitz continuous functions from \mathbb{R}^d to \mathbb{R}^m and $C_b^{0,\beta}(\mathbb{R}^d, \mathbb{R}^m)$ denotes the space of bounded β -Hölder continuous functions. Observe as well that, if $f \in C^{k,\beta}$, $k \geq 1$, then f is Lipschitz continuous.

We will as well use the notation $[[n, p]]$, $(n, p) \in (\mathbb{N}_0)^2$, $n \leq p$, for the set of integers being between n and p . Also, for a given Borel function $f : \mathbb{R}^d \rightarrow E$, where E can be $\mathbb{R}, \mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^q$, $q \in \{r, d\}$, we set for $k \in \mathbb{N}_0$:

$$f_k := f(X_k),$$

where $(X_k)_{k \geq 0}$ is the scheme introduced in (S). Eventually, for $k \in \mathbb{N}_0$, we denote by $\mathcal{F}_k := \sigma((X_j)_{j \in [0, k]})$.

2. Main results

2.1. Result of non-asymptotic Gaussian concentration

Our main concentration result is given by the following theorem. In this theorem, we consider a slightly more general situation than for the CLT recalled in Theorem 1. We only assume $\varphi \in C^{3,\beta}(\mathbb{R}^d, \mathbb{R})$, $\beta \in (0, 1]$ instead of $\varphi \in C^4(\mathbb{R}^d, \mathbb{R})$ with existing bounded partial derivatives up to order four (which in particular implies that in Theorem 1 $\varphi \in C^{3,1}(\mathbb{R}^d, \mathbb{R})$).

Theorem 2. Assume (C1), (GC), (C2), (L ν), (U), (S) hold. Consider a Lipschitz continuous (possibly unbounded) function $\varphi \in C^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ for some $\beta \in (0, 1]$. Let us furthermore suppose that:

$$\exists C_{V,\varphi} > 0, \forall x \in \mathbb{R}^d, \quad |\varphi(x)| \leq C_{V,\varphi}(1 + \sqrt{V(x)}). \tag{G ν }$$

Let $\theta \in [1/(2 + \beta), 1]$ and assume the step sequence $(\gamma_k)_{k \geq 1}$ is of the form $\gamma_k \asymp k^{-\theta}$.

(a) Unbiased Case (sub-optimal convergence rate): Let $\theta \in (\frac{1}{2+\beta}, 1]$.

(i) Assume that the mapping $x \mapsto \langle \nabla\varphi(x), b(x) \rangle$ is Lipschitz continuous.

Then, there exist two explicit monotonic deterministic sequences $c_n \leq 1 \leq C_n, n \geq 1$, with $\lim_n C_n = \lim_n c_n = 1$, such that for all $n \geq 1$ and for every $a > 0$:

$$\mathbb{P}[|\sqrt{\Gamma_n}v_n(\mathcal{A}\varphi)| \geq a] \leq 2C_n \exp\left(-c_n \frac{a^2}{2\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}\right).$$

(ii) When the mapping $x \mapsto \langle \nabla\varphi(x), b(x) \rangle$ is not Lipschitz continuous. The above result still holds for $0 < a \leq \chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$ for a positive sequence $\chi_n \rightarrow 0$ arbitrarily slowly, so that $\chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}} \xrightarrow{n} +\infty$. In particular, for a fixed $a > 0$, the above concentration inequality holds for n large enough.

(b) Biased Case (Optimal Convergence Rate): Let $\theta = \frac{1}{2+\beta}$. We set for any $(k, t, u, x) \in \llbracket 1, n \rrbracket \times [0, 1]^2 \times \mathbb{R}^d$:

$$\Lambda_{k-1}^\beta(t, u, x) := \mathbb{E}[\text{Tr}((D^3\varphi(x + \gamma_k b(x) + ut\sqrt{\gamma_k}\sigma(x)U_k)\sigma(x)U_k)(\sigma(x)U_k \otimes U_k\sigma(x)^*))], \tag{2.1}$$

keeping in mind that, since $\varphi \in C^{3,\beta}(\mathbb{R}^d, \mathbb{R})$, $[D^3\varphi]_\beta < +\infty$. We define subsequently:

$$E_n^\beta := \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^{3/2} \int_0^1 dt (1-t)t \int_0^1 du \Lambda_{k-1}^\beta(t, u, X_{k-1}). \tag{2.2}$$

Set now

$$\begin{aligned} \mathcal{B}_{n,\beta} &:= E_n^\beta, \quad \text{if } \beta \in (0, 1), \\ \mathcal{B}_{n,\beta} &:= E_n^\beta + \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^2 \int_0^1 (1-t) \text{Tr}(D^2\varphi(X_{k-1} + t\gamma_k b_{k-1})b_{k-1} \otimes b_{k-1}) dt \\ &\quad + \frac{1}{2\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k \text{Tr}((D^2\varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2\varphi(X_{k-1}))\Sigma_{k-1}), \quad \text{if } \beta = 1. \end{aligned} \tag{2.3}$$

There exist two explicit monotonic deterministic sequences $c_n \leq 1 \leq C_n, n \geq 1$, with $\lim_n C_n = \lim_n c_n = 1$ such that for all $n \geq 1$ and $a > 0$:

$$\mathbb{P}[|\sqrt{\Gamma_n}v_n(\mathcal{A}\varphi) + \mathcal{B}_{n,\beta}| \geq a] \leq 2C_n \exp\left(-c_n \frac{a^2}{2\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}\right).$$

For $\beta \in (0, 1)$, the random variables $|\mathcal{B}_{n,\beta}| = |E_n^\beta| \leq \frac{[\varphi^{(3)}]_\beta \|\sigma\|_\infty^{(3+\beta)} \mathbb{E}[|U_1|^{3+\beta}]}{(1+\beta)(2+\beta)(3+\beta)} \frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} \stackrel{a.s.}{\asymp} a_{\beta,\infty} > 0$. Also, for $\beta = 1$, the $(\mathcal{B}_{n,1})_{n \geq 1}$ are exponentially integrable and if, furthermore, $D^3\varphi$ is C^1 , $\mathcal{B}_{n,1} \asymp -\tilde{\gamma}m$ a.s. where $\tilde{\gamma}m$ is as in Theorem 1. In any case, a bias appears in our deviation controls when we consider, for a given smoothness of order $\beta \in (0, 1]$ for $D^3\varphi$, the fastest associated time steps $\gamma_k \asymp k^{-\theta}$, $\theta = \frac{1}{2+\beta}$.

Remark 5. Observe that, when $\beta = 1$, the above result provides a non-asymptotic counterpart of the limit Theorem 1. In particular, the concentration constants appearing in Theorem 2 asymptotically match those of the centered CLT recalled in Theorem 1, up to a substitution of the asymptotic variance $\int_{\mathbb{R}^d} |\sigma^* \nabla\varphi(x)|^2 \nu(dx)$ by its natural upper bound $\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2$.

Importantly, these bounds do not require ‘‘a priori’’ non-degeneracy conditions and only depend on the diffusion coefficient through the sup-norm of the diffusion matrix Σ , assumption (C2). It will anyhow be very natural to consider

a non-degeneracy condition ([37,41]), or a confluence condition ([35]), when investigating the deviations for a given function f , in order to ensure the solvability of the corresponding Poisson equation $\mathcal{A}\varphi = f - \nu(f)$ and to derive explicit upper bounds for $\|\nabla\varphi\|_\infty$ in terms of the coefficients b, σ and the source f which turn out to be crucial to design computable non-asymptotic confidence intervals. These aspects are discussed in Section 2.2 below.

The alternative form of the asymptotic variance (cf. Remark 3) $\int_{\mathbb{R}^d} |\sigma^* \nabla\varphi(x)|^2 \nu(dx) = -2 \int_{\mathbb{R}^d} [f(x) - \nu(f)] \varphi(x) \nu(dx)$ suggests that for bounded source terms f , an associated natural variance bound would be $2\|f\|_\infty \|\varphi\|_\infty$. Such a control would a priori require less regularity on φ than assumed in Theorem 2. One could for instance try to exploit suitable regularization procedures, like for instance the one proposed in Section 5.3 for the proof of Theorem 7 below, to establish non-asymptotic deviation results under weaker assumptions. Our main objective being to capture unbounded Lipschitz functions f , these aspects will concern further research.

Remark 6 (Smoothness and Convergence Rate). Observe that, in coherence with the asymptotic setting of the CLT recalled in Theorem 1, for a given $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$, $\beta \in (0, 1]$, the fastest convergence rate for the deviations is attained for $\theta = \frac{1}{2+\beta}$. A bias appears, which can be difficult to estimate in practice since φ is usually unknown.

Remark 7 (On the smoothness property of $x \mapsto \langle b(x), \nabla\varphi(x) \rangle$). The Lipschitz continuity assumption on the above mapping appearing in case (i) might seem awkward at first sight. It is non-intrinsic in the sense that it involves both the drift b of the model and the test function φ . However, this condition naturally appears when φ is a smooth solution to the Poisson equation $\mathcal{A}\varphi = f - \nu(f)$. Indeed, recalling the definition of \mathcal{A} in (\mathcal{L}_ν) (iii), we can rewrite:

$$\langle \nabla\varphi(x), b(x) \rangle = f(x) - \nu(f) - \frac{1}{2} \text{Tr}(\Sigma(x) D^2\varphi(x)).$$

Hence, the Lipschitz continuity of the function in the above left hand side readily follows as soon as the source f is Lipschitz and if $D^2\varphi$ is bounded and Lipschitz continuous (since σ is also bounded and Lipschitz). Note that with the previous notations for function spaces the previous conditions are implied if $f \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R}) \subset \mathcal{C}^{0,1}(\mathbb{R}^d, \mathbb{R})$, $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R}) \Rightarrow D^2\varphi \in \mathcal{C}_b^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d) \subset \mathcal{C}_b^{0,1}(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)$. We refer to Section 5.1.2 for details.

We now state an improvement of the previous concentration bound when $\|\sigma\|^2 - \nu(\|\sigma\|^2)$ is itself a coboundary, i.e. when the Poisson problem $\mathcal{A}\vartheta = \|\sigma\|^2 - \nu(\|\sigma\|^2)$ can be solved with ϑ satisfying the assumptions required for φ in Theorem 2. Precisely, we have the following result.

Theorem 3.

(a) *Under the assumptions of Theorem 2 and with the notations introduced therein, provided that ϑ , solution to the Poisson equation $\mathcal{A}\vartheta = \|\sigma\|^2 - \nu(\|\sigma\|^2)$, satisfies the same smoothness and growth conditions as φ , for $\beta \in (0, 1]$ and $\theta \in (\frac{1}{2+\beta}, 1]$ (unbiased case), there exist two explicit monotonic deterministic sequences $\tilde{c}_n \leq 1 \leq \tilde{C}_n$, $n \geq 1$, with $\lim_n \tilde{C}_n = \lim_n \tilde{c}_n = 1$ such that for all $n \geq 1$ and $0 < a \leq \chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$ with $(\chi_n)_{n \geq 1}$ a positive sequence $\chi_n \xrightarrow{n} 0$ arbitrarily slowly, satisfying $\chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}} \xrightarrow{n} +\infty$:*

$$\mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] \leq 2\tilde{C}_n \exp\left(-\tilde{c}_n \frac{a^2}{2\nu(\|\sigma\|^2) \|\nabla\varphi\|_\infty^2}\right). \tag{2.4}$$

(b) *If ϑ solve the Poisson equation $\mathcal{A}\vartheta = \|\sigma\|^2 - \nu(\|\sigma\|^2)$ mutatis mutandis for a matrix norm dominating the operator norm ($\|\sigma(x)\| \geq \|\sigma(x)\|$), then the above bound (2.4) still holds with $\nu(\|\sigma\|^2)$ instead of $\nu(\|\sigma\|^2)$.*

Importantly, the above result allows to improve the natural variance bound $\|\nabla\varphi\|_\infty^2 \|\sigma\|_\infty^2$ of Theorem 2 by a more refined, namely $\|\nabla\varphi\|_\infty^2 \nu(\|\sigma\|^2)$. Such a bound can be particularly interesting when the supremum norm of σ is high but its average w.r.t. the invariant distribution ν significantly lower. We refer to Section 4, Theorem 8 (general form of Theorem 3) and 6.2 (numerical results) for further discussions on that topic.

Of course Claim (b) is less sharp than (a) stated with the operator norm $\|\cdot\|$ but solving the Poisson equation for $\|\sigma(x)\|$ seems highly non trivial. By contrast, if $\|\sigma(x)\| = \|\sigma(x)\|_F := [\text{Tr}(\sigma\sigma^*(x))]^{1/2}$ stands for the Fröbenius norm, Theorem 4 below yields the expected smoothness properties on $\|\sigma\|_F^2 - \nu(\|\sigma\|_F^2)$ that ensure the existence of a solution to $\mathcal{A}\vartheta = \|\sigma\|_F^2 - \nu(\|\sigma\|_F^2)$ meeting the required smoothness conditions. The price to pay with such computable norms being that they usually induce some dependence on the dimension d on the estimates (observe e.g. for the identity matrix I_d of $\mathbb{R}^d \otimes \mathbb{R}^d$, $\|I_d\|_F = d^{1/2}$).

2.2. Uniqueness of the invariant distribution and regularity issues for the Poisson problem

For our deviation analysis to work, we need to have the uniqueness of the invariant distribution ν and to establish some pointwise controls on the solution of the associated Poisson equation. Namely, we need to have quantitative bounds on its derivatives and the associated Hölder continuity modulus up to order 3.

To do so, additionally to our main assumptions introduced for Theorem 2, we will work in the confluent setting. In dimension one, any ergodic diffusion is in *some sense* confluent (see [23], Appendix of the English translation, Theorem 2.2 p. 308 and its alternative proof in [29] Theorem 2). Here, we will suppose that the following condition holds:

• *Confluence Conditions.*

(\mathbf{D}_α^p) We assume that there exists $\alpha > 0$ and $p \in (1, 2]$ such that for all $x \in \mathbb{R}^d, \xi \in \mathbb{R}^d$

$$\left\langle \frac{Db(x) + Db(x)^*}{2} \xi, \xi \right\rangle + \frac{1}{2} \sum_{j=1}^r \left((p-2) \frac{|\langle D\sigma_{\cdot j}(x) \xi, \xi \rangle|^2}{|\xi|^2} + |D\sigma_{\cdot j} \xi|^2 \right) \leq -\alpha |\xi|^2,$$

where Db stands here for the Jacobian of b , $\sigma_{\cdot j}$ stands for the j th column of the diffusion matrix σ and $D\sigma_{\cdot j}$ for its Jacobian matrix.

Within the confluent framework, we will consider from now on two kinds of assumptions which first give the uniqueness of ν and then can lead to the required smoothness and to computable gradient bounds, which are crucial since they are precisely the quantities appearing in the non-asymptotic Gaussian deviation controls as emphasized in the statement of Theorem 2.

- Strong Confluence condition and regularity of the coefficients, which means that the drift is sufficiently dominant in the dynamics and the coefficients are *smooth* (see assumption (\mathbf{C}_R) below). Note that these conditions may hold for degenerate diffusion coefficients.
- Non-degeneracy of the diffusion coefficient and mild confluence condition and smoothness on the coefficients (see assumption (\mathbf{C}_{UE}) below).

Under a sufficiently strong confluence condition, i.e. when α is large enough in (\mathbf{D}_α^p), and provided that the coefficients b, σ, f are sufficiently smooth, it is quite direct to derive, through stochastic flow techniques *à la Kunita*, the required pointwise bounds for the derivatives of the Feynman–Kac representation of the solution to the Poisson equation (see [35] and Section 5.1).

In the non-degenerate case, the main advantage is that we can alleviate some restrictions on α and the smoothness assumptions on b, σ, f to benefit from an elliptic regularity bootstrap deriving from suitable Schauder estimates available in the current setting from the work by Krylov and Priola [24].

We now introduce a smoothness assumption on b, σ, f that will be useful in both considered cases.

• *Smoothness of the coefficients and the source.* For $k \in \{1, 3\}$ and $\beta \in (0, 1)$ define

($\mathbf{R}_{k,\beta}$) The coefficients in equation (1.1) are s.t. $b \in C^{k,\beta}(\mathbb{R}^d, \mathbb{R}^d), \sigma \in C_b^{k,\beta}(\mathbb{R}^d, \mathbb{R}^d)$. Also, the source f for which we want to estimate $\nu(f)$ belong to $C^{k,\beta}(\mathbb{R}^d, \mathbb{R})$.

With these assumptions at hand, we now introduce the first setting to be considered.

◇ *The confluent and regular assumption (\mathbf{C}_R),* holds if (\mathbf{D}_α^p), ($\mathbf{R}_{3,\beta}$), for some $\beta \in (0, 1]$, are in force and $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{2(3+\beta)-p}$ where $\|D\sigma\|_\infty := \sup_{x \in \mathbb{R}^d} (\sum_{j=1}^d \|D\sigma_{\cdot j}(x)\|^2)^{\frac{1}{2}}$ recalling that, for every $j \in \llbracket 1, d \rrbracket, \|D\sigma_{\cdot j}(x)\|$ stands for the operator norm of $D\sigma_{\cdot j}(x)$.

In particular, we do not impose in this case any additional structure condition on σ which can degenerate.

In our second main framework, we will assume some uniform ellipticity conditions.

• *Non-degeneracy Conditions.*

(\mathbf{UE}) Uniform ellipticity. We assume that w.l.o.g. that $r = d$ ($r \geq d$ could also be considered) in (1.1) and that the diffusion coefficient σ is such that

$$\exists \underline{\sigma} > 0, \forall (\xi, x) \in \mathbb{R}^d \times \mathbb{R}^d, \quad \langle \sigma \sigma^*(x) \xi, \xi \rangle \geq \underline{\sigma} |\xi|^2.$$

We now introduce our second main setting:

◇ *The confluent and non-degenerate assumption (\mathbf{C}_{UE}),* holds if (\mathbf{D}_α^p), ($\mathbf{R}_{1,\beta}$), for some $\beta \in (0, 1]$, are in force. If $d > 1$, we also assume that $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{2(1+\beta)-p}$ and that the diffusion matrix Σ satisfies a structure condition: for each $(i, j) \in \llbracket 1, d \rrbracket^2, \Sigma_{i,j}(x) = \Sigma_{i,j}(x_{i \wedge j}, \dots, x_d)$.

Theorem 4. Assume that (\mathcal{L}_V) and either (\mathbf{C}_R) or (\mathbf{C}_{UE}) are in force. Then there exists a unique invariant distribution for the solution of (1.1), i.e. assumption (\mathbf{U}) holds.

The associated Poisson equation

$$\forall x \in \mathbb{R}^d, \quad \mathcal{A}\varphi(x) = f(x) - \nu(f), \tag{2.5}$$

admits a unique solution $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$, $\beta \in (0, 1)$ centered w.r.t. ν . Furthermore, the following gradient bound holds

$$\|\nabla\varphi\|_\infty \leq \frac{[f]_1}{\alpha},$$

and the mapping $x \mapsto \langle \nabla\varphi(x), b(x) \rangle$ is Lipschitz continuous.

2.2.1. About the regularity of the coefficients

Under (\mathbf{C}_R) , the derivatives can be expressed using iterated tangent processes and we cannot expect, without *a priori* any non-degeneracy condition, for a smoothing effect to hold. To have $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$, we need to consider a source $f \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ and the same smoothness on b, σ (Assumption $(\mathbf{R}_{3,\beta})$). We refer to Section 5.1 for the proof of Theorem 4 under (\mathbf{C}_R) .

In the non-degenerate case, the solvability of the Poisson problem is usually studied in a Sobolev setting, see e.g. [37]. Let us also indicate that pointwise gradient bounds have been obtained by the same authors in [38] for bounded drifts and diffusion coefficients which are additionally supposed to be *smooth*, i.e. at least $\mathcal{C}_b^{2,\beta}$ with the notations introduced in Section 1.3. We point out that these estimates do not apply in our current setting in which the drift has typically linear growth.

We eventually mention the last paper by these authors, namely [39]. They derive therein the uniqueness of the martingale solution to the Poisson equation in a potentially degenerate setting under suitable local Doeblin conditions. In that framework, pointwise controls are obtained as well for the solution itself but not for its *derivatives*.

To obtain the required smoothness, we use here in the non-degenerate framework of (\mathbf{C}_{UE}) some Schauder estimates, deriving from the work of Krylov and Priola [24], which allow to benefit from the elliptic regularity. Namely, to obtain the mentioned smoothness on φ solving $\mathcal{A}\varphi = f - \nu(f)$, that we expect to be in $\mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$, $\beta \in (0, 1)$, we can take a source $f \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R})$ and $b \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d)$, $\sigma \in \mathcal{C}_b^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d)$.

We would eventually like to emphasize that the structure condition on Σ might seem weird at first sight. It is actually needed to decouple the PDEs formally satisfied by $(\partial_{x_i}\varphi)_{i \in \llbracket 1, d \rrbracket}$ in order to exploit the *a priori* estimates of [24] established for scalar valued PDEs. We refer to Section 5.1 for a proof and details.

2.2.2. About the confluence condition and the restrictions on σ

We work here in the confluent setting of (\mathbf{D}_α^p) . This assumption will allow, through a pathwise analysis associated with the tangent flow, to derive a pointwise gradient bound. Another possibility to obtain such a bound is to assume a so-called Bakry and Émery curvature criterion, see [3,4]. Under this condition, the gradient and semi-group commute up to an exponential multiplicative factor (see equation (2.7) below).

▷ *Bakry and Émery curvature criterion.* First, we recall that the “*carré du champ*” operator Γ of a Markov process with generator \mathcal{A} reads, for every f, g in its domain $\mathcal{D}(\mathcal{A})$

$$\Gamma(f, g) := \frac{1}{2}(\mathcal{A}(fg) - f\mathcal{A}g - g\mathcal{A}f) \quad \text{and} \quad \Gamma(f) := \Gamma(f, f).$$

We also need to define the Γ_2 operator

$$\Gamma_2(f) = \frac{1}{2}(\mathcal{A}\Gamma(f) - 2\Gamma(f, \mathcal{A}f)).$$

In our Brownian diffusion setting, we have

$$\forall x \in \mathbb{R}^d, \quad \Gamma(f)(x) = |\sigma^* \nabla f(x)|^2.$$

whereas the computation of Γ_2 is significantly more involved. However, if the diffusion matrix $\Sigma = \sigma\sigma^*$ is constant then:

$$\Gamma_2(f)(x) := \text{Tr}((D^2 f(x)\Sigma)^2) - \langle \nabla f, Db\Sigma\nabla f \rangle(x).$$

With these notations at hand, we say that *the semi-group* $(P_t)_{t \geq 0}$ of \mathcal{A} satisfies the Bakry and Émery curvature criterion with parameter $\rho > 0$ if

$$(\mathbf{BE}_\rho) \quad \forall f \in \mathcal{D}(\mathcal{A}), \quad \Gamma_2(f) \geq \rho \Gamma(f).$$

Observe that for $\Sigma = I_d$ the condition (\mathbf{BE}_ρ) is actually equivalent to (\mathbf{D}_α^p) with $\alpha = \rho$ (and any $p \in (1, 2]$ since $D\sigma = 0$) and reads

$$\left\langle \frac{Db(x) + Db(x)^*}{2} \xi, \xi \right\rangle \leq -\rho |\xi|^2.$$

The computation of the Γ_2 for a general non-degenerate diffusion of the form (1.1) is not easy and is discussed in [1]. In particular, in full generality, the computation of Γ_2 requires the coefficients of the operator itself to be *smooth* (i.e. at least C^2). We also refer to [2] if the diffusion matrix is scalar diagonal, i.e. $\Sigma(x) = \varsigma(x)I_d$, $x \in \mathbb{R}^d$, where ς is real valued. In that case, it is then shown that (\mathbf{BE}_ρ) holds if and only if:

$$-\frac{1}{2} \langle (M(x) + M^*(x)) \xi, \xi \rangle \leq -\rho \varsigma(x) |\xi|^2, \tag{2.6}$$

where

$$M(x) = \frac{1}{2} (\varsigma(x) \Delta \varsigma(x) + \langle b(x), \nabla \varsigma(x) \rangle - \|\nabla \varsigma(x)\|^2) I_d + \left(\frac{1}{2} - \frac{d}{4} \right) \nabla \varsigma \otimes \nabla \varsigma(x) - \varsigma(x)^2 Db(x).$$

An important property when (\mathbf{BE}_ρ) holds, see again [3,4], is that the following commutation inequality holds:

$$\forall t \geq 0, \forall x \in \mathbb{R}^d, \quad \Gamma(P_t f)(x) \leq \exp(-2\rho t) P_t \Gamma(f). \tag{2.7}$$

To conclude, let us say that the Bakry–Emery curvature condition is a very powerful tool to derive pointwise gradient bounds. In our framework, this is unfortunately not enough as soon as $d > 1$, because additionally to this kind of bounds we also need, to enter in the framework of Schauder estimates under (\mathbf{C}_{UE}) , a control of the β -Hölder modulus of the gradient (see Section 5.1.2). It does not seem that the condition (\mathbf{BE}_ρ) helps to get such controls. The restrictions on the variations of $D\sigma$ appearing in both assumptions (\mathbf{C}_{UE}) and (\mathbf{C}_R) are precisely needed to derive in the first case the bounds on $[D\varphi]_\beta$ and in the second one to prove that the derivatives exist up to order 3 and that $[D^3\varphi]_\beta$ is controlled as well. This explains why the conditions on $D\sigma$ are more stringent in the potentially degenerate setting (\mathbf{C}_R) . In each case, those bounds are obtained through pathwise analysis and the restrictions on $D\sigma$ ensure the time integrability of the iterated tangent flows, see again Section 5.1.2 and Appendix A in [35] for details.

2.3. Practical deviation bounds

2.3.1. A first non-asymptotic confidence interval result

Theorem 5 (Non-asymptotic confidence intervals without bias). *Let the assumptions of Theorem 4 be in force. Then, there exists a unique invariant distribution ν for (1.1), i.e. (\mathbf{U}) holds. Also, φ satisfies (\mathbf{G}_ν) introduced in Theorem 2 for $V(x) \asymp 1 + |x|^2$.*

Assume that $(\mathbf{C1})$ (sub-gaussian tails of the innovation) holds and that the step sequence $(\gamma_k)_{k \geq 1}$ is such that $\gamma_k \asymp k^{-\theta}$, $\theta \in (\frac{1}{2+\beta}, 1]$. Then, for $(c_n)_{n \geq 1}$, $(C_n)_{n \geq 1}$ like in Theorem 2 with $\lim_n c_n = \lim_n C_n = 1$, we have that for all $n \geq 1$, $a > 0$ and for any matrix norm $\|\cdot\|$ dominating the operator norm $\|\cdot\|$:

$$\mathbb{P}[\sqrt{\Gamma_n} |v_n(f) - \nu(f)| > a] \leq 2C_n \exp\left(-c_n \frac{a^2 \alpha^2}{2 \|\sigma\|_\infty^2 [f]_1^2}\right) \quad \text{with } \|\sigma\|_\infty := \sup_{x \in \mathbb{R}^r} \|\sigma(x)\|, \tag{2.8}$$

$$\mathbb{P}\left[v(f) \in \left[v_n(f) - \frac{a \|\sigma\|_\infty [f]_1}{\alpha \sqrt{\Gamma_n}}, v_n(f) + \frac{a \|\sigma\|_\infty [f]_1}{\alpha \sqrt{\Gamma_n}}\right]\right] \geq 1 - 2C_n \exp\left(-c_n \frac{a^2}{2}\right), \tag{2.9}$$

where the parameter α is the same as in the pointwise gradient bound of Theorem 4.

Proof. Equation (2.8) is a direct consequence of Theorem 2 and the gradient bound in Theorem 4. Indeed, the mean-value Theorem readily yields that (\mathbf{G}_ν) holds. It then suffices to observe that $v_n(f) - \nu(f) = v_n(\mathcal{A}\varphi)$. To prove (2.9), setting

$a_{\sigma, f, \alpha} := a \|\sigma\|_{\infty} \frac{[f]_1}{\alpha}$, it suffices to write:

$$\mathbb{P}\left[v(f) \in \left[v_n(f) - \frac{a_{\sigma, f, \alpha}}{\sqrt{\Gamma_n}}, v_n(f) + \frac{a_{\sigma, f, \alpha}}{\sqrt{\Gamma_n}}\right]\right] = 1 - \mathbb{P}\left[\sqrt{\Gamma_n} |v_n(f) - v(f)| \geq a_{\sigma, f, \alpha}\right]$$

and conclude by (2.8). □

2.3.2. A more refined non-asymptotic confidence interval when $\|\sigma\|^2 - v(\|\sigma\|^2)$ is a coboundary

We provide in Theorem 6 below a kind of *Slutsky's Lemma* when, for a matrix norm $\|\cdot\|$ dominating the operator norm (i.e. $\|\sigma(x)\| \leq \|\sigma(x)\|$), $\|\sigma\|^2 - v(\|\sigma\|^2)$ is a coboundary.

Theorem 6 (Slutsky type concentration result for the coboundary case). *Under the assumptions of Theorem 5, for $\beta \in (0, 1]$ and $\theta \in (\frac{1}{2+\beta}, 1]$ (unbiased case), assuming as well that there is a unique solution ϑ to $A\vartheta = \|\sigma\|^2 - v(\|\sigma\|^2)$ satisfying the same assumptions as φ in Theorem 5, there exist two explicit monotonic deterministic sequences $c_n \leq 1 \leq C_n$, $n \geq 1$, with $\lim_n C_n = \lim_n c_n = 1$ such that for all $n \geq 1$, $a > 0$, the following bounds hold: if $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$ (Gaussian deviations) then,*

$$\mathbb{P}\left[\left|\sqrt{\Gamma_n} \frac{v_n(f) - v(f)}{\sqrt{v_n(\|\sigma\|^2)}}\right| \geq a\right] \leq 2C_n \exp\left(-c_n \frac{a^2 \alpha^2}{2[f]_1^2}\right), \tag{2.10}$$

$$\mathbb{P}\left[v(f) \in \left[v_n(f) - \frac{a\sqrt{v_n(\|\sigma\|^2)}[f]_1}{\alpha\sqrt{\Gamma_n}}, v_n(f) + \frac{a\sqrt{v_n(\|\sigma\|^2)}[f]_1}{\alpha\sqrt{\Gamma_n}}\right]\right] \geq 1 - 2C_n \exp\left(-c_n \frac{a^2}{2}\right). \tag{2.11}$$

Again, the non-asymptotic confidence interval is explicitly computable as a function of the given source f , the coefficients in the dynamics and the chosen (computable) matrix norm $\|\cdot\|$. It is also sharper than the one in (2.9).

2.3.3. Towards Lipschitz sources in the non-degenerate case

We conclude this section stating a non-asymptotic deviation result for Lipschitz sources under some non-degeneracy conditions (assumption (C_{UE}) of Theorem 4 replacing the condition stated there for f by a Lipschitz condition).

Theorem 7 (Non-asymptotic concentration bounds for Lipschitz functions). *Let the assumptions of Theorem 4 with (C_{UE}) hold except that f is here only a Lipschitz continuous function. For a time step sequence $(\gamma_k)_{k \geq 1}$ of the form $\gamma_k \asymp k^{-\theta}$, $\theta \in (1/2, 1]$, we have that, there exist two explicit monotonic sequences $c_n \leq 1 \leq C_n$, $n \geq 1$, with $\lim_n C_n = \lim_n c_n = 1$ such that for all $n \geq 1$, $a > 0$:*

$$\mathbb{P}\left[|\sqrt{\Gamma_n}(v_n(f) - v(f))| \geq a\right] \leq 2C_n \exp\left(-c_n \frac{a^2 \alpha^2}{2\|\sigma\|_{\infty}^2 [f]_1^2}\right), \tag{2.12}$$

where α is as in Theorem 4.

Such estimates are important since they allow to get rather close to the natural framework which appear in functional inequalities (that mainly deal with Wasserstein distances and their possible deviations). Indeed, through the Monge–Kantorovich formulation, the Wasserstein distance \mathbf{W}_1 involves Lipschitz functions, since it is precisely achieved taking the minimum over Lipschitz functions for any possible coupling with marginal corresponding to the arguments of the distance (see [4]).

In the literature, some non-asymptotic bounds can be found for the deviations of the Wasserstein distance between the empirical measure of a homogeneous Markov chain and its stationary distribution (see Boissard [9]). Here, we manage to get directly the non-asymptotic deviation bounds over all possible Lipschitz functions for the empirical measure of the scheme aiming directly to approximate the target stationary distribution of the diffusion. Handling the Wasserstein distance in our framework would amount to consider the supremum over the Lipschitz functions in the probability in (2.12). This will concern further research.

We eventually point out that Theorem 7 is obtained through regularization arguments of the source f exploiting the previous results of Theorems 2 and 4 (see Section 5.3 for details). This leads to a constraint on the steps, i.e. $\gamma_n \asymp n^{-\theta}$, $\theta \in (\frac{1}{2}, 1]$. This is the price to pay, indeed a bigger θ yields a lower convergence rate, to handle less regular Lipschitz sources. Also, to perform the approximation procedure, we precisely need a kind of elliptic bootstrap (like in Theorem 4 under (C_{UE})). This is why we impose the non-degeneracy assumptions.

3. Proof of the concentration results (Theorem 2)

For notational convenience, we say that assumption **(A)** holds whenever **(C1)**, **(GC)**, **(C2)**, **(L_v)**, **(U)** and **(S)** are fulfilled. We assume throughout this section that **(A)** is in force and that the function φ appearing in the lemmas satisfies the smoothness assumptions of Theorem 2.

3.1. Strategy

To control the deviations of $v_n(\mathcal{A}\varphi)$ we first give a decomposition lemma, obtained by a standard Taylor expansion. The idea is to perform a kind of *splitting* between the deterministic contributions in the transitions and the random innovations. Doing so, we manage to prove that the contributions involving the innovations can be gathered into conditionally Lipschitz continuous functions of the noise, with small Lipschitz constant (functions $(\psi_k(X_{k-1}, \cdot))_{k \in \llbracket 1, n \rrbracket}$ below). These functions precisely give the Gaussian concentration, see Lemma 2. The other terms, that we will call from now on “remainders”, will be shown to be uniformly controlled w.r.t. n and do not give any asymptotic contribution in the “fast decreasing” case $\theta > 1/(2 + \beta)$ (with the terminology of Theorem 2), see Lemmas 3, 4 and 5.

Lemma 1 (Local Decomposition of the empirical measure). *For all $n \geq 1$ and $k \in \llbracket 0, n - 1 \rrbracket$:*

$$\begin{aligned} \varphi(X_k) - \varphi(X_{k-1}) &= \gamma_k \mathcal{A}\varphi(X_{k-1}) + \left[\gamma_k \int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla \varphi(X_{k-1}), b_{k-1} \rangle dt \right. \\ &\quad \left. + \frac{1}{2} \gamma_k \text{Tr}((D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2 \varphi(X_{k-1})) \Sigma_{k-1}) + \psi_k(X_{k-1}, U_k) \right] \\ &=: \gamma_k \mathcal{A}(X_{k-1}) + (\psi_k(X_{k-1}, U_k) + R_{n,k}^1(X_{k-1})), \end{aligned} \quad (3.1)$$

where for each $k \in \llbracket 1, n \rrbracket$, conditionally to \mathcal{F}_{k-1} , the mapping $u \mapsto \psi_k(X_{k-1}, u)$ is Lipschitz continuous in u with constant $\sqrt{\gamma_k} \|\sigma_{k-1}\| \|\nabla \varphi\|_\infty$.

Introducing for a given k , the mapping $u \mapsto \Delta_k(X_{k-1}, u) := \psi_k(X_{k-1}, u) - \mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}]$, we then rewrite:

$$\varphi(X_k) - \varphi(X_{k-1}) = \gamma_k \mathcal{A}\varphi(X_{k-1}) + \Delta_k(X_{k-1}, U_k) + R_{n,k}(X_{k-1}),$$

with $R_{n,k}(X_{k-1}) := R_{n,k}^1(X_{k-1}) + \mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}]$. The contribution $\Delta_k(X_{k-1}, U_k)$ can be viewed as a martingale increment. Introduce now the associated (true) martingale

$$M_n := \sum_{k=1}^n \Delta_k(X_{k-1}, U_k). \quad (3.2)$$

Summing over k yields:

$$\varphi(X_n) - \varphi(X_0) = \Gamma_n v_n(\mathcal{A}\varphi) + M_n + \sum_{k=1}^n R_{n,k}(X_{k-1}). \quad (3.3)$$

Defining $R_n := \sum_{k=1}^n R_{n,k}(X_{k-1}) + \varphi(X_0) - \varphi(X_n)$ we obtain the following decomposition of the empirical measure:

$$v_n(\mathcal{A}\varphi) = -\frac{1}{\Gamma_n} (M_n + R_n). \quad (3.4)$$

– *Unbiased Case (Sub-Optimal Convergence Rate).* This case corresponds to *fast* decreasing steps of the form $\gamma_k \asymp k^{-\theta}$, $\theta > 1/(2 + \beta)$. To investigate the non-asymptotic deviations of the empirical measure, the idea is now to write for $a, \lambda > 0$:

$$\begin{aligned} \mathbb{P}[\sqrt{\Gamma_n} v_n(\mathcal{A}\varphi) \geq a] &\leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \mathbb{E}\left[\exp\left(-\frac{\lambda}{\Gamma_n} (M_n + R_n)\right)\right] \\ &\leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \mathbb{E}\left[\exp\left(-\frac{q\lambda}{\Gamma_n} M_n\right)\right]^{1/q} \mathbb{E}\left[\exp\left(\frac{p\lambda}{\Gamma_n} |R_n|\right)\right]^{1/p}, \quad \frac{1}{p} + \frac{1}{q} = 1, p, q > 1. \end{aligned} \quad (3.5)$$

We actually aim to choose $q := q(n) \rightarrow 1$. For a suitable choice of q satisfying the previous condition, we manage, in the fast decreasing case, to show that $\mathcal{R}_n := \mathbb{E}[\exp(\frac{p\lambda}{\Gamma_n}|R_n|)]^{1/p} \rightarrow 1$. For the term involving the martingale M_n we actually use the Gaussian concentration property (**GC**) of the innovation on its increments $(\Delta_k(X_{k-1}, U_k))_{k \in \llbracket 1, n \rrbracket}$. Namely, using the control of the Lipschitz constant of $\Delta_k(X_{k-1}, \cdot)$ stated in Lemma 1, we derive:

$$\begin{aligned} \mathbb{E}\left[\exp\left(-\frac{q\lambda}{\Gamma_n}M_n\right)\right] &= \mathbb{E}\left[\exp\left(-\frac{q\lambda}{\Gamma_n}M_{n-1}\right)\mathbb{E}\left[\exp\left(-\frac{q\lambda}{\Gamma_n}\Delta_{n-1}(X_{n-1}, U_n)\right)\middle|\mathcal{F}_{n-1}\right]\right] \\ &\leq \mathbb{E}\left[\exp\left(-\frac{q\lambda}{\Gamma_n}M_{n-1}\right)\right]\exp\left(\frac{\lambda^2q^2}{2\Gamma_n^2}\gamma_n\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2\right) \\ &\leq \exp\left(\frac{\lambda^2q^2}{2\Gamma_n}\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2\right), \end{aligned} \tag{3.6}$$

iterating the procedure to derive the last identity. From (3.5), we thus get:

$$\mathbb{P}[\sqrt{\Gamma_n}v_n(\mathcal{A}\varphi) \geq a] \leq \mathcal{R}_n \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}} + \frac{\lambda^2q}{2\Gamma_n}\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2\right).$$

Keeping in mind that we manage to find $q := q(n) \downarrow_n 1$ such that the remainder $\mathcal{R}_n \downarrow_n 1$, the result of Theorem 2 in the considered case then follows from a quadratic optimization over the parameter λ .

– *Biased Case (Optimal Convergence Rate)*. This case corresponds to *slow* decreasing steps of the form $\gamma_k \asymp k^{-\theta}$, $\theta = 1/(2 + \beta)$. In this setting, some terms of the remainder R_n in (3.4) give a non trivial asymptotic contribution. We choose to subtract them before studying the deviation (term $\mathcal{B}_{n,\beta}$ in (2.3)).

3.2. Explicit controls on the remainders

Summing the increments appearing in (3.1), we now choose for the analysis to write for a given $n \in \mathbb{N}$ the remainder R_n defined after (3.3) as

$$R_n = \sum_{k=1}^n R_{n,k}(X_{k-1}) + \varphi(X_0) - \varphi(X_n) = (D_{2,b,n} + D_{2,\Sigma,n}) + \bar{G}_n - L_n,$$

where:

$$\begin{aligned} D_{2,b,n} &:= \sum_{k=1}^n \gamma_k \int_0^1 \langle \nabla\varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla\varphi(X_{k-1}), b_{k-1} \rangle dt, \\ D_{2,\Sigma,n} &:= \frac{1}{2} \sum_{k=1}^n \gamma_k \text{Tr}((D^2\varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2\varphi(X_{k-1}))\Sigma_{k-1}), \\ \bar{G}_n &:= \sum_{k=1}^n \mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}], \\ L_n &:= \varphi(X_n) - \varphi(X_0). \end{aligned} \tag{3.7}$$

We refer to the proof of Lemma 1 to check that the above definition of \bar{G}_n actually matches the term $\sqrt{\Gamma_n}E_n^\beta$ introduced in equation (2.2) of Theorem 2. We rewrite from (3.4)

$$v_n(\mathcal{A}\varphi) = -\frac{1}{\Gamma_n}(M_n + R_n) = -\frac{1}{\Gamma_n}(M_n + (D_{2,b,n} + D_{2,\Sigma,n}) + \bar{G}_n - L_n). \tag{3.8}$$

We now split the analysis according to the cases (a) and (b) introduced in Theorem 2.

(a) $\theta \in (1/(2 + \beta), 1]$, $\beta \in (0, 1]$. From (3.8), the exponential Tchebychev and Hölder inequalities yield that, for all $\lambda \in \mathbb{R}_+$ and $p, q \in (1, +\infty)$, $\frac{1}{p} + \frac{1}{q} = 1$,

$$\begin{aligned} & \mathbb{P}[\sqrt{\Gamma_n} v_n(\mathcal{A}\varphi) \geq a] \\ & \leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \left(\mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n} M_n\right)\right)^{\frac{1}{q}} \\ & \quad \times \left(\mathbb{E} \exp\left(\frac{2p\lambda}{\Gamma_n} (|L_n| + |\bar{G}_n|)\right)\right)^{\frac{1}{2p}} \left(\mathbb{E} \exp\left(\frac{4p\lambda}{\Gamma_n} |D_{2,b,n}|\right)\right)^{\frac{1}{4p}} \left(\mathbb{E} \exp\left(\frac{4p\lambda}{\Gamma_n} |D_{2,\Sigma,n}|\right)\right)^{\frac{1}{4p}}. \end{aligned} \tag{3.9}$$

(b) $\theta = \frac{1}{2+\beta}$, $\beta \in (0, 1]$. If $\beta = 1$, denoting, $D_{2,n} := D_{2,b,n} + D_{2,\Sigma,n}$, we have from (3.7) and with the notations of (2.3), $(\bar{G}_n + D_{2,n}) = \sqrt{\Gamma_n} \mathcal{B}_{n,1}$. We study the deviations of:

$$\begin{aligned} \mathbb{P}[\sqrt{\Gamma_n} v_n(\mathcal{A}\varphi) + \mathcal{B}_{n,\beta} \geq a] &= \mathbb{P}\left[v_n(\mathcal{A}\varphi) + \frac{\bar{G}_n + D_{2,n}}{\Gamma_n} \geq \frac{a}{\sqrt{\Gamma_n}}\right] \\ &\leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \left(\mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n} M_n\right)\right)^{\frac{1}{q}} \left(\mathbb{E} \exp\left(\frac{p\lambda}{\Gamma_n} |L_n|\right)\right)^{\frac{1}{p}}. \end{aligned} \tag{3.10}$$

For $\beta \in (0, 1)$, the contributions of $D_{2,n}$ do not yield any asymptotic bias. Recalling from (2.3) that $\mathcal{B}_{n,\beta} = E_n^\beta = \frac{\bar{G}_n}{\sqrt{\Gamma_n}}$, we write:

$$\begin{aligned} \mathbb{P}[\sqrt{\Gamma_n} v_n(\mathcal{A}\varphi) + \mathcal{B}_{n,\beta} \geq a] &= \mathbb{P}\left[v_n(\mathcal{A}\varphi) + \frac{\bar{G}_n}{\Gamma_n} \geq \frac{a}{\sqrt{\Gamma_n}}\right] \\ &\leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \left(\mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n} M_n\right)\right)^{\frac{1}{q}} \left(\mathbb{E} \exp\left(\frac{2p\lambda}{\Gamma_n} |L_n|\right)\right)^{\frac{1}{2p}} \\ &\quad \times \left(\mathbb{E} \exp\left(\frac{4p\lambda}{\Gamma_n} |D_{2,b,n}|\right)\right)^{\frac{1}{4p}} \left(\mathbb{E} \exp\left(\frac{4p\lambda}{\Gamma_n} |D_{2,\Sigma,n}|\right)\right)^{\frac{1}{4p}}. \end{aligned} \tag{3.11}$$

Remark 8. Observe that in case (a), the “small steps” and the corresponding sufficient smoothness of φ prevent from the appearance of a bias. As a result, the concentration bound is, at the non-asymptotic level, the same as in Theorem 1, up to the additional upper-bound for the variance. In case (b), we subtract the terms $\mathcal{B}_{n,\beta}$ that asymptotically give a bias. When $\beta = 1$, this is the case for both terms $\frac{\bar{G}_n}{\Gamma_n}, \frac{D_{2,n}}{\Gamma_n}$. Also, for $D^3\varphi \in \mathcal{C}^1$, $\mathcal{B}_{n,1} = \frac{\bar{G}_n + D_{2,n}}{\sqrt{\Gamma_n}} \xrightarrow[n]{n} -\tilde{\gamma}m$ introduced in Theorem 1. For $\beta \in (0, 1)$ and $\varphi \in \mathcal{C}^3(\mathbb{R}^d, \mathbb{R})$, $[\varphi^{(3)}]_\beta < +\infty$, the only term giving a bias is $\mathcal{B}_{n,\beta} = E_n^\beta = \frac{\bar{G}_n}{\sqrt{\Gamma_n}}$.

The lemma below provides the Gaussian contribution to be exploited in inequalities (3.9)–(3.11).

Lemma 2 (Gaussian concentration). For $a > 0, q \in (1, +\infty)$, setting

$$\lambda_n := \frac{a}{q \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2} \sqrt{\Gamma_n}, \tag{3.12}$$

we derive:

$$\exp\left(-\lambda_n \frac{a}{\sqrt{\Gamma_n}}\right) \left(\mathbb{E} \exp\left(-\frac{q\lambda_n}{\Gamma_n} M_n\right)\right)^{\frac{1}{q}} \leq \exp\left(-\frac{a^2}{2q \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}\right).$$

Lemma 3 (Bounds for the Conditional Expectations). With the above notations, we have that for $\beta \in (0, 1], \theta \in [\frac{1}{2+\beta}, 1]$:

$$|E_n^\beta| = \frac{|\bar{G}_n|}{\sqrt{\Gamma_n}} \leq a_n := \frac{[\varphi^{(3)}]_\beta \|\sigma\|_\infty^{(3+\beta)} \mathbb{E}[|U_1|^{3+\beta}] \Gamma_n^{(\frac{3+\beta}{2})}}{(1 + \beta)(2 + \beta)(3 + \beta) \sqrt{\Gamma_n}}, \quad a.s.$$

Moreover, $a_n \rightarrow a_\infty$, with $a_\infty = 0$ if $\theta \in (\frac{1}{2+\beta}, 1]$ and $a_\infty > 0$ if $\theta = \frac{1}{2+\beta}$. Also, for $\beta \in (0, 1]$, $\theta \in (\frac{1}{2+\beta}, 1]$:

$$\left(\mathbb{E} \exp\left(\frac{2p\lambda_n}{\Gamma_n} |\bar{G}_n|\right) \right)^{\frac{1}{2p}} \leq \exp\left(\frac{\lambda_n}{\sqrt{\Gamma_n}} a_n\right) \leq \exp\left(\frac{\lambda_n^2}{2\Gamma_n p} + \frac{a_n^2 p}{2}\right), \quad \forall p > 1. \tag{3.13}$$

As indicated before, we now aim at controlling the remainders. In particular, from (3.5) and (3.7), we are led to handle terms of the form

$$\mathbb{E} \exp\left(c \sum_{k=1}^n \gamma_k^2 |b(X_{k-1})|^2\right) \stackrel{(\mathcal{L}_V)}{\leq} \mathbb{E} \exp\left(c C_V \sum_{k=1}^n \gamma_k^2 |V(X_{k-1})|\right)$$

for small enough real constants $c > 0$.

To this end, we will thoroughly rely on the following important integrability result for the Lyapunov function.

Proposition 1. Under (A) there is a constant $c_V := c_V((\mathbf{A})) > 0$ such that for all $\lambda \in [0, c_V]$, $\xi \in [0, 1]$:

$$I_V^\xi := \sup_{n \geq 0} \mathbb{E}[\exp(\lambda V_n^\xi)] < +\infty.$$

We now have the following results for the terms appearing in (3.7).

Lemma 4 (Initial term). Let $q \in (1, +\infty)$ be fixed and λ_n be as in (3.12) in Lemma 2. For functions φ satisfying (G_V), i.e. there exists $C_{V,\varphi} > 0$ such that for any $x \in \mathbb{R}^d$, $|\varphi(x)| \leq C_{V,\varphi}(1 + \sqrt{V(x)})$, for $p := \frac{q}{q-1}$ and $j \in \{1, 2\}$:

$$\left(\mathbb{E} \exp\left(j p \lambda_n \frac{|L_n|}{\Gamma_n}\right) \right)^{\frac{1}{j p}} \leq (I_V^1)^{\frac{1}{j p}} \exp\left(\frac{(j+1)p C_{V,\varphi}^2 \lambda_n^2}{c_V \Gamma_n^2} + \frac{c_V}{p}\right) = (I_V^1)^{\frac{1}{j p}} \exp\left(\frac{(j+1)p C_{V,\varphi}^2 a^2}{c_V q^2 \|\sigma\|_\infty^4 \|\nabla \varphi\|_\infty^4 \Gamma_n} + \frac{c_V}{p}\right),$$

with c_V, I_V^1 like in Proposition 1.

Lemma 5 (Remainders). Let $q \in (1, +\infty)$ be fixed and λ_n be as in Lemma 2. Then, there exists $C_{3.14} := C_{3.14}((\mathbf{A}), \varphi)$ such that for $p = \frac{q}{q-1}$:

$$\left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,\Sigma,n}|\right) \right)^{\frac{1}{4p}} \leq \exp\left(C_{3.14} \frac{p \lambda_n^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2}\right) (I_V^1)^{\frac{1}{4p}}. \tag{3.14}$$

We also have:

– If the mapping $x \mapsto \langle \nabla \varphi(x), b(x) \rangle$ is Lipschitz continuous, then there exists $C_{3.15} := C((\mathbf{A}), \varphi) > 0$ such that

$$\left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}|\right) \right)^{\frac{1}{4p}} \leq \exp\left(C_{3.15} \frac{p \lambda_n^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2}\right) (I_V^1)^{\frac{1}{4p}}. \tag{3.15}$$

– For $a \leq \frac{c_V \|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2}{4C_V \|D^2 \varphi\|_\infty^2} \frac{q}{p} \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$, there exists an \mathbb{R}^+ -valued sequence $(v_n)_{n \geq 1}$ such that $|v_n| \leq C_{3.16} := C_{3.16}((\mathbf{A}), \varphi)$ and

$$\left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}|\right) \right)^{\frac{1}{4p}} \leq (I_V^1)^{v_n}. \tag{3.16}$$

Also, $v_n \rightarrow v_\infty$ where $v_\infty = 0$ if $\theta > 1/3$ and $v_\infty > 0$ for $\theta = 1/3$.

Proof of Theorem 2. From Lemma 2 we get:

$$\left(\mathbb{E} \exp\left(-q \lambda_n \frac{M_n}{\Gamma_n}\right) \right)^{\frac{1}{q}} \exp\left(-\frac{a \lambda_n}{\sqrt{\Gamma_n}}\right) \leq \exp\left(-\frac{a^2}{2q \|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2}\right). \tag{3.17}$$

(a) We deal with the case $\theta \in (\frac{1}{2+\beta}, 1]$, $\beta \in (0, 1]$.

(i) We suppose that the mapping $x \mapsto \langle \nabla\varphi(x), b(x) \rangle$ is Lipschitz continuous. Plugging in (3.9) the controls from (3.17), Lemma 3 equation (3.13) ($|E_n^\beta| = \frac{|\tilde{G}_n|}{\sqrt{\Gamma_n}}$ is a.s. bounded), Lemma 4 (with $j = 2$) and Lemma 5 (equations (3.14), (3.15)), we get:

$$\begin{aligned} & \mathbb{P}\left[v_n(\mathcal{A}\varphi) \geq \frac{a}{\sqrt{\Gamma_n}}\right] \\ & \leq \exp\left(-\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\right) \exp\left(\frac{\lambda_n^2}{2\Gamma_n p} + \frac{pa_n^2}{2}\right) \exp\left(\frac{3pC_{V,\varphi}^2\lambda_n^2}{c_V\Gamma_n^2} + \frac{c_V}{p}\right) (I_V^1)^{\frac{1}{2p}} \\ & \quad \times \exp\left(C_{3.14}\frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{\Gamma_n^2}\right) (I_V^1)^{\frac{1}{4p}} \times \exp\left(C_{3.15}\frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{\Gamma_n^2}\right) (I_V^1)^{\frac{1}{4p}} \\ & \leq (I_V^1)^{\frac{1}{p}} \exp\left(-\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\left(1 - \frac{1}{q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\left\{\frac{p}{\Gamma_n}\left(\frac{6C_{V,\varphi}^2}{c_V} + 2[C_{3.14} + C_{3.15}](\Gamma_n^{(2)})^2\right) + \frac{1}{p}\right\}\right)\right) \\ & \quad \times \exp\left(\frac{c_V}{p} + \frac{pa_n^2}{2}\right). \end{aligned} \tag{3.18}$$

Recall now that for $\theta > \frac{1}{2+\beta} \geq 1/3$, $\Gamma_n^{(\frac{3+\beta}{2})}/\sqrt{\Gamma_n} \xrightarrow[n]{n} 0$, $\Gamma_n^{(2)}/\sqrt{\Gamma_n} \xrightarrow[n]{n} 0$ (see Lemma 3 and Remark 2). We now take $p := p_n \xrightarrow[n]{n} +\infty$, and therefore $q := q_n \xrightarrow[n]{n} 1$, such that $p_n^{1/2}\frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} \xrightarrow[n]{n} 0$ so that from Lemma 3, $p_n a_n^2 \xrightarrow[n]{n} 0$. Since $\frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} \geq \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}}$ this in turn implies:

$$d_n := \frac{1}{q_n\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2} \left\{ \frac{p_n}{\Gamma_n} \left(\frac{6C_{V,\varphi}^2}{c_V} + 2[C_{3.14} + C_{3.15}](\Gamma_n^{(2)})^2 \right) + \frac{1}{p_n} \right\} \xrightarrow[n]{n} 0. \tag{3.19}$$

We conclude from (3.18) setting $c_n = q_n^{-1}(1 - d_n)$, $C_n := (I_V^1)^{\frac{1}{p_n}} \exp\left(\frac{c_V}{p_n} + \frac{p_n a_n^2}{2}\right) \xrightarrow[n]{n} 1$. Observe that taking an increasing sequence $(p_n)_{n \geq 1}$ readily yields $C_n \downarrow_n 1$, and $q_n \downarrow_n 1$. Also, the sequence $(p_n)_{n \geq 1}$ can be chosen in order to have, for n large enough, $d_n \downarrow_n 0$ so that $c_n \uparrow_n 1$.

(ii) Assume $a \leq \frac{c_V q}{4C_{V,\varphi}} \frac{\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}{\|D^2\varphi\|_\infty^2} \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$. Plugging in (3.9) the controls from (3.17), Lemma 3, equation (3.13), Lemmas 4 (with $j = 2$), 5 (equations (3.14), (3.16)) we then derive:

$$\begin{aligned} & \mathbb{P}\left[v_n(\mathcal{A}\varphi) \geq \frac{a}{\sqrt{\Gamma_n}}\right] \\ & \leq \exp\left(-\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\right) \exp\left(\frac{\lambda_n^2}{2\Gamma_n p} + \frac{pa_n^2}{2}\right) \exp\left(\frac{3pC_{V,\varphi}^2\lambda_n^2}{c_V\Gamma_n^2} + \frac{c_V}{p}\right) (I_V^1)^{\frac{1}{2p}} \\ & \quad \times \exp\left(C_{3.14}\frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{\Gamma_n^2}\right) (I_V^1)^{\frac{1}{4p}} (I_V^1)^{v_n} \\ & \leq (I_V^1)^{v_n + \frac{3}{4p}} \exp\left(\frac{c_V}{p} + \frac{pa_n^2}{2}\right) \exp\left(-\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\left(1 - \frac{1}{q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\right.\right. \\ & \quad \left.\left. \times \left\{ \frac{p}{\Gamma_n} \left(\frac{6C_{V,\varphi}^2}{c_V} + 2C_{3.14}(\Gamma_n^{(2)})^2 \right) + \frac{1}{p} \right\} \right)\right), \end{aligned} \tag{3.20}$$

where the sequence $(v_n)_{n \geq 1}$ is defined in Lemma 5. Since $\theta > \frac{1}{2+\beta} \geq 1/3$ (see Remark 2), we again take $p := p_n \uparrow_n +\infty$ so that $p_n^{1/2} a_n \xrightarrow[n]{n} 0$ which also guarantees:

$$d_n := \frac{1}{q_n\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2} \left\{ p_n \left(\frac{6C_{V,\varphi}^2}{c_V\Gamma_n} + \frac{2C_{3.14}(\Gamma_n^{(2)})^2}{\Gamma_n} \right) + \frac{1}{p_n} \right\} \xrightarrow[n]{n} 0. \tag{3.21}$$

In this case, we derive the result by setting $c_n := q_n^{-1}(1 - d_n) \xrightarrow[n]{n} 1$, $C_n := (I_V^1)^{v_n + \frac{3}{4p_n}} \exp\left(\frac{c_V}{p_n} + \frac{p_n a_n^2}{2}\right) \xrightarrow[n]{n} 1$ (see the limits of v_n following equation (3.16) and (3.34)). Again, $(p_n)_{n \geq 1}$ can be chosen in order to have the stated monotonicity

for n large enough. Set now

$$\chi_n := \frac{c_V \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2 q_n}{4C_V \|D^2\varphi\|_\infty^2 p_n}, \tag{3.22}$$

so that $a \leq \chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$. Thus, the slower p_n goes to infinity, the wider the domain of validity for the estimate in the parameter a .

(b) It remains to analyze the case $\beta \in (0, 1]$, $\theta = \frac{1}{2+\beta}$. Let us deal with $\beta = 1$. From (3.10), the controls of (3.17) and Lemma 4 (with $j = 1$) we get:

$$\mathbb{P}\left[v_n(\mathcal{A}\varphi) + \frac{\bar{G}_n + D_{2,n}}{\Gamma_n} \geq \frac{a}{\sqrt{\Gamma_n}} \right] \leq \exp\left(-\frac{a^2}{2q \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}\right) \exp\left(\frac{2pC_{V,\varphi}^2 \lambda_n^2}{c_V \Gamma_n^2} + \frac{c_V}{p}\right) (I_V^1)^{\frac{1}{p}}.$$

Recalling the definition of λ_n in (3.12), we conclude as previously with obvious modifications of $(c_n)_{n \geq 1}$, $(C_n)_{n \geq 1}$. The case $\beta \in (0, 1)$ is handled similarly starting from (3.11).

Also, when $D^3\varphi \in \mathcal{C}^1$, we derive similarly to the proof of Theorem 10 in [25] that $\mathcal{B}_{n,1} \xrightarrow[n]{} -\tilde{\gamma}m$.

Eventually, the final control involving the two sided deviation is derived by symmetry. □

3.3. Proof of the technical lemmas

This section is devoted to the proof of the previously used Lemmas 1–5 and Proposition 1 which were the key ingredients to derive Theorem 2.

Proof of Lemma 1. For $k \in \llbracket 1, n \rrbracket$, we first write:

$$\begin{aligned} \varphi(X_k) - \varphi(X_{k-1}) &= (\varphi(X_k) - \varphi(X_{k-1} + \gamma_k b_{k-1})) + (\varphi(X_{k-1} + \gamma_k b_{k-1}) - \varphi(X_{k-1})) \\ &=: T_{k-1,r}(\varphi) + T_{k-1,d}(\varphi), \end{aligned} \tag{3.23}$$

in order to split the random and deterministic contributions in the transitions of the scheme (S).

We then perform a Taylor expansion with integral remainder at order 2 for the function φ in the two terms of the r.h.s. of (3.23). Namely, with the above notations:

$$\begin{aligned} T_{k-1,d}(\varphi) &= \gamma_k b_{k-1} \cdot \nabla\varphi(X_{k-1}) + \gamma_k \int_0^1 \langle \nabla\varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla\varphi(X_{k-1}), b_{k-1} \rangle dt, \\ T_{k-1,r}(\varphi) &= \sqrt{\gamma_k} \sigma_{k-1} U_k \cdot \nabla\varphi(X_{k-1} + \gamma_k b_{k-1}) \\ &\quad + \gamma_k \int_0^1 (1-t) \text{Tr}(D^2\varphi(X_{k-1} + \gamma_k b_{k-1} + t\sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^*) dt. \end{aligned}$$

Hence,

$$\begin{aligned} &\varphi(X_k) - \varphi(X_{k-1}) \\ &= \gamma_k \mathcal{A}\varphi(X_{k-1}) \\ &\quad + \gamma_k \int_0^1 \langle \nabla\varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla\varphi(X_{k-1}), b_{k-1} \rangle dt + \sqrt{\gamma_k} \sigma_{k-1} U_k \cdot \nabla\varphi(X_{k-1} + \gamma_k b_{k-1}) \\ &\quad + \gamma_k \int_0^1 (1-t) \text{Tr}(D^2\varphi(X_{k-1} + \gamma_k b_{k-1} + t\sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* - D^2\varphi(X_{k-1}) \Sigma_{k-1}) dt \\ &= \gamma_k \mathcal{A}\varphi(X_{k-1}) + \gamma_k \int_0^1 \langle \nabla\varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla\varphi(X_{k-1}), b_{k-1} \rangle dt \\ &\quad + \gamma_k \int_0^1 (1-t) \text{Tr}((D^2\varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2\varphi(X_{k-1})) \Sigma_{k-1}) dt + \psi_k(X_{k-1}, U_k) \\ &=: \gamma_k \mathcal{A}\varphi(X_{k-1}) + D_{2,b}^k + D_{2,\Sigma}^k + \psi_k(X_{k-1}, U_k), \end{aligned} \tag{3.24}$$

where

$$\begin{aligned} \psi_k(X_{k-1}, U_k) &= \sqrt{\gamma_k} \sigma_{k-1} U_k \cdot \nabla \varphi(X_{k-1} + \gamma_k b_{k-1}) \\ &\quad + \gamma_k \int_0^1 (1-t) \operatorname{Tr}(D^2 \varphi(X_{k-1} + \gamma_k b_{k-1} + t \sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^*) \\ &\quad - D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \Sigma_{k-1}) dt. \end{aligned} \quad (3.25)$$

Observe now that, conditionally to \mathcal{F}_{k-1} , the mapping $u \mapsto \psi_k(X_{k-1}, u)$ is Lipschitz continuous: indeed, the innovation U_k does not appear in the other contributions of the right side of (3.24). Consequently, as φ is Lipschitz continuous we derive, for any $(u, u') \in (\mathbb{R}^d)^2$:

$$|\psi_k(X_{k-1}, u) - \psi_k(X_{k-1}, u')| \leq \sqrt{\gamma_k} \|\sigma_{k-1}\| \|\nabla \varphi\|_\infty |u - u'|.$$

The result is obtained by summing up the previous identities from $k = 1$ to n , observing, with the notations of (3.7), that $L_n = \sum_{k=1}^n \varphi(X_k) - \varphi(X_{k-1})$, $D_{2,b,n} = \sum_{k=1}^n D_{2,b}^k$, $D_{2,\Sigma,n} = \sum_{k=1}^n D_{2,\Sigma}^k$, $G_n := \sum_{k=1}^n \psi_k(X_{k-1}, U_k)$. \square

Proof of Lemma 2. The idea is to use conditionally and iteratively the Gaussian concentration property (**GC**) of the innovation. Let us note that this strategy was already the key ingredient in [17]. In the current framework, we exploit that the functions $u \mapsto \Delta_k(X_{k-1}, u) := \psi_k(X_{k-1}, u) - \mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}]$ are conditionally independent w.r.t. \mathcal{F}_{k-1} and Lipschitz continuous with constant $\sqrt{\gamma_k} \|\sigma\|_\infty \|\nabla \varphi\|_\infty$ by Lemma 1. We thus write:

$$\begin{aligned} \mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n} M_n\right) &= \mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n} \sum_{k=1}^n \Delta_k(X_{k-1}, U_k)\right) \\ &= \mathbb{E} \left[\exp\left(-\frac{q\lambda}{\Gamma_n} \sum_{k=1}^{n-1} \Delta_k(X_{k-1}, U_k)\right) \mathbb{E} \left[\exp\left(-\frac{q\lambda}{\Gamma_n} \Delta_n(X_{n-1}, U_n)\right) \middle| \mathcal{F}_{n-1} \right] \right] \\ &\leq \mathbb{E} \left[\exp\left(-\frac{q\lambda}{\Gamma_n} \sum_{k=1}^{n-1} \Delta_k(X_{k-1}, U_k)\right) \exp\left(\frac{q^2 \lambda^2}{2\Gamma_n^2} \gamma_n \|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2\right) \right], \end{aligned} \quad (3.26)$$

where we used (**GC**) in the third line recalling as well that $\mathbb{E}[\Delta_n(X_{n-1}, U_n) | \mathcal{F}_{n-1}] = 0$.

Iterating the process over k , we obtain:

$$\left(\mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n} M_n\right)\right)^{\frac{1}{q}} = \left(\mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n} \sum_{k=1}^n \Delta_k(X_{k-1}, U_k)\right)\right)^{\frac{1}{q}} \leq \exp\left(\frac{q\lambda^2 \|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2}{2\Gamma_n}\right). \quad (3.27)$$

Finally,

$$\exp\left(-\frac{\lambda a}{\sqrt{\Gamma_n}}\right) \left(\mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n} M_n\right)\right)^{\frac{1}{q}} \leq \exp\left(\frac{g(\lambda)}{\sqrt{\Gamma_n}}\right),$$

where $g : \mathbb{R}^+ \rightarrow \mathbb{R}$ is defined by $g(\lambda) = -\frac{a}{\sqrt{\Gamma_n}} \lambda + \frac{q\lambda^2}{2\Gamma_n} \|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2$. As $a > 0$, the function attains its minimum at λ_n given in (3.12). This eventually yields the expected bound. \square

Proof of Lemma 3. From the definition in (3.25) and the Fubini theorem, we have that for each $k \in \llbracket 1, n \rrbracket$:

$$\begin{aligned} \mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}] &= \gamma_k \int_0^1 (1-t) \operatorname{Tr}(\mathbb{E}[D^2 \varphi(X_{k-1} + \gamma_k b_{k-1} + t \sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* \\ &\quad - D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \Sigma_{k-1} | \mathcal{F}_{k-1}]) dt. \end{aligned} \quad (3.28)$$

Recalling that U_k has the same moments as the standard Gaussian random variable up to order three (see (**GC**)) and is independent of \mathcal{F}_{k-1} , a Taylor expansion yields:

$$\begin{aligned} &\mathbb{E}[\operatorname{Tr}(D^2 \varphi(X_{k-1} + \gamma_k b_{k-1} + t \sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* - D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \Sigma_{k-1}) | \mathcal{F}_{k-1}] \\ &= \operatorname{Tr}(D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \sigma_{k-1} \mathbb{E}[U_k \otimes U_k] \sigma_{k-1}^*) \end{aligned}$$

$$\begin{aligned}
 & + \int_0^1 \mathbb{E} \left[\text{Tr} \left((D^3 \varphi(X_{k-1} + \gamma_k b_{k-1} + ut \sqrt{\gamma_k} \sigma_{k-1} U_k) t \sqrt{\gamma_k} \sigma_{k-1} U_k) (\sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^*) \right) \middle| \mathcal{F}_{k-1} \right] du \\
 & - \text{Tr} (D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \Sigma_{k-1}) \\
 = & \text{Tr} (D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \sigma_{k-1} \underbrace{(\mathbb{E}[U_k \otimes U_k] - I)}_{=0} \sigma_{k-1}^*) \\
 & + t \sqrt{\gamma_k} \int_0^1 \mathbb{E} \left[\text{Tr} \left(([D^3 \varphi(X_{k-1} + \gamma_k b_{k-1} + ut \sqrt{\gamma_k} \sigma_{k-1} U_k) - D^3 \varphi(X_{k-1} + \gamma_k b_{k-1})] \sigma_{k-1} U_k) \right. \right. \\
 & \left. \left. \times (\sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^*) \right) \middle| \mathcal{F}_{k-1} \right] du,
 \end{aligned}$$

recalling from **(GC)** that for each $(i, j, l) \in \llbracket 1, r \rrbracket^3$, $\mathbb{E}[U_k^i U_k^j U_k^l | \mathcal{F}_{k-1}] = \mathbb{E}[U_1^i U_1^j U_1^l] = 0$ (cancellation argument). Hence,

$$\begin{aligned}
 |\mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}]| & \leq \gamma_k \int_0^1 (1-t) t^{1+\beta} [\varphi^{(3)}]_\beta \mathbb{E} \left[\gamma_k^{\frac{1+\beta}{2}} \|\sigma_{k-1}\|^{3+\beta} |U_k|^{3+\beta} \int_0^1 u^\beta du \middle| \mathcal{F}_{k-1} \right] dt \\
 & = \frac{[\varphi^{(3)}]_\beta \gamma_k^{\frac{3+\beta}{2}} \|\sigma_{k-1}\|^{3+\beta} \mathbb{E}[|U_k|^{3+\beta}]}{(1+\beta)(2+\beta)(3+\beta)},
 \end{aligned}$$

recalling that the third derivatives of φ are β -Hölder continuous for the first inequality. We thus derive:

$$|E_n^\beta| = \frac{|\bar{G}_n|}{\sqrt{\Gamma_n}} \leq \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^{n-1} |\mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}]| \leq \frac{[\varphi^{(3)}]_\beta \|\sigma\|_\infty^{3+\beta} \mathbb{E}[|U_1|^{3+\beta}] \Gamma_n^{\frac{3+\beta}{2}}}{(1+\beta)(2+\beta)(3+\beta) \sqrt{\Gamma_n}} =: a_n.$$

Proof of Proposition 1. First of all, let us decompose the Lyapunov function V with a Taylor expansion like in Lemma 1. We again use a splitting between the deterministic contributions and those involving the innovation. We write for each $n \in \mathbb{N}$:

$$\begin{aligned}
 & V(X_n) - V(X_{n-1}) \\
 = & \gamma_n \mathcal{A}V(X_{n-1}) + \gamma_n^2 \int_0^1 (1-t) \text{Tr}(D^2 V(X_{n-1} + t \gamma_n b_{n-1}) b_{n-1} \otimes b_{n-1}) dt \\
 & - \frac{\gamma_n}{2} \text{Tr}(D^2 V(X_{n-1}) \Sigma_{n-1}) + \sqrt{\gamma_n} \sigma_{n-1} U_n \cdot \nabla V(X_{n-1} + \gamma_n b_{n-1}) \\
 & + \gamma_n \int_0^1 (1-t) \text{Tr}(D^2 V(X_{n-1} + \gamma_n b_{n-1} + t \sqrt{\gamma_n} \sigma_{n-1} U_n) \sigma_{n-1} U_n \otimes U_n \sigma_{n-1}^*) dt \\
 \leq & -\gamma_n \alpha_V V(X_{n-1}) + \gamma_n \beta_V + C_V \frac{\gamma_n^2}{2} \|D^2 V\|_\infty V(X_{n-1}) \\
 & + \frac{\gamma_n}{2} \|D^2 V\|_\infty \|\sigma\|_\infty^2 + \sqrt{\gamma_n} \sigma_{n-1} U_n \cdot \nabla V(X_{n-1} + \gamma_n b_{n-1}) + \frac{\gamma_n}{2} \|D^2 V\|_\infty \|\sigma\|_\infty^2 |U_n|^2 \\
 \leq & \gamma_n \left(-\frac{\alpha_V}{2} V(X_{n-1}) + \tilde{c} \right) + \sqrt{\gamma_n} \sigma_{n-1} U_n \cdot \nabla V(X_{n-1} + \gamma_n b_{n-1}) + \frac{\gamma_n}{2} \|D^2 V\|_\infty \|\sigma\|_\infty^2 |U_n|^2 \tag{3.29}
 \end{aligned}$$

for a constant $\tilde{c} := \tilde{c}(V, \sigma, \beta_V)$. We have in fact considered the time steps sufficiently small, indeed in **(S)**, we have chosen for each $n \in \mathbb{N}$, $\gamma_n < \min(\frac{1}{2\sqrt{C_V \tilde{c}}}, \frac{\alpha_V}{2C_V \|D^2 V\|_\infty})$. The two terms involving the innovation U_n in the above decomposition can be controlled thanks to the Gaussian concentration hypothesis **(GC)**. Let us define for all $x \in \mathbb{R}^d$ and $\gamma, \lambda > 0$ the quantities:

$$I_1(\gamma, \lambda, x) := \mathbb{E} \left[\exp \left(\lambda \sqrt{\gamma} \sigma(x) U_1 \cdot \nabla V(x + \gamma b(x)) \right) \right], \quad I_2(\gamma, \lambda) := \mathbb{E} \left[\exp \left(\lambda \frac{\gamma}{2} \|D^2 V\|_\infty \|\sigma\|_\infty^2 |U_1|^2 \right) \right].$$

The first one is directly controlled owing to hypothesis **(GC)**:

$$I_1(\gamma_n, \lambda, x) \leq \exp\left(\frac{\lambda^2 \gamma_n |\sigma^*(x) \nabla V(x + \gamma_n b(x))|^2}{2}\right) \underset{(\mathcal{L}_V)}{\leq} \exp\left(\frac{\lambda^2 \gamma_n C_V \|\sigma\|_\infty^2 V(x + \gamma_n b(x))}{2}\right). \tag{3.30}$$

Furthermore, under **(GC)**, for any $c < \frac{1}{2}$, $I_c := \mathbb{E}[\exp(c|U_n|^2)] < +\infty$. Hence, for any $\lambda < \frac{2c}{\|D^2V\|_\infty \|\sigma\|_\infty^2 \gamma_1}$, Jensen's inequality yields:

$$I_2(\gamma_n, \lambda) \leq \left[\mathbb{E} \exp(c|U_n|^2)\right]^{\frac{\lambda \gamma_n \|D^2V\|_\infty \|\sigma\|_\infty^2}{2c}} = \exp\left(\gamma_n \ln(I_c) \frac{\lambda \|D^2V\|_\infty \|\sigma\|_\infty^2}{2c}\right). \tag{3.31}$$

These controls allow to prove the integrability statement of the proposition by induction. For $n = 0$, recalling from assumption **(C1)** that for any $\lambda < \lambda_0$, $\mathbb{E} \exp(\lambda|X_0|^2) < +\infty$ and from **(L_V)**(i) that $V(x) \leq \bar{c}|x|^2$ outside of a compact set, we derive that for all $\lambda \in (0, \frac{\lambda_0}{\bar{c}})$, there exists $C_{V,\lambda}^0 \in (1, +\infty)$ such that

$$\mathbb{E} \exp(\lambda V(X_0)) \leq C_{V,\lambda}^0.$$

Set now $\tilde{\beta}_V := \tilde{c} + \ln(I_c) \frac{\|D^2V\|_\infty \|\sigma\|_\infty^2}{2c}$ and $\tilde{\alpha}_V := \min(\frac{1}{\gamma_1}, \frac{\alpha_V}{2} - \lambda C_V \|\sigma\|_\infty^2 (1 + \gamma_1 C_V [1 + \frac{\gamma_1 \|D^2V\|_\infty}{2}])) \in (0, \frac{1}{\gamma_1}]$, for $\lambda < \frac{2C_V \|\sigma\|_\infty^2 (1 + \gamma_1 C_V [1 + \frac{\gamma_1 \|D^2V\|_\infty}{2}])}{\alpha_V}$.

Let us assume that for any $\lambda < \lambda_V := \min(\frac{\lambda_0}{2\bar{c}}, \frac{\alpha_V}{2C_V \|\sigma\|_\infty^2 (1 + \gamma_1 C_V [1 + \frac{\gamma_1 \|D^2V\|_\infty}{2}])}, \frac{c}{\|D^2V\|_\infty \|\sigma\|_\infty^2 \gamma_1})$, the property

$$\forall k \in \llbracket 0, n-1 \rrbracket, \quad \mathbb{E} \exp(\lambda V(X_k)) \leq C_{V,\lambda}^0 \vee \exp\left(\frac{\lambda \tilde{\beta}_V}{\tilde{\alpha}_V}\right), \tag{P_{n-1}}$$

holds for a fixed $n - 1 \in \mathbb{N}_0$ and let us prove (P_n) . By inequalities (3.29), (3.30) and (3.31) and the Cauchy–Schwarz inequality, we derive that for any $\lambda < \lambda_V$,

$$\begin{aligned} \mathbb{E} \exp(\lambda V(X_n)) &= \mathbb{E}[\exp(\lambda V(X_{n-1})) \mathbb{E}[\exp(\lambda(V(X_n) - V(X_{n-1}))) | \mathcal{F}_{n-1}]] \\ &\leq \mathbb{E}\left[\exp\left(\lambda \left[V(X_{n-1}) \left(1 - \frac{\alpha_V}{2} \gamma_n\right) + \tilde{c} \gamma_n\right]\right) I_1(\gamma_n, 2\lambda, X_{n-1})^{1/2} I_2(\gamma_n, 2\lambda)^{1/2}\right] \\ &= \exp(\lambda \gamma_n \tilde{\beta}_V) \mathbb{E}\left[\exp\left(\lambda \left(1 - \frac{\alpha_V}{2} \gamma_n\right) V(X_{n-1}) + \lambda^2 \gamma_n C_V \|\sigma\|_\infty^2 V(X_{n-1} + \gamma_n b_{n-1})\right)\right]. \end{aligned}$$

Recall now that $V(X_{n-1} + \gamma_n b_{n-1}) \leq V(X_{n-1}) + \gamma_n |\nabla V(X_{n-1})| |b_{n-1}| + \frac{\gamma_n^2}{2} \|D^2V\|_\infty |b_{n-1}|^2 \stackrel{(\mathcal{L}_V)(ii)}{\leq} V(X_{n-1}) (1 + \gamma_n C_V [1 + \frac{\gamma_n \|D^2V\|_\infty}{2}])$. Thus,

$$\begin{aligned} \mathbb{E}[\exp(\lambda V(X_n))] &\leq \exp(\lambda \gamma_n \tilde{\beta}_V) \mathbb{E}[\underbrace{\exp(\lambda (1 - \gamma_n \tilde{\alpha}_V) V(X_{n-1}))}_{\in [0,1]}] \\ &\stackrel{(\text{Jensen})}{\leq} \exp(\lambda \gamma_n \tilde{\beta}_V) \mathbb{E}[\exp(\lambda V(X_{n-1}))]^{(1 - \gamma_n \tilde{\alpha}_V)} \leq \exp(\lambda \gamma_n \tilde{\beta}_V) C_{V,\lambda}^{(1 - \gamma_n \tilde{\alpha}_V)} \end{aligned}$$

using (P_{n-1}) for the last inequality. From the above equation and the previous definition of $C_{V,\lambda}$ we have:

$$\exp(\lambda \gamma_n \tilde{\beta}_V) C_{V,\lambda}^{(1 - \gamma_n \tilde{\alpha}_V)} \leq C_{V,\lambda} \iff C_{V,\lambda} \geq \exp\left(\frac{\lambda \tilde{\beta}_V}{\tilde{\alpha}_V}\right).$$

Hence, (P_n) holds. Taking $c_V < \lambda_V$ completes the proof. □

Remark 9. Noting that $v^* := \inf_{x \in \mathbb{R}^d} V(x) > 0$, we get that for all $(n, \xi) \in \mathbb{N} \times [0, 1]$, and $\lambda < \lambda_V (v^*)^{1-\xi}$:

$$\mathbb{E} \exp(\lambda V_n^\xi) = \mathbb{E} \exp\left(\lambda (v^*)^\xi \underbrace{\left(\frac{V_n}{v^*}\right)^\xi}_{\geq 1}\right) \leq \mathbb{E} \exp(\lambda (v^*)^{\xi-1} V_n) \leq C_{V,\lambda (v^*)^{\xi-1}} < +\infty.$$

Thus, we readily get as a by-product of Proposition 1 that, for all $\xi \in [0, 1]$, $\lambda < \lambda_V(v^*)^{1-\xi}$, $\sup_{n \in \mathbb{N}} \mathbb{E} \exp(\lambda V_n^\xi) < +\infty$. We refer to Lemaire (see e.g. Theorem 17 in [28]) for additional results in that direction.

Proof of Lemma 4. Recalling from (\mathbf{G}_V) that there exists $C_{V,\varphi} > 0$ such that for any $x \in \mathbb{R}^d$, $|\varphi(x)| \leq C_{V,\varphi}(1 + \sqrt{V(x)})$, we get for $j \in \{1, 2\}$:

$$\begin{aligned} & \left[\mathbb{E} \exp\left(jp\lambda_n \frac{|\varphi(X_0) - \varphi(X_n)|}{\Gamma_n}\right) \right]^{\frac{1}{jp}} \\ & \leq \left[\mathbb{E} \exp\left(jp\lambda_n \frac{C_{V,\varphi}(2 + \sqrt{V(X_0)} + \sqrt{V(X_n)})}{\Gamma_n}\right) \right]^{\frac{1}{jp}} \\ & \leq \exp\left(2C_{V,\varphi} \frac{\lambda_n}{\Gamma_n}\right) \left[\mathbb{E} \exp\left(2jpC_{V,\varphi}\lambda_n \frac{\sqrt{V(X_0)}}{\Gamma_n}\right) \right]^{\frac{1}{2jp}} \left[\mathbb{E} \exp\left(2jpC_{V,\varphi}\lambda_n \frac{\sqrt{V(X_n)}}{\Gamma_n}\right) \right]^{\frac{1}{2jp}}. \end{aligned}$$

Write now for $i \in \{0, n\}$ by the Young inequality:

$$2jpC_{V,\varphi}\lambda_n \frac{\sqrt{V(X_i)}}{\Gamma_n} \leq c_V V(X_i) + \frac{(jp)^2 C_{V,\varphi}^2 \lambda_n^2}{c_V \Gamma_n^2},$$

where c_V is the positive real constant such that $I_V^1 = \sup_{n \geq 0} \mathbb{E}[\exp(c_V V(X_n))] < +\infty$ (see Proposition 1). We then get

$$\begin{aligned} & \left[\mathbb{E} \exp\left(jp\lambda_n \frac{|\varphi(X_0) - \varphi(X_n)|}{\Gamma_n}\right) \right]^{\frac{1}{jp}} \\ & \leq \exp\left(2C_{V,\varphi} \frac{\lambda_n}{\Gamma_n}\right) \exp\left(\frac{jpC_{V,\varphi}^2 \lambda_n^2}{c_V \Gamma_n^2}\right) (\mathbb{E} \exp(c_V V(X_0)))^{\frac{1}{2jp}} (\mathbb{E} \exp(c_V V(X_n)))^{\frac{1}{2jp}} \\ & \leq \exp\left(\frac{(j+1)pC_{V,\varphi}^2 \lambda_n^2}{c_V \Gamma_n^2}\right) \exp\left(\frac{c_V}{p}\right) (I_V^1)^{\frac{1}{jp}}. \end{aligned}$$

□

Proof of Lemma 5.

• *Proof of inequalities (3.15) and (3.16).*

– If $x \mapsto \langle \nabla \varphi(x), b(x) \rangle$ is Lipschitz continuous. We first rewrite from the definition of $D_{2,b,n}$ in (3.7):

$$\begin{aligned} D_{2,b,n} &= \sum_{k=1}^n \gamma_k \int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla \varphi(X_{k-1}), b_{k-1} \rangle dt \\ &= \sum_{k=1}^n \gamma_k \left[\int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}), b_{k-1} - b(X_{k-1} + t\gamma_k b_{k-1}) \rangle dt \right. \\ & \quad \left. + \int_0^1 (\langle \nabla \varphi, b \rangle(X_{k-1} + t\gamma_k b_{k-1}) - \langle \nabla \varphi, b \rangle(X_{k-1})) dt \right]. \end{aligned}$$

From the boundedness of $\nabla \varphi$, and the Lipschitz property of the mappings $x \mapsto b(x)$ (which has been assumed from the very beginning) and $x \mapsto \langle \nabla \varphi(x), b(x) \rangle$ (assumed for the current inequality), recalling that $b_{k-1} = b(X_{k-1})$, one derives that:

$$|D_{2,b,n}| \leq \sum_{k=1}^n \gamma_k^2 (\|\nabla \varphi\|_\infty [b]_1 + [\langle \nabla \varphi, b \rangle]_1) \frac{|b_{k-1}|}{2} \leq C \sum_{k=1}^n \gamma_k^2 |b_{k-1}|, \quad C := C(b, \varphi). \tag{3.32}$$

From this inequality, assumption (\mathcal{L}_V) (ii) and the Jensen inequality (applied to the exponential function for the measure $\frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 \delta_k$), we derive:

$$\left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}|\right) \right)^{\frac{1}{4p}} \leq \left(\frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 \mathbb{E} \exp\left(\frac{4p\lambda_n \Gamma_n^{(2)}}{\Gamma_n} C \sqrt{C_V} \sqrt{V_{k-1}}\right) \right)^{\frac{1}{4p}}.$$

From the Young inequality we obtain:

$$\mathbb{E} \exp\left(\frac{4p\lambda_n \Gamma_n^{(2)}}{\Gamma_n} C \sqrt{C_V} \sqrt{V_{k-1}}\right) \leq \exp\left(\left(\frac{2\sqrt{2}p\lambda_n \Gamma_n^{(2)}}{\Gamma_n} \frac{C \sqrt{C_V}}{\sqrt{c_V}}\right)^2\right) \mathbb{E}[\exp(c_V V_{k-1})].$$

We finally derive with the notations of Proposition 1:

$$\left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}| \right)\right)^{\frac{1}{4p}} \leq \exp\left(\frac{2p\lambda_n^2 (\Gamma_n^{(2)})^2 (C \sqrt{C_V})^2}{\Gamma_n^2 c_V}\right) (I_V^1)^{\frac{1}{4p}} \leq \exp\left(C_{3.15} \frac{p\lambda_n^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2}\right) (I_V^1)^{\frac{1}{4p}},$$

setting $C_{3.15} := 2 \frac{(C \sqrt{C_V})^2}{c_V}$ with $C = \frac{1}{2}(\|\nabla\varphi\|_\infty [b]_1 + [(\nabla\varphi, b)]_1)$ as in (3.32).

– If $a \leq \frac{c_V \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2 q_n \sqrt{\Gamma_n}}{4C_V \|D^2\varphi\|_\infty^2 p_n \Gamma_n^{(2)}} = \chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$ with the notation introduced in (3.22). Write first from (3.7) (definition of $D_{2,b,n}$), using a Taylor expansion on $\nabla\varphi$:

$$\begin{aligned} & \left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}| \right)\right)^{\frac{1}{4p}} \\ & \leq \left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} \sum_{k=1}^n \gamma_k^2 \int_0^1 (1-t) |\text{Tr}(D^2\varphi(X_{k-1} + t\gamma_k b_{k-1}) b_{k-1} \otimes b_{k-1})| dt\right)\right)^{\frac{1}{4p}}. \end{aligned} \tag{3.33}$$

We first easily get from the assumptions on φ and point (ii) of (\mathcal{L}_V) that:

$$\left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}| \right)\right)^{\frac{1}{4p}} \leq \left(\mathbb{E} \exp\left(\frac{2p\lambda_n}{\Gamma_n} \sum_{k=1}^n \gamma_k^2 C_V V_{k-1} \|D^2\varphi\|_\infty\right)\right)^{\frac{1}{4p}}.$$

From the Jensen inequality, we derive:

$$\left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}| \right)\right)^{\frac{1}{4p}} \leq \left(\frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 \mathbb{E} \exp\left(\frac{2p\lambda_n \Gamma_n^{(2)}}{\Gamma_n} \|D^2\varphi\|_\infty C_V V_{k-1}\right)\right)^{\frac{1}{4p}}.$$

We then have from the definition of λ_n in (3.12) that:

$$\bar{v}_n := \frac{2p\lambda_n \Gamma_n^{(2)}}{\Gamma_n} \|D^2\varphi\|_\infty \frac{C_V}{c_V} = \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \frac{2C_V p}{c_V q} \frac{\|D^2\varphi\|_\infty}{\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2} a \leq 1.$$

The Jensen inequality for concave functions yields for each $k \in [1, n]$:

$$\mathbb{E} \exp\left(\frac{2p\lambda_n \Gamma_n^{(2)}}{\Gamma_n} \|D^2\varphi\|_\infty C_V V_{k-1}\right) = \mathbb{E} \exp(\bar{v}_n c_V V_{k-1}) \leq (\mathbb{E} \exp(c_V V_{k-1}))^{\bar{v}_n}.$$

Thus, setting

$$v_n := \frac{\bar{v}_n}{4p} = \frac{\lambda_n \Gamma_n^{(2)}}{2\Gamma_n} \|D^2\varphi\|_\infty \frac{C_V}{c_V}, \tag{3.34}$$

we finally derive,

$$\left[\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}| \right)\right]^{\frac{1}{4p}} \leq \left[\frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 \left(\sup_{l \geq 1} \mathbb{E}[\exp(c_V V_{l-1})]\right)^{\bar{v}_n}\right]^{\frac{1}{4p}} = (I_V^1)^{v_n} =: C_n,$$

using again the notations of Proposition 1. This gives (3.16). □

• *Proof of inequality (3.14).* We proceed as for the proof of (3.16) and (3.15). Write:

$$\left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,\Sigma,n}| \right)\right)^{\frac{1}{4p}}$$

$$\begin{aligned} &\leq \left(\mathbb{E} \exp \left(\frac{4p\lambda_n}{\Gamma_n} \sum_{k=1}^n \frac{\gamma_k}{2} \left| \text{Tr} \left((D^2\varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2\varphi(X_{k-1})) \Sigma_{k-1} \right) \right| \right) \right)^{\frac{1}{4p}} \\ &\leq \left(\mathbb{E} \exp \left(\frac{2p\lambda_n}{\Gamma_n} \|\sigma\|_\infty^2 [\varphi^{(2)}]_1 \sum_{k=1}^n \gamma_k^2 |b_{k-1}| \right) \right)^{\frac{1}{4p}} \leq \left(\mathbb{E} \exp \left(\frac{2p\lambda_n}{\Gamma_n} \|\sigma\|_\infty^2 [\varphi^{(2)}]_1 C_V^{\frac{1}{2}} \sum_{k=1}^n \gamma_k^2 |V_{k-1}|^{\frac{1}{2}} \right) \right)^{\frac{1}{4p}} \\ &\leq \left(\frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 \mathbb{E} \exp \left(\frac{2p\lambda_n \Gamma_n^{(2)}}{\Gamma_n} \|\sigma\|_\infty^2 [\varphi^{(2)}]_1 C_V^{\frac{1}{2}} |V_{k-1}|^{\frac{1}{2}} \right) \right)^{\frac{1}{4p}}. \end{aligned}$$

Using once again the Young inequality and setting $C_{3.14} := \frac{\|\sigma\|_\infty^4 [\varphi^{(2)}]_1^2 C_V}{4 c_V}$, we obtain:

$$\left(\mathbb{E} \exp \left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,\Sigma,n}| \right) \right)^{\frac{1}{4p}} \leq \exp \left(\frac{p\lambda_n^2}{4} \left(\frac{\Gamma_n^{(2)}}{\Gamma_n} \right)^2 \|\sigma\|_\infty^4 [\varphi^{(2)}]_1^2 \frac{C_V}{c_V} \right) (I_V^1)^{\frac{1}{4p}} \leq \exp \left(C_{3.14} p\lambda_n^2 \left(\frac{\Gamma_n^{(2)}}{\Gamma_n} \right)^2 \right) (I_V^1)^{\frac{1}{4p}}. \quad \square$$

4. A refinement when $\|\sigma\|^2 - \nu(\|\sigma\|^2)$ is a coboundary

We will assume in this section that there exists a solution ϑ of the Poisson problem $\mathcal{A}\vartheta = \|\sigma\|^2 - \nu(\|\sigma\|^2)$, where $\|\cdot\|$ is a matrix norm such that $\|\cdot\| \leq \|\cdot\|$, satisfying the assumptions stated for φ in Theorem 2. This is in particular the case for the Fröbenius norm $\|\cdot\|_F$ under the assumptions of the previous Theorem 4.

In this special case, we have a slightly different concentration result improving our previous ones for a certain deviation range.

Theorem 8. *Under the assumptions of Theorem 2 and with the notations introduced therein, we have that:*

- (a) For $\beta \in (0, 1]$ and $\theta \in (\frac{1}{2+\beta}, 1]$, there exist two explicit monotonic sequences $\tilde{c}_n \leq 1 \leq \tilde{C}_n$, $n \geq 1$, with $\lim_n \tilde{C}_n = \lim_n \tilde{c}_n = 1$ such that for all $n \geq 1$ and $a > 0$:

$$\begin{aligned} \mathbb{P} [|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)| \geq a] &\leq 2\tilde{C}_n \exp \left(- \frac{\tilde{c}_n}{2\nu(\|\sigma\|^2) \|\nabla\varphi\|_\infty^2} \Phi_n(a) \right), \\ \Phi_n(a) &:= \left[\left(a^2 \left(1 - \frac{2}{1 + \sqrt{1 + 4\tilde{c}_n^3 \frac{\Gamma_n}{a^2}}} \right) \right) \vee \left(a^{\frac{4}{3}} \Gamma_n^{\frac{1}{3}} \tilde{c}_n \left(1 - \frac{2}{3} \tilde{c}_n \left(\frac{\Gamma_n}{a^2} \right)^{\frac{1}{3}} \right)_+ \right) \right], \end{aligned}$$

where $x_+ = \max(x, 0)$ and $\tilde{c}_n := (\frac{[\varphi]_1}{[\vartheta]_1})^{2/3} \nu(\|\sigma\|^2) \|\sigma\|_\infty^{-2/3} \check{c}_n$ with \check{c}_n being an explicit positive sequence s.t. $\check{c}_n \downarrow 1$.

- (b) For $\beta \in (0, 1]$, $\theta = \frac{1}{2+\beta}$, there exist two explicit monotonic sequences $\tilde{c}_n \leq 1 \leq \tilde{C}_n$, $n \geq 1$, with $\lim_n \tilde{C}_n = \lim_n \tilde{c}_n = 1$ such that for all $n \geq 1$ and $a > 0$:

$$\mathbb{P} [|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) + \mathcal{B}_{n,\beta}| \geq a] \leq 2\tilde{C}_n \exp \left(- \frac{\tilde{c}_n}{2\nu(\|\sigma\|^2) \|\nabla\varphi\|_\infty^2} \Phi_n(a) \right).$$

Remark 10 (About deviation rates). Observe that in order to derive global deviation bounds (valid for every $a > 0$) two concentration regimes appear in the previous bounds. For an arbitrary fixed $a > 0$, we have that for n large enough (depending on a), the Gaussian concentration regime will give the fastest decay, since $\frac{2}{1 + \sqrt{1 + 4\tilde{c}_n^3 \frac{\Gamma_n}{a^2}}} \rightarrow 0$. Also, when

$a \asymp \sqrt{\Gamma_n}$ the two above contributions give a Gaussian bound, with suboptimal constants. Eventually, when $a \gg \sqrt{\Gamma_n}$, for a fixed n , we have that the first term is “stuck” at the threshold Γ_n whatever level a is considered, i.e. $a^2 (1 - \frac{2}{1 + \sqrt{1 + 4\tilde{c}_n^3 \frac{\Gamma_n}{a^2}}}) \xrightarrow{a \rightarrow \infty} \tilde{c}_n^3 \Gamma_n$ whereas the second clearly becomes bigger.

To summarize, when the Gaussian regime prevails (i.e. when $\frac{a}{\sqrt{\Gamma_n}}$ is small), the results of Theorem 2 have been improved in the sense that the variance in the deviations is a sharper upper bound of the “carré du champ” $\int_{\mathbb{R}^d} |\sigma^* \nabla\varphi(x)|^2 \nu(dx)$ appearing in the asymptotic Theorem 1. Indeed, we managed to replace the supremum norm $\|\sigma\|_\infty^2$ deriving from Theorem 2 and the domination condition on the matrix norms by $\nu(\|\sigma\|^2)$. However, our martingale approach naturally leads to a bound in $\|\nabla\varphi\|_\infty^2$.

On the other hand, the global double regime seems to be the price to pay to benefit from the better approximation of the “carré du champ” in the Gaussian regime.

Eventually, Theorem 3 is a direct consequence of the previous theorem in the Gaussian regime.

Proof. We focus on case (a) for $\beta \in (0, 1)$, $\theta \in (1/(2 + \beta), 1]$. Case (b) could be derived similarly following the proof of Theorem 2. We restart from the computations of Section 3.1 that give for any $\lambda > 0$ the control in equation (3.9). Let us now deal with the term giving the concentration and write for any $\rho > 1$:

$$\begin{aligned} \mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n} M_n\right) &\leq \left(\mathbb{E} \exp\left(-\rho \frac{q\lambda}{\Gamma_n} M_n - \frac{\rho^2(q\lambda)^2[\varphi]_1^2}{2\Gamma_n^2} \sum_{k=1}^n \gamma_k \mathcal{A}\vartheta(X_{k-1})\right)\right)^{\frac{1}{\rho}} \\ &\quad \times \left(\mathbb{E} \exp\left(\frac{\rho^2(q\lambda)^2[\varphi]_1^2}{2(\rho-1)\Gamma_n^2} \sum_{k=1}^n \gamma_k \mathcal{A}\vartheta(X_{k-1})\right)\right)^{1-\frac{1}{\rho}} =: \mathcal{T}_1^{\frac{1}{\rho}} \mathcal{T}_2^{1-\frac{1}{\rho}}. \end{aligned} \tag{4.1}$$

Since for any $x \in \mathbb{R}^d$, $\mathcal{A}\vartheta(x) = \|\sigma(x)\|^2 - \nu(\|\sigma\|^2)$, we obtain:

$$\mathcal{T}_1 = \exp\left(\frac{\rho^2(q\lambda)^2[\varphi]_1^2 \nu(\|\sigma\|^2)}{2\Gamma_n}\right) \mathbb{E} \exp\left(-\rho \frac{q\lambda}{\Gamma_n} M_n - \frac{\rho^2(q\lambda)^2[\varphi]_1^2}{2\Gamma_n^2} \sum_{k=1}^n \gamma_k \|\sigma(X_{k-1})\|^2\right).$$

The key idea is that we have exploited the Poisson equation solved by ϑ to replace the previous *rough* control $\exp(\frac{(q\lambda)^2[\varphi]_1^2 \|\sigma\|_\infty^2}{2\Gamma_n})$, coming from the martingale increment obtained in equation (3.6) and the domination condition on the matrix norms, by the above term $\exp(\frac{\rho^2(q\lambda)^2[\varphi]_1^2 \nu(\|\sigma\|^2)}{2\Gamma_n})$. This last contribution will be part of the optimization procedure over λ . This improvement will be all the more significant that neighborhoods of the points where the norm of the diffusion coefficient σ attains its supremum are not *very much* charged by the invariant distribution. The point for \mathcal{T}_1 is then to prove that the remaining expectation is less than 1. It will be shown by exhibiting an appropriate underlying supermartingale.

Set to this end $\tilde{\mathcal{T}}_1 := \exp(-\frac{\rho^2(q\lambda)^2[\varphi]_1^2 \nu(\|\sigma\|^2)}{2\Gamma_n}) \mathcal{T}_1$. Define now, for a given $n \in \mathbb{N}$ and $m \in \mathbb{N}_0$, $S_m := \exp(-\rho \frac{q\lambda}{\Gamma_n} M_m - \frac{\rho^2(q\lambda)^2[\varphi]_1^2}{2\Gamma_n^2} \sum_{k=1}^m \gamma_k \|\sigma(X_{k-1})\|^2)$. From the definition of the martingale $(M_k)_{k \geq 1}$ in (3.2) and the controls of the Lipschitz constants of the functions $(\psi_k(X_{k-1}, \cdot))_{k \in [1, n]}$ in Lemma 1, we get by iterated conditioning:

$$\begin{aligned} \tilde{\mathcal{T}}_1 &\leq \mathbb{E} \left[S_{n-1} \exp\left(-\frac{\rho^2(q\lambda)^2[\varphi]_1^2}{2\Gamma_n^2} \gamma_n \|\sigma(X_{n-1})\|^2\right) \mathbb{E} \left[\exp\left(-\rho \frac{q\lambda}{\Gamma_n} (M_n - M_{n-1})\right) \middle| \mathcal{F}_{n-1} \right] \right] \\ &\stackrel{\text{(GC)}}{\leq} \mathbb{E} \left[S_{n-1} \exp\left(-\frac{\rho^2(q\lambda)^2[\varphi]_1^2}{2\Gamma_n^2} \gamma_n \|\sigma(X_{n-1})\|^2\right) \exp\left(\frac{\rho^2(q\lambda)^2}{2\Gamma_n^2} \gamma_n [\varphi]_1^2 \|\sigma(X_{n-1})\|^2\right) \right] \leq \mathbb{E}[S_{n-1}] \leq 1. \end{aligned}$$

In other words, $(S_m)_{m \geq 0}$ is a positive supermartingale. We finally get that, for any $\rho > 1$:

$$\mathcal{T}_1^{\frac{1}{\rho}} \leq \exp\left(\frac{\rho(q\lambda)^2[\varphi]_1^2 \nu(\|\sigma\|^2)}{2\Gamma_n}\right). \tag{4.2}$$

For the term \mathcal{T}_2 , we have that setting $\mu := \mu(q, n, \rho, \lambda) = \frac{\rho^2(q\lambda)^2[\varphi]_1^2}{2(\rho-1)\Gamma_n}$,

$$\mathcal{T}_2 = \mathbb{E} \exp\left(\frac{\mu}{\Gamma_n} \sum_{k=1}^n \gamma_k \mathcal{A}\vartheta(X_{k-1})\right),$$

so that this contribution can be controlled from the previous expansion of Lemma 1 exploiting the technical lemmas of Section 3.1 replacing λ by μ and φ by ϑ .

In case (a), for $\theta \in (1/(2 + \beta), 1]$, $\beta \in (0, 1]$, the Hölder inequalities yield that for all $\mu \in \mathbb{R}_+$ and $\bar{p}, \bar{q} \in (1, +\infty)$, $\frac{1}{\bar{p}} + \frac{1}{\bar{q}} = 1$, similarly to (3.9),

$$\mathcal{T}_2 = \mathbb{E} \exp\left(\frac{\mu}{\Gamma_n} \sum_{k=1}^n \gamma_k \mathcal{A}\vartheta(X_{k-1})\right)$$

$$\begin{aligned} &\leq \left(\mathbb{E} \exp\left(-\frac{\bar{q}\mu}{\Gamma_n} M_n^\vartheta\right) \right)^{\frac{1}{q}} \left(\mathbb{E} \exp\left(\frac{2\bar{p}\mu}{\Gamma_n} (|L_n^\vartheta| + |\bar{G}_n^\vartheta|)\right) \right)^{\frac{1}{2p}} \left(\mathbb{E} \exp\left(\frac{4\bar{p}\mu}{\Gamma_n} |D_{2,b,n}^\vartheta|\right) \right)^{\frac{1}{4p}} \\ &\quad \times \left(\mathbb{E} \exp\left(\frac{4\bar{p}\mu}{\Gamma_n} |D_{2,\Sigma,n}^\vartheta|\right) \right)^{\frac{1}{4p}}, \end{aligned} \tag{4.3}$$

where the superscripts in ϑ emphasize that the contributions to be analyzed are those associated with the solution ϑ of the Poisson problem with source $\|\|\sigma\|^2 - \nu(\|\|\sigma\|^2)$.

Still for simplicity, we assume as well (case (i)) that the mapping $x \mapsto \langle b(x), \nabla \vartheta(x) \rangle$ is Lipschitz continuous. Plugging in (4.3) the controls established in Lemma 3 (inequality (3.13) with $a_n^\vartheta := \frac{[\vartheta^{(3)}]_\beta \|\sigma\|_\infty^{(3+\beta)} \mathbb{E}[|U_1|^{3+\beta}] \Gamma_n^{\frac{(3+\beta)}{2}}}{(1+\beta)(2+\beta)(3+\beta) \sqrt{\Gamma_n}}$), Lemma 4 (with $j = 2$), Lemma 5 (inequalities (3.14) and (3.15)) and (3.27), then replacing λ_n by μ , we get similarly to the first inequality of (3.18) and with the notations of Lemma 3:

$$\begin{aligned} \mathcal{F}_2 &\leq \exp\left(\frac{\bar{q}\mu^2 \|\|\sigma\|^2_\infty [\vartheta]_1^2}{2\Gamma_n}\right) \exp\left(\frac{\mu^2}{2\Gamma_n \bar{p}} + \frac{\bar{p}(a_n^\vartheta)^2}{2}\right) \exp\left(\frac{3\bar{p}C_{V,\vartheta}^2 \mu^2}{c_V \Gamma_n^2} + \frac{c_V}{\bar{p}}\right) (I_V^1)^{\frac{1}{2\bar{p}}} \\ &\quad \times \exp\left(C_{3.14} \frac{\bar{p}\mu^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2}\right) (I_V^1)^{\frac{1}{4\bar{p}}} \times \exp\left(C_{3.15} \frac{\bar{p}\mu^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2}\right) (I_V^1)^{\frac{1}{4\bar{p}}} \\ &\leq \exp\left(\frac{\mu^2}{\Gamma_n} \left(\frac{\bar{q}\|\|\sigma\|^2_\infty [\vartheta]_1^2}{2} + \bar{p} \left(\frac{(\Gamma_n^{(2)})^2}{\Gamma_n} [C_{3.14} + C_{3.15}] + \frac{3C_{V,\vartheta}^2}{c_V \Gamma_n} + \frac{1}{2\bar{p}}\right)\right)\right) \times \exp\left(\frac{c_V}{\bar{p}} + \frac{\bar{p}(a_n^\vartheta)^2}{2}\right) (I_V^1)^{\frac{1}{\bar{p}}}. \end{aligned}$$

Set now

$$\begin{aligned} \bar{C}_n &:= \exp\left(\frac{c_V}{\bar{p}} + \frac{\bar{p}(a_n^\vartheta)^2}{2}\right) (I_V^1)^{\frac{1}{\bar{p}}}, \\ \bar{e}_n &:= \bar{p} \left(\frac{(\Gamma_n^{(2)})^2}{\Gamma_n} [C_{3.14} + C_{3.15}] + \frac{3C_{V,\vartheta}^2}{c_V \Gamma_n}\right) + \frac{1}{2\bar{p}}. \end{aligned} \tag{4.4}$$

In the considered case, the exponent $\bar{p} := \bar{p}_n$ can again be taken such that $\bar{p}_n \xrightarrow{n} +\infty$ and $\bar{p}_n \frac{(\Gamma_n^{(2)})^2}{\Gamma_n} \xrightarrow{n} 0$ in order to have, $\bar{e}_n \xrightarrow{n} 0, \bar{C}_n \xrightarrow{n} 1$ with the indicated monotonicity for large enough n .

We derive from the above control and (4.2) that for all $q, \rho > 1$:

$$\left(\mathbb{E} \exp\left(-\frac{\lambda q}{\Gamma_n} M_n\right) \right)^{\frac{1}{q}} \leq (\mathcal{F}_1^{\frac{1}{\rho}} \mathcal{F}_2^{1-\frac{1}{\rho}})^{\frac{1}{q}} \leq \exp\left(\frac{\rho q \lambda^2 [\varphi]_1^2 \nu(\|\|\sigma\|^2)}{2\Gamma_n}\right) \bar{C}_n^{\frac{\rho-1}{\rho q}} \exp\left(\frac{\rho-1}{\rho q} \frac{\mu^2}{\Gamma_n} \left(\frac{\bar{q}\|\|\sigma\|^2_\infty [\vartheta]_1^2}{2} + \bar{e}_n\right)\right).$$

Plugging this bound in (3.9), using again the controls of Lemmas 4 and 5, eventually yields:

$$\begin{aligned} &\mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] \\ &\leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \exp\left(\frac{\lambda^2}{2\Gamma_n} \left(\rho q [\varphi]_1^2 \nu(\|\|\sigma\|^2) + \frac{1}{p}\right)\right) \bar{C}_n^{\frac{\rho-1}{\rho q}} \exp\left(\frac{\rho-1}{\rho q} \frac{\mu^2}{\Gamma_n} \left(\frac{\bar{q}\|\|\sigma\|^2_\infty [\vartheta]_1^2}{2} + \bar{e}_n\right)\right) \\ &\quad \times \exp\left(\frac{\lambda^2}{\Gamma_n} p \left(\frac{(\Gamma_n^{(2)})^2}{\Gamma_n} [C_{3.14} + C_{3.15}] + \frac{3C_{V,\varphi}^2}{c_V \Gamma_n}\right)\right) \exp\left(\frac{c_V}{p} + \frac{p a_n^2}{2}\right) (I_V^1)^{\frac{1}{p}}. \end{aligned}$$

Choosing $p := p_n \xrightarrow{n} +\infty$ and such that $p_n \frac{(\Gamma_n^{(2)})^2}{\Gamma_n} \xrightarrow{n} 0$, we get by a standard symmetry and with the notations introduced in the proof of Theorem 2:

$$\begin{aligned} \mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)| \geq a] &\leq 2C_n \bar{C}_n^{\frac{\rho-1}{\rho q}} \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \exp\left(\frac{\lambda^2}{\Gamma_n} \left(\frac{\rho q [\varphi]_1^2 \nu(\|\|\sigma\|^2)}{2} + e_n\right)\right) \\ &\quad \times \exp\left(\frac{\rho-1}{\rho q} \frac{\mu^2}{\Gamma_n} \left(\frac{\bar{q}\|\|\sigma\|^2_\infty [\vartheta]_1^2}{2} + \bar{e}_n\right)\right), \end{aligned}$$

where e_n is defined similarly to \bar{e}_n in (4.4) replacing \bar{p} by p . In particular $e_n \xrightarrow{n} 0$. Note that for the previous choices of p, \bar{p} , we have that $\tilde{C}_n := C_n \bar{C}_n^{\frac{\rho-1}{\rho q}} \xrightarrow{n} 1$ uniformly in $\rho > 1$. Recalling that $\mu = \frac{(q\lambda)^2 \rho^2 [\varphi]_1^2}{2(\rho-1)\Gamma_n}$, we are thus led to minimize the polynomial function

$$P : \lambda \mapsto -\frac{a\lambda}{\sqrt{\Gamma_n}} + \frac{\lambda^2}{\Gamma_n} A_n + \frac{\lambda^4}{\Gamma_n^3} B_n,$$

where $A_n = A_n(\rho) = \rho \tilde{A}_n$ and $B_n = B_n(\rho) = \frac{\rho^3}{\rho-1} \tilde{B}_n$ with

$$\tilde{A}_n := \frac{q[\varphi]_1^2 \nu(\|\sigma\|^2)}{2} + e_n \quad \text{and} \quad \tilde{B}_n := \frac{q^3[\varphi]_1^4}{4} \left(\frac{\bar{q} \|\sigma\|_\infty^2 [\vartheta]_1^2}{2} + \bar{e}_n \right). \quad (4.5)$$

Note that both sequences $(\tilde{A}_n)_{n \geq 1}$ and $(\tilde{B}_n)_{n \geq 1}$ are bounded and bounded away from zero sequences (and do not depend on ρ). The function P is clearly convex and coercive so it attains its minimum at λ_{\min} , unique zero of the equation $P'(\lambda_{\min}) = 0$. This equation reads

$$\lambda^3 + \frac{A_n \Gamma_n^2}{2B_n} \lambda - \frac{a \Gamma_n^{\frac{5}{2}}}{4B_n} = 0 \quad (4.6)$$

which is the canonical form of this third degree equation to apply the Cardan–Tartaglia formula¹ so that

$$\lambda_{\min}(\rho) = \frac{\Gamma_n}{2} \left[\left(\frac{a}{\sqrt{\Gamma_n} B_n} + \sqrt{\left(\frac{2A_n}{3B_n} \right)^3 + \frac{a^2}{\Gamma_n B_n^2}} \right)^{\frac{1}{3}} + \left(\frac{a}{\sqrt{\Gamma_n} B_n} - \sqrt{\left(\frac{2A_n}{3B_n} \right)^3 + \frac{a^2}{\Gamma_n B_n^2}} \right)^{\frac{1}{3}} \right]. \quad (4.7)$$

In order to derive our non-asymptotic bound, we select two “regimes” based on a first order expansion of λ_{\min} in two cases $\frac{a}{B_n \sqrt{\Gamma_n}} \rightarrow 0$ and $\frac{B_n \sqrt{\Gamma_n}}{a} \rightarrow 0$, assuming that the free parameter $\rho = \rho_n$ to be specified later on remains bounded, *e.g.* $\rho \in (1, 3]$ (which implies that both quantities $\frac{A_n}{B_n}$ and $\frac{1}{B_n}$ remain bounded as well). Also, note that if $\rho \rightarrow 1$, then $\frac{1}{B_n}$ and $\frac{A_n}{B_n} \rightarrow 0$. First, one easily checks that if $(x_n)_{n \geq 1}$ and $(y_n)_{n \geq 1}$ are two sequences of positive real numbers where $(y_n)_{n \geq 1}$ is bounded, then

$$(x_n + \sqrt{y_n^3 + x_n^2})^{\frac{1}{3}} + (x_n - \sqrt{y_n^3 + x_n^2})^{\frac{1}{3}} \sim \begin{cases} \frac{2}{3} \frac{x_n}{y_n} & \text{if } x_n = o(y_n^{\frac{3}{2}}) \text{ (then } x_n \rightarrow 0), \\ (2x_n)^{\frac{1}{3}} & \text{if } y_n = o(x_n^{\frac{2}{3}}) \text{ (then } x_n \rightarrow +\infty). \end{cases} \quad (4.8)$$

- If $\frac{a}{B_n \sqrt{\Gamma_n}} = o\left(\left(\frac{A_n}{B_n}\right)^{\frac{3}{2}}\right)$ (hence goes to 0), setting then $x_n = \frac{a}{B_n \sqrt{\Gamma_n}}$ and $y_n = \frac{2A_n}{3B_n}$ yields

$$\lambda_{\min}(\rho) \sim \lambda^*(\rho) := \frac{a \sqrt{\Gamma_n}}{2A_n} \quad \text{as } n \rightarrow +\infty.$$

Note that $\lambda^* := \lambda^*(\rho)$ corresponds to the optimization of the quadratic part of P . Then

$$P(\lambda^*) = -\frac{a^2}{4A_n} \left(1 - \frac{a^2}{4A_n^3} \frac{B_n}{\Gamma_n} \right) = -\frac{a^2}{4A_n \rho} \left(1 - \frac{a^2}{4A_n^3 (\rho-1)} \frac{\tilde{B}_n}{\Gamma_n} \right).$$

Set now $\xi_n := \frac{\alpha_n(a)}{\rho-1}$ with $\alpha_n(a) = \frac{\tilde{B}_n}{4A_n^3} \frac{a^2}{\Gamma_n}$. Then

$$P(\lambda^*) = -\frac{a^2}{4\tilde{A}_n} \frac{1 - \xi_n}{1 + \frac{\alpha_n(a)}{\xi_n}}.$$

¹If the equation $z^3 + pz + q = 0$ has a unique real zero z_* then its discriminant $\Delta = 4p^3 + 27q^2 > 0$ and $z_* = \left(\frac{1}{2}(-q + \sqrt{\frac{\Delta}{27}})\right)^{\frac{1}{3}} + \left(\frac{1}{2}(-q - \sqrt{\frac{\Delta}{27}})\right)^{\frac{1}{3}}$.

It remains to maximize the mapping $\xi \mapsto \frac{1-\xi}{1+\alpha_n(a)\xi^{-1}}$ over $(0, 1)$. Its optimum is attained for $\xi_n^* = \frac{1}{1+\sqrt{1+\frac{1}{\alpha_n(a)}}}$, which in turn yields

$$P(\lambda^*) = -\frac{a^2}{4\tilde{A}_n} \left(1 - \frac{2}{1 + \sqrt{1 + 4\frac{\tilde{A}_n^3 \Gamma_n}{B_n a^2}}} \right). \tag{4.9}$$

Let us remark that, with the resulting specification of $\rho = \rho_n^* := 1 + \frac{\alpha_n(a)}{\xi_n^*} \in (1, 3]$ (at least for large enough n), the above condition $x_n = o(y_n^{\frac{3}{2}})$ in (4.8) is satisfied *a posteriori*.

- If $\frac{a}{B_n \sqrt{\Gamma_n}} \rightarrow +\infty$, then, still owing to (4.8),

$$\lambda_{\min}(\rho) \sim \bar{\lambda}^*(\rho) = \frac{\Gamma_n}{2} \left(\frac{2a}{B_n \sqrt{\Gamma_n}} \right)^{\frac{1}{3}} = \left(\frac{a\Gamma_n}{4B_n} \right)^{\frac{1}{3}} \sqrt{\Gamma_n} \quad \text{as } n \rightarrow +\infty.$$

The value $\bar{\lambda}^*(\rho)$ corresponds to the *quartic* pseudo-optimum of P (i.e. obtained by neglecting the quadratic term). This yields, when reintroducing the parameter ρ ,

$$P(\bar{\lambda}^*(\rho)) = -a^{\frac{4}{3}} \Gamma_n^{\frac{1}{3}} \frac{(\rho - 1)^{\frac{1}{3}}}{\rho (4\tilde{B}_n)^{\frac{1}{3}}} \left(\frac{3}{4} - \frac{\tilde{A}_n}{(4\tilde{B}_n)^{\frac{1}{3}}} \frac{\Gamma_n^{\frac{1}{3}}}{a^{\frac{2}{3}}} (\rho - 1)^{\frac{1}{3}} \right).$$

The right hand side of this equality is a function of $\rho \in (1, +\infty)$. Its analysis yields that the optimum is attained in $(1, 3/2]$ and that it tends asymptotically in n to $3/2$ in our considered regime. Taking as suboptimal $\rho = 3/2$ gives:

$$P(\bar{\lambda}^*(\rho)) \leq -\frac{a^{\frac{4}{3}}}{4} \left(\frac{\Gamma_n}{B_n} \right)^{\frac{1}{3}} \left(1 - \frac{2}{3} \frac{\tilde{A}_n}{B_n} \left(\frac{\Gamma_n}{a^2} \right)^{\frac{1}{3}} \right). \tag{4.10}$$

From (4.10), (4.9) and (4.5), we conclude the proof of case (a) by setting $\bar{c}_n := \tilde{A}_n \tilde{B}_n^{-\frac{1}{3}}$ which matches with the definition in the statement of the Theorem.

In the biased case, the result follows similarly from the corresponding analysis performed in Section 2 taking $A_n(\rho) = \frac{\rho q[\varphi]_{\Gamma}^2 v(\|\sigma\|^2)}{2}$. □

Remark 11. • When $a \asymp \sqrt{\Gamma_n}$, one checks that $\lambda_{\min}(\rho) \asymp \Gamma_n$ and $P(\lambda_{\min}(\rho)) \asymp -\Gamma_n$. This behavior is consistent with our non-asymptotic bound. However, for practical and numerical purposes observe that the optimum can be estimated. Namely, plugging the identity (4.6) satisfied by $\lambda_{\min}(\rho)$ in (4.7) into the definition of P , yields

$$\begin{aligned} P(\lambda_{\min}(\rho)) &= -\frac{\lambda_{\min}(\rho)}{2\sqrt{\Gamma_n}} \left(\frac{3a}{2} - \frac{\lambda_{\min}(\rho)\rho\tilde{A}_n}{\sqrt{\Gamma_n}} \right) \\ &= -\frac{\sqrt{\Gamma_n}}{4} \frac{(\rho - 1)^{\frac{1}{3}}}{\rho} \Phi_n(a, \rho) \left(\frac{3a}{2} - \frac{\sqrt{\Gamma_n}}{2} (\rho - 1)^{\frac{1}{3}} \tilde{A}_n \Phi_n(a, \rho) \right), \end{aligned}$$

where

$$\Phi_n(a, \rho) = \left(\frac{a}{\sqrt{\Gamma_n} \tilde{B}_n} + \left((\rho - 1) \left(\frac{2\tilde{A}_n}{3\tilde{B}_n} \right)^3 + \frac{a^2}{\tilde{B}_n^2 \Gamma_n} \right)^{\frac{1}{2}} \right)^{\frac{1}{3}} + \left(\frac{a}{\sqrt{\Gamma_n} \tilde{B}_n} - \left((\rho - 1) \left(\frac{2\tilde{A}_n}{3\tilde{B}_n} \right)^3 + \frac{a^2}{\tilde{B}_n^2 \Gamma_n} \right)^{\frac{1}{2}} \right)^{\frac{1}{3}}.$$

Then, an optimization in $\rho \in (1, +\infty)$ for given a, Γ_n can be performed (noting that $\rho \mapsto (\rho - 1)^{i/3} \rho^{-1}$, $i \in \{1, 2\}$ are bounded functions over $(1, +\infty)$).

5. Smoothness results for the Poisson problems (proof of Theorem 4)

We first prove here Theorem 4 which allows to derive from the deviation results of Theorems 2 and 3 the practical deviation bounds of Section 2.3 (i.e. Theorems 5, 6 and 7). We recall that we work in the confluent setting of (\mathbf{D}_α^p) and that we additionally consider two main types of assumptions:

- Strong confluence conditions and smoothness (\mathbf{C}_R) . Namely, assumptions (\mathcal{L}_V) , (\mathbf{D}_α^p) and $(\mathbf{R}_{3,\beta})$ introduced in Sections 1.2 and 2.2 with the condition $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{2(3+\beta)-p}$.
- Mild confluence conditions and non-degeneracy (\mathbf{C}_{UE}) . Namely, assumptions (\mathcal{L}_V) , (\mathbf{D}_α^p) , $(\mathbf{R}_{1,\beta})$ and (\mathbf{UE}) introduced in Sections 1.2 and 2.2 together, when $d > 1$, with the condition $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{2(1+\beta)-p}$ and the technical structure assumption on the diffusion coefficient that for each $(i, j) \in \llbracket 1, d \rrbracket^2$, $\Sigma_{i,j}(x) = \Sigma_{i,j}(x_{i \wedge j}, \dots, x_d)$.

It is well known that when (\mathbf{C}_R) or (\mathbf{C}_{UE}) are in force, there exists a unique invariant distribution for (1.1), i.e. assumption (\mathbf{U}) holds. We refer to [23,33,35,37] for proofs of this assertion. The next step consists precisely in investigating the smoothness of the corresponding Poisson problem as well as some associated quantitative pointwise bounds on the gradient of its solution, which is one of the key terms appearing in the deviation bounds of Theorems 2 and 3.

Let us indicate that the conditions appearing in (\mathbf{C}_R) depend on pure pathwise properties, whereas the case (\mathbf{C}_{UE}) takes advantage of the regularity of the underlying semi-group which allows to alleviate some smoothness assumptions on the coefficients and some restrictions on the variations of σ . When the dimension increases, it becomes useful to benefit from the smoothing effects of a non-degenerate semi-group, especially if we keep in mind that one of our goals is to handle Lipschitz continuous sources.

5.1. Proof of Theorem 4

Under (\mathbf{C}_{UE}) or (\mathbf{C}_R) , it is well known that the Poisson equation (2.5) that we now recall:

$$\forall x \in \mathbb{R}^d, \quad \mathcal{A}\varphi(x) = f(x) - \nu(f),$$

admits a unique solution centered w.r.t. ν and with linear growth, in $W_{p,\text{loc}}^2(\mathbb{R}^d, \mathbb{R})$ for any $p > 1$ under (\mathbf{C}_{UE}) (see [37]), or in $\mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ under (\mathbf{C}_R) (see Proposition A.8 in [35]). In both cases, we have the following representation:

$$\varphi(x) = - \int_{\mathbb{R}_+} (P_t f(x) - \nu(f)) dt \quad \text{where } P_t f(x) := \mathbb{E}[f(Y_t^{0,x})] \tag{5.1}$$

and $Y_t^{0,x}$ solves (1.1) with $Y_0^{0,x} = x$. To comply with the framework of the above Theorems 5 and 6, the first step is to establish a pointwise gradient control.

5.1.1. Gradient control

Under (\mathbf{C}_{UE}) or (\mathbf{C}_R) we manage to obtain pointwise gradient bounds for φ . In our current confluent setting, these estimates are obtained through controls on the tangent flow, again without any *a priori* uniform ellipticity condition of type (\mathbf{UE}) .

Lemma 6 (Pointwise Gradient Bounds). *Assume that (\mathbf{C}_{UE}) or (\mathbf{C}_R) holds. Then*

$$\|\nabla\varphi\|_\infty \leq \frac{[f]_1}{\alpha},$$

with α as in (\mathbf{D}_α^p) .

Proof. *Gradient Control in the Confluent framework.* Assume now that (\mathbf{D}_α^p) holds. Observe that, as soon as $(\mathbf{R}_{1,\beta})$ holds, it is well known that that $\nabla_x Y_t^{0,x}$ is well defined and belongs to $L^2(\mathbb{P})$, see [20]. Hence, for $t > 0$, $i \in \llbracket 1, d \rrbracket$:

$$\begin{aligned} \partial_{x_i} \mathbb{E}[f(Y_t^{0,x})] &= \mathbb{E}[\langle \nabla f(Y_t^{0,x}), \partial_{x_i} Y_t^{0,x} \rangle], \\ \partial_{x_i} Y_t^{0,x} &= e_i + \int_0^t Db(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x} ds + \sum_{j=1}^d \int_0^t D\sigma_{\cdot j}(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x} dW_s^j, \end{aligned}$$

where e_i stands for the i th canonical vector and $Db, D\sigma_{\cdot j} \in \mathbb{R}^d \otimes \mathbb{R}^d$.

Let $p \in (1, 2]$ be given such that (\mathbf{D}_α^p) holds. Considering the mapping $y \in \mathbb{R}^d \mapsto |y|^p$, where $|\cdot|$ stands for the Euclidean norm of \mathbb{R}^d , it is easily seen from Itô's formula that:

$$\begin{aligned} |\partial_{x_i} Y_t^{0,x}|^p &= 1 + p \int_0^t \left\langle \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|}, Db(Y_s^{0,x}) \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|} \right\rangle |\partial_{x_i} Y_s^{0,x}|^p ds \\ &\quad + p \sum_{j=1}^d \int_0^t \left\langle \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|}, D\sigma_{\cdot j}(Y_s^{0,x}) \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|} \right\rangle |\partial_{x_i} Y_s^{0,x}|^p dW_s^j \\ &\quad + \frac{p}{2} \sum_{j=1}^d \int_0^t \left(\frac{|D\sigma_{\cdot j}(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x}|^2}{|\partial_{x_i} Y_s^{0,x}|^2} + (p-2) \frac{|\langle \partial_{x_i} Y_s^{0,x}, D\sigma_{\cdot j}(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x} \rangle|^2}{|\partial_{x_i} Y_s^{0,x}|^4} \right) \\ &\quad \times |\partial_{x_i} Y_s^{0,x}|^p ds \\ &= \exp\left(p \int_0^t \left\langle \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|}, Db(Y_s^{0,x}) \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|} \right\rangle ds\right) \times \mathcal{E}(M)_t \\ &\quad \times \exp\left(\frac{p}{2} \sum_{j=1}^d \int_0^t \left(\frac{|D\sigma_{\cdot j}(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x}|^2}{|\partial_{x_i} Y_s^{0,x}|^2} + (p-2) \frac{|\langle \partial_{x_i} Y_s^{0,x}, D\sigma_{\cdot j}(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x} \rangle|^2}{|\partial_{x_i} Y_s^{0,x}|^4} \right) ds\right), \end{aligned} \tag{5.2}$$

where $(M_t)_{t \geq 0} := (p \sum_{j=1}^d \int_0^t \langle \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|}, D\sigma_{\cdot j}(Y_s^{0,x}) \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|} \rangle dW_s^j)_{t \geq 0}$ is a square integrable martingale with bounded integrand and $\mathcal{E}(M)_t := \exp(M_t - \frac{1}{2} \langle M \rangle_t)$ denotes the associated Doléans exponential martingale. From condition (\mathbf{D}_α^p) , we thus get:

$$|\partial_{x_i} Y_t^{0,x}|^p \leq \exp(-\alpha pt) \times \mathcal{E}(M)_t. \tag{5.3}$$

We eventually derive:

$$\int_0^{+\infty} |\mathbb{E}[\langle \nabla f(Y_t^{0,x}), \partial_{x_i} Y_t^{0,x} \rangle]| dt \leq [f]_1 \int_0^{+\infty} \mathbb{E}[|\partial_{x_i} Y_t^{0,x}|^p]^{1/p} dt \leq [f]_1 \int_0^{+\infty} \exp(-\alpha t) dt = \frac{[f]_1}{\alpha}.$$

From the above control and equation (5.1), we thus derive:

$$\forall i \in \llbracket 1, d \rrbracket, \forall x \in \mathbb{R}^d, \quad |\partial_{x_i} \varphi(x)| \leq \frac{[f]_1}{\alpha}. \tag{5.4}$$

Similarly, for any $x \in \mathbb{R}^d$, $\nabla \varphi(x) = \int_0^{+\infty} \mathbb{E}[(\nabla Y_t^{0,x})^* \nabla f(Y_t^{0,x})] dt$ where $\nabla Y_t^{0,x} = (\partial_{x_1} Y_t^{0,x} \dots \partial_{x_d} Y_t^{0,x})$ so that

$$(\nabla Y_t^{0,x})^* = \begin{pmatrix} (\partial_{x_1} Y_t^{0,x})^* \\ \vdots \\ (\partial_{x_d} Y_t^{0,x})^* \end{pmatrix}.$$

Hence, recalling that $|\cdot|$ stands for the Euclidean norm, $|\nabla \varphi(x)| \leq \int_0^{+\infty} \mathbb{E}[\|(\nabla Y_t^{0,x})^*\| \|\nabla f(Y_t^{0,x})\|] dt$ where we recall that for $A \in \mathbb{R}^d \otimes \mathbb{R}^d$, $\|A\| := \sup_{|z| \leq 1, z \in \mathbb{R}^d} |Az|$ denotes the operator (or spectral) matrix norm. Thus, $|\nabla \varphi(x)| \leq \|\nabla f\|_\infty \int_0^{+\infty} \mathbb{E}[\|(\nabla Y_t^{0,x})^*\|^p]^{1/p} dt = \|\nabla f\|_\infty \int_0^{+\infty} \mathbb{E}[\|\nabla Y_t^{0,x}\|^p]^{1/p} dt$. Now,

$$\|\nabla Y_t^{0,x}\| = \sup_{|z| \leq 1} |\nabla Y_t^{0,x} z|.$$

For any $z \in \mathbb{R}^d$, $|z| \leq 1$, setting $Z_t^{0,x,z} := \nabla Y_t^{0,x} z$, one has the following dynamics for the \mathbb{R}^d -valued process $(Z_s^{0,x,z})_{s \in [0,t]}$:

$$Z_t^{0,x,z} := z + \int_0^t Db(Y_s^{0,x}) Z_s^{0,x,z} ds + \sum_{j=1}^d \int_0^t D\sigma_{\cdot j}(Y_s^{0,x}) Z_s^{0,x,z} dW_s^j.$$

Hence, we derive similarly to (5.3) that $|Z_t^{0,x,z}|^p \leq |z|^p \exp(-p\alpha t)\mathcal{E}(M_t)$, where $\mathcal{E}(M_t)$ does not depend on z . Write now,

$$\mathbb{E}[\|\nabla Y_t^{0,x}\|^p]^{1/p} = \mathbb{E}\left[\sup_{|z|\leq 1} |Z_t^{0,x,z}|^p\right]^{1/p} \leq \mathbb{E}\left[\sup_{|z|\leq 1} |z|^p \exp(-p\alpha t)\mathcal{E}(M_t)\right]^{1/p} \leq \exp(-\alpha t). \tag{5.5}$$

This eventually proves the claim $\|\nabla\varphi\|_\infty := \sup_{x\in\mathbb{R}^d} |\nabla\varphi(x)| \leq \frac{\|\nabla f\|_\infty}{\alpha}$. □

5.1.2. *Additional smoothness*

- Theorem 4 can be derived under (C_R) , by iterating computations similar to the ones performed in Lemma 6. On the other hand, to have the required smoothness, since we cannot expect some smoothing effect from a non-degenerate diffusion coefficient, we have to impose that b, σ, f themselves lie in $C^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ and the restriction on the variations of σ which ensures exponential integrability in time for the expectations of the iterated tangent flows, see Lemma A.8 in [35] for details (see the parallel between the above condition on $D\sigma$ and assumption (AC_p) appearing p. 559 in [35]).
- Proving Theorem 4 under (C_{UE}) requires more sophisticated tools (Schauder estimates for operators with unbounded coefficients).

Proof of Theorem 4 under (C_{UE}) . Let us begin with the scalar case. For $d = 1$, set for any $x \in \mathbb{R}$,

$$\begin{aligned} v(x) &:= - \int_0^{+\infty} dt \mathbb{E}[\Psi(Y_t^{0,x})\partial_x Y_t^{0,x}] \\ &= - \int_0^{+\infty} dt \mathbb{E}\left[\Psi(Y_t^{0,x}) \exp\left(\int_0^t b'(Y_s^{0,x}) ds\right) \mathcal{E}\left(\int_0^t \sigma'(Y_s^{0,x}) dW_s\right)\right], \end{aligned} \tag{5.6}$$

where for any $y \in \mathbb{R}$, $\Psi(y) := \partial_y f(y)$. We observe that $\partial_x \varphi(x) = v(x)$. Also, from our assumptions on f, b, σ , we have that $\Psi, b', \sigma' \in C_b^{0,\beta}(\mathbb{R}^d, \mathbb{R})$. Theorems 2.4-2.6 in Krylov and Priola, [24] then yield the existence of a unique solution to the PDE:

$$\tilde{\mathcal{A}}w(x) + b'(x)w(x) = \Psi(x), \quad \text{where } \tilde{\mathcal{A}}w(x) = \mathcal{A}w(x) + \sigma\sigma'(x)w'(x), \tag{5.7}$$

belonging to $C_b^{2,\beta}(\mathbb{R}^d, \mathbb{R})$ and such that the following Schauder estimate holds:

$$\exists C \geq 1, \quad \|w\|_{2,\beta} \leq C(1 + \|\Psi\|_{0,\beta}), \tag{5.8}$$

where $\|\cdot\|_{2,\beta}$ and $\|\cdot\|_{0,\beta}$ are the associated norms of the Hölder spaces $C_b^{2,\beta}$ and $C_b^{0,\beta}$ defined in (1.7). Indeed, from (D_α^b) , we get that $b'(x) \leq -\alpha < 0$ and the potential in (5.7) has the good sign. From (5.6) and the Girsanov theorem, we also get:

$$v(x) = - \int_0^{+\infty} dt \mathbb{E}\left[\Psi(\tilde{Y}_t^{0,x}) \exp\left(\int_0^t b'(\tilde{Y}_s^{0,x}) ds\right)\right],$$

where $d\tilde{Y}_s^{0,x} = (b(\tilde{Y}_s^{0,x}) + \sigma\sigma'(\tilde{Y}_s^{0,x})) ds + \sigma(\tilde{Y}_s^{0,x}) dW_s$. Note that \tilde{Y} has generator $\tilde{\mathcal{A}}$. A simple identification procedure, similar to the proof of Theorem II.1.1 in Bass [5] then gives $v = w$. The result follows from (5.8). Let us emphasize that this is a quite deep and involved result for unbounded coefficients.

In the multi-dimensional setting, recalling the technical condition that for all $i \in \llbracket 1, d \rrbracket, j \geq i, \Sigma_{i,j}(x) = \Sigma_{i,j}(x_1, \dots, x_d)$, we have that differentiating formally the PDE (2.5) in the space variable $x_i, i \in \llbracket 1, d \rrbracket$ yields that $\partial_{x_i} \varphi = v_i$ should satisfy:

$$\begin{aligned} \tilde{\mathcal{A}}w_i(x) + \partial_{x_i} b_i(x)w_i(x) &= \Psi_i(x) - \sum_{j \in \llbracket 1, d \rrbracket \setminus \{i\}} \partial_{x_i} b_j(x)v_j(x) \\ &\quad - \frac{1}{2} \sum_{j \in \llbracket 1, i-1 \rrbracket} \partial_{x_i} \Sigma_{j,j}(x)\partial_{x_j} v_j(x) - \sum_{j \in \llbracket 1, i-1 \rrbracket} \sum_{k \in \llbracket j+1, d \rrbracket \setminus \{i\}} \partial_{x_i} \Sigma_{j,k}(x)\partial_{x_j} v_k(x), \end{aligned} \tag{5.9}$$

with $\Psi_i(x) := \partial_{x_i} f(x)$ and

$$\tilde{\mathcal{A}}w_i(x) := \mathcal{A}w_i(x) + \frac{1}{2} \partial_{x_i} \Sigma_{i,i}(x) \partial_{x_i} w_i(x) + \sum_{j \in \llbracket 1, d \rrbracket \setminus \{i\}} \partial_{x_i} \Sigma_{i,j}(x) \partial_{x_j} w_i(x).$$

We would now like to enter the previous framework of Schauder estimates. To do so, we first observe from (\mathbf{D}_α^p) and the Cauchy–Schwarz inequality that $\partial_{x_i} b_i(x) \leq -\alpha < 0$. Consider now $i = 1$ in (5.9). From our current assumptions on f, b and the previous computations on the gradient for the multi-dimensional case, it remains to prove $\tilde{\Psi}_1(x) := \Psi_1(x) - \sum_{j \neq 1} \partial_{x_1} b_j(x) v_j(x) \in \mathcal{C}_b^{0,\beta}(\mathbb{R}^d, \mathbb{R})$. This will be the case, once we will have proved that $\nabla \varphi$ is β -Hölder continuous, which is a priori not direct. This property is assumed for the remaining of the proof and shown below. In particular, it leads to the restriction concerning the variations of σ when $d > 1$. Hence, Theorems 2.4–2.6 in Krylov and Priola, [24] still apply and give that there exists a unique solution $w_1 \in \mathcal{C}_b^{2,\beta}(\mathbb{R}^d, \mathbb{R})$ to (5.9) which also satisfies:

$$\exists C \geq 1, \|w_1\|_{2,\beta} := \sum_{\alpha, |\alpha| \in \llbracket 0, 2 \rrbracket} \|D^\alpha w_1\|_\infty + [D^{(2)} w_1]_\beta \leq C(1 + \|\tilde{\Psi}_1\|_{0,\beta}) =: \bar{C}((\mathcal{L}_V), (\mathbf{R}_{1,\beta}), (\mathbf{UE})). \tag{5.10}$$

The identification $w_1 = \partial_{x_1} \varphi = v_1$ is standard. The control (5.10) allows to iterate, since it gives that $\nabla w_1 = (\partial_{x_1} v_1, \dots, \partial_{x_d} v_1) = (\partial_{x_1, x_1} \varphi_1, \dots, \partial_{x_d, x_1} \varphi)$ is β -Hölder. We thus get by induction, from the specific chosen structure on σ and by Theorems 2.4–2.6 in Krylov and Priola, [24], that for each $i \in \llbracket 1, d \rrbracket$ there exists a unique solution $w_i \in \mathcal{C}_b^{2,\beta}(\mathbb{R}^d, \mathbb{R})$ to (5.9) such that:

$$\begin{aligned} \exists C \geq 1, \|w_i\|_{2,\beta} &\leq C(1 + \|\tilde{\Psi}_i\|_\beta) =: \bar{C}((\mathcal{L}_V), (\mathbf{R}_{1,\beta}), (\mathbf{UE})), \\ \tilde{\Psi}_i(x) &:= \Psi_i(x) - \sum_{j \in \llbracket 1, d \rrbracket \setminus \{i\}} \partial_{x_i} b_j(x) v_j(x) - \frac{1}{2} \sum_{\substack{1 \leq j < k \\ k \in \llbracket 1, d \rrbracket \setminus \{i\}}} \partial_{x_i} \Sigma_{j,k}(x) \partial_{x_j} v_k(x). \end{aligned} \tag{5.11}$$

The Lipschitz property of the mapping $x \mapsto \langle \nabla \varphi(x), b(x) \rangle$ is eventually derived following the procedure described in Remark 7. □

Remark 12 (Structure of σ). We emphasize that the structure condition on σ assumed in Theorem 4 under $(\mathbf{C}_{\mathbf{UE}})$ is mainly technical. It is of course always verified in dimension $d = 1$. For $d > 1$ it is motivated by the fact that, differentiating (2.5) without this assumption yields to consider a system of coupled linear PDEs with growing coefficients for which the Schauder estimates have not been established yet. Following the existing literature for Schauder estimates for systems (see e.g. Boccia [8]), we think that the results of Krylov and Priola should extend to this case. This would allow to get rid of the indicated condition. Here, the condition simply allows to decouple the system.

Let us mention too that the results by Priola [40] could also be a starting point to investigate the smoothness of the Poisson problem for degenerate kinetic models.

These aspects will concern further research.

Additional Smoothness continued: β -Hölder continuity of the gradient through pathwise analysis. We control here, under (\mathbf{D}_α^p) , $p \in (1, 2]$ and $(\mathbf{R}_{1,\beta})$, $\beta \in (0, 1]$, the β -Hölder modulus of continuity of the gradient. We will progressively see how the restrictions on $D\sigma$ come out. For $(x, x') \in \mathbb{R}^{2d}$, write for each $i \in \llbracket 1, d \rrbracket$:

$$\begin{aligned} |\partial_{x_i} \varphi(x) - \partial_{x_i} \varphi(x')| &= \left| \int_0^{+\infty} (\mathbb{E}[\langle \nabla f(Y_t^{0,x}), \partial_{x_i} Y_t^{0,x} \rangle] - \mathbb{E}[\langle \nabla f(Y_t^{0,x'}), \partial_{x_i} Y_t^{0,x'} \rangle]) dt \right| \\ &\leq \left| \int_0^{+\infty} ([\nabla f]_\beta \mathbb{E}[|Y_t^{0,x} - Y_t^{0,x'}|^\beta] |\partial_{x_i} Y_t^{0,x}|] + \|\nabla f\|_\infty \mathbb{E}[|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|]) dt \right| \\ &=: (G_1^\beta + G_2^\beta)(x, x'). \end{aligned} \tag{5.12}$$

Let us first deal with the expectation in G_1^β . Namely, write

$$\mathbb{E}[|Y_t^{0,x} - Y_t^{0,x'}|^\beta |\partial_{x_i} Y_t^{0,x}|] \leq \mathbb{E}[|Y_t^{0,x} - Y_t^{0,x'}|^{\bar{p}\beta}]^{\frac{1}{\bar{p}}} \mathbb{E}[|\partial_{x_i} Y_t^{0,x}|^{\bar{q}}]^{\frac{1}{\bar{q}}}, \quad \bar{p}, \bar{q} > 1, \bar{p}^{-1} + \bar{q}^{-1} = 1.$$

Take now $\bar{p}\beta = \bar{q} \iff \bar{p} = \frac{1+\beta}{\beta}$, $\bar{q} = 1 + \beta$ which leads to the same integrability constraints on the flows.

If $\beta + 1 \leq p$ in (\mathbf{D}_α^p) , then we readily get similarly to (5.3) that $\mathbb{E}[|\partial_{x_i} Y_t^{0,x} | \bar{q}]^{\frac{1}{\bar{q}}} \leq \exp(-\alpha t)$.

If now $\beta + 1 > p$, as soon as $(\mathbf{D}_{\bar{\alpha}}^{1+\beta})$ holds for some $\bar{\alpha} > 0$, which is actually the case provided that

$$\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{1 + \beta - p}, \tag{5.13}$$

for $\bar{q} = 1 + \beta$, we again get similarly to (5.3) that $\mathbb{E}[|\partial_{x_i} Y_t^{0,x} | \bar{q}]^{\frac{1}{\bar{q}}} \leq \exp(-\bar{\alpha} t)$. On the other hand, the mean value theorem yields:

$$\begin{aligned} \mathbb{E}[|Y_t^{0,x} - Y_t^{0,x'} | \bar{p}\beta]^{\frac{1}{\bar{p}}} &\leq |x - x'|^\beta \mathbb{E}\left[\int_0^1 d\lambda \|\nabla Y_t^{0,x'+\lambda(x-x')}\| \bar{p}\beta\right]^{\frac{1}{\bar{p}}} \\ &\leq |x - x'|^\beta \left(\int_0^1 d\lambda \mathbb{E}[\|\nabla Y_t^{0,x'+\lambda(x-x')}\| \bar{p}\beta]\right)^{\frac{1}{\bar{p}}} \\ &\leq |x - x'|^\beta [\exp(-\alpha\beta t)\mathbb{I}_{1+\beta \leq p} + \exp(-\bar{\alpha}\beta t)\mathbb{I}_{1+\beta > p}], \end{aligned}$$

exploiting (5.5) for the last inequality provided that (5.13), which in turn implies that $(\mathbf{D}_{\bar{\alpha}}^{1+\beta})$ for some $\bar{\alpha} > 0$, holds if $1 + \beta > p$. Plugging these bounds in (5.12) gives that:

$$\forall (x, x') \in (\mathbb{R}^d)^2, \quad |G_1^\beta(x, x')| \leq \frac{|\nabla f|_\beta}{(1 + \beta)} \left[\frac{\mathbb{I}_{1+\beta \leq p}}{\alpha} + \frac{\mathbb{I}_{1+\beta > p}}{\bar{\alpha}} \right] |x - x'|^\beta. \tag{5.14}$$

We already see that, when $1 + \beta > p$, for the parameter p of the initial confluence condition (\mathbf{D}_α^p) , a first constraint on the variations of σ , namely (5.13) appears.

Let us now turn to G_2^β . Following the expansion of (5.2) write:

$$\begin{aligned} |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2 &= 2 \int_0^t \langle \partial_{x_i} Y_s^{0,x} - \partial_{x_i} Y_s^{0,x'}, Db(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x} - Db(Y_s^{0,x'}) \partial_{x_i} Y_s^{0,x'} \rangle ds \\ &\quad + 2 \sum_{j=1}^d \int_0^t \langle \partial_{x_i} Y_s^{0,x} - \partial_{x_i} Y_s^{0,x'}, D\sigma_{\cdot j}(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x} - D\sigma_{\cdot j}(Y_s^{0,x'}) \partial_{x_i} Y_s^{0,x'} \rangle dW_s^j \\ &\quad + \sum_{j=1}^d \int_0^t |D\sigma_{\cdot j}(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x} - D\sigma_{\cdot j}(Y_s^{0,x'}) \partial_{x_i} Y_s^{0,x'}|^2 ds. \end{aligned}$$

Let $u(t) := \mathbb{E}|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2, t \geq 0$. First note that $u(0) = 0$. Taking now the expectation and interchanging expectation and time integration yields

$$u(t) = \int_0^t \mathbb{E} \Xi_s ds,$$

where $(\Xi_t)_{t \geq 0}$ is a pathwise continuous process clearly determined by the terms inside the above time integrals. One readily checks that, $t \mapsto \mathbb{E} \Xi_s$ is continuous so that u is continuously differentiable and satisfies

$$\begin{aligned} u'(t) &= 2\mathbb{E} \langle \partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}, Db(Y_t^{0,x}) \partial_{x_i} Y_t^{0,x} - Db(Y_t^{0,x'}) \partial_{x_i} Y_t^{0,x'} \rangle \\ &\quad + \sum_{j=1}^d \mathbb{E} |D\sigma_{\cdot j}(Y_t^{0,x}) \partial_{x_i} Y_t^{0,x} - D\sigma_{\cdot j}(Y_t^{0,x'}) \partial_{x_i} Y_t^{0,x'}|^2. \end{aligned}$$

Using the Young inequality for a parameter $\varepsilon \in (0, 1]$, small enough and to be chosen further, we derive:

$$u'(t) \leq 2\mathbb{E} \left\langle \left[\frac{\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}}{|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|}, Db(Y_t^{0,x}) \frac{\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}}{|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|} \right] \right\rangle |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2$$

$$\begin{aligned}
 & + \int_0^t \|Db(Y_t^{0,x}) - Db(Y_t^{0,x'})\| |\partial_{x_i} Y_t^{0,x'}| |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}| \Big] \\
 & + \mathbb{E} \left[(1 + \varepsilon) \sum_{j=1}^d \left| D\sigma_{\cdot j}(Y_t^{0,x}) \frac{\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}}{|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|} \right|^2 |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2 \right. \\
 & \left. + (1 + \varepsilon^{-1}) \sum_{j=1}^d \|D\sigma_{\cdot j}(Y_t^{0,x}) - D\sigma_{\cdot j}(Y_t^{0,x'})\|^2 |\partial_{x_i} Y_t^{0,x'}|^2 \right].
 \end{aligned}$$

From this computation, the point is now to make the confluence condition (\mathbf{D}_α^p) appear and to separate the components for which we will exploit the β -Hölder continuity, namely $Db, (D\sigma_{\cdot j})_{j \in \llbracket 1, n \rrbracket}$. To do so we first observe that:

$$\begin{aligned}
 & \left\langle \frac{\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}}{|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|}, Db(Y_t^{0,x}) \frac{\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}}{|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|} \right\rangle + \frac{1}{2} \sum_{j=1}^d \left(\left| D\sigma_{\cdot j}(Y_t^{0,x}) \frac{\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}}{|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|} \right|^2 \right) \\
 & \leq -\alpha + (2 - p) \frac{1}{2} \|D\sigma\|_\infty^2 = -\tilde{\alpha},
 \end{aligned}$$

where we suppose from now on that

$$-\tilde{\alpha} := -\alpha + (2 - p) \frac{1}{2} \|D\sigma\|_\infty^2 < 0 \iff \|D\sigma\|_\infty^2 < \frac{2\alpha}{2 - p}. \tag{5.15}$$

Hence,

$$\begin{aligned}
 u'(t) & \leq 2\mathbb{E} \left[\left(-\tilde{\alpha} + \frac{\varepsilon}{2} \|D\sigma\|_\infty^2 \right) |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2 \right] \\
 & + 2\mathbb{E} \left[\|Db(Y_t^{0,x}) - Db(Y_t^{0,x'})\| |\partial_{x_i} Y_t^{0,x'}| |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}| \right] \\
 & + \mathbb{E} \left[(1 + \varepsilon^{-1}) \sum_{j=1}^d \|D\sigma_{\cdot j}(Y_t^{0,x}) - D\sigma_{\cdot j}(Y_t^{0,x'})\|^2 |\partial_{x_i} Y_t^{0,x'}|^2 \right].
 \end{aligned}$$

Using now again the Young inequality, with $\eta \in (0, 1]$ small enough, for the middle term of the above r.h.s., we obtain:

$$\begin{aligned}
 u'(t) & \leq 2 \left(-\tilde{\alpha} + \frac{\varepsilon}{2} \|D\sigma\|_\infty^2 + \frac{\eta}{2} \right) \mathbb{E} [|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2] \\
 & + \eta^{-1} \mathbb{E} [\|Db(Y_t^{0,x}) - Db(Y_t^{0,x'})\|^2 |\partial_{x_i} Y_t^{0,x'}|^2] \\
 & + (1 + \varepsilon^{-1}) \sum_{j=1}^d \mathbb{E} [\|D\sigma_{\cdot j}(Y_t^{0,x}) - D\sigma_{\cdot j}(Y_t^{0,x'})\|^2 |\partial_{x_i} Y_t^{0,x'}|^2] \\
 & \leq 2 \left(-\tilde{\alpha} + \frac{\varepsilon}{2} \|D\sigma\|_\infty^2 + \frac{\eta}{2} \right) \mathbb{E} [|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2] \\
 & + \eta^{-1} [Db]_\beta^2 \mathbb{E} [|Y_t^{0,x} - Y_t^{0,x'}|^{2\beta} |\partial_{x_i} Y_t^{0,x'}|^2] + (1 + \varepsilon^{-1}) [D\sigma]_\beta^2 \mathbb{E} [|Y_t^{0,x} - Y_t^{0,x'}|^{2\beta} |\partial_{x_i} Y_t^{0,x'}|^2]. \tag{5.16}
 \end{aligned}$$

Denote:

$$-\tilde{\alpha}_{\varepsilon, \eta, \sigma} := -\tilde{\alpha} + \frac{\varepsilon}{2} \|D\sigma\|_\infty^2 + \frac{\eta}{2} < 0,$$

for ε, η small enough. Setting for every $t \geq 0$,

$$r(t) := \eta^{-1} [Db]_\beta^2 \mathbb{E} [|Y_t^{0,x} - Y_t^{0,x'}|^{2\beta} |\partial_{x_i} Y_t^{0,x'}|^2] + (1 + \varepsilon^{-1}) [D\sigma]_\beta^2 \mathbb{E} [|Y_t^{0,x} - Y_t^{0,x'}|^{2\beta} |\partial_{x_i} Y_t^{0,x'}|^2],$$

hence equation (5.16) can be regarded as an ordinary differential inequation:

$$u'(t) \leq -2\tilde{\alpha}_{\varepsilon, \eta, \sigma} u(t) + r(t), \quad u(0) = 0.$$

We derive from the Gronwall lemma that

$$u(t) = \mathbb{E}[|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2] \leq \exp(-2\tilde{\alpha}_{\varepsilon,\eta,\sigma} t) \int_0^t \exp(2\tilde{\alpha}_{\varepsilon,\eta,\sigma} s) r(s) ds.$$

Reproducing as well the computations that led to (5.14), we derive:

$$u(t) \leq C_{\eta,\varepsilon,\beta} |x - x'|^{2\beta} \int_0^t \exp(-2\tilde{\alpha}_{\varepsilon,\eta,\sigma}(t-s)) \left(\mathbb{E}[|\partial_{x_i} Y_s^{0,x'}|^{2(1+\beta)}] \right)^{\frac{1}{1+\beta}} + \int_0^1 d\lambda \mathbb{E}[\|\nabla Y_s^{0,x'+\lambda(x-x')}\|^{2(1+\beta)}]^{\frac{\beta}{1+\beta}} ds.$$

From the analysis leading to (5.3), (5.5) we now derive:

$$\begin{aligned} & \mathbb{E}[|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2] \\ & \leq \frac{C_{\eta,\varepsilon,\beta}}{2} |x - x'|^{2\beta} \exp(-2\tilde{\alpha}_{\varepsilon,\eta,\sigma} t) \left[\frac{\exp(2(\tilde{\alpha}_{\varepsilon,\eta,\sigma} - \tilde{\alpha}_{2(1+\beta)})t)}{\tilde{\alpha}_{\varepsilon,\eta,\sigma} - \tilde{\alpha}_{2(1+\beta)}} + \frac{\exp(2(\tilde{\alpha}_{\varepsilon,\eta,\sigma} - \beta\tilde{\alpha}_{2(1+\beta)})t)}{\tilde{\alpha}_{\varepsilon,\eta,\sigma} - \beta\tilde{\alpha}_{2(1+\beta)}} \right], \end{aligned}$$

and

$$-\tilde{\alpha}_{2(1+\beta)} \leq -\alpha + (2(1 + \beta) - p) \frac{1}{2} \|D\sigma\|_{\infty}^2.$$

Thus, $\tilde{\alpha}_{2(1+\beta)} < 0$ as soon as

$$\|D\sigma\|_{\infty}^2 < \frac{2\alpha}{2(1 + \beta) - p}, \tag{5.17}$$

which is precisely the restriction on the variations of σ appearing in $(\mathbf{C}_{\mathbf{UE}})$ when $d > 1$, then $\tilde{\alpha}_{2(1+\beta)} > 0$ and:

$$\mathbb{E}[|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2] \leq \bar{C}_{\eta,\varepsilon,\beta} |x - x'|^{2\beta} \exp(-\tilde{\alpha}_{2(1+\beta)} t).$$

This last control then gives the expected bound for the β -Hölder modulus of the gradient. Namely, from (5.12), (5.14),

$$[\partial_{x_i} \varphi]_{\beta} < \frac{[\nabla f]_{\beta}}{(1 + \beta)} \left[\frac{\mathbb{I}_{1+\beta \leq p}}{\alpha} + \frac{\mathbb{I}_{1+\beta > p}}{\bar{\alpha}} \right] + \frac{\|\nabla f\|_{\infty} \bar{C}_{\eta,\varepsilon,\beta}}{\tilde{\alpha}_{2(1+\beta)}}.$$

5.2. Proof of the practical results of Section 2.3

We first begin with the proof of the

5.2.1. Slutsky like Theorem 6

We keep here for simplicity the generic notation $\|\cdot\|$ for any admissible matrix norm according to the assumptions of the theorem. We first write:

$$\mathbb{P}\left[\frac{\sqrt{\Gamma_n} v_n(f) - v(f)}{\sqrt{v_n(\|\sigma\|^2)}} \geq a\right] = \mathbb{P}\left[v_n(\mathcal{A}\varphi) \geq \frac{a}{\sqrt{\Gamma_n}} \sqrt{v_n(\|\sigma\|^2)}\right]. \tag{5.18}$$

We then proceed similarly to Theorem 8, with an exponential Bienaymé–Tchebychev inequality, for any $\lambda > 0$ we have:

$$\begin{aligned} & \mathbb{P}\left[\frac{\sqrt{\Gamma_n} v_n(f) - v(f)}{\sqrt{v_n(\|\sigma\|^2)}} \geq a\right] \\ & \leq \mathbb{E}\left[\exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}} \sqrt{v_n(\|\sigma\|^2)}\right) \exp(\lambda v_n(\mathcal{A}\varphi))\right] \\ & = \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}} \sqrt{v(\|\sigma\|^2)}\right) \mathbb{E}\left[\exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}} [\sqrt{v_n(\|\sigma\|^2)} - \sqrt{v(\|\sigma\|^2)}]\right) \exp(\lambda v_n(\mathcal{A}\varphi))\right] \end{aligned}$$

$$\begin{aligned} &= \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\sqrt{\nu(\|\sigma\|^2)}\right)\mathbb{E}\left[\exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\frac{\nu_n(\|\sigma\|^2)-\nu(\|\sigma\|^2)}{\sqrt{\nu_n(\|\sigma\|^2)}+\sqrt{\nu(\|\sigma\|^2)}}\right)\exp(\lambda\nu_n(\mathcal{A}\varphi))\right] \\ &= \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\sqrt{\nu(\|\sigma\|^2)}\right)\mathbb{E}\left[\exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu_n(\|\sigma\|^2)}+\sqrt{\nu(\|\sigma\|^2)}}\right)\exp(\lambda\nu_n(\mathcal{A}\varphi))\right]. \end{aligned}$$

By the Hölder inequality, for $\tilde{p}, \tilde{q} > 1$, such that $\frac{1}{\tilde{p}} + \frac{1}{\tilde{q}} = 1$:

$$\begin{aligned} &\mathbb{P}\left[\sqrt{\Gamma_n}\frac{\nu_n(f)-\nu(f)}{\sqrt{\nu_n(\|\sigma\|^2)}}\geq a\right] \\ &\leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\sqrt{\nu(\|\sigma\|^2)}\right)\left[\mathbb{E}\exp\left(-\frac{a\tilde{p}\lambda}{\sqrt{\Gamma_n}}\frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu_n(\|\sigma\|^2)}+\sqrt{\nu(\|\sigma\|^2)}}\right)\right]^{1/\tilde{p}}\left[\mathbb{E}\exp(\lambda\tilde{q}\nu_n(\mathcal{A}\varphi))\right]^{1/\tilde{q}}. \end{aligned}$$

The proof of Theorem 8 yields:

$$\begin{aligned} \mathbb{P}\left[\sqrt{\Gamma_n}\frac{\nu_n(f)-\nu(f)}{\sqrt{\nu_n(\|\sigma\|^2)}}\geq a\right] &\leq \mathcal{R}_n \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\sqrt{\nu(\|\sigma\|^2)}\right)\exp\left(\frac{\rho\tilde{q}\lambda^2}{\Gamma_n}\tilde{A}_n + \frac{\rho^3\tilde{q}^3\lambda^4}{(\rho-1)\Gamma_n}\tilde{B}_n\right) \\ &\quad \times \left(\mathbb{E}\exp\left(-\frac{a\tilde{p}\lambda}{\sqrt{\Gamma_n}}\frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu_n(\|\sigma\|^2)}+\sqrt{\nu(\|\sigma\|^2)}}\right)\right)^{1/\tilde{p}}, \end{aligned} \tag{5.19}$$

where we recall from identity (4.5):

$$\tilde{A}_n = \frac{q[\varphi]_1^2\nu(\|\sigma\|^2)}{2} + e_n \quad \text{and} \quad \tilde{B}_n = \frac{q^3[\varphi]_1^4}{4}\left(\frac{\bar{q}\|\sigma\|_\infty^2[\vartheta]_1^2}{2} + \bar{e}_n\right).$$

Also, $\mathcal{R}_n \xrightarrow[n]{}$ 1 denotes a “generic” remainder. Observe that thanks to the bounds of Theorem 4 (stated in the above Lemma 6), we get:

$$\tilde{A}_n \leq \frac{q[f]_1^2\nu(\|\sigma\|^2)}{2\alpha} + e_n. \tag{5.20}$$

Let us now handle the remainder $[\mathbb{E}\exp(-\frac{a\tilde{p}\lambda}{\sqrt{\Gamma_n}}\frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu_n(\|\sigma\|^2)}+\sqrt{\nu(\|\sigma\|^2)}})]^{1/\tilde{p}}$:

$$\begin{aligned} &\left[\mathbb{E}\exp\left(-\frac{a\tilde{p}\lambda}{\sqrt{\Gamma_n}}\frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu_n(\|\sigma\|^2)}+\sqrt{\nu(\|\sigma\|^2)}}\right)\right]^{1/\tilde{p}} \\ &= \left[\mathbb{E}\left(\exp\left[-\frac{a\tilde{p}\lambda}{\sqrt{\Gamma_n}}\frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu_n(\|\sigma\|^2)}+\sqrt{\nu(\|\sigma\|^2)}}\right](\mathbf{1}_{\nu_n(\mathcal{A}\vartheta)\geq 0} + \mathbf{1}_{\nu_n(\mathcal{A}\vartheta)< 0})\right)\right]^{1/\tilde{p}} \\ &\leq \left(\left[\mathbb{E}\exp\left(\frac{a\tilde{p}^2\lambda}{\sqrt{\Gamma_n}}\frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu(\|\sigma\|^2)}}\right)\right]^{1/\tilde{p}}\mathbb{P}[\nu_n(\mathcal{A}\vartheta)\geq 0]^{1/\tilde{q}}\right. \\ &\quad \left. + \left[\mathbb{E}\exp\left(-\frac{a\tilde{p}^2\lambda}{\sqrt{\Gamma_n}}\frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu(\|\sigma\|^2)}}\right)\right]^{1/\tilde{p}}\mathbb{P}[\nu_n(\mathcal{A}\vartheta)< 0]^{1/\tilde{q}}\right)^{1/\tilde{p}}. \end{aligned}$$

Let us mention that we introduced the above partition in order to get a sharper constant in the final inequality, 1 below instead of 2, which would follow getting rid of the indicator functions. Now, by Theorem 2 we easily get:

$$\begin{aligned} &\mathbb{E}\left[\exp\left(-\frac{a\tilde{p}\lambda}{\sqrt{\Gamma_n}}\frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu_n(\|\sigma\|^2)}+\sqrt{\nu(\|\sigma\|^2)}}\right)\right]^{1/\tilde{p}} \\ &\leq \mathcal{R}_n \exp\left(\frac{a^2\tilde{p}^2}{\Gamma_n\nu(\|\sigma\|^2)}\frac{\lambda^2}{\Gamma_n}\left(\frac{\|\sigma\|_\infty^2\|\nabla\vartheta\|_\infty^2}{2} + e_n\right)\right)\left[\mathbb{P}[\nu_n(\mathcal{A}\vartheta)\geq 0]^{1/\tilde{q}} + \mathbb{P}[\nu_n(\mathcal{A}\vartheta)< 0]^{1/\tilde{q}}\right]^{1/\tilde{p}}. \end{aligned} \tag{5.21}$$

We choose $\tilde{p} := \tilde{p}(n) \rightarrow +\infty$, such that $\frac{\tilde{p}^2 a^2}{\Gamma_n} \rightarrow 0$, and so $(\mathbb{P}[v_n(\mathcal{A}\vartheta) \geq 0]^{1/\tilde{q}} + \mathbb{P}[v_n(\mathcal{A}\vartheta) < 0]^{1/\tilde{q}})^{1/\tilde{p}} \leq 2^{1/\tilde{p}} \rightarrow 1$. Moreover, exploiting again that for the *Gaussian regime*, $\frac{\tilde{p}^2 a^2}{\Gamma_n} \rightarrow 0$, we obtain by (5.19) and (5.21):

$$\mathbb{P}\left[\sqrt{\Gamma_n} \frac{v_n(f) - v(f)}{\sqrt{v_n(\|\sigma\|^2)}} \geq a\right] \leq \mathcal{R}_n \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}} \sqrt{v(\|\sigma\|^2)}\right) \exp\left(\frac{\rho\tilde{q}\lambda^2}{\Gamma_n} (\tilde{A}_n + e_n) + \frac{\rho^3\tilde{q}^3\lambda^4}{(\rho-1)\Gamma_n} \tilde{B}_n\right). \tag{5.22}$$

From identity (5.22), the optimization over λ is similar to the one performed in the proof of Theorem 8. This yields the deviation bound (2.10). The non-asymptotic confidence interval in (2.11) is derived as for Theorem 5 from the gradient bounds of Theorem 4 and (2.10).

5.3. Regularization of Lipschitz sources

We assume here that assumptions **(C2)**, **(L ν)**, **(UE)** are in force. We suppose as well that the following smoothness holds for b, σ :

(R $_{b,\sigma}$) Regularity and Structure. We assume that there exists $\beta \in (0, 1)$ such that b, σ in (1.1) belong to $C^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d)$ and $C^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)$ respectively. Also, for each $(i, j) \in \llbracket 1, d \rrbracket^2$, $\Sigma_{i,j}(x) = \Sigma_{i,j}(x_{i \wedge j}, \dots, x_d)$.

Importantly, we are interested, under assumptions **(C2)**, **(L ν)**, **(UE)**, **(R $_{b,\sigma}$)**, in giving controls for the estimation of $v(f)$ when the source f is simply Lipschitz continuous. This is indeed the natural framework for the source which can be handled through functional inequality techniques, see [9,30].

To comply with our previous framework, namely to exploit the smoothness result of Theorem 4 under **(C $_{UE}$)**, we need to regularize the source. Let η be a mollifier (i.e. a non-negative compactly supported function such that $\int_{\mathbb{R}^d} \eta(x) dx = 1$). Define for $\delta > 0$, $\eta_\delta(x) = \frac{1}{\delta^d} \eta(\frac{x}{\delta})$. We regularize f introducing $f_\delta := f \star \eta_\delta$ where \star stands for the convolution on \mathbb{R}^d . From usual estimates, we obtain:

$$\begin{aligned} \exists C_\eta > 0, \forall x \in \mathbb{R}^d, \quad |f_\delta(x) - f(x)| &\leq C_\eta \delta [f]_1, \\ \forall \beta \in (0, 1), \quad [\nabla f_\delta]_\beta &\leq C_\eta [f]_1 \delta^{-\beta}. \end{aligned} \tag{5.23}$$

We emphasize here that we will choose β later in order to be compatible with a certain range of step sequences. We assume for simplicity that $\theta \in (1/3, 1]$ (no bias). Recall that we want to investigate:

$$\begin{aligned} \mathbb{P}\left[\sqrt{\Gamma_n} (v_n(f) - v(f)) \geq a\right] &= \mathbb{P}\left[(v_n(f_\delta) - v(f_\delta)) + R_{n,\delta}(f) \geq \frac{a}{\sqrt{\Gamma_n}}\right], \\ R_{n,\delta}(f) &:= [(v_n(f) - v(f)) - (v_n(f_\delta) - v(f_\delta))]. \end{aligned} \tag{5.24}$$

From (5.23), one readily gets:

$$|R_{n,\delta}(f)| \leq 2C_\eta \delta [f]_1. \tag{5.25}$$

On the other hand, the coefficients b, σ and the source f_δ satisfy assumption **(R $_{1,\beta}$)** (observe indeed that the mollified function $f_\delta \in C^{1,\beta}(\mathbb{R}^d, \mathbb{R})$). Hence, Theorem 4 yields that there exists a unique solution $\varphi_\delta \in C^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ to the equation:

$$\mathcal{A}\varphi_\delta = f_\delta - v(f_\delta). \tag{5.26}$$

Observe from the proof of Theorem 4 under **(C $_{UE}$)** (see equations (5.4) and (5.11)) and (5.23) that:

$$\begin{aligned} \|\nabla \varphi_\delta\|_\infty &\leq \alpha^{-1} [f]_1, \quad \forall \beta \in (0, 1), \exists C_\beta > 0, \forall i \in \{1, 2\}, \\ [\varphi_\delta^{(i)}]_1 &\leq C_\beta (1 + \|\nabla f_\delta\|_{C^\beta}) \leq C_\beta \delta^{-\beta}, \quad [\varphi_\delta^{(3)}]_\beta \leq C_\beta \delta^{-\beta}, \\ [(\nabla \varphi_\delta, b)]_1 &\leq C_\beta \delta^{-\beta}. \end{aligned} \tag{5.27}$$

Now, from (5.26) the deviation in (5.24) rewrites:

$$\mathbb{P}\left[\sqrt{\Gamma_n} (v_n(f) - v(f)) \geq a\right] = \mathbb{P}\left[v_n(\mathcal{A}\varphi_\delta) + R_{n,\delta}(f) \geq \frac{a}{\sqrt{\Gamma_n}}\right]. \tag{5.28}$$

From (5.25), the term $R_{n,\delta}(f)$ can be seen as a remainder as soon as $\frac{a}{\sqrt{\Gamma_n}} \gg 2C_\eta\delta[f]_1 \geq |R_{n,\delta}(f)|$. On the other hand, the deviations of $v_n(\mathcal{A}\varphi_\delta)$ can be analyzed as above, reproducing the proofs of Theorems 2 and 8, replacing the bounds on $([\varphi^{(i)}]_1)_{i \in \{1,2\}}$, $[\varphi^{(3)}]_\beta$ appearing therein by those of equation (5.27). Precisely, we get from (5.25), similarly to (3.18) (replacing the controls on φ by those on φ_δ in the proofs of Lemmas 3 and 5):

$$\begin{aligned} \mathbb{P}\left[|v_n(\mathcal{A}\varphi_\delta) + R_{n,\delta}(f)| \geq \frac{a}{\sqrt{\Gamma_n}}\right] &\leq 2\left[\mathbb{E}\exp\left(-\frac{q\lambda_n}{\Gamma_n}M_n\right)\right]^{\frac{1}{q}} \exp\left(-\frac{a\lambda_n}{\sqrt{\Gamma_n}}\left(1 - \frac{\sqrt{\Gamma_n}2C_\eta[f]_1\delta}{a}\right)\right) \\ &\quad \times \exp\left(\frac{\lambda_n^2}{2\Gamma_n p} + \frac{p(a_n^\delta)^2}{2}\right) \exp\left(\frac{3pC_{V,\varphi}^2\lambda_n^2}{c_V\Gamma_n^2} + \frac{c_V}{p}\right)(I_V^1)^{\frac{1}{2p}} \\ &\quad \times \exp\left(C_{3.14}^\delta \frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{\Gamma_n^2}\right)(I_V^1)^{\frac{1}{4p}} \times \exp\left(C_{3.15}^\delta \frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{\Gamma_n^2}\right)(I_V^1)^{\frac{1}{4p}}, \end{aligned} \tag{5.29}$$

where $C_{3.15}^\delta := C_\beta\delta^{-2\beta} \frac{(C\sqrt{c_V})^2}{c_V}$ and $C_{3.14}^\delta := \frac{\|\sigma\|_\infty^4 C_\beta^2 \delta^{-2\beta}}{4} \frac{c_V}{c_V}$ precisely correspond to the modifications of the constants $C_{3.15} = 2 \frac{(C\sqrt{c_V})^2}{c_V}$ with $C = \frac{1}{2}(\|\nabla\varphi\|_\infty[b]_1 + \langle \nabla\varphi, b \rangle_1)$ and $C_{3.14} = \frac{\|\sigma\|_\infty^4 [\varphi^{(2)}]_1^2}{4} \frac{c_V}{c_V}$ introduced in the proof of Lemma 5 when replacing $\|D^2\varphi\|_\infty$ by $\|D^2\varphi_\delta\|_\infty \leq C_\beta\delta^{-\beta}$ and $\langle \nabla\varphi, b \rangle_1 \leq C$ by $\langle \nabla\varphi_\delta, b \rangle_1 \leq C_\beta\delta^{-\beta}C$. Similarly,

$$a_n^\delta := \frac{[\varphi_\delta^{(3)}]_\beta \|\sigma\|_\infty^{(3+\beta)} \mathbb{E}[|U_1|^{3+\beta}] \Gamma_n^{(\frac{3+\beta}{2})}}{(1+\beta)(2+\beta)(3+\beta) \sqrt{\Gamma_n}} \leq \frac{C_\beta\delta^{-\beta} \|\sigma\|_\infty^{(3+\beta)} \mathbb{E}[|U_1|^{3+\beta}] \Gamma_n^{(\frac{3+\beta}{2})}}{(1+\beta)(2+\beta)(3+\beta) \sqrt{\Gamma_n}},$$

is obtained from the definition of a_n in Lemma 3 replacing $[\varphi^{(3)}]_\beta$ by $[\varphi_\delta^{(3)}]_\beta$. From the above equation and Lemma 2 we get similarly to (3.18):

$$\begin{aligned} \mathbb{P}\left[|v_n(f) - v(f)| \geq \frac{a}{\sqrt{\Gamma_n}}\right] &= \mathbb{P}\left[|v_n(\mathcal{A}\varphi_\delta) + R_{n,\delta}(f)| \geq \frac{a}{\sqrt{\Gamma_n}}\right] \\ &\leq 2(I_V^1)^{\frac{1}{p}} \exp\left(\frac{c_V}{p} + \frac{p(a_n^\delta)^2}{2}\right) \exp\left(-\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi_\delta\|_\infty^2}\left(1 - \frac{\sqrt{\Gamma_n}4C_\eta[f]_1\delta}{a}\right.\right. \\ &\quad \left.\left.- \frac{1}{q\|\sigma\|_\infty^2\|\nabla\varphi_\delta\|_\infty^2}\left\{\frac{p}{\Gamma_n}\left(\frac{6C_{V,\varphi}^2}{c_V} + 2[C_{3.14}^\delta + C_{3.15}^\delta](\Gamma_n^{(2)})^2\right) + \frac{1}{p}\right\}\right)\right), \end{aligned}$$

where $C_\eta > 0$ is defined in (5.23). The Young inequality yields that for any $\varepsilon_n > 0$:

$$\begin{aligned} \mathbb{P}\left[|v_n(f) - v(f)| \geq \frac{a}{\sqrt{\Gamma_n}}\right] &\leq 2(I_V^1)^{\frac{1}{p}} \exp\left(\frac{c_V}{p} + \frac{p(a_n^\delta)^2}{2} + \varepsilon_n^{-1}\Gamma_n\delta^2\right) \exp\left(-\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi_\delta\|_\infty^2}\right. \\ &\quad \left.\times \underbrace{\left(1 - \frac{1}{q\|\sigma\|_\infty^2\|\nabla\varphi_\delta\|_\infty^2}\left\{2\varepsilon_n C_\eta^2[f]_1^2 + \frac{p}{\Gamma_n}\left(\frac{6C_{V,\varphi}^2}{c_V} + 2[C_{3.14}^\delta + C_{3.15}^\delta](\Gamma_n^{(2)})^2\right) + \frac{1}{p}\right\}\right)}_{=: d_n^\delta}\right). \end{aligned} \tag{5.30}$$

We now want to let $p := p(n) \rightarrow +\infty$, $\varepsilon_n \rightarrow 0$ so that the associated contributions in the above equation can be viewed as remainders. From the previous definitions of $C_{3.15}^\delta$, $C_{3.14}^\delta$, we see that, to achieve this goal, two constraints need to be fulfilled: namely, we must choose δ , p such that

$$\varepsilon_n^{-1}\Gamma_n\delta^2 \xrightarrow{n} 0 \quad \text{and} \quad p(a_n^\delta)^2 \xrightarrow{n} 0.$$

If $\theta \in (1/2, 1]$ there exists $\beta \in (0, 1)$ such that $\Gamma_n^{(\frac{3+\beta}{2})} \leq C$. In that case: $a_n^\delta \leq \frac{C}{\sqrt{\Gamma_n}} \delta^{-\beta} = \Gamma_n^{-(\frac{1}{2}(1-\beta)-\beta\varepsilon)} \xrightarrow[n]{} 0$ for $\delta = \Gamma_n^{-(\frac{1}{2}+\varepsilon)}$ and $\varepsilon < \frac{1-\beta}{2\beta}$. Taking $p := p(n) = \Gamma_n^{(\frac{1}{2}(1-\beta)-\beta\varepsilon)}$ yields $p(a_n^\delta)^2 \xrightarrow[n]{} 0$. On the other hand, $\varepsilon_n = \Gamma_n^{-\varepsilon}$ also yields $\varepsilon_n^{-1} \Gamma_n \delta^2 = \Gamma_n^{-\varepsilon} \xrightarrow[n]{} 0$.

For $\theta \in (1/3, 1/2)$, $\Gamma_n^{(\frac{3+\beta}{2})}$ diverges for any $\beta \in (0, 1)$, we then have $\frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} \leq C n^{\frac{1}{2}-\theta(1+\frac{\beta}{2})}$. Hence, there exists $\beta \in (0, 1)$ such that $\frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} \leq C n^{\frac{1}{2}-\theta(1+\frac{\beta}{2})} \xrightarrow[n]{} 0$. However, taking $\delta = \Gamma_n^{-(\frac{1}{2}+\varepsilon)}$, which seems to be an almost “necessary” choice to satisfy the first constraint $\varepsilon_n^{-1} \Gamma_n \delta^2 \xrightarrow[n]{} 0$, yields:

$$a_n^\delta \asymp \delta^{-\beta} \frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} \asymp n^{(1+\beta)(\frac{1}{2}-\theta)+\varepsilon\beta(1-\theta)} \xrightarrow[n]{} +\infty,$$

so that the second constraint cannot be fulfilled. This means that the regularization induces a constraint on the time steps which must not be too large. In other words, under the sole Lipschitz assumption on the source f , the fastest convergence regime is out of reach.

Summing up the previous computations, we complete the proof of Theorem 7.

6. Applications

6.1. Non-asymptotic deviation bounds in the almost sure CLT

Let $(U_n)_{n \geq 1}$ be an i.i.d sequence of centered d -dimensional random variables with unit covariance matrix. We define the sequence of normalized partial sums by $Z_0 = 0$ and

$$Z_n := \frac{\sum_{k=1}^n U_k}{\sqrt{n}}, \quad n \geq 1.$$

The almost sure Central Limit Theorem (denoted from now on *a.s. CLT*) describes how the weighted sum of the renormalized sums Z_n which appear in the *usual asymptotic CLT*, behaves viewed as a random measure. Precisely, it states that setting for $k \geq 1$, $\gamma_k = 1/k$:

$$v_n^Z := \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k \delta_{Z_k} \xrightarrow[n]{w, a.s.} G, \quad G(dx) := \exp\left(-\frac{|x|^2}{2}\right) \frac{dx}{(2\pi)^{d/2}}. \tag{6.1}$$

The above convergence had been established in [25], as a by-product of their results concerning the approximation of invariant distributions, under the minimal moment condition $U_i \in L^2(\mathbb{P})$, thus weakening the initial assumptions by Brosamler and Schatte (see [11] and [42]). The underlying idea is to use a reformulation of the dynamics of $(Z_n)_{n \geq 0}$ in terms of a discretization scheme appearing as a perturbation of (S). One indeed easily checks that, for $n \geq 0$:

$$Z_{n+1} = Z_n - \frac{\gamma_{n+1}}{2} Z_n + \sqrt{\gamma_{n+1}} U_{n+1} + r_n Z_n, \quad r_n := \sqrt{1 - \frac{1}{n+1}} - 1 + \frac{1}{2(n+1)} = O\left(\frac{1}{n^2}\right). \tag{6.2}$$

Thus, the sequence $(Z_n)_{n \geq 0}$ appears as a perturbed Euler scheme with decreasing step $\gamma_n = \frac{1}{n}$ of the Ornstein–Uhlenbeck process $dX_t = -\frac{1}{2}X_t dt + dW_t$ whose invariant distribution is G . Then the regular Euler scheme

$$X_{n+1} = X_n - \frac{\gamma_{n+1}}{2} X_n + \sqrt{\gamma_{n+1}} U_{n+1}, \tag{6.3}$$

satisfies (1.3) with $\nu = G$. The *a.s.* weak convergence (6.1) established in [25] follows as a consequence of the (fast enough) convergence of Z_n towards X_n as n goes to infinity.

Moreover, this rate is fast enough to guarantee that the sequence v_n^Z satisfies the conclusion of Theorem 1 point (a) (when $\gamma_n = \frac{1}{n}$, $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \xrightarrow[n]{} 0$), *i.e.* its convergence rate is ruled by a CLT at rate $\sqrt{\log(n)}$. In fact this holds under a lower moment assumption $U_1 \in L^3(\mathbb{P})$.

Let us mention that the convergence rates related to the *a.s.* CLT had already been investigated by several authors. Let us quote among relevant works, Csörgő and Horváth [13], for real valued i.i.d. random variables, Chaâbane and Maâouia [12], who investigate the convergence rate of the strong quadratic law of large numbers for some extensions to vector-valued martingales, and Heck [19], for large deviation results. As an application of our previous results, we will derive some new non-asymptotic Gaussian deviation bounds for the *a.s.* CLT, when the involved random variables $(U_n)_{n \geq 1}$ satisfy **(GC)**. We insist here that the sub-Gaussianity of the innovations is crucial to get a non-asymptotic Gaussian deviation bound. The result readily extends to the wider class of innovations satisfying the general sub-Gaussian exponential deviation inequality (1.4). Also, we slightly weaken the regularity assumptions needed on the function f in [25] for the associated *a.s.* CLT to hold.

6.1.1. *Non-asymptotic deviation bounds*

Theorem 9. *Assume the innovation sequence $(U_n)_{n \geq 1}$ satisfies **(GC)** and let f be a Lipschitz continuous function such that $G(f) = \int_{\mathbb{R}^d} f(x)G(dx) = 0$. Then, there exist two explicit monotonic sequences $c_n \leq 1 \leq C_n, n \geq 1$, with $\lim_n C_n = \lim_n c_n = 1$ such that for all $a > 0$ and $n \geq 1$:*

$$\mathbb{P}[\sqrt{\log(n)+1}|v_n^Z(f)| \geq a] \leq 2C_n \exp\left(-c_n \frac{a^2}{2\|\nabla\varphi\|_\infty^2}\right), \tag{6.4}$$

where φ denotes the solution of the Poisson equation:

$$\forall x \in \mathbb{R}^d, \quad \frac{1}{2}\Delta\varphi(x) - \frac{1}{2}x \cdot \nabla\varphi(x) = f(x), \tag{6.5}$$

which, under the current assumptions, is unique and belongs to $W_{p,\text{loc}}^2(\mathbb{R}^d, \mathbb{R})$, for any $p > 1$, with $\|\nabla\varphi\|_\infty \leq 2[f]_1$.

Proof. For $(Z_n)_{n \geq 0}$ as in (6.2), and $(X_n)_{n \geq 0}$ as in (6.3) we introduce:

$$\Delta_n := Z_n - X_n.$$

With the definition of v_n^Z in (6.1), write $v_n^Z(f) = \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k f(Z_{k-1})$. We also have similarly $v_n^X(f) := \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k \times f(X_{k-1})$. For any $\lambda > 0$, we derive similarly to (3.9) (see as well (5.29)) and with the notations of (5.24):

$$\begin{aligned} &\mathbb{P}[\sqrt{\Gamma_n}|v_n^Z(f)| \geq a] \\ &= \mathbb{P}\left[\sqrt{\Gamma_n}\left|\frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k (f(Z_{k-1}) - f(X_{k-1})) + v_n^X(f)\right| \geq a\right] \\ &\leq \mathbb{P}\left[\sqrt{\Gamma_n}\left|\frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k (f(Z_{k-1}) - f(X_{k-1})) + v_n^X(\mathcal{A}\varphi_\delta) + R_{n,\delta}(f)\right| \geq a\right] \\ &\leq 2 \exp\left(-\frac{\lambda a}{\sqrt{\Gamma_n}}\left(1 - \frac{2\sqrt{\Gamma_n}C_\eta[f]_1\delta}{a}\right)\right) \left(\mathbb{E} \exp(\bar{p}[f]_1 \lambda v_n^\Delta(|\cdot|))\right)^{\frac{1}{\bar{p}}} \left(\mathbb{E} \exp\left(-\frac{q\bar{q}\lambda}{\Gamma_n} M_n\right)\right)^{\frac{1}{q\bar{q}}} \\ &\quad \times \left(\mathbb{E} \exp\left(\frac{2p\bar{q}\lambda}{\Gamma_n} (|L_n| + |\bar{G}_n|)\right)\right)^{\frac{1}{2p\bar{q}}} \left(\mathbb{E} \exp\left(\frac{4p\bar{q}\lambda}{\Gamma_n} |D_{2,b,n}|\right)\right)^{\frac{1}{4p\bar{q}}} \left(\mathbb{E} \exp\left(\frac{4p\bar{q}\lambda}{\Gamma_n} |D_{2,\Sigma,n}|\right)\right)^{\frac{1}{4p\bar{q}}} \end{aligned} \tag{6.6}$$

for $\bar{q}, q \in (1, +\infty)$, $\bar{p} = \frac{\bar{q}}{\bar{q}-1}$, $p = \frac{q}{q-1}$. Also, φ_δ corresponds to the solution of the Poisson equation (6.5) obtained replacing f by its mollified version f_δ . Now, we need the following lemma to control $v_n^\Delta(|\cdot|) := \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k |\Delta_{k-1}|$.

Lemma 7. *There is a non-negative constant $C_{6.7}$ such that for any $\lambda > 0$:*

$$\mathbb{E} \exp(\lambda v_n^\Delta(|\cdot|)) = \mathbb{E} \exp\left(\frac{\lambda}{\Gamma_n} \sum_{k=1}^n \gamma_k |\Delta_{k-1}|\right) \leq \exp\left(\frac{C_{6.7}\lambda \mathbb{E}[|U_1|] \Gamma_n^{\frac{3}{2}}}{\Gamma_n} + \frac{C_{6.7}^2 \lambda^2 \Gamma_n^{(3)}}{2\Gamma_n^2}\right). \tag{6.7}$$

For clarity, we postpone the proof to the end of the current section.

On the other hand, from Section 5.3 we have that $\varphi_\delta \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ for any $\beta \in (0, 1)$. We derive from (6.6), (6.7) similarly to the proof of Theorem 7 by setting $\bar{\lambda}_n := \frac{a\sqrt{\Gamma_n}}{q\bar{q}\|\nabla\varphi\|_\infty^2}$, recalling that $\|\nabla\varphi\|_\infty = \|\nabla\varphi_\delta\|_\infty$:

$$\begin{aligned} \mathbb{P}[\sqrt{\Gamma_n}|v_n^Z(f)| \geq a] &\leq 2 \exp\left(-\frac{a^2}{2q\bar{q}\|\nabla\varphi_\delta\|_\infty^2}\left(1 - \frac{4\sqrt{\Gamma_n}C_\eta[f]_1\delta}{a}\right)\right) \exp\left(\frac{C_{6.7}\bar{\lambda}_n[f]_1\mathbb{E}[|U_1|]\Gamma_n^{(\frac{3}{2})}}{\Gamma_n}\right) \\ &\quad \times \exp\left(\frac{C_{6.7}^2\bar{p}[f]_1^2\bar{\lambda}_n^2\Gamma_n^{(3)}}{2\Gamma_n^2}\right) (I_V^1)^{\frac{1}{p\bar{q}}} \exp\left(\frac{c_V}{p\bar{q}} + \frac{p(a_n^\delta)^2}{2\bar{q}}\right) \\ &\quad \times \exp\left(\bar{\lambda}_n\left(p\bar{q}\left(\frac{3C_{V,\varphi}^2}{c_V\Gamma_n^2} + [C_{3.14}^\delta + C_{3.15}^\delta]\frac{(\Gamma_n^{(2)})^2}{\Gamma_n^2}\right) + \frac{1}{2p\bar{q}}\right)\right) \\ &\leq 2(I_V^1)^{\frac{1}{p\bar{q}}} \exp\left(\frac{c_V}{p\bar{q}} + \frac{p(a_n^\delta)^2}{2\bar{q}} + \varepsilon_n^{-1}\Gamma_n\delta^2\right) \\ &\quad \times \exp\left(-\frac{a^2}{2q\bar{q}\|\nabla\varphi\|_\infty^2}\left(1 - d_n^\delta - \frac{\bar{p}}{q\bar{q}\|\nabla\varphi\|_\infty^2}\frac{[f]_1^2C_{6.7}^2(\Gamma_n^{(3)} + \mathbb{E}[|U_1|]^2(\Gamma_n^{(\frac{3}{2})})^2)}{\Gamma_n}\right)\right), \end{aligned}$$

for $\varepsilon_n > 0$ and d_n^δ as in (5.30). Choose again $(p_n)_{n \geq 1}$ and δ as in Section 5.3 so that $q_n \xrightarrow{n} 1$, $d_n^\delta \xrightarrow{n} 0$ with the indicated monotonicity for n large enough. We can now take $\bar{p} := \bar{p}_n \xrightarrow{n} +\infty$ such that $\frac{\bar{p}}{\Gamma_n} \xrightarrow{n} 0$. The above inequality then gives the result up to a direct modification of the sequences $(C_n)_{n \geq 1}$, $(c_n)_{n \geq 1}$. \square

Proof of Lemma 7. The definition of Δ_n implies:

$$\Delta_{n+1} = \Delta_n \left(1 - \frac{\gamma_{n+1}}{2}\right) + r_n Z_n,$$

where we recall from (6.2) that $r_n := \sqrt{1 - \frac{1}{n+1}} - 1 + \frac{1}{2(n+1)} = O(\frac{1}{n^2})$. In particular, there exists $\bar{C}_1 > 0$ such that for each $n \geq 1$,

$$|r_n| \leq \frac{\bar{C}_1}{n^2}. \tag{6.8}$$

Setting now $\rho_0 = 1$ and for $n \geq 1$:

$$\rho_n := \left[\prod_{k=1}^n \left(1 - \frac{\gamma_k}{2}\right)\right]^{-1} = \prod_{k=1}^n \frac{2k}{2k-1},$$

a direct induction on Δ_n yields:

$$\Delta_n = \frac{1}{\rho_n} \sum_{k=1}^n r_k \rho_k Z_k = \frac{1}{\rho_n} \sum_{k=1}^n r_k \rho_k \left(\sum_{l=1}^k \frac{U_l}{\sqrt{k}}\right) = \frac{1}{\rho_n} \sum_{l=1}^n \left(\sum_{k=l}^n \frac{r_k \rho_k}{\sqrt{k}}\right) U_l. \tag{6.9}$$

Also, from the Wallis formula $\rho_n \sim_n \sqrt{\pi n}$, which implies that there exists $\bar{C}_2 \geq 1$ such that for each $n \geq 1$:

$$\bar{C}_2^{-1}\sqrt{n} \leq \rho_n \leq \bar{C}_2\sqrt{n}. \tag{6.10}$$

We now get from (6.9) and the Fubini theorem:

$$\Gamma_n v_n^\Delta(|\cdot|) = \sum_{k=1}^n \gamma_k |\Delta_{k-1}| \leq \sum_{k=1}^n \frac{\gamma_k}{\rho_{k-1}} \sum_{l=1}^{k-1} \left(\sum_{m=l}^{k-1} \frac{|r_m| \rho_m}{\sqrt{m}}\right) |U_l| = \sum_{l=1}^{n-1} \left[\sum_{k=l+1}^n \frac{\gamma_k}{\rho_{k-1}} \left(\sum_{m=l}^{k-1} \frac{|r_m| \rho_m}{\sqrt{m}}\right)\right] |U_l|. \tag{6.11}$$

Combining (6.8) and (6.10), we get that there exist constants $\bar{C}_3, \bar{C}_4 > 0$ such that for each $k \in \llbracket l+1, n \rrbracket$.

$$\frac{\gamma_k}{\rho_{k-1}} \sum_{m=l}^{k-1} \frac{|r_m| \rho_m}{\sqrt{m}} \leq \frac{\bar{C}_3}{k^{3/2}l}, \quad \sum_{k=l+1}^n \frac{\gamma_k}{\rho_{k-1}} \sum_{m=l}^{k-1} \frac{|r_m| \rho_m}{\sqrt{m}} \leq \frac{\bar{C}_4}{l^{3/2}}. \tag{6.12}$$

Plugging this inequality in (6.11), we derive:

$$\nu_n^\Delta(|\cdot|) \leq \frac{1}{\Gamma_n} \sum_{l=1}^{n-1} \left[\sum_{k=l+1}^n \frac{\gamma_k}{\rho_{k-1}} \sum_{m=l}^{k-1} \frac{|r_m| \rho_m}{\sqrt{m}} \right] |U_l| \leq \frac{\bar{C}_4}{\Gamma_n} \sum_{l=1}^{n-1} \frac{|U_l|}{l^{3/2}}. \tag{6.13}$$

For any $\lambda > 0$, equation (6.13) and the Gaussian concentration property (GC) of the innovation entail:

$$\begin{aligned} \mathbb{E} \exp(\lambda \nu_n^\Delta(|\cdot|)) &\leq \prod_{k=1}^{n-1} \mathbb{E} \exp\left(\frac{\bar{C}_4 \lambda}{\Gamma_n k^{\frac{3}{2}}} |U_k|\right) \leq \prod_{k=1}^{n-1} \exp\left(\frac{\bar{C}_4 \lambda}{\Gamma_n k^{\frac{3}{2}}} \mathbb{E}[|U_1|] + \frac{1}{2} \left(\frac{\bar{C}_4 \lambda}{\Gamma_n k^{\frac{3}{2}}}\right)^2\right) \\ &= \exp\left(\frac{\bar{C}_4 \lambda \mathbb{E}[|U_1|] \Gamma_n^{\frac{3}{2}}}{\Gamma_n} + \frac{\bar{C}_4^2 \lambda^2 \Gamma_n^{(3)}}{2 \Gamma_n^2}\right). \end{aligned}$$

This completes the proof. □

6.2. Numerical results

We present in this section numerical results associated with the computation of the empirical measure ν_n illustrating our previous theorems.

6.2.1. Sub-Gaussian tails

We first consider $d = r = 1$. Also, for simplicity, the innovations $(U_i)_{i \geq 1}$ and X_0 are Bernoulli variables with $\mathbb{P}(U_1 = -1) = \mathbb{P}(U_1 = 1) = \frac{1}{2}$. We illustrate here Theorem 2 taking $b(x) = -\frac{x}{2}$, and $\sigma(x) = \cos(x)$ in (1.1). This is a (weakly) hypoelliptic example. Indeed, setting for $x \in \mathbb{R}$, $X_1(x) = \cos(x) \partial_x$ and $X_0(x) = -\frac{x}{2} \partial_x$, we have $\text{span}\{X_1, [X_1, X_0]\} = \mathbb{R}$. We choose as well to compute $\nu_n(\mathcal{A}\varphi)$ for $\varphi(x) = x + \varepsilon \cos(x)$ for $\varepsilon = 0.01$, and $\varphi(x) = \cos(x)$. The function φ is here given. The assumptions of Theorems 3 and 8 (for the coboundary condition on $\|\sigma\|^2 - \nu(\|\sigma\|^2)$) follow from Theorem 18 in Rotschild and Stein [41] (up to the introduction of a suitable partition of unity). From Theorem 2, for steps of the form $(\gamma_k)_{k \geq 1} = (k^{-\theta})_{k \geq 1}$, $\theta \in [1/3, 1]$ (corresponding to $\beta = 1$ in Theorem 2), the function

$$a \in \mathbb{R}^+ \mapsto g_{n,\theta}(a) := \begin{cases} \log(\mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)| \geq a]), & \theta \in (1/3, 1], \\ \log(\mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) + (\mathcal{B}_{n,1} - E_n^1)| \geq a]), & \theta = 1/3, \end{cases}$$

is such that for $a > a_n := a_n(\theta)$ where for $\theta \in (1/3, 1]$, $a_n(\theta) = 0$ and for $\theta = 1/3$, $a_n(\theta) = \frac{[\varphi^{(3)}]_\beta \|\sigma\|_\infty^{(3+\beta)} \mathbb{E}[|U_1|^{3+\beta}]}{(1+\beta)(2+\beta)(3+\beta)} \frac{\Gamma_n^{\frac{(3+\beta)}{2}}}{\sqrt{\Gamma_n}}$:

$$g_{n,\theta}(a) \leq -c_n \frac{(a - a_n)^2}{2\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2} + \log(2C_n).$$

We plot in Figure 1 the curves of $g_{n,\theta}$ for θ varying as $\theta_j = \frac{1}{3} + (1 - \frac{1}{3})\frac{j}{5}$, for $j \in \llbracket 1, 5 \rrbracket$, $\varphi(x) = x + \varepsilon \cos(x)$ and in Figure 2 the curve of $g_{n,\theta}$ for $\theta = \theta_0 = \frac{1}{3}$ and $\varphi(x) = \cos(x)$. The simulations have been performed for $n = 5 \times 10^4$ in Figure 1, $n = 5 \times 10^6$ in Figure 2, and the probability estimated by Monte Carlo simulation for $MC = 10^4$ realizations of the random variable $|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)|$ in the unbiased case and in the biased case of the random variable $|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) + (\mathcal{B}_{n,1} - E_n^1)^M|$, where $(\mathcal{B}_{n,1} - E_n^1)^M$ is obtained from $\mathcal{B}_{n,1} - E_n^1$ replacing the integral over $[0, 1]$, that needs to be evaluated at every time step, by a quantization of the uniform law on $[0, 1]$ with $M = 10$ points. We refer to [18] or [32] for details on quantization. We point out that this is one drawback that appears to obtain the fastest convergence rate, the bias needs to be estimated and therefore the function φ in some sense known (since the approximation of the bias requires to compute its derivatives). The corresponding 95% confidence intervals have size at most of order 0.016. To compare with, we also introduce the functions $S_{n,\theta}(a) := -\frac{(a - a_n(\theta))^2}{2\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}$, $S_{n,\theta,c}(a) := -\frac{(a - a_n(\theta))^2}{2\nu_{nc}(\sigma^2) \|\nabla\varphi\|_\infty^2}$, $S_{n,\theta,A}(a) := -\frac{(a - a_n(\theta))^2}{2\nu_{nc}(\sigma \nabla\varphi^2)}$ and the optimal concentration $P(\lambda_{\min})(n, \theta, a, \rho)$, obtained in Remark 11, optimizing numerically in ρ . The quantities $\nu_{nc}(\sigma^2)$, $\nu_{nc}(\sigma \nabla\varphi^2)$ in the previous expressions actually correspond to the numerical estimation, for $n_c = 10^4$ and $(\gamma_k^c)_{k \geq 1} = (k^{-\theta^c})_{k \geq 1}$ with $\theta^c = \frac{1}{3} + 10^{-3}$, of $\nu(\sigma^2)$, $\nu(\sigma \nabla\varphi^2)$ appearing respectively in the sharper concentration bound of Theorem 8 when $\sigma^2 - \nu(\sigma^2)$ is a coboundary and in the asymptotic Theorem 1. In the unbiased case of Figure 1, we plot the maximum in j of the $(S_{n,\theta_j})_{j \in \llbracket 1, 5 \rrbracket}$, $(S_{n,\theta_j,c})_{j \in \llbracket 1, 5 \rrbracket}$, $(S_{n,\theta_j,A})_{j \in \llbracket 1, 5 \rrbracket}$, $(P(\lambda_{\min})(n, \theta_j, a, \rho))_{j \in \llbracket 1, 5 \rrbracket}$ corresponding to $j = 1$. The associated curves are denoted by S_n , $S_{n,c}$, $S_{n,A}$ and $P(\lambda_{\min})(n)$.

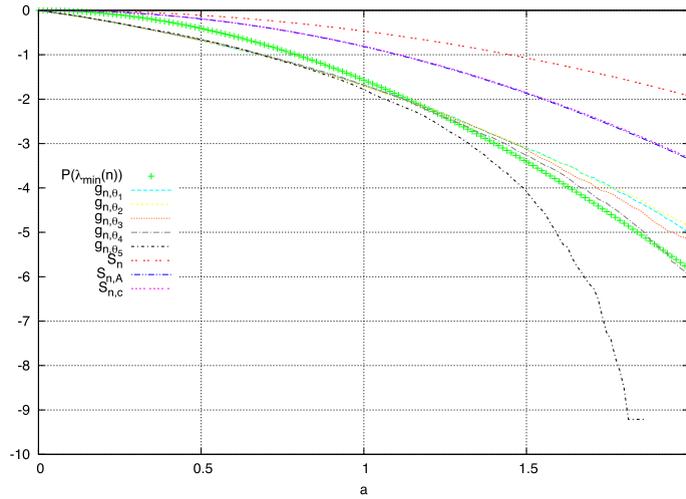


Fig. 1. Unbiased Case. Plot of $a \mapsto g_{n,\theta}(a)$, for $(\theta_k)_{k \in [1,5]}$, with $\varphi(x) = x + \varepsilon \cos(x)$, $\varepsilon = 0.01$.

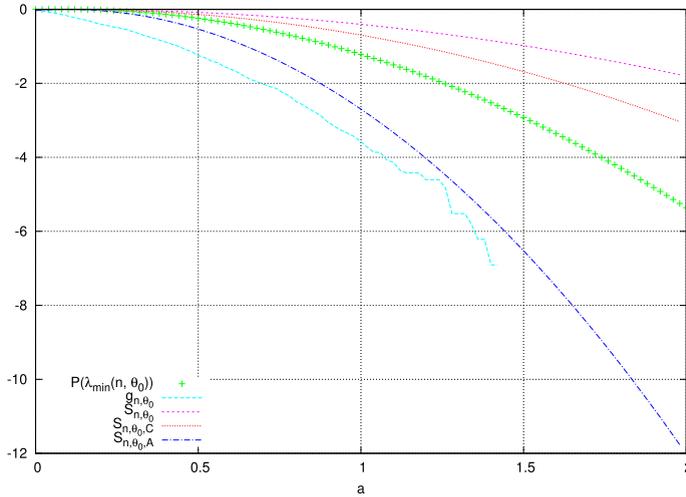


Fig. 2. Biased Case. Plot of $a \mapsto g_{n,\theta}(a)$, for $\theta_0 = \frac{1}{3}$, with $\varphi(x) = \sigma(x) = \cos(x)$.

Figures 1 and 2 correspond to the unbiased and biased cases respectively. In the unbiased case, we observe that the curves almost overlay, the optimal deviation rate $P(\lambda_{\min})$ is very close to the empirical data. It is also below the numerical estimation of the asymptotic threshold given by $S_{n,\theta,A}$ which is, for our considered example, almost indistinguishable from the coboundary $S_{n,\theta,c}$ (indeed, since $\varepsilon = 0.01$, $\|\nabla\varphi\|_\infty^2 \leq 1 + \varepsilon^2$ and $\nu(\sigma^2)\|\nabla\varphi\|_\infty^2 \simeq \nu(|\sigma\nabla\varphi|^2)$) and far below from the bounds of $S_{n,\theta}$. In the biased case, $P(\lambda_{\min})$ stays very close to the theoretical asymptotic bound given by $S_{n,\theta,A}$ up to a certain deviation level a , namely for $a \in [0, 0.5]$. It then remains the best bound provided by our results. In this example, the improvement associated with $S_{n,\theta,c}$ is also notable. It is precisely because the source term has a more oscillating gradient that we have also considered a larger running time, corresponding to $n = 10^6$, for the empirical curves. For this choice, we see relatively good agreement w.r.t. to the asymptotic deviation bounds of $S_{n,\theta_0,A}$.

Figures 3 and 4 thus illustrate that the explicit *optimal rate* of Remark 11 seems rather appropriate to capture the deviations of the empirical random measures.

We eventually plot below the deviation curves with source $\varphi(x) = \cos(x)$ adding a last curve obtained replacing in the formula for $P(\lambda_{\min})$ of Remark 11 the $\|\nabla\varphi\|_\infty^2 \nu(\sigma^2)$ by $\nu(|\sigma\nabla\varphi|^2)$. For practical purposes, this last quantity is again estimated numerically with the same previous parameters. Even if the analysis of Theorem 8 cannot be extended to justify such a choice, the empirical evidence is rather striking.

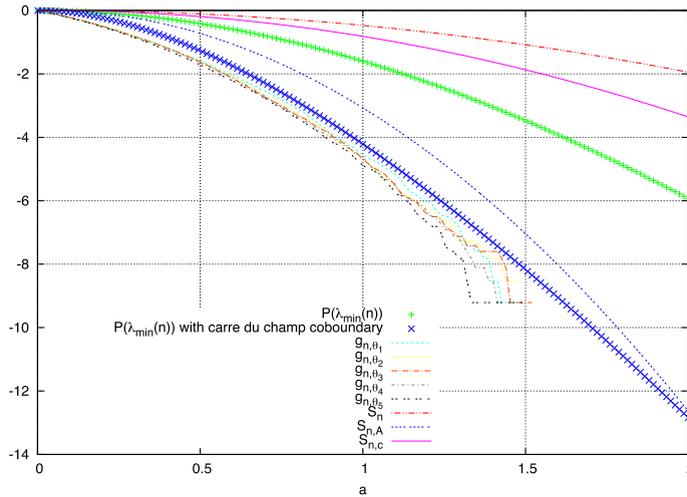


Fig. 3. Plot of $a \mapsto g_{n,\theta}(a)$, for $(\theta_k)_{k \in [1,5]}$, with $\varphi(x) = \sigma(x) = \cos(x)$.

6.2.2. Slutsky like result

In this paragraph, we illustrate our results from Theorem 6, which can be viewed as an extension of the usual Slutsky’s Lemma to our current framework, for a multidimensional process, precisely for $r = d = 2$ in the case $\beta \in (0, 1)$. In order to converge as fast as possible without bias, we take $\theta = \frac{1}{2+\beta} + \frac{1}{1000}$. We also choose a model which satisfies the assumptions of Theorem 4 under (\mathbf{C}_{UE}) and Lemma 6. We consider:

$$f(x) = \frac{|x|^{1+\beta}}{1 + |x|^\beta}, \quad b(x) = \begin{pmatrix} -4x_1 + 6x_2 \\ -5x_1 - 5x_2 \end{pmatrix}, \quad \sigma\sigma^*(x) = \begin{pmatrix} \frac{\cos(x_1+x_2)}{2} + 1 & \frac{\sin(x_1)\sin(x_2)}{4} \\ \frac{\sin(x_1)\sin(x_2)}{4} & 1 - \frac{\sin(x_2)}{2} \end{pmatrix}.$$

Remark that the non-degeneracy condition (\mathbf{UE}) is fulfilled by $\Sigma = \sigma\sigma^*$, as well as the condition set in Theorem 4 under (\mathbf{C}_{UE}) , $\Sigma_{i,j}(x) = \Sigma_{i,j}(x_1, \dots, x_d)$, for all $1 \leq i \leq j \leq d$. Furthermore, from the Cholesky decomposition, we write:

$$\sigma(x) = \begin{pmatrix} \sqrt{\frac{\cos(x_1+x_2)}{2} + 1} & 0 \\ \frac{\sin(x_1)\sin(x_2)}{4\sqrt{\frac{\cos(x_1+x_2)}{2} + 1}} & \sqrt{-\frac{\sin(x_1)^2\sin(x_2)^2}{16(\frac{\cos(x_1+x_2)}{2} + 1)} + 1 - \frac{\sin(x_2)}{2}} \end{pmatrix}.$$

Let us check that (\mathbf{D}_α^p) is satisfied. Firstly, remark that $\frac{Db+Db^*}{2}$ is a constant matrix whose eigenvalues are $\{-\frac{\sqrt{2}+9}{2}, \frac{\sqrt{2}-9}{2}\}$. Direct computations yield that, for all $x \in \mathbb{R}^d, \xi \in \mathbb{R}^d$:

$$\left\langle \frac{Db(x) + Db(x)^*}{2} \xi, \xi \right\rangle + \frac{1}{2} \sum_{j=1}^r |D\sigma_{\cdot j}(x)\xi|^2 \leq -3.085|\xi|^2.$$

It can be checked similarly that the condition $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{2(1+\beta)-p}$ is satisfied for $\alpha = 3.085$ and $\beta = 0.5$ which we consider below. Also, the condition $(\mathbf{R}_{1,\beta})$ clearly holds. In other words, all assumptions of Theorem 6 are in force. We set for the following plot:

$$g_n^\sigma(a) = \log \mathbb{P}[\sqrt{\Gamma_n} |v_n(f)| \geq a], \quad S_\sigma(a) = -\frac{a^2\alpha^2}{2[f]_1^2},$$

with $\alpha = 3.085$, and $[f]_1 = 1$.

Unlike in the previous simulations, we do not know here the value of $v(f)$. In fact, in Section 6.2.1 we had chosen to compute the deviation of $\mathcal{A}\varphi$ from $0 = v(\mathcal{A}\varphi)$. Here, we estimate from the ergodic theorem $v(f)$, taking $\beta = 0.5$, by $v_{n^c}(f) \approx 0.71308$ for $n^c = 5 \cdot 10^5$. Running $MC = 10^2$ samples, we find that the size of the associated 95% confidence interval is $3.208 \cdot 10^{-4}$. Finally, the simulations are performed for $n = 5 \times 10^4$, and the probability is calculated by Monte Carlo algorithm for $MC = 10^3$ realizations. The maximum size of the associated 95% confidence interval is $4.75054 \cdot 10^{-5}$. The innovations are Gaussian random variables.

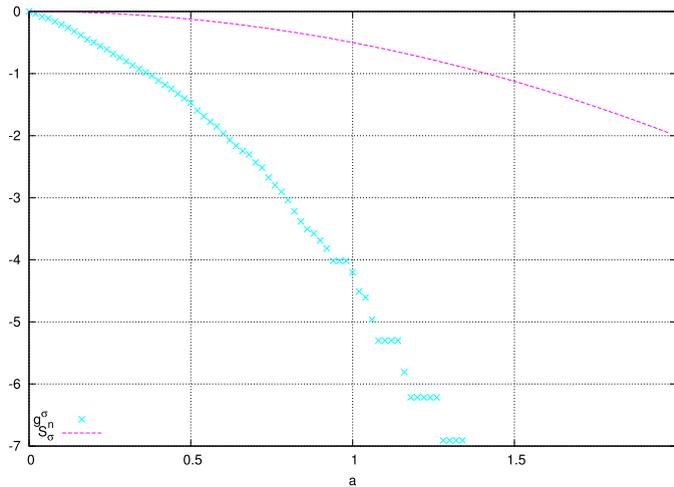


Fig. 4. Plot of $a \mapsto g_n(a)$ with $f(x) = \frac{|x|^\beta}{1+|x|^\beta}$, $\beta = 0.5$.

In Figure 4, we observe that the curve S_σ stays above g_n^σ as proved in Theorem 6. However, remark that the graphs are quite spaced. This can be explained, among other things, by the difference between $\nu(\|\sigma\|^2) \frac{\|f\|_1}{\alpha}$ and the asymptotic variance $\nu(|\sigma^* \nabla \varphi|^2)$. Furthermore we have represented S_σ which is a kind of asymptotic version of $P(\lambda_{\min}(n))$ in the previous plots.

Acknowledgements

For the second author, the article was prepared within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

References

- [1] A. Arnold, A. Carlen and Q. Ju. Large-time behavior of non-symmetric Fokker–Planck type equations. *Commun. Stoch. Anal.* **2** (1) (2008) 153–175. MR2446997 <https://doi.org/10.31390/cosa.2.1.11>
- [2] A. Arnold, P. Markowich, G. Toscani and A. Unterreiter. On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker–Planck type equations. *Commun. Partial Differ. Equ.* **26** (1–2) (2001) 43–100. MR1842428 <https://doi.org/10.1081/PDE-100002246>
- [3] D. Bakry and M. Émery. Diffusions hypercontractives. *Sémin. Probab.* **XIX** (1985) 177–206. MR0889476 <https://doi.org/10.1007/BFb0075847>
- [4] D. Bakry, I. Gentil and M. Ledoux. *Analysis and Geometry of Markov Diffusion Operators. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **348**. Springer, Cham, 2014. MR3155209 <https://doi.org/10.1007/978-3-319-00227-9>
- [5] R. F. Bass. *Diffusions and Elliptic Operators*. Springer, New York, 1997. MR1483890
- [6] R. N. Bhattacharya. On the functional central limit theorem and the law of the iterated logarithm for Markov processes. *Z. Wahrsch. Verw. Gebiete* **60** (2) (1982) 185–201. MR0663900 <https://doi.org/10.1007/BF00531822>
- [7] G. Blower and F. Bolley. Concentration inequalities on product spaces with applications to Markov processes. *Studia Math.* **175** (1) (2006) 47–72. MR2261699 <https://doi.org/10.4064/sm175-1-3>
- [8] S. Boccia. Schauder estimates for solutions of higher-order parabolic systems. *Methods Appl. Anal.* **20** (1) (2013) 47–67. MR3085781 <https://doi.org/10.4310/MAA.2013.v20.n1.a3>
- [9] E. Boissard. Simple bounds for the convergence of empirical and occupation measures in 1-Wasserstein distance. *Electron. J. Probab.* **16** (2011) 2296–2333. MR2861675 <https://doi.org/10.1214/EJP.v16-958>
- [10] F. Bolley, A. Guillin and C. Villani. Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Related Fields* **137** (2007) 541–593. MR2280433 <https://doi.org/10.1007/s00440-006-0004-7>
- [11] G. A. Brosamler. An almost everywhere central limit theorem. *Math. Proc. Cambridge Philos. Soc.* **104** (3) (1988) 561–574. MR0957261 <https://doi.org/10.1017/S0305004100065750>
- [12] F. Chaabane and F. Maaouia. Théorèmes limites avec poids pour les martingales vectorielles. *ESAIM Probab. Stat.* **4** (2000) 137–189 (electronic). MR1808925 <https://doi.org/10.1051/ps:2000103>
- [13] M. Csörgő and L. Horváth. Invariance principles for logarithmic averages. *Math. Proc. Cambridge Philos. Soc.* **112** (1) (1992) 195–205. MR1162944 <https://doi.org/10.1017/S0305004100070870>
- [14] J. Dedecker and S. Gouëzel. Subgaussian concentration inequalities for geometrically ergodic Markov chains. *Electron. Commun. Probab.* **20** (2015) Article 64. MR3407208 <https://doi.org/10.1214/ECP.v20-3966>
- [15] M. Duflo. *Méthodes Récursives Aléatoires. Techniques Stochastiques. [Stochastic Techniques]*. Masson, Paris, 1990. MR1082344
- [16] E. Ethier and T. Kurtz. *Markov Processes. Characterization and Convergence*. Wiley, New York, 1986. MR0838085 <https://doi.org/10.1002/9780470316658>

- [17] N. Frikha and S. Menozzi. Concentration bounds for stochastic approximations. *Electron. Commun. Probab.* **17** (2012) no. 47. MR2988393 <https://doi.org/10.1214/ecp.v17-1952>
- [18] S. Graf and H. Lushgy. *Foundations of Quantization for Random Vectors. Lecture Notes in Mathematics* **1730**. Springer, Berlin, 2000. MR1764176 <https://doi.org/10.1007/BFb0103945>
- [19] M. K. Heck. The principle of large deviations for the almost everywhere central limit theorem. *Stochastic Process. Appl.* **76** (1) (1998) 61–75. MR1638007 [https://doi.org/10.1016/S0304-4149\(98\)00023-4](https://doi.org/10.1016/S0304-4149(98)00023-4)
- [20] N. Ikeda and S. Watanabe. *Stochastic Differential Equations*. North Holland, Amsterdam, 1980. MR0637061
- [21] A. Joulin and Y. Ollivier. Curvature, concentration and error estimates for Markov chain Monte Carlo. *Ann. Probab.* **38** (6) (2010) 2418–2442. MR2683634 <https://doi.org/10.1214/10-AOP541>
- [22] R. Khasminskii. *Stochastic Stability of Differential Equations*, 2nd edition. *Stochastic Modelling and Applied Probability* **66**. Springer, Heidelberg, 2012. With contributions by G. N. Milstein and M. B. Nevelson. MR2894052 <https://doi.org/10.1007/978-3-642-23280-0>
- [23] R. Z. Khasminskii. *Stochastic Stability of Differential Equations. Monographs and Textbooks on Mechanics of Solids and Fluids: Mechanics and Analysis* **7**. Sijthoff & Noordhoff, Alphen aan den Rijn—Germantown Md., 1980. Translated from the Russian by D. Louvish. MR0600653
- [24] N. V. Krylov and E. Priola. Elliptic and parabolic second-order PDEs with growing coefficients. *Comm. Partial Differential Equations* **35** (1) (2010) 1–22. MR2748616 <https://doi.org/10.1080/03605300903424700>
- [25] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion. *Bernoulli* **8** (3) (2002) 367–405. MR1913112 <https://doi.org/10.1142/S0219493703000838>
- [26] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion: The case of a weakly mean reverting drift. *Stoch. Dyn.* **3** (4) (2003) 435–451. MR2030742 <https://doi.org/10.1142/S0219493703000838>
- [27] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXIII* 120–216. *Lecture Notes in Math.* **1709**. Springer, Berlin, 1999. MR1767995 <https://doi.org/10.1007/BFb0096511>
- [28] V. Lemaire. An adaptive scheme for the approximation of dissipative systems. *Stochastic Process. Appl.* **117** (10) (2007) 1491–1518. MR2353037 <https://doi.org/10.1016/j.spa.2007.02.004>
- [29] V. Lemaire, G. Pagès and F. Panloup. Invariant measure of duplicated diffusions and application to Richardson–Romberg extrapolation. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** (4) (2015) 1562–1596. MR3414458 <https://doi.org/10.1214/13-AIHP591>
- [30] F. Malrieu and D. Talay. Concentration inequalities for Euler schemes. In *Monte Carlo and Quasi-Monte Carlo Methods 2004* 355–371. H. Niederreiter and D. Talay (Eds). Springer, Berlin, Heidelberg, 2006. MR2208718 https://doi.org/10.1007/3-540-31186-6_21
- [31] F. Merlevède, M. Peligrad and E. Rio. A Bernstein type inequality and moderate deviations for weakly dependent sequences. *Probab. Theory Related Fields* **151** (3) (2011) 435–474. MR2851689 <https://doi.org/10.1007/s00440-010-0304-9>
- [32] G. Pagès. A space vector quantization method for numerical integration. *J. Comput. Appl. Math.* **89** (1997) 1–38. MR1625987 [https://doi.org/10.1016/S0377-0427\(97\)00190-8](https://doi.org/10.1016/S0377-0427(97)00190-8)
- [33] G. Pagès. Sur quelques algorithmes récurrents pour les probabilités numériques. *ESAIM Probab. Stat.* **5** (2001) 141–170. MR1875668 <https://doi.org/10.1051/ps:2001106>
- [34] G. Pagès and F. Panloup. Ergodic approximation of the distribution of a stationary diffusion: Rate of convergence. *Ann. Appl. Probab.* **22** (3) (2012) 1059–1100. MR2977986 <https://doi.org/10.1214/11-AAP779>
- [35] G. Pagès and F. Panloup. A mixed-step algorithm for the approximation of the stationary regime of a diffusion. *Stochastic Process. Appl.* **124** (1) (2014) 522–565. MR3131304 <https://doi.org/10.1016/j.spa.2013.07.011>
- [36] G. Pagès and C. Rey. Recursive computation of invariant distributions of Feller processes. *Stochastic Process. Appl.* **130** (1) (2020) 328–365. MR4035032 <https://doi.org/10.1016/j.spa.2019.03.008>
- [37] E. Pardoux and A. Veretennikov. On the Poisson equation and diffusion approximation. I. *Ann. Probab.* **29** (3) (2001) 1061–1085. MR1872736 <https://doi.org/10.1214/aop/1015345596>
- [38] E. Pardoux and A. Y. Veretennikov. On Poisson equation and diffusion approximation. II. *Ann. Probab.* **31** (3) (2003) 1166–1192. MR1988467 <https://doi.org/10.1214/aop/1055425774>
- [39] E. Pardoux and A. Y. Veretennikov. On the Poisson equation and diffusion approximation. III. *Ann. Probab.* **33** (3) (2005) 1111–1133. MR2135314 <https://doi.org/10.1214/009117905000000062>
- [40] E. Priola. Global Schauder estimates for a class of degenerate Kolmogorov equations. *Studia Math.* **194** (2) (2009) 117–153. MR2534181 <https://doi.org/10.4064/sm194-2-2>
- [41] L. P. Rothschild and E. M. Stein. Hypocoelliptic differential operators and nilpotent groups. *Acta Math.* **137** (3–4) (1976) 247–320. MR0436223 <https://doi.org/10.1007/BF02392419>
- [42] P. Schatte. On strong versions of the Central Limit Theorem. *Math. Nachr.* **137** (1988) 249–256. MR0968997 <https://doi.org/10.1002/mana.19881370117>
- [43] D. Talay. Stochastic Hamiltonian dissipative systems: Exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Related Fields* **8** (2) (2002) 163–198. MR1924934
- [44] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch. Anal. Appl.* **8** (4) (1990) 94–120. MR1091544 <https://doi.org/10.1080/07362999008809220>
- [45] C. Villani. Hypocoercivity. *Mem. Amer. Math. Soc.* **202** (950) (2009) iv+141. MR2562709 <https://doi.org/10.1090/S0065-9266-09-00567-5>