Non-Asymptotic Gaussian Estimates for the Recursive Approximation of the Invariant Measure of a Diffusion

Igor Honoré, Stephane Menozzi, Gilles Pagès

NON-ASYMPTOTIC GAUSSIAN ESTIMATES FOR THE RECURSIVE APPROXIMATION OF THE INVARIANT MEASURE OF A DIFFUSION

I. HONORÉ, S. MENOZZI, AND G. PAGÈS

Abstract. We obtain non-asymptotic Gaussian concentration bounds for the difference between the invariant measure 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 the (squared) Fröbenius norm of the diffusion coefficient also lies in 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.

1. Introduction

1.1. **Setting.** The aim of this article is to approach the invariant measure of the solution of the diffusion equation:

$$dY_t = b(Y_t)dt + \sigma(Y_t)dW_t, \tag{1.1}$$

where $(W_t)_{t\geq 0}$ is a Wiener process of dimension r on a given filtered probability space $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t\geq 0}, \mathbb{P})$, $b : \mathbb{R}^d \to \mathbb{R}^d$, and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^r$ are assumed to be Lipschitz continuous functions and to satisfy a mean-reverting assumption. Typically, if $\mathcal{A}$ denotes the infinitesimal generator of the diffusion (1.1), there exists a twice continuously differentiable Lyapunov function $V : \mathbb{R}^d \to (0, +\infty)$ such that $\lim_{|x|\to+\infty} V(x) = +\infty$ 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. We will also assume uniqueness of the invariant measure, denoted from now by ν. We refer to the monographs by Khasminski and Milstein [KM11], Ethier and Kurtz [EK86] or Villani [Vil09], for in-depth discussions on the conditions yielding such existence and uniqueness results.

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:

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

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) measure of the scheme in the following way. For all $A \in \mathcal{B}(\mathbb{R}^d)$ (where $\mathcal{B}(\mathbb{R}^d)$ denotes the Borel σ-field on $\mathbb{R}^d$):

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

The measure ν_n is here defined accordingly to the *intrinsic* time scale of the scheme. Since we are interested in long time approximation, we consider steps $(\gamma_k)_{k\geq 1}$ such that $\Gamma_n := \sum_{k=1}^n \gamma_k \underset{n}{\to} +\infty$. We also assume $\gamma_k \underset{k}{\downarrow} 0$. Observe that, for a bounded ν-*a.s.* continuous function f, it can be shown, see *e.g.* Theorem 1 in [LP02], that:

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

Date: July 10, 2017.

Key words and phrases. Invariant distribution, diffusion processes, inhomogeneous Markov chains, non-asymptotic Gaussian concentration, almost sure Central Limit Theorem.

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 [Duf90]. 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 measure 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.* [TT90], [Tal02] 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 [LP02] 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 convergence in law 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 [PV01], Rothschield and Stein [RS76] or Villani [Vil09] who discuss the solvability of the Poisson problem under some ellipticity or hypoellipticity assumptions, or to Pagès and Panloup [PP12] who exploit some confluence conditions allowing to handle for instance the case of an Ornstein-Uhlenbeck process with degenerate covariance matrix.

In the current paper, our goal is to establish, in a similar framework, a non-asymptotic Gaussian control for the deviations of the quantity $\nu_n(f)-\nu(f)$. 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 non-asymptotic confidence intervals that can be *explicitly* computed. A specific feature of such non-asymptotic deviation inequality 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, 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 Section 4 and 6.2 for details about these points. Eventually, this allows to provide deviations inequalities for plain Lipschitz continuous sources f in the ergodic approximation, by using a suitable regularization procedure, see Section 5.4.

The main difficulty, compared to the extensive literature on deviation inequalities for ergodic homogeneous Markov chains is to take into account the inhomogeneous feature of this problem, combined with the fact that our target is the true invariant distribution of the continuous time diffusion. Such kind of approximation produce some remainder terms that should be controlled accurately enough in a non-asymptotic way to produce tractable deviation inequalities asymptotically close to their counterparts for the diffusion itself. This a major difficulty compared to a CLT where these remainder terms are simply requested to go to 0.

As mentioned above and like for the CLT (see [Bha82] for the diffusion or [LP02] 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 requires to solve the Poisson equation in that spirit, that means deriving pointwise regularity results on φ from those made on f.

In particular, we will not rely on the Sobolev regularity but rather on some Schauder estimates in line with the works by Krylov and Priola [KP10], which allow to benefit from the parabolic regularity. 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:

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

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 $\forall n \geq n_0,\ C^{-1}v_n \leq u_n \leq C v_n$.

Hypotheses.

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

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

(GC) The i.i.d. innovation sequence $(U_n)_{n\geq 1}$ is such that $\mathbb{E}[U_1] = 0$ and for all $(i,j,k) \in \{1,\cdots,r\}^3$, $\mathbb{E}[U_1^i U_1^j] = \delta_{ij}$, $\mathbb{E}[U_1^i U_1^j U_1^k] = 0$. 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 $g : \mathbb{R}^r \to \mathbb{R}$ and every $\lambda > 0$:

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

Observe that if $U_1 \overset{\text{(law)}}{=} \mathcal{N}(0, I_r)$ or $U_1 \overset{\text{(law)}}{=} (\frac{1}{2}(\delta_1 + \delta_{-1}))^{\otimes r}$, *i.e.* for Gaussian or symmetrized 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 there exists $\varpi > 0$ such that for all $\lambda > 0$:

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

yielding that for all $r \geq 0$, $\mathbb{P}[|U_1| \geq r] \leq 2\exp(-\frac{r^2}{\varpi})$ (sub-Gaussian concentration of the innovation). Setting $\varpi = 2$ corresponds to the standard Gaussian concentration. This is also the constant in the logarithmic Sobolev inequality satisfied by the standard Gaussian measure.

(C2) There exists a positive constant κ such that, defining for all $x \in \mathbb{R}^d,\ \Sigma(x) := \sigma\sigma^*(x)$:

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

where $\|\sigma(x)\|$ stands for the Fröbenius norm of $\sigma(x)$. We then set $\|\sigma\|_\infty := \sup_{x\in\mathbb{R}^d} \|\sigma(x)\|$.

($\mathcal{L}_\mathbf{V}$) We have the following Lyapunov like stability condition:

There exists $V : \mathbb{R}^d \longrightarrow [v^*, +\infty[$ with $v^* > 0$ such that

i) V is a $\mathcal{C}^2$ function, $\|D^2 V\|_\infty < +\infty$, and $\lim_{|x|\to\infty} V(x) = +\infty$.

ii) There exists $C_{_V} \in (0, +\infty)$ such that for all $x \in \mathbb{R}^d$:

$$|\nabla V(x)|^2 + |b(x)|^2 \leq C_{_V} V(x).$$

iii) 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 for all $x \in \mathbb{R}^d$ by:

$$\mathcal{A}\varphi(x) = b(x)\cdot\nabla\varphi(x) + \frac{1}{2}\mathrm{Tr}\big(\Sigma(x) D^2\varphi(x)\big),$$

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

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

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

(U) There is a unique invariant measure ν to Equation (1.1).

(S) For V satisfying $(\mathcal{L}_V)$:

i) As a consequence of $(\mathcal{L}_V)$ i), there exist constants K and $\bar{c}$ such that for $|x| \geq K$, $|V(x)| \leq \bar{c}|x|^2$ and $|b(x)| \leq \sqrt{C_V \bar{c}}|x|$.

ii) We assume that the sequence $(\gamma_k)_{k\geq 1}$ satisfies for all $k \geq 1, \gamma_k \leq \frac{1}{2}\min(\frac{1}{\sqrt{C_V \bar{c}}}, \frac{\alpha_V}{C_V \|D^2 V\|_\infty})$.

Condition ii) in **(S)** means 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* $(\mathcal{L}_\mathbf{V})$ *yields existence (see* [EK86]*). Additional structure conditions ((hypo-)ellipticity* [KM11], [PV01], [Vil09] *or confluence* [PP12]*) 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 Lemma 5 below).*

We say that assumption **(A)** holds whenever **(C1)**, **(GC)**, **(C2)**, $(\mathcal{L}_\mathbf{V})$, **(U)** and **(S)** are fulfilled. Except when explicitly indicated, we assume throughout the paper that assumption **(A)** is in force.

Observe that, as soon as conditions **(C2)**, $(\mathcal{L}_\mathbf{V})$, **(U)** are satisfied and $\mathbb{E}[U_1] = 0$, the following Central Limit Theorem (CLT) holds (see Theorems 9, 10 in [LP02]).

Theorem 1 (CLT)**.** *Under* **(C2)**, $(\mathcal{L}_\mathbf{V})$, **(U)**, *if* $\mathbb{E}[U_1] = 0$, $\mathbb{E}[U_1^{\otimes 3}] = 0$, *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* $D^2\varphi$ *and* $D^3\varphi$ *bounded, one has*

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

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

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

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

where

$$\begin{aligned} m &:= -\int_{\mathbb{R}^d} \Big(\frac{1}{2} D^2\varphi(x) b(x)^{\otimes 2} + \Phi_4(x)\Big)\nu(dx), \\ \Phi_4(x) &:= \int_{\mathbb{R}^r} \Big(\frac{1}{2}\langle D^3\varphi(x) b(x), (\sigma(x)u)^{\otimes 2}\rangle + \frac{1}{24} D^4\varphi(x)(\sigma(x)u)^{\otimes 4}\Big)\mu(du) \end{aligned}$$

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* $(\partial^3_{x_i,x_j,x_k}\varphi)_{(i,j,k)\in[\![1,d]\!]^3}$ *and we denote, for all* $x \in \mathbb{R}^d$, *by* $D^3\varphi(x)b(x)$ *the* $\mathbb{R}^d \otimes \mathbb{R}^d$ *matrix with entries* $\big(D^3\varphi(x)b(x)\big)_{ij} = \sum_{k=1}^d (D^3\varphi(x))_{ijk} b_k(x)$, $(i,j) \in [\![1,d]\!]^2$.

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* $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \underset{n}{\to} 0$. *In case (b), that is for* $\theta \in (0, \frac{1}{3}]$, $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \underset{n}{\to} \widetilde{\gamma}$, *with* $\widetilde{\gamma} < +\infty$ *for* $\theta = \frac{1}{3}$ *and* $\widetilde{\gamma} = +\infty$ *for* $\theta \in (0, \frac{1}{3})$.

Let us mention that, when $\frac{\Gamma_n^{(3/2)}}{\sqrt{\Gamma_n}} < +\infty$, *i.e.* $\gamma_n \asymp n^{-\theta}, \theta \in (1/2, 1]$, the statement of point *(a)* holds without the condition $\mathbb{E}[U_1^{\otimes 3}] = 0$ and as soon as $\mathbb{E}[|U_1|^4] < +\infty$ (see Theorem 9 in [LP02]). Moreover, the boundedness condition **(C2)** can be relaxed to derive the CLT, which holds provided $\lim_{|x|\to+\infty} \frac{|\sigma^*\nabla\varphi(x)|^2}{V(x)} = 0$ (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 [LP02] for further considerations.

Remark 3. *Note that the asymptotic variance corresponds to the usual integral of the "carré du champ" w.r.t. to the invariant measure, see Bhattacharya* [Bha82] *or the monograph by Bakry et al.* [BGL14]*, i.e.:*

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

Remark 4. *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)**, $(\mathcal{L}_{\mathbf{V}})$ *(see Bhattacharya* [Bha82]*). In fact* **(C2)** *can be partially relaxed as well, like mentioned above. Precisely,* $\frac{1}{\sqrt{t}} \int_0^t f(Y_s) ds \xrightarrow{\mathcal{L}} \mathcal{N}(0, \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi|^2 d\nu)$ *as* $t \to +\infty$*. In both cases, the normalization is the same: the square root of the considered running time,* t *for the diffusion and* Γ_n *for the scheme. In Theorem 1, the fastest convergence is obtained for* $\theta = 1/3$ *(critical value between "fast" and "slow" settings), for which a bias induced by the discretization steps in. When* $\theta \in (0, 1/3)$ *(slow decreasing step), this discretization effect becomes prominent and "hides" the CLT.*

Remark 5. *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 from a practical viewpoint to choose* γ_0 *in an appropriate way, namely by minimizing the function* $\gamma_0 \mapsto c_1 \gamma_0 + c_2 \gamma_0^{-1/2}, c_1 = \lim_n \frac{\sum_{k=1}^n k^{-2/3}}{(\sum_{k=1}^n k^{-1/3})^{1/2}}, c_2 = \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi|^2 d\nu$*, 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, i.e.* $c_2 \leq \|\sigma\|_\infty \|\nabla \varphi\|_\infty$*.*

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 [MT06] 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 [Led99] or Bakry *et al.* [BGL14] 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 [LP02] and have been successfully used in Frikha and Menozzi [FM12] 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 $\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)$ 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 [DG15] 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 [BGV07] 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.* [MPR11]. From a more dynamical viewpoint, let us mention the work of Joulin and Ollivier [JO10], 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 measure. We also mention a work of Blower and Bolley [BB06], who obtain Gaussian concentration properties for deviations of functional of the path for metric space valued homogeneous Markov chains or Boissard [Boi11] 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 transition law involved, see *e.g.* condition *(ii)* in Theorems 1.2 and 2.1 of [BB06]. Checking such continuity conditions can be difficult in practice. Sufficient conditions, which require absolute continuity and smoothness of the transition laws are given in Proposition 2.2 of [BB06].

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 [PP12]. Also, the approach could possibly extend to diffusions with less stringent Lyapunov conditions, like the weakly mean reverting drifts considered in [LP03], or even to more general ergodic Markov processes. 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 φ is $\mathcal{C}^3(\mathbb{R}^d,\mathbb{R})$, Lipschitz continuous with $(D^i\varphi)_{i\in\{2,3\}}$ bounded, such that there exists $C_\varphi>0$ satisfying for all $x\in\mathbb{R}^d$, $|D^2\varphi(x)|\leq C_\varphi(1+|x|)^{-1}$ and $D^3\varphi$ is Lipschitz continuous, we then establish that there are explicit sequences $c_n\leq 1\leq C_n$ converging to 1 such that for all $n\in\mathbb{N}$, for all $a>0$ and $\gamma_k\asymp k^{-\theta},\ \theta\in(\frac{1}{3},1]$,

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

When the diffusion coefficient σ is such that $\|\sigma\|^2-\nu(\|\sigma\|^2)$ is itself a coboundary, the previous bound improves 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 3 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 [Bro88] and [Sch88]) and revisited through the ergodic discretization schemes viewpoint in [LP02].

Both applications require a careful investigation of the corresponding Poisson problem.

The paper is organized as follows. We conclude this section by introducing some notations. We then state and prove in Section 2 our main concentration results. To this end, we state various technical Lemmas whose proofs are postponed to Section 3. Section 4 is devoted to the case where $\|\sigma\|^2-\nu(\|\sigma\|^2)$ is a coboundary, We then establish in Section 5 some precise and quantitative deviation bounds for $\sqrt{\Gamma_n}\big(\nu_n(f)-\nu(f)\big)$ provided some good controls on the Poisson equation holds, including a Slutsky like theorem in view of practical applications. We prove that these controls are achieved in two frameworks: when some non-degeneracy or confluent conditions hold. The associated non-asymptotic confidence intervals are discussed as well. 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 constant that may change from line to line and depend, uniformly in time, on known parameters appearing in **(A)**. Other possible dependencies will be explicitly specified. We will also denote by $\mathscr{R}_n$ 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 additionally 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 $f \in \mathcal{C}^p(\mathbb{R}^d, \mathbb{R}),\ p \in \mathbb{N}$, 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 $\mathbb{N}_0^d \backslash \{0\}$ 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

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

We also denote by $\mathcal{C}_b^{k,\beta}$ the subset of $\mathcal{C}^{k,\beta}$ for which the functions themselves ares bounded. In particular, $\mathcal{C}^{0,1}(\mathbb{R}^d, \mathbb{R}^m)$ is the space of Lipshitz continuous functions from $\mathbb{R}^d$ to $\mathbb{R}^m$. 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 $f : \mathbb{R}^d \to E$, where E can be $\mathbb{R}, \mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^q,\ q \in \{r, d\}$, we set for $k \in \mathbb{N}_0$:

$$f_k \ := \ f(X_k).$$

Eventually, for $k \in \mathbb{N}_0$, we denote by $\mathcal{F}_k := \sigma\big((X_j)_{j \in [\![0,k]\!]}\big)$.

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 $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$, $\beta \in (0,1]$ instead of $\varphi \in \mathcal{C}^4(\mathbb{R}^d, \mathbb{R})$ with existing bounded partial derivatives up to order four (which in particular implies that in Theorem 1 $\varphi \in \mathcal{C}^{3,1}(\mathbb{R}^d, \mathbb{R})$).

Theorem 2. *Assume* **(A)** *holds. Consider a Lipschitz continuous (possibly unbounded) function $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ for some $\beta \in (0,1]$. Let us assume furthermore that:*

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

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

(*a*) ***Unbiased Case (sub-optimal convergence rate):*** *Assume $\beta \in (0,1]$ and $\theta \in (\frac{1}{2+\beta}, 1]$*

(i) Assume that there exists $C_\varphi > 0$ such that for all $x \in \mathbb{R}^d$, $|D^2\varphi(x)| \leq C_\varphi/(1+|x|)$.
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 $n \geq 1$ and for every $a > 0$:

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

(ii) If the function φ does not fulfill the condition $|D^2\varphi(x)| \leq C_\varphi/(1+|x|)$, the above result still holds for $0 < a \leq \chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$ for a positive sequence $\chi_n \underset{n}{\to} 0$ arbitrarily slowly, so that $\chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}} \underset{n}{\to} +\infty$. In particular, for a fixed $a > 0$, the above concentration inequality holds for n large enough.

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

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

and define subsequently[1]*:*

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

1. The dependency of E_n^β in β is hidden in the presence of the β-Hölder function $D^3\varphi$ in Λ_{k-1}

Set now for $\beta \in (0,1]$:

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

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

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

For $\beta \in (0,1)$, the random variables $|\mathcal{B}_{n,\beta}| = |E_n^\beta| \le \frac{[\varphi^{(3)}]_\beta \|\sigma\|_\infty^{(3+\beta)} \mathbb{E}\big[|U_1|^{3+\beta}\big]}{(1+\beta)(2+\beta)(3+\beta)} \frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} \underset{n}{\longrightarrow} a_{\beta,\infty} > 0$ a.s. Also, for $\beta = 1$, the $(\mathcal{B}_{n,1})_{n\ge1}$ are exponentially integrable and if, furthermore, $D^3\varphi$ is $\mathcal{C}^1$, $\mathcal{B}_{n,1} \underset{n}{\to} -\widetilde{\gamma}m$ a.s. where $\widetilde{\gamma}m$ 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 6. *Observe that, when $\beta = 1$, the above result provides the exact 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 $\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2$. We refer to Section 4 for an improvement when $\|\sigma\|^2 - \nu(\|\sigma\|^2)$ is itself a coboundary.*

Importantly, these bounds do no require 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 (*[PV01], [RS76], [Vil09]*), or a confluence condition (*[PP12]*), when investigating the deviations for a given function f, in order to ensure the solvability of the corresponding Poisson equation $\mathcal{A}\varphi = f - \nu(f)$.*

We discuss in Section 5 further on some specific cases that yield the required smoothness in Theorem 2 for the solution φ of the Poisson equation $\mathcal{A}\varphi = f - \nu(f)$ corresponding to a given source f. Roughly speaking, if $\sigma\sigma^*$ is non-degenerate, $b \in \mathcal{C}^{1,\beta}(\mathbb{R}^d,\mathbb{R}^d), \sigma \in \mathcal{C}_b^{1,\beta}(\mathbb{R}^d,\mathbb{R}^d\otimes\mathbb{R}^d)$ and the source $f \in C^{1,\beta}(\mathbb{R}^d,\mathbb{R})$ for $\beta \in (0,1)$, we have that under **(A)**, *the solution φ of the Poisson problem satisfies the above assumptions (a) (i). Under confluence conditions, which allow degeneracies, we cannot hope to rely on an elliptic bootstrap effect and the required smoothness on φ actually imposes that the source itself is smooth.*

Remark 7 (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.*

Observe that we could derive from Theorem 2, up to the resolution of the Poisson equation $\mathcal{A}\varphi = f - \nu(f)$, some non-asymptotic confidence interval for $\nu_n(f)$. This point is discussed in Section 5.

2.2. **Strategy.** To control the deviations of $\nu_n(\mathcal{A}\varphi)$ 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 $(\psi_k(X_{k-1},\cdot))_{k\in[\![1,n]\!]}$ 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 $k \in [\![0, n-1]\!]$:*

$$\begin{aligned}\varphi(X_k) - \varphi(X_{k-1}) &= \gamma_k \mathcal{A}\varphi(X_{k-1}) + \Big[\gamma_k^2 \int_0^1 (1-t)\mathrm{Tr}\Big(D^2\varphi(X_{k-1} + t\gamma_k b_{k-1})b_{k-1} \otimes b_{k-1}\Big)dt \\ &\quad + \frac{1}{2}\gamma_k \,\mathrm{Tr}\Big(\big(D^2\varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2\varphi(X_{k-1})\big)\Sigma_{k-1}^2\Big) + \psi_k(X_{k-1}, U_k)\Big] \\ &=: \gamma_k \mathcal{A}(X_{k-1}) + \Big(\psi_k(X_{k-1}, U_k) + R_{n,k}^1(X_{k-1})\Big), \qquad (2.4)\end{aligned}$$

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

$$M_n := \sum_{k=1}^n \Delta_k(X_{k-1}, U_k). \qquad (2.5)$$

Summing over k yields:

$$\varphi(X_n) - \varphi(X_0) = \Gamma_n \nu_n(\mathcal{A}\varphi) + M_n + \sum_{k=1}^n R_{n,k}(X_{k-1}). \qquad (2.6)$$

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:

$$\nu_n(\mathcal{A}\varphi) = -\frac{1}{\Gamma_n}(M_n + R_n). \qquad (2.7)$$

- *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 $a, \lambda > 0$:

$$\begin{aligned}\mathbb{P}\big[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) \geq a\big] &\leq \exp\Big(-\frac{a\lambda}{\sqrt{\Gamma_n}}\Big)\mathbb{E}\left[\exp\Big(-\frac{\lambda}{\Gamma_n}(M_n + R_n)\Big)\right] \\ &\leq \exp\Big(-\frac{a\lambda}{\sqrt{\Gamma_n}}\Big)\mathbb{E}\left[\exp\Big(-\frac{q\lambda}{\Gamma_n}M_n\Big)\right]^{1/q}\mathbb{E}\left[\exp\Big(\frac{p\lambda}{\Gamma_n}|R_n|\Big)\right]^{1/p}, \frac{1}{p}+\frac{1}{q} = 1,\ p, q > 1. \qquad (2.8)\end{aligned}$$

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

$$\begin{aligned}\mathbb{E}\left[\exp\Big(-\frac{q\lambda}{\Gamma_n}M_n\Big)\right] &= \mathbb{E}\left[\exp\Big(-\frac{q\lambda}{\Gamma_n}M_{n-1}\Big)\mathbb{E}\left[\exp\Big(-\frac{q\lambda}{\Gamma_n}\Delta_{n-1}(X_{n-1}, U_n)\Big)\,|\,\mathcal{F}_{n-1}\right]\right] \\ &\leq \mathbb{E}\left[\exp\Big(-\frac{q\lambda}{\Gamma_n}M_{n-1}\Big)\right]\exp\left(\frac{\lambda^2 q^2}{2\Gamma_n^2}\gamma_n\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2\right) \leq \exp\left(\frac{\lambda^2 q^2}{2\Gamma_n}\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2\right), \qquad (2.9)\end{aligned}$$

iterating the procedure to derive the last identity. From (2.8), we thus get:

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

Keeping in mind that we manage to find $q := q(n) \downarrow_n 1$ such that the remainder $\mathscr{R}_n \downarrow_n 1$, 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 (2.7) give a non trivial asymptotic contribution. We choose to substract them before studying the deviation (term $\mathcal{B}_{n,\beta}$ in (2.3)).

2.3. **Explicit controls on the remainders.** Summing the increments appearing in (2.4), we now choose for the analysis to write for a given $n \in \mathbb{N}$ the remainder R_n defined after (2.6) 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:

$$\begin{aligned} D_{2,b,n} &:= \sum_{k=1}^n \gamma_k^2 \int_0^1 (1-t)\,\mathrm{Tr}\Big(D^2\varphi(X_{k-1} + t\gamma_k b_{k-1}) b_{k-1} \otimes b_{k-1}\Big) dt, \\ D_{2,\Sigma,n} &:= \frac{1}{2}\sum_{k=1}^n \gamma_k \mathrm{Tr}\Big(\big(D^2\varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2\varphi(X_{k-1})\big)\Sigma_{k-1}^2\Big), \\ \bar{G}_n &:= \sum_{k=1}^n \mathbb{E}\left[\psi_k(X_{k-1}, U_k)|\mathcal{F}_{k-1}\right], \\ L_n &:= \varphi(X_n) - \varphi(X_0). \end{aligned} \tag{2.10}$$

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 (2.7)

$$\nu_n(\mathcal{A}\varphi) = -\frac{1}{\Gamma_n}(M_n + R_n) = -\frac{1}{\Gamma_n}\big(M_n + (D_{2,b,n} + D_{2,\Sigma,n}) + \bar{G}_n - L_n\big). \tag{2.11}$$

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 (2.11), the exponential Tchebychev and Hölder inequalities yield that, for all $\lambda \in \mathbb{R}_+$ and all $p, q \in (1, +\infty)$, $\frac{1}{p} + \frac{1}{q} = 1$,

$$\begin{aligned} \mathbb{P}\big[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) \geq a\big] &\leq \exp\Big(-\frac{a\lambda}{\sqrt{\Gamma_n}}\Big)\left(\mathbb{E}\exp\Big(-\frac{q\lambda}{\Gamma_n}M_n\Big)\right)^{\frac{1}{q}} \\ &\times \left(\mathbb{E}\exp\Big(\frac{2p\lambda}{\Gamma_n}\big(|L_n| + |\bar{G}_n|\big)\Big)\right)^{\frac{1}{2p}} \left(\mathbb{E}\exp\Big(\frac{4p\lambda}{\Gamma_n}|D_{2,b,n}|\Big)\right)^{\frac{1}{4p}} \left(\mathbb{E}\exp\Big(\frac{4p\lambda}{\Gamma_n}|D_{2,\Sigma,n}|\Big)\right)^{\frac{1}{4p}}. \end{aligned} \tag{2.12}$$

(*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 (2.10) 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:

$$\begin{aligned} \mathbb{P}\big[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) + \mathcal{B}_{n,\beta} \geq a\big] &= \mathbb{P}\Big[\nu_n(\mathcal{A}\varphi) + \frac{\bar{G}_n + D_{2,n}}{\Gamma_n} \geq \frac{a}{\sqrt{\Gamma_n}}\Big] \\ &\leq \exp\Big(-\frac{a\lambda}{\sqrt{\Gamma_n}}\Big)\left(\mathbb{E}\exp\Big(-\frac{q\lambda}{\Gamma_n}M_n\Big)\right)^{\frac{1}{q}} \left(\mathbb{E}\exp\Big(\frac{p\lambda}{\Gamma_n}|L_n|\Big)\right)^{\frac{1}{p}}. \end{aligned} \tag{2.13}$$

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:

$$\begin{aligned} \mathbb{P}\big[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) + \mathcal{B}_{n,\beta} \geq a\big] &= \mathbb{P}\Big[\nu_n(\mathcal{A}\varphi) + \frac{\bar{G}_n}{\Gamma_n} \geq \frac{a}{\sqrt{\Gamma_n}}\Big] \leq \exp\Big(-\frac{a\lambda}{\sqrt{\Gamma_n}}\Big)\left(\mathbb{E}\exp\Big(-\frac{q\lambda}{\Gamma_n}M_n\Big)\right)^{\frac{1}{q}} \\ &\times \left(\mathbb{E}\exp\Big(\frac{2p\lambda}{\Gamma_n}|L_n|\Big)\right)^{\frac{1}{2p}} \left(\mathbb{E}\exp\Big(\frac{4p\lambda}{\Gamma_n}|D_{2,b,n}|\Big)\right)^{\frac{1}{4p}} \left(\mathbb{E}\exp\Big(\frac{4p\lambda}{\Gamma_n}|D_{2,\Sigma,n}|\Big)\right)^{\frac{1}{4p}}. \end{aligned} \tag{2.14}$$

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 $D^3\varphi \in \mathcal{C}^1$, $\mathcal{B}_{n,1} = \frac{\bar{G}_n + D_{2,n}}{\sqrt{\Gamma_n}} \underset{n}{\to} -\widetilde{\gamma}m$ 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 (2.12) - (2.14).

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

$$\lambda_n := \frac{a}{q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\sqrt{\Gamma_n}, \tag{2.15}$$

we derive:

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

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

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

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

$$\Big(\mathbb{E}\exp\Big(\frac{2p\lambda_n}{\Gamma_n}|\bar{G}_n|\Big)\Big)^{\frac{1}{2p}} \leq \exp\Big(\frac{\lambda_n}{\sqrt{\Gamma_n}}a_n\Big) \leq \exp\Big(\frac{\lambda_n^2}{2\Gamma_n p} + \frac{a_n^2 p}{2}\Big),\ \forall p > 1. \tag{2.16}$$

As indicated before, we now aim at controlling the remainders. In particular, from (2.8) and (2.10), we are led to handle terms of the form

$$\mathbb{E}\exp\Big(c\sum_{k=1}^n\gamma_k^2|b(X_{k-1})|^2\Big) \underset{(\mathcal{L}_{\mathbf{V}})}{\leq} \mathbb{E}\exp\Big(c\,C_{_V}\sum_{k=1}^n\gamma_k^2|V(X_{k-1})|\Big)$$

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 $\lambda \in [0, c_V]$, $\xi \in [0;1]$:*

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

We now have the following results for the terms appearing in (2.10).

Lemma 4 (Initial term)**.** *Let $q \in (1, +\infty)$ be fixed and λ_n be as in (2.15) in Lemma 2. For functions φ satisfying (G_V), i.e. there exists $C_{V,\varphi} > 0$ such that for all $x \in \mathbb{R}^d$, $|\varphi(x)| \leq C_{V,\varphi}(1 + \sqrt{V(x)})$, for $p := \frac{q}{q-1}$ and $j \in \{1,2\}$:*

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

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 $C_{2.17} := C_{2.17}((\mathbf{A}), \varphi)$ such that for $p = \frac{q}{q-1}$:*

$$\Big(\mathbb{E}\exp\Big(\frac{4p\lambda_n}{\Gamma_n}\big|D_{2,\Sigma,n}\big|\Big)\Big)^{\frac{1}{4p}} \leq \exp\left(C_{2.17}\frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{\Gamma_n^2}\right)(I_V^1)^{\frac{1}{4p}}. \tag{2.17}$$

We also have:

- *If for all $x \in \mathbb{R}^d$, $\big|D^2\varphi(x)\big| \leq \frac{C_\varphi}{1+|x|}$ for some constant $C_\varphi \geq 0$, then there exists $C_{2.18} := C((\mathbf{A}), \varphi) > 0$ such that*

$$\Big(\mathbb{E}\exp\Big(\frac{4p\lambda_n}{\Gamma_n}\big|D_{2,b,n}\big|\Big)\Big)^{\frac{1}{4p}} \leq \exp\left(C_{2.18}\Big(\frac{3p\lambda_n^2(\Gamma_n^{(2)})^2}{2\Gamma_n^2} + \frac{1}{2p}\Big)\right)(I_V^1)^{\frac{1}{4p}}. \tag{2.18}$$

- *For* $a \leq \frac{c_v q}{4C_V p}\frac{\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}{\|D^2\varphi\|_\infty^2}\frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$, *there exists an* $\mathbb{R}^+$*-valued sequence* $(v_n)_{n\geq 1}$ *such that* $\big|v_n\big| \leq C_{2.19} := C_{2.19}((\mathbf{A}),\varphi)$ *and*

$$\Big(\mathbb{E}\exp\Big(\frac{4p\lambda_n}{\Gamma_n}\big|D_{2,b,n}\big|\Big)\Big)^{\frac{1}{4p}} \leq (I_V^1)^{v_n}. \tag{2.19}$$

Also, $v_n \underset{n}{\to} v_\infty$ *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:

$$\Big(\mathbb{E}\exp\Big(-q\lambda_n\frac{M_n}{\Gamma_n}\Big)\Big)^{\frac{1}{q}}\exp\Big(-\frac{a\lambda_n}{\sqrt{\Gamma_n}}\Big) \leq \exp\Big(-\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\Big). \tag{2.20}$$

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

(i) We suppose that there exists $C_\varphi > 0$ such that for all $x \in \mathbb{R}^d$, $|D^2\varphi(x)| \leq C_\varphi/(1+|x|)$. Plugging in (2.12) the controls from (2.20), Lemma 3 equation (2.16), Lemma 4 (with $j=2$) and Lemma 5 (equations (2.17), (2.18)), we get:

$$\begin{aligned}
\mathbb{P}\Big[\nu_n(\mathcal{A}\varphi) \geq \frac{a}{\sqrt{\Gamma_n}}\Big] &\leq \exp\Big(-\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\Big)\exp\Big(\frac{\lambda_n^2}{2\Gamma_n p}+\frac{pa_n^2}{2}\Big)\exp\left(\frac{3pC_{V,\varphi}^2\lambda_n^2}{c_V\Gamma_n^2}+\frac{c_V}{p}\right)(I_V^1)^{\frac{1}{2p}} \\
&\quad\times\exp\left(C_{2.17}\frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{\Gamma_n^2}\right)(I_V^1)^{\frac{1}{4p}}\times\exp\left(C_{2.18}\Big(\frac{3p\lambda_n^2(\Gamma_n^{(2)})^2}{2\Gamma_n^2}+\frac{1}{2p}\Big)\right)(I_V^1)^{\frac{1}{4p}} \\
&\leq (I_V^1)^{\frac{1}{p}}\exp\left(-\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\Big(1-\frac{1}{q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\Big\{\frac{p}{\Gamma_n}\Big(\frac{6C_{V,\varphi}^2}{c_V}+\big[2C_{2.17}+3C_{2.18}\big](\Gamma_n^{(2)})^2\Big)+\frac{1}{p}\Big\}\Big)\right) \\
&\quad\times\exp\Big(\frac{1}{p}\big[c_V+\frac{C_{2.18}}{2}\big]+\frac{pa_n^2}{2}\Big).
\end{aligned} \tag{2.21}$$

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

$$d_n := \frac{1}{q_n\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\Big\{\frac{p_n}{\Gamma_n}\Big(\frac{6C_{V,\varphi}^2}{c_V}+\big[2C_{2.17}+3C_{2.18}\big](\Gamma_n^{(2)})^2\Big)+\frac{1}{p_n}\Big\} \underset{n}{\to} 0. \tag{2.22}$$

We conclude from (2.21) setting $c_n = q_n^{-1}(1-d_n), C_n := (I_V^1)^{\frac{1}{p_n}}\exp(\frac{1}{p_n}[c_V+\frac{C_{2.18}}{2}]+\frac{p_n a_n^2}{2}) \underset{n}{\to} 1$. 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 $a \leq \frac{c_V q}{4C_V p}\frac{\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}{\|D^2\varphi\|_\infty^2}\frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$. Plugging in (2.12) the controls from (2.20), Lemma 3, equation (2.16) , Lemmas 4 (with $j=2$), 5 (equations (2.17), (2.19)) we then derive:

$$\begin{aligned}
\mathbb{P}\Big[\nu_n(\mathcal{A}\varphi) \geq \frac{a}{\sqrt{\Gamma_n}}\Big] &\leq \exp\Big(-\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\Big)\exp\Big(\frac{\lambda_n^2}{2\Gamma_n p}+\frac{pa_n^2}{2}\Big)\exp\left(\frac{3pC_{V,\varphi}^2\lambda_n^2}{c_V\Gamma_n^2}+\frac{c_V}{p}\right)(I_V^1)^{\frac{1}{2p}} \\
&\quad\times\exp\left(C_{2.17}\frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{\Gamma_n^2}\right)(I_V^1)^{\frac{1}{4p}}(I_V^1)^{v_n} \\
&\leq (I_V^1)^{v_n+\frac{3}{4p}}\exp\left(\frac{c_V}{p}+\frac{pa_n^2}{2}\right)\exp\left(-\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\Big(1-\frac{1}{q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\Big\{\frac{p}{\Gamma_n}\Big(\frac{6C_{V,\varphi}^2}{c_V}+2C_{2.17}(\Gamma_n^{(2)})^2\Big)+\frac{1}{p}\Big\}\Big)\right).
\end{aligned} \tag{2.23}$$

Since $\theta > \frac{1}{2+\beta} \geq 1/3$ (see Remark 2), we again take $p := p_n \uparrow_n +\infty$ so that $p_n^{1/2} a_n \underset{n}{\to} 0$ which also guarantees:

$$d_n := \frac{1}{q_n \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2} \Big\{ p_n \Big(\frac{6C_{V,\varphi}^2}{c_V \Gamma_n} + \frac{2C_{2.17}(\Gamma_n^{(2)})^2}{\Gamma_n} \Big) + \frac{1}{p_n} \Big\} \underset{n}{\to} 0. \tag{2.24}$$

In this case, we derive the result by setting $c_n := q_n^{-1}(1-d_n) \underset{n}{\to} 1$, $C_n := (I_V^1)^{v_n + \frac{3}{4p_n}} \exp(\frac{c_V}{p_n} + \frac{p_n a_n^2}{2}) \underset{n}{\to} 1$ (see the limits of v_n following equation (2.19) and (3.11)). Again, $(p_n)_{n\geq 1}$ can be chosen in order to have the stated monotonicity for n large enough. Set now

$$\chi_n := \frac{c_V \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}{4C_{_V} \|D^2\varphi\|_\infty^2} \frac{q_n}{p_n}, \tag{2.25}$$

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 (2.13), the controls of (2.20) and Lemma 4 (with $j=1$) we get:

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

Recalling the definition of λ_n in (2.15), 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 (2.14).

Also, when $D^3\varphi \in \mathcal{C}^1$, we derive similarly to the proof of Theorem 10 in [LP02] that $\mathcal{B}_{n,1} \underset{n}{\to} -\widetilde{\gamma} m$.

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

3. Proof of the Technical Lemmas

Proof of Lemma 1. For $k \in [\![1,n]\!]$, we first write:

$$\begin{aligned} \varphi(X_k) - \varphi(X_{k-1}) &= (\varphi(X_k) - \varphi(X_{k-1} + \gamma_k b_{k-1})) + (\varphi(X_{k-1} + \gamma_k b_{k-1}) - \varphi(X_{k-1})) \\ &=: T_{k-1,r}(\varphi) + T_{k-1,d}(\varphi), \end{aligned} \tag{3.1}$$

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 (3.1). Namely, with the above notations:

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

Hence,

$$\begin{aligned} \varphi(X_k) - \varphi(X_{k-1}) &= \gamma_k \mathcal{A}\varphi(X_{k-1}) \\ &\quad + \gamma_k^2 \int_0^1 (1-t) \mathrm{Tr}\Big(D^2\varphi(X_{k-1} + t\gamma_k b_{k-1}) b_{k-1} \otimes b_{k-1} \Big) dt + \sqrt{\gamma_k}\sigma_{k-1}U_k \cdot \nabla\varphi(X_{k-1} + \gamma_k b_{k-1}) \\ &\quad + \gamma_k \int_0^1 (1-t) \mathrm{Tr}\Big(D^2\varphi(X_{k-1} + \gamma_k b_{k-1} + t\sqrt{\gamma_k}\sigma_{k-1}U_k) \sigma_{k-1}U_k \otimes U_k \sigma_{k-1}^* - D^2\varphi(X_{k-1})\Sigma_{k-1} \Big) dt \\ &= \gamma_k \mathcal{A}\varphi(X_{k-1}) + \gamma_k^2 \int_0^1 (1-t) \mathrm{Tr}\Big(D^2\varphi(X_{k-1} + t\gamma_k b_{k-1}) b_{k-1} \otimes b_{k-1} \Big) dt \\ &\quad + \gamma_k \int_0^1 (1-t) \mathrm{Tr}\Big(\big(D^2\varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2\varphi(X_{k-1}) \big) \Sigma_{k-1} \Big) dt + \psi_k(X_{k-1}, U_k) \\ &=: \gamma_k \mathcal{A}\varphi(X_{k-1}) + D_{2,b}^k + D_{2,\Sigma}^k + \psi_k(X_{k-1}, U_k), \end{aligned} \tag{3.2}$$

where

$$\psi_k(X_{k-1},U_k)=\sqrt{\gamma_k}\sigma_{k-1}U_k\cdot\nabla\varphi(X_{k-1}+\gamma_k b_{k-1}) \tag{3.3}$$
$$+\gamma_k\int_0^1(1-t)\mathrm{Tr}\Big(D^2\varphi(X_{k-1}+\gamma_k b_{k-1}+t\sqrt{\gamma_k}\sigma_{k-1}U_k)\sigma_{k-1}U_k\otimes U_k\sigma_{k-1}^*-D^2\varphi(X_{k-1}+\gamma_k b_{k-1})\Sigma_{k-1}\Big)dt.$$

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 (3.2). Consequently, as φ is Lispchitz continuous we derive, for all $(u,u')\in(\mathbb{R}^d)^2$:

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

The result is obtained by summing up the previous identities from $k=1$ to n, observing, with the notations of (2.10), that $L_n=\sum_{k=1}^n\varphi(X_k)-\varphi(X_{k-1}),D_{2,b,n}=\sum_{k=1}^n D_{2,b}^k$, $D_{2,\Sigma,n}=\sum_{k=1}^n D_{2,\Sigma}^k$, $G_n:=\sum_{k=1}^n\psi_k(X_{k-1},U_k)$. □

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 [FM12]. In the current framework, we exploit that the functions $u\mapsto\Delta_k(X_{k-1},u):=\psi_k(X_{k-1},u)-\mathbb{E}\left[\psi_k(X_{k-1},U_k)|\mathcal{F}_{k-1}\right]$ 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:

$$\begin{aligned}\mathbb{E}\exp\Big(-\frac{q\lambda}{\Gamma_n}M_n\Big) &= \mathbb{E}\exp\left(-\frac{q\lambda}{\Gamma_n}\sum_{k=1}^n\Delta_k(X_{k-1},U_k)\right)\\ &= \mathbb{E}\Big[\exp\Big(-\frac{q\lambda}{\Gamma_n}\sum_{k=1}^{n-1}\Delta_k(X_{k-1},U_k)\Big)\mathbb{E}\Big[\exp\Big(-\frac{q\lambda}{\Gamma_n}\Delta_n(X_{n-1},U_n)\Big)|\mathcal{F}_{n-1}\Big]\Big]\\ &\le \mathbb{E}\Big[\exp\Big(-\frac{q\lambda}{\Gamma_n}\sum_{k=1}^{n-1}\Delta_k(X_{k-1},U_k)\Big)\exp\Big(\frac{q^2\lambda^2}{2\Gamma_n^2}\gamma_n\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2\Big)\Big], \qquad (3.4)\end{aligned}$$

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

Iterating the process over k, we obtain:

$$\Big(\mathbb{E}\exp\Big(-\frac{q\lambda}{\Gamma_n}M_n\Big)\Big)^{\frac{1}{q}}=\left(\mathbb{E}\exp\Big(-\frac{q\lambda}{\Gamma_n}\sum_{k=1}^n\Delta_k(X_{k-1},U_k)\Big)\right)^{\frac{1}{q}}\le\exp\Big(\frac{q\lambda^2\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}{2\Gamma_n}\Big). \tag{3.5}$$

Finally,

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

where $g:\mathbb{R}^+\to\mathbb{R}$ 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 (2.15). This eventually yields the expected bound. □

Proof of Lemma 3. From the definition in (3.3) and the Fubini theorem, we have that for all $k\in[\![1,n]\!]$:

$$\begin{aligned}\mathbb{E}\left[\psi_k(X_{k-1},U_k)|\mathcal{F}_{k-1}\right] &= \gamma_k\int_0^1(1-t)\mathrm{Tr}\Big(\mathbb{E}\big[D^2\varphi(X_{k-1}+\gamma_k b_{k-1}+t\sqrt{\gamma_k}\sigma_{k-1}U_k)\sigma_{k-1}U_k\otimes U_k\sigma_{k-1}^*\\ &\qquad -D^2\varphi(X_{k-1}+\gamma_k b_{k-1})\Sigma_{k-1}|\mathcal{F}_{k-1}\big]\Big)dt. \qquad (3.6)\end{aligned}$$

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:

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

recalling from **(GC)** that for all $(i,j,l)\in[\![1,r]\!]$, $\mathbb{E}\left[U_k^iU_k^jU_k^l|\mathcal{F}_{k-1}\right]=\mathbb{E}\left[U_1^iU_1^jU_1^l\right]=0$ (cancellation argument). Hence,

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

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

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

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 all $n\in\mathbb{N}$:

$$\begin{aligned}
V(X_n)-V(X_{n-1}) &= \gamma_n\mathcal{A}V(X_{n-1})+\gamma_n^2\int_0^1(1-t)\mathrm{Tr}\Big(D^2V(X_{n-1}+t\gamma_n b_{n-1})b_{n-1}\otimes b_{n-1}\Big)dt\\
&\quad-\frac{\gamma_n}{2}\mathrm{Tr}\big(D^2V(X_{n-1}))\Sigma_{n-1}\big)+\sqrt{\gamma_n}\sigma_{n-1}U_n\cdot\nabla V(X_{n-1}+\gamma_n b_{n-1})\\
&\quad+\gamma_n\int_0^1(1-t)\mathrm{Tr}\Big(D^2V(X_{n-1}+\gamma_n b_{n-1}+t\sqrt{\gamma_n}\sigma_{n-1}U_n)\sigma_{n-1}U_n\otimes U_n\sigma_{n-1}^*\Big)dt\\
&\le -\gamma_n\alpha_{_V}V(X_{n-1})+\gamma_n\beta_V+C_{_V}\frac{\gamma_n^2}{2}\|D^2V\|_\infty V(X_{n-1})\\
&\quad+\frac{\gamma_n}{2}\|D^2V\|_\infty\|\sigma\|_\infty^2+\sqrt{\gamma_n}\sigma_{n-1}U_n\cdot\nabla V(X_{n-1}+\gamma_n b_{n-1})+\frac{\gamma_n}{2}\|D^2V\|_\infty\|\sigma\|_\infty^2|U_n|^2\\
&\le \gamma_n\Big(-\frac{\alpha_{_V}}{2}V(X_{n-1})+\widetilde{c}\Big)+\sqrt{\gamma_n}\sigma_{n-1}U_n\cdot\nabla V(X_{n-1}+\gamma_n b_{n-1})+\frac{\gamma_n}{2}\|D^2V\|_\infty\|\sigma\|_\infty^2|U_n|^2 \qquad (3.7)
\end{aligned}$$

for a constant $\widetilde{c}:=\widetilde{c}(V,\sigma,\beta_V)$. We have in fact considered the time steps sufficiently small (in **(S)**, we have chosen for all $n\in\mathbb{N}$, $\gamma_n<\min(\frac{1}{2\sqrt{C_{_V}\widetilde{c}}},\frac{\alpha_{_V}}{2C_{_V}\|D^2V\|_\infty})$). 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 all $\gamma,\lambda>0$ the quantities:

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

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

$$I_1(\gamma_n,\lambda,x) \;\leq\; \exp\Big(\frac{\lambda^2\gamma_n|\sigma^*(x)\nabla V(x+\gamma_n b(x))|^2}{2}\Big) \underset{(\mathcal{L}_{\mathbf{V}})}{\leq} \exp\Big(\frac{\lambda^2\gamma_n C_{_V}\|\sigma\|_\infty^2 V(x+\gamma_n b(x))}{2}\Big). \tag{3.8}$$

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

$$I_2(\gamma_n,\lambda) \;\leq\; \Big[\mathbb{E}\exp\big(c|U_n|^2\big)\Big]^{\frac{\lambda\gamma_n\|D^2V\|_\infty\|\sigma\|_\infty^2}{2c}} = \exp\Big(\gamma_n\ln(I_c)\frac{\lambda\|D^2V\|_\infty\|\sigma\|_\infty^2}{2c}\Big). \tag{3.9}$$

These controls allow to prove the integrability statement of the proposition by induction. For $n=0$, recalling from assumption **(C1)** that for all $\lambda<\lambda_0, \mathbb{E}\exp(\lambda|X_0|^2)<+\infty$ and from **(S)** 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 $C^0_{V,\lambda}\in(1,+\infty)$ such that

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

Set now $\widetilde{\beta}_V:=\widetilde{c}+\ln(I_c)\frac{\|D^2V\|_\infty\|\sigma\|_\infty^2}{2c}$ and $\widetilde{\alpha}_V:=\min\big(\frac{1}{\gamma_1},\frac{\alpha_V}{2}-\lambda C_{_V}\|\sigma\|_\infty^2(1+\gamma_1 C_{_V}[1+\frac{\gamma_1\|D^2V\|_\infty}{2}])\big)\in(0,\frac{1}{\gamma_1}]$, for $\lambda<\frac{\alpha_V}{2C_{_V}\|\sigma\|_\infty^2(1+\gamma_1C_{_V}[1+\frac{\gamma_1\|D^2V\|_\infty}{2}])}$.

Let us assume that for all $\lambda<\lambda_V:=\min\Big(\frac{\lambda_0}{2\bar{c}},\frac{\alpha_V}{2C_{_V}\|\sigma\|_\infty^2(1+\gamma_1C_{_V}[1+\frac{\gamma_1\|D^2V\|_\infty}{2}])},\frac{c}{\|D^2V\|_\infty\|\sigma\|_\infty^2\gamma_1}\Big)$, the property

$$(\mathcal{P}_{n-1})\qquad \forall k\in[\![0,n-1]\!],\ \mathbb{E}\exp\big(\lambda V(X_k)\big)\leq C_{V,\lambda}:=C^0_{V,\lambda}\vee\exp\Big(\frac{\lambda\widetilde{\beta}_V}{\widetilde{\alpha}_V}\Big),$$

holds for a fixed $n-1\in\mathbb{N}_0$ and let us prove $(\mathcal{P}_n)$. By inequalities (3.7), (3.8) and (3.9) and the Cauchy-Schwarz inequality, we derive that for all $\lambda<\lambda_V$,

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

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

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

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

$$\exp\big(\lambda\gamma_n\widetilde{\beta}_V\big)C_{V,\lambda}^{(1-\gamma_n\widetilde{\alpha}_V)}\leq C_{V,\lambda}\iff C_{V,\lambda}\geq\exp\Big(\frac{\lambda\widetilde{\beta}_V}{\widetilde{\alpha}_V}\Big).$$

Hence, $(\mathcal{P}_n)$ 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 for all $\lambda<\lambda_V(v^*)^{1-\xi}$:*

$$\mathbb{E}\exp(\lambda V_n^\xi)=\mathbb{E}\exp\Big(\lambda(v^*)^\xi\underbrace{\Big(\frac{V_n}{v^*}\Big)^\xi}_{\geq1}\Big)\leq\mathbb{E}\exp\big(\lambda(v^*)^{\xi-1}V_n\big)\leq C_{V,\lambda(v^*)^{\xi-1}}<+\infty.$$

Thus, we readily get as a byproduct 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 [Lem05] *for additional results in that direction.*

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

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

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

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

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

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

Proof of Lemma 5.

- *Proof of inequalities* (2.19) *and* (2.18).
  Write first

$$\Big(\mathbb{E}\exp\Big(\frac{4p\lambda_n}{\Gamma_n}|D_{2,b,n}|\Big)\Big)^{\frac{1}{4p}} \leq \left(\mathbb{E}\exp\Big(\frac{4p\lambda_n}{\Gamma_n}\sum_{k=1}^n\gamma_k^2\int_0^1(1-t)\Big|\mathrm{Tr}\Big(D^2\varphi(X_{k-1}+t\gamma_k b_{k-1})b_{k-1}\otimes b_{k-1}\Big)\Big|dt\Big)\right)^{\frac{1}{4p}}. \tag{3.10}$$

- *If* $a \leq \frac{c_V q}{4C_{V}p}\frac{\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}{\|D^2\varphi\|_\infty^2}\frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}} = \chi_n\frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$ *with the notation introduced in* (2.25). We first easily get from the assumptions on φ and point ii) of $(\mathcal{L}_V)$ that:

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

From 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:

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

We then have from the definition of λ_n in (2.15) 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 all $k \in [\![1,n]\!]$:

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

Thus, setting

$$v_n := \frac{\bar{v}_n}{4p} = \frac{\lambda_n\Gamma_n^{(2)}}{2\Gamma_n}\|D^2\varphi\|_\infty\frac{C_{V}}{c_V}, \tag{3.11}$$

we finally derive,

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

using again the notations of Proposition 1. This gives (2.19).

- *If* $\|D^2\varphi(x)\| \le \frac{C}{1+|x|}$. In this case, restarting from (3.10), the idea consists in introducing the partition $\{|X_k|\le K\}\cup\{|X_k|>K\}$ for a given constant $K>0$. The Lyapunov function V is then bounded on the compact set and for $\{|X_k|>K\}$ we write:

$$|D^2\varphi(X_k+t\gamma_k b_k)| \le \frac{C}{1+|X_k+t\gamma_k b_k|}.$$

This control will allow to equilibrate the growth of V. Recall indeed from $(\mathcal{L}_{\mathbf{V}})$ that for a given threshold $K>0$ there are constants $C_{_V}, \bar c := \bar c(V,K)>0$ such that for every $x\in\mathbb{R}^d$, $|x|\ge K$, $|V(x)|\le \bar c|x|^2$, $|b(x)|\le \sqrt{\bar c\, C_{_V}}|x|$ (see also Remark 1). Hence, since $t\in[0,1]$,

$$|X_k+t\gamma_k b_k| \ge (|X_k|-\gamma_k|b_k|)_+ \ge |X_k|(1-\gamma_k\sqrt{\bar c\, C_{_V}})_+.$$

From assumption **(S)**, we have that for $k\ge 1$, $(1-\gamma_k\sqrt{\bar c\, C_{_V}})\ge 1/2$. Thus:

$$|D^2\varphi(X_k+t\gamma_k b_k)| \le \|D^2\varphi\|_\infty \mathbf{1}_{\{|X_k|\le K\}} + \frac{2C}{|X_k|}\mathbf{1}_{\{|X_k|>K\}}. \tag{3.12}$$

From this inequality, equation (3.10) and applying as above the Jensen inequality, we derive:

$$\begin{aligned}\Big(\mathbb{E}\exp\Big(\frac{4p\lambda_n}{\Gamma_n}|D_{2,b,n}|\Big)\Big)^{\frac{1}{4p}} &\le \Big(\frac{1}{\Gamma_n^{(2)}}\sum_{k=1}^n\gamma_k^2\mathbb{E}\exp\Big(\frac{4p\lambda_n\Gamma_n^{(2)}}{\Gamma_n}C_{_V}V_{k-1}\Big(\frac{\|D^2\varphi\|_\infty}{2}\mathbb{1}_{|X_{k-1}|\le K}+\frac{C}{|X_{k-1}|}\mathbb{1}_{|X_{k-1}|>K}\Big)\Big)\Big)^{\frac{1}{4p}}\\ &\le \exp\Big(\frac{\lambda_n\Gamma_n^{(2)}}{2\Gamma_n}C_{_V}\big(\sup_{|x|\le K}V(x)\big)\|D^2\varphi\|_\infty\Big)\Big(\frac{1}{\Gamma_n^{(2)}}\sum_{k=1}^n\gamma_k^2\mathbb{E}\exp\Big(\frac{4p\lambda_n\Gamma_n^{(2)}}{\Gamma_n}CC_{_V}\sqrt{\bar c}\sqrt{V_{k-1}}\Big)\Big)^{\frac{1}{4p}},\end{aligned}$$

recalling that for $|x|>K$, $\sqrt{V(x)}\le\sqrt{\bar c}|x|$. From the Young inequality we obtain:

$$\mathbb{E}\exp\Big(\frac{4p\lambda_n\Gamma_n^{(2)}}{\Gamma_n}CC_{_V}\sqrt{\bar c}\sqrt{V_{k-1}}\Big) \le \exp\Big(\Big(\frac{2p\lambda_n\Gamma_n^{(2)}}{\Gamma_n}\frac{CC_{_V}\sqrt{\bar c}}{\sqrt{c_V}}\Big)^2\Big)\mathbb{E}\left[\exp(c_V V_{k-1})\right].$$

We finally derive with the notations of Proposition 1:

$$\begin{aligned}\Big(\mathbb{E}\exp\Big(\frac{4p\lambda_n}{\Gamma_n}|D_{2,b,n}|\Big)\Big)^{\frac{1}{4p}} &\le \exp\Big(\frac{\lambda_n\Gamma_n^{(2)}}{2\Gamma_n}C_{_V}\big(\sup_{|x|\le K}V(x)\big)\|D^2\varphi\|_\infty\Big)\exp\Big(\frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{\Gamma_n^2}\frac{(CC_{_V})^2\bar c}{c_V}\Big)(I_V^1)^{\frac{1}{4p}}\\ &\le \exp\Big(C_{2.18}\Big(\frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{\Gamma_n^2}+\frac{\lambda_n\Gamma_n^{(2)}}{\Gamma_n}\Big)\Big)(I_V^1)^{\frac{1}{4p}},\end{aligned}$$

setting $C_{2.18} := \frac{C_{_V}}{2}\big(\sup_{|x|\le K}V(x)\big)\|D^2\varphi\|_\infty \vee \frac{(CC_{_V})^2\bar c}{c_V}$. From the Young inequality we finally have $\frac{\lambda_n\Gamma_n^{(2)}}{\Gamma_n}\le \frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{2\Gamma_n^2}+\frac{1}{2p}$ which, plugged in the above control, completes the proof of (2.18). □

- *Proof of inequality* (2.17). We proceed as for the proof of (2.19) and (2.18). Write:

$$\begin{aligned}
&\Big(\mathbb{E}\exp\Big(\frac{4p\lambda_n}{\Gamma_n}|D_{2,\Sigma,n}|\Big)\Big)^{\frac{1}{4p}}\\
&\qquad\le\Big(\mathbb{E}\exp\Big(\frac{4p\lambda_n}{\Gamma_n}\sum_{k=1}^n\frac{\gamma_k}{2}\Big|\mathrm{Tr}\Big(\big(D^2\varphi(X_{k-1}+\gamma_k b_{k-1})-D^2\varphi(X_{k-1})\big)\Sigma_{k-1}\Big)\Big|\Big)\Big)^{\frac{1}{4p}}\\
&\qquad\le\Big(\mathbb{E}\exp\Big(\frac{2p\lambda_n}{\Gamma_n}\|\sigma\|_\infty^2[\varphi^{(2)}]_1\sum_{k=1}^n\gamma_k^2|b_{k-1}|\Big)\Big)^{\frac{1}{4p}}\le\Big(\mathbb{E}\exp\Big(\frac{2p\lambda_n}{\Gamma_n}\|\sigma\|_\infty^2[\varphi^{(2)}]_1C_V^{\frac12}\sum_{k=1}^n\gamma_k^2|V_{k-1}|^{\frac12}\Big)\Big)^{\frac{1}{4p}}\\
&\qquad\le\Big(\frac{1}{\Gamma_n^{(2)}}\sum_{k=1}^n\gamma_k^2\mathbb{E}\exp\Big(\frac{2p\lambda_n\Gamma_n^{(2)}}{\Gamma_n}\|\sigma\|_\infty^2[\varphi^{(2)}]_1C_V^{\frac12}|V_{k-1}|^{\frac12}\Big)\Big)^{\frac{1}{4p}}.
\end{aligned}$$

Using once again the Young inequality and setting $C_{2.17}:=\frac{\|\sigma\|_\infty^4[\varphi^{(2)}]_1^2}{4}\frac{C_V}{c_V}$, we obtain:

$$\Big(\mathbb{E}\exp\Big(\frac{4p\lambda_n}{\Gamma_n}|D_{2,\Sigma,n}|\Big)\Big)^{\frac{1}{4p}}\le\exp\Big(\frac{p\lambda_n^2}{4}\Big(\frac{\Gamma_n^{(2)}}{\Gamma_n}\Big)^2\|\sigma\|_\infty^4[\varphi^{(2)}]_1^2\frac{C_V}{c_V}\Big)(I_V^1)^{\frac{1}{4p}}\le\exp\Big(C_{2.17}p\lambda_n^2\Big(\frac{\Gamma_n^{(2)}}{\Gamma_n}\Big)^2\Big)(I_V^1)^{\frac{1}{4p}}.\qquad\square$$

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 $\mathcal{A}\vartheta=\|\sigma\|^2-\nu(\|\sigma\|^2)$ satisfying the assumptions stated in Theorem 2, which is in particular the case if σ is smooth enough (namely $\sigma\in C^{1,\beta}(\mathbb{R}^d,\mathbb{R}^d\otimes\mathbb{R}^r),\beta\in(0,1]$) and the coefficients in (1.1) satisfy some non-degeneracy conditions, see *e.g.* Theorem 6 below.

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

Theorem 3. *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 $\tilde c_n\le1\le\tilde C_n$, $n\ge1$, with $\lim_n\tilde C_n=\lim_n\tilde c_n=1$ such that for all $n\ge1$ for all $a>0$:

$$\begin{aligned}
\mathbb{P}\big[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)|\ge a\big]&\le 2\,\widetilde C_n\exp\Big(-\frac{\widetilde c_n}{2\nu(\|\sigma\|^2)\|\nabla\varphi\|_\infty^2}\,\Phi_n(a)\Big),\\
\Phi_n(a)&:=\left[\Big(a^2\Big(1-\frac{2}{1+\sqrt{1+4\,\bar c_n^3\,\frac{\Gamma_n}{a^2}}}\Big)\Big)\vee\Big(a^{\frac43}\Gamma_n^{\frac{\bar c_n}{3}}\Big(1-\frac23\bar c_n\Big(\frac{\Gamma_n}{a^2}\Big)^{\frac13}\Big)_+\Big)\right],
\end{aligned}$$

where $x_+=\max(x,0)$ and $\bar c_n:=\Big(\frac{[\varphi]_1}{[\vartheta]_1}\Big)^{2/3}\nu(\|\sigma\|^2)\|\sigma\|_\infty^{-2/3}\check c_n$ with $\check c_n$ being an explicit positive sequence s.t. $\check c_n\downarrow_n1$.
(b) For $\beta\in(0,1],\theta=\frac{1}{2+\beta}$, there exist two explicit monotonic sequences $\tilde c_n\le1\le\tilde C_n$, $n\ge1$, with $\lim_n\tilde C_n=\lim_n\tilde c_n=1$ such that for all $n\ge1$ for all $a>0$:

$$\mathbb{P}\big[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)+\mathcal{B}_{n,\beta}|\ge a\big]\le2\,\widetilde C_n\exp\Big(-\frac{\widetilde c_n}{2\nu(\|\sigma\|^2)\|\nabla\varphi\|_\infty^2}\Phi_n(a)\Big).$$

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 $\frac{2}{1+\sqrt{1+4\bar c_n^3\frac{\Gamma_n}{a^2}}}\underset{n}{\to}0$. 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. $a^2\big(1-\frac{2}{1+\sqrt{1+4\bar c_n^3\frac{\Gamma_n}{a^2}}}\big)\underset{a\to\infty}{\longrightarrow}\bar c_n^3\Gamma_n$ 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 $\|\sigma\|_\infty^2$ appearing in Theorem 2 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.

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 2.2 that give for all $\lambda > 0$ the control in equation (2.12). Let us now deal with the term giving the concentration and write for all $\rho > 1$:

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

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

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

The key idea is that we have exploited the Poisson equation solved by ϑ to replace the previous *rough* control $\exp\Big(\frac{(q\lambda)^2[\varphi]_1^2\|\sigma\|_\infty^2}{2\Gamma_n}\Big)$ coming from the martingale increment obtained in equation (2.9) by the above term $\exp\Big(\frac{\rho^2(q\lambda)^2[\varphi]_1^2\nu(\|\sigma\|^2)}{2\Gamma_n}\Big)$. 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 measure. The point for $\mathscr{T}_1$ 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 $\widetilde{\mathscr{T}_1} := \exp\Big(-\frac{\rho^2(q\lambda)^2[\varphi]_1^2\nu(\|\sigma\|^2)}{2\Gamma_n}\Big)\mathscr{T}_1$. Define now, for a given $n \in \mathbb{N}$ and $m \in \mathbb{N}_0$, $S_m := \exp\Big(-\rho\frac{q\lambda}{\Gamma_n}M_m - \frac{\rho^2(q\lambda)^2[\varphi]_1^2}{2\Gamma_n^2}\sum_{k=1}^m \gamma_k\|\sigma(X_{k-1})\|^2\Big)$. From the definition of the martingale $(M_k)_{k\ge1}$ in (2.5) and the controls of the Lipschitz constants of the functions $\big(\psi_k(X_{k-1},\cdot)\big)_{k\in[\![1,n]\!]}$ in Lemma 1, we get by iterated conditioning:

$$\begin{aligned}\widetilde{\mathscr{T}_1} \;&\le\; \mathbb{E}\Big[S_{n-1}\exp\Big(-\frac{\rho^2(q\lambda)^2[\varphi]_1^2}{2\Gamma_n^2}\gamma_n\|\sigma(X_{n-1})\|^2\Big)\mathbb{E}\Big[\exp\Big(-\rho\frac{q\lambda}{\Gamma_n}(M_n - M_{n-1})\Big)\Big|\mathcal{F}_{n-1}\Big]\Big] \\ &\underset{(\mathbf{GC})}{\le}\; \mathbb{E}\Big[S_{n-1}\exp\Big(-\frac{\rho^2(q\lambda)^2[\varphi]_1^2}{2\Gamma_n^2}\gamma_n\|\sigma(X_{n-1})\|^2\Big)\exp\Big(\frac{\rho^2(q\lambda)^2}{2\Gamma_n^2}\gamma_n[\varphi]_1^2\|\sigma(X_{n-1})\|^2\Big)\Big] \le \mathbb{E}\left[S_{n-1}\right] \le 1.\end{aligned}$$

In other words, $(S_m)_{m\ge0}$ is a positive supermartingale. We finally get that, for all $\rho > 1$:

$$\mathscr{T}_1^{\frac{1}{\rho}} \le \exp\Big(\frac{\rho(q\lambda)^2[\varphi]_1^2\nu(\|\sigma\|^2)}{2\Gamma_n}\Big). \qquad (4.2)$$

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

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

so that this contribution can be controlled from the previous expansion of Lemma 1 exploiting the technical lemmas of Section 2.2 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 all $\bar p, \bar q \in (1,+\infty)$, $\frac{1}{\bar p}+\frac{1}{\bar q}=1$, similarly to (2.12),

$$\mathscr{T}_2 = \mathbb{E}\exp\Big(\frac{\mu}{\Gamma_n}\sum_{k=1}^n \gamma_k \mathcal{A}\vartheta(X_{k-1})\Big) \le \Big(\mathbb{E}\exp\Big(-\frac{\bar q\mu}{\Gamma_n}M_n^\vartheta\Big)\Big)^{\frac{1}{\bar q}}$$

$$\times \Big(\mathbb{E}\exp\Big(\frac{2\bar p\mu}{\Gamma_n}|L_n^\vartheta|\Big)\Big)^{\frac{1}{2\bar p}} \Big(\mathbb{E}\exp\Big(\frac{4\bar p\mu}{\Gamma_n}|D_{2,b,n}^\vartheta|\Big)\Big)^{\frac{1}{4\bar p}} \Big(\mathbb{E}\exp\Big(\frac{4\bar p\mu}{\Gamma_n}|D_{2,\Sigma,n}^\vartheta|\Big)\Big)^{\frac{1}{4\bar p}}, \tag{4.3}$$

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 there exists $C_\vartheta>0$ such that for all $x\in\mathbb{R}^d$, $|D^2\vartheta(x)| \le C_\vartheta/(1+|x|)$. Plugging in (4.3) the controls established in Lemma 4 (with $j=2$), Lemma 5 (equations (2.17) and (2.18)) and (3.5), then replacing λ_n by μ, we get similarly to the first inequality of (2.21) and with the notations of Lemma 3:

$$\begin{aligned}
\mathscr{T}_2 &\le \exp\Big(\frac{\bar q\mu^2\|\sigma\|_\infty^2[\vartheta]_1^2}{2\Gamma_n}\Big)\exp\Big(\frac{\mu^2}{2\Gamma_n\bar p}+\frac{\bar p a_n^2}{2}\Big)\exp\Big(\frac{3\bar p C_{V,\vartheta}^2\mu^2}{c_V\Gamma_n^2}+\frac{c_V}{\bar p}\Big)(I_V^1)^{\frac{1}{2\bar p}}\\
&\quad\times\exp\Big(C_{2.17}\frac{\bar p\mu^2(\Gamma_n^{(2)})^2}{\Gamma_n^2}\Big)(I_V^1)^{\frac{1}{4\bar p}}\times\exp\Big(C_{2.18}\Big(\frac{3\bar p\mu^2(\Gamma_n^{(2)})^2}{2\Gamma_n^2}+\frac{1}{2\bar p}\Big)\Big)(I_V^1)^{\frac{1}{4\bar p}}.\\
&\le \exp\Big(\frac{\mu^2}{\Gamma_n}\Big(\frac{\bar q\|\sigma\|_\infty^2[\vartheta]_1^2}{2}+\bar p\Big(\frac{(\Gamma_n^{(2)})^2}{\Gamma_n}[C_{2.17}+\frac{3}{2}C_{2.18}]+\frac{3C_{V,\vartheta}^2}{c_V\Gamma_n}\Big)+\frac{1}{2\bar p}\Big)\Big)\exp\Big(\frac{1}{\bar p}\big(c_V+\frac{C_{2.18}}{2}\big)+\frac{\bar p a_n^2}{2}\Big)(I_V^1)^{\frac{1}{\bar p}}.
\end{aligned}$$

Set now

$$\begin{aligned}
\bar C_n &:= \exp\Big(\frac{1}{\bar p}\big(c_V+\frac{C_{2.18}}{2}\big)+\frac{\bar p a_n^2}{2}\Big)(I_V^1)^{\frac{1}{\bar p}},\\
\bar e_n &:= \bar p\Big(\frac{(\Gamma_n^{(2)})^2}{\Gamma_n}[C_{2.17}+\frac{3}{2}C_{2.18}]+\frac{3C_{V,\vartheta}^2}{c_V\Gamma_n}\Big)+\frac{1}{2\bar p}.
\end{aligned} \tag{4.4}$$

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

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

$$\Big(\mathbb{E}\exp\Big(-\frac{\lambda q}{\Gamma_n}M_n\Big)\Big)^{\frac{1}{q}} \le \Big(\mathscr{T}_1^{\frac{1}{\rho}}\mathscr{T}_2^{1-\frac{1}{\rho}}\Big)^{\frac{1}{q}} \le \exp\Big(\frac{\rho q\lambda^2[\varphi]_1^2\nu(\|\sigma\|^2)}{2\Gamma_n}\Big)\bar C_n^{\frac{\rho-1}{\rho q}}\exp\Big(\frac{\rho-1}{\rho q}\frac{\mu^2}{\Gamma_n}\Big(\frac{\bar q\|\sigma\|_\infty^2[\vartheta]_1^2}{2}+\bar e_n\Big)\Big).$$

Plugging this bound in (2.12), using again the controls of Lemmas 4 and 5, eventually yields:

$$\begin{aligned}
\mathbb{P}\Big[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)\ge a\Big] &\le \exp\Big(-\frac{a\lambda}{\sqrt{\Gamma_n}}\Big)\exp\Big(\frac{\lambda^2}{2\Gamma_n}\big(\rho q[\varphi]_1^2\nu(\|\sigma\|^2)+\frac{1}{p}\big)\Big)\bar C_n^{\frac{\rho-1}{\rho q}}\exp\Big(\frac{\rho-1}{\rho q}\frac{\mu^2}{\Gamma_n}\Big(\frac{\bar q\|\sigma\|_\infty^2[\vartheta]_1^2}{2}+\bar e_n\Big)\Big)\\
&\quad\times\exp\Big(\frac{\lambda^2}{\Gamma_n}p\Big(\frac{(\Gamma_n^{(2)})^2}{\Gamma_n}\big(C_{2.17}+\frac{3}{2}C_{2.18}\big)+\frac{3C_{V,\varphi}^2}{c_V\Gamma_n}\Big)\Big)\exp\Big(\frac{1}{p}\big(c_V+\frac{C_{2.18}}{2}\big)+\frac{pa_n^2}{2}\Big)(I_V^1)^{\frac{1}{p}}.
\end{aligned}$$

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

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

where e_n is defined similarly to $\bar e_n$ in (4.4) replacing $\bar p$ by p. In particular $e_n\underset{n}{\to}0$. Note that for the previous choices of $p,\bar p$, we have that $\widetilde C_n := C_n\bar C_n^{\frac{\rho-1}{\rho q}}\underset{n}{\to}1$ 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

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

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

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

Note that both sequences $(\widetilde{A}_n)_{n \geq 1}$ and $(\widetilde{B}_n)_{n \geq 1}$ 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 $\lambda_{\min}$, unique zero of the equation $P'(\lambda_{\min}) = 0$. This equation reads

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

which is the canonical form of this third degree equation to apply the Cardan-Tartaglia formula ([2]) so that

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

In order to derive our non-asymptotic bound, we select two "regimes" based on a first order expansion of $\lambda_{\min}$ in two cases $\frac{a}{B_n \sqrt{\Gamma_n}} \to 0$ and $\frac{B_n \sqrt{\Gamma_n}}{a} \to 0$, 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 $\rho \to 1$, then $\frac{1}{B_n}$ and $\frac{A_n}{B_n} \to 0$. First, one easily checks that if $(x_n)_{\geq 1}$ and $(a_n)_{n \geq 1}$ are two sequences of positive real numbers where $(a_n)_{n \geq 1}$ is bounded, then

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

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

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

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

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

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

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

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

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

---

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

Note 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 $x_n = o\big(a_n^{\frac{3}{2}}\big)$ in (4.8) is satisfied *a posteriori*.

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

$$\lambda_{\min}(\rho) \sim \bar\lambda^*(\rho) = \frac{\Gamma_n}{2}\Big(\frac{2a}{B_n\sqrt{\Gamma_n}}\Big)^{\frac13} = \left(\frac{a\Gamma_n}{4B_n}\right)^{\frac13}\sqrt{\Gamma_n} \quad \text{as } n \to +\infty.$$

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,

$$P\big(\bar\lambda^*(\rho)\big) = -a^{\frac43}\Gamma_n^{\frac13}\frac{(\rho-1)^{\frac13}}{\rho(4\widetilde B_n)^{\frac13}}\left(\frac34 - \frac{\widetilde A_n}{(4\widetilde B_n)^{\frac13}}\frac{\Gamma_n^{\frac13}}{a^{\frac23}}(\rho-1)^{\frac13}\right).$$

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:

$$P\big(\bar\lambda^*(\rho)\big) \le -\frac{a^{\frac43}}{4}\left(\frac{\Gamma_n}{\widetilde B_n}\right)^{\frac13}\left(1 - \frac23\frac{\widetilde A_n}{\widetilde B_n}\Big(\frac{\Gamma_n}{a^2}\Big)^{\frac13}\right). \tag{4.10}$$

From (4.10), (4.9) and (4.5), we conclude the proof of case *(a)* by setting $\bar c_n := \widetilde A_n \widetilde B_n^{-\frac13}$ 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 $A_n(\rho) = \frac{\rho q[\varphi]_1^2\nu(\|\sigma\|^2)}{2}$. □

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*

$$\begin{aligned} P\big(\lambda_{\min}(\rho)\big) &= -\frac{\lambda_{\min}(\rho)}{2\sqrt{\Gamma_n}}\left(\frac{3a}{2} - \frac{\lambda_{\min}(\rho)\rho\widetilde A_n}{\sqrt{\Gamma_n}}\right) \\ &= -\frac{\sqrt{\Gamma_n}}{4}\frac{(\rho-1)^{\frac13}}{\rho}\Phi_n(a,\rho)\left(\frac{3a}{2} - \frac{\sqrt{\Gamma_n}}{2}(\rho-1)^{\frac13}\widetilde A_n\Phi_n(a,\rho)\right), \end{aligned}$$

where $\Phi_n(a,\rho) = \left(\frac{a}{\sqrt{\Gamma_n}\widetilde B_n} + \left((\rho-1)\Big(\frac{2\widetilde A_n}{3\widetilde B_n}\Big)^3 + \frac{a^2}{\widetilde B_n^2\Gamma_n}\right)^{\frac12}\right)^{\frac13} + \left(\frac{a}{\sqrt{\Gamma_n}\widetilde B_n} - \left((\rho-1)\Big(\frac{2\widetilde A_n}{3\widetilde B_n}\Big)^3 + \frac{a^2}{\widetilde B_n^2\Gamma_n}\right)^{\frac12}\right)^{\frac13}.$

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. Practical Deviation Bounds and Related Smoothness Issues for the Poisson Problems

We begin giving some practical concentration results for the quantity $\sqrt{\Gamma_n}\big(\nu_n(f) - \nu(f)\big)$, where the source f is a given input, and related controls on the associated confidence intervals. For these controls to hold, some smoothness conditions for the Poisson problem $\mathcal{A}\varphi = f - \nu(f)$ and/or $\mathcal{A}\vartheta = \|\sigma\|^2 - \nu(\|\sigma\|^2)$ are needed. These results are almost direct corollaries of Theorems 2 and 3 above. We then discuss how to obtain the required smoothness in two cases:

- when some non-degeneracy on the coefficients holds (uniform ellipticity).
- when some confluence conditions are available.

We eventually discuss, in the non-degenerate case, how a smoothing procedure can be developed to get closer to the assumptions of Lipschitz source which is usually made when using functional inequalities.

5.1. **A first Non-asymptotic confidence interval result.**

Theorem 4 (Non-asymptotic confidence intervals without bias)**.** *Let f be in $\mathcal{C}^{1,\beta}(\mathbb{R}^d,\mathbb{R})$ for some $\beta\in(0,1)$. Assume that the Poisson equation:*

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

admits a unique solution $\varphi\in\mathcal{C}^{3,\beta}(\mathbb{R}^d,\mathbb{R})$. Suppose as well that there exists a positive constant C_φ such that, for all $x\in\mathbb{R}^d$, $|D^2\varphi(x)|\le C_\varphi/(1+|x|)$ and the following gradient bound holds:

$$\|\nabla\varphi\|_\infty\le\frac{[f]_1}{\alpha},\ \alpha>0. \tag{5.1}$$

We suppose finally that φ satisfies (G_V) introduced in Theorem 2.

Assume the step sequence $(\gamma_k)_{k\ge1}$ is such that $\gamma_k\asymp k^{-\theta}$, $\theta\in(\frac{1}{2+\beta},1]$. Then, for $(c_n)_{n\ge1},(C_n)_{n\ge1}$ like in Theorem 2 with $\lim_n c_n=\lim_n C_n=1$, we have that for all $n\ge1$ and $a>0$:

$$\mathbb{P}\big[\sqrt{\Gamma_n}|\nu_n(f)-\nu(f)|>a\big]\le 2\,C_n\exp\Big(-c_n\frac{a^2\alpha^2}{2\|\sigma\|_\infty^2[f]_1^2}\Big), \tag{5.2}$$

$$\mathbb{P}\left[\nu(f)\in\Big[\nu_n(f)-\frac{a\|\sigma\|_\infty[f]_1}{\alpha\sqrt{\Gamma_n}},\nu_n(f)+\frac{a\|\sigma\|_\infty[f]_1}{\alpha\sqrt{\Gamma_n}}\Big]\right]\ge1-2C_n\exp\left(-c_n\frac{a^2}{2}\right). \tag{5.3}$$

Proof. Equation (5.2) is a direct consequence of Theorem 2 and the gradient bound (5.1) observing that $\nu_n(f)-\nu(f)=\nu_n(\mathcal{A}\varphi)$. To prove (5.3), setting $a_{\sigma,f,\alpha}:=a\|\sigma\|_\infty\frac{[f]_1}{\alpha}$, it suffices to write:

$$\mathbb{P}\big[\sqrt{\Gamma_n}\big|\nu_n(f)-\nu(f)\big|\ge a_{\sigma,f,\alpha}\big] \;=\; \mathbb{P}\big[\sqrt{\Gamma_n}\big|\nu_n(f-\nu(f))\big|\ge a_{\sigma,f,\alpha}\big]=\mathbb{P}\big[\sqrt{\Gamma_n}\big|\nu_n(\mathcal{A}\varphi)\big|\ge a_{\sigma,f,\alpha}\big].$$

Hence:

$$\mathbb{P}\left[\nu(f)\in\Big[\nu_n(f)-\frac{a_{\sigma,f,\alpha}}{\sqrt{\Gamma_n}},\nu_n(f)+\frac{a_{\sigma,f,\alpha}}{\sqrt{\Gamma_n}}\Big]\right] \;=\; 1-\mathbb{P}[\sqrt{\Gamma_n}\big|\nu_n(\mathcal{A}\varphi)\big|\ge a_{\sigma,f,\alpha}].$$

Conclude by noting that Theorem 2 (a) yields:

$$\mathbb{P}[\sqrt{\Gamma_n}\big|\nu_n(\mathcal{A}\varphi)\big|\ge a_{\sigma,f,\alpha}] \;\le\; 2C_n\exp\left(-c_n\frac{a^2}{2}\right). \qquad\square$$

5.2. **A more refined non-asymptotic confidence interval when $\|\sigma\|^2$ is a coboundary.** We provide here a kind of *Slutsky's Lemma* when $\|\sigma\|^2$ is a coboundary. The key idea is to apply Theorem 3 discarding a better concentration rate and to replace the term $\nu(\|\sigma\|^2)$ appearing there by its ergodic approximation $\nu_n(\|\sigma\|^2)$.

Theorem 5 (Slutsky type concentration result for the coboundary case)**.** *Under the assumptions of Theorem 4, for $\beta\in(0,1]$ and $\theta\in(\frac{1}{2+\beta},1]$, assuming as well that there is a unique solution ϑ to $\mathcal{A}\vartheta=\|\sigma\|^2-\nu(\|\sigma\|^2)$ satisfying the same assumptions as φ in Theorem 4, there exist two explicit monotonic sequences $c_n\le1\le C_n$, $n\ge1$, with $\lim_n C_n=\lim_n c_n=1$ such that for all $n\ge1$, for all $a>0$, the following bounds hold: if $\frac{a}{\sqrt{\Gamma_n}}\to0$ (Gaussian deviations) then,*

$$\mathbb{P}\big[|\sqrt{\Gamma_n}\frac{\nu_n(f)-\nu(f)}{\sqrt{\nu_n(\|\sigma\|^2)}}|\ge a\big]\le 2\,C_n\exp\left(-c_n\frac{a^2\alpha^2}{2[f]_1^2}\right), \tag{5.4}$$

$$\mathbb{P}\left[\nu(f)\in\Big[\nu_n(f)-\frac{a\sqrt{\nu_n(\|\sigma\|^2)}[f]_1}{\alpha\sqrt{\Gamma_n}},\nu_n(f)+\frac{a\sqrt{\nu_n(\|\sigma\|^2)}[f]_1}{\alpha\sqrt{\Gamma_n}}\Big]\right]\ge1-2C_n\exp\left(-c_n\frac{a^2}{2}\right). \tag{5.5}$$

Again, the non-asymptotic confidence interval is explicitly computable. It is also sharper than the one in (5.3).

Proof. First we write:

$$\mathbb{P}\left[\sqrt{\Gamma_n}\frac{\nu_n(f)-\nu(f)}{\sqrt{\nu_n\left(\|\sigma\|^2\right)}}\ge a\right]=\mathbb{P}\left[\nu_n(\mathcal{A}\varphi)\ge\frac{a}{\sqrt{\Gamma_n}}\sqrt{\nu_n\left(\|\sigma\|^2\right)}\right]. \tag{5.6}$$

We then proceed similarly to Theorem 3, with an exponential Bienaymé-Tchebychev inequality, for all $\lambda > 0$ we have:

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

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

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

The proof of Theorem 3 yields:

$$\begin{aligned}
\mathbb{P}\left[\sqrt{\Gamma_n}\frac{\nu_n(f)-\nu(f)}{\sqrt{\nu_n\left(\|\sigma\|^2\right)}} \geq a\right] \ &\leq \ \mathscr{R}_n\exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\sqrt{\nu\left(\|\sigma\|^2\right)}\right)\exp\left(\frac{\rho\widetilde{q}\lambda^2}{\Gamma_n}\widetilde{A}_n+\frac{\rho^3\widetilde{q}^3\lambda^4}{(\rho-1)\Gamma_n}\widetilde{B}_n\right) \\
&\quad\times\left(\mathbb{E}\exp\left(-\frac{a\widetilde{p}\lambda}{\sqrt{\Gamma_n}}\frac{\nu_n\left(\mathcal{A}\vartheta\right)}{\sqrt{\nu_n\left(\|\sigma\|^2\right)}+\sqrt{\nu\left(\|\sigma\|^2\right)}}\right)\right)^{1/\widetilde{p}}, \qquad (5.7)
\end{aligned}$$

where we recall from identity (4.5):

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

Also, $\mathscr{R}_n \underset{n}{\to} 1$ denotes a "generic" remainder. Observe that thanks to (5.1), we get:

$$\widetilde{A}_n \leq \frac{q[f]_1^2\nu(\|\sigma\|^2)}{2\alpha}+e_n. \qquad (5.8)$$

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

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

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:

$$\mathbb{E}\left[\exp\left(-\frac{a\widetilde{p}\lambda}{\sqrt{\Gamma_n}}\frac{\nu_n\left(\mathcal{A}\vartheta\right)}{\sqrt{\nu_n\left(\|\sigma\|^2\right)}+\sqrt{\nu\left(\|\sigma\|^2\right)}}\right)\right]^{1/\widetilde{p}}$$

$$\leq \mathscr{R}_n \exp\left(\frac{a^2\widetilde{p}^2}{\Gamma_n\nu(\|\sigma\|^2)}\frac{\lambda^2}{\Gamma_n}\Big(\frac{\|\sigma\|_\infty^2\|\nabla\vartheta\|_\infty^2}{2}+e_n\Big)\right)\Big[\mathbb{P}\big[\nu_n(\mathcal{A}\vartheta)\geq 0\big]^{1/\widetilde{q}}+\mathbb{P}\big[\nu_n(\mathcal{A}\vartheta)<0\big]^{1/\widetilde{q}}\Big]^{1/\widetilde{p}}. \tag{5.9}$$

We choose $\widetilde{p}:=\widetilde{p}(n)\to+\infty$, such that $\frac{\widetilde{p}^2a^2}{\Gamma_n}\to 0$, and so $(\mathbb{P}\big[\nu_n(\mathcal{A}\vartheta)\geq 0\big]^{1/\widetilde{q}}+\mathbb{P}[\nu_n(\mathcal{A}\vartheta)<0]^{1/\widetilde{q}})^{1/\widetilde{p}}\leq 2^{1/\widetilde{p}}\to 1$. Moreover, exploiting again that for the *Gaussian regime*, $\frac{\widetilde{p}^2a^2}{\Gamma_n}\to 0$, we obtain by (5.9) and (5.7):

$$\mathbb{P}\left[\sqrt{\Gamma_n}\frac{\nu_n(f)-\nu(f)}{\sqrt{\nu_n\left(\|\sigma\|^2\right)}}\geq a\right]\leq \mathscr{R}_n\exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\sqrt{\nu\left(\|\sigma\|^2\right)}\right)\exp\left(\frac{\rho\widetilde{q}\lambda^2}{\Gamma_n}(\widetilde{A}_n+e_n)+\frac{\rho^3\widetilde{q}^3\lambda^4}{(\rho-1)\Gamma_n}\widetilde{B}_n\right). \tag{5.10}$$

From identity (5.10), the optimization over λ is similar to the one performed in the proof of Theorem 3. This yields the deviation bound (5.4). The non-asymptotic confidence interval in (5.5) is derived as for Theorem 4 from (5.1) and (5.4). □

5.3. **Regularity Issues for the Poisson Problems.** We mention carefully that, for our deviation analysis to work, we need to establish some pointwise controls on the solution of the previously mentioned Poisson equations. Namely, we need to have quantitative bounds on its derivatives and the associated Hölder continuity modulus up to order 3. In the non-degenerate case, the solvability is usually studied in a Sobolev setting, see *e.g.* [PV01]. To obtain the required smoothness, we use here in the non-degenerate framework some Schauder estimates, deriving from the work of Krylov and Priola [KP10], which allow to benefit from the parabolic 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 $f\in C^{1,\beta}(\mathbb{R}^d,\mathbb{R})$. The other typical assumption that will yield the required smoothness on φ is a confluence condition provided that the source f and the coefficients b,σ are smooth enough. In this case, the derivatives can be expressed using iterated tangent processes and we cannot hope, 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 sufficiently smooth coefficients b,σ.

We now introduce various types of assumptions that we are going to consider for the remaining of the section:

- **Non-degeneracy Conditions**. By that, we mean that we assume:

(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

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

- **Smoothness of the coefficients and the source**. Concerning the smoothness of the coefficients b,σ and the source f, we introduce for $k\in\{1,3\}$ and $\beta\in(0,1)$ the following condition.

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

- **Confluence Conditions**.

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

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

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.

In the following, we say that **(ND)** (non-degeneracy) holds when, additionally to **(A)**, **(UE)** and **($\mathbf{R}_{1,\beta}$)** are in force. Similarly, we say that **(C)** holds when **(A)**, **($\mathbf{R}_{3,\beta}$)** and **(D)** are satisfied. Under **(ND)** or **(C)**, it is well known that the Poisson equation:

$$\forall x\in\mathbb{R}^d,\ \mathcal{A}\varphi(x)=f(x)-\nu(f), \tag{5.11}$$

admits a unique solution, in $H^1(\mathbb{R}^d,\mathbb{R})$ under **(ND)** (see [PV01]), or in $\mathcal{C}^{3,\beta}(\mathbb{R}^d,\mathbb{R})$ under **(C)** (see [PP12]). In both cases, we have the following representation:

$$\varphi(x) = -\int_{\mathbb{R}_+} \big(P_t f(x) - \nu(f)\big)dt \quad \text{where} \quad P_t f(x) := \mathbb{E}\,[f(Y_t^{0,x})] \tag{5.12}$$

and $Y_t^{0,x}$ solves (1.1) with $Y_0^{0,x} = x$. To comply with the framework of the above Theorems 4 and 5, the first step is to establish a pointwise gradient control.

5.3.1. *Gradient Control.* When the coefficients b, σ satisfy at least $(\mathbf{R}_{1,\beta})$ in (1.1), there are two kinds of assumptions leading to explicit pointwise gradient controls. The first one is confluence condition **(C)** described above that gives the expected bound, through controls on the tangent flow, without any *a priori* uniform ellipticity condition of type **(UE)**. On the other hand, in the non-degenerate framework, that is under **(ND)**, when the so-called Bakry and Émery curvature criterion holds, see [BE85, BGL14], the gradient and semi-group commute up to an exponential multiplicative factor (see equation (5.13) below). It then readily gives the result.

▷ *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 $\mathcal{D}(\mathcal{A})$

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

We also need to define the Γ_2 operator

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

In our Brownian diffusion setting, we have

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

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

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

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

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

Observe that for $\Sigma = I_d$ the condition $\mathbf{(BE}_\rho\mathbf{)}$ is actually equivalent to **(D)** with $\alpha = \rho$ 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 [ACJ08]. We also refer to [AMTU01] if the diffusion matrix is scalar diagonal, *i.e.* $\Sigma(x) = \varsigma^2(x) I_d$, $x \in \mathbb{R}^d$, where ς is real valued.

An important property when $\mathbf{(BE}_\rho\mathbf{)}$ holds, see again [BE85], [BGL14], is that the following commutation inequality holds:

$$\forall t \geq 0,\ \forall x \in \mathbb{R}^d,\ \Gamma(P_t f)(x) \leq \exp(-2\rho t) P_t \Gamma(f). \tag{5.13}$$

Lemma 6 (Pointwise Gradient Bounds)**.** *Assume that* **(C)** *or* **(ND)** *holds. Under* **(ND)**, *assume furthermore that either* **(D)** *or* $\mathbf{(BE}_\rho\mathbf{)}$ *(for some ρ) is in force. Then*

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

with α as in **(D)** *when this condition holds, and where, under* **(ND)** *and* $\mathbf{(BE}_\rho\mathbf{)}$, $\alpha = \rho\big(\frac{\kappa}{\underline{\sigma}}\big)^{-1/2}$ *for $d \geq 2$ and $\alpha = \rho$ for $d = 1$.*

Proof. Uniformly elliptic framework. Let us start with the proof under **(ND)** and **(BE$_\rho$)**. In that case, from the previous expression of Γ and (5.13):

$$\forall t \geq 0,\ \forall x \in \mathbb{R}^d, |\sigma^*(x)\nabla_x \mathbb{E}\left[f(Y_t^{0,x})\right]| \leq \exp(-\rho t)\Big(\mathbb{E}\left[|\sigma^*\nabla f(Y_t^{0,x})|^2\right]\Big)^{1/2}.$$

Using the ellipticity and the boundedness of $\sigma\sigma^*$ (see assumptions **(UE)** p. 26 and **(C2)** p. 3) yields:

$$\forall t \geq 0,\ \forall x \in \mathbb{R}^d, \|\nabla_x \mathbb{E}\left[f(Y_t^{0,x})\right]\|_\infty \leq \exp(-\rho t)\Big(\frac{\kappa}{\underline{\sigma}}\Big)^{1/2}\|\nabla f\|_\infty.$$

The representation formula (5.12) then yields the announced result with $\alpha = \left(\frac{\kappa}{\underline{\sigma}}\right)^{-1/2}\rho$.

Confluent framework. Assume now that **(D)** holds. Observe that, under **(A)**, whenever **(ND)** or **(C)** holds, it is well known that that $\nabla_x Y_t^{0,x}$ is well defined and belongs to $L^2(\mathbb{P})$, see [IW80]. Hence, for $t > 0,\ i \in [\![1,d]\!]$:

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

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 **(D)** 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:

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

where $(M_t)_{t\geq 0} := \big(p\sum_{j=1}^d \int_0^t \langle \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|}, D\sigma_{\cdot j}(Y_s^{0,x})\frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|}\rangle dW_s^j\big)_{t\geq 0}$ 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 **(D)**, we thus get:

$$|\partial_{x_i} Y_t^{0,x}|^p \leq \exp(-\alpha p t) \times \mathcal{E}(M_t).$$

We eventually derive:

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

From the above control and equation (5.12), we thus derive:

$$\forall i \in [\![1,d]\!],\ \forall x \in \mathbb{R}^d,\ |\partial_{x_i}\varphi(x)| \leq \frac{[f]_1}{\alpha}. \tag{5.14}$$

□

5.3.2. *Additional Smoothness.* We have the following regularity result.

Theorem 6. *Assume one of the two assumptions below hold:*

– **(ND)** *and either* **(D)** *or* **(BE$_\rho$)** *for some $\rho > 0$ are in force. Suppose as well that, for every $(i,j) \in [\![1,d]\!]^2$, $\Sigma_{i,j}(x) = \Sigma_{i,j}(x_{i\wedge j}, \cdots, x_d)$.*

– **(C)** *is in force (without additional structure condition on σ).*

Then, the Poisson Equation (5.11) *admits a unique solution $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$, $\beta \in (0,1)$. Furthermore, there exists $C_\varphi > 0$ such that for all $x \in \mathbb{R}^d$, $|D^2\varphi(x)| \leq \frac{C_\varphi}{1+|x|}$.*

Proof. Under **(C)**, the statement can be derived 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 $\mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$.

We now give the proof under **(ND)** which requires more sophisticated tools (Schauder estimates for operators with unbounded coefficients). Let us begin with the scalar case. For $d = 1$, set for all $x \in \mathbb{R}$,

$$v(x) := -\int_0^{+\infty} dt\, \mathbb{E}[\Psi(Y_t^{0,x})\partial_x Y_t^{0,x}] = -\int_0^{+\infty} dt\, \mathbb{E}\Big[\Psi(Y_t^{0,x}) \exp\Big(\int_0^t b'(Y_s^{0,x})ds\Big)\mathcal{E}\Big(\int_0^t \sigma'(Y_s^{0,x})dW_s\Big)\Big] \tag{5.15}$$

where for all $y \in \mathbb{R}$, $\Psi(y) := \partial_y f(y)$. We observe that $\partial_x\varphi(x) = v(x)$. Also, from our assumptions on f, b, σ, we have that $\Psi, b', \sigma' \in \mathcal{C}^\beta(\mathbb{R}^d, \mathbb{R})$. Theorems 2.4-2.6 in Krylov and Priola, [KP10] then yield the existence of a unique solution to the PDE:

$$\widetilde{\mathcal{A}}w(x) + b'(x)w(x) = \Psi(x), \quad \text{where} \quad \widetilde{\mathcal{A}}w(x) = \mathcal{A}w(x) + \sigma\sigma'(x)w'(x), \tag{5.16}$$

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

$$\exists\, C \geq 1,\ \|w\|_{2,\beta} \leq C(1 + \|\Psi\|_\beta). \tag{5.17}$$

Indeed, from **(D)** or **(BE$_\rho$)**, we get that $b'(x) \leq -\alpha < 0$ and the potential in (5.16) has the good sign. From (5.15) and the Girsanov theorem, we also get:

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

where $d\widetilde{Y}_s^{0,x} = \Big(b(\widetilde{Y}_s^{0,x}) + \sigma\sigma'(\widetilde{Y}_s^{0,x})\Big)ds + \sigma(\widetilde{Y}_s^{0,x})dW_s$. Note that $\widetilde{Y}$ has generator $\widetilde{\mathcal{A}}$. A simple identification procedure, similar to the proof of Theorem II.1.1 in Bass [Bas97] then gives $v = w$. The result follows from (5.17). 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 $i \in [\![1,d]\!]$, $j \geq i$, $\Sigma_{i,j}(x) = \Sigma_{i,j}(x_i, \cdots, x_d)$, we have that differentiating formally the PDE (5.11) in the space variable x_i, $i \in [\![1,d]\!]$ yields that $\partial_{x_i}\varphi = v_i$ should satisfy:

$$\begin{aligned}\widetilde{\mathcal{A}}\, w_i(x) + \partial_{x_i} b_i(x) w_i(x) &= \Psi_i(x) - \sum_{j \in [\![1,d]\!]\setminus\{i\}} \partial_{x_i} b_j(x) v_j(x) \\ &\quad -\frac{1}{2}\sum_{1\leq j<i} \partial_{x_i}\Sigma_{j,j}(x)\partial_{x_j} v_j(x) - \sum_{\substack{1 \leq j < i,\\ k > j,\, k \neq i}} \partial_{x_i}\Sigma_{j,k}(x)\partial_{x_j} v_k(x),\end{aligned}$$

$$\text{with} \quad \Psi_i(x) = \partial_{x_i} f(x) \quad \text{and} \quad \widetilde{\mathcal{A}} w_i(x) = \mathcal{A}\, w_i(x) + \frac{1}{2}\partial_{x_i}\Sigma_{i,i}(x)\partial_{x_i} w_i(x) + \sum_{j\neq i}^d \partial_{x_i}\Sigma_{i,j}(x)\partial_{x_j} w_i(x). \tag{5.18}$$

From our current assumptions on f, b and the previous computations on the gradient for the multi-dimensional case, we easily get that $\widetilde{\Psi}_1(x) := \Psi_1(x) - \sum_{j\neq 1}\partial_{x_1} b_j(x) v_j(x) \in \mathcal{C}^\beta(\mathbb{R}^d, \mathbb{R})$. Indeed, it is easily checked from the previous computations that $\nabla\varphi$ is β-Hölder continuous. Hence, Theorems 2.4-2.6 in Krylov and Priola, [KP10] still apply and give that there exists a unique solution $w_1 \in \mathcal{C}^{2,\beta}(\mathbb{R}^d, \mathbb{R})$ to (5.18) which also satisfies:

$$\exists\, C \geq 1,\ \|w_1\|_{2,\beta} := \sum_{\alpha, |\alpha| \in [\![0,2]\!]} \|D^\alpha w_1\|_\infty + [D^{(2)} w_1]_\beta \leq C(1 + \|\widetilde{\Psi}_1\|_\beta) =: \bar{C}(f, (\mathbf{A}), \beta). \tag{5.19}$$

The identification $w_1 = \partial_{x_1}\varphi = v_1$ is standard. The control (5.19) allows to iterate, since it gives that $\nabla w_1 = (\partial_{x_1} v_1, \cdots, \partial_{x_d} v_1) = (\partial_{x_1,x_1}\varphi_1, \cdots, \partial_{x_d,x_1}\varphi)$ 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, [KP10], that for all $i \in [\![1,d]\!]$ there exists a unique solution $w_i \in \mathcal{C}^\beta(\mathbb{R}^d,\mathbb{R})$ to (5.18) such that:

$$\exists\, C \geq 1,\ \|w_i\|_{2,\beta} \leq C(1+\|\widetilde{\Psi}_i\|_\beta) =: \bar{C}(f,(\mathbf{A}),\beta),$$
$$\widetilde{\Psi}_i(x) := \Psi_i(x) - \sum_{j\in[\![1,d]\!]\setminus\{i\}} \partial_{x_i} b_j(x) v_j(x) - \sum_{\substack{1\leq j<i,\\ k\in[\![1,d]\!]\setminus\{i\}}} \partial_{x_i}\Sigma_{j,k}(x)\partial_{x_j} v_k(x). \tag{5.20}$$

The condition on $|D^2\varphi(x)|$ follows from the Remark p.7 in [KP10]. Let us anyhow provide a direct proof. Recalling that $\partial_{x_i}\varphi = w_i$, observe from (5.18) that for all $x \in \mathbb{R}^d$:

$$\big\langle b(x), \partial_{x_i}\nabla\varphi(x)\big\rangle = \widetilde{\Psi}_i(x) - \frac{1}{2}\mathrm{Tr}\big(\Sigma(x)D^2(\partial_{x_i}\varphi(x))\big). \tag{5.21}$$

Now, from (5.14) and (5.20), φ is such that $\nabla\varphi$, $D^2\varphi$, and $D^3\varphi$ are bounded. This is also the case for $\widetilde{\Psi}_i$. Equation (5.21) then gives that $|b(x)\cdot\partial_{x_k}\nabla\varphi(x)|$ is bounded. The Lyapunov condition on b, assumption $(\mathcal{L}_{\mathbf{V}})$, iii), gives the stated decay.

□

Remark 12 (Practical Application)**.** *Under the assumptions of Theorem 6 for the coefficients b,σ and the source f, the required smoothness for φ needed in Theorem 4 holds. The non-asymptotic confidence interval follows. Observe that, in this case, we again have a quantitative control on $\|\nabla\varphi\|_\infty \leq \frac{[f]_1}{\alpha}$ from Lemma 6. In this framework, given the model* (1.1) *and its coefficients b,σ, a fixed source f, the non-asymptotic confidence interval* (5.3) *is explicitly computable. Observe as well that, since we also assumed $\sigma \in \mathcal{C}_b^{1,\beta}(\mathbb{R}^d,\mathbb{R}^d\otimes\mathbb{R}^d)$, then since σ is bounded, $\|\sigma\|^2 \in \mathcal{C}_b^{1,\beta}(\mathbb{R}^d,\mathbb{R})$. Therefore, still from Theorem 6, the required smoothness for ϑ needed in Theorem 5 holds. Hence, the improved bound* (5.5) *applies.*

Let us also indicate that, under **(UE)** *and either* **(D)** *or* **($\mathbf{BE}_\rho$)** *for some $\rho > 0$, if f is simply Lipschitz continuous and b,σ as in* **($\mathbf{R}_{1,\beta}$)**, *we derive below in Theorem 7, following the proof of Theorem 4, that* (5.3) *still holds for $\gamma_k \asymp k^{-\theta}$, $k \in (1/2,1]$. The non-asymptotic confidence interval in* (5.3) *is again explicitly computable.*

Remark 13 (Structure of σ)**.** *We emphasize that the structure condition on σ assumed in Theorem 6 is mainly technical. It is of course always verified in dimension $d = 1$. For $d > 1$ it is motivated by the fact that, differentiating* (5.11) *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* [Boc13]*), 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 [Pri09] *could also be a starting point to investigate the smoothness of the Poisson problem for degenerate kinetic models.*

These aspects will concern further research.

5.4. **Regularization of Lipschitz Sources.** We assume here that assumptions **(A)**, **(UE)** are in force. We suppose as well that the following smoothness holds for b,σ:

($\mathbf{R}_{b,\sigma}$) Regularity and Structure. We assume that there exists $\beta \in (0,1)$ such that b,σ in (1.1) belong to $\mathcal{C}^{1,\beta}(\mathbb{R}^d,\mathbb{R}^d)$ and $\mathcal{C}_b^{1,\beta}(\mathbb{R}^d,\mathbb{R}^d\otimes\mathbb{R}^d)$ respectively. Also, for all $(i,j)\in[\![1,d]\!]^2$, $\Sigma_{i,j}(x) = \Sigma_{i,j}(x_{i\wedge j},\cdots,x_d)$.

Importantly, we are interested, under assumptions **(A)**, **(UE)**, **($\mathbf{R}_{b,\sigma}$)**, in giving controls for the estimation of $\nu(f)$ 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 [MT06], [Boi11].

To comply with our previous framework, namely to exploit the smoothness result of Theorem 6, 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:

$$\exists\, C_\eta > 0,\ \forall x\in\mathbb{R}^d,\qquad |f_\delta(x) - f(x)| \leq C_\eta\delta[f]_1,$$
$$\forall\beta\in(0,1),\qquad [\nabla f_\delta]_\beta \leq C_\eta[f]_1\delta^{-\beta}. \tag{5.22}$$

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:

$$\begin{aligned}\mathbb{P}[\sqrt{\Gamma_n}(\nu_n(f)-\nu(f)) \geq a] &= \mathbb{P}\Big[(\nu_n(f_\delta)-\nu(f_\delta)) + R_{n,\delta}(f) \geq \frac{a}{\sqrt{\Gamma_n}}\Big], \\ R_{n,\delta}(f) &:= [(\nu_n(f)-\nu(f)) - (\nu_n(f_\delta)-\nu(f_\delta))]. \end{aligned} \tag{5.23}$$

From (5.22), one readily gets:

$$|R_{n,\delta}(f)| \leq 2C_\eta \delta [f]_1. \tag{5.24}$$

On the other hand, the coefficients b, σ and the source f_δ satisfy assumption $\mathbf{(R}_{1,\beta}\mathbf{)}$ (observe indeed that the mollified function $f_\delta \in C^{1,\beta}(\mathbb{R}^d, \mathbb{R})$). Hence, Theorem 6 yields that there exists a unique solution $\varphi_\delta \in C^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ to the equation:

$$\mathcal{A}\varphi_\delta = f_\delta - \nu(f_\delta). \tag{5.25}$$

Observe from the proof of Theorem 6 (see equations (5.14) and (5.20)) and (5.22) that:

(5.26)

$$\|\nabla\varphi_\delta\|_\infty \leq \alpha^{-1}[f]_1,\ \forall \beta \in (0,1),\ \exists C_\beta > 0,\ \forall i \in \{1,2\},\ [\varphi_\delta^{(i)}]_1 \leq C_\beta(1+\|\nabla f_\delta\|_{\mathcal{C}^\beta}) \leq C_\beta \delta^{-\beta},\ [\varphi_\delta^{(3)}]_\beta \leq C_\beta \delta^{-\beta},$$
$$\forall x \in \mathbb{R}^d,\ |D_x^2 \varphi_\delta(x)| \leq \frac{C_\beta \delta^{-\beta}}{1+|x|}.$$

Now, from (5.25) the deviation in (5.23) rewrites:

$$\mathbb{P}\big[\sqrt{\Gamma_n}(\nu_n(f)-\nu(f)) \geq a\big] = \mathbb{P}\Big[\nu_n(\mathcal{A}\varphi_\delta) + R_{n,\delta}(f) \geq \frac{a}{\sqrt{\Gamma_n}}\Big]. \tag{5.27}$$

From (5.24), 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 $\nu_n(\mathcal{A}\varphi_\delta)$ can be analyzed as above, reproducing the proofs of Theorems 2 and 3, replacing the bounds on $([\varphi^{(i)}]_1)_{i\in\{1,2\}}, [\varphi^{(3)}]_\beta$ appearing therein by those of equation (5.26). Precisely, we get from (5.24), similarly to (2.21) (replacing the controls on φ by those on φ_δ in the proofs of Lemmas 3 and 5):

$$\begin{aligned}\mathbb{P}\left[\Big|\nu_n(\mathcal{A}\varphi_\delta) + R_{n,\delta}(f)\Big| \geq \frac{a}{\sqrt{\Gamma_n}}\right] &\leq 2\left[\mathbb{E}\exp\Big(-\frac{q\lambda_n}{\Gamma_n}M_n\Big)\right]^{\frac{1}{q}} \exp\Big(-\frac{a\lambda_n}{\sqrt{\Gamma_n}}(1-\frac{\sqrt{\Gamma_n}2C_\eta[f]_1\delta}{a})\Big) \\ &\quad\times \exp\Big(\frac{\lambda_n^2}{2\Gamma_n p} + \frac{p(a_n^\delta)^2}{2}\Big)\exp\left(\frac{3pC_{V,\varphi}^2\lambda_n^2}{c_V\Gamma_n^2} + \frac{c_V}{p}\right)(I_V^1)^{\frac{1}{2p}} \\ &\quad\times \exp\left(C_{2.17}^\delta \frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{\Gamma_n^2}\right)(I_V^1)^{\frac{1}{4p}} \times \exp\left(C_{2.18}^\delta\Big(\frac{3p\lambda_n^2(\Gamma_n^{(2)})^2}{2\Gamma_n^2} + \frac{1}{2p}\Big)\right)(I_V^1)^{\frac{1}{4p}}\end{aligned} \tag{5.28}$$

where $C_{2.18}^\delta := C_\beta\delta^{-\beta}\Big(\frac{C_V}{2}\big(\sup_{|x|\leq K}V(x)\big) \vee \frac{(CC_V)^2\bar{c}}{c_V}\Big)$ and $C_{2.17}^\delta := \frac{\|\sigma\|_\infty^4 C_\beta^2\delta^{-2\beta}}{4}\frac{C_V}{c_V}$ precisely correspond to the modifications of the constants $C_{2.18}$ and $C_{2.17} := \frac{\|\sigma\|_\infty^4[\varphi^{(2)}]_1^2}{4}\frac{C_V}{c_V}$ 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 $|D^2\varphi(x)| \leq \frac{C_\varphi}{1+|x|}$ by $|D^2\varphi_\delta(x)| \leq \frac{C_\beta\delta^{-\beta}C_\varphi}{1+|x|}$. Similarly,

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

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:

$$\begin{aligned}\mathbb{P}\left[|\nu_n(f)-\nu(f)| \geq \frac{a}{\sqrt{\Gamma_n}}\right] &= \mathbb{P}\left[\Big|\nu_n(\mathcal{A}\varphi_\delta) + R_{n,\delta}(f)\Big| \geq \frac{a}{\sqrt{\Gamma_n}}\right] \leq 2(I_V^1)^{\frac{1}{p}}\exp\Big(\frac{1}{p}\big[c_V + \frac{C_{2.18}^\delta}{2}\big] + \frac{p(a_n^\delta)^2}{2}\Big) \\ \times\exp\left(-\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\Big(1 - \frac{\sqrt{\Gamma_n}4C_\eta[f]_1\delta}{a} - \frac{1}{q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\Big\{\frac{p}{\Gamma_n}\Big(\frac{6C_{V,\varphi}^2}{c_V} + \big[2C_{2.17}^\delta + 3C_{2.18}^\delta\big](\Gamma_n^{(2)})^2\Big) + \frac{1}{p}\Big\}\Big)\right).\end{aligned}$$

The Young inequality yields that for all $\varepsilon_n > 0$:

$$\mathbb{P}\left[|\nu_n(f) - \nu(f)| \geq \frac{a}{\sqrt{\Gamma_n}}\right] \leq 2(I_V^1)^{\frac{1}{p}} \exp\Big(\frac{1}{p}\big[c_V + \frac{C^\delta_{2.18}}{2}\big] + \frac{p(a_n^\delta)^2}{2} + \varepsilon_n^{-1}\Gamma_n\delta^2\Big)$$

$$\times \exp\left(-\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\Big(1 - \underbrace{\frac{1}{q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\Big\{2\varepsilon_n C_\eta^2[f]_1^2 + \frac{p}{\Gamma_n}\Big(\frac{6C_{V,\varphi}^2}{c_V} + \big[2C^\delta_{2.17} + 3C^\delta_{2.18}\big](\Gamma_n^{(2)})^2\Big) + \frac{1}{p}\Big\}}_{=:d_n^\delta}\Big)\right). \tag{5.29}$$

We now want to let $p := p(n) \underset{n}{\to} +\infty,\ \varepsilon_n \underset{n}{\to} 0$ so that the associated contributions in the above equation can be viewed as remainders. From the previous definitions of $C^\delta_{2.18},\ C^\delta_{2.17}$, we see that, to achieve this goal, two constraints need to be fulfilled: namely, we must choose $\delta,\ p$ such that

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

Now, if $\theta \in (1/2, 1]$ there exists $\beta \in (0,1)$ such that $\Gamma^{(\frac{3+\beta}{2})} \leq C$. In that case: $a_n^\delta \leq \frac{C}{\sqrt{\Gamma_n}}\delta^{-\beta} = \Gamma_n^{-(\frac{1}{2}(1-\beta)-\beta\varepsilon)} \underset{n}{\to} 0$ 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 $p(a_n^\delta)^2 \underset{n}{\to} 0$. On the other hand, $\varepsilon_n = \Gamma_n^{-\varepsilon}$ also yields $\varepsilon_n^{-1}\Gamma_n\delta^2 = \Gamma_n^{-\varepsilon} \underset{n}{\to} 0$.

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

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

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 have established the following concentration result.

Theorem 7 (Non-asymptotic concentration bounds for Lipschitz functions)**.** *Assume* **(A)**, **(UE)**, **($\mathbf{R}_{b,\sigma}$)** *and either* **(D)** *or* **($\mathbf{BE}_\rho$)** *for some $\rho > 0$ are in force. Let f be 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 $n \geq 1$ and for every $a > 0$:*

$$\mathbb{P}\big[|\sqrt{\Gamma_n}\big(\nu_n(f) - \nu(f)\big)| \geq a\big] \leq 2C_n \exp\left(-c_n\frac{a^2\alpha^2}{2\|\sigma\|_\infty^2[f]_1^2}\right)$$

where α is as in Lemma 6.

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

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

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 $k \geq 1,\ \gamma_k = 1/k$:

$$\nu_n^Z := \frac{1}{\Gamma_n}\sum_{k=1}^n \gamma_k \delta_{Z_k} \overset{w,\ a.s.}{\underset{n}{\longrightarrow}} G,\ G(dx) := \exp\Big(-\frac{|x|^2}{2}\Big)\frac{dx}{(2\pi)^{d/2}}. \tag{6.1}$$

The above convergence had been established in [LP02], as a by-product of their results concerning the approximation of invariant measures, under the minimal moment condition $U_i \in L^2(\mathbb{P})$, thus weakening the initial assumptions by Brosamler and Schatte (see [Bro88] and [Sch88]). 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 $n \geq 0$:

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

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 [LP02] 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 ν_n^Z satisfies the conclusion of Theorem 1 point *(a)* (when $\gamma_n = \frac{1}{n}$, $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \underset{n}{\to} 0$), *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 [CH92], for real valued i.i.d. random variables, Chaâbane and Maâouia [CM00], who investigate the convergence rate of the strong quadratic law of large numbers for some extensions to vector-valued martingales, and Heck [Hec98], 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 [LP02] for the associated *a.s.* CLT to hold.

6.1.1. *Non-Asymptotic Deviation Bounds.*

Theorem 8. *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 $n \geq 1$:*

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

where φ denotes the solution of the Poisson equation:

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

which, under the current assumptions, is unique and belongs to $H^1(\mathbb{R}^d, \mathbb{R})$ 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 ν_n^Z in (6.1), write $\nu_n^Z(f) = \frac{1}{\Gamma_n}\sum_{k=1}^n \gamma_k f(Z_{k-1})$. We also have similarly $\nu_n^X(f) := \frac{1}{\Gamma_n}\sum_{k=1}^n \gamma_k f(X_{k-1})$. For all $\lambda > 0$, we derive similarly to (2.12) (see as well (5.28)) and with the notations of

(5.23):

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

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 $\nu_n^\Delta(|\cdot|):=\frac{1}{\Gamma_n}\sum_{k=1}^n\gamma_k|\Delta_{k-1}|$.

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

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

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

On the other hand, from Section 5.4 we have that $\varphi_\delta\in\mathcal{C}^{3,\beta}(\mathbb{R}^d,\mathbb{R})$ for all $\beta\in(0,1)$. We derive from (6.6), (6.7) similarly to the proof of Theorem 7 by setting $\bar{\lambda}_n:=\frac{a\sqrt{\Gamma_n}}{q\bar{q}\|\nabla\varphi\|_\infty^2}$:

$$\begin{aligned}
\mathbb{P}\big[\sqrt{\Gamma_n}|\nu_n^Z(f)|\geq a\big] &\leq 2\exp\Big(-\frac{a^2}{2q\bar{q}\|\nabla\varphi\|_\infty^2}\Big(1-\frac{4\sqrt{\Gamma_n}C_\eta[f]_1\delta}{a}\Big)\Big)\exp\Big(\frac{C_{6.7}\bar{\lambda}_n[f]_1\mathbb{E}\left[|U_1|\right]\Gamma_n^{(\frac{3}{2})}}{\Gamma_n}\Big)\\
&\quad\exp\Big(\frac{C_{6.7}^2\bar{p}[f]_1^2\bar{\lambda}_n^2\Gamma_n^{(3)}}{2\Gamma_n^2}\Big)(I_V^1)^{\frac{1}{p\bar{q}}}\exp\Big(\frac{1}{p\bar{q}}\big(c_V+\frac{C_{2.18}^\delta}{2}\big)+\frac{(pa_n^\delta)^2}{2\bar{q}}\Big)\\
&\quad\times\exp\Big(\bar{\lambda}_n^2\Big(p\bar{q}\Big(\frac{3C_{V,\varphi}^2}{c_V\Gamma_n^2}+\big[C_{2.17}^\delta+\frac{3}{2}C_{2.18}^\delta\big]\frac{(\Gamma_n^{(2)})^2}{\Gamma_n^2}\Big)+\frac{1}{2p\bar{q}}\Big)\Big)\\
&\leq 2(I_V^1)^{\frac{1}{p\bar{q}}}\exp\Big(\frac{1}{p\bar{q}}\big(c_V+\frac{C_{2.18}^\delta}{2}\big)+\frac{p(a_n^\delta)^2}{2\bar{q}}+\varepsilon_n^{-1}\Gamma_n\delta^2\Big)\\
&\quad\times\exp\left(-\frac{a^2}{2q\bar{q}\|\nabla\varphi\|_\infty^2}\Big(1-d_n^\delta-\frac{\bar{p}}{q\bar{q}\|\nabla\varphi\|_\infty^2}\frac{[f]_1^2C_{6.7}^2\Big(\Gamma_n^{(3)}+\mathbb{E}\left[|U_1|\right]^2(\Gamma_n^{(\frac{3}{2})})^2\Big)}{\Gamma_n}\Big)\right),
\end{aligned}$$

for $\varepsilon_n>0$ and d_n^δ as in (5.29). Choose again $(p_n)_{n\geq1}$ and δ as in Section 5.4 so that $q_n\underset{n}{\to}1, d_n^\delta\underset{n}{\to}0$ with the indicated monotonicity for n large enough. We can now take $\bar{p}:=\bar{p}_n\underset{n}{\to}+\infty$ such that $\frac{\bar{p}}{\Gamma_n}\underset{n}{\to}0$. The above inequality then gives the result up to a direct modification of the sequences $(C_n)_{n\geq1},(c_n)_{n\geq1}$. □

Proof of Lemma 7. The definition of Δ_n implies:

$$\Delta_{n+1}=\Delta_n\Big(1-\frac{\gamma_{n+1}}{2}\Big)+r_nZ_n,$$

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 all $n\geq1$,

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

Setting now $\rho_0=1$ and for $n\geq1$:

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

a direct induction on Δ_n yields:

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

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 all $n \geq 1$:

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

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

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

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

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

Plugging this inequality in (6.11), we derive:

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

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

$$\begin{aligned}\mathbb{E}\exp\Big(\lambda\nu_n^{\Delta}(|\cdot|)\Big) &\leq \prod_{k=1}^{n-1}\mathbb{E}\exp\Big(\frac{C_4\lambda}{\Gamma_n k^{\frac{3}{2}}}|U_k|\Big) \leq \prod_{k=1}^{n-1}\exp\Big(\frac{\bar{C}_4\lambda}{\Gamma_n k^{\frac{3}{2}}}\mathbb{E}\left[|U_1|\right] + \frac{1}{2}\Big(\frac{\bar{C}_4\lambda}{\Gamma_n k^{\frac{3}{2}}}\Big)^2\Big)\\ &= \exp\Big(\frac{\bar{C}_4\lambda\mathbb{E}\left[|U_1|\right]\Gamma_n^{(\frac{3}{2})}}{\Gamma_n} + \frac{\bar{C}_4^2\lambda^2\Gamma_n^{(3)}}{2\Gamma_n^2}\Big).\end{aligned}$$

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 $\mathrm{span}\{X_1, [X_1, X_0]\} = \mathbb{R}$. 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 Theorem 2 follow from Theorem 18 in Rotschild and Stein [RS76] (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 $a \in \mathbb{R}^+ \mapsto g_{n,\theta}(a) := \begin{cases}\log\big(\mathbb{P}[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)| \geq a]\big), \ \theta \in (1/3, 1],\\ \log\big(\mathbb{P}[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) + (\mathcal{B}_{n,1} - E_n^1)| \geq a]\big), \ \theta = 1/3,\end{cases}$ is such that for $a > a_n := a_n(\theta)$ where for $\theta \in (1/3, 1]$, $a_n(\theta) = 0$ and for $\theta = 1/3$, $a_n(\theta) = \frac{[\varphi^{(3)}]_\beta\|\sigma\|_\infty^{(3+\beta)}\mathbb{E}\left[|U_1|^{3+\beta}\right]}{(1+\beta)(2+\beta)(3+\beta)}\frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}}$:

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

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 $j \in [\![1, 5]\!]$, $\varphi(x) = x + \varepsilon\cos(x)$ 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 [GL00] or [Pag97] 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 φin

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 $S_{n,\theta}(a) := -\frac{(a-a_n(\theta))^2}{2\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}$, $S_{n,\theta,c}(a) := -\frac{(a-a_n(\theta))^2}{2\nu_{n_c}(\sigma^2)\|\nabla\varphi\|_\infty^2}$, $S_{n,\theta,A}(a) := -\frac{(a-a_n(\theta))^2}{2\nu_{n_c}(|\sigma\nabla\varphi|^2)}$ and the optimal concentration $P(\lambda_{\min})(n,\theta,a,\rho)$, obtained in Remark 11, optimizing numerically in ρ. The quantities $\nu_{n_c}(\sigma^2), \nu_{n_c}(|\sigma\nabla\varphi|^2)$ 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 3 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 $(S_{n,\theta_j})_{j\in[\![1,5]\!]}, (S_{n,\theta_j,c})_{j\in[\![1,5]\!]}, (S_{n,\theta_j,A})_{j\in[\![1,5]\!]}, \big(P(\lambda_{\min})(n,\theta_j,a,\rho)\big)_{j\in[\![1,5]\!]}$ corresponding to $j = 1$. The associated curves are denoted by $S_n, S_{n,c}, S_{n,A}$ and $P(\lambda_{\min})(n)$.

The 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 $S_{n,\theta_0,A}$.

The figures below thus illustrate that the explicit *optimal rate* of Remark 11 seems rather appropriate to capture the deviations of the empirical random measures.

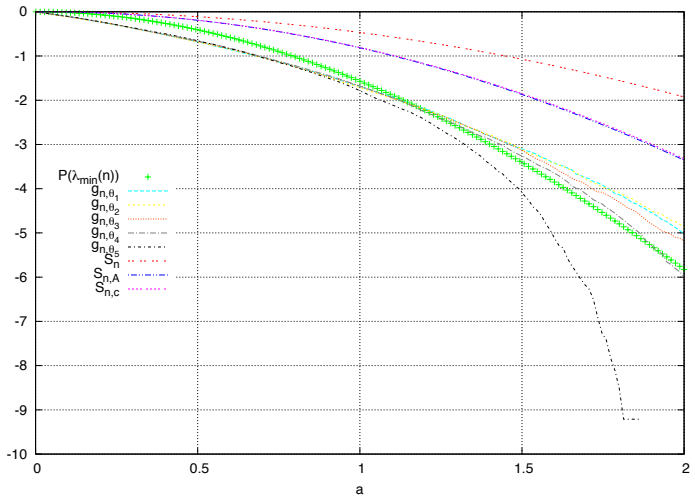

FIGURE 1. Unbiased Case. Plot of $a \mapsto g_{n,\theta}(a)$, for $(\theta_k)_{k\in[\![1,5]\!]}$, with $\varphi(x) = \sigma(x) = x + \varepsilon\cos(x)$, $\varepsilon = 0.01$.

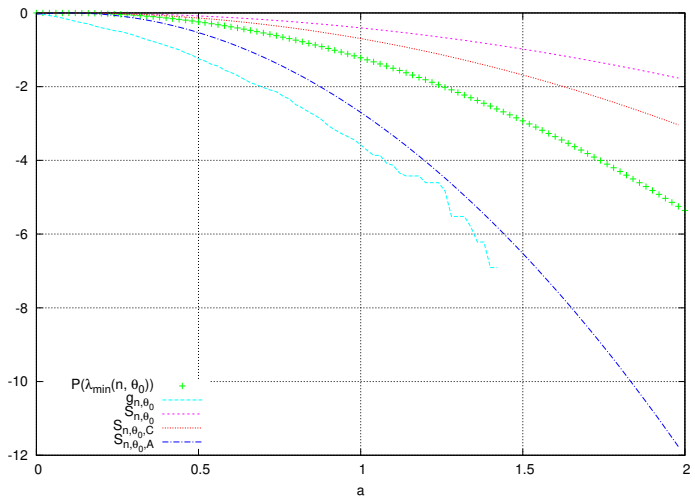

FIGURE 2. Biased Case. Plot of $a \mapsto g_{n,\theta}(a)$, for $\theta_0 = \frac{1}{3}$, with $\varphi(x) = \sigma(x) = \cos(x)$.

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 3 cannot be extended to justify such a choice, the empirical evidence is rather striking.

6.2.2. *Slutsky like result.* In this paragraph, we illustrate our results from Theorem 5, which can be viewed as an extension of the usual Slustky'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 6 and Lemma 6. We consider:

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

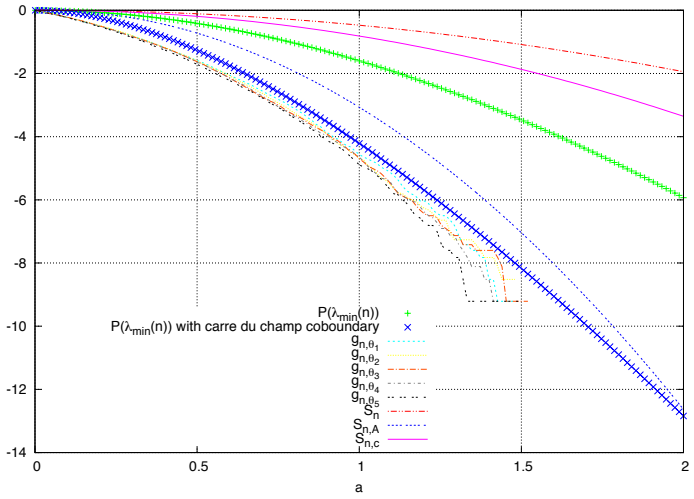

FIGURE 3. Plot of $a \mapsto g_{n,\theta}(a)$, for $(\theta_k)_{k\in[\![1,5]\!]}$, with $\varphi(x) = \sigma(x) = \cos(x)$.

Remark that the non-degeneracy condition **(UE)** is fulfilled by $\Sigma = \sigma\sigma^*$, as well as the condition set in Theorem 6, $\Sigma_{i,j}(x) = \Sigma_{i,j}(x_i, \ldots, x_d)$, for all $1 \le i \le j \le d$. 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 **(D)** 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 $x \in \mathbb{R}^d$, $\xi \in \mathbb{R}^d$:

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

Also, the condition $\mathbf{(R}_{1,\beta}\mathbf{)}$ is clearly satisfied. In other words, all assumptions of Theorem 5 hold. We set for the following plot:

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

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

Unlike in the previous simulations, we do not know here the value of $\nu(f)$. In fact, in paragraph 6.2.1 we had chosen to compute the deviation of $\mathcal{A}\varphi$ from $0 = \nu(\mathcal{A}\varphi)$. Here, we estimate from the ergodic theorem $\nu(f)$, taking $\beta = .5$, by $\nu_{n^c}(f) \approx 0.71308$ 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.

In Figure 4, we observe that the curve S_σ stays above g_n^σ as proved in Theorem 5. 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.

ACKNOWLEDGMENTS

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.

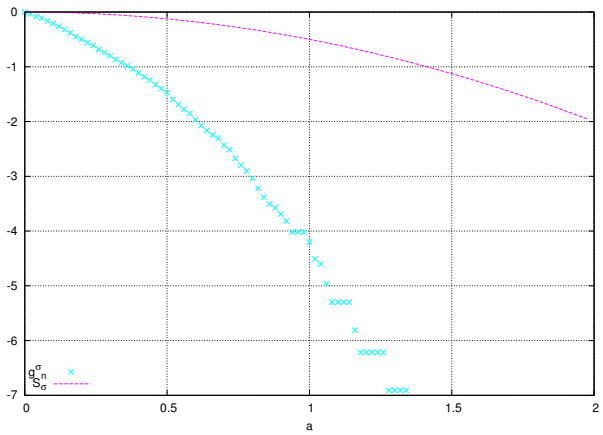

FIGURE 4. Plot of $a \mapsto g_n(a)$ with $f(x) = \frac{|x|^\beta}{1+|x|^\beta}$, $\beta = .5$.

REFERENCES

[ACJ08] A. Arnold, A. Carlen, and Q. Ju. Large-time behavior of non-symmetric Fokker-Planck type equations. *Comm. Stoch. Analysis*, 2–1:153–175, 2008.

[AMTU01] 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. *Comm. Partial Diff. Equations*, 26-(1-2):43–100, 2001.

[Bas97] R. F. Bass. *Diffusions and Elliptic Operators*. Springer, 1997.

[BB06] G. Blower and F. Bolley. Concentration inequalities on product spaces with applications to Markov processes. *Studia Mathematica*, 175-1:47–72, 2006.

[BE85] D. Bakry and M. Émery. Diffusions hypercontractives. *Séminaire de probabilités*, XIX:177–206, 1985.

[BGL14] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.

[BGV07] F. Bolley, A. Guillin, and C. Villani. Quantitative concentration inequalities for empirical measures on non-compact spaces. *Prob. Th. Rel. Fields*, 137:541–593, 2007.

[Bha82] 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):185–201, 1982.

[Boc13] S. Boccia. Schauder estimates for solutions of higher-order parabolic systems. *Methods Appl. Anal.*, 20(1):47–67, 2013.

[Boi11] E. Boissard. Simple bounds for the convergence of empirical and occupation measures in 1-Wasserstein distance. *Electronic Journal of Probability*, 16, 2011.

[Bro88] G. A. Brosamler. An almost everywhere central limit theorem. *Math. Proc. Cambridge Philos. Soc.*, 104(3):561–574, 1988.

[CH92] M. Csörgő and L. Horváth. Invariance principles for logarithmic averages. *Math. Proc. Cambridge Philos. Soc.*, 112(1):195–205, 1992.

[CM00] F. Chaabane and F. Maaouia. Théorèmes limites avec poids pour les martingales vectorielles. *ESAIM Probab. Statist.*, 4:137–189 (electronic), 2000.

[DG15] J. Dedecker and S. Gouëzel. Subgaussian concentration inequalities for geometrically ergodic markov chains. *Electronic Communications in Probability*, 20, Article 64:1–12, 2015.

[Duf90] M. Duflo. *Méthodes récursives aléatoires*. Techniques Stochastiques. [Stochastic Techniques]. Masson, Paris, 1990.

[EK86] E. Ethier and T. Kurtz. *Markov Processes. Characterization and Convergence*. Wiley, 1986.

[FM12] N. Frikha and S. Menozzi. Concentration bounds for stochastic approximations. *Electron. Commun. Probab.*, 17:no. 47, 15, 2012.

[GL00] S. Graf and H. Lushgy. *Foundations of quantization for random vectors*. LNM-1730, Springer-Verlag, 2000.

[Hec98] M. K. Heck. The principle of large deviations for the almost everywhere central limit theorem. *Stochastic Process. Appl.*, 76(1):61–75, 1998.

[IW80] N. Ikeda and S. Watanabe. *Stochastic differential equations*. North Holland, 1980.

[JO10] A. Joulin and Y. Ollivier. Curvature, concentration and error estimates for Markov chain Monte Carlo. *Ann. Probab.*, 38- 6:2418–2442, 2010.

[KM11] R. Khasminskii and G.N. Milstein. *Stochastic Stability of Differential Equations*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2011.

[KP10] N. V. Krylov and E. Priola. Elliptic and parabolic second-order PDEs with growing coefficients. *Comm. Partial Differential Equations*, 35(1):1–22, 2010.

[Led99] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Math.*, pages 120–216. Springer, Berlin, 1999.

[Lem05] V. Lemaire. An adaptive scheme for the approximation of dissipative systems. February 2005.

[LP02] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion. *Bernoulli*, 8–3:367–405, 2002.

[LP03] 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):435–451, 2003.

[MPR11] F. Merlevède, M. Peligrad, and E. Rio. A Bernstein type inequality and moderate deviations for weakly dependent sequences. *Probability Theory and Related Fields*, 151-3:435–474, 2011.

[MT06] F. Malrieu and D. Talay. Concentration inequalities for Euler Schemes. In H. Niederreiter and D. Talay, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2004*, pages 355–371. Springer Berlin Heidelberg, 2006.

[Pag97] G. Pagès. A space vector quantization method for numerical integration. *Journal of Applied and computational Mathematics*, 89:1–38, 1997.

[PP12] G. Pagès and F. Panloup. Ergodic approximation of the distribution of a stationary diffusion: rate of convergence. *Ann. Appl. Probab.*, 22(3):1059–1100, 2012.

[Pri09] E. Priola. Global Schauder estimates for a class of degenerate Kolmogorov equations. *Studia Math.*, 194(2):117–153, 2009.

[PV01] E. Pardoux and A. Veretennikov. On the Poisson Equation and Diffusion Approximation. I. *Ann. Probab.*, 29–3:1061–1085, 2001.

[RS76] L. P. Rothschild and E. M. Stein. Hypoelliptic differential operators and nilpotent groups. *Acta Math.*, 137(3-4):247–320, 1976.

[Sch88] P. Schatte. On strong versions of the Central Limit Theorem. *Math. Nachr.*, 137:249–256, 1988.

[Tal02] D. Talay. Stochastic Hamiltonian dissipative systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Processes and Related Fields*, 8–2:163–198, 2002.

[TT90] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch. Anal. and App.*, 8-4:94–120, 1990.

[Vil09] C. Villani. Hypocoercivity. *Mem. Amer. Math. Soc.*, 202(950):iv+141, 2009.

Université d'Evry Val d'Essonne
E-mail address: igor.honore@univ-evry.fr
Current address: Laboratoire de Mathématiques et Modélisation d'Évry (LaMME), 23 Boulevard de France, 91037, Evry, France.

Université d'Évry Val d'Essonne and Higher School of Economics
E-mail address: stephane.menozzi@univ-evry.fr
Current address: Laboratoire de Mathématiques et Modélisation d'Évry (LaMME), 23 Boulevard de France, 91037, Évry, France and Laboratory of Stochastic Analysis, HSE, Shabolovka 31, Moscow, Russian Federation.

Université Pierre et Marie Curie, Paris 6
E-mail address: gilles.pages@upmc.fr
Current address: Laboratoire de Probabilités et Modèles Aléatoires. Campus de Jussieu, case 188, 4 pl. Jussieu 75252 Paris Cedex 5, France.