



**HAL**  
open science

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

Igor Honoré, Stephane Menozzi, Gilles Pagès

► **To cite this version:**

Igor Honoré, Stephane Menozzi, Gilles Pagès. Non-Asymptotic Gaussian Estimates for the Recursive Approximation of the Invariant Measure of a Diffusion. 2016. hal-01321645v2

**HAL Id: hal-01321645**

**<https://hal.science/hal-01321645v2>**

Preprint submitted on 29 Jun 2016 (v2), last revised 23 May 2018 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# 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  $\nu$  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 bounds 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:

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

where  $(W_t)_{t \geq 0}$  is a Wiener process of dimension  $r$  on a given filtered probability space  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0}, \mathbb{P})$ ,  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ , and  $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^r$  are assumed to be Lipschitz continuous functions and to satisfy a suitable Lyapunov like condition ensuring existence of an invariant measure. We will also assume uniqueness of the invariant measure, denoted from now by  $\nu$ . We refer to the monographs by Khasminski and Milstein [KM11], Ethier and Kurtz [EK86] or Villani [Vil09], for a thorough discussion 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:

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

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  $\sigma$ -field on  $\mathbb{R}^d$ ):

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

The measure  $\nu_n$  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  $\Gamma_n := \sum_{k=1}^n \gamma_k \xrightarrow{n} +\infty$ . We also assume  $\gamma_k \xrightarrow{k} 0$ . Observe that, for a bounded  $\nu$ -a.s. continuous function  $f$ , it can be shown, see e.g. Theorem 1 in [LP02], that:

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

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

---

*Date:* June 29, 2016.

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

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  $\nu^h$  denoting the invariant measure of the scheme. 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 previously investigated by Lamberton and Pagès [LP02] for functions  $f$  writing as  $f - \nu(f) = \mathcal{A}\varphi$ , where  $\mathcal{A}$  stands for the infinitesimal generator of (1.1), i.e.  $f - \nu(f)$  is a *coboundary*. The specific reason of investigating such kind of functions simply follows from the fact that the invariant measure  $\nu$  solves, at least in the distributional sense, the Poisson Equation  $\mathcal{A}^*\nu = 0$  (where  $\mathcal{A}^*$  stands for the adjoint of  $\mathcal{A}$ ). Thus, for a smooth function  $\varphi$ , one gets that  $\nu(\mathcal{A}\varphi) = \int_{\mathbb{R}^d} \mathcal{A}\varphi(x)\nu(dx) = 0$ . The authors then investigate the convergence in law of the renormalized difference  $\nu_n(\mathcal{A}\varphi) - \nu(\mathcal{A}\varphi) = \nu_n(\mathcal{A}\varphi)$ . We emphasize that the CLT for functions of that type does not require non-degeneracy conditions. These assumptions naturally appear to investigate the limit behavior of a suitable renormalization of the quantity  $\nu_n(f) - \nu(f)$  for a general and possibly unbounded given  $f$ . In that case, the Poisson Equation  $\mathcal{A}\varphi = f - \nu(f)$  needs to be solved and this is precisely for this step that some structure conditions are needed. We refer for instance to the work of Pardoux and Veretennikov [PV01], Rothschild 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 will be to establish, in such cases, a non-asymptotic Gaussian control for the deviations of the quantity  $\nu_n(f) - \nu(f)$ . 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. Also, when  $\|\sigma\|^2 - \nu(\|\sigma\|^2)$  is a coboundary, we manage to improve our analysis, to derive better concentration bounds in a certain deviation range. This requires to investigate the auxiliary Poisson problem  $\mathcal{A}\vartheta = \|\sigma\|^2 - \nu(\|\sigma\|^2)$  and as a matter of fact, some additional deviation regimes appear. Also, this additional study seems rather efficient to capture the numerical behavior of the empirical deviations. We refer to Section 4 and 7 for details about these points.

**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  $\exists n_0 \in \mathbb{N}$ ,  $\exists C \geq 1$  s.t.  $\forall n \geq n_0$ ,  $C^{-1}v_n \leq u_n \leq Cv_n$ .

### 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:

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

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

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

Observe that if  $U_1 \stackrel{(\text{law})}{=} \mathcal{N}(0, I_r)$  or  $U_1 \stackrel{(\text{law})}{=} \left(\frac{1}{2}(\delta_1 + \delta_{-1})\right)^{\otimes r}$ , 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  $\varpi > 0$  s.t. for all  $\lambda > 0$ :

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

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  $\kappa$  s.t., defining for all  $x \in \mathbb{R}^d$ ,  $\Sigma(x) := \sigma \sigma^*(x)$ :

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

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

**(L<sub>V</sub>)** We have the following Lyapunov like stability condition:

There exists  $V : \mathbb{R}^d \rightarrow [v^*, +\infty[$  with  $v^* > 0$  s.t.

- i)  $V$  is a  $\mathcal{C}^2$  function,  $\|D^2V\|_\infty < \infty$ , and  $\lim_{|x| \rightarrow \infty} V(x) = +\infty$ .
- ii) There exists  $C_V \in (0, +\infty)$  s.t. 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:

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

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  $M \in \mathbb{R}^d \otimes \mathbb{R}^d$ ,  $\text{Tr}(M)$  denotes the trace of the matrix  $M$ .

There exist  $\alpha_V > 0$ ,  $\beta_V \in \mathbb{R}^+$  s.t. 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  $\nu$  to equation (1.1).

**(S)** For  $V$  satisfying **(L<sub>V</sub>)**:

- i) There exist constants  $K$  and  $\bar{c}$  s.t. for  $|x| \geq K$ ,  $|V(x)| \leq \bar{c}|x|^2$ .
- ii) We assume that the sequence  $(\gamma_k)_{k \geq 1}$  satisfies for all  $k \geq 1$ ,  $\gamma_k \leq \frac{1}{2} \min(\frac{1}{\sqrt{C_V \bar{c}}}, \frac{\alpha_V}{C_V \|D^2V\|_\infty})$ .

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.** *The above conditions actually imply that the drift coefficient  $b$  lies, out of a compact set, between two hyperplanes separated from 0. Also, the Lyapunov function is equivalent to the square norm.*

*We have assumed **(U)** without imposing some specific non degeneracy conditions. Observe that **(L<sub>V</sub>)** 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 3 (see especially the proof of Lemma 5 below).*

We say that assumption **(A)** holds whenever **(C1)**, **(GC)**, **(C2)**, **(L<sub>V</sub>)**, **(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)**, **(L<sub>V</sub>)**, **(U)** are satisfied and  $\mathbb{E}[U_1] = 0$ , the following Central Limit Theorem (CLT) holds (see Theorems 9, 10 in [LP02]).

**Theorem 1.** *Under **(C2)**, **(L<sub>V</sub>)**, **(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 all globally Lipschitz function  $\varphi$  in  $\mathcal{C}^3(\mathbb{R}^d, \mathbb{R})$  with  $D^2\varphi$  and  $D^3\varphi$  bounded, one has:

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

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

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

where

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

and  $\mu$  denotes the law of the innovations  $(U_k)_{k \geq 1}$ . In the above definition of  $\Phi_4$ , the term  $D^3\varphi$  stands for the order 3 tensor  $(\partial_{x_i, x_j, x_k}^3 \varphi)_{(i,j,k) \in \llbracket 1, d \rrbracket^3}$  and we denote, for all  $x \in \mathbb{R}^d$ , by  $D^3\varphi(x)b(x)$  the  $\mathbb{R}^d \otimes \mathbb{R}^d$  matrix with entries  $(D^3\varphi(x)b(x))_{ij} = \sum_{k=1}^d (D^3\varphi(x))_{ijk} b_k(x)$ ,  $(i, j) \in \llbracket 1, d \rrbracket^2$ .

**Remark 2.** Let us specify that for a step sequence  $(\gamma_n)_{n \in \mathbb{N}}$  s.t.  $\gamma_n \asymp n^{-\theta}$ ,  $\theta \in (0, 1]$ , it is easily checked that case (a) occurs for  $\theta \in (\frac{1}{3}, 1]$  for which  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow 0$ . In case (b), that is for  $\theta \in (0, \frac{1}{3}]$ ,  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow \tilde{\gamma}$ , with  $\tilde{\gamma} < +\infty$  for  $\theta = \frac{1}{3}$  and  $\tilde{\gamma} = +\infty$  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  $\lim_{|x| \rightarrow +\infty} \frac{|\sigma^*\nabla\varphi(x)|^2}{V(x)} = 0$  (strictly sublinear diffusion) in case (a) and  $\sup_{x \in \mathbb{R}^d} \frac{|\sigma^*\nabla\varphi(x)|^2}{V(x)} < +\infty$  (sublinear diffusion) in case (b). We refer again to Theorems 9 and 10 in [LP02] for further considerations.

**Remark 3.** Observe that the result is stated for functions of the form  $\mathcal{A}\varphi$ . Let us recall that the invariant measure  $\nu$  solves, at least in the distributional sense, the Fokker-Planck equation  $\mathcal{A}^*\nu = 0$ , where  $\mathcal{A}^*$  stand for the adjoint operator of  $\mathcal{A}$ . Thus, for a smooth function  $\varphi$ ,

$$\nu(\mathcal{A}\varphi) := \int \mathcal{A}\varphi(x)\nu(dx) = \int \varphi\mathcal{A}^*\nu(dx) = 0.$$

Note as well that the asymptotic variance corresponds to the usual integral of the ‘‘carré du champ’’ w.r.t. to the invariant measure, see 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)**, **(L<sub>V</sub>)** (see Bhattacharya [Bha82]). In fact **(C2)** can be partially relaxed as well, like mentioned above. Precisely,  $\frac{1}{\sqrt{t}} \int_0^t f(Y_s) ds \xrightarrow{\mathcal{L}} \mathcal{N}(0, \int_{\mathbb{R}^d} |\sigma^*\nabla\varphi(x)|^2 \nu(dx))$  as  $t \rightarrow +\infty$ . In both cases, the normalization is the same: the square root of the considered running time,  $t$  for the diffusion and  $\Gamma_n$  for the scheme. Anyhow, the fastest convergence is obtained for  $\theta = 1/3$  (critical value between ‘‘fast’’ and ‘‘slow’’ settings), for which the discretization bias steps in. This bias might be seen as a discretization effect. When  $\theta \in (0, 1/3)$  (slow decreasing step), this discretization effect becomes prominent and ‘‘hides’’ the CLT.

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 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. Let us as well mention the recent work by Dedecker and Gouëzel [DG15] who also use this 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, we think that the martingale-based approach we adopt in this work is rather simple, robust and supports very naturally 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,...). Also, in order to attain the fastest convergence rate, we will thoroughly discuss how the associated bias appearing in Theorem 1 and Theorems 2, 3 below can be numerically estimated.
- 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]. This application also requires to investigate carefully the associated 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. In Section 4, we prove that when  $\|\sigma\|^2 - \nu(\|\sigma\|^2)$  is itself a coboundary, a mixed regime appears in the non-asymptotic deviation bounds which dramatically improves 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*. Section 5 and 6 are respectively dedicated to the two applications mentioned above, the

practical derivation of some non-asymptotic confidence intervals, and the non-asymptotic deviation bounds for the almost-sure CLT. Eventually, we conclude in Section 7 with some numerical results.

**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.

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

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

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

$$(1.5) \quad [f]_{\beta'} \leq [f]_\beta^{\frac{\beta}{\beta'}} (2\|f\|_\infty)^{1 - \frac{\beta}{\beta'}}.$$

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

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

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

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

## 2. MAIN RESULTS

**2.1. Result of non-asymptotic Gaussian concentration.** Our first main result is the following Theorem which provides the non-asymptotic counterpart of the limit Theorem 1.

**Theorem 2.** *Assume **(A)** holds. Let  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  be a Lipschitz continuous function with  $\mathcal{C}^3$  regularity and Lipschitz continuous bounded existing partial derivatives (up to order 3). We suppose furthermore that:*

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

Assume the step sequence  $(\gamma_k)_{k \geq 1}$  is of the form  $\gamma_k \asymp k^{-\theta}$ ,  $\theta \in [1/3, 1]$ .

(a) Fast Decreasing Step (case  $\theta \in (\frac{1}{3}, 1]$ ):

Assume that there exists  $C_\varphi > 0$  s.t. for all  $x \in \mathbb{R}^d$ ,  $|D^2 \varphi(x)| \leq C_\varphi/(1 + |x|)$ .

Then, there exist explicit non-negative sequences  $(c_n)_{n \geq 1}$  and  $(C_n)_{n \geq 1}$ , respectively increasing and decreasing for  $n$  large enough, with  $\lim_n C_n = \lim_n c_n = 1$  such that for every  $n \geq 1$  and every  $a \geq a_n$ :

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

$$\text{where } a_n = \frac{[\varphi^{(3)}]_1 \|\sigma\|_\infty^4 \mathbb{E}[|U_1|^4] \Gamma_n^{(2)}}{4! \sqrt{\Gamma_n}} = \begin{cases} O(\ln(n)^{-\frac{1}{2}}), & \text{if } \theta = 1, \\ O(n^{\frac{\theta-1}{2}}), & \text{if } \theta \in (1/2, 1), \\ O(\ln(n)n^{-\frac{1}{4}}), & \text{if } \theta = 1/2, \\ O(n^{\frac{1-3\theta}{2}}) & \text{if } \theta \in (1/3, 1/2). \end{cases} \quad \text{In any case, we have that } a_n \xrightarrow{n} 0.$$

(b) Critical Decreasing Step (case  $\theta = \frac{1}{3}$ ): Assume  $\varphi$  is  $\mathcal{C}^4$ . There exist explicit non-negative sequences  $(c_n)_{n \geq 1}$  and  $(C_n)_{n \geq 1}$ , respectively increasing and decreasing for  $n$  large enough, with  $\lim_n C_n = \lim_n c_n = 1$  such that for all  $n \geq 1$  and for every  $a > 0$ :

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

where  $(E_n)_{n \geq 1}$  is a sequence of exponentially integrable random variables<sup>1</sup>, i.e.

$$\exists \lambda_0 > 0, \forall \lambda \leq \lambda_0, \sup_{n \geq 1} \mathbb{E} \exp(\lambda |E_n|) < +\infty$$

satisfying  $E_n \xrightarrow[n]{n} 0$  a.s (see (2.4) further on for a precise definition of  $E_n$ ).

**Remark 5.** Let us observe that 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 not require non-degeneracy conditions and only depend on the diffusion coefficient through the upper bound of the diffusion matrix  $\Sigma$ , 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  $\mathcal{A}\varphi = f - \nu(f)$ .

Theorem 2 can actually be derived from the more general Theorem 3 below which gives concentration results for functions  $\varphi$  satisfying the previous assumptions up to order two and  $[\varphi^{(3)}]_\beta < +\infty$ ,  $\beta \in (0, 1]$ .

**Theorem 3.** Assume **(A)** holds. Consider a globally Lipschitz continuous (possibly unbounded) function  $\varphi \in \mathcal{C}^3(\mathbb{R}^d, \mathbb{R})$  with  $(D^\alpha \varphi)_{|\alpha| \in \{2,3\}}$  bounded and s.t.  $[\varphi^{(3)}]_\beta < +\infty$  for some  $\beta \in (0, 1]$ . We suppose furthermore that:

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

Assume the step sequence  $(\gamma_k)_{k \geq 1}$  is of the form  $\gamma_k \asymp k^{-\theta}$ ,  $\theta \in [1/3, 1]$ .

Set for all  $(k, t, u, x) \in \llbracket 1, n \rrbracket \times [0, 1]^2 \times \mathbb{R}^d$ :

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

and define subsequently:

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

(a) Assume  $(\beta \neq 1 \text{ and } \theta \in [\frac{1}{2+\beta}, 1])$  or  $(\beta = 1 \text{ and } \theta \in (\frac{1}{3}, 1])$

(i) Assume that there exists  $C_\varphi > 0$  s.t. for all  $x \in \mathbb{R}^d$ ,  $|D^2 \varphi(x)| \leq C_\varphi/(1 + |x|)$ .

Then, there exist explicit non-negative sequences  $(c_n)_{n \geq 1}$  and  $(C_n)_{n \geq 1}$ , respectively increasing and decreasing for  $n$  large enough, with  $\lim_n C_n = \lim_n c_n = 1$ , such that for all  $n \geq 1$  and for every  $a > 0$ :

$$\mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) + E_n^1| \geq a] \leq 2C_n \exp \left( -c_n \frac{a^2}{2\|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2} \right).$$

Moreover, we have  $|E_n^1| \leq a_n$  a.s. where

$$(2.3) \quad a_n = \frac{[\varphi^{(3)}]_\beta \|\sigma\|_\infty^{(3+\beta)} \mathbb{E}[|U_1|^{3+\beta}] \Gamma_n^{(\frac{3+\beta}{2})}}{(1+\beta)(2+\beta)(3+\beta) \sqrt{\Gamma_n}} \xrightarrow[n]{} a_\infty = \begin{cases} 0 & \text{if } \theta > \frac{1}{2+\beta}, \\ > 0 & \text{if } \theta = \frac{1}{2+\beta} \text{ (potential bias if } \beta \neq 1). \end{cases}$$

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

(b) Assume  $\beta = 1, \theta = \frac{1}{3}$ , we set

$$(2.4) \quad E_n := \tilde{\gamma} m + E_n^1 + \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^2 \int_0^1 (1-t) \text{Tr} \left( D^2 \varphi(X_{k-1} + t \gamma_k b_{k-1}) b_{k-1} \otimes b_{k-1} \right) dt \\ + \frac{1}{2\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k \text{Tr} \left( (D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2 \varphi(X_{k-1})) \Sigma_{k-1} \right).$$

1. If there exists  $C_\varphi > 0$ ,  $\forall x \in \mathbb{R}^d$ ,  $|D^2 \varphi(x)| \leq 1/(1 + |x|)$ ,  $(E_n)_{n \geq 1}$  is even square exponentially integrable, (see condition **(C1)**).



There exist explicit non-negative sequences  $(c_n)_{n \geq 1}$  and  $(C_n)_{n \geq 1}$ , respectively increasing and decreasing for  $n$  large enough, with  $\lim_n C_n = \lim_n c_n = 1$  such that for all  $n \geq 1$  and for every  $a > 0$ :

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

The random variables  $(E_n)_{n \geq 1}$  are exponentially integrable. If furthermore  $D^3\varphi$  is  $\mathcal{C}^1$ ,  $E_n \xrightarrow[n]{a.s.} 0$ . Basically, we control the deviations of the empirical mean adding a contribution that asymptotically annihilates the bias.

**Remark 6.** Observe that, for  $\beta \in (0, 1)$ , the choice of the time-step sequence as  $\gamma_k \asymp k^{-\theta}$  gives that for  $\theta > 1/(2 + \beta)$ ,  $\frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} \xrightarrow[n]{a.s.} 0$  so that  $a_n \xrightarrow[n]{a.s.} 0$  in such cases.

On the other hand, for  $\theta = 1/(2 + \beta)$  we obtain a bias in the previous deviation bounds. This generalizes the controls of Theorems 1 and 2, established for  $\beta = 1$ , and emphasizes the importance of the Hölder regularity of the function  $D^3\varphi$  in the appearance of a bias. We also think that, in this case, a corresponding biased CLT should hold.

Observe that we could derive from Theorem 3, up to the resolution of the Poisson Equation  $\mathcal{A}\varphi = f - \nu(f)$ , some non-asymptotic confidence interval for  $\nu_n(f)$ . However, in order to compute them in practice, some centering terms, namely  $E_n^1$  or  $E_n$ , need to be estimated as well. This induces a double difficulty: it firstly requires to explicitly know the solution  $\varphi$  of the Poisson Equation. Indeed, both  $E_n^1$  and  $E_n$  read as weighted sums of derivatives of  $\varphi$ . Secondly, since these terms contain conditional expectations, they need to be approximated as well. 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 Theorems 2 and 3), see Lemmas 3, 4 and 5.

**Lemma 1** (Decomposition of the empirical measure). *For all  $n \geq 1$ :*

$$\begin{aligned} \Gamma_n \nu_n(\mathcal{A}\varphi) &= \varphi(X_n) - \varphi(X_0) - \left[ \sum_{k=1}^n \gamma_k^2 \int_0^1 (1-t) \text{Tr}\left(D^2\varphi(X_{k-1} + t\gamma_k b_{k-1}) b_{k-1} \otimes b_{k-1}\right) dt \right. \\ &\quad \left. + \frac{1}{2} \sum_{k=1}^n \gamma_k \text{Tr}\left((D^2\varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2\varphi(X_{k-1})) \Sigma_{k-1}^2\right) + \sum_{k=1}^n \psi_k(X_{k-1}, U_k) \right], \end{aligned}$$

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  $\sqrt{\gamma_k} \|\sigma_{k-1}\| \|\nabla\varphi\|_\infty$ .

Note from the above expansion that a key tool to control the remainders is the integrability of the Lyapunov function. We are indeed led to handle terms of the form

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

for small enough real constants  $c > 0$ . The required integrability conditions to control such quantities are given by Proposition 1 (see also the Proof of Lemma 5).

For notational convenience we introduce for a given  $n \in \mathbb{N}^*$  the following quantities:

$$\begin{aligned}
L_n &:= \varphi(X_n) - \varphi(X_0), \\
D_{2,b,n} &:= \sum_{k=1}^n \gamma_k^2 \int_0^1 (1-t) \operatorname{Tr} \left( D^2 \varphi(X_{k-1} + t\gamma_k b_{k-1}) b_{k-1} \otimes b_{k-1} \right) dt, \\
D_{2,\Sigma,n} &:= \frac{1}{2} \sum_{k=1}^n \gamma_k \operatorname{Tr} \left( (D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2 \varphi(X_{k-1})) \Sigma_{k-1}^2 \right), \\
(2.5) \quad G_n &:= \sum_{k=1}^n \psi_k(X_{k-1}, U_k).
\end{aligned}$$

We also introduce  $\bar{G}_n$  the compensator of  $G_n$  and  $M_n$  the associated martingale. Namely:

$$(2.6) \quad M_n := G_n - \bar{G}_n, \quad \bar{G}_n := \sum_{k=1}^n \mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}].$$

We refer to the proof of Lemma 1 to check that the above definition of  $\bar{G}_n$  actually matches the term  $\sqrt{\Gamma_n} E_n^1$  introduced in equation (2.2) of Theorem 3. From (2.6),  $(M_k)_{k \geq 1}$  is a martingale. Importantly, from Lemma 1, writing as well  $R_n := L_n - D_{2,b,n} - D_{2,\Sigma,n}$ , we derive:

$$(2.7) \quad \nu_n(\mathcal{A}\varphi) = \frac{1}{\Gamma_n} (R_n - G_n) = \frac{1}{\Gamma_n} (R_n - M_n - \bar{G}_n).$$

We now split the analysis according to the cases (a) and (b) introduced in Theorem 3.

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

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

(b) In that case  $\theta = \frac{1}{3}$ ,  $\beta = 1$ . Denoting,  $D_{2,n} = D_{2,b,n} + D_{2,\Sigma,n}$ , we have from (2.5) and with the notations of (2.4),  $D_{2,n} = \sqrt{\Gamma_n} (E_n - E_n^1 - \tilde{\gamma}m)$ . We study the deviations of:

$$\begin{aligned}
(2.9) \quad \mathbb{P} [ |\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) - \tilde{\gamma}m + E_n| \geq a ] &= \mathbb{P} \left[ \left| \nu_n(\mathcal{A}\varphi) + \frac{\bar{G}_n + D_{2,n}}{\Gamma_n} \right| \geq \frac{a}{\sqrt{\Gamma_n}} \right] \\
&\leq 2 \exp \left( - \frac{a\lambda}{\sqrt{\Gamma_n}} \right) \left( \mathbb{E} \exp \left( \frac{q\lambda}{\Gamma_n} M_n \right) \right)^{\frac{1}{q}} \left( \mathbb{E} \exp \left( \frac{p\lambda}{\Gamma_n} |L_n| \right) \right)^{\frac{1}{p}}.
\end{aligned}$$

**Remark 7.** Observe that in case (a) we have subtracted the contributions that potentially give some bias at the limit, which is the case for  $\theta = \frac{1}{2+\beta}$ . Also, since  $\beta < 1$ , the other terms in  $R_n$  do not contribute to the potential bias. We insist that for  $\theta \in (\frac{1}{2+\beta}, 1]$ , there is no asymptotic bias in that case and  $\frac{\bar{G}_n}{\Gamma_n} \xrightarrow[n]{a.s.} 0$  (see Lemma 3 below). Note carefully that we keep this strategy in the case (b). Indeed, both terms  $\frac{\bar{G}_n}{\Gamma_n}, \frac{D_{2,n}}{\Gamma_n}$  give asymptotically some bias, actually for  $D^3\varphi \in \mathcal{C}^1$ ,  $\frac{\bar{G}_n + D_{2,n}}{\sqrt{\Gamma_n}} \xrightarrow[n]{} -\tilde{\gamma}m$  introduced in Theorem 1, and we again subtract them for our deviation study.

We also mention that, from a practical viewpoint, the subtracted terms can be rather efficiently computed along the ergodic mean, see Section 5.

The Lemma below provides the Gaussian contribution deriving from inequalities (2.8) and (2.9).

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

$$(2.10) \quad \lambda_n := \frac{a}{q \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2} \sqrt{\Gamma_n},$$

we derive:

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

**Lemma 3** (Bounds for the Bias). *With the above notations, we have that:*

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

Moreover,  $a_n \xrightarrow{n} a_\infty$  where  $a_\infty = 0$  if  $\theta > 1/(2+\beta)$  and  $a_\infty > 0$  for  $\theta = 1/(2+\beta)$  (bias).

As indicated before, we now aim at controlling the remainders. To this end, we will thoroughly rely on the following important integrability result for the Lyapunov function.

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

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

Let us mention that additional integrability results concerning possibly more general Lyapunov functions can be found in Lemaire (see [Lem05]).

We now have the following results for the terms appearing in (2.5).

**Lemma 4** (Initial term). *Let  $q \in (1, +\infty)$  be fixed and  $\lambda_n$  be as in (2.10) in Lemma 2. For functions  $\varphi$  satisfying  $(G_V)$ , i.e. there exists  $C_{V,\varphi} > 0$  s.t. for all  $x \in \mathbb{R}^d$ ,  $|\varphi(x)| \leq C_{V,\varphi}(1 + \sqrt{V(x)})$ , for  $p := \frac{q}{q-1}$  and  $j \in \{1, 2\}$ :*

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

with  $c_V, I_V^1$  like in Proposition 1.

**Lemma 5** (Remainders). *Let  $q \in (1, +\infty)$  be fixed and  $\lambda_n$  be as in Lemma 2. Then, there exists  $C_{2.11} := C_{2.11}(\mathbf{A}, \varphi)$  s.t. for  $p = \frac{q}{q-1}$ :*

$$(2.11) \quad \left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,\Sigma,n}|\right)\right)^{\frac{1}{4p}} \leq \exp\left(C_{2.11} \frac{p\lambda_n^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2}\right) (I_V^1)^{\frac{1}{4p}}.$$

We also have:

- If for all  $x \in \mathbb{R}^d$ ,  $|D^2\varphi(x)| \leq \frac{C_\varphi}{1+|x|}$  for some constant  $C_\varphi \geq 0$ , then there exists  $C_{2.12} := C(\mathbf{A}, \varphi) > 0$  s.t.

$$(2.12) \quad \left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}|\right)\right)^{\frac{1}{4p}} \leq \exp\left(C_{2.12} \left(\frac{3p\lambda_n^2 (\Gamma_n^{(2)})^2}{2\Gamma_n^2} + \frac{1}{2p}\right)\right) (I_V^1)^{\frac{1}{4p}}.$$

- For  $a \leq \frac{c_V q}{4C_V p} \frac{\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2 \sqrt{\Gamma_n}}{\|D^2\varphi\|_\infty^2 \Gamma_n^{(2)}}$ , there exists an  $\mathbb{R}^+$ -valued sequence  $(v_n)_{n \geq 1}$  s.t.  $|v_n| \leq C_{2.13} := C_{2.13}(\mathbf{A}, \varphi)$  and

$$(2.13) \quad \left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}|\right)\right)^{\frac{1}{4p}} \leq (I_V^1)^{v_n}.$$

Also,  $v_n \xrightarrow{n} v_\infty$  where  $v_\infty = 0$  if  $\theta > 1/3$  and  $v_\infty > 0$  for  $\theta = 1/3$ .

**Proof of Theorem 3.** From Lemma 2 we get:

$$(2.14) \quad \left(\mathbb{E} \exp\left(q\lambda_n \frac{M_n}{\Gamma_n}\right)\right)^{\frac{1}{q}} \exp\left(-\frac{a\lambda_n}{\sqrt{\Gamma_n}}\right) \leq \exp\left(-\frac{a^2}{2q\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}\right).$$

(a) We deal with the case  $\beta \in (0, 1), \theta \in [\frac{1}{2+\beta}, 1]$ . Indeed, the second setting  $\beta = 1, \theta \in (\frac{1}{3}, 1]$  will readily follow from this case by (1.5).

(i) We suppose that there exists  $C_\varphi > 0$  s.t. for all  $x \in \mathbb{R}^d$ ,  $|D^2\varphi(x)| \leq C_\varphi/(1+|x|)$ . Plugging in (2.8) the controls from (2.14), Lemma 4 (with  $j = 2$ ) and Lemma 5 (equations (2.11), (2.12)), we get:

$$\begin{aligned}
\mathbb{P} \left[ \left| \nu_n(\mathcal{A}\varphi) + \frac{\bar{G}_n}{\Gamma_n} \right| \geq \frac{a}{\sqrt{\Gamma_n}} \right] &\leq 2 \exp \left( -\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2} \right) \exp \left( \frac{3pC_{V,\varphi}^2\lambda_n^2}{c_V\Gamma_n^2} + \frac{c_V}{p} \right) (I_V^1)^{\frac{1}{2p}} \\
&\quad \times \exp \left( C_{2.11} \frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{\Gamma_n^2} \right) (I_V^1)^{\frac{1}{4p}} \times \exp \left( C_{2.12} \left( \frac{3p\lambda_n^2(\Gamma_n^{(2)})^2}{2\Gamma_n^2} + \frac{1}{2p} \right) \right) (I_V^1)^{\frac{1}{4p}} \\
&\leq 2(I_V^1)^{\frac{1}{p}} \exp \left( -\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2} \left( 1 - \frac{p}{q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2\Gamma_n} \left( \frac{6C_{V,\varphi}^2}{c_V} + [2C_{2.11} + 3C_{2.12}](\Gamma_n^{(2)})^2 \right) \right) \right) \\
&\quad \times \exp \left( \frac{1}{p} [c_V + \frac{C_{2.12}}{2}] \right).
\end{aligned} \tag{2.15}$$

Recall now that for  $\theta \geq \frac{1}{2+\beta} > 1/3$ ,  $\Gamma_n^{(2)}/\sqrt{\Gamma_n} \xrightarrow[n]{\rightarrow} 0$  (see Remark 2). We now take  $p := p_n \xrightarrow[n]{\rightarrow} +\infty$ , and therefore  $q := q_n \xrightarrow[n]{\rightarrow} 1$ , s.t.

$$d_n := \frac{p_n}{q_n\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2\Gamma_n} \left( \frac{6C_{V,\varphi}^2}{c_V} + [2C_{2.11} + 3C_{2.12}](\Gamma_n^{(2)})^2 \right) \xrightarrow[n]{\rightarrow} 0, \tag{2.16}$$

and conclude from (2.15) setting  $c_n = q_n^{-1}(1 - d_n)$ ,  $C_n := (I_V^1)^{\frac{1}{p_n}} \exp(\frac{C_{2.12}}{2p_n}) \xrightarrow[n]{\rightarrow} 1$ . Observe that taking an increasing sequence  $(p_n)_{n \geq 1}$  readily yields  $C_n \downarrow 1$ , and  $q_n \downarrow 1$ . Also, the sequence  $(p_n)_{n \geq 1}$  can be chosen in order to have, for  $n$  large enough,  $d_n \downarrow 0$  so that  $c_n \uparrow 1$ .

(ii) Assume  $a \leq \frac{c_V q \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2 \sqrt{\Gamma_n}}{4C_V p \|D^2\varphi\|_\infty^2 \Gamma_n^{(2)}}$ . Plugging in (2.8) the controls from (2.14), Lemmas 4 (with  $j = 2$ ), 5 (equations (2.11), (2.13)) we then derive:

$$\begin{aligned}
\mathbb{P} \left[ \left| \nu_n(\mathcal{A}\varphi) + \frac{\bar{G}_n}{\Gamma_n} \right| \geq \frac{a}{\sqrt{\Gamma_n}} \right] &\leq 2 \exp \left( -\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2} \right) \exp \left( \frac{3pC_{V,\varphi}^2\lambda_n^2}{c_V\Gamma_n^2} + \frac{c_V}{p} \right) (I_V^1)^{\frac{1}{2p}} \\
&\quad \times \exp \left( C_{2.11} \frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{\Gamma_n^2} \right) (I_V^1)^{\frac{1}{4p}} (I_V^1)^{v_n} \\
&\leq 2(I_V^1)^{v_n + \frac{3}{4p}} \exp \left( \frac{c_V}{p} \right) \exp \left( -\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2} \left( 1 - \frac{p}{q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2\Gamma_n} \left( \frac{6C_{V,\varphi}^2}{c_V} + 2C_{2.11}(\Gamma_n^{(2)})^2 \right) \right) \right).
\end{aligned} \tag{2.17}$$

Since  $\theta \geq \frac{1}{2+\beta} > 1/3$  (see Remark 2), we again take  $p := p_n \uparrow +\infty$  so that:

$$d_n := \frac{p_n}{q_n\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2} \left( \frac{6C_{V,\varphi}^2}{c_V\Gamma_n} + \frac{2C_{2.11}(\Gamma_n^{(2)})^2}{\Gamma_n} \right) \xrightarrow[n]{\rightarrow} 0. \tag{2.18}$$

In this case, we derive the result by setting  $c_n := q_n^{-1}(1 - d_n) \xrightarrow[n]{\rightarrow} 1$ ,  $C_n := (I_V^1)^{v_n + \frac{3}{4p_n}} \exp(\frac{c_V}{p_n}) \xrightarrow[n]{\rightarrow} 1$  (see the limits of  $v_n$  following equation (2.13) and (3.10)). Again,  $(p_n)_{n \geq 1}$  can be chosen in order to have the stated monotonicity for  $n$  large enough. Set now  $\chi_n := \frac{c_V \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2 q_n}{4C_V \|D^2\varphi\|_\infty^2 p_n}$  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  $\beta = 1, \theta = 1/3$ . From (2.9), the controls of (2.14) and Lemma 4 (with  $j = 1$ ) we get:

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

Recalling the definition of  $\lambda_n$  in (2.10), we conclude as previously with obvious modifications of  $(c_n)_{n \geq 1}$ ,  $(C_n)_{n \geq 1}$ . Eventually, we derive similarly to the proof of Theorem 10 in [LP02] that  $E_n \xrightarrow[n]{\rightarrow} 0$ .  $\square$

**Proof of Theorem 2.** To derive point (a) it suffices to write:

$$\begin{aligned} \mathbb{P}[\sqrt{\Gamma_n}|\nu_n(\mathcal{A}\varphi)| \geq a] &\leq \mathbb{P}\left[|E_n^1| + |\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) + E_n^1| \geq a\right] \\ &\leq \mathbb{P}\left[a_n + |\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) + E_n^1| \geq a\right] = \mathbb{P}\left[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) + E_n^1| \geq a - a_n\right], \end{aligned}$$

using the bound on  $E_n^1$  following equation (2.3) for the penultimate inequality. Since  $a > a_n$ , the result readily follows from point (a),(i) of Theorem 3 and (1.5), having in mind that the partial derivatives of  $\varphi$  are bounded up to order three.

Point (b) of Theorem 2 is a direct consequence of point (b) of Theorem 3 since we assumed  $\varphi$  to be  $\mathcal{C}^4$ .  $\square$

### 3. PROOF OF THE TECHNICAL LEMMAS

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

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

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

We then perform a Taylor expansion with integral remainder at order 2 for the function  $\varphi$  in the two terms of the r.h.s. of (3.1). Namely, with the above notations:

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

Hence,

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

where

$$(3.3) \quad \begin{aligned} \psi_k(X_{k-1}, U_k) &= \sqrt{\gamma_k} \sigma_{k-1} U_k \cdot \nabla \varphi(X_{k-1} + \gamma_k b_{k-1}) \\ &\quad + \gamma_k \int_0^1 (1-t) \text{Tr}\left(D^2 \varphi(X_{k-1} + \gamma_k b_{k-1} + t\sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* - D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \Sigma_{k-1}\right) dt. \end{aligned}$$

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  $\varphi$  is globally Lipschitz continuous we derive, for all  $(u, u') \in (\mathbb{R}^d)^2$ :

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

The result is obtained by summing up the previous identities from  $k = 1$  to  $n$ , observing, with the notations of (2.5), that  $L_n = \sum_{k=1}^n \varphi(X_k) - \varphi(X_{k-1})$ ,  $D_{2,b,n} = \sum_{k=1}^n D_{2,b}^k$ ,  $D_{2,\Sigma,n} = \sum_{k=1}^n D_{2,\Sigma}^k$ ,  $G_n := \sum_{k=1}^n \psi_k(X_{k-1}, U_k)$ .  $\square$

**Proof of Lemma 2.** The idea is to use conditionally and iteratively the Gaussian concentration property (GC) of the innovation. Let us note that this strategy was already the key ingredient in [FM12]. In the current framework, we exploit that the functions  $u \mapsto \Delta_k(X_{k-1}, u) := \psi_k(X_{k-1}, u) - \mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}]$  are

conditionally independent w.r.t.  $\mathcal{F}_{k-1}$  and Lipschitz continuous with constant  $\sqrt{\gamma_k}\|\sigma\|_\infty\|\nabla\varphi\|_\infty$  by Lemma 1. We thus write:

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

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

Iterating the process over  $k$ , we obtain:

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

Finally,

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

where  $g: \mathbb{R}^+ \rightarrow \mathbb{R}$  is defined by  $g(\lambda) = -\frac{a}{\sqrt{\Gamma_n}}\lambda + \frac{q\lambda^2}{2\Gamma_n} \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2$ . As  $a > 0$ , the function attains its minimum at  $\lambda_n$  given in (2.10). This eventually yields the expected bound.  $\square$

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

$$(3.6) \quad \mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}] = \gamma_k \int_0^1 (1-t) \text{Tr} \left( \mathbb{E} \left[ D^2\varphi(X_{k-1} + \gamma_k b_{k-1} + t\sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* - D^2\varphi(X_{k-1} + \gamma_k b_{k-1}) \Sigma_{k-1} \right] \middle| \mathcal{F}_{k-1} \right) dt.$$

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:

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

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

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

recalling that the third derivatives of  $\varphi$  are  $\beta$ -Hölder continuous for the first inequality. We thus derive:

$$|E_1^n| = \frac{|\bar{G}_n|}{\sqrt{\Gamma_n}} \leq \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^{n-1} |\mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}]| \leq \frac{[\varphi^{(3)}]_\beta \|\sigma\|_\infty^{3+\beta} \mathbb{E}[|U_1|^{3+\beta}] \Gamma_n^{(\frac{3+\beta}{2})}}{(1+\beta)(2+\beta)(3+\beta) \sqrt{\Gamma_n}} =: a_n.$$

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

$$\begin{aligned} V(X_n) - V(X_{n-1}) &= \gamma_n \mathcal{A}V(X_{n-1}) + \gamma_n^2 \int_0^1 (1-t) \text{Tr} \left( D^2V(X_{n-1} + t\gamma_n b_{n-1}) b_{n-1} \otimes b_{n-1} \right) dt \\ &\quad - \frac{\gamma_n}{2} \text{Tr} (D^2V(X_{n-1})) \Sigma_{n-1} + \sqrt{\gamma_n} \sigma_{n-1} U_n \cdot \nabla V(X_{n-1} + \gamma_n b_{n-1}) \\ &\quad + \gamma_n \int_0^1 (1-t) \text{Tr} \left( D^2V(X_{n-1} + \gamma_n b_{n-1} + t\sqrt{\gamma_n} \sigma_{n-1} U_n) \sigma_{n-1} U_n \otimes U_n \sigma_{n-1}^* \Sigma_{n-1} \right) dt \\ &\leq -\gamma_n \alpha_V V(X_{n-1}) + \gamma_n \beta_V + C_V \frac{\gamma_n^2}{2} \|D^2V\|_\infty V(X_{n-1}) \\ &\quad + \frac{\gamma_n}{2} \|D^2V\|_\infty \|\sigma\|_\infty^2 + \sqrt{\gamma_n} \sigma_{n-1} U_n \cdot \nabla V(X_{n-1} + \gamma_n b_{n-1}) + \frac{\gamma_n}{2} \|D^2V\|_\infty \|\sigma\|_\infty^2 |U_n|^2 \\ (3.7) \quad &\leq \gamma_n \left( -\frac{\alpha_V}{2} V(X_{n-1}) + \tilde{c} \right) + \sqrt{\gamma_n} \sigma_{n-1} U_n \cdot \nabla V(X_{n-1} + \gamma_n b_{n-1}) + \frac{\gamma_n}{2} \|D^2V\|_\infty \|\sigma\|_\infty^2 |U_n|^2 \end{aligned}$$

for a constant  $\tilde{c} := \tilde{c}(V, \sigma, \beta_V)$ . We have in fact considered the time steps sufficiently small (in **(S)**), we have chosen for all  $n \in \mathbb{N}$ ,  $\gamma_n < \min(\frac{1}{2\sqrt{C_V \tilde{c}}}, \frac{\alpha_V}{2C_V \|D^2V\|_\infty})$ . 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:

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

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

$$(3.8) \quad I_1(\gamma_n, \lambda, x) \leq \exp \left( \frac{\lambda^2 \gamma_n |\sigma^*(x) \nabla V(x + \gamma_n b(x))|^2}{2} \right) \stackrel{(\mathcal{L}_V)}{\leq} \exp \left( \frac{\lambda^2 \gamma_n C_V \|\sigma\|_\infty^2 V(x + \gamma_n b(x))}{2} \right).$$

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

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

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  $C_{V,\lambda}^0 \in (1, +\infty)$  s.t.

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

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

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

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

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  $\lambda < \lambda_V$ ,

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

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

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

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

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

Hence,  $(\mathcal{P}_n)$  holds. Taking  $c_V < \lambda_V$  completes the proof.  $\square$

**Remark 8.** 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  $\lambda < \lambda_V (v^*)^{1-\xi}$ :

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

Thus, we readily get as a 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 again refer to Lemaire [Lem05] for additional results in that direction.

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

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

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

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

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

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

**Proof of Lemma 5.**



• *Proof of inequalities (2.13) and (2.12).* We first easily get from the assumptions on  $\varphi$  that:

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

recalling point ii) of  $(\mathcal{L}_V)$  for the last inequality. 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:

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

- If  $a \leq \frac{c_V q}{4C_V p} \frac{\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2 \sqrt{\Gamma_n}}{\|D^2\varphi\|_\infty^2 \Gamma_n^{(2)}}$ . In that case we have from (2.10) 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  $k \in \llbracket 1, n \rrbracket$ :

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

Thus, setting

$$(3.10) \quad v_n := \frac{\bar{v}_n}{4p} = \frac{\lambda_n \Gamma_n^{(2)}}{\Gamma_n} \|D^2\varphi\|_\infty \frac{C_V}{c_V},$$

we finally derive,

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

using again the notations of Proposition 1. This gives (2.13).

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

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

This control will allow to equilibrate the growth of  $V$ . Recall indeed from  $(\mathcal{L}_V)$  that for a given threshold  $K > 0$  there are constants  $C_V, \bar{c} := \bar{c}(V, K) > 0$  s.t.  $\forall |x| \geq K, |V(x)| \leq \bar{c}|x|^2, |b(x)| \leq \sqrt{\bar{c}C_V}|x|$  (see also Remark 1). Hence, since  $t \in [0, 1]$ ,

$$|X_k + t\gamma_k b_k| \geq (|X_k| - \gamma_k |b_k|)_+ \geq |X_k| (1 - \gamma_k \sqrt{\bar{c}C_V})_+.$$

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

$$(3.11) \quad |D^2\varphi(X_k + t\gamma_k b_k)| \leq \|D^2\varphi\|_\infty \mathbb{1}_{\{|X_k| \leq K\}} + \frac{2C}{|X_k|} \mathbb{1}_{\{|X_k| > K\}}.$$

From this inequality and applying as above the Jensen inequality, we derive:

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

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

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

We finally derive with the notations of Proposition 1:

$$\begin{aligned} \left( \mathbb{E} \exp \left( \frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}| \right) \right)^{\frac{1}{4p}} &\leq \exp \left( \frac{\lambda_n \Gamma_n^{(2)}}{2\Gamma_n} C_V \left( \sup_{|x| \leq K} V(x) \right) \|D^2 \varphi\|_\infty \right) \exp \left( \frac{p\lambda_n^2 (\Gamma_n^{(2)})^2 (C C_V)^2 \bar{c}}{\Gamma_n^2 c_V} \right) (I_V^1)^{\frac{1}{4p}} \\ &\leq \exp \left( C_{2.12} \left( \frac{p\lambda_n^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2} + \frac{\lambda_n \Gamma_n^{(2)}}{\Gamma_n} \right) \right) (I_V^1)^{\frac{1}{4p}}, \end{aligned}$$

setting  $C_{2.12} := \frac{C_V}{2} \left( \sup_{|x| \leq K} V(x) \right) \|D^2 \varphi\|_\infty \vee \frac{(C C_V)^2 \bar{c}}{c_V}$ . From the Young inequality we finally have  $\frac{\lambda_n \Gamma_n^{(2)}}{\Gamma_n} \leq \frac{p\lambda_n^2 (\Gamma_n^{(2)})^2}{2\Gamma_n^2} + \frac{1}{2p}$  which, plugged in the above control, completes the proof of (2.12).

- *Proof of inequality (2.11).* We proceed as for the proof of (2.13) and (2.12). Write:

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

Using once again the Young inequality and setting  $C_{2.11} := \frac{\|\sigma\|_\infty^4 [\varphi^{(2)}]_1^2 C_V}{4 c_V}$ , we obtain:

$$\left( \mathbb{E} \exp \left( \frac{4p\lambda_n}{\Gamma_n} |D_{2,\Sigma,n}| \right) \right)^{\frac{1}{4p}} \leq \exp \left( \frac{p\lambda_n^2}{4} \left( \frac{\Gamma_n^{(2)}}{\Gamma_n} \right)^2 \|\sigma\|_\infty^4 [\varphi^{(2)}]_1^2 \frac{C_V}{c_V} \right) (I_V^1)^{\frac{1}{4p}} \leq \exp \left( C_{2.11} p \lambda_n^2 \left( \frac{\Gamma_n^{(2)}}{\Gamma_n} \right)^2 \right) (I_V^1)^{\frac{1}{4p}}. \quad \square$$

#### 4. A REFINEMENT WHEN $\|\sigma\|^2 - \nu(\|\sigma\|^2)$ IS A COBOUNDARY

We will assume in this Section that there exists a solution  $\vartheta$  of the Poisson problem  $\mathcal{A}\vartheta = \|\sigma\|^2 - \nu(\|\sigma\|^2)$  satisfying the assumptions stated in Theorems 2 and 3, which is in particular the case if  $\sigma$  is smooth enough and the coefficients in (1.1) satisfy some non-degeneracy conditions, see e.g. Lorenzi and Bertoldi [LB06]. In a non-degenerate setting, those assumptions should hold as well provided  $\sigma \in C^{1+\beta}(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^r)$ ,  $\beta \in (0, 1]$  (see also Section 6.2 and Appendix A).

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

**Theorem 4.** *Under the Assumptions of Theorem 3 and with the notations introduced therein, we have that:*

- (a) *For  $(\beta \neq 1 \text{ and } \theta \in [\frac{1}{2+\beta}, 1])$  or  $(\beta = 1 \text{ and } \theta \in (\frac{1}{3}, 1])$ , there exist explicit non-negative sequences  $(\tilde{c}_n)_{n \geq 1}$  and  $(\tilde{C}_n)_{n \geq 1}$ , respectively increasing and decreasing for  $n$  large enough, with  $\lim_n \tilde{C}_n = \lim_n \tilde{c}_n = 1$  and an auxiliary sequence  $(\bar{c}_n)_{n \geq 1}$ , decreasing for  $n$  large enough, with  $\lim_n \bar{c}_n = \bar{c}_\infty > 0$  s.t. for all  $n \geq 1$  for all  $a > 0$ :*

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

where  $x_+ = \max(x, 0)$ .

(b) For  $\beta = 1, \theta = \frac{1}{3}$ , there exist explicit non-negative sequences  $(\tilde{c}_n)_{n \geq 1}$  and  $(\tilde{C}_n)_{n \geq 1}$ , respectively increasing and decreasing for  $n$  large enough, with  $\lim_n \tilde{C}_n = \lim_n \tilde{c}_n = 1$  and an auxiliary sequence  $(\bar{c}_n)_{n \geq 1}$ , decreasing for  $n$  large enough, with  $\lim_n \bar{c}_n = \bar{c}_\infty > 0$  s.t. for all  $n \geq 1$  for all  $a > 0$ :

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

**Remark 9** (About deviation rates). Observe that in order to derive **global** deviation bounds (valid for every  $a > 0$ ) two concentration regimes appear in the previous bounds. For an arbitrary fixed  $a > 0$ , we have that for  $n$  large enough (depending on  $a$ ), the Gaussian concentration regime will give the fastest decay, since  $\frac{2}{1 + \sqrt{1 + 4\tilde{c}_n^3 \frac{\Gamma_n}{a^2}}} \rightarrow 0$ . Also, when  $a \asymp \sqrt{\Gamma_n}$  the two above contributions give a Gaussian bound, with suboptimal constants. Eventually, when  $a \gg \sqrt{\Gamma_n}$ , for a fixed  $n$ , we have that the first term is “stuck” at the threshold  $\Gamma_n$  whatever level  $a$  is considered, i.e.  $a^2 \left( 1 - \frac{2}{1 + \sqrt{1 + 4\tilde{c}_n^3 \frac{\Gamma_n}{a^2}}} \right) \xrightarrow{a \rightarrow \infty} \tilde{c}_n^3 \Gamma_n$  whereas the second clearly becomes bigger.

To summarize, when the Gaussian regime prevails (i.e. when  $\frac{a}{\sqrt{\Gamma_n}}$  is small), the results of Theorems 2, 3 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 Theorems 2, 3 by  $\nu(\|\sigma\|^2)$ . However, our martingale approach intrinsically 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  $\beta \in (0, 1), \theta \in [1/(2+\beta), 1]$ . The second setting  $\beta = 1, \theta \in (\frac{1}{3}, 1]$  can be derived from (1.5) as in the proof of Theorem 3. Case (b) could be derived similarly following the proof of Theorem 3. We restart from the computations of Section 2.2 that give for all  $\lambda > 0$  the control in equation (2.8). Let us now focus on the term giving the concentration and write for all  $\rho > 1$ :

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

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

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

Set  $\tilde{\mathcal{F}}_1 := \exp \left( - \frac{\rho^2 (q\lambda)^2 [\varphi]_1^2 \nu(\|\sigma\|^2)}{2\Gamma_n^2} \right) \mathcal{F}_1$ . Define now, for a given  $n \in \mathbb{N}$  and  $m \in \mathbb{N}_0$ ,  $S_m := \exp \left( \rho \frac{q\lambda}{\Gamma_n} M_m - \frac{\rho^2 (q\lambda)^2 [\varphi]_1^2}{2\Gamma_n^2} \sum_{k=1}^m \gamma_k \|\sigma(X_{k-1})\|^2 \right)$ . From the definition of the martingale  $M_n$  in (2.6) and the controls of the Lipschitz constants of the functions  $(\Psi_k(X_{k-1}, \cdot))_{k \in [1, n]}$  in Lemma 1, we get by iterated conditioning:

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

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

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

For the term  $C_2$ , we have that setting  $\mu := \mu(q, n, \rho, \lambda) = \frac{(q\lambda)^2 \rho^2 [\varphi]_1^2}{2(\rho-1)\Gamma_n}$ ,

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

so that this contribution can be controlled from the previous expansion of Lemma 1 exploiting the technical Lemmas of Section 2.2 replacing  $\lambda$  by  $\mu$  and  $\varphi$  by  $\vartheta$ .

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 (2.8),

$$\begin{aligned} \mathcal{F}_2 &= \mathbb{E} \exp \left( \frac{\mu}{\Gamma_n} \sum_{k=1}^n \gamma_k \mathcal{A}^\vartheta(X_{k-1}) \right) \leq \left( \mathbb{E} \exp \left( \frac{\bar{q}\mu}{\Gamma_n} M_n^\vartheta \right) \right)^{\frac{1}{\bar{q}}} \\ (4.3) \quad &\quad \times \left( \mathbb{E} \exp \left( \frac{2\bar{p}\mu}{\Gamma_n} |L_n^\vartheta| \right) \right)^{\frac{1}{2\bar{p}}} \left( \mathbb{E} \exp \left( \frac{4\bar{p}\mu}{\Gamma_n} |D_{2,b,n}^\vartheta| \right) \right)^{\frac{1}{4\bar{p}}} \left( \mathbb{E} \exp \left( \frac{4\bar{p}\mu}{\Gamma_n} |D_{2,\Sigma,n}^\vartheta| \right) \right)^{\frac{1}{4\bar{p}}} \end{aligned}$$

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

Still for simplicity, we assume as well (case (i)) that there exists  $C_\vartheta > 0$  s.t. for all  $x \in \mathbb{R}^d$ ,  $|D^2\vartheta(x)| \leq C_\vartheta/(1+|x|)$ . Plugging in (4.3) the controls established in Lemma 4 (with  $j=2$ ), Lemma 5 (equations (2.11) and (2.12)) and (3.5), then replacing  $\lambda_n$  by  $\mu$ , we get:

$$\begin{aligned} \mathcal{F}_2 &\leq \exp \left( \frac{\bar{q}\mu^2 \|\sigma\|_\infty^2 [\vartheta]_1^2}{2\Gamma_n} \right) \exp \left( \frac{3\bar{p}C_{V,\vartheta}^2 \mu^2}{c_V \Gamma_n^2} + \frac{c_V}{\bar{p}} \right) (I_V^1)^{\frac{1}{2\bar{p}}} \\ &\quad \times \exp \left( C_{2.11} \frac{\bar{p}\mu^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2} \right) (I_V^1)^{\frac{1}{4\bar{p}}} \times \exp \left( C_{2.12} \left( \frac{3\bar{p}\mu^2 (\Gamma_n^{(2)})^2}{2\Gamma_n^2} + \frac{1}{2\bar{p}} \right) \right) (I_V^1)^{\frac{1}{4\bar{p}}}. \\ &\leq \exp \left( \frac{\mu^2}{\Gamma_n} \left( \frac{\bar{q}\|\sigma\|_\infty^2 [\vartheta]_1^2}{2} + \bar{p} \left( \frac{(\Gamma_n^{(2)})^2}{\Gamma_n} [C_{2.11} + \frac{3}{2}C_{2.12}] + \frac{3C_{V,\vartheta}^2}{c_V \Gamma_n} \right) \right) \right) \exp \left( \frac{1}{\bar{p}} (c_V + \frac{C_{2.12}}{2}) \right) (I_V^1)^{\frac{1}{\bar{p}}}. \end{aligned}$$

Set now

$$\begin{aligned} \bar{C}_n &:= \exp \left( \frac{1}{\bar{p}} (c_V + \frac{C_{2.12}}{2}) \right) (I_V^1)^{\frac{1}{\bar{p}}}, \\ (4.4) \quad \bar{e}_n &:= \bar{p} \left( \frac{(\Gamma_n^{(2)})^2}{\Gamma_n} [C_{2.11} + \frac{3}{2}C_{2.12}] + \frac{3C_{V,\vartheta}^2}{c_V \Gamma_n} \right). \end{aligned}$$

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

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

$$\left( \mathbb{E} \exp \left( \frac{\lambda q}{\Gamma_n} M_n \right) \right)^{\frac{1}{q}} \leq \left( T_1^{\frac{1}{\rho}} T_2^{1-\frac{1}{\rho}} \right)^{\frac{1}{q}} \leq \exp \left( \frac{\rho q \lambda^2 [\varphi]_1^2 \nu(\|\sigma\|^2)}{2\Gamma_n} \right) \bar{C}_n^{\frac{\rho-1}{\rho q}} \exp \left( \frac{\rho-1}{\rho q} \frac{\mu^2}{\Gamma_n} \left( \frac{\bar{q}\|\sigma\|_\infty^2 [\vartheta]_1^2}{2} + \bar{e}_n \right) \right).$$

Plugging this bound in (2.8), using again the controls of Lemmas 4 and 5, eventually yields:

$$\begin{aligned} \mathbb{P} \left[ \left| \sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) + E_n^1 \right| \geq a \right] &\leq 2 \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \right) \exp \left( \frac{\rho q \lambda^2 [\varphi]_1^2 \nu(\|\sigma\|^2)}{2\Gamma_n} \right) \bar{C}_n^{\frac{\rho-1}{\rho q}} \exp \left( \frac{\rho-1}{\rho q} \frac{\mu^2}{\Gamma_n} \left( \frac{\bar{q}\|\sigma\|_\infty^2 [\vartheta]_1^2}{2} + \bar{e}_n \right) \right) \\ &\quad \times \exp \left( \frac{\lambda^2}{\Gamma_n} p \left( \frac{(\Gamma_n^{(2)})^2}{\Gamma_n} [C_{2.11} + \frac{11}{4}C_{2.12}] + \frac{3C_{V,\varphi}^2}{c_V \Gamma_n} \right) \right) \exp \left( \frac{1}{p} (c_V + \frac{C_{2.12}}{2}) \right) (I_V^1)^{\frac{1}{p}}. \end{aligned}$$

Choosing  $p := p_n \rightarrow +\infty$  and s.t.  $p_n \frac{(\Gamma_n^{(2)})^2}{\Gamma_n} \rightarrow 0$ , we get with the notations introduced in the proof of Theorem 3:

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

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

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

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

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

Note that both sequences  $(\tilde{A}_n)_{n \geq 1}$  and  $(\tilde{B}_n)_{n \geq 1}$  are bounded and bounded away from zero sequences (and do not depend on  $\rho$ ). 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

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

which is the canonical form of this third degree equation to apply the Cardan-Tartaglia formula <sup>(2)</sup> so that

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

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

$$(4.7) \quad \left( x_n + \sqrt{a_n^3 + x_n^2} \right)^{\frac{1}{3}} + \left( x_n - \sqrt{a_n^3 + x_n^2} \right)^{\frac{1}{3}} \sim \begin{cases} \frac{2}{3} \frac{x_n}{a_n} & \text{if } x_n = o(a_n^{\frac{3}{2}}) \text{ (then } x_n \rightarrow 0), \\ (2x_n)^{\frac{1}{3}} & \text{if } a_n = o(x_n^{\frac{2}{3}}) \text{ (then } x_n \rightarrow +\infty). \end{cases}$$

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

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

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

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

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

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

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  $z_* = \left( \frac{1}{2}(-q + \sqrt{\frac{\Delta}{27}}) \right)^{\frac{1}{3}} + \left( \frac{1}{2}(-q - \sqrt{\frac{\Delta}{27}}) \right)^{\frac{1}{3}}$ .

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

$$(4.8) \quad P(\lambda^*) = -\frac{a^2}{4\tilde{A}_n} \left( 1 - \frac{2}{1 + \sqrt{1 + 4\frac{\tilde{A}_n^3 \Gamma_n}{\tilde{B}_n a^2}}} \right).$$

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  $x_n = o(a_n^{\frac{3}{2}})$  in (4.7) is satisfied *a posteriori*.

- If  $\frac{a}{B_n \sqrt{\Gamma_n}} \rightarrow +\infty$ , then, still owing to (4.7),

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

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

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

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

$$(4.9) \quad P(\bar{\lambda}^*) \leq -\frac{a^{\frac{4}{3}}}{4} \left( \frac{\Gamma_n}{\tilde{B}_n} \right)^{\frac{1}{3}} \left( 1 - \frac{2\tilde{A}_n}{3\tilde{B}_n} \left( \frac{\Gamma_n}{a^2} \right)^{\frac{1}{3}} \right).$$

From (4.9), (4.8) and (4.5), we conclude the proof of case (a) by setting  $\bar{c}_n := \tilde{A}_n \tilde{B}_n^{-\frac{1}{3}}$ .

In the biased case, the result follows similarly from the corresponding analysis performed in Section 2 taking  $A_n(\rho) = \frac{\rho q[\varphi]_{1\nu}^2 (\|\sigma\|^2)}{2}$ .  $\square$

**Remark 10.** • When  $a \asymp \sqrt{\Gamma_n}$ , one checks that  $\lambda_{\min} \asymp \Gamma_n$  and  $P(\lambda_{\min}) \asymp -\Gamma_n$ . 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  $\lambda_{\min}$  into the definition of  $P$ , yields

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

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

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

## 5. TRACTABLE NON-ASYMPTOTIC CONFIDENCE INTERVALS.

We give in this section, as an application of our main Theorems 2 and 3, some *computable* non-asymptotic confidence intervals.

**5.1. First Non-asymptotic Confidence Intervals and Approximation Issues.** To begin with, we first give, as a direct corollary of Theorem 3, a non-asymptotic confidence interval result which does not require numerical approximations of the bias. Anyhow, it does not benefit from the fastest convergence rate which induces a bias.

**Theorem 5** (Non-asymptotic confidence intervals without bias). *Let  $f$  be a Lipschitz continuous function in  $\mathcal{C}^1(\mathbb{R}^d, \mathbb{R})$  s.t.  $[f^{(1)}]_\beta < +\infty$  for some  $\beta \in (0, 1)$ . Assume that the Poisson Equation:*

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

*admits a unique solution  $\varphi$ , supposed to be globally Lipschitz continuous and in  $\mathcal{C}^3(\mathbb{R}^d, \mathbb{R})$ . Moreover, assume that  $(D^\alpha \varphi)_{|\alpha| \in \{2, 3\}}$  are bounded and that there exists a positive constant  $C_\varphi$  s.t., for all  $x \in \mathbb{R}^d$ ,  $|D^2 \varphi(x)| \leq C_\varphi/(1 + |x|)$  and  $[\varphi^{(3)}]_\beta < +\infty$ . We suppose finally that  $\varphi$  satisfies  $(G_V)$ .*

*Assume the step sequence  $(\gamma_k)_{k \geq 1}$  is s.t.  $\gamma_k \asymp k^{-\theta}$ ,  $\theta \in [\frac{1}{2+\beta}, 1]$ . Then, for  $(c_n)_{n \geq 1}, (C_n)_{n \geq 1}$  like in Theorem 3 with  $\lim_n c_n = \lim_n C_n = 1$ , we have that for all  $n \geq 1$  and  $a > \bar{a}_n := \frac{a_n}{\|\sigma\|_\infty \|\nabla \varphi\|_\infty}$  with  $a_n$  given by (2.3):*

$$\mathbb{P} \left[ \nu(f) \in \left[ \nu_n(f) - \frac{a \|\sigma\|_\infty \|\nabla \varphi\|_\infty}{\sqrt{\Gamma_n}}, \nu_n(f) + \frac{a \|\sigma\|_\infty \|\nabla \varphi\|_\infty}{\sqrt{\Gamma_n}} \right] \right] \geq 1 - 2C_n \exp \left( -c_n \frac{(a - \bar{a}_n)^2}{2} \right).$$

*Proof.* Setting  $a_{\sigma, \varphi} := a \|\sigma\|_\infty \|\nabla \varphi\|_\infty$ , it suffices to write:

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

Hence:

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

Conclude by noting that Theorem 3 (a) yields:

$$\begin{aligned} \mathbb{P}[\sqrt{\Gamma_n} |\nu_n(\mathcal{A}\varphi)| \geq a_{\sigma, \varphi}] &\leq \mathbb{P}[|E_n^1| + |\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) + E_n^1| \geq a_{\sigma, \varphi}] \leq \mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) + E_n^1| \geq a_{\sigma, \varphi} - a_n] \\ &\leq 2C_n \exp \left( -c_n \frac{(a - \bar{a}_n)^2}{2} \right). \quad \square \end{aligned}$$

The previous result is simple to use and has the major advantage that no *a priori* explicit knowledge of the solution  $\varphi$  of the associated Poisson Equation is required (but a bound on the quantity  $\|\nabla \varphi\|_\infty$  which can usually be obtained through the Lipschitz constant  $[f]_1$  of the source, see e.g. Section 6.2 and Appendix A).

Note however that, when  $\varphi$  is explicitly known, the previous result does not allow to attain the fastest admissible concentration rate in Theorems 2 and 3 corresponding to  $\theta = \frac{1}{3}, \beta = 1$ . To this end, we need to have some *tractable* approximation of the bias  $E_n^1, E_n$  appearing in Theorems 2 and 3. Observe that with the notations of the indicated theorems:

$$\begin{aligned} E_n^1 &= \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^{3/2} \left[ \tilde{\mathbb{E}}[(1 - T_1) T_1 \Lambda_{k-1}(T_1, T_2, x)] \right]_{|x=X_{k-1}}, \\ E_n - \tilde{\gamma} m &= E_n^1 + \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^2 \left[ \tilde{\mathbb{E}} \left[ (1 - T_1) \text{Tr} \left( D^2 \varphi(x + T_1 \gamma_k b(x)) b(x) \otimes b(x) \right) \right] \right]_{|x=X_{k-1}} \\ (5.1) \quad &+ \frac{1}{2\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k \text{Tr} \left( (D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2 \varphi(X_{k-1})) \Sigma_{k-1} \right), \end{aligned}$$

where  $T_1, T_2$  are two independent uniform random variables on  $[0, 1]$  defined on some auxiliary probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ .

In practice, even when  $\varphi$  is explicitly known, the approximation of the above quantities, is a notable issue. Indeed, they directly appear in the non-asymptotic confidence intervals that would readily follow from Theorem 3. The derivation of a computable confidence interval therefore requires a quick and well controlled estimation of the above expectations w.r.t.  $\tilde{\mathbb{P}}$  along the discretization scheme. Since we have assumed in Theorem 3 some regularity on the function  $\varphi$ , we now choose to approximate the bias by *quantization* (which takes advantage of the regularity).

**5.2. Numerical Approximation of the Bias: a Quantization Approach.** The basic idea of (quadratic) optimal quantization consists in approximating (in  $L^2$ ) a given random variable by a discrete one taking finitely many values. For a square integrable random vector  $Z : (\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{\mathbb{P}}) \rightarrow \mathbb{R}^r$  and a fixed integer  $M \in \mathbb{N}$ , a quadratic optimal quantization  $Z^M$  of  $Z$  is a discrete random variable defined as  $Z^M = \pi_{\Upsilon^M}(Z)$  where  $\pi_{\Upsilon^M} : \mathbb{R}^r \rightarrow \mathbb{R}^r$  is a Borel nearest neighbor projection on an *optimal quantizer* (or *quantization grid*)  $\Upsilon^M \subset \mathbb{R}^r$  of size (at most)  $M$ , which minimizes the quadratic distortion:

$$D_2(\Upsilon) := \tilde{\mathbb{E}}[(Z - \pi_{\Upsilon}(Z))^2] \text{ over all grids } \Upsilon \subset \mathbb{R}^r \text{ of size at most } M.$$

Such an optimal grid always exists and can be computed by stochastic optimization techniques (see [PP05]). Furthermore  $\Upsilon^M$  satisfies a *self-consistency* or *stationary* property, namely  $\tilde{\mathbb{E}}(Z | Z^M) = Z^M$ , and a sharp rate of convergence to zero of  $\|Z - \pi_{\Upsilon^M}(Z)\|_2 \sim c_Z \cdot M^{-\frac{1}{r}}$  as  $M \rightarrow +\infty$  holds (Zador's Theorem). We refer to the monograph of Graf and Luschgy [GL00] for a comprehensive presentation. In the framework of numerical probability, optimal quantization has been widely used in various applicative fields since it provides a weighted *cubature formula* of the form  $\mathbb{E}g(Z^M) = \sum_{\xi \in \Upsilon^*} \mathbb{P}(\pi_{\Upsilon^M}(Z) = \xi)g(\xi)$  as an approximation of  $\mathbb{E}f(Z)$ . Let us also mention, among others, the works of Pagès [Pag97] for numerical integration, of Bally *et al.* [BPP05], Pagès and Printems [PP05] for option pricing or Delarue and Menozzi [DM08] for the numerical approximation of coupled forward-backward SDEs.

Here, we choose to quantize separately the random variables  $T_1, T_2$  appearing in (5.1). Indeed, as  $T_1$  is the scalar uniform law over the unit interval, the *optimal* quadratic quantizer is explicitly known since for  $M \in \mathbb{N}$ ,  $\Upsilon^M = \{\frac{2i-1}{2M}, i = 1 : M\}$  (midpoint quantizers *at level*  $M$ ), so that

$$(5.2) \quad T_1^M = \frac{2\lfloor MT_1 \rfloor + 1}{2M} \text{ so that } T_1^M \stackrel{(\text{law})}{=} \frac{1}{M} \sum_{k=1}^M \delta_{\frac{2k-1}{2M}}.$$

In particular, it satisfies the stationarity property  $\tilde{\mathbb{E}}[T_1 | T_1^M] = T_1^M$ . It is also readily seen that

$$(5.3) \quad \tilde{\mathbb{E}}[|T_1 - T_1^M|^2] = \frac{1}{12}M^{-2},$$

that is the most elementary case of the general Zador theorem. Observe that, in the considered case, the cubature formula of interest is but the mid-point quadrature formula for  $M$  points. We will only take advantage of its synthetic formulation.

We can now define the *minimal computable* counterparts of the expressions appearing in (5.1). Namely, we set for  $M \in \mathbb{N}$ :

$$(5.4) \quad \begin{aligned} E_n^{1,M} &= \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^{3/2} \left[ \tilde{\mathbb{E}}[(1 - T_1^M)T_1^M \Lambda_{k-1}(T_1^M, T_2^M, x)] \right]_{|x=X_{k-1}}, \\ E_n^M - \tilde{\gamma}m &= E_n^{1,M} + \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^2 \left[ \tilde{\mathbb{E}}[(1 - T_1^M)\text{Tr}(D^2\varphi(x + T_1^M \gamma_k b(x))b(x) \otimes b(x))] \right]_{|x=X_{k-1}} \\ &+ \frac{1}{2\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k \text{Tr} \left( (D^2\varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2\varphi(X_{k-1}))\Sigma_{k-1} \right) =: E_n^{1,M} + E_n^{2,M} \end{aligned}$$

where  $T_1^M, T_2^M$  are independent random variables whose laws are given in (5.2).

Observe that  $E_n^M - \tilde{\gamma}m$  can be explicitly computed provided  $\varphi$  is known and the innovations  $(U_k)_{k \geq 1}$  are discrete (typically a Bernoulli law). This last point, permits to compute for all  $k \in \llbracket 1, n \rrbracket$  and all  $(t, u, x) \in [0, 1]^2 \times \mathbb{R}^d$ ,  $\Lambda_{k-1}(t, u, x)$  directly as a finite sum. When the innovations have absolutely continuous distributions<sup>(3)</sup>, typically normal distributions, another approximation procedure is needed to evaluate  $\Lambda_{k-1}$ . We again choose to proceed by optimal quantization, assuming that an optimal  $M$ -quantizer, denoted  $\Upsilon^M$ , of the innovation  $U$  is available, as well as its companion weights  $p^M(\xi) = \mathbb{P}(\pi_{\Upsilon^M}(U) = \xi)$ ,  $\xi \in \Upsilon^M$  (have in mind that  $U^M = \pi_{\Upsilon^M}(U)$ ). Thus, optimal quantizers (an their companion weight vectors) of multivariate normal distributions can be downloaded for various sizes  $M$  and dimensions  $r$  at the website

www.quantize.maths-fi.com

3. or more generally have an absolutely continuous component. We restrict here to absolutely continuous laws for simplicity.



In particular, it follows from Zador's Theorem that there exists  $C_U > 0$  s.t. for all  $M \in \mathbb{N}$ :

$$(5.5) \quad \mathbb{E}[|U - U^M|^2] \leq C_U M^{-2/r} \quad \text{and} \quad \mathbb{E}[U | U^M] = U^M.$$

Thus, introducing for all  $k \in \llbracket 1, n \rrbracket$  and all  $(t, u, x) \in [0, 1]^2 \times \mathbb{R}^d$ ,

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

leads to consider for absolutely continuous innovations:

$$(5.6) \quad \begin{aligned} E_n^{1,M,C} &= \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^{3/2} \left[ \tilde{\mathbb{E}}[(1 - T_1^M) T_1^M \Lambda_{k-1}^M(T_1^M, T_2^M, x)] \right]_{|x=X_{k-1}}, \\ E_n^{M,C} - \tilde{\gamma}m &= E_n^{1,M,C} + E_n^{2,M}. \end{aligned}$$

Note in particular from the definitions in (5.4) and (5.6) that the computation of  $E_n^M - \tilde{\gamma}m$ ,  $E_n^{M,C} - \tilde{\gamma}m$  does not require to know explicitly neither  $\tilde{\gamma}$  nor  $m$ . For notational convenience we introduce:

$$(5.7) \quad E_n^{1,M,U} := \begin{cases} E_n^{1,M}, & \text{for discrete innovations } (U_k)_{k \geq 1}, \\ E_n^{1,M,C}, & \text{for absolutely continuous innovations } (U_k)_{k \geq 1}. \end{cases}$$

We set correspondingly  $E_n^{M,U} - \tilde{\gamma}m = E_n^{1,M,U} + E_n^{2,M}$ .

We now state a useful Lemma that controls the differences  $E_n^{1,M,U} - E_n^1$ ,  $E_n^{M,U} - E_n$ .

**Lemma 6** (Quantization Error). *There exists a constant  $\bar{C}_1 > 0$  s.t. for all  $n \geq 1, M \in \mathbb{N}$ :*

$$(5.8) \quad |E_n^{1,M,U} - E_n^1| \leq \bar{C}_1 \Psi_U(M) a_n =: a_n^{1,M,U}, \quad \text{a.s.}$$

with  $a_n$  as in (2.3) and  $\psi_U(M) := \begin{cases} M^{-\beta}, & \text{for discrete } (U_i)_{i \geq 1}, \\ M^{-\frac{\beta}{r}}, & \text{for absolutely continuous } (U_i)_{i \geq 1}. \end{cases}$

Assume as well that there exists  $C > 0$  s.t. for  $j \in \{2, 3\}, x \in \mathbb{R}^d$ ,  $|D^j \varphi(x)| \leq \frac{C}{1+|x|^{j-1}}$ . Then, there exist  $\bar{c}_2, \bar{C}_2 > 0$  s.t. for all  $n \geq 1, M \in \mathbb{N}, a \geq a_n^{M,U} := a_n^{1,M,U} + \bar{C}_2 M^{-1} \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}}$ :

$$(5.9) \quad \mathbb{P}[|E_n^{M,U} - E_n| \geq a] \leq \exp \left( -\bar{c}_2 M^2 \frac{\Gamma_n}{(\Gamma_n^{(2)})^2} (a - a_n^{M,U})^2 \right) I_V^1.$$

Recall that we have considered steps  $(\gamma_k)_{k \geq 1}$  s.t.  $\gamma_k \asymp k^{-\theta}$ ,  $\theta \in [1/3, 1]$ . Observe now that for  $\theta > 1/3$ , the quantity  $\frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}} \rightarrow_n +\infty$ . Also, if  $\theta \geq \frac{1}{2+\beta}$ ,  $\beta \in (0, 1)$ , we have  $a_n^{M,U} \rightarrow_M 0$  (we even have from (2.3) that  $a_n^{M,U} \rightarrow_n 0$  uniformly in  $M$  if  $\theta > \frac{1}{2+\beta}$ ). This will allow to improve the validity range of our non-asymptotic confidence intervals for  $\theta \in (1/3, 1]$  w.r.t. the result without bias of Theorem 5. In that case, the above controls in some sense emphasize that the quantization error is *negligible* w.r.t. to the concentration rates appearing in Theorem 3.

On the other hand, for  $\theta = 1/3$ , which implies  $\beta = 1$ ,  $\frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}} \rightarrow_n c > 0$ , the bound in (5.9) will be uniformly controlled in  $n$ , but depends explicitly on  $M$ , i.e. there exist positive constants  $\bar{C}, \bar{c}$ , s.t.  $\forall a \geq a_n^{M,U} := \bar{C} \Psi_U(M)$ :

$$\mathbb{P}[|E_n^{M,U} - E_n| \geq a] \leq \exp \left( -\bar{c} M^2 (a - a_n^{M,U})^2 \right) I_V^1.$$

Again  $a_n^{M,U} \rightarrow_M 0$ , but the global concentration rate will, even asymptotically in  $n$ , involve  $M$  as well.

Those facts are thoroughly quantified in the next subsection. For the sake of clarity, we postpone the proof of Lemma 6 to Section 5.4.

### 5.3. Associated Non-asymptotic Confidence Intervals.

**Theorem 6** (Non-asymptotic confidence intervals). *Let  $f$  be a Lipschitz continuous function in  $\mathcal{C}^1(\mathbb{R}^d, \mathbb{R})$  s.t.  $[f^{(1)}]_\beta < +\infty$  for some  $\beta \in (0, 1]$ . Assume that the Poisson Equation:*

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

*admits a unique solution  $\varphi$ , supposed to be globally Lipschitz continuous and in  $\mathcal{C}^3(\mathbb{R}^d, \mathbb{R})$ . Moreover, assume that  $(D^\alpha \varphi)_{|\alpha| \in \{2, 3\}}$  are bounded and that there exists a positive constant  $C_\varphi$  s.t. for all  $x \in \mathbb{R}^d$ ,  $|D^2 \varphi(x)| \leq C_\varphi / (1 + |x|)$  and  $[\varphi^{(3)}]_\beta < +\infty$ . We suppose finally that  $\varphi$  satisfies  $(G_V)$ .*

We use the same dichotomy as in Theorem 3.

- In case (a), for fixed  $n \geq 1, M \geq 2$  and  $a > \frac{a_n^{1,M,U}}{\|\sigma\|_\infty \|\nabla\varphi\|_\infty} =: \bar{a}_n^{1,M,U}$  with  $a_n^{1,M,U}$  as in Lemma 6:

$$\begin{aligned} \mathbb{P} \left[ \nu(f) \in \left[ \nu_n(f) + \frac{E_n^{1,M,U} - a\|\sigma\|_\infty \|\nabla\varphi\|_\infty}{\sqrt{\Gamma_n}}, \nu_n(f) + \frac{E_n^{1,M,U} + a\|\sigma\|_\infty \|\nabla\varphi\|_\infty}{\sqrt{\Gamma_n}} \right] \right] \\ \geq 1 - 2C_n \exp \left( -c_n \frac{(a - \bar{a}_n^{1,M,U})^2}{2} \right). \end{aligned}$$

- In case (b), assuming as well that for all  $x \in \mathbb{R}^d$ ,  $|D^3\varphi(x)| \leq \frac{C_\varphi}{1+|x|^2}$ , we obtain for  $n \geq 1, M \geq 2, (\alpha_1, \alpha_2) \in (0, 1)^2, \alpha_1 + \alpha_2 = 1, a > \frac{a_n^{M,U}}{\|\sigma\|_\infty \|\nabla\varphi\|_\infty}$  and  $a_n^{M,U}$  defined in Lemma 6:

$$\begin{aligned} \mathbb{P} \left[ \nu(f) \in \left[ \nu_n(f) + \frac{E_n^{1,M,U} + E_n^{2,M} - a\|\sigma\|_\infty \|\nabla\varphi\|_\infty}{\sqrt{\Gamma_n}}, \nu_n(f) + \frac{E_n^{1,M,U} + E_n^{2,M} + a\|\sigma\|_\infty \|\nabla\varphi\|_\infty}{\sqrt{\Gamma_n}} \right] \right] \\ \geq 1 - \left( 2C_n \exp \left( -c_n \frac{(a\alpha_1)^2}{2} \right) + \exp \left( -\bar{c}_2 M^2 \frac{\Gamma_n}{(\Gamma_n^{(2)})^2} (a\|\sigma\|_\infty \|\nabla\varphi\|_\infty \alpha_2 - a_n^{M,U})^2 \right) \right). \end{aligned}$$

*Proof.* Part (a) can be derived from Theorem 3 similarly to Theorem 5 using equation (5.8) from Lemma 6.

For part (b), recalling that  $E_n^{1,M,U} + E_n^{2,M} = E_n^{M,U} - \tilde{\gamma}m$ , we get that the quantity to control also writes:

$$\begin{aligned} \mathbb{P} \left[ \nu(f) \in \left[ \nu_n(f) + \frac{E_n^{M,U} - \tilde{\gamma}m - a\|\sigma\|_\infty \|\nabla\varphi\|_\infty}{\sqrt{\Gamma_n}}, \nu_n(f) + \frac{E_n^{M,U} - \tilde{\gamma}m + a\|\sigma\|_\infty \|\nabla\varphi\|_\infty}{\sqrt{\Gamma_n}} \right] \right] \\ = 1 - \mathbb{P} \left[ \left| \sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) + E_n^{M,U} - \tilde{\gamma}m \right| \geq a_{\sigma,\varphi} \right] \geq 1 - \mathbb{P} \left[ \left| \sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) + E_n - \tilde{\gamma}m \right| + \left| E_n^{M,U} - E_n \right| \geq a_{\sigma,\varphi} \right] \\ \geq 1 - (\mathbb{P} \left[ \left| \sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) + E_n - \tilde{\gamma}m \right| \geq \alpha_1 a_{\sigma,\varphi} \right] + \mathbb{P} \left[ \left| E_n^{M,U} - E_n \right| \geq \alpha_2 a_{\sigma,\varphi} \right]). \end{aligned}$$

Claim (b) of Theorem 3 and Equation (5.9) in Lemma 6 yield the statement.  $\square$

**5.4. Proof of Lemma 6.** We begin with the proof of (5.8). Let us write from (5.1), (5.4):

$$\begin{aligned} |E_n^1 - E_n^{1,M}| &= \left| \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^{3/2} \left[ \tilde{\mathbb{E}} \left[ (1 - T_1) T_1 \Lambda_{k-1}(T_1, T_2, x) - (1 - T_1^M) T_1^M \Lambda_{k-1}(T_1^M, T_2^M, x) \right] \right]_{x=X_{k-1}} \right| \\ &\leq \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^{3/2} \left( \left| \left[ \tilde{\mathbb{E}} \left[ ((1 - T_1) T_1 - (1 - T_1^M) T_1^M) \Lambda_{k-1}(T_1^M, T_2^M, x) \right] \right]_{x=X_{k-1}} \right| + \right. \\ &\quad \left. + \frac{1}{4} \left[ \tilde{\mathbb{E}} \left[ \left| \Lambda_{k-1}(T_1^M, T_2^M, x) - \Lambda_{k-1}(T_1, T_2, x) \right| \right] \right]_{x=X_{k-1}} \right) \\ (5.10) \quad &=: \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^{3/2} (I_k^{M,1} + I_k^{M,2}). \end{aligned}$$

Recalling that  $\mathbb{E}[U_1^{\otimes 3}] = 0$  and that  $T_1$  and  $T_1^M$  are independent of  $U_1$ , we get by cancellation that for all  $k \in \llbracket 1, n \rrbracket$  with the notations of (2.1):

$$\begin{aligned} I_k^{M,1} &= \left| \left[ \tilde{\mathbb{E}} \left[ ((1 - T_1) T_1 - (1 - T_1^M) T_1^M) \times (\Lambda_{k-1}(T_1^M, T_2^M, x) - \Lambda_{k-1}(0, 0, x)) \right] \right]_{x=X_{k-1}} \right| \\ &= \left| \left[ \tilde{\mathbb{E}} \left[ ((T_1^M)^2 - T_1^2) \times (\Lambda_{k-1}(T_1^M, T_2^M, x) - \Lambda_{k-1}(0, 0, x)) \right] \right]_{x=X_{k-1}} \right|, \end{aligned}$$

conditioning w.r.t.  $\sigma(T_1^M, T_2^M)$  (recalling that  $T_1^M, T_2^M$  are independent) and using the stationarity property in (5.2) for the last equality. Observe that the stationarity also yields:

$$\tilde{\mathbb{E}}[(T_1^M)^2 - T_1^2 | T_1^M] = -\tilde{\mathbb{E}}[(T_1^M)^2 - 2T_1 T_1^M + T_1^2 | T_1^M] = -\mathbb{E}[(T_1^M - T_1)^2 | T_1^M].$$

Thus:

$$(5.11) \quad I_k^{M,1} = \left| \left[ \tilde{\mathbb{E}} \left[ (T_1^M - T_1)^2 \times (\Lambda_{k-1}(T_1^M, T_2^M, x) - \Lambda_{k-1}(0, 0, x)) \right] \right]_{x=X_{k-1}} \right|.$$

From the  $\beta$ -Hölder continuity of  $D^3\varphi$  and the definition of  $\Lambda_{k-1}$  in (2.1), we derive:

$$|\Lambda_{k-1}(T_1^M, T_2^M, x) - \Lambda_{k-1}(0, 0, x)| \leq [\varphi^{(3)}]_\beta \mathbb{E}[|U_1|^{3+\beta}] \gamma_k^{\beta/2} \|\sigma\|_\infty^{3+\beta},$$

which, combined together with (5.11) and (5.3), (with  $a_n$  defined in (2.3)):

$$(5.12) \quad \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^{3/2} I_k^{M,1} \leq \frac{1}{12} M^{-2} [\varphi^{(3)}]_\beta \mathbb{E}[|U_1|^{3+\beta}] \|\sigma\|_\infty^{3+\beta} \frac{\Gamma_n^{(3+\beta)/2}}{\sqrt{\Gamma_n}} = \frac{1}{12} M^{-2} a_n (1+\beta)(2+\beta)(3+\beta).$$

We obtain similarly that for all  $k \in \llbracket 1, n \rrbracket$ :

$$\begin{aligned} |I_k^{M,2}| &= \frac{1}{4} \left[ \tilde{\mathbb{E}}[|\Lambda_{k-1}(T_1^M, T_2^M, x) - \Lambda_{k-1}(T_1, T_2, x)|] \right]_{x=X_{k-1}} \\ &\leq \frac{1}{4} [\varphi^{(3)}]_\beta \|\sigma\|_\infty^{3+\beta} \mathbb{E}[|U_1|^{3+\beta}] \gamma_k^{\beta/2} \tilde{\mathbb{E}}[|T_1^M T_2^M - T_1 T_2|^\beta] \\ &= \frac{1}{4} [\varphi^{(3)}]_\beta \|\sigma\|_\infty^{3+\beta} \mathbb{E}[|U_1|^{3+\beta}] \gamma_k^{\beta/2} \tilde{\mathbb{E}}[|T_2^M - T_2|^\beta + |T_1^M - T_1|^\beta] \\ &\leq [\varphi^{(3)}]_\beta \frac{\|\sigma\|_\infty^{3+\beta} \mathbb{E}[|U_1|^{3+\beta}] M^{-\beta} \gamma_k^{\beta/2}}{2} \times \frac{1}{12^{\beta/2}}, \end{aligned}$$

using the Hölder inequality and (5.3) for the last inequality. We finally get:

$$(5.13) \quad \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^{3/2} I_k^{M,2} \leq \frac{1}{2 \times 12^{\beta/2}} M^{-\beta} a_n (1+\beta)(2+\beta)(3+\beta).$$

We get the result plugging (5.13) and (5.12) in (5.10) for  $\bar{C}_1 = \frac{1}{12^{\beta/2}} (\frac{1}{12^{1-\beta/2}} + \frac{1}{2}) (1+\beta)(2+\beta)(3+\beta)$ .

The term  $E_n^{1,M,C} - E_n^1$  is handled similarly, up to the additional term  $\mathbb{E}[(1-T_1^M)T_1^M(\Lambda_{k-1}^M - \Lambda_{k-1})(T_1^M, T_2^M, x)]$  which gives the stated bound, using again centering arguments (recall that  $\mathbb{E}[U^{\otimes 3}] = \mathbb{E}[(U^M)^{\otimes 3}] = 0$ ), the Hölder continuity of  $D^3\varphi$  and (5.5).

Let us now turn to the proof of (5.9). By definition we get:

$$\begin{aligned} E_n^{M,U} - E_n &= E_n^{1,M,U} - E_n^{1,M} + \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^2 \left[ \tilde{\mathbb{E}} \left[ \left[ ((1-T_1)D^2\varphi(x + T_1\gamma_k b(x)) \right. \right. \right. \\ &\quad \left. \left. \left. - (1-T_1^M)D^2\varphi(x + T_1^M\gamma_k b(x))) b(x) \otimes b(x) \right] \right] \right]_{x=X_{k-1}} \\ &= E_n^{1,M,U} - E_n^{1,M} + \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^2 \left[ \tilde{\mathbb{E}} \left[ \left[ ((1-T_1) - (1-T_1^M))D^2\varphi(x + T_1\gamma_k b(x)) \right. \right. \right. \\ &\quad \left. \left. \left. + (1-T_1^M)[D^2\varphi(x + T_1\gamma_k b(x)) - D^2\varphi(x + T_1^M\gamma_k b(x))] b(x) \otimes b(x) \right] \right] \right]_{x=X_{k-1}}. \end{aligned}$$

Let us now define for  $k \in \llbracket 1, n \rrbracket$ :

$$\begin{aligned} \bar{I}_k^{M,1} &:= \left[ \tilde{\mathbb{E}} \left[ \left[ (T_1^M - T_1)D^2\varphi(x + T_1\gamma_k b(x)) b(x) \otimes b(x) \right] \right] \right]_{x=X_{k-1}}, \\ \bar{I}_k^{M,2} &:= \left[ \tilde{\mathbb{E}} \left[ \left[ (1-T_1^M)[D^2\varphi(x + T_1\gamma_k b(x)) - D^2\varphi(x + T_1^M\gamma_k b(x))] b(x) \otimes b(x) \right] \right] \right]_{x=X_{k-1}}. \end{aligned}$$

Let us now investigate, for  $\lambda > 0$ , the quantity:

$$\begin{aligned} \mathbb{E} \exp \left( \lambda |E_n^{M,U} - E_n| \right) &\leq \mathbb{E} \exp \left( \lambda |E_n^{1,M,U} - E_n^1| + \frac{\lambda}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^2 (|\bar{I}_k^{M,1}| + |\bar{I}_k^{M,2}|) \right) \\ &\leq \exp \left( \lambda a_n^{1,M,U} \right) \\ &\quad \times \mathbb{E} \exp \left( \frac{M^{-1}\lambda}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^2 \sum_{j=2}^3 \left( \sup_{u \in [0,1]} \|D^j\varphi(X_{k-1} + u\gamma_k b_{k-1})\|_\infty \gamma_k^{j-2} (C_V V_{k-1})^{\frac{j}{2}} \right) \right). \end{aligned}$$

Since we have assumed that there exists  $C > 0$  s.t. for all  $j \in \{2, 3\}, x \in \mathbb{R}^d$ ,  $|D^j \varphi(x)| \leq \frac{C}{1+|x|^{j-1}}$  we derive, as in the proof of Lemma 5, from  $(\mathcal{L}_V)$ ,  $(S)$  and the Jensen inequality that there exists  $\bar{C} > 0$  s.t.:

$$\mathbb{E} \exp \left( \lambda |E_n^{M,U} - E_n| \right) \leq \exp \left( \lambda a_n^{1,M,U} \right) \left( \frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 \mathbb{E} \exp \left( \frac{M^{-1} \lambda \Gamma_n^{(2)}}{\sqrt{\Gamma_n}} C \bar{C} (1 + 2V_{k-1}^{\frac{1}{2}}) \right) \right).$$

From the Young inequality we obtain:

$$(5.14) \quad \mathbb{E} \exp \left( \lambda |E_n^{M,U} - E_n| \right) \leq \exp \left( \lambda a_n^{M,U} \right) \exp \left( \lambda^2 \frac{M^{-2} (C \bar{C})^2 (\Gamma_n^{(2)})^2}{c_V \Gamma_n} \right) I_V^1, \\ a_n^{M,U} := a_n^{1,M,U} + C \bar{C} M^{-1} \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}}.$$

Now, for all  $\lambda > 0, a \geq a_n^{M,U}$ :

$$\mathbb{P} [ |E_n^{M,U} - E_n| \geq a ] \leq \exp(-\lambda a) \mathbb{E} \exp \left( \lambda |E_n^{M,U} - E_n| \right) \\ \stackrel{(5.14)}{\leq} \exp \left( -\lambda (a - a_n^{M,U}) \right) \exp \left( \lambda^2 \frac{M^{-2} (C \bar{C})^2 (\Gamma_n^{(2)})^2}{c_V \Gamma_n} \right) I_V^1.$$

An optimization over  $\lambda$  yields:

$$\mathbb{P} [ |E_n^{M,U} - E_n| \geq a ] \leq \exp \left( -\bar{c}_2 M^2 \frac{\Gamma_n}{(\Gamma_n^{(2)})^2} (a - a_n^{M,U})^2 \right) I_V^1, \quad \bar{c}_2 := \frac{c_V}{4(C \bar{C})^2}.$$

This concludes the proof.  $\square$

Let us conclude saying that, when the function  $\varphi$  is not explicitly known, it is then still possible to estimate the bias  $m$  using point (b) in Theorem 1 for  $\tilde{\gamma} = +\infty$ , running the ergodic mean for a time step sequence of the form  $\gamma_n \asymp n^{-\theta}$ ,  $\theta < 1/3$ . On the other hand, the explicit knowledge of  $\varphi$  can occur to consider ergodic control variates in the simulation.

## 6. APPLICATION: 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  $k \geq 1$ ,  $\gamma_k = 1/k$ :

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

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  $n \geq 0$ :

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

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

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

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  $\nu_n^Z$  satisfies the conclusion of Theorem 1 point (a) (when  $\gamma_n = \frac{1}{n}$ ,  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow 0$ ), 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. Non-Asymptotic Deviation Bounds.

**Theorem 7.** *Assume the innovation sequence  $(U_n)_{n \geq 1}$  satisfies **(GC)** and let  $f \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R})$  be a globally Lipschitz continuous function s.t.  $[f^{(1)}]_\beta < +\infty$  for  $\beta \in (0, 1)$  and  $G(f) = \int_{\mathbb{R}^d} f(x)G(dx) = 0$ . Then, there exist explicit non-negative sequences  $(c_n)_{n \geq 1}$  and  $(C_n)_{n \geq 1}$ , respectively increasing and decreasing for  $n$  large enough, with  $\lim_n C_n = \lim_n c_n = 1$  s.t. for all  $a > a_n = \frac{[\varphi^{(3)}]_\beta \|\sigma\|_\infty^{3+\beta} \mathbb{E}[|U_1|^{3+\beta}] \Gamma_n^{(\frac{3+\beta}{2})}}{(1+\beta)(2+\beta)(3+\beta) \sqrt{\Gamma_n}}$  and  $n \geq 1$ :*

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

where  $\varphi$  denotes the solution of the Poisson Equation:

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

which, under the current assumptions, is unique and belongs to  $\mathcal{C}^3(\mathbb{R}^d, \mathbb{R})$  with  $[\varphi^{(3)}]_\beta < +\infty$ . Observe that for the specific considered time step  $\gamma_k = \frac{1}{k}$ ,  $a_n = O\left(\frac{1}{\sqrt{\log(n)}}\right) \rightarrow 0$ .

*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  $\nu_n^Z$  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 (2.8):

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

for  $\bar{q}, q \in (1, +\infty)$ ,  $\bar{p} = \frac{\bar{q}}{\bar{q}-1}$ ,  $p = \frac{q}{q-1}$ . 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  $C_{6.7}$  s.t. for all  $\lambda > 0$ :*

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

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

On the other hand, assuming the indicated smoothness for the solution of the Poisson Equation (6.5) (this result is proved in Section 6.2), we derive from (6.6), (6.7) similarly to the proof of Theorem 3 by setting  $\bar{\lambda}_n := \frac{(a-a_n)\sqrt{\Gamma_n}}{q\bar{q}\|\nabla\varphi\|_\infty^2}$ :

$$\begin{aligned} \mathbb{P}[\sqrt{\Gamma_n}|\nu_n^Z(f)| \geq a] &\leq 2 \exp\left(-\frac{(a-a_n)^2}{2q\bar{q}\|\nabla\varphi\|_\infty^2}\right) \exp\left(\frac{C_{6.7}\bar{\lambda}_n[f]_1\mathbb{E}[|U_1|]\Gamma_n^{\frac{3}{2}}}{\Gamma_n}\right) \exp\left(\frac{C_{6.7}^2\bar{p}[f]_1^2\bar{\lambda}_n^2\Gamma_n^3}{2\Gamma_n^2}\right) \\ &\quad \times (I_V^1)^{\frac{1}{p\bar{q}}} \exp\left(\frac{1}{p\bar{q}}(c_V + \frac{C_{2.12}}{2})\right) \exp\left(\bar{\lambda}_n^2 p\bar{q}\left(\frac{3C_{V,\varphi}^2}{c_V\Gamma_n^2} + [C_{2.11} + \frac{3}{2}C_{2.12}]\frac{(\Gamma_n^{(2)})^2}{\Gamma_n^2}\right)\right) \\ &\leq 2(I_V^1)^{\frac{1}{p\bar{q}}} \exp\left(\frac{1}{p\bar{q}}(c_V + \frac{C_{2.12}}{2}) + \frac{1}{2\bar{p}}\right) \exp\left(\frac{\bar{p}[f]_1^2\bar{\lambda}_n^2 C_{6.7}^2}{2\Gamma_n^2}\left(\Gamma_n^{(3)} + \mathbb{E}[|U_1|]^2(\Gamma_n^{\frac{3}{2}})^2\right)\right) \\ &\quad \times \exp\left(-\frac{(a-a_n)^2}{2q\bar{q}\|\nabla\varphi\|_\infty^2}\left(1-d_n - \frac{\bar{p}}{q\bar{q}\|\nabla\varphi\|_\infty^2}\frac{[f]_1^2 C_{6.7}^2\left(\Gamma_n^{(3)} + \mathbb{E}[|U_1|]^2(\Gamma_n^{\frac{3}{2}})^2\right)}{\Gamma_n}\right)\right) \end{aligned}$$

with  $d_n$  as in (2.16). Choose again  $(p_n)_{n \geq 1}$  as in (2.16) so that  $q_n \xrightarrow{n} 1, d_n \xrightarrow{n} 0$  with the indicated monotonicity for  $n$  large enough. We can now take  $\bar{p} := \bar{p}_n \xrightarrow{n} +\infty$  s.t.  $\frac{\bar{p}}{\Gamma_n} \xrightarrow{n} 0$ . The above inequality then gives the result up to a direct modification of the sequences  $(C_n)_{n \geq 1}, (c_n)_{n \geq 1}$ .  $\square$

*Proof of Lemma 7.* The definition of  $\Delta_n$  implies:

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

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

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

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

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

a direct induction on  $\Delta_n$  yields:

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

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

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

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

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

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

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

Plugging this inequality in (6.11), we derive:

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

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

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

This completes the proof.  $\square$

**6.2. Regularity Results for the Poisson Equation.** In this Section we prove the following regularity result.

**Lemma 8.** *Let  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  be as in Theorem 7. Then the Poisson Equation (6.5) admits a unique solution  $\varphi$  which is in  $\mathcal{C}^3(\mathbb{R}^d, \mathbb{R})$  globally Lipschitz continuous and s.t.  $[\varphi^{(3)}]_\beta < +\infty$ ,  $\beta \in (0, 1)$ . Furthermore, there exists  $C_\varphi > 0$  s.t. for all  $x \in \mathbb{R}^d$ ,  $|D^2 \varphi(x)| \leq \frac{C_\varphi}{1+|x|}$ .*

*Proof.* Under the considered assumptions, it is well known, see e.g. Pardoux and Veretennikov [PV01], that the solution of the Poisson Equation (6.5) writes:

$$(6.14) \quad \varphi(x) = - \int_{\mathbb{R}_+} \mathbb{E}[f(X_t^{0,x})] dt,$$

where  $X_t^{0,x}$  is the process with generator  $\mathcal{A} = -\frac{1}{2}x \cdot \nabla + \frac{1}{2}\Delta$ , that is the Ornstein-Uhlenbeck diffusion equation:

$$X_t^{0,x} = x - \frac{1}{2} \int_0^t X_s^{0,x} ds + B_t = \exp\left(-\frac{t}{2}\right)x + \int_0^t \exp\left(\frac{s-t}{2}\right) dB_s \stackrel{(\text{law})}{=} \exp\left(-\frac{t}{2}\right)x + \mathcal{N}(0, \sigma^2(t)I_d),$$

where  $\sigma^2(t) = 1 - \exp(-t)$ . For any given  $t > 0$ , the density therefore explicitly writes

$$(6.15) \quad p(t, x, y) = \frac{1}{(2\pi\sigma^2(t))^{d/2}} \exp\left(-\frac{|y - e^{-t/2}x|^2}{2\sigma^2(t)}\right).$$

Observe as well that for every  $k \in \llbracket 1, d \rrbracket$ :

$$\partial_{x_k} p(t, x, y) = e^{-t/2} y_k - \frac{e^{-t/2} x_k}{\sigma^2(t)} p(t, x, y) = -e^{-t/2} \partial_{y_k} p(t, x, y).$$

So we readily derive that  $\nabla \varphi$  is bounded and as a consequence that  $\varphi$  has linear growth, i.e. there exists  $c \geq 1$  s.t. for all  $x \in \mathbb{R}^d$ ,  $|\varphi(x)| \leq c(1 + |x|)$ . To be more precise, for all  $k \in \llbracket 1, d \rrbracket$ ,

$$\begin{aligned} |\partial_{x_k} \varphi(x)| &= \left| \int_{\mathbb{R}_+} \partial_{x_k} \mathbb{E}[f(X_t^{0,x})] dt \right| = \left| \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} \partial_{x_k} p(t, x, y) f(y) dy dt \right| \\ &= \left| \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} -e^{-t/2} \partial_{y_k} p(t, x, y) f(y) dy dt \right| = \left| \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} e^{-t/2} p(t, x, y) \partial_{y_k} f(y) dy dt \right| \\ &\leq \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} e^{-t/2} p(t, x, y) |\partial_{y_k} f(y)| dy dt \leq 2[f]_1. \end{aligned}$$

Now, setting for all  $x \in \mathbb{R}^d$ ,  $k \in \llbracket 1, d \rrbracket$ ,  $v_k(x) := \int_0^{+\infty} dt \exp(-t/2) \mathbb{E}[\Psi_k(X_t^{0,x})]$ , where for all  $y \in \mathbb{R}^d$ ,  $\Psi_k(y) := -\partial_{y_k} f(y)$ , we observe that  $\partial_{x_k} \varphi(x) = v_k(x)$ . Also, from our assumption on  $f$ , we have that  $\Psi_k \in \mathcal{C}^\beta(\mathbb{R}^d, \mathbb{R})$ . Theorems 2.4-2.6 in Krylov and Priola, [KP10] then yield that there exists a unique solution to the PDE:

$$\mathcal{A} w_k(x) - \frac{1}{2} w_k(x) = -\Psi_k(x),$$

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

$$(6.16) \quad \exists C \geq 1, \|w_k\|_{2+\beta} \leq C(1 + \|\Psi_k\|_\beta).$$

A simple identification procedure following the proof of Theorem II.1.1 in Bass [Bas97] gives  $w_k = v_k$ . The result follows from (6.16). Let us emphasize that this is a quite deep and involved result. To give some intuition about how such estimates can be derived we provide a proof of a slightly weaker statement in Appendix A using direct computations.

The condition on  $|D^2\varphi(x)|$  follows from the Remark p.7 in [KP10]. Let us anyhow provide a direct proof. Observe that since  $\varphi, f, b(x) = -x/2$  are Lipschitz continuous and continuously differentiable, we readily get differentiating equation (6.5) that for all  $k \in \llbracket 1, d \rrbracket$  and all  $x \in \mathbb{R}^d$ :

$$b(x) \cdot \partial_{x_k} \nabla \varphi(x) - \frac{1}{2} \partial_{x_k} \varphi(x) + \frac{1}{2} \partial_{x_k} \Delta \varphi(x) = \partial_{x_k} f(x),$$

so that

$$(6.17) \quad b(x) \cdot \partial_{x_k} \nabla \varphi(x) = \partial_{x_k} f(x) + \frac{1}{2} \partial_{x_k} \varphi(x) - \frac{1}{2} \partial_{x_k} \Delta \varphi(x).$$

Now,  $\varphi$  is such that  $\nabla \varphi, D^2\varphi$ , and  $D^3\varphi$  are bounded. This is also the case for  $\nabla f$ . Equation (6.17) then gives that  $|b(x) \cdot \partial_{x_k} \nabla \varphi(x)|$  is bounded. The specific form of  $b(x) = -x/2$  (Lyapunov condition) gives the stated decay.  $\square$

## 7. NUMERICAL RESULTS

We present in this section numerical results associated with the computation of the empirical measure  $\nu_n$  illustrating our previous theorems. For all simulations we took  $d = r = 1$ . Also, for simplicity, the innovations  $(U_i)_{i \geq 1}$  and  $X_0$  are Bernoulli variables with  $\mathbb{P}(U_1 = -1) = \mathbb{P}(U_1 = 1) = \frac{1}{2}$ .

**7.1. Sub-Gaussian tails.** We first illustrate Theorem 2 taking  $b(x) = -\frac{x}{2}$ , and  $\sigma(x) = \cos(x)$  in (1.1). This is a (weakly) hypoelliptic example. Indeed, setting for  $x \in \mathbb{R}$ ,  $X_1(x) = \cos(x)\partial_x$  and  $X_0(x) = -\frac{x}{2}\partial_x$ , we have  $\text{span}\{X_1, [X_1, X_0]\} = \mathbb{R}$ . We choose as well to compute  $\nu_n(\mathcal{A}\varphi)$  for  $\varphi(x) = x + \varepsilon \cos(x)$  for  $\varepsilon = 0.01$ , and  $\varphi = \cos(x)$ . 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  $(\gamma_k)_{k \geq 1} = (k^{-\theta})_{k \geq 1}$ ,  $\theta \in [1/3, 1]$ , the function  $a \in \mathbb{R}^+ \mapsto g_{n,\theta}(a) := \log(\mathbb{P}[\sqrt{\Gamma_n}|\nu_n(\mathcal{A}\varphi)| \geq a])$  is s.t. for  $a > a_n := a_n(\theta)$  defined in (2.3):

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

We plot in Figure 1 the curves of  $g_{n,\theta}$  for  $\theta$  varying as  $\theta_j = \frac{1}{3} + (1 - \frac{1}{3})\frac{j}{5}$ , for  $j \in \llbracket 1, 5 \rrbracket$ ,  $\varphi(x) = \cos(x) + \varepsilon \cos(x)$  and in Figure 2 the curve of  $g_{n,\theta}$  for  $\theta_0 = \frac{1}{3}$  and  $\varphi(x) = \cos(x)$ . The simulations have been performed for  $n = 5 \times 10^4$  in Figure 1,  $n = 10^6$  in Figure 2, and the probability estimated by Monte Carlo simulation for  $MC = 10^4$  realizations of the random variable  $\sqrt{\Gamma_n}|\nu_n(\mathcal{A}\varphi)|$ . 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 10, optimizing numerically in  $\rho$ . 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 4 when  $\sigma^2 - \nu(\sigma^2)$  is a coboundary and in the asymptotic Theorem 1. In the unbiased case of Figure 1, the  $a_n(\theta)$  have, for the considered parameters, almost the same order for  $j \in \llbracket 1, 5 \rrbracket$  (at most  $2.7 \cdot 10^{-2}$ ). For the associated graph, we plot the maximum in  $j$  of the  $(S_{n,\theta_j})_{j \in \llbracket 1, 5 \rrbracket}, (S_{n,\theta_j,c})_{j \in \llbracket 1, 5 \rrbracket}, (S_{n,\theta_j,A})_{j \in \llbracket 1, 5 \rrbracket}, (P(\lambda_{\min})(n, \theta_j, a, \rho))_{j \in \llbracket 1, 5 \rrbracket}$  corresponding to  $j = 1$ . The associated curves are denoted by  $S_n, S_{n,c}, S_{n,A}$  and  $P(\lambda_{\min})(n)$ .

The Figures 1 and 2 correspond to unbiased and biased cases respectively. In the biased 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 unbiased 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 10 seems rather appropriate to capture the deviations of the empirical random measures.

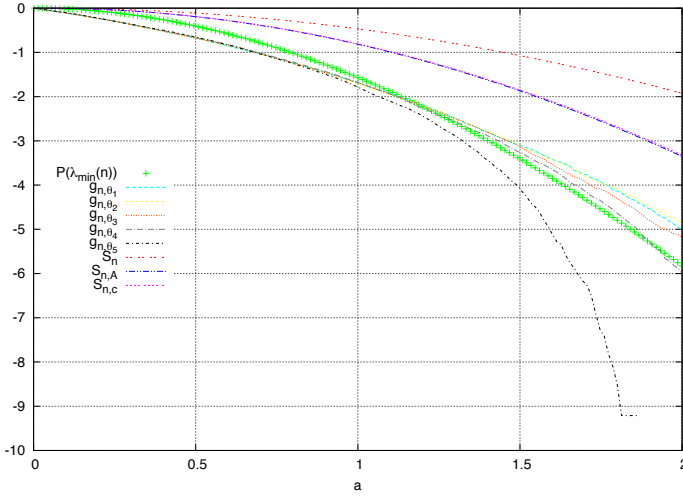


FIGURE 1. Unbiased Case. Plot of  $a \mapsto g_{n,\theta}(a)$ , for  $(\theta_k)_{k \in [1,5]}$ , with  $\varphi(x) = \sigma(x) = x + \varepsilon \cos(x)$ ,  $\varepsilon = 0.01$ .

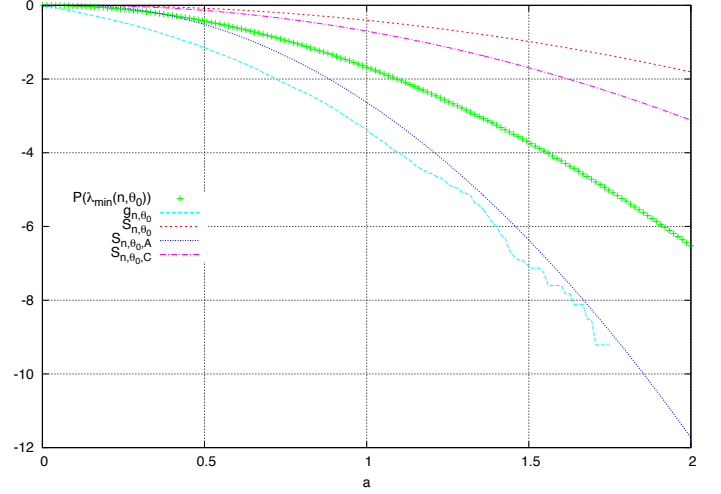


FIGURE 2. Biased Case. Plot of  $a \mapsto g_{n,\theta}(a)$ , for  $\theta_0 = \frac{1}{3}$ , with  $\varphi(x) = \sigma(x) = \cos(x)$ .

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 10 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

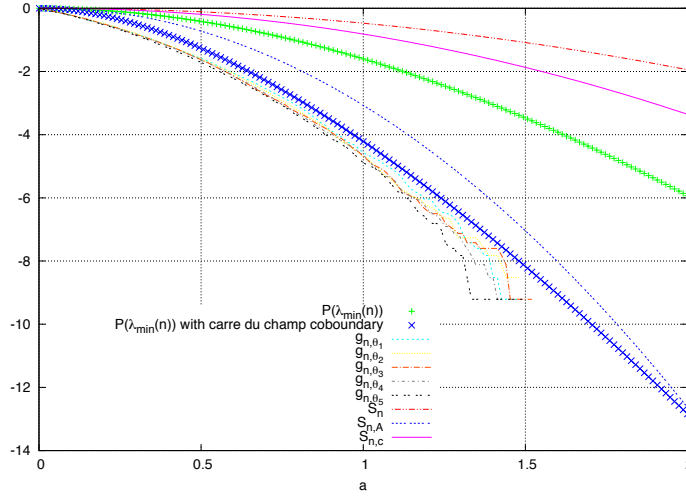


FIGURE 3. Plot of  $a \mapsto g_{n,\theta}(a)$ , for  $(\theta_k)_{k \in [1,5]}$ , with  $\varphi(x) = \sigma(x) = \cos(x)$ .

Theorem 4 cannot be extended to justify such a choice, the empirical evidence is rather striking.

**7.2. Estimation of the Bias.** In Theorem 2, to attain the fastest convergence for  $\theta = \frac{1}{3}$ ,  $\beta = 1$  a bias  $-\tilde{\gamma}m + E_n$  appears. We choose to compute it numerically by quantization (see Section 5). We emphasize how this approximation behaves for the Ornstein-Uhlenbeck process corresponding to  $b(x) = -\frac{x}{2}$  and  $\sigma(x) = 1$  in (1.1). In this case, the invariant measure  $\nu$  of the SDE is the standard Gaussian law. Also, for  $\varphi(x) = \cos(x)$ , the mean  $\tilde{\gamma}m$  and the asymptotic variance  $\Sigma^2 := \int_{\mathbb{R}} |\nabla\varphi(x)|^2 \nu(dx)$  of the error are explicitly computable (with  $\tilde{\gamma} = \lim_n \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} = \sqrt{6}\gamma_0^{3/2}$  and  $m$  is specified in point (b) of Theorem 1). Also  $m = \exp(-\frac{1}{2})\frac{11}{24} \simeq 0,2780$ ,  $\Sigma^2 = \frac{1}{2}(1 - \exp(-2))$ .

On the other hand, the mean square error  $\mathcal{E}rr_2(\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)) := \mathbb{E}\left[(\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi))^2\right] \xrightarrow{n} \tilde{\gamma}^2 m^2 + \Sigma^2$ . So far, we have always chosen  $(\gamma_k)_{k \geq 1} = \left(\frac{\gamma_0}{k^\theta}\right)_{k \geq 1}$  without paying much attention to  $\gamma_0$ . Our goal is now to choose  $\gamma_0$  in order to minimize the previous mean square error. To this end, observe that:

$$\gamma_0^{-1} \mathcal{E}rr_2(\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)) \xrightarrow{n} \gamma_0^2 \tilde{\gamma}_1^2 m^2 + \frac{\Sigma^2}{\gamma_0},$$

with  $\tilde{\gamma}_1 := \lim_n \frac{\bar{\Gamma}_n^{(2)}}{\sqrt{\Gamma_n}} = \sqrt{6}$  where  $\bar{\Gamma}_n, \bar{\Gamma}_n^{(2)}$  are associated with the sequence  $(\tilde{\gamma}_k)_{k \geq 1} := (k^{-\frac{1}{3}})_{k \geq 1}$  corresponding to  $\gamma_0 = 1$ . An optimisation of the previous r.h.s. with respect to  $\gamma_0$  then yields

$$(7.1) \quad \gamma_0 = \left(\frac{\Sigma^2}{2m^2\tilde{\gamma}_1^2}\right)^{\frac{1}{3}} = \left(\frac{\Sigma^2}{12m^2}\right)^{\frac{1}{3}}.$$

We plot below in Figure 7.2 the empirical density of  $\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)$ , the corrected empirical density of  $\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) + E_n^{1,M,U} + E_n^{2,M}$ , where  $E_n^{1,M,U} + E_n^{2,M} = E_n^{M,U} - \tilde{\gamma}m$  is an explicitly computable quantity which does not require the explicit knowledge of  $m$  (see again Theorems 2, 3 and Section 5), and the density of  $\mathcal{N}(\tilde{\gamma}m, \Sigma^2)$  (theoretical limit law of the error). We considered for the simulation  $n = 10^4$ ,  $MC = 10^4$  realizations and took for the quantization method  $M = 10$ .

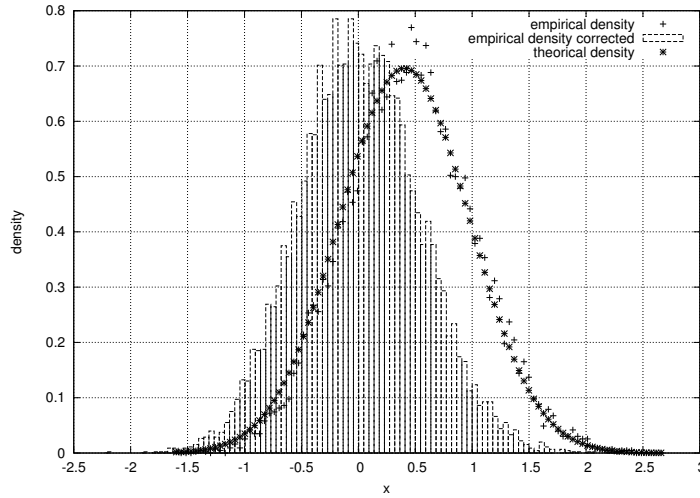


FIGURE 4. Empirical and theoretical density of  $\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)$

The Figure 7.2 emphasizes that the variance of the empirical densities matches the theoretical one. Also, the numerical approximation  $E_n^{1,M,U} + E_n^{2,M} = E_n^{M,U} - \tilde{\gamma}m$  of the bias correction  $-\tilde{\gamma}m + E_n$  seems efficient even for the small number of quantization points considered: without correction the empirical mean is 0.465 (the *theoretical* mean reads  $\sqrt{6}\gamma_0^{3/2}m = 0.4649$ ) whereas with the correction it reads 0.0317. For the previous simulations, the 95% confidence interval associated with the Monte Carlo error is uniformly bounded by 0.022.

**7.3. Regularity and Bias.** The aim of this subsection is to emphasize numerically the bias induced by either the maximum velocity in the deviation rates, corresponding to steps  $\gamma_n \asymp n^{-1/3}$ , or by the *low* regularity of the third order derivatives of the solutions to the Poisson Equation, as indicated by Theorems 2, 3.

We consider, for  $\beta \in \{0.9, 1\}$ , the function  $\varphi(x) = 10 \frac{|x|^{3+\beta}}{1+|x|^{2+\beta}}$  which satisfies the criteria of Theorem 3. The multiplication by 10 is meant to amplify the effects of the *low* regularity (i.e.  $\beta$ -Hölder) at 0.

In both cases, to minimize the mean square error, we choose to optimize  $\gamma_0$  following (7.1). The case  $\beta = 1$ ,  $\theta = \frac{1}{3}$  is exactly covered by Theorem 2. The parameters  $m, \Sigma^2$  needed in (7.1) are estimated by a Monte-Carlo procedure with  $n = MC = 10^4$  for the step sequence  $(\tilde{\gamma}_k)_{k \geq 1} = (k^{-1/3})_{k \geq 1}$ . The width of the associated 95% confidence interval is 0.018. Keep as well in mind that  $\tilde{\gamma}_1 = \sqrt{6}$ .

For  $\beta = 0.9$ , we have not proven a limit theorem similar to Theorem 1, giving the asymptotic expression of the bias. From our previous analysis, considering the bias appearing in Theorem 3 for a  $\beta$ -Hölder  $D^3\varphi$ , we

replace the parameter  $m$  in (7.1) by  $m_n := \frac{\Gamma_n}{\Gamma_n^{(\frac{3+\beta}{2})}} \nu_n(\mathcal{A}\varphi)$  for  $n = 10^6$ . Indeed, we are tempted to say, that in the current case, the following law of large numbers (extending by analogy the slow decreasing case  $b$ ) in Theorem 1) holds:

$$(7.2) \quad \frac{\Gamma_n}{\Gamma_n^{(\frac{3+\beta}{2})}} \nu_n(\mathcal{A}\varphi) \xrightarrow{\mathbb{P}} m.$$

Numerical experiments tend to confirm this conjecture. Indeed, running a Monte Carlo procedure associated with a sequence  $(m_n^i)_{i \in [1, MC]}$  of independent random variables with law  $m_n$ , for steps of the form  $\gamma_n \asymp n^{-\theta}$ ,  $\theta = 0.1$ , taking  $MC = 100$ , the corresponding empirical mean writes  $-0.032$  and the empirical variance reads  $2 \times 10^{-5}$ . We chose a *coarse* time step in order to have a longer running time for the estimation of the l.h.s. in (7.2).

Eventually, the parameters  $\tilde{\gamma}_1$  and  $\Sigma^2$  in (7.1) are computed as above for the same parameters  $n, MC$ .

In order to *spend* a relevant amount of time around 0 (*singularity* of  $D^3\varphi$  which is  $\beta$ -Hölder at that point), we take  $b(x) = -0.15x$ ,  $\sigma(x) = 0.3$  (so that the limit law is  $\mathcal{N}(0, 0.3)$ ) and  $X_0 \sim \mathcal{N}(0, 0.5)$ . We have computed simulations for  $n = 10^4$  and for Monte Carlo simulations we have performed  $MC = 10^4$  realizations and  $M = 30$  for the quantization points. For Figure 5,  $\beta = 1$ , and for Figure 6,  $\beta = 0.9$ . The choice of  $\beta = 0.9$  was here motivated in order to keep a *good* convergence rate for the quantization (see Section 5).

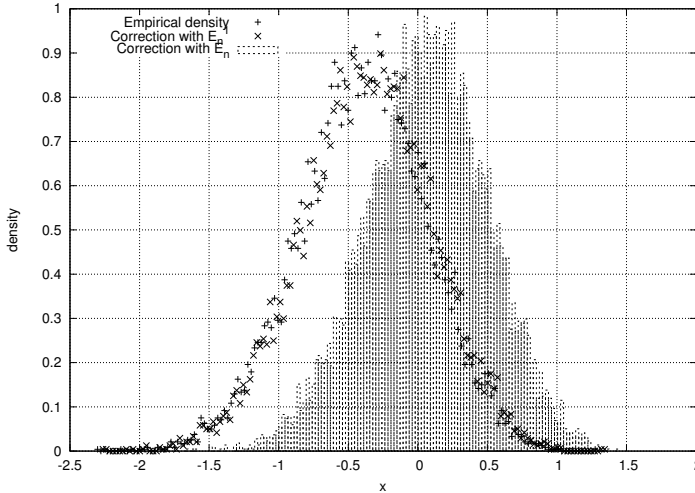


FIGURE 5. Empirical density of  $\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)$  for  $\beta = 1$ .

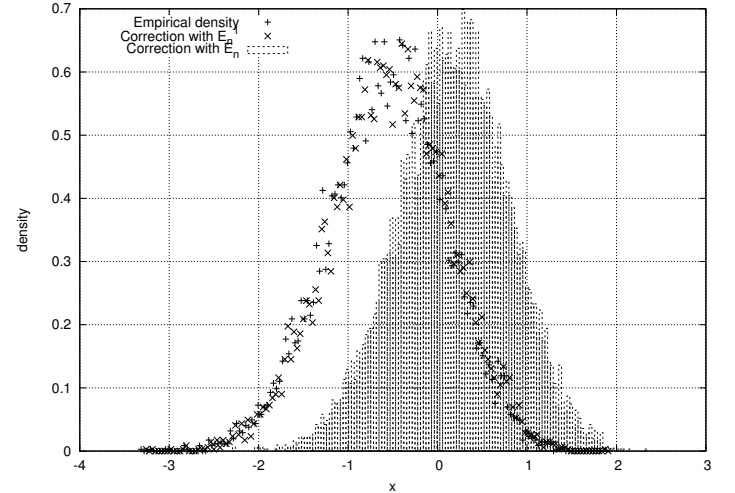


FIGURE 6. Empirical density of  $\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)$  for  $\beta = 0.9$ .

In Figure 5, the empirical mean without correction is  $-0.382765$ . With the correction  $E_n$ , the average is reduced to  $0.0798771$ . In Figure 6, for  $\beta = 0.9$ , we observe another bias due to the lack of regularity for  $\theta = \frac{1}{2+\beta}$ . Indeed, without correction the average is  $-0.556343$ . With correction  $E_n^1$  it goes down slightly to  $-0.516002$ , whereas with the *full* correction  $E_n$  it is reduced to  $0.194132$  (though now overestimated). The test function is only *locally* Hölder, around 0. Numerically, this local irregularity is difficult to catch, in spite of the presence of a non negligible bias. This is why the correction with  $E_n^1$  seems ineffective to annihilate the bias.

Furthermore, the complete correction with  $E_n$ , which appears in Theorem 2 for smoother third derivatives, only partially reduces the bias, seemingly. For both Figures 5 and 6 the confident interval has a length around to 0.026.

It could be interesting, in order to investigate the efficiency of the correction with  $E_n^1$  only, to observe the behavior of the empirical measure for a function  $\varphi$  with almost everywhere Hölder continuous third derivatives, like e.g. the Weierstrass function (see e.g. [Zyg36]).

## APPENDIX A. SCHAUDER ESTIMATES FOR THE POISSON PROBLEM: A DIRECT PROOF

We prove in this section a result *weaker* than Lemma 8, which can be anyhow established by direct computations. We will use the notations already introduced in Section 6.2.

**Lemma 9.** *Let  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  be as in Theorem 7 with  $\beta \in (0, 1]$ . Then the Poisson Equation (6.5) admits a unique solution  $\varphi \in \mathcal{C}^3(\mathbb{R}^d, \mathbb{R})$ , globally Lipschitz continuous and s.t. for all  $\varepsilon > 0$ ,  $C_{\beta, \varepsilon} := [\varphi^{(3)}]_{\beta - \varepsilon} < +\infty$  with  $C_{\beta, \varepsilon} \xrightarrow{\varepsilon \rightarrow 0} +\infty$ . Furthermore, there exists  $C_\varphi > 0$  s.t. for all  $x \in \mathbb{R}^d$ ,  $|D^2\varphi(x)| \leq \frac{C_\varphi}{1+|x|}$ .*

*Proof.* We recall that the solution of (6.5) writes for all  $x \in \mathbb{R}^d$ ,  $\varphi(x) = -\int_0^{+\infty} \mathbb{E}[f(X_t^{0,x})]dt$ , where  $X_t^{0,x} = \exp(-\frac{t}{2})x + \int_0^t \exp(-\frac{t-s}{2})dB_s$  is the Ornstein-Uhlenbeck process whose invariant law is the standard  $d$ -dimensional Gaussian law and whose Gaussian density at time  $t > 0$  is denoted by  $p(t, x, \cdot)$ . The differentiability of  $\varphi$  as well as the control  $\|\nabla\varphi\|_\infty \leq C[f]_1$  were obtained directly.

Also, we had defined in the proof of Lemma 8, for all  $k \in \llbracket 1, d \rrbracket$ ,

$$v_k(x) := \partial_{x_k}\varphi(x) = \int_0^{+\infty} \exp\left(-\frac{t}{2}\right)\mathbb{E}[\Psi_k(X_t^{0,x})]dt,$$

where  $\Psi_k = -\partial_{y_k}f \in \mathcal{C}^\beta(\mathbb{R}^d, \mathbb{R})$ .

The statement of the lemma will follow provided we establish that, for all  $k \in \llbracket 1, d \rrbracket$ ,  $v_k$  solves the Poisson Equation

$$(A.1) \quad \mathcal{A}v_k(x) - \frac{1}{2}v_k(x) = -\Psi_k(x),$$

and satisfies:

$$(A.2) \quad \exists C_{\beta, \varepsilon} \geq 1, \forall k \in \llbracket 1, d \rrbracket, \|v_k\|_{2+\beta-\varepsilon} \leq C_{\beta, \varepsilon}.$$

To prove this control, we focus on the second derivatives  $(\partial_{x_i, x_j}v_k)_{(i,j) \in \llbracket 1, d \rrbracket}$  which yield the biggest singularity when differentiating the density  $p(t, x, \cdot)$  of  $X_t^{0,x}$ . We will establish that these derivatives are bounded and  $(\beta - \varepsilon)$ -Hölder continuous. The other contributions in the  $\mathcal{C}^{2+\beta-\varepsilon}$ -Hölder norm in (A.2) can be handled more directly.

Direct computations carried out on the explicit density of the Ornstein-Uhlenbeck process give that for a given  $m \in \mathbb{N}$ , there exists  $C_m \geq 1$  s.t. for all multi-index  $\alpha \in \mathbb{N}_0^d$ ,  $|\alpha| := \sum_{i=1}^d \alpha_i \leq m$  and all  $(t, x, y) \in \mathbb{R}^+ \times (\mathbb{R}^d)^2$ :

$$(A.3) \quad |\partial_x^\alpha p(t, x, y)| \leq \frac{C_m \exp\left(-\frac{|\alpha|t}{2}\right)}{\sigma(t)^{d+|\alpha|}} \exp\left(-C_m^{-1} \frac{|y - \exp(-\frac{t}{2})x|^2}{2\sigma^2(t)}\right), \quad \sigma^2(t) = 1 - \exp(-t).$$

The above control gives that, for  $\eta > 0$  and  $(i, j, x) \in \llbracket 1, d \rrbracket^2 \times \mathbb{R}^d$  the quantity:

$$D_{i,j}^\eta(x) := \int_\eta^{+\infty} \exp\left(-\frac{t}{2}\right) \int_{\mathbb{R}^d} \partial_{x_i, x_j}^2 p(t, x, y) \Psi_k(y) dy dt$$

is well defined. The first point is to prove that this quantity is uniformly bounded w.r.t.  $\eta$ . To this end we use a cancellation property and the Hölder continuity of  $\Psi_k$ . Note indeed that, since  $\int_{\mathbb{R}^d} p(t, x, y) dy = 1$  then  $\int_{\mathbb{R}^d} \partial_{x_i, x_j} p(t, x, y) dy = 0$ . So, we also have:

$$D_{i,j}^\eta(x) := \int_\eta^{+\infty} \exp\left(-\frac{t}{2}\right) \int_{\mathbb{R}^d} \partial_{x_i, x_j}^2 p(t, x, y) \left(\Psi_k(y) - \Psi_k\left(\exp\left(-\frac{t}{2}\right)x\right)\right) dy dt.$$

The  $\beta$ -Hölder continuity of  $\Psi_k$ , (A.3) and the boundedness of the mapping  $u \mapsto \exp(-cu^2)|u|^\beta, c > 0$ , on the whole real line, then yield:

$$\begin{aligned} |D_{i,j}^\eta(x)| &\leq C_2 \int_\eta^{+\infty} \exp\left(-\frac{3t}{2}\right) \int_{\mathbb{R}^d} \frac{1}{\sigma(t)^{d+2}} \exp\left(-C_2^{-1} \frac{|y - \exp(-\frac{t}{2})x|^2}{2\sigma^2(t)}\right) |y - \exp\left(-\frac{t}{2}\right)x|^\beta dy dt \\ &\leq C_2 \int_\eta^{+\infty} \exp\left(-\frac{3t}{2}\right) \frac{1}{\sigma(t)^{2-\beta}} \int_{\mathbb{R}^d} \frac{1}{\sigma(t)^d} \exp\left(-C_{2,\beta}^{-1} \frac{|y - \exp(-\frac{t}{2})x|^2}{2\sigma^2(t)}\right) dy dt =: \bar{C}_\beta < +\infty, \end{aligned}$$

where  $\bar{C}_\beta$  does not depend on  $x$  (integration, up to multiplicative constants, of a Gaussian density). Indeed, the only integrability issue in the above integral is for  $t$  in a neighborhood of 0. In that case  $\sigma(t)^{2-\beta} \sim t^{1-\beta/2}$  giving an integrable singularity. We thus derive that

$$\forall (i, j, x) \in \llbracket 1, d \rrbracket^2 \times \mathbb{R}^d, \partial_{x_i, x_j} v_k(x) = \lim_{\eta \rightarrow 0} D_{i,j}^\eta(x) \text{ and } \exists \bar{C}_\beta > 0, \forall (i, j) \in \llbracket 1, d \rrbracket^2, \|\partial_{x_i, x_j} v_k\|_\infty \leq \bar{C}_\beta.$$

Let us remark that  $\nabla_x v_k(x)$  could be directly bounded, i.e. without centering. Indeed, from (A.3),  $\nabla_x p(t, x, y)$  yields an integrable singularity in  $\sigma(t)^{-1} \sim t^{-1/2}$  in a neighborhood of 0. We thus have, up to a modification

of  $\bar{C}_\beta$ ,  $\|\nabla v_k\|_\infty + \|D^2 v_k\|_\infty \leq \bar{C}_\beta$ . Recalling that  $(\partial_t + \mathcal{A})p(t, x, y) = 0$  we readily derive that  $v_k$  satisfies (A.1). The equation and the bounds then also readily give that there exists  $C_\varphi > 0$  s.t. for all  $x \in \mathbb{R}^d$ ,  $|\nabla v_k(x)| \leq \frac{C}{1+|x|}$ .

Let us now turn to the regularity of  $\partial_{x_i, x_j} v_k$ . To investigate  $[\partial_{x_i, x_j}^2 v]_{\beta-\varepsilon} := \sup_{x \neq x'} \frac{|\partial_{x_i, x_j}^2 v(x) - \partial_{x_i, x_j}^2 v(x')|}{|x-x'|^{\beta-\varepsilon}}$ , we restrict w.l.o.g. to the case  $|x - x'| \leq 1$ . Set  $\Delta(x, x') := \partial_{x_i, x_j}^2 v(x) - \partial_{x_i, x_j}^2 v(x')$ , we have:

$$\begin{aligned} \Delta(x, x') &= \left( \int_0^1 \exp\left(-\frac{t}{2}\right) \int_{\mathbb{R}^d} (\partial_{x_i, x_j}^2 p(t, x, y) - \partial_{x_i, x_j}^2 p(t, x', y)) \Psi_k(y) dy dt \right) \\ &\quad + \left( \int_1^{+\infty} \exp\left(-\frac{t}{2}\right) \int_{\mathbb{R}^d} (\partial_{x_i, x_j}^2 p(t, x, y) - \partial_{x_i, x_j}^2 p(t, x', y)) \Psi_k(y) dy dt \right) =: \Delta_S(x, x') + \Delta_L(x, x'), \end{aligned}$$

i.e.  $\Delta_S(x, x')$  (resp.  $\Delta_L(x, x')$ ) is the contribution associated with *short* (resp. *long*) times. For  $\Delta_L(x, x')$ , there are no time-singularities. The mean-value theorem and equation (A.3) yield that there exists  $C > 0$  s.t.:

$$(A.4) \quad |\Delta_L(x, x')| \leq C|x - x'| \leq C|x - x'|^{\beta-\varepsilon},$$

since we have assumed  $|x - x'| \leq 1$ .

For  $\Delta_S(x, x')$ , for fixed  $x, x' \in \mathbb{R}^d$ ,  $|x - x'| \leq 1$ , we split the time interval  $[0, 1] := D_1 \cup D_2$ , where  $D_1 := \{t \in [0, 1] : |x - x'| \leq t^{1/2}\}$  and  $D_2 := \{t \in [0, 1] : |x - x'| > t^{1/2}\}$ . We introduce through this partitioning a cutting-threshold corresponding to the usual parabolic scaling.

Set now, for  $i \in \{1, 2\}$ ,

$$\Delta_{S,i}(x, x') := \int_{D_i} \exp\left(-\frac{t}{2}\right) \int_{\mathbb{R}^d} (\partial_{x_i, x_j}^2 p(t, x, y) - \partial_{x_i, x_j}^2 p(t, x', y)) \Psi_k(y) dy dt.$$

For  $\Delta_{S,1}$ , using again a centering argument and the mean-value theorem yields:

$$\begin{aligned} \Delta_{S,1}(x, x') &= \int_{D_1} \exp\left(-\frac{t}{2}\right) \int_{\mathbb{R}^d} (\partial_{x_i, x_j}^2 p(t, x, y) - \partial_{x_i, x_j}^2 p(t, x', y)) \left( \Psi_k(y) - \Psi_k\left(\exp\left(-\frac{t}{2}\right)x'\right) \right) dy dt \\ &= \int_{D_1} \exp\left(-\frac{t}{2}\right) \int_{\mathbb{R}^d} \left( \int_0^1 \nabla(\partial_{x_i, x_j}^2 p(t, x' + \lambda(x - x'), y)) \cdot (x - x') d\lambda \right) \\ &\quad \times \left( \Psi_k(y) - \Psi_k\left(\exp\left(-\frac{t}{2}\right)x'\right) \right) dy dt. \end{aligned}$$

Equation (A.3) then gives:

$$\begin{aligned} |\Delta_{S,1}(x, x')| &\leq C_3 \int_{D_1} \exp(-2t) \int_{\mathbb{R}^d} \frac{1}{t^{(d+3)/2}} \sup_{\lambda \in [0, 1]} \exp\left(-C_3^{-1} \frac{|y - (x' + \lambda(x - x')) \exp(-\frac{t}{2})|^2}{2t}\right) \\ &\quad \times |x - x'| |y - x' \exp(-\frac{t}{2})|^\beta dy dt \\ &\leq C_3 \int_{D_1} \exp(-2t) \int_{\mathbb{R}^d} \frac{1}{t^{(d+3)/2}} \exp\left(-C_3^{-1} \frac{|y - x' \exp(-\frac{t}{2})|^2}{4t}\right) \exp\left(C_3^{-1} \frac{|x - x'|^2}{2t}\right) \\ &\quad \times |x - x'|^{\beta-\varepsilon} t^{(1-(\beta-\varepsilon))/2} |y - x' \exp(-\frac{t}{2})|^\beta dy dt \\ &\leq \bar{C}_3 \int_{D_1} \exp(-2t) t^{-\frac{3}{2} + \frac{1-(\beta-\varepsilon)}{2} + \frac{\beta}{2}} \left( \int_{\mathbb{R}^d} \frac{1}{t^{d/2}} \exp\left(-C_3^{-1} \frac{|y - x' \exp(-\frac{t}{2})|^2}{4t}\right) dy \right) dt \\ &\quad \times |x - x'|^{\beta-\varepsilon} \\ (A.5) \quad &\leq \bar{C}_3 \int_{[0, 1]} t^{-1+\varepsilon/2} dt |x - x'|^{\beta-\varepsilon} \leq \frac{\bar{C}_3 |x - x'|^{\beta-\varepsilon}}{\varepsilon}, \end{aligned}$$

up to a modification of  $\bar{C}_3$  in the last two inequalities.

For  $\Delta_{S,2}(x, x')$  we use again centering techniques to obtain:

$$\begin{aligned} |\Delta_{S,2}(x, x')| &\leq \left| \int_{D_2} \exp\left(-\frac{t}{2}\right) \int_{\mathbb{R}^d} \partial_{x_i, x_j}^2 p(t, x, y) \left( \Psi_k(y) - \Psi_k\left(x \exp\left(-\frac{t}{2}\right)\right) \right) dy dt \right| \\ &\quad + \left| \int_{D_2} \exp\left(-\frac{t}{2}\right) \int_{\mathbb{R}^d} \partial_{x_i, x_j}^2 p(t, x', y) \left( \Psi_k(y) - \Psi_k\left(x' \exp\left(-\frac{t}{2}\right)\right) \right) dy dt \right|. \end{aligned}$$

From (A.3) we get:

$$\begin{aligned}
|\Delta_{S,2}(x, x')| &\leq C_2 \left( \int_{D_2} \exp\left(-\frac{3t}{2}\right) \int_{\mathbb{R}^d} \frac{1}{t^{d/2+1}} \exp\left(-C_2^{-1} \frac{|y - x \exp(-\frac{t}{2})|^2}{4t}\right) |y - x \exp(-\frac{t}{2})|^\beta dy dt \right. \\
&\quad \left. + \int_{D_2} \exp\left(-\frac{3t}{2}\right) \int_{\mathbb{R}^d} \frac{1}{t^{d/2+1}} \exp\left(-C_2^{-1} \frac{|y - x' \exp(-\frac{t}{2})|^2}{4t}\right) |y - x' \exp(-\frac{t}{2})|^\beta dy dt \right) \\
&\leq 2\bar{C}_2 \int_{D_2} t^{-1+\beta/2} dt \leq 2\bar{C}_2 \int_{D_2} t^{-1+\beta/2} \left(\frac{|x - x'|}{t^{1/2}}\right)^{\beta-\varepsilon} dt \\
\text{(A.6)} \quad &\leq \frac{\bar{C}_2 |x - x'|^{\beta-\varepsilon}}{\varepsilon},
\end{aligned}$$

up to a modification of  $\bar{C}_2$  for the last inequalities. The control (A.2) then follows from (A.4), (A.5) and (A.6). Eventually, since we have now established that for all  $k \in \llbracket 1, d \rrbracket$ ,  $v_k \in C^{2+\beta-\varepsilon}(\mathbb{R}^d, \mathbb{R})$ , we derive from Theorem II.I.1 in [Bas97] (usual verification procedure) that  $v_k$  is the unique solution of (A.1). This completes the proof.  $\square$

#### 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.

#### REFERENCES

- [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.
- [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.
- [Boi11] E. Boissard. Simple bounds for the convergence of empirical and occupation measures in 1-Wasserstein distance. *Electronic Journal of Probability*, 16, 2011.
- [BPP05] V. Bally, G. Pagès, and J. Printems. A quantization tree method for pricing and hedging multi-dimensional American options. *Math. Fin.*, 15-1:119–168, 2005.
- [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.
- [DM08] F. Delarue and S. Menozzi. An interpolated stochastic algorithm for quasi-linear PDEs. *Math. Comp.*, 77(261):125–158 (electronic), 2008.
- [Duf90] Marie Dufflo. *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.
- [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.
- [LB06] L. Lorenzi and M. Bertoldi. *Analytical Methods for Markov Semigroups*. CRC-Press, 2006.
- [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.
- [PP05] G. Pagès and J. Printems. Functional quantization for numerics with an application to option pricing. *Monte Carlo Methods And Applications*, 11:407–446, 2005.
- [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.
- [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.
- [Zyg36] A. Zygmund. *Trigonometric Series*. Cambridge University Press, 1936.

UNIVERSITÉ D'ÉVRY 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.menzozi@univ-evry.fr

*Current address:* Laboratoire de Mathématiques et Modélisation d'Évry (LaMME), 23 Boulevard de France, 91037, Evry, 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.