DOI: 10.1214/EJP.v15-814

Electronic Journal of Probability

Vol. 15 (2010), Paper no. 53, pages 1645–1681.

Journal URL
http://www.math.washington.edu/~ejpecp/

On some Non Asymptotic Bounds for the Euler Scheme

V. Lemaire* S. Menozzi†

Abstract

We obtain non asymptotic bounds for the Monte Carlo algorithm associated to the Euler discretization of some diffusion processes. The key tool is the Gaussian concentration satisfied by the density of the discretization scheme. This Gaussian concentration is derived from a Gaussian upper bound of the density of the scheme and a modification of the so-called "Herbst argument" used to prove Logarithmic Sobolev inequalities. We eventually establish a Gaussian lower bound for the density of the scheme that emphasizes the concentration is sharp.

Key words: Non asymptotic Monte Carlo bounds, Discretization schemes, Gaussian concentration.

AMS 2000 Subject Classification: Primary 60H35,65C30,65C05, 60E15.

Submitted to EJP on January 8, 2010, final version accepted September 24, 2010.

*LPMA, CNRS UMR 7599, UPMC Univ Paris 06, 4 place Jussieu 75005 Paris, vincent.lemaire@upmc.fr
†LPMA, CNRS UMR 7599, Université Denis Diderot, 175 Rue du Chevaleret 75013 Paris, menozzi@math.jussieu.fr

1 Introduction

1.1 Statement of the problem

Let the $\mathbb{R}^d$-valued process $(X_t)_{t\geqslant 0}$ satisfy the dynamics

$$X_t = x + \int_0^t b(s,X_s)\mathrm{d}s + \int_0^t B\sigma(s,X_s)\mathrm{d}W_s, \tag{1.1}$$

where $(W_t)_{t\geqslant 0}$ is a d'-dimensional ($d' \leqslant d$) standard Brownian motion defined on a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbb{P})$ satisfying the usual assumptions. The matrix $B = \begin{pmatrix} \mathbf{I_{d'\times d'}} \\ \mathbf{0_{(d-d')\times d'}} \end{pmatrix}$ is the embedding matrix from $\mathbb{R}^{d'}$ into $\mathbb{R}^d$. The coefficients $b : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d, \sigma : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d'} \otimes \mathbb{R}^{d'}$ are assumed to be Lipschitz continuous in space, 1/2-Hölder continuous in time so that there exists a unique strong solution to (1.1).

Let us fix $T > 0$ and introduce for $(t,x) \in [0,T] \times \mathbb{R}^d$, $Q(t,x) := \mathbb{E}[f(T, X_T^{t,x})]$, where f is a measurable function, bounded in time and with polynomial growth in space. The numerical approximation of $Q(t,x)$ appears in many applicative fields. In mathematical finance, $Q(t,x)$ can be related to the price of an option when the underlying asset follows the dynamics (1.1). In this framework we consider two important cases:

(a) If $d = d'$, $Q(t,x)$ corresponds to the price at time t when $X_t = x$ of the vanilla option with maturity T and pay-off f.

(b) If $d' = d/2$, $b(x) = \begin{pmatrix} b_1(x) \\ b_2(x) \end{pmatrix}$ where $b_1(x) \in \mathbb{R}^{d'}, b_2(x) = (x_1, \cdots, x_{d'})^*$, $Q(t,x)$ corresponds to the price of an Asian option.

It is also well known, see e.g. Friedman [8], that $Q(t,x)$ is the Feynman-Kac representation of the solution of the parabolic PDE

$$\begin{cases} \partial_t Q(t,x) + LQ(t,x) = 0,\ (t,x) \in [0,T) \times \mathbb{R}^d, \\ Q(T,x) = f(T,x),\ x \in \mathbb{R}^d, \end{cases} \tag{1.2}$$

where L stands for the infinitesimal generator of (1.1). Hence, the quantity $Q(t,x)$ can also be related to problems of heat diffusion with Cauchy boundary conditions (case (a)) or to kinetic systems (case (b)).

The natural probabilistic approximation of $Q(t,x)$ consists in considering the Monte Carlo algorithm. This approach is particularly relevant compared to deterministic methods if the dimension d is large. To this end we introduce some discretization schemes. For case (a) we consider the Euler scheme with time step $\Delta := T/N$, $N \in \mathbb{N}^*$. Set $\forall i \in \mathbb{N}$, $t_i = i\Delta$ and for $t \geqslant 0$, define $\phi(t) = t_i$ for $t_i \leqslant t < t_{i+1}$. The Euler scheme writes

$$X_t^\Delta = x + \int_0^t b(\phi(s), X_{\phi(s)}^\Delta)\mathrm{d}s + \int_0^t \sigma(\phi(s), X_{\phi(s)}^\Delta)\mathrm{d}W_s. \tag{1.3}$$

For case (b) we define

$$X_t^\Delta = x + \int_0^t \begin{pmatrix} b_1(\phi(s), X_{\phi(s)}^\Delta) \\ (X_s^\Delta)^{1,d'} \end{pmatrix} \mathrm{d}s + \int_0^t B\sigma(\phi(s), X_{\phi(s)}^\Delta)\mathrm{d}W_s, \tag{1.4}$$

where $(X_s^\Delta)^{1,d'} := \big((X_s^\Delta)^1, \cdots, (X_s^\Delta)^{d'}\big)^*$. Equation (1.4) defines a completely simulatable scheme with Gaussian increments. On every time step, the last d' components are the integral of a Gaussian process.

The weak error for the above problems has been widely investigated in the literature. Under suitable assumptions on the coefficients b, σ and f (namely smoothness) it is shown in Talay and Tubaro [24] that $E_D(\Delta) := \mathbb{E}_x[f(T, X_T^\Delta)] - \mathbb{E}_x[f(T, X_T)] = C\Delta + O(\Delta^2)$. Bally and Talay [1] then extended this result to the case of bounded measurable functions f in a hypoelliptic setting for time homogeneous coefficients b, σ. Also, still for time homogeneous coefficients, similar expansions have been derived for the difference of the densities of the process and the discretization scheme, see Konakov and Mammen [14] in case (a), Konakov *et al.* [15] in case (b) for a uniformly elliptic diffusion coefficient $\sigma\sigma^*$, and eventually Bally and Talay [2] for a hypoelliptic diffusion and a slight modification of the Euler scheme. The constant C in the above development involves the derivatives of Q and therefore depends on f, b, σ, x.

The expansion of $E_D(\Delta)$ gives a good control on the impact of the discretization procedure of the initial diffusion, and also permits to improve the convergence rate using e.g. Richardson-Romberg extrapolation (see [24]). Anyhow, to have a global sharp control of the numerical procedure it remains to consider the quantities

$$E_{MC}(M, \Delta) = \frac{1}{M}\sum_{i=1}^{M} f(T, (X_T^\Delta)^i) - \mathbb{E}_x\left[f(T, X_T^\Delta)\right]. \tag{1.5}$$

In the previous quantities M stands for the number of independent samples in the Monte Carlo algorithm and $\big((X_t^\Delta)^i_{t\geqslant 0}\big)_{i\in[\![1,M]\!]}$ are independent sample paths. Indeed, the global error associated to the Monte Carlo algorithm writes:

$$E(M, \Delta) = E_D(\Delta) + E_{MC}(M, \Delta),$$

where $E_D(\Delta)$ is the *discretization error* and $E_{MC}(M, \Delta)$ is the pure *Monte Carlo* error.

The convergence of $E_{MC}(M, \Delta)$, to 0 when $M \to \infty$ is ensured under the above assumptions on f by the strong law of large numbers. A speed of convergence can also be derived from the central limit theorem, but these results are asymptotic, i.e. they hold for a sufficiently large M. On the other hand, a non asymptotic result is provided by the Berry-Esseen Theorem that compares the distribution function of the normalized Monte Carlo error to the distribution function of the normal law at order $O(M^{-1/2})$.

In the current work we are interested in giving, for Lipschitz continuous in space functions f, non asymptotic error bounds for the quantity $E_{MC}(M, \Delta)$. Similar issues had previously been studied by Malrieu and Talay [18]. In that work, the authors investigated the concentration properties of the Euler scheme and obtained Logarithmic Sobolev inequalities, that imply Gaussian concentration see e.g. Ledoux [17], for multi-dimensional Euler schemes with constant diffusion coefficients. Their goal was in some sense different than ours since they were mainly interested in ergodic simulations. In that framework we also mention the recent work of Joulin and Ollivier for Markov chains [11].

Our strategy is here different. We are interested in the approximation of $Q(t,x),\ t \leqslant T$ where $T > 0$ is fixed. It turns out that the log-Sobolev machinery is in some sense too rigid and too ergodic oriented. Also, as far as approximation schemes are concerned it seems really difficult to obtain log-Sobolev inequalities in dimension greater or equal than two without the constant diffusion assumption, see [18]. Anyhow, under suitable assumptions on b, σ (namely uniform ellipticity of $\sigma\sigma^*$ and mild space regularity), the discretization schemes (1.3), (1.4) can be shown to have a density admitting a Gaussian upper bound. From this *a priori* control we can modify Herbst's argument to obtain an expected Gaussian concentration as well as the tensorization property (see [17]) that will yield for $r > 0$ and a Lipschitz continuous in space f, $\mathbb{P}\big[|E_{MC}(M,\Delta)| \geqslant r + \delta\big] \leqslant 2\exp(-\frac{M}{\alpha(T)}r^2)$ for $\alpha(T) > 0$ independent of M uniformly in $\Delta = T/N$. Here $\delta \geqslant 0$ is a bias term (independent of M) depending on the constants appearing in the Gaussian domination (see Theorem 2.1) and on the Wasserstein distance between the law of the discretization scheme and the Gaussian upper bound. We also prove that a Gaussian lower bound holds true for the density of the scheme. Hence, the Gaussian concentration is sharp, i.e. for a function f with suitable non vanishing behavior at infinity, the concentration is at most Gaussian, i.e. $\mathbb{P}\big[|E_{MC}(M,\Delta)| \geqslant r - \bar{\delta}\big] \geqslant 2\exp(-\frac{M}{\bar{\alpha}(T)}r^2)$, for r large enough, $\bar{\delta}$ depending on f, and the Gaussian upper and lower bounds, $\bar{\alpha}(T) > 0$ independent of M uniformly in $\Delta = T/N$.

The paper is organized as follows, we first give our standing assumptions and some notations in Section 1.2. We state our main results in Section 2. Section 3 is dedicated to concentration properties and non asymptotic Monte Carlo bounds for random variables whose law admits a density dominated by a probability density satisfying a log-Sobolev inequality. We prove our main deviations results at the end of that section as well. In Section 4 we show how to obtain the previously mentioned Gaussian bounds in the two cases introduced above. The main tool for the upper bound is a discrete parametrix representation of Mc Kean-Singer type for the density of the scheme, see [19] and Konakov and Mammen [13] or [14]. The lower bound is then derived through suitable chaining arguments adapted to our non Markovian setting.

1.2 Assumptions and Notations

We first specify some assumptions on the coefficients. Namely, we assume:

(UE) The diffusion coefficient is uniformly elliptic. There exists $\lambda_0 \geqslant 1$ s.t. for $(t,x,\xi) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^{d'}$ we have $\lambda_0^{-1}|\xi|^2 \leqslant \langle a(t,x)\xi,\xi\rangle \leqslant \lambda_0|\xi|^2$ where $a(t,x) := \sigma\sigma^*(t,x)$, and $|.|$ stands for the Euclidean norm.

(SB) The drift b is bounded and the diffusion coefficient σ is uniformly η-Hölder continuous in space, $\eta > 0$, uniformly in time. That is there exists $L_0 > 0$ s.t.

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |b(t,x)| + \sup_{t\in[0,T],(x,y)\in\mathbb{R}^{2d},\ x\neq y} \frac{|\sigma(t,x)-\sigma(t,y)|}{|x-y|^\eta} \leqslant L_0.$$

Throughout the paper we assume that **(UE)**, **(SB)** are in force.

In the following we will denote by C a generic positive constant that can depend on $L_0, \lambda_0, \eta, d, T$. We reserve the notation c for constant depending on L_0, λ_0, η, d but not on T. In particular the constants c, C are uniform w.r.t the discretization parameter $\Delta = T/N$ and eventually the value of both c, C may change from line to line.

To establish concentration properties, we will work with the class of Lipschitz continuous functions $F : \mathbb{R}^d \to \mathbb{R}$ satisfying $|\nabla F|_\infty = \text{esssup}_x |\nabla F(x)| \leqslant 1$ where $|\nabla F|$ denotes the Euclidean norm of the gradient ∇F, defined almost everywhere, of F. From now on, for a given function F (possibly not Lipschitz), ∇F stands for the usual "weak" gradient.

For a given probability measure μ on $\mathbb{R}^d$ we write $\mu(F)$ or $\mathbb{E}_\mu[F(X)]$ for the expectation of $F(X)$ where X is a random variable with law μ.

Denote now by S^{d-1} the unit sphere of $\mathbb{R}^d$. For $z \in \mathbb{R}^d \backslash \{0\}$, $\pi_{S^{d-1}}(z)$ stands for the uniquely defined projection on S^{d-1}. For given $\rho_0 > 0$, $\beta > 0$, we introduce the following growth assumption in space for F in the above class of functions:

$(\mathbf{G}_{\rho_0,\beta})$ There exists $A \subset S^{d-1}$ such that

$$\forall y \in \mathbb{R}^d \backslash B(\rho_0), \pi_{S^{d-1}}(y) \in A, \; y_0 := \rho_0 \pi_{S^{d-1}}(y), F(y) - F(y_0) \geqslant \beta |y - y_0|,$$

with A of non empty interior and $|A| \geqslant \varepsilon > 0$ for $d \geqslant 2$ ($|.|$ standing here for the Lebesgue measure of S^{d-1}), and $A \subset \{-1, 1\}$ for $d = 1$. In the above equation $B(\rho_0)$ stands for the Euclidean ball of $\mathbb{R}^d$ of radius ρ_0, and $\langle ., . \rangle$ denotes the scalar product in $\mathbb{R}^d$.

Remark 1.1. *The above assumption simply means that for $|y| \geqslant \rho_0$ the graph of F stays above a given hyperplane. In particular, for all $z \in A, F(rz) \underset{r \to +\infty}{\to} +\infty$.*

The bounds of the quantities $E_{MC}(M, \Delta)$ will be established for real valued functions f that are uniformly Lipschitz continuous in space and measurable bounded in time, such that for a fixed T, $F(.) := f(T, .)$ will be Lispchitz continuous satisfying $|\nabla F|_\infty \leqslant 1$. Moreover, for the lower bounds, we will suppose that the above F satisfies $(\mathbf{G}_{\rho_0,\beta})$.

2 Results

Before dealing with the numerical schemes, let us specify that under **(UE)** and **(SB)** the nature of the diffusion (1.1) is very different in cases (a) and (b).

- In case (a), we are in the framework of uniformly elliptic diffusion processes under which Aronson's estimates are well understood (cf. Sheu [22]).
- Case (b) is degenerate. If the coefficients are Lipschitz, the diffusion satifies a weak Hörmander assumption that guarantees the existence of the density (see eg. [20]). In this framework, Aronson's estimates have been investigated more recently in [6] and extended to the case of bounded drift b_1 and η-Hölder diffusion coefficient σ for $\eta > 1/2$.

One of the main results of the paper (Theorem 2.1) is to prove that similar estimates hold for the discretization schemes (1.3) and (1.4) without any restriction on η in case (b).

Let us first justify that under the assumptions **(UE)**, **(SB)**, the discretization schemes admit a density. For all $x \in \mathbb{R}^d$, $0 \leqslant j < j' \leqslant N, A \in \mathscr{B}(\mathbb{R}^d)$ (where $\mathscr{B}(\mathbb{R}^d)$ stands for the Borel σ-field of $\mathbb{R}^d$) we get

$$\mathbb{P}\big[X^\Delta_{t_{j'}} \in A \big| X^\Delta_{t_j} = x\big] = \int_{(\mathbb{R}^d)^{j'-j-1} \times A} p^\Delta(t_j, t_{j+1}, x, x_{j+1}) p^\Delta(t_{j+1}, t_{j+2}, x_{j+1}, x_{j+2}) \times \cdots \times p^\Delta(t_{j'-1}, t_{j'}, x_{j'-1}, x_{j'}) \mathrm{d}x_{j+1} \mathrm{d}x_{j+2} \cdots \mathrm{d}x_{j'}, \quad (2.1)$$

where the notation $p^\Delta(t_i,t_{i+1},x_i,x_{i+1})$, $i\in[\![0,N-1]\!]$ stands in case (a) for the density at point x_{i+1} of a Gaussian random variable with mean $x_i+b(t_i,x_i)\Delta$ and non degenerated covariance matrix $a(t_i,x_i)\Delta$, whereas in case (b) it stands for the density of a Gaussian random variable with mean $\begin{pmatrix} x_i^{1,d'}+b_1(t_i,x_i)\Delta, \\ x_i^{d'+1,d}+x_i^{1,d'}\Delta+b_1(t_i,x_i)\Delta^2/2 \end{pmatrix}$ and non degenerated as well covariance matrix $\begin{pmatrix} a(t_i,x_i)\Delta & a(t_i,x_i)\Delta^2/2 \\ a(t_i,x_i)\Delta^2/2 & a(t_i,x_i)\Delta^3/3 \end{pmatrix}$, where $\forall y\in\mathbb{R}^d$, $y^{1,d'}=(y^1,\ldots,y^{d'})^*$ and $y^{d'+1,d}=(y^{d'+1},\ldots,y^d)^*$.

Equation (2.1) therefore guarantees the existence of the density for the discretization schemes. From now on, we denote by $p^\Delta(t_j,t_{j'},x,\cdot)$ the transition densities between times t_j and $t_{j'}$, $0\leqslant j<j'\leqslant N$, of the discretization schemes (1.3), (1.4). Let us denote by $\mathbb{P}_x$ (resp. $\mathbb{P}_{t_j,x}$, $0\leqslant j<N$) the conditional probability given $\left\{X_0^\Delta=x\right\}$ (resp. $\{X_{t_j}^\Delta=x\}$), so that in particular $\mathbb{P}_x\left[X_T^\Delta\in A\right]=\int_A p^\Delta(0,T,x,x')\mathrm{d}x'$. We have the following Gaussian estimates for the densities of the schemes.

Theorem 2.1 ("Aronson" Gaussian estimates for the discrete Euler scheme)**.** *Assume* **(UE)**, **(SB)**. *There exist constants $c>0, C\geqslant 1$, s.t. for every $0\leqslant j<j'\leqslant N$:*

$$C^{-1}p_{c^{-1}}(t_{j'}-t_j,x,x')\leqslant p^\Delta(t_j,t_{j'},x,x')\leqslant Cp_c(t_{j'}-t_j,x,x'),\tag{2.2}$$

where for all $0\leqslant s<t\leqslant T$, in case (a), $p_c(t-s,x,x'):=\left(\frac{c}{2\pi(t-s)}\right)^{d/2}\exp\left(-c\frac{|x'-x|^2}{2(t-s)}\right)$ and in case (b)

$$p_c(t-s,x,x'):=\left(\frac{\sqrt{3}c}{2\pi(t-s)^2}\right)^{d/2}\exp\left(-c\left\{\frac{|(x')^{1,d'}-x^{1,d'}|^2}{4(t-s)}+3\frac{|(x')^{d'+1,d}-x^{d'+1,d}-\frac{x^{1,d'}+(x')^{1,d'}}{2}(t-s)|^2}{(t-s)^3}\right\}\right).$$

Note that p_c enjoys the semigroup property, i.e. $\forall 0<s<t$, $\int_{\mathbb{R}^d}p_c(t-s,x,u)p_c(s,u,x')\mathrm{d}u=p_c(t,x,x')$ (see Kolmogorov [12] or [15] for case (b)).

Remark 2.1. *Note that in case (a), the above upper bound can be found in [14] in the case of time homogeneous Lipschitz continuous coefficients. Also, for time dependent coefficients, the upper bound is given by Proposition 3.5 of Gobet and Labart [9] and the lower bound on the diagonal (*i.e. *$x=x'$) can be derived from their Corollary 2.4. Note that since they use Malliavin Calculus, stronger smoothness assumptions on the coefficients are needed. Here, in case (a) our framework is the one of the "standard" PDE assumptions to derive Aronson's estimates for the fundamental solution of non degenerated non-divergence form second order operators, see e.g. Sheu [22]. In particular no regularity in time is needed. In this case, the above theorem provides a technical improvement of existing results.*

On the other hand, in the degenerated hypoelliptic framework of case (b), the result is to our best knowledge new and extends to numerical schemes the results of [6].

In particular, the parametrix techniques we use to prove Theorem 2.1 can be applied to both cases under minimal natural assumptions (see Section 4 for details).

Our second result is the Gaussian concentration of the Monte Carlo error $E_{MC}(M,\Delta)$ defined in (1.5) for a fixed M uniformly in $\Delta=T/N$, $N\geqslant 1$.

Theorem 2.2 (Gaussian concentration). *Assume* **(UE)**, **(SB)**. *For the constants c and C of Theorem 2.1, we have for every $\Delta = T/N$, $N \geqslant 1$, and every Lipschitz continuous function in space and measurable bounded in time $f : \mathbb{R}^d \to \mathbb{R}$ satisfying $\left|\nabla f(T,.)\right|_\infty \leqslant 1$ in (1.5),*

$$\forall r > 0, \quad \forall M \geqslant 1, \quad \mathbb{P}_x\left[\left|E_{MC}(M,\Delta)\right| \geqslant r + \delta_{C,\alpha(T)}\right] \leqslant 2e^{-\frac{M}{\alpha(T)}r^2}, \tag{2.3}$$

with

$$\frac{1}{\alpha(T)} = \begin{cases} \frac{c}{2T} & \text{in case } (a), \\ \frac{c}{2T}\left(1+\frac{3}{T^2}\left(1-\sqrt{1+\frac{T^2}{3}+\frac{T^4}{9}}\right)\right) & \text{in case } (b), \end{cases} \tag{2.4}$$

and $\delta_{C,\alpha(T)} = 2\sqrt{\alpha(T)\log C}$.
Moreover, if $F(.) := f(T,.) \geqslant 0$ satisfies for a given $\rho_0 > 0$ and $\beta > 0$, the growth assumption $(\mathbf{G}_{\rho_0,\beta})$,

$$\forall r > 0, \quad \forall M \geqslant 1, \quad \mathbb{P}_x\left[\left|E_{MC}(M,\Delta)\right| \geqslant r - \bar{\delta}_{c,C,T,f}\right] \geqslant 2\exp\left(-\frac{M}{\bar{\alpha}(T)}\left[\frac{r}{\beta} \vee \rho_0\right]^2\right), \tag{2.5}$$

where $\bar{\delta}_{c,C,T,f} = (1+\sqrt{2})\sqrt{\alpha(T)\log C} + \gamma_{c^{-1},T}(F) + \rho_0\beta - \underline{F}$, $\gamma_{c^{-1},T}(\mathrm{d}x') = p_{c^{-1}}(T,x,x')\mathrm{d}x'$, and $\underline{F} := \inf_{s\in S^{d-1}} F(s\rho_0)$. The constant $\bar{\alpha}(T)^{-1}$ appearing in (2.5) writes in case (a)

$$\frac{1}{\bar{\alpha}(T)} \quad = \quad \bar{\Lambda} + \chi := \begin{cases} \frac{c^{-1}}{2T} + \frac{1}{\rho_0^2}\log\left(\frac{\pi^{d/2}C}{K(d,A)}\right)_+ & \text{for } d \text{ even}, \\ \frac{c^{-1}\theta}{2T} + \frac{1}{\rho_0^2}\log\left(\frac{\pi^{d/2}C}{\arccos(\theta^{-1/2})K(d,A)}\right)_+ & \text{for } d \text{ odd}, \ \theta \in (1,+\infty), \end{cases}$$

where for all $d \in \mathbb{N}^*$, $A \subset S^{d-1}$ appearing in $(\mathbf{G}_{\rho_0,\beta})$,

$$K(d,A) = \begin{cases} \frac{|A|(d/2-1)!}{2}, \ d \text{ even}, \\ \frac{|A|\prod_{j=1}^{\frac{d-1}{2}}(j-1/2)}{\pi^{1/2}}, \ d \text{ odd}. \end{cases} \tag{2.6}$$

In case (b), d is even and

$$\frac{1}{\bar{\alpha}(T)} = \bar{\Lambda} + \chi := \frac{c^{-1}}{2T}\left(1+\frac{3}{T^2}\left[1+\sqrt{1+\frac{T^2}{3}+\frac{T^4}{9}}\right]\right) + \frac{1}{\rho_0^2}\log\left(\frac{\left(\frac{\pi}{\sqrt{3}T}\right)^{d/2}[T^2+3(1+\sqrt{1+\frac{T^2}{3}+\frac{T^4}{9}})]^{d/2}C}{K(d,A)}\right)_+ .$$

From Theorem 2.1 and our current assumptions on f, we can deduce from the central limit theorem that $M^{1/2}E_{MC}(M,\Delta) \overset{(\text{law})}{\underset{M}{\to}} \mathcal{N}(0,\sigma^2(f,\Delta))$, $\sigma^2(f,\Delta) := \mathbb{E}_x[f(X_T^\Delta)^2] - \mathbb{E}_x[f(X_T^\Delta)]^2$. From this asymptotic regime, we thus derive that for large M the typical deviation rate r (i.e. the size of the confidence interval) in (2.3) has order $c\sigma(f,\Delta)M^{-1/2}$ where for a given threshold $\alpha \in (0,1)$, $c := c(\alpha)$ can be deduced from the inverse of the Gaussian distribution function. In other words, r

is typically small for large M. On the other hand, we have a systematic bias $\delta_{C,\alpha(T)}$, independently of M. In whole generality, this bias is inherent to the concentration arguments used to derive the above bounds, see Section 3, and cannot be avoided. Hence, those bounds turn out to be particularly relevant to derive non asymptotic confidence intervals when r and $\delta_{C,\alpha(T)}$ have the same order. In particular, the parameter M is not meant to go to infinity. This kind of result can be useful if for instance it is particularly heavy to simulate the underlying Euler scheme and that only a relatively small number M of samples is reasonably allowed. On the other hand, the smaller T is the bigger M can be. Precisely, one can prove that the constant C of Theorem 2.1 is bounded by $\bar{c}\exp(\bar{c}L_0^2 T)$ (see Section 4). Hence from (2.4), we have $\delta_{C,\alpha(T)} = O(T^{1/2})$ for T small.

Remark 2.2. *For the lower bound, the "expected" value for $\bar{\alpha}(T)^{-1}$ would be $\bar{\lambda}$ corresponding to the largest eigenvalue of one half the inverse of the covariance matrix of the random variable with density $p_{c^{-1}}(T,x,.)$ appearing in the lower bound of Theorem 2.1. There are two corrections with respect to this intuitive approach. First, there is in case (a) an additional multiplicative term $\theta > 1$ (that can be optimized) when d is odd. This correction is unavoidable for $d = 1$, anyhow for odd $d > 1$, it can be avoided up to an additional additive factor like the above χ (see the proof of Proposition 3.3 for details). We kept this presentation to be homogeneous for all odd dimensions.*

Also, an additive correction (or penalty) factor χ appears. It is mainly due to our growth assumption $(\mathbf{G}_{\rho_0,\beta})$.Observe anyhow that, for given $T > 0, C \geqslant 1, \varepsilon > 0$ s.t. $|A| \geqslant \varepsilon$, if the dimension d is large enough, by definition of $K(d,A)$, we have $\chi = 0$. Still, for $d = 1$ (which can only occur in case (a)) we cannot avoid the correction factor χ.

Remark 2.3. *Let us also specify that in the above definition of χ, ρ_0 is not meant to go to zero, even though some useful functions like $|.|$ satisfy $(\mathbf{G}_{\rho_0,\mathbf{1}})$ with any $\rho_0 > 0$. Actually, the bound is particularly relevant in "large regimes", that is when r/β is not assumed to be small. Also, we could replace in the above definition of χ, ρ_0 by $R > 0$ as soon as $r/\beta \geqslant R$. In particular, if F satisfies $(\mathbf{G}_{\rho_0,\beta})$, for $R \geqslant \rho_0$ it also satisfies $(\mathbf{G}_{R,\beta})$. We gave the statement with ρ_0 in order to be uniform w.r.t. the threshold ρ_0 appearing in the growth assumption of F but the correction term can be improved in function of the deviation factor r/β.*

Remark 2.4. *Note that under* **(UE)**, **(SB)**, *in case (a), the martingale problem in the sense of Stroock and Varadhan is well posed for equation (1.1), see Theorem 7.2.1 in [23]. Also, from Theorem 2.1 and the estimates of Section 4, one can deduce that the unique weak solution of the martingale problem has a smooth density that satisfies Aronson like bounds. The well-posedness of the martingale problem in case (b) remains to our best knowledge an open question and will concern further research.*

Remark 2.5. *In case (b), the concentration regime in the above bounds highly depends on T. Since the two components do not have the same scale we have that, in short time, the concentration regime is the one of the non degenerated component in the upper bound (resp. of the degenerated component in the lower bound). For large T, it is the contrary.*

We now consider an important case for applications in case (b). Namely, in kinetic models (resp. in financial mathematics) it is often useful to evaluate the expectation of functions that involve the difference of the first component and its normalized average (which corresponds to a time normalization of the second component). This allows to compare the velocity (resp. the price) at a given time T and the averaged velocity (resp. averaged price) on the associated time interval. Obviously, the normalization is made so that the two components have time-homogeneous scales. We have the following result.

Corollary 2.1. *In case (b), if f in (1.5) writes $f(T,x) = g(T,\mathbb{T}_T^{-1}x)$ where $\mathbb{T}_T^{-1} = \begin{pmatrix} \mathbf{I_{d'\times d'}} & \mathbf{0_{d'\times d'}} \\ \mathbf{0_{d'\times d'}} & T^{-1}\mathbf{I_{d'\times d'}} \end{pmatrix}$ and g is a Lipschitz continuous function in space and measurable bounded in time satisfying $\left|\nabla g(T,.)\right|_\infty \leqslant 1$ then we have for every $\Delta = T/N,\ N \geqslant 1$,*

$$\forall M \geqslant 1, \quad \mathbb{P}_x\left[\left|E_{MC}(M,\Delta)\right| \geqslant r + \delta_{C,\alpha(T)}\right] \leqslant 2e^{-\frac{M}{\alpha(T)}r^2},$$

with $\alpha(T)^{-1} = (4-\sqrt{13})\frac{c}{T}$ and $\delta_{C,\alpha(T)} = 2\sqrt{\alpha(T)\log C}$.

A lower bound could be derived similarly to Theorem 2.2.

The proof of Theorems 2.1 and 2.2 (as well as Corollary 2.1) are respectively postponed to Sections 4.4 and 3.3.

3 Gaussian concentration and non asymptotic Monte Carlo bounds

3.1 Gaussian concentration - Upper bound

We recall that a probability measure γ on $\mathbb{R}^d$ satisfies a logarithmic Sobolev inequality with constant $\alpha > 0$ if for all $f \in H^1(\mathrm{d}\gamma) := \{g \in L^2(\mathrm{d}\gamma) : \int \left|\nabla g\right|^2 \mathrm{d}\gamma < +\infty\}$ such that $f \geqslant 0$, one has

$$\mathrm{Ent}_\gamma(f^2) \leqslant \alpha \int \left|\nabla f\right|^2 \mathrm{d}\gamma, \qquad (\mathbf{LSI}_\alpha)$$

where $\mathrm{Ent}_\gamma(\phi) = \int \phi \log(\phi)\mathrm{d}\gamma - \left(\int \phi \mathrm{d}\gamma\right)\log\left(\int \phi \mathrm{d}\gamma\right)$ denotes the entropy of the measure γ. In particular, we have the following result (see [17] Section 2.2 eq. (2.17)).

Proposition 3.1. *Let V be a $\mathscr{C}^2$ convex function on $\mathbb{R}^d$ with $\mathrm{Hess}V \geqslant \lambda \mathbf{I_{d\times d}}$, $\lambda > 0$ and such that e^{-V} is integrable with respect to the Lebesgue measure. Let $\gamma(\mathrm{d}x) = \frac{1}{Z}e^{-V(x)}\mathrm{d}x$ be a probability measure (Gibbs measure). Then γ satisfies a logarithmic Sobolev inequality with constant $\alpha = \frac{2}{\lambda}$.*

Throughout this section we consider a probability measure μ with density m with respect to the Lebesgue measure λ_K on $\mathbb{R}^K$ (here we have in mind $K = d$ or $K = Md$, M being the number of Monte Carlo paths). We assume that μ is dominated by a probability measure γ in the following sense

$$\gamma(\mathrm{d}x) = q(x)\mathrm{d}x \text{ satisfies } (\mathbf{LSI}_\alpha) \quad \text{and} \quad \exists \kappa \geqslant 1, \quad \forall x \in \mathbb{R}^K, \quad m(x) \leqslant \kappa q(x). \qquad (\mathbf{H}_{\kappa,\alpha})$$

Proposition 3.2. *Assume that μ and γ satisfy $(\mathbf{H}_{\kappa,\alpha})$. Then for all Lipschitz continuous function $F : \mathbb{R}^K \to \mathbb{R}$ s.t. $|\nabla F|_\infty \leqslant 1$,*

$$\forall r > 0, \quad \mathbb{P}_\mu\Big[F(Y) - \mu(F) \geqslant r + W_1(\mu,\gamma)\Big] \leqslant \kappa e^{-\frac{r^2}{\alpha}},$$

where $W_1(\mu,\gamma) = \sup_{|\nabla F|_\infty \leqslant 1} \left|\mu(F) - \gamma(F)\right|$ (Wasserstein distance W_1 between μ and γ).

Proof. By the Markov inequality, one has for every $\lambda > 0$,

$$\mathbb{P}_\mu\Big[F(Y) - \mu(F) \geqslant r + W_1(\mu,\gamma)\Big] \leqslant e^{-\lambda(\mu(F)+r+W_1(\mu,\gamma))}\mathbb{E}_\mu\Big[e^{\lambda F(Y)}\Big], \tag{3.1}$$

and by $(\mathbf{H}_{\kappa,\alpha})$, $\mathbb{E}_\mu\big[e^{\lambda F(Y)}\big] \leqslant \kappa \mathbb{E}_\gamma\big[e^{\lambda F(Y)}\big]$. Since γ satisfies a logarithmic Sobolev inequality with constant $\alpha > 0$, the Herbst argument (see e.g. Ledoux [17] section 2.3) gives

$$\mathbb{E}_\gamma\Big[e^{\lambda F(Y)}\Big] \leqslant e^{\lambda\gamma(F)+\frac{\alpha}{4}\lambda^2},$$

so that $\mathbb{E}_\mu\big[e^{\lambda F(Y)}\big] \leqslant \kappa e^{\lambda\mu(F)+\frac{\alpha}{4}\lambda^2+\lambda(\gamma(F)-\mu(F))}$ and

$$\mathbb{E}_\mu\Big[e^{\lambda F(Y)}\Big] \leqslant \kappa e^{\lambda\mu(F)+\frac{\alpha}{4}\lambda^2+\lambda W_1(\mu,\gamma)}, \tag{3.2}$$

since owing to the definition of W_1 one has $W_1(\mu,\gamma) \geqslant \gamma(F) - \mu(F)$. Plugging the above control (3.2) into (3.1) yields

$$\mathbb{P}_\mu\Big[F(Y) - \mu(F) \geqslant r + W_1(\mu,\gamma)\Big] \leqslant \kappa e^{-\lambda r+\frac{\alpha}{4}\lambda^2}.$$

An optimization on λ gives the result. □

Lemma 3.1. *Assume that μ with density m and γ with density q satisfy the domination condition*

$$\exists \kappa \geqslant 1, \quad \forall x \in \mathbb{R}^d, \quad m(x) \leqslant \kappa q(x)$$

and that there exist $(\alpha,\beta_1,\beta_2) \in (\mathbb{R}_+)^3$ such that for all Lipschitz continuous function F satisfying $|\nabla F|_\infty \leqslant 1$ and for all $\lambda > 0$, $\mathbb{E}_\gamma\left[e^{\lambda F(Y)}\right] \leqslant e^{\lambda\gamma(F)+\frac{\alpha}{4}\lambda^2+\beta_1\lambda+\beta_2}$. Then we have, $W_1(\mu,\gamma) \leqslant \beta_1 + \sqrt{\alpha\left(\beta_2+\log(\kappa)\right)}$.

Proof. Recall first that for a non-negative function f, we have the following variational formulation of the entropy:

$$\mathrm{Ent}_\gamma(f) = \sup\left\{\mathbb{E}_\gamma\left[fh\right];\ \mathbb{E}_\gamma\left[e^h\right] \leqslant 1\right\}. \tag{3.3}$$

W.l.o.g. we consider F such that $\mu(F) \geqslant \gamma(F)$. Let $\lambda > 0$ and $h := \lambda F - \lambda\gamma(F) - \frac{\alpha}{4}\lambda^2 - \beta_1\lambda - \beta_2$ so that $\mathbb{E}_\gamma\left[e^h\right] \leqslant 1$ and

$$\mathbb{E}_\mu[h] = \mathbb{E}_\gamma\left[\frac{m}{q}h\right] = \lambda\left(\mu(F)-\gamma(F)\right) - \frac{\alpha}{4}\lambda^2 - \beta_1\lambda - \beta_2.$$

We then have

$$\begin{aligned}\mu(F)-\gamma(F) &= \frac{\alpha}{4}\lambda + \beta_1 + \frac{1}{\lambda}\left(\beta_2 + \mathbb{E}_\gamma\left[\frac{m}{q}h\right]\right),\\ &\overset{(3.3)}{\leqslant} \frac{\alpha}{4}\lambda + \beta_1 + \frac{1}{\lambda}\left(\beta_2 + \mathrm{Ent}_\gamma\left(\frac{m}{q}\right)\right).\end{aligned}$$

An optimization in λ yields

$$\mu(F)-\gamma(F) \leqslant \beta_1 + \sqrt{\alpha\left(\beta_2+\mathrm{Ent}_\gamma\left(\frac{m}{q}\right)\right)}. \tag{3.4}$$

Now using the domination condition, one has $\mathrm{Ent}_\gamma\left(\frac{m}{q}\right) = \int \frac{m}{q}\log\left(\frac{m}{q}\right)\mathrm{d}\gamma \leqslant \log(\kappa)$ and the results follows. □

Remark 3.1. *Note that if γ satisfies an* ($\mathbf{LSI}_\alpha$) *we have $\beta_1 = \beta_2 = 0$ and the result* (3.4) *of Lemma 3.1 reads $W_1(\mu,\gamma) \leqslant \sqrt{\alpha \operatorname{Ent}_\gamma\left(\frac{m}{q}\right)} \leqslant \sqrt{\alpha \log(\kappa)}$. For similar controls concerning the W_2 Wasserstein distance see Theorem 1 of Otto and Villani [21] or Bobkov et al. [4].*

Using the tensorization property of the logarithmic Sobolev inequality we derive the following Corollary.

Corollary 3.1. *Let $Y^1,\dots,Y^M$ be* i.i.d. *$\mathbb{R}^d$-valued random variables with law μ. Assume there exist $\alpha > 0$, $\kappa \geqslant 1$ and γ such that* ($\mathbf{H}_{\kappa,\alpha}$) *holds on $\mathbb{R}^d$. Then, for all Lipschitz continuous function $f : \mathbb{R}^d \to \mathbb{R}$ satisfying $\left|\nabla f\right|_\infty \leqslant 1$, we have*

$$\forall r > 0,\ M \geqslant 1, \quad \mathbb{P}\left[\left|\frac{1}{M}\sum_{k=1}^{M} f(Y^k) - \mathbb{E}_\mu\left[f(Y^1)\right]\right| \geqslant r + \delta_{\kappa,\alpha}\right] \leqslant 2e^{-M\frac{r^2}{\alpha}}, \tag{3.5}$$

with $\delta_{\kappa,\alpha} = 2\sqrt{\alpha\log(\kappa)} \geqslant 0$.

Proof. Let $r > 0$ and $M \geqslant 1$. Clearly, changing f into $-f$, it suffices to prove that

$$\mathbb{P}\left[\frac{1}{M}\sum_{k=1}^{M} f(Y^k) - \mathbb{E}_\mu\left[f(Y^1)\right] \geqslant r + \delta_{\kappa,\alpha}\right] \leqslant e^{-M\frac{r^2}{\alpha}}.$$

By tensorization, the measure $\gamma^{\otimes M}$ satisfies an ($\mathbf{LSI}_\alpha$) with the same constant α as γ, and then the probabilities $\mu^{\otimes M}$ and $\gamma^{\otimes M}$ satisfy ($\mathbf{H}_{\kappa^M,\alpha}$) on $\mathbb{R}^K, K = Md$. In this case, Lemma 3.1 gives $\sqrt{M}\delta_{\kappa,\alpha} \geqslant W_1(\mu^{\otimes M},\gamma^{\otimes M}) + \sqrt{M\alpha\log(\kappa)}$ and then

$$\begin{aligned}&\mathbb{P}\left[\frac{1}{M}\sum_{k=1}^{M} f(Y^k) - \mathbb{E}_\mu\left[f(Y^1)\right] \geqslant r + \delta_{\kappa,\alpha}\right]\\ &\qquad\leqslant \mathbb{P}\left[\frac{1}{\sqrt{M}}\sum_{k=1}^{M} f(Y^k) - \sqrt{M}\mathbb{E}_\mu\left[f(Y^1)\right] \geqslant \sqrt{M}\left(r + \sqrt{\alpha\log(\kappa)}\right) + W_1(\mu^{\otimes M},\gamma^{\otimes M})\right].\end{aligned}$$

Applying Proposition 3.2 with the measures $\mu^{\otimes M}$ and $\gamma^{\otimes M}$, the function $F(x_1,\dots,x_M) = \frac{1}{\sqrt{M}}\sum_{k=1}^{M} f(x_k)$ (which satisfies $|\nabla F|_\infty \leqslant 1$) and $\tilde{r} = \sqrt{M}(r + \sqrt{\alpha\log(\kappa)})$ we obtain

$$\mathbb{P}\left[\frac{1}{M}\sum_{k=1}^{M} f(Y^k) - \mathbb{E}_\mu\left[f(Y^1)\right] \geqslant r + \delta_{\kappa,\alpha}\right] \leqslant \kappa^M e^{-M\frac{\left(r+\sqrt{\alpha\log(\kappa)}\right)^2}{\alpha}},$$

and we easily conclude. □

Remark 3.2. *Note that the term $\delta_{\kappa,\alpha}$ can be seen as a penalty term due on the one hand to the transport between μ and γ, and on the other hand to the explosion of the domination constant κ^M between $\mu^{\otimes M}$ and $\gamma^{\otimes M}$ when M tends to infinity. We emphasize that the bias $\delta_{\kappa,\alpha}$ is independent of M. Hence, the result below is especially relevant when r and $\delta_{\kappa,\alpha}$ have the same order. In particular, the non-asymptotic confidence interval given by* (3.5) *cannot be compared to the asymptotic confidence interval deriving from the central limit theorem whose size has order $O(M^{-1/2})$.*

Remark 3.3. *Note that to obtain the non-asymptotic bounds of the Monte Carlo procedure* (3.5)*, we successively used the concentration properties of the reference measure γ, the control of the distance $W_1(\mu,\gamma)$ given by the variational formulation of the entropy (see Lemma 3.1) and the tensorization property of the functional inequality satisfied by γ. The same arguments can therefore be applied to a reference measure γ satisfying a Poincaré inequality.*

3.2 Gaussian concentration - Lower bound

Concerning the previous deviation rate of Proposition 3.2, a natural question consists in understanding whether it is sharp or not. Namely, for a given function f satisfying suitable growth conditions at infinity, otherwise we cannot see the asymptotic growth, do we have a lower bound of the same order, i.e. with Gaussian decay at infinity? The next proposition gives a positive answer to that question.

Proposition 3.3. *Let $f:\mathbb{R}^d\to\mathbb{R}_+$ be a Lipschitz continuous function satisfying $\left|\nabla f\right|_\infty \leqslant 1$ and assumption $(\mathbf{G}_{\rho_0,\beta})$ for given $\rho_0,\beta>0$.*

For a $\mathscr{C}^2$ function V on $\mathbb{R}^d$ such that e^{-V} is integrable with respect to λ_d and s.t. $\exists\bar\lambda\geqslant 1,\ \bar\lambda\mathbf{I_{d\times d}}\geqslant$ $\mathrm{Hess}(V)\geqslant 0$, let $\gamma(\mathrm{d}x)=e^{-V(x)}Z^{-1}\mathrm{d}x$ be the associated Gibbs probability measure. We assume that $\exists\kappa\geqslant 1$ s.t. for $|x|\geqslant\rho_0$ the measures $\mu(\mathrm{d}x)=m(x)\mathrm{d}x$ and $\gamma(\mathrm{d}x)$ satisfy

$$m(x)\geqslant\kappa^{-1}e^{-V(x)}Z^{-1}.$$

Let $\bar\Lambda:=\frac{\bar\lambda}{2}+\frac{\sup_{s\in S^{d-1}}|V(s\rho_0)|}{\rho_0^2}+\frac{\sup_{s\in S^{d-1}}|\nabla V(s\rho_0)|}{\rho_0}$.

We have

$$\forall r>0,\quad \mathbb{P}_\mu\big[f(Y)-\mu(f)\geqslant r-(W_1(\mu,\gamma)+\delta(f,\gamma))\big]$$
$$\geqslant\begin{cases}\frac{K(d,A)}{Z\bar\Lambda^{d/2}\kappa}\exp\left(-\bar\Lambda\left[\frac{r}{\beta}\vee\rho_0\right]^2\right),\ d\text{ even},\\ \frac{\arccos(\theta^{-1/2})K(d,A)}{Z\bar\Lambda^{d/2}\kappa}\exp\left(-\theta\bar\Lambda\left[\frac{r}{\beta}\vee\rho_0\right]^2\right),\ \forall\theta>1,\ d\text{ odd},\end{cases}$$

with $\delta(f,\gamma)=\gamma(f)+\beta\rho_0-\underline{f}$, $\underline{f}:=\inf_{s\in S^{d-1}}f(s\rho_0)$, and $K(d,A)$ defined in (2.6) *where $A\subset S^{d-1}$ appears in $(\mathbf{G}_{\rho_0,\beta})$.*

Proof. Set $E:=\{Y\in A\times[\rho_0,\infty)\}$. Here we use the convention that for $d=1$, $A\times[\rho_0,+\infty)\subset(-\infty,-\rho_0]\cup[\rho_0,+\infty)$. Write now

$$\begin{aligned}\mathbb{P}_\mu\big[f(Y)-\mu(f)\geqslant r-(W_1(\mu,\gamma)+\delta(f,\gamma))\big]&\geqslant\mathbb{P}_\mu\big[f(Y)-\mu(f)\geqslant r-(W_1(\mu,\gamma)+\delta(f,\gamma)),E\big]\\&\geqslant\kappa^{-1}\mathbb{P}_\gamma\Big[f(Y)\geqslant r-\beta\rho_0+\underline{f},E\Big]:=\kappa^{-1}\mathscr{P}.\end{aligned}\tag{3.6}$$

Denoting $Y_0=\rho_0\pi_{S^{d-1}}(Y)$, we have

$$\begin{aligned}\mathscr{P}&\geqslant\mathbb{P}_\gamma\Big[f(Y_0)+\big(f(Y)-f(Y_0)\big)\geqslant r-\beta\rho_0+\underline{f},E\Big],\\&\overset{(\mathbf{G}_{\rho_0,\beta})}{\geqslant}\mathbb{P}_\gamma\Big[\beta|Y-Y_0|\geqslant r-\beta\rho_0+\underline{f}-f(Y_0),E\Big]\geqslant\mathbb{P}_\gamma\left[|Y|\geqslant\frac{r-\beta\rho_0}{\beta}+|Y_0|,E\right]\\&\geqslant\mathbb{P}_\gamma\left[|Y|\geqslant\frac{r}{\beta}\vee\rho_0,\pi_{S^{d-1}}(Y)\in A\right].\end{aligned}\tag{3.7}$$

Write

$$\mathscr{P} \geqslant \int_A \sigma(\mathrm{d}s) \int_{\rho_0 \vee \frac{r}{\beta}}^{+\infty} \mathrm{d}\rho \rho^{d-1} \exp(-V(s\rho)) Z^{-1},$$

where $\sigma(\mathrm{d}s)$ stands for the Lebesgue measure of S^{d-1}. Now, $\mathrm{Hess}(V) \leqslant \bar{\lambda} \mathbf{I_{d\times d}}$ yields $\forall \rho \geqslant \rho_0 \vee \frac{r}{\beta}$, $|V(s\rho)|/\rho^2 \leqslant \bar{\Lambda}$, $\bar{\Lambda} := \frac{\bar{\lambda}}{2} + \frac{\sup_{s\in S^{d-1}} |V(s\rho_0)|}{\rho_0^2} + \frac{\sup_{s\in S^{d-1}} |\nabla V(s\rho_0)|}{\rho_0}$ and therefore

$$\begin{aligned} \mathscr{P} &\geqslant |A| \int_{\rho_0 \vee \frac{r}{\beta}}^{+\infty} \mathrm{d}\rho \rho^{d-1} \exp(-\bar{\Lambda}\rho^2) Z^{-1} \geqslant \frac{|A|}{Z(2\bar{\Lambda})^{d/2}} \int_{(\rho_0 \vee \frac{r}{\beta})(2\bar{\Lambda})^{1/2}}^{+\infty} \mathrm{d}\rho \rho^{d-1} \exp(-\frac{\rho^2}{2}) \\ &= \frac{|A|}{Z(2\bar{\Lambda})^{d/2}} Q_d\left((\rho_0 \vee \frac{r}{\beta})(2\bar{\Lambda})^{1/2}\right). \end{aligned} \tag{3.8}$$

We now have the following explicit expression:

$$\forall x > 0,\ Q_d(x) := \exp(-\frac{x^2}{2}) M(d,x),$$

$$M(d,x) := \begin{cases} \sum_{i=0}^{\frac{d}{2}-1} x^{2i} \prod_{j=i+1}^{\frac{d}{2}-1} 2j,\ d \text{ even}, \\ \sum_{i=0}^{\frac{d-1}{2}-1} x^{2i+1} \prod_{j=i+1}^{\frac{d-1}{2}-1} (2j+1) + \prod_{j=0}^{\frac{d-1}{2}-1} (2j+1) \exp(\frac{x^2}{2}) \int_x^{+\infty} \exp(-\frac{\rho^2}{2}) \mathrm{d}\rho,\ d \text{ odd}, \end{cases}$$

with the convention that $\sum_{i=0}^{-1} = 0,\ \forall k \in \mathbb{N},\ \prod_{j=k}^{k-1} j = 1$.

Observe now that $\int_x^{\infty} \exp(-\rho^2/2) d\rho = (2\pi)^{1/2} \mathbb{P}[\mathscr{N}(0,1) \geqslant x] \geqslant (2\pi)^{1/2} \mathbb{P}[Y \in \mathscr{K}, |Y| \geqslant x/\cos(\tilde{\theta})] := (2\pi)^{1/2} \mathscr{Q}(x)$, where $Y \sim \mathscr{N}(\mathbf{0_{2\times 1}}, \mathbf{I_{2\times 2}})$ is a standard bidimensional Gaussian vector and $\mathscr{K} := \{z \in \mathbb{R}^2, \langle z, e_1 \rangle \geqslant \cos(\tilde{\theta})|z|\},\ \tilde{\theta} \in (0, \frac{\pi}{2}),\ e_1 = (1,0)$. Since $\mathscr{Q}(x) = \frac{\tilde{\theta}}{\pi} \exp(-\frac{x^2}{2\cos^2(\tilde{\theta})})$, we derive that

$$Q_d(x) \geqslant \begin{cases} 2^{d/2-1}(d/2-1)! \exp(-\frac{x^2}{2}),\ d \text{ even}, \\ \frac{\tilde{\theta} 2^{d/2}}{\pi^{1/2}} \prod_{j=1}^{\frac{d-1}{2}} (j - \frac{1}{2}) \exp(-\frac{x^2}{2\cos^2(\tilde{\theta})}),\ d \text{ odd}, \end{cases}$$

which plugged into (3.8) yields:

$$\mathscr{P} \geqslant \begin{cases} \frac{K(d,A)}{Z\bar{\Lambda}^{d/2}} \exp\left(-\bar{\Lambda}\left[\frac{r}{\beta} \vee \rho_0\right]^2\right),\ d \text{ even}, \\ \frac{\tilde{\theta} K(d,A)}{Z\bar{\Lambda}^{d/2}} \exp\left(-\frac{\bar{\Lambda}}{\cos^2(\tilde{\theta})}\left[\frac{r}{\beta} \vee \rho_0\right]^2\right),\ d \text{ odd}. \end{cases}$$

□

Corollary 3.2. *Under the assumptions of Proposition 3.3, let $Y^1, \cdots, Y^M$ be* i.i.d. *$\mathbb{R}^d$-valued random variables with law μ. We have $\forall r > 0,\ \forall M \geqslant 1$,*

$$\mathbb{P}\left[\left|\frac{1}{M}\sum_{k=1}^M f(Y^k) - \mathbb{E}_\mu\left[f(Y^1)\right]\right| \geqslant r - (W_1(\mu,\gamma) + \delta(f,\gamma))\right] \geqslant 2\exp\left(-M(\theta\bar{\Lambda} + \chi)\left[\frac{r}{\beta} \vee \rho_0\right]^2\right)$$

where $\chi = \frac{1}{\rho_0^2}\log\left(\frac{Z\bar{\Lambda}^{d/2}\kappa}{K(d,A)}\right)_+$, $\theta = 1$ *for d even, and* $\chi = \frac{1}{\rho_0^2}\log\left(\frac{Z\bar{\Lambda}^{d/2}\kappa}{K(d,A)\arccos(\theta^{-1/2})}\right)_+$, $\theta \in (1,+\infty)$ *for d odd, and with* $K(d,A)$ *defined in* (2.6).

Proof. We only consider d even. By independence of the $((Y)^k)_{k\in[\![1,M]\!]}$, exploiting $\bigcap_{k=1}^M\{f(Y^k) - \mathbb{E}_\mu\left[f(Y^1)\right] \geqslant r-(W_1(\mu,\gamma)+\delta(f,\gamma))\} \subset \{\frac{1}{M}\sum_{k=1}^M f(Y^k) - \mathbb{E}_\mu\left[f(Y^1)\right] \geqslant r-(W_1(\mu,\gamma)+\delta(f,\gamma))\}$, we have

$$\mathbb{P}\left[\frac{1}{M}\sum_{k=1}^{M} f(Y^k)-\mathbb{E}_\mu\left[f(Y^1)\right] \geqslant r-(W_1(\mu,\gamma)+\delta(f,\gamma))\right] \geqslant \left(\frac{K(d,A)}{Z\bar{\Lambda}^{d/2}\kappa}\right)^M \exp\left(-M\bar{\Lambda}\left[\frac{r}{\beta}\vee\rho_0\right]^2\right).$$

For $\chi = \frac{1}{\rho_0^2}\log\left(\frac{Z\bar{\Lambda}^{d/2}\kappa}{K(d,A)}\right)_+$, we thus obtain

$$\mathbb{P}\left[\frac{1}{M}\sum_{k=1}^{M} f(Y^k)-\mathbb{E}_\mu\left[f(Y^1)\right] \geqslant r-(W_1(\mu,\gamma)+\delta(f,\gamma))\right] \geqslant \exp\left(-M(\bar{\Lambda}+\chi)\left[\frac{r}{\beta}\vee\rho_0\right]^2\right),$$

which completes the proof.

□.

3.3 Proofs of Theorem 2.2 and Corollary 2.1

- *Theorem 2.2 - Upper bound* (2.3).

In case (a), the Gaussian probability $\gamma_{c,T}$ with density $p_c(T,x,.)$ defined in Theorem 2.1 satisfies a logarithmic Sobolev inequality with constant $\alpha(T) = \frac{2T}{c}$ (see Proposition 3.1). The result then follows from Theorem 2.1 and Corollary 3.1.

In case (b), $\gamma_{c,T}(\mathrm{d}x') = p_c(T,x,x')\mathrm{d}x' = Z^{-1}e^{-V_{T,x}(x')}\mathrm{d}x'$ where

$$V_{T,x}(x') = c\left(\frac{|(x')^{1,d'}-x^{1,d'}|^2}{4T} + 3\frac{|(x')^{d'+1,d}-x^{d'+1,d}-\frac{x^{1,d'}+(x')^{1,d'}}{2}T|^2}{T^3}\right). \tag{3.9}$$

The Hessian matrix of $V_{T,x}$ satisfies

$$\forall x' \in \mathbb{R}^d, \quad \mathrm{Hess}V_{T,x}(x') = \begin{pmatrix} \frac{2c}{T}\mathbf{I_{d'\times d'}} & \frac{-3c}{T^2}\mathbf{I_{d'\times d'}} \\ \frac{-3c}{T^2}\mathbf{I_{d'\times d'}} & \frac{6c}{T^3}\mathbf{I_{d'\times d'}} \end{pmatrix} \geqslant \lambda\mathbf{I_{d\times d}},$$

with $\lambda = \frac{c}{T} + \frac{3c}{T^3}\left(1-\sqrt{1+\frac{T^2}{3}+\frac{T^4}{9}}\right) > 0$. By Proposition 3.1, the probability $\gamma_{c,T}$ satisfies a logarithmic Sobolev inequality with constant $\alpha(T) = \frac{2T}{c}\frac{1}{1+\frac{3}{T^2}\left(1-\sqrt{1+\frac{T^2}{3}+\frac{T^4}{9}}\right)}$. We still conclude by Theorem 2.1 and Corollary 3.1.

- *Theorem 2.2 - Lower bound* (2.5).

With the notation $p_{c^{-1}}(t-s,x,x') = Z^{-1}e^{-V_{t-s,x}(x')}$, the Hessian of the potential $V_{T,x}$ satisfies $\forall x' \in \mathbb{R}^d$, $\mathrm{Hess}V_{T,x}(x') \leqslant \bar{\lambda}\mathbf{I_{d\times d}}$ where $\bar{\lambda} = \frac{c^{-1}}{T}$ in case (a) and $\bar{\lambda} = \frac{c^{-1}}{T} + \frac{3c^{-1}}{T^3}\left(1+\sqrt{1+\frac{T^2}{3}+\frac{T^4}{9}}\right)$ in case (b). Set $\gamma_{c^{-1},T}(\mathrm{d}x') = p_{c^{-1}}(T,x,x')\mathrm{d}x'$ and $\mu_T(\mathrm{d}x') = p^{\Delta}(0,T,x,x')\mathrm{d}x'$. Since μ_T and $\gamma_{c,T}$ satisfy $(\mathbf{H}_{\kappa,\alpha})$ with $\kappa = C$ and $\alpha = \alpha(T)$ defined in (2.4), the probability μ_T satisfies (3.2), and Lemma 3.1 yields $W_1(\mu_T,\gamma_{c,T}) \leqslant \sqrt{\alpha(T)\log(C)}$. Now, $\gamma_{c^{-1},T}$ and $\gamma_{c,T}$ satisfy $(\mathbf{H}_{\kappa,\alpha})$ with $\kappa = C^2$ and $\alpha = \alpha(T)$. We therefore get from Lemma 3.1, $W_1(\gamma_{c^{-1},T},\gamma_{c,T}) \leqslant \sqrt{2\alpha(T)\log(C)}$. Hence, $W_1(\gamma_{c^{-1},T},\mu_T) \leqslant W_1(\mu_T,\gamma_{c,T}) + W_1(\gamma_{c^{-1},T},\gamma_{c,T}) \leqslant (1+\sqrt{2})\sqrt{\alpha(T)\log(C)}$. Now, by definition of $\bar{\delta}_{c,C,T,f}$ we have $\bar{\delta}_{c,C,T,f} \geqslant W_1(\gamma_{c^{-1},T},\mu_T) + \delta(f,\gamma_{c^{-1},T})$, ($\delta(f,\gamma_{c^{-1},T})$ introduced in Proposition 3.3) and Corollary 3.2 yields

$$\mathbb{P}_x\left[\frac{1}{M}\sum_{k=1}^{M} f(T,(X_T^{\Delta})^k) - \mathbb{E}_x\left[f(T,X_T^{\Delta})\right] \geqslant r - \bar{\delta}_{c,C,T,f}\right] \geqslant \exp\left(-\frac{M}{\bar{\alpha}(T)}\left[\frac{r}{\beta}\vee\rho_0\right]^2\right),$$

where observing that for our Gaussian bounds $\bar{\Lambda} = \frac{\bar{\lambda}}{2}$, and

$$\frac{1}{\bar{\alpha}(T)} = \begin{cases} \frac{\bar{\lambda}}{2} + \chi,\ \chi = \frac{1}{\rho_0^2}\log\left(\frac{2^{-d/2}Z\bar{\lambda}^{d/2}C}{K(d,A)}\right)_+ & \text{for } d \text{ even,}\\ \theta\frac{\bar{\lambda}}{2} + \chi,\ \chi = \frac{1}{\rho_0^2}\log\left(\frac{2^{-d/2}Z\bar{\lambda}^{d/2}C}{K(d,A)\arccos(\theta^{-1/2})}\right)_+ & \text{for } d \text{ odd, } \theta > 1,\end{cases}$$

and $K(d,A)$ defined in (2.6).

Observe now that in case (a), the normalization factor $Z = Z(T,d)$ associated to $p_{c^{-1}}(T,x,.)$ writes $Z = (2\pi cT)^{d/2}$. Hence, recalling that $\bar{\lambda} = (cT)^{-1}$, we obtain in this case

$$\chi = \begin{cases} \frac{1}{\rho_0^2}\log\left(\frac{\pi^{d/2}C}{K(d,A)}\right)_+ & \text{for } d \text{ even,}\\ \frac{1}{\rho_0^2}\log\left(\frac{\pi^{d/2}C}{K(d,A)\arccos(\theta^{-1/2})}\right) & \text{for } d \text{ odd, } \theta > 1.\end{cases}$$

In case (b), we have $Z = \left(\frac{2\pi}{\sqrt{3}}c\right)^{d/2} T^d, \bar{\lambda} = \frac{1}{cT}\left(1 + \frac{3}{T^2}\left[1+\sqrt{1+\frac{T^2}{3}+\frac{T^4}{9}}\right]\right)$ so that $2^{-d/2}Z\bar{\lambda}^{d/2} = \left(\frac{\pi}{\sqrt{3}T}\right)^{d/2}[T^2 + 3(1+\sqrt{1+\frac{T^2}{3}+\frac{T^4}{9}})]^{d/2}$. Eventually, since in case (b) we always have d even, the correction writes

$$\chi = \frac{1}{\rho_0^2}\log\left(\frac{\left(\frac{\pi}{\sqrt{3}T}\right)^{d/2}[T^2+3(1+\sqrt{1+\frac{T^2}{3}+\frac{T^4}{9}})]^{d/2}C}{K(d,A)}\right)_+.$$

This completes the proof. □

- *Proof of Corollary 2.1.*

Note that the random variable $Y_T^{\Delta} = \mathbb{T}_T^{-1}X_T^{\Delta}$ admits the density $p_Y^{\Delta}(T,y,y') = T^{d'}p^{\Delta}(0,T,\mathbb{T}_Ty,\mathbb{T}_Ty')$ with respect to $\lambda_d(\mathrm{d}y')$. By Theorem 2.1 this density is dominated by $(ZT^{d'})^{-1}e^{-V_{T,\mathbb{T}_Ty}(\mathbb{T}_Ty')}$ where $V_{T,x}$ is defined in (3.9). The Hessian of $y' \mapsto V_{T,\mathbb{T}_Ty}(\mathbb{T}_Ty')$ satisfies

$$\forall y' \in \mathbb{R}^d,\quad \mathrm{Hess}V_{T,\mathbb{T}_Ty}(\mathbb{T}_Ty') = \begin{pmatrix} \frac{2c}{T}\mathbf{I}_{\mathbf{d'}\times\mathbf{d'}} & \frac{-3c}{T}\mathbf{I}_{\mathbf{d'}\times\mathbf{d'}}\\ \frac{-3c}{T}\mathbf{I}_{\mathbf{d'}\times\mathbf{d'}} & \frac{6c}{T}\mathbf{I}_{\mathbf{d'}\times\mathbf{d'}}\end{pmatrix} \geqslant \lambda\mathbf{I}_{\mathbf{d}\times\mathbf{d}},$$

with $\lambda = \frac{c}{T}(4-\sqrt{13})$. We still conclude by Proposition 3.1, and Corollary 3.1.

□

4 Derivation of the Gaussian bounds for the discretization schemes

In order to clarify the reading, we begin with a section dedicated to the presentation of the parametrix techniques, initially developed in [19], in the continuous case. For simplicity, we restrict ourselves to case (a) for the description of the method and the specification of the various steps needed to derive Aronson's bounds. However, the method can be extended to the degenerate case (b). This has been done successfully in the continuous case in [15], [6].

We will adapt the methods developed therein to the discrete case in Sections 4.3, 4.4.

4.1 Parametrix techniques in the non degenerate continuous case

For simplicity, we suppose in this section that the coefficients b, σ are $C_b^\infty([0,T]\times\mathbb{R}^d,\mathbb{R})$ (bounded together with their derivatives at every order) and $\sigma\sigma^*$ is uniformly elliptic. We also assume that they satisfy assumptions **(UE)**,**(SB)**. We denote by L_t the infinitesimal generator of X at time $t \geqslant 0$, i.e.:

$$\forall \varphi \in C^2(\mathbb{R}^d,\mathbb{R}),\ \forall x\in\mathbb{R}^d,\ L_t\varphi(x) = \langle b(t,x), D_x\varphi(x)\rangle + \frac{1}{2}\mathrm{Tr}(a(t,x)D_x^2\varphi(x)).$$

The non-degeneracy and the smoothness of the coefficients guarantee that the solution $(X_t)_{t\geqslant 0}$ of (1.1) admits a smooth transition density p satisfying

$$\begin{aligned}
&\partial_t p(s,t,x,y) = L_t^* p(s,t,x,y),\ t>s, (x,y)\in(\mathbb{R}^d)^2,\ p(s,t,x,.) \underset{t\downarrow s}{\to} \delta_x(.),\\
&\partial_s p(s,t,x,y) = -L_s p(s,t,x,y),\ t>s, (x,y)\in(\mathbb{R}^d)^2,\ p(s,t,.,y) \underset{s\uparrow t}{\to} \delta_y(.),
\end{aligned} \tag{4.1}$$

where L_t^* stands for the adjoint of L_t (see e.g. [7]). We will show that the constants in the Aronson bounds only depend on the fixed final time $T>0$, the upper bounds of the coefficients, the uniform ellipticity constant and the η-Hölder modulus of continuity appearing in **(SB) but not** on the derivatives of the coefficients as in the Malliavin calculus approach.

Remark 4.1. *Under the sole assumptions* **(UE)**, **(SB)**, *we can consider a sequence of mollified equations whose coefficients uniformly satisfy* **(UE)**, **(SB)** *for which Aronson's bounds hold. From a smoothness viewpoint, since the bounds only depend on the η-Hölder modulus of continuity, letting the mollifying parameter go to zero, we derive that they remain valid for the density of X thanks to uniqueness in law (deriving from the well posedness of the martingale problem, see Theorem 7.2.1 in [23]) and Radon-Nikodym's theorem.*

For a fixed starting point $(s,x)\in[0,T)\times\mathbb{R}^d$ and a given $x'\in\mathbb{R}^d$, introduce the Gaussian process

$$\forall t\in[s,T],\ \tilde{X}_t = x + \int_s^t b(u,x')\mathrm{d}u + \int_s^t \sigma(u,x')\mathrm{d}W_u. \tag{4.2}$$

Its infinitesimal generator $\tilde{L}_t^{x'}$ at time $t\in[s,T]$ writes:

$$\forall \varphi\in C^2(\mathbb{R}^d,\mathbb{R}),\ \forall x\in\mathbb{R}^d,\ \tilde{L}_t^{x'}\varphi(x) = \langle b(t,x'), D_x\varphi(x)\rangle + \frac{1}{2}\mathrm{Tr}(a(t,x')D_x^2\varphi(x)).$$

Observe that $(\tilde{X}_t)_{t\in[s,T]} \equiv (\tilde{X}_t^{x'})_{t\in[s,T]}$ and that its density $\tilde{p}^{x'}$ satisfies:

$$\partial_s \tilde{p}^{x'}(s,t,x,y) = -\tilde{L}_s^{x'} \tilde{p}^{x'}(s,t,x,y),\ t > s, (x,y) \in (\mathbb{R}^d)^2,\ \tilde{p}^{x'}(s,t,.,y) \underset{s\uparrow t}{\to} \delta_y(.). \tag{4.3}$$

Hence, from (4.1) and (4.3)

$$\begin{aligned}
(p - \tilde{p}^{x'})(s,t,x,x') &= \int_s^t du \partial_u \int_{\mathbb{R}^d} dz p(s,u,x,z) \tilde{p}^{x'}(u,t,z,x') \\
&= \int_s^t du \int_{\mathbb{R}^d} dz \left(\partial_u p(s,u,x,z) \tilde{p}^{x'}(u,t,z,x') + p(s,u,x,z) \partial_u \tilde{p}^{x'}(u,t,z,x') \right) \\
&= \int_s^t du \int_{\mathbb{R}^d} dz \left(L_u^* p(s,u,x,z) \tilde{p}^{x'}(u,t,z,x') - p(s,u,x,z) \tilde{L}_u^{x'} \tilde{p}^{x'}(u,t,z,x') \right) \\
&= \int_s^t du \int_{\mathbb{R}^d} dz p(s,u,x,z)(L_u - \tilde{L}_u^{x'}) \tilde{p}^{x'}(u,t,z,x'). \qquad (4.4)
\end{aligned}$$

Remark 4.2. *Observe that the parameter x' appears here twice. We actually freeze the coefficients in the Gaussian process at the final point where we want to estimate the density.*

Introducing the notation $f \otimes g(s,t,x,x') = \int_s^t du \int_{\mathbb{R}^d} dz f(s,u,x,z) g(u,t,z,x')$ for the time-space convolution and setting $\tilde{p}(s,t,x,x') := \tilde{p}^{x'}(s,t,x,x')$, $H(s,t,x,x') := (L_s - \tilde{L}_s^{x'})\tilde{p}(s,t,x,x')$, equation (4.4) writes:

$$(p - \tilde{p})(s,t,x,x') = p \otimes H(s,t,x,x').$$

The main idea then consists in iterating this procedure for $p(s,u,x,z)$ in (4.4) exploiting the transition density of the process $\tilde{X}^z$, z being the integration variable. In such a way, we obtain the iterated convolutions of the kernel H as well as the formal expansion:

$$p(s,t,x,x') = \sum_{r=0}^{\infty} \tilde{p} \otimes H^{(r)}(s,t,x,x'), \tag{4.5}$$

with $H^{(0)} = I, H^{(r)} = H \otimes H^{(r-1)},\ r \geqslant 1$.

Under **(UE)**, **(SB)**, for fixed $T > 0$, since $\tilde{p}$ is a Gaussian density, $\exists C := C(T, \lambda_0, L_0),\ c := c(\lambda_0)$ s.t. for any multi-index $\alpha, |\alpha| \leqslant 4$, and all $0 < s < t < T$, $(x,x') \in (\mathbb{R}^d)^2$,

$$\begin{aligned}
&|\partial_x^\alpha \tilde{p}(s,t,x,x')| \leqslant \frac{C}{(t-s)^{|\alpha|/2}} p_c(t-s,x,x'), \\
&\qquad \text{with } p_c(t-s,x,x') := \left(\frac{c}{2\pi(t-s)} \right)^{d/2} \exp\left(-c \frac{|x-x'|^2}{t-s} \right). \qquad (4.6)
\end{aligned}$$

The kernel writes:

$$H(s,t,x,x') = \langle b(s,x) - b(s,x'), D_x \tilde{p}(s,t,x,x') \rangle + \frac{1}{2} \mathrm{tr}\left((a(s,x) - a(s,x')) D_x^2 \tilde{p}(s,t,x,x') \right). \tag{4.7}$$

In particular, equation (4.6) and **(SB)** give:

$$\begin{aligned}|H(s,t,x,x')| &\leqslant C\left(\frac{2|b|_\infty}{(t-s)^{1/2}}+\frac{|a(s,x)-a(s,x')|}{(t-s)}\right)p_c(t-s,x,x')\\ &\leqslant C((t-s)^{-1/2}+(t-s)^{-1+\frac{\eta}{2}})p_c(t-s,x,x').\end{aligned}$$

The η-Hölder continuity of a is used to compensate the non-integrable time singularity of second order terms (up to a modification of c for the last inequality). The kernel H is therefore regularizing in the sense that:
$\exists C := C(T,\lambda_0,L_0,\eta),\ c := c(\lambda_0,\eta),\ \forall r \geqslant 1,$

$$|\tilde{p}\otimes H^{(r)}(s,t,x,x')| \leqslant C^{r+1}(t-s)^{r\eta/2}\prod_{i=1}^{r+1} B\left(1+\frac{(i-1)\eta}{2},\frac{\eta}{2}\right)p_c(t-s,x,x'), \qquad (4.8)$$

where $B(m,n) := \int_0^1 s^{m-1}(1-s)^{n-1}ds$ stands for the β function. For small $t-s$, this term decreases geometrically with r.

From (4.8), the r.h.s. of equation (4.5) is well defined. On the other hand, direct computations show that this r.h.s. satisfies (4.1). Hence, from the uniqueness of the solution to (4.1) we indeed deduce that (4.5) provides a series expansion of the density, i.e. the previous formal derivation is fully justified.

Remark 4.3. *Observe that to extend the previous approach to case* (b)*, we have to consider a suitable frozen Gaussian process that takes into account the metric structure of the problem. The idea is once again to compensate the non integrable time singularity deriving from the second order terms. To this end, the difference of the diffusion coefficients in the convolution kernel has to be homogeneous to the off-diagonal terms of the frozen density in case* (b)*. The associated frozen process is different from case* (a) *mainly because of the transport of the initial condition (see [15] for the continuous case and equation* (4.11) *below for the scheme).*

4.2 Steps for Aronson's bounds in the non degenerate continuous case

We specify below what are the crucial steps to obtain Aronson's estimates from the controls of the previous section.

- **Upper bound**. Equations (4.8) and (4.5) and the asymptotics of the β function directly give the upper bound, namely $\exists C := C(T,\lambda_0,L_0,\eta),\ c := c(\lambda_0,\eta),\ \forall 0 \leqslant s < t \leqslant T, (x,x') \in (\mathbb{R}^d)^2,\ p(s,t,x,x') \leqslant Cp_c(t-s,x,x')$.

- **Lower bound**. There are three steps for the lower bound:

* *Step 1: lower bound on the diagonal in short time*. Equations (4.8) and (4.5) give that there exist $C := C(T,\lambda_0,L_0,\eta), c_0 := c_0(\lambda_0)$

$$p(s,t,x,x') \geqslant \tilde{p}(s,t,x,x') - C(t-s)^{\frac{\eta}{2}}p_c(s,t,x,x') \geqslant c_0{}^{-1}p_{c^{-1}}(s,t,x,x') - C(t-s)^{\frac{\eta}{2}}p_c(s,t,x,x').$$

Fix now $R_0 > 0$. For $|x-x'|^2/(2(t-s)) \leqslant R_0$ and $|t-s| \leqslant C_1 := C_1(C, c_0, c, \eta, R_0)$ sufficiently small[1], the above equation gives:

$$p(s,t,x,x') \geqslant \frac{c_0^{-1}}{2(2\pi c(t-s))^{d/2}} \exp(-c^{-1}R_0). \tag{4.9}$$

We thus have, from the parametrix expansion of the density and the regularizing property of the kernel, the lower bound on the compact sets of the parabolic metric in short time. In the current framework, this idea goes back to Il'in *et al.* [10].

* *Step 2: chaining and global lower bound in short time*. It consists in deriving the lower bound in short time when $|x-x'|^2/(2(t-s)) \geqslant R_0$. To this end a chaining is needed. The idea is to successively apply, along a path joining x to x' in time $t-s$, the lower bound on the compact sets of the metric. The Markov property of the underlying process is also thoroughly exploited. This procedure is rather standard in the non degenerate case, see eg. Chapter VII of Bass [3], for which the natural path turns out to be the straight line between x and x'. This gives the global lower bound in short time.

* *Step 3: lower bound for fixed arbitrary time*. The global lower bound, i.e. for fixed time $T > 0$ can then be derived by convolution or scaling arguments. The convolution argument is standard. We refer to Section 2.3 of [6] for the scaling arguments.

Remark 4.4. *In case* (b)*, Aronson's bounds are derived following the same steps from the associated parametrix representation and controls on the convolution kernel. Anyhow, the situation is much more subtle for the chaining. It consists in applying the lower bound for the compact sets along the optimal path of the deterministic control problem associated to equation* (1.1) *joining* x *to* x' *in time* $t-s$ *(namely when replacing the Brownian entry by a deterministic control), see Section 4.3 and Appendix A in [6]. Let us briefly mention that in case* (a) *the straight line is as well the optimal solution of the associated deterministic control problem when* $\sigma \equiv 1$. *The uniform ellipticity assumption makes it an admissible path for the chaining in case* (a).

4.3 Parametrix representation of the densities for the schemes

We will follow the previous program for the discretization schemes in both cases (a) and (b). The main advantage of the McKean-Singer parametrix is that it admits a natural discrete counterpart, cf. [13, 14, 15]. Similarly to the previous section we first introduce a "frozen" inhomogeneous Markov chain and discrete convolution kernels from which we derive an expansion for the density of the schemes similar to (4.5), cf. Proposition 4.1 below.

We then establish some smoothing properties of our discrete convolution kernel similar to (4.8). This is done in Lemma 4.1. This is a very technical part that involves tedious computations. The proof is postponed to Appendix A

As in the continuous case, Proposition 4.1 and Lemma 4.1 directly give the upper bound of Theorem 2.1 and the lower bound in small time for the compact sets of the associated metric, which in case (b) is different from the one considered in the previous section, see equation (4.16) below.

[1]One can check that $C_1 := \left(\frac{c_0^{-1}}{2c^d C}\exp(-c^{-1}R_0)\right)^{2/\eta}$

The chaining arguments are adapted to our current non-Markovian framework and turn out to be, even in case (a), not so standard, see Section 4.4.2.

Remark 4.5. *Let us mention that in case* (a) *for smoother coefficients (if* b,σ *are* $C_b^{0,2}([0,T]\times\mathbb{R}^d)$*) a much more direct proof of the upper bound can be obtained via Malliavin calculus, see Proposition 3.5 in [9].*

We first need to introduce some objects and notations. Let us begin with the "frozen" inhomogeneous scheme. For fixed $x,x'\in\mathbb{R}^d$, $0\leqslant j<j'\leqslant N$, we define $(\widetilde{X}_{t_i}^{\Delta})_{i\in[\![j,j']\!]}(\equiv(\widetilde{X}_{t_i}^{\Delta,x'})_{i\in[\![j,j']\!]})$ by

$$\widetilde{X}_{t_j}^{\Delta}=x,\quad \forall i\in[\![j,j'),\ \widetilde{X}_{t_{i+1}}^{\Delta}=\widetilde{X}_{t_i}^{\Delta}+b(t_i,x')\Delta+\sigma(t_i,x')(W_{t_{i+1}}-W_{t_i}) \tag{4.10}$$

for case (a). Note that in the above definition the coefficients of the process are frozen at x', but we omit this dependence for notational convenience. In case (b) we define $(\widetilde{X}_{t_i}^{\Delta})_{i\in[\![j,j']\!]}(=(\widetilde{X}_{t_i}^{\Delta,x',j'})_{i\in[\![j,j']\!]})$ by $\widetilde{X}_{t_j}^{\Delta}=x$, and $\forall i\in[\![j,j')$,

$$\widetilde{X}_{t_{i+1}}^{\Delta}=\widetilde{X}_{t_i}^{\Delta}+\begin{pmatrix} b_1(t_i,x')\Delta\\ \displaystyle\int_{t_i}^{t_{i+1}}(\widetilde{X}_s^{\Delta})^{1,d'}\mathrm{d}s\end{pmatrix}+B\sigma\left(t_i,x'-\begin{pmatrix}\mathbf{0}_{d'\times 1}\\ (x')^{1,d'}\end{pmatrix}(t_{j'}-t_i)\right)(W_{t_{i+1}}-W_{t_i}). \tag{4.11}$$

That is, in case (b) the frozen process also depends on j' through an additional term in the diffusion coefficient. This correction term is needed, in order to have good continuity properties w.r.t. the underlying metric associated to p_c when performing differences of the form $a(t_j,x)-a(t_j,x'-\begin{pmatrix}\mathbf{0}_{d'\times 1}\\ (x')^{1,d'}\end{pmatrix}(t_{j'}-t_i))$, see the definition (4.16), Section 4.4 and Appendix A for details.

From now on, $p^{\Delta}(t_j,t_{j'},x,\cdot)$ and $\widetilde{p}^{\Delta,t_{j'},x'}(t_j,t_{j'},x,\cdot)$ denote the transition densities between times t_j and $t_{j'}$ of the discretization schemes (1.3), (1.4) and the "frozen" schemes (4.10), (4.11) respectively.

Let us introduce a discrete "analogue" to the inhomogeneous infinitesimal generators of the continuous objects from which we derive the kernel of the discrete parametrix representation. For a sufficiently smooth function $\psi:\mathbb{R}^d\to\mathbb{R}$ and fixed $x'\in\mathbb{R}^d$, $j'\in(0,N]\!]$, define the family of operators $(L_{t_j}^{\Delta})_{j\in[\![0,j')}$ and $(\widetilde{L}_{t_j}^{\Delta})_{j\in[\![0,j')}(=(\widetilde{L}_{t_j}^{\Delta,t_{j'},x'})_{j\in[\![0,j')})$ by

$$L_{t_j}^{\Delta}\psi(x)=\frac{\mathbb{E}[\psi(X_{t_j+\Delta}^{\Delta})|X_{t_j}^{\Delta}=x]-\psi(x)}{\Delta},\quad\text{and}\quad \widetilde{L}_{t_j}^{\Delta}\psi(x)=\frac{\mathbb{E}[\psi(\widetilde{X}_{t_j+\Delta}^{\Delta})|\widetilde{X}_{t_j}^{\Delta}=x]-\psi(x)}{\Delta}.$$

Using the notation $\widetilde{p}^{\Delta}(t_j,t_{j'},x,x')=\widetilde{p}^{\Delta,t_{j'},x'}(t_j,t_{j'},x,x')$, we now define the discrete kernel H^{Δ} by

$$H^{\Delta}(t_j,t_{j'},x,x')=\left(L_{t_j}^{\Delta}-\widetilde{L}_{t_j}^{\Delta}\right)\widetilde{p}^{\Delta}(t_j+\Delta,t_{j'},x,x'),\quad 0\leqslant j<j'\leqslant N. \tag{4.12}$$

Remark 4.6. *Note carefully that the fixed variable* x' *appears here twice: as the final point where we consider the density and as freezing point in the previous schemes* (4.10), (4.11). *We also mention that, because of the discretization, there is a slight "shift" in time in the definition of* H^{Δ}. *Namely we have* $t_j+\Delta$ *in* $\widetilde{p}^{\Delta}$ *instead of the somehow expected* t_j *that would be the "exact" discrete counterpart of the continuous kernel* (4.7).

Note also that if $j' = j+1$ *i.e.* $t_{j'} = t_j + \Delta$, the transition probability $\widetilde{p}^{\Delta,t_{j'},x'}(t_{j+1}, t_{j+1}, ., x')$ is the Dirac measure $\delta_{x'}$ so that

$$
\begin{aligned}
H^{\Delta}(t_j, t_{j+1}, x, x') &= \Delta^{-1}\left(\mathbb{E}\big[\delta_{x'}(X^{\Delta}_{t_{j+1}})\big|X^{\Delta}_{t_j} = x\big] - \mathbb{E}\big[\delta_{x'}(\widetilde{X}^{\Delta}_{t_{j+1}})\big|\widetilde{X}^{\Delta}_{t_j} = x\big]\right), \\
&= \Delta^{-1}\left(p^{\Delta}(t_j, t_{j+1}, x, x') - \widetilde{p}^{\Delta,t_{j'},x'}(t_j, t_{j+1}, x, x')\right).
\end{aligned}
$$

From the previous definition (4.12), for all $0 \leqslant j < j' \leqslant N$,

$$
H^{\Delta}(t_j, t_{j'}, x, x') = \Delta^{-1}\int_{\mathbb{R}^d}\left[p^{\Delta} - \widetilde{p}^{\Delta,t_{j'},x'}\right](t_j, t_{j+1}, x, u)\widetilde{p}^{\Delta,t_{j'},x'}(t_{j+1}, t_{j'}, u, x')\mathrm{d}u.
$$

Analogously to Lemma 3.6 in [13] we obtain the following result.

Proposition 4.1 (Parametrix for the density of the Euler scheme)**.**
Assume **(UE)**, **(SB)** *are in force. Then, for* $0 \leqslant t_j < t_{j'} \leqslant T$,

$$
p^{\Delta}(t_j, t_{j'}, x, x') = \sum_{r=0}^{j'-j}\left(\widetilde{p}^{\Delta} \otimes_{\Delta} H^{\Delta,(r)}\right)(t_j, t_{j'}, x, x'), \tag{4.13}
$$

where the discrete time convolution type operator $\otimes_{\Delta}$ *is defined by*

$$
(g \otimes_{\Delta} f)(t_j, t_{j'}, x, x') = \sum_{k=0}^{j'-j-1}\Delta\int_{\mathbb{R}^d} g(t_j, t_{j+k}, x, u)f(t_{j+k}, t_{j'}, u, x')\mathrm{d}u,
$$

where $g \otimes_{\Delta} H^{\Delta,(0)} = g$ *and for all* $r \geqslant 1$, $H^{\Delta,(r)} = H^{\Delta} \otimes_{\Delta} H^{\Delta,(r-1)}$ *denotes the r-fold discrete convolution of the kernel* H^{Δ}*. W.r.t. the above definition, we use the convention that* $\widetilde{p}^{\Delta}\otimes_{\Delta}H^{\Delta,(r)}(t_j, t_j, x, x') = 0, r \geqslant 1$.

The following lemma gives the smoothing properties of our discrete convolution kernel which is, as indicated in Section 4.2, a key argument for the proof of Aronson's bounds.

Lemma 4.1. *There exists* $c > 0, C \geqslant 1$ *s.t. for all* $0 \leqslant j < j' \leqslant N$*, for all* $r \in [\![0, j'-j]\!], \forall (x, x') \in \mathbb{R}^d$,

$$
|\widetilde{p}^{\Delta} \otimes_{\Delta} H^{\Delta,(r)}(t_j, t_{j'}, x, x')| \leqslant C^{r+1}(t_{j'} - t_j)^{r\eta/2}\prod_{i=1}^{r+1}B\left(1 + \frac{(i-1)\eta}{2}, \frac{\eta}{2}\right)p_c(t_{j'} - t_j, x, x'). \tag{4.14}
$$

In the above equation $B(m, n) := \int_0^1 s^{m-1}(1-s)^{n-1}\mathrm{d}s$ *stands for the* β *function.*

The proof of the lemma is postponed to Appendix A.

4.4 Proof of Theorem 2.1: Aronson's bounds for the schemes

4.4.1 Proof of the upper bound.

The upper bound in (2.2) directly follows from Proposition 4.1 and the asymptotics of the β function. It is also useful to achieve the first step of the lower bound.

4.4.2 Proof of the lower bound.

We provide in this section the global lower bound in short time. W.l.o.g. we assume that $T \leqslant 1$. This allows to substitute the constant C appearing in (4.14) by a constant $c_0 \leqslant c\exp(c|b|_\infty^2)$ uniformly for $t_{j'} - t_j \leqslant T$. From the upper bound, we derive the lower bound in short time, on the compact sets of the underlying metric, see (4.16) below. This gives the diagonal decay. To get the whole bound in short time it remains to obtain the "off-diagonal" bound. To this end a chaining argument is needed. In case (a) it is quite standard in the Markovian framework, see Chapter VII of Bass [3] or Kusuoka and Stroock [16]. In case (b), the chaining in the appendix of [6] can be adapted to our discrete framework. We adapt below these arguments to our non Markovian setting for the sake of completeness.

Eventually, to derive the lower bound for an arbitrary fixed $T > 0$ it suffices to use the bound in short time and the semigroup property of $p_{c^{-1}}$. Naturally, the biggest is T, the worse is the constant in the global lower bound.

* *Step 1: lower bound on the diagonal in short time.*
From Proposition 4.1 we have

$$
\begin{aligned}
p^\Delta(t_j, t_{j'}, x, x') &\geqslant \widetilde{p}^\Delta(t_j, t_{j'}, x, x') - \sum_{r=1}^{j'-j} |\widetilde{p}^\Delta \otimes_\Delta H^{\Delta,(r)}(t_j, t_{j'}, x, x')| \\
&\geqslant c_0^{-1} p_{c^{-1}}(t_{j'} - t_j, x, x') - c_0(t_{j'} - t_j)^{\eta/2} p_c(t_{j'} - t_j, x, x'),
\end{aligned} \tag{4.15}
$$

exploiting $\widetilde{p}^\Delta(t_j, t_{j'}, x, x') \geqslant c_0^{-1} p_{c^{-1}}(t_{j'} - t_j, x, x')$ (cf. Lemma 3.1 of [15] in case (b)) and (4.14) (replacing C by c_0) for the last inequality. Equation (4.15) provides a lower bound on compact sets provided that T is small enough. Precisely, denoting

$$
d^2_{t_{j'}-t_j}(x, x') = \begin{cases} \frac{|x - x'|^2}{t_{j'} - t_j} & \text{in case } (a), \\ \frac{\left|(x')^{1,d'} - x^{1,d'}\right|^2}{2(t_{j'} - t_j)} + 6\frac{\left|(x')^{d'+1,d} - x^{d'+1,d} - \frac{x^{1,d'} + (x')^{1,d'}}{2}(t_{j'} - t_j)\right|^2}{(t_{j'} - t_j)^3} & \text{in case } (b), \end{cases} \tag{4.16}
$$

we have that, for a given $R_0 \geqslant 1/2$, if $d^2_{t_{j'}-t_j}(x, x') \leqslant 2R_0$ and $(t_{j'} - t_j) \leqslant T \leqslant \left(\frac{1}{2c_0^2 c^d}\exp(-c^{-1}R_0)\right)^{2/\eta}$,

$$
\begin{aligned}
p^\Delta(t_j, t_{j'}, x, x') &\geqslant \frac{1}{(2\pi)^{d/2}(t_{j'} - t_j)^{\mathbf{S}}} \left(\frac{c_0^{-1}}{c^{d/2}} \exp(-c^{-1}R_0) - c_0 c^{d/2} T^{\eta/2} \right) \\
&\geqslant \frac{c_0^{-1}}{(2\pi c)^{d/2} 2(t_{j'} - t_j)^{\mathbf{S}}} \exp(-c^{-1}R_0)
\end{aligned}
$$

where the parameter $\mathbf{S}$ is the intrinsic scale of the scheme. In case (a) $\mathbf{S} = d/2$, in case (b) $\mathbf{S} = d$. Hence, up to a modification of c_0^{-1} we have that

$$
\exists c_0 \geqslant 1,\ \forall 0 \leqslant j < j' \leqslant N,\ \forall (x, x') \in (\mathbb{R}^d)^2,\ d^2_{t_{j'}-t_j}(x, x') \leqslant 2R_0,\ p^\Delta(t_j, t_{j'}, x, x') \geqslant \frac{c_0^{-1}}{(t_{j'} - t_j)^{\mathbf{S}}}. \tag{4.17}
$$

In particular $\exists c>0$, $c_0\geqslant 1$, $\forall 0\leqslant j<j'\leqslant N$, $\forall (x,x')\in(\mathbb{R}^d)^2$, $d^2_{t_{j'}-t_j}(x,x')\leqslant 2R_0$, $p^\Delta(t_j,t_{j'},x,x')\geqslant c_0^{-1}p_{c^{-1}}(t_{j'}-t_j,x,x')$.

* *Step 2 : chaining and global lower bound in short time.*

Case *(a)*. Let us introduce: $\forall 0\leqslant s<t\leqslant T$, $(x,x',y)\in(\mathbb{R}^d)^3$, $p^{\Delta,y}(s,t,x,x')\mathrm{d}x':=\mathbb{P}[X_t^\Delta\in \mathrm{d}x'|X_s^\Delta=x, X^\Delta_{\phi(s)}=y]$. Equation (4.17) provides a lower bound for the density of the scheme when s,t correspond to discretization times. For the chaining, the first step consists in extending this result to arbitrary times $0\leqslant s<t\leqslant T$.Precisely, if $d^2_{t-s}(x,x')\leqslant R_0/12$ we prove that

$$\exists c_0\geqslant 1,\ \forall 0\leqslant s<t\leqslant T,\ \forall y,\ p^{\Delta,y}(s,t,x,x')\geqslant c_0^{-1}(t-s)^{-d/2}. \tag{4.18}$$

If $\phi(t)=\phi(s)$ or $t=\phi(t)=\phi(s)+\Delta$, the above density is Gaussian and (4.18) holds. If $t\neq\phi(t)=(\phi(s)+\Delta)$ or $t=\phi(t)=\phi(s)+2\Delta$, equation (4.18) directly follows from a convolution argument between two Gaussian random variables. Note anyhow carefully that the “crude” convolution argument cannot be iterated L times for an arbitrary large L. Indeed, in that case the constants would have a geometric decay. Thus, for $\phi(t)-(\phi(s)+\Delta)\geqslant\Delta$, supposing w.l.o.g. that s and t do not belong to the time grid $\left\{t_j\right\}_{j\in[\![0,N]\!]}$[2], we write

$$\begin{aligned}p^{\Delta,y}(s,t,x,x')&=\int_{(\mathbb{R}^d)^2}p^{\Delta,y}(s,\phi(s)+\Delta,x,x_1)p^\Delta(\phi(s)+\Delta,\phi(t),x_1,x_2)p^\Delta(\phi(t),t,x_2,x')\mathrm{d}x_1\mathrm{d}x_2\\&\geqslant\int_{B_R(s,t,x,x')}p^{\Delta,y}(s,\phi(s)+\Delta,x,x_1)p^\Delta(\phi(s)+\Delta,\phi(t),x_1,x_2)p^\Delta(\phi(t),t,x_2,x')\mathrm{d}x_1\mathrm{d}x_2\end{aligned} \tag{4.19}$$

where $B_R(s,t,x,x'):=\{x_1\in\mathbb{R}^d: d^2_{\phi(s)+\Delta-s}(x,x_1)\leqslant R\}\times\{x_2\in\mathbb{R}^d: d^2_{t-\phi(t)}(x_2,x')\leqslant R\}$ for $R>0$ to be specified later on. Now, for $(x_1,x_2)\in B_R(s,t,x,x')$,

$$d^2_{\phi(t)-(\phi(s)+\Delta)}(x_1,x_2)=\frac{|x_1-x_2|^2}{\phi(t)-(\phi(s)+\Delta)}\leqslant\frac{2|x_1-x|^2+4|x-x'|^2+4|x_2-x'|^2}{\phi(t)-(\phi(s)+\Delta)}\leqslant 6R+R_0,$$

where we used that for $\phi(t)-(\phi(s)+\Delta)\geqslant\Delta$, $\frac{1}{\phi(t)-(\phi(s)+\Delta)}\leqslant\frac{3}{t-s}$ in the last inequality. Taking $R=R_0/6$ we obtain that $\forall(x_1,x_2)\in B_R(s,t,x,x')$, $d^2_{\phi(t)-(\phi(s)+\Delta)}(x_1,x_2)\leqslant 2R_0$. We therefore derive from (4.17) and (4.19) that $\exists c_0>0$,

$$\begin{aligned}&p^{\Delta,y}(s,t,x,x')\\&\quad\geqslant c_0^{-1}(\phi(s)+\Delta-s)^{-d/2}(t-\phi(t))^{-d/2}(\phi(t)-(\phi(s)+\Delta))^{-d/2}\int_{(\mathbb{R}^d)^2}\mathbb{I}_{(x_1,x_2)\in B_R(s,t,x,x')}\mathrm{d}x_1\mathrm{d}x_2.\end{aligned}$$

Since $\phi(t)-(\phi(s)+\Delta)\leqslant t-s$ and there exists $\tilde c>0$ s.t. $|\{x_1\in\mathbb{R}^d: d^2_{\phi(s)+\Delta-s}(x,x_1)\leqslant R\}|\geqslant\tilde c(\phi(s)+\Delta-s)^{d/2}$, $|\{x_2\in\mathbb{R}^d: d^2_{t-\phi(t)}(x_2,x')\leqslant R\}|\geqslant\tilde c(t-\phi(t))^{d/2}$ where $|.|$ stands for the Lebesgue measure of a given set in $\mathbb{R}^d$, we derive (4.18) from the above equation up to a modification of c_0.

[2] Indeed, if (s,t) both belong to the grid, equation (4.17) already gives the bound. If s or (exclusive) t belongs to the grid, the arguments below can be easily adapted.

It now remains to do the chaining when for $0 \leqslant j < j' \leqslant N$, $(x,x') \in (\mathbb{R}^d)^2$ we have $d^2_{t_{j'}-t_j}(x,x') \geqslant 2R_0 \geqslant 1$. Set $L = \lceil K d^2_{t_{j'}-t_j}(x,x') \rceil$, for $K \geqslant 1$ to be specified later on and $h := (t_{j'} - t_j)/L$. Note that $L \geqslant 1$. For all $i \in [\![0,L]\!]$ we denote $s_i = t_j + ih, y_i = x + \frac{i}{L}(x'-x)$ so that $s_0 = t_j, s_L = t_{j'}$, $y_0 = x, y_L = x'$. Introduce now $\rho := d_{t_{j'}-t_j}(x,x')(t_{j'}-t_j)^{1/2}/L = |x'-x|/L$ and for all $i \in [\![1,L-1]\!]$, $B_i := \{x \in \mathbb{R}^d : |x - y_i| \leqslant \rho\}$. Note that with the previous definitions $\forall i \in [\![0,L-1]\!]$, $|y_{i+1} - y_i| = |x'-x|/L = \rho$. Thus,

$$\begin{aligned} \forall x_1 \in B_1,\ |x - x_1| \leqslant 2\rho,\ \forall i \in [\![1,L-2]\!],\ (x_i,x_{i+1}) \in B_i \times B_{i+1},\ |x_i - x_{i+1}| \leqslant 3\rho, \\ \forall x_{L-1} \in B_{L-1},\ |x_{L-1} - x'| \leqslant 2\rho. \end{aligned} \tag{4.20}$$

We can now choose K large enough s.t.

$$3\rho/\sqrt{h} = 3d_{t_{j'}-t_j}(x,x')/\sqrt{L} \leqslant (R_0/12)^{1/2} \tag{4.21}$$

so that according to (4.18), denoting $x_0 = x, x_L = x'$, for all $i \in [\![0,L-1]\!]$, $\forall y \in \mathbb{R}^d, (x_i,x_{i+1}) \in B_i \times B_{i+1}$, $p^{\Delta,y}(s_i,s_{i+1},x_i,x_{i+1}) \geqslant c_0^{-1}h^{-d/2}$ (with the slight abuse of notation $B_0 = \{x\}, B_L = \{x'\}$ and $p^{\Delta,y}(0,h,x_0,\, x_1) = p^{\Delta}(0,h,x,x_1)$).

We have

$$p^{\Delta}(t_j,t_{j'},x,x') \geqslant \mathbb{E}_{t_j,x}\left[\mathbb{I}_{\cap_{i=1}^{L-1} X^{\Delta}_{s_i} \in B_i} p^{\Delta, X^{\Delta}_{\phi(s_{L-1})}}(s_{L-1},t_{j'},X^{\Delta}_{s_{L-1}},x')\right]. \tag{4.22}$$

To proceed we have to distinguish two cases: $h \geqslant \Delta$ and $h < \Delta$.

- If $h \geqslant \Delta$, write from (4.22),

$$p^{\Delta}(t_j,t_{j'},x,x') \geqslant \mathbb{E}_{t_j,x}\left[\mathbb{I}_{\cap_{i=1}^{L-1} X^{\Delta}_{s_i} \in B_i} \mathbb{E}[p^{\Delta, X^{\Delta}_{\phi(s_{L-1})}}(s_{L-1},t_{j'},X^{\Delta}_{s_{L-1}},x') | X^{\Delta}_{s_{L-1}},\, X^{\Delta}_{\phi(s_{L-1})}]\right].$$

Since we consider the events $X^{\Delta}_{s_{L-1}} \in B_{L-1}$, we derive from (4.20), (4.21) that $|X^{\Delta}_{s_{L-1}} - x'|/\sqrt{h} \leqslant 2\rho/\sqrt{h} \leqslant 3d_{t_{j'}-t_j}(x,x')/\sqrt{L} \leqslant (R_0/12)^{1/2}$. Hence, from (4.18)

$$\begin{aligned} p^{\Delta}(t_j,t_{j'},x,x') &\geqslant c_0^{-1}h^{-d/2}\mathbb{E}_{t_j,x}\left[\mathbb{I}_{\cap_{i=1}^{L-1} X^{\Delta}_{s_i} \in B_i}\right] \\ &= c_0^{-1}h^{-d/2}\mathbb{E}_{t_j,x}\left[\mathbb{I}_{\cap_{i=1}^{L-2} X^{\Delta}_{s_i} \in B_i}\mathbb{P}[X^{\Delta}_{s_{L-1}} \in B_{L-1} | X^{\Delta}_{s_{L-2}}, X^{\Delta}_{\phi(s_{L-2})}]\right]. \end{aligned}$$

Now $\mathbb{P}[X^{\Delta}_{s_{L-1}} \in B_{L-1} | X^{\Delta}_{s_{L-2}}, X^{\Delta}_{\phi(s_{L-2})}] = \int_{B_{L-1}} p^{\Delta, X^{\Delta}_{\phi(s_{L-2})}}(s_{L-2},s_{L-1},X^{\Delta}_{s_{L-2}},y)\mathrm{d}y$, but since we restrict to $X^{\Delta}_{s_{L-2}} \in B_{L-2}$, according to (4.20), we have for all $y \in B_{L-1}$, $|X^{\Delta}_{s_{L-2}} - y|/\sqrt{h} \leqslant 3\rho/\sqrt{h} \leqslant (R_0/12)^{1/2}$ for the previous R and therefore (4.18) yields

$$p^{\Delta}(t_j,t_{j'},x,x') \geqslant (c_0^{-1}h^{-d/2})^2 |B_{L-1}| \mathbb{E}_{t_j,x}[\mathbb{I}_{\cap_{i=1}^{L-2} X^{\Delta}_{s_i} \in B_i}].$$

Iterating the process we finally get

$$p^{\Delta}(t_j,t_{j'},x,x') \geqslant (c_0^{-1}h^{-d/2})^L \prod_{i=1}^{L-1} |B_i|.$$

Observing that

$$\exists \tilde{c} > 0,\ \forall i \in [\![1, L-1]\!],\ |B_i| \geqslant \tilde{c}\rho^d, \tag{4.23}$$

we obtain from the previous definition of h and L:

$$\begin{aligned} p^{\Delta}(t_j, t_{j'}, x, x') &\geqslant (c_0^{-1}h^{-d/2})^L(\tilde{c}\rho^d)^{L-1} \\ &\geqslant c_0^{-1}(t_{j'} - t_j)^{-d/2}\exp((L-1)\log(c_0^{-1}\tilde{c}(\rho/\sqrt{h})^d)) \\ &\geqslant c_0^{-1}(t_{j'} - t_j)^{-d/2}\exp(-cd^2_{t_{j'}-t_j}(x, x')) \end{aligned} \tag{4.24}$$

for a suitable c up to a modification of c_0.

- If $h < \Delta$. We have to introduce for all $k \in [\![j, j')$, $I_k := \{l \in [\![0, L-1]\!],\ s_l \in [t_k, t_{k+1}[\}$. Rewrite from (4.22)

$$p^{\Delta}(t_j, t_{j'}, x, x') \quad \geqslant \quad \mathbb{E}_{t_j,x}[\mathbb{I}_{\cap_{k=j}^{j'-1}\cap_{i\in I_k}X^{\Delta}_{s_i}\in B_i}p^{\Delta, X^{\Delta}_{\phi(s_{L-1})}}(s_{L-1}, t_{j'}, X^{\Delta}_{s_{L-1}}, x')].$$

Define for all $k \in [\![j, j')$, $i \in [\![1, \sharp I_k]\!], I_k^i \in I_k$ and $t_k \leqslant s_{I_k^1} < s_{I_k^2} < \cdots < s_{I_k^{\sharp I_k}} < t_{k+1}$. In particular, for all $i \in [\![1, \sharp I_k - 1]\!]$, $s_{I_k^{i+1}} - s_{I_k^i} = h$. Rewrite now,

$$p^{\Delta}(t_j, t_{j'}, x, x') \quad \geqslant \quad \mathbb{E}_{t_j,x}[\mathbb{I}_{\cap_{k=j}^{j'-2}\cap_{i\in I_k}X^{\Delta}_{s_i}\in B_i}\mathbb{E}[\mathbb{I}_{\cap_{i\in I_{j'-1}}X^{\Delta}_{s_i}\in B_i}p^{\Delta, X^{\Delta}_{t_{j'-1}}}(s_{L-1}, t_{j'}, X^{\Delta}_{s_{L-1}}, x')|\mathscr{F}_{s_{I_{j'-2}^{\sharp I_{j'-2}}}}]]. \tag{4.25}$$

Introducing

$$\begin{aligned} P_{j'-1,j} &:= \mathbb{E}[\mathbb{I}_{\cap_{i\in I_{j'-1}}X^{\Delta}_{s_i}\in B_i}p^{\Delta, X^{\Delta}_{t_{j'-1}}}(s_{L-1}, t_{j'}, X^{\Delta}_{s_{L-1}}, x')|\mathscr{F}_{s_{I_{j'-2}^{\sharp I_{j'-2}}}}] \\ &= \mathbb{E}[\mathbb{I}_{X^{\Delta}_{s_{I^1_{j'-1}}}\in B_{I^1_{j'-1}}}\int_{\prod_{i=2}^{\sharp I_{j'-1}}B_{I^i_{j'-1}}} p^{\Delta, X^{\Delta}_{t_{j'-1}}}(s_{I^1_{j'-1}}, s_{I^2_{j'-1}}, X^{\Delta}_{s_{I^1_{j'-1}}}, x_2) \\ &\quad \times \prod_{i=2}^{\sharp I_{j'-1}-1} p^{\Delta, X^{\Delta}_{t_{j'-1}}}(s_{I^i_{j'-1}}, s_{I^{i+1}_{j'-1}}, x_i, x_{i+1})p^{\Delta, X^{\Delta}_{t_{j'-1}}}(s_{I_{j'-1}^{\sharp I_{j'-1}}}, t_{j'}, x_{\sharp I_{j'-1}}, x')\prod_{i=2}^{\sharp I_{j'-1}}\mathrm{d}x_i|\mathscr{F}_{s_{I_{j'-2}^{\sharp I_{j'-2}}}}], \end{aligned}$$

we derive from (4.20), (4.21) and (4.18)

$$\begin{aligned} P_{j'-1,j} \quad &\geqslant \quad (c_0^{-1}h^{-d/2})^{\sharp I_{j'-1}}\prod_{i=2}^{\sharp I_{j'-1}}|B_{I^i_{j'-1}}|\int_{B_{I^1_{j'-1}}} p^{\Delta, X^{\Delta}_{t_{j'-2}}}(s_{I_{j'-2}^{\sharp I_{j'-2}}}, s_{I^1_{j'-1}}, X^{\Delta}_{s_{I_{j'-2}^{\sharp I_{j'-2}}}}, x_1)\mathrm{d}x_1 \\ &\overset{(4.23)}{\geqslant} \quad (c_0^{-1}h^{-d/2})^{\sharp I_{j'-1}}(\tilde{c}\rho^d)^{\sharp I_{j'-1}-1}\int_{B_{I^1_{j'-1}}} p^{\Delta, X^{\Delta}_{t_{j'-2}}}(s_{I_{j'-2}^{\sharp I_{j'-2}}}, s_{I^1_{j'-1}}, X^{\Delta}_{s_{I_{j'-2}^{\sharp I_{j'-2}}}}, x_1)\mathrm{d}x_1. \end{aligned}$$

Plugging this estimate in (4.25) we obtain

$$\begin{aligned}
p^{\Delta}(t_j,t_{j'},x,x') &\geqslant (c_0^{-1}h^{-d/2})^{\sharp I_{j'-1}}(\tilde c\rho^d)^{\sharp I_{j'-1}-1}\\
&\quad\times\mathbb{E}_{t_j,x}[\mathbb{I}_{\cap_{k=j}^{j'-2}\cap_{i\in I_k}X^{\Delta}_{s_i}\in B_i}\int_{B_{I^1_{j'-1}}} p^{\Delta,X^{\Delta}_{t_{j'-2}}}(s_{I_{j'-2}^{\sharp I_{j'-2}}},s_{I^1_{j'-1}},X^{\Delta}_{s_{I_{j'-2}^{\sharp I_{j'-2}}}},x_1)\mathrm{d}x_1]\\
&\geqslant (c_0^{-1}h^{-d/2})^{\sharp I_{j'-1}+1}(\tilde c\rho^d)^{\sharp I_{j'-1}}\mathbb{E}_{t_j,x}[\mathbb{I}_{\cap_{k=j}^{j'-2}\cap_{i\in I_k}X^{\Delta}_{s_i}\in B_i}]
\end{aligned}$$

using once again (4.21), (4.18) for the last inequality. Iterating this procedure we still obtain (4.24) and can conclude as in the previous case.

Case *(b)*. If $d^2_{t-s}(x,x')\leqslant \tilde c^{-1}R_0$, for $\tilde c$ large enough, we can show similarly to case (a) that

$$\exists c_0>0,\ \forall 0\leqslant s<t\leqslant T,\ \forall y,\ p^{\Delta,y}(s,t,x,x')\geqslant c_0^{-1}(t-s)^{-d}. \tag{4.26}$$

As in the previous paragraph we reduce to the case $\phi(t)-(\phi(s)+\Delta)\geqslant\Delta$, $(s,t)\notin\{t_j\}^2_{j\in[\![0,N]\!]}$. Then, equation (4.19) still holds for the previous set B_R with the current definition of $d^2_{\cdot}(.,.)$. From standard computations, we derive taking a suitable R that $\forall(x_1,x_2)\in B_R(s,t,x,x')$, $d^2_{\phi(t)-(\phi(s)+\Delta)}(x_1,x_2)\leqslant 2R_0$. Therefore,

$$p^{\Delta,y}(s,t,x,x')\geqslant c_0^{-1}(\phi(s)+\Delta-s)^{-d}(t-\phi(t))^{-d}(\phi(t)-(\phi(s)+\Delta))^{-d}\int_{(\mathbb{R}^d)^2}\mathbb{I}_{(x_1,x_2)\in B_R(s,t,x,x')}\mathrm{d}x_1\mathrm{d}x_2. \tag{4.27}$$

Define now $\forall(u,y)\in(0,T]\times\mathbb{R}^d,\ R>0$,

$$\tilde B_R(u,y):=\{z\in\mathbb{R}^d:|z^{1,d'}-y^{1,d'}|^2/u\leqslant R/7,|z^{d'+1,d}-y^{d'+1,d}-y^{1,d'}u|^2/u^3\leqslant R/24\}.$$

We have that $\forall z\in\tilde B_R(u,y)$:

$$\begin{aligned}
d^2_u(y,z) &:= \frac{|z^{1,d'}-y^{1,d'}|^2}{2u}+6\frac{|z^{d'+1,d}-y^{d'+1,d}-\frac{y^{1,d'}+z^{1,d'}}{2}u|^2}{u^3}\\
&\leqslant 12\frac{|z^{d'+1,d}-y^{d'+1,d}-y^{1,d'}u|^2}{u^3}+7\frac{|z^{1,d'}-y^{1,d'}|^2}{2u}\leqslant R.
\end{aligned}$$

Hence $\tilde B_R(\phi(s)+\Delta-s,x)\times\tilde B_R(t-\phi(t),x')\subset B_R(s,t,x,x')$ and therefore $\exists\tilde c>0,|B_R(s,t,x,x')|\geqslant\tilde c(t-\phi(t))^d(\phi(s)+\Delta-s)^d$ which plugged into (4.27) yields (4.26).

It now remains to do the chaining when $d^2_{t_{j'}-t_j}(x,x')\geqslant 2R_0$. The crucial point is to choose a "good" path between x and x'. In the non degenerated case it was naturally the straight line between the two points (Euclidean geodesic). In our current framework we can relate $d^2_{t_{j'}-t_j}(x,x')$ to a deterministic control problem. Introduce:

$$I(t_{j'}-t_j,x,x')=\inf\{\int_0^{t_{j'}-t_j}|\varphi(s)|^2\mathrm{d}s,\phi(0)=x,\ \phi(t_{j'}-t_j)=x'\},\ \dot\phi_t=A\phi_t+B\varphi_t, \tag{\textbf{CD}}$$

with $A = \begin{pmatrix} \mathbf{0_{d'\times d'}} & \mathbf{0_{d'\times d'}} \\ \mathbf{I_{d'\times d'}} & \mathbf{0_{d'\times d'}} \end{pmatrix}, B = \begin{pmatrix} \mathbf{1_{d'\times d'}} \\ \mathbf{0_{d'\times d'}} \end{pmatrix}$, $\varphi \in L^2([0,t_{j'}-t_j],\mathbb{R}^{d'})$. Problem (**CD**) is a linear deterministic controllability problem that has a unique solution corresponding to

$$\varphi_s = B^*[R(t_{j'}-t_j,s)]^*[Q^{-1}_{t_{j'}-t_j}](x'-R(t_{j'}-t_j,0)x), \tag{4.28}$$

where R stands for the resolvent, i.e. $\forall 0 \leqslant t, t_0 \leqslant t_{j'}-t_j, \partial_t R(t,t_0) = AR(t,t_0)$, $R(t_0,t_0) = \mathbf{I_{d\times d}}$ and $Q_{t_{j'}-t_j} = \int_0^{t_{j'}-t_j} R(t_{j'}-t_j,s)BB^*R(t_{j'}-t_j,s)^* \mathrm{d}s$ is the Gram matrix, see e.g. Theorem 1.11 Chapter 1 in Coron [5]. For (**CD**) the resolvent writes $R(t,t_0) = \begin{pmatrix} \mathbf{I_{d'\times d'}} & \mathbf{0_{d'\times d'}} \\ (t-t_0)\mathbf{I_{d'\times d'}} & \mathbf{I_{d'\times d'}} \end{pmatrix}$ and therefore the Gram matrix of the control problem corresponds to the covariance matrix of the process $X_t = x + \int_0^t AX_s \mathrm{d}s + BW_t$ at time $t_{j'}-t_j$, that is $Q_{t_{j'}-t_j} = \begin{pmatrix} (t_{j'}-t_j)\mathbf{I_{d'\times d'}} & (t_{j'}-t_j)^2/2\mathbf{I_{d'\times d'}} \\ (t_{j'}-t_j)^2/2\mathbf{I_{d'\times d'}} & (t_{j'}-t_j)^3/3\mathbf{I_{d'\times d'}} \end{pmatrix}$. Hence, explicit computations give: $\forall s \in [0,t_{j'}-t_j]$,

$$\varphi_s = \frac{(x')^{1,d'} - x^{1,d'}}{(t_{j'}-t_j)^2}[6s - 2(t_{j'}-t_j)] + 6\frac{(x')^{d'+1,d} - x^{d'+1,d} - x^{1,d'}(t_{j'}-t_j)}{(t_{j'}-t_j)^3}[t_{j'}-t_j-2s], \tag{4.29}$$

and thus, $\frac{1}{2}I(t_{j'}-t_j,x,x') = d^2_{t_{j'}-t_j}(x,x')$ defined in (4.16). Now we have a candidate for a deterministic curve along which we can do the chaining. It is simply the deterministic curve $(\phi_s)_{s\in[0,t_{j'}-t_j]}$ solution of (**CD**) for the above control $(\varphi_s)_{s\in[0,t_{j'}-t_j]}$.

To complete the proof of the chaining it remains to specify how to define the $(s_i)_{i\geqslant 1}, (y_i)_{i\geqslant 1}$ and the associated sets. Recall that $2R_0 \geqslant 1$. We set here $L := \lceil K d^2_{t_{j'}-t_j}(x,x') \rceil$ for an integer $K \geqslant 3$ to be specified later on. In term of the new distance, L is similar in its definition to the one of the previous paragraph. Define $s_0 = 0$, $s_i := \inf\{t \in [s_{i-1}, t_{j'}-t_j] : \int_{s_{i-1}}^t |\varphi_s|^2 \mathrm{d}s = I(t_{j'}-t_j,x,x')/L\} \wedge (s_{i-1} + (t_{j'}-t_j)/L)\mathbb{I}_{s_{i-1}<(t_{j'}-t_j)(1-\frac{2}{L})} + (t_{j'}-t_j)\mathbb{I}_{s_{i-1}\geqslant(t_{j'}-t_j)(1-\frac{2}{L})}$, $i \geqslant 1$. The previous conditions on R_0, K give the well posedness of this definition.

Lemma 4.2 (Controls on the time step)**.** *Set for all $i \geqslant 0$, $\varepsilon_i := s_{i+1} - s_i$. There exist a constant $c_1 \leqslant 1$ and an integer $\bar{L} \in [L-1, L/c_1]$, s.t. $s_{\bar{L}} = t_{j'}-t_j$ and*

$$\forall i \in [\![0, \bar{L}-2]\!],\ c_1\frac{t_{j'}-t_j}{L} \leqslant \varepsilon_i \leqslant \frac{t_{j'}-t_j}{L},\ \frac{t_{j'}-t_j}{L} \leqslant \varepsilon_{\bar{L}-1} \leqslant 2\frac{t_{j'}-t_j}{L}. \tag{4.30}$$

Proof. We first set $\bar{L} = \inf\{k \geq 1 : s_k = t_{j'}-t_j\}$. The set $\{k \geq 1 : s_k = t_{j'}-t_j\}$ is clearly non-empty. The upper bound in (4.30) then follows from the definition of the family $(s_i)_{i\geqslant 1}$. Suppose now that $s_i < (t_{j'}-t_j)(1-2/L)$ for a given $0 \leq i \leq \bar{L}-2$. Assume also that $s_{i+1} - s_i < (t_{j'}-t_j)/L$ (otherwise $\varepsilon_i = (t_{j'}-t_j)/L$). Then, $\int_{s_i}^{s_{i+1}} |\varphi_s|^2 \mathrm{d}s = I(t_{j'}-t_j,x,x')/L$. From (4.28), (4.29), (4.16), we deduce that

$$\exists c_2 > 0,\ \sup_{0\leq s\leq t_{j'}-t_j} |\varphi_s| \leq c_2(t_{j'}-t_j)^{-1/2} d_{t_{j'}-t_j}(x,x').$$

Hence, we obtain

$$\int_{s_i}^{s_{i+1}} |\varphi_s|^2 \mathrm{d}s := \frac{I(t_{j'}-t_j,x,x')}{L} \leqslant c_2^2 \varepsilon_i \frac{d^2_{t_{j'}-t_j}(x,x')}{(t_{j'}-t_j)}.$$

Recalling that $I(t_{j'}-t_j,x,x') = 2d^2_{t_{j'}-t_j}(x,x')$, the lower bound in (4.30) follows for all i s.t. $s_i < T(1-2/L)$. The bound for $\bar{L}$ and the last time step are then easily derived. □

Define now for all $i \in [\![0,\bar{L}]\!]$, $y_i = \phi_{s_i}$ (in particular $y_0 = x$ and $y_{\bar{L}} = x'$), and for all $i \in [\![1,\bar{L}-1]\!]$,

$$B_i := \{z \in \mathbb{R}^d : |Q_{K\rho^2}^{-1/2}(R(s_i,s_{i-1})y_{i-1}-z)| + |Q_{K\rho^2}^{-1/2}(z-R(s_i,s_{i+1})y_{i+1})| \leqslant 2R_0K^{-1/2}\},$$

where $\rho := d_{t_{j'}-t_j}(x,x')(t_{j'}-t_j)^{1/2}/L$. Because of the transport term, we are led to consider sets that involve the forward transport from the previous point on the optimal curve and the backward transport of the next point in the above definition. Equation (4.22) still holds with L replaced by $\bar{L}$. Following the strategy of the previous paragraph concerning the conditioning, the end of the proof relies on the following

Lemma 4.3 (Controls for the chaining). *With the previous assumptions and definitions we have that for K large enough:*

$$\begin{aligned}
\forall i \in [\![1,\bar{L}-2]\!],\ \forall (x_i,x_{i+1}) \in B_i \times B_{i+1}, d^2_{\varepsilon_i}(x_i,x_{i+1}) \leqslant 2R_0,\\
\forall x_1 \in B_1,\ d^2_{s_1}(x,x_1) \leqslant 2R_0,\\
\forall x_{\bar{L}-1} \in B_{\bar{L}-1}, d^2_{\varepsilon_{\bar{L}-1}}(x_{\bar{L}-1},x') \leqslant 2R_0.
\end{aligned} \tag{4.31}$$

For the same c_1 as in Lemma 4.2,

$$\forall i \in [\![1,\bar{L}-1]\!], |B_i| \geqslant c_1\rho^{2d}, \tag{4.32}$$

where $|B_i|$ stands for the Lebesgue measure of the set B_i.

Indeed, exploiting, (4.30), (4.31) (resp. (4.32)) instead of (4.20), (4.21) (resp. (4.23)), the proof remains unchanged. The proof of Lemma 4.3 is postponed to Appendix A.

A Proof of the technical Lemmas

A.1 Proof of Lemma 4.1.

The key estimate is the following control of the convolution kernel H^Δ. There exist $c > 0, C \geqslant 1$, s.t. for all $0 \leqslant j < j' \leqslant N,\ x,x' \in \mathbb{R}^d$,

$$|H^\Delta(t_j,t_{j'},x,x')| \leqslant C(t_{j'}-t_j)^{-1+\eta/2}p_c(t_{j'}-t_j,x,x'). \tag{A.1}$$

Indeed this bound yields that for all $0 \leqslant j < j' \leqslant N,\ x,x' \in \mathbb{R}^d$

$$\begin{aligned}
|\widetilde{p}^\Delta \otimes_\Delta H^\Delta(t_j,t_{j'},x,x')| &\leqslant \Delta \sum_{k=0}^{j'-j-1} \int_{\mathbb{R}^d} \widetilde{p}^\Delta(t_j,t_{j+k},x,u)|H^\Delta(t_{j+k},t_{j'},u,x')|du\\
&\leqslant C^2\Delta \sum_{k=0}^{j'-j-1} (t_{j'}-t_{j+k})^{-1+\eta/2}p_c(t_{j'}-t_j,x,x')\\
&\leqslant C^2(t_{j'}-t_j)^{\eta/2}B\left(1,\frac{\eta}{2}\right)p_c(t_{j'}-t_j,x,x')
\end{aligned}$$

using the inequality $\widetilde{p}^\Delta(t_{j+k}-t_j,x,u)\leqslant Cp_c(t_j,t_{j+k},x,u)$ (cf. Lemma 3.1 of [15] in case (b)) and the semigroup property of p_c for the last but one inequality. The bound (4.14) then follows from the above control and (A.1) by induction.

Proof of (A.1). We consider two cases.

- $\mathbf{j'=j+1}$. From (4.12) we have in this case, $\forall x,x'\in\mathbb{R}^d$,

$$H^\Delta(t_j,t_{j'},x,x')=\Delta^{-1}(p^\Delta-\widetilde{p}^\Delta)(t_j,t_{j'},x,x')$$

which are Gaussian densities. In case (a) we have

$$H^\Delta(t_j,t_{j'},x,x')=\Delta^{-1}\Bigg(\frac{G\big((\sqrt{\Delta}\sigma(t_j,x))^{-1}(x'-x-b(t_j,x)\Delta)\big)}{\big(\Delta^d\det(a(t_j,x))\big)^{1/2}} - \frac{G\big((\sqrt{\Delta}\sigma(t_j,x'))^{-1}(x'-x-b(t_j,x')\Delta)\big)}{\big(\Delta^d\det(a(t_j,x'))\big)^{1/2}}\Bigg),$$

where $\forall z\in\mathbb{R}^d$, $G(z)=\exp(-|z|^2/2)(2\pi)^{-d/2}$ stands for the density of the standard Gaussian vector of $\mathbb{R}^d$. In case (b) we get

$$H^\Delta(t_j,t_{j'},x,x')=\Delta^{-1}(2\sqrt{3})^{d'}\times\left(\frac{G\left(\begin{pmatrix}(\Delta^{1/2}\sigma(t_j,x))^{-1}((x')^{1,d'}-x^{1,d'}-b_1(t_j,x)\Delta)\\ 2\sqrt{3}(\Delta^{3/2}\sigma(t_j,x))^{-1}((x')^{d'+1,d}-x^{d'+1,d}-\frac{x^{1,d'}+(x')^{1,d'}}{2}\Delta)\end{pmatrix}\right)}{\Delta^d\det(a(t_j,x))}\right.$$
$$\left.-\frac{G\left(\begin{pmatrix}(\Delta^{1/2}\sigma(t_j,(x')^\Delta))^{-1}((x')^{1,d'}-x^{1,d'}-b_1(t_j,x')\Delta)\\ 2\sqrt{3}(\Delta^{3/2}\sigma(t_j,(x')^\Delta))^{-1}((x')^{d'+1,d}-x^{d'+1,d}-\frac{x^{1,d'}+(x')^{1,d'}}{2}\Delta)\end{pmatrix}\right)}{\Delta^d\det(a(t_j,(x')^\Delta))}\right),$$

where $(x')^\Delta := x'-\begin{pmatrix}\mathbf{0}_{d'\times 1}\\ (x')^{1,d'}\Delta\end{pmatrix}$ allows to have good continuity properties to equilibrate the singularities coming from the difference $|x-(x')^\Delta|\leqslant|(x')^{1,d'}-x^{1,d'}|(1+\frac{\Delta}{2})+|(x')^{d'+1,d}-x^{d'+1,d}-\frac{x^{1,d'}+(x')^{1,d'}}{2}\Delta|$ with the terms appearing in the exponential. In all cases, tedious but elementary computations involving the mean value theorem yield that $\exists c>0, C\geqslant 1$ s.t. $|H^\Delta(t_j,t_{j'},x,x')|\leqslant C\Delta^{-1+\eta/2}p_c(\Delta,x,x')$.

- $\mathbf{j'>j+1}$. **Case *(a)***.

We first define for all $(j,x,z)\in[\![0,N]\!]\times(\mathbb{R}^d)^2$, the transition

$$\mathrm{T}_\Delta(t_j,x,z):=b(t_j,x)\Delta+\Delta^{1/2}\sigma(t_j,x)z.$$

The discrete convolution kernel then writes

$$H^{\Delta}(t_j,t_{j'},x,x')=\Delta^{-1}\int_{\mathbb{R}^d}G(z)\Big\{\Big(\widetilde{p}^{\Delta}(t_{j+1},t_{j'},x+\mathrm{T}_{\Delta}(t_j,x,z),x')-\widetilde{p}^{\Delta}(t_{j+1},t_{j'},x,x')\Big)$$
$$-\Big(\widetilde{p}^{\Delta}(t_{j+1},t_{j'},x+\mathrm{T}_{\Delta}(t_j,x',z),x')-\widetilde{p}^{\Delta}(t_{j+1},t_{j'},x,x')\Big)\Big\}\,\mathrm{d}z:=T_1^{(a)}-T_2^{(a)}.$$

Now exploiting that $\int_{\mathbb{R}^d}G(z)z\mathrm{d}z=0$, a Taylor expansion at order 3 of $T_1^{(a)},T_2^{(a)}$ yields

$$\begin{aligned}H^{\Delta}(t_j,t_{j'},x,x')&=\langle b(t_j,x)-b(t_j,x'),D_x\widetilde{p}^{\Delta}(t_{j+1},t_{j'},x,x')\rangle\\&\quad+\frac{1}{2}\mathrm{Tr}\Big(\big(a(t_j,x)-a(t_j,x')\big)D_x^2\widetilde{p}^{\Delta}(t_{j+1},t_{j'},x,x')\Big)\\&\quad+R^{\Delta}(t_j,t_{j'},x,x')\\&:=(H+R^{\Delta})(t_j,t_{j'},x,x').\end{aligned}\tag{A.2}$$

In the above equation H is the difference of the infinitesimal generators at time t_j of the processes $(X_t)_{t\geqslant 0}$ satisfying (1.1) and the Gaussian process $\tilde{X}_t=x+\int_{t_j}^t b(s,x')\mathrm{d}s+\int_{t_j}^t\sigma(s,x')\mathrm{d}W_s,\ t\geqslant t_j$, which can be seen as the continuous version of the frozen Markov chain $(\tilde{X}_{t_i}^{\Delta})_{i\in[\![j,N]\!]}$ introduced in (4.10), applied to the Gaussian density $\widetilde{p}^{\Delta}(t_{j+1},t_{j'},\cdot,x')$ at point x. The remainder term then writes

$$R^{\Delta}(t_j,t_{j'},x,x')=\frac{\Delta}{2}\mathrm{Tr}\Big(\big(bb^*(t_j,x)-bb^*(t_j,x')\big)D_x^2\widetilde{p}^{\Delta}(t_{j+1},t_{j'},x,x')\Big)+$$
$$3\Delta^{-1}\sum_{|\nu|=3}\int_{\mathbb{R}^d}\mathrm{d}zG(z)\int_0^1\mathrm{d}\delta(1-\delta)^2\Bigg[D_x^{\nu}\widetilde{p}^{\Delta}(t_{j+1},t_{j'},x+\delta\,\mathrm{T}_{\Delta}(t_j,x,z),x')\frac{\big(\mathrm{T}_{\Delta}(t_j,x,z)\big)^{\nu}}{\nu!}$$
$$-D_x^{\nu}\widetilde{p}^{\Delta}(t_{j+1},t_{j'},x+\delta\,\mathrm{T}_{\Delta}(t_j,x',z),x')\frac{\big(\mathrm{T}_{\Delta}(t_j,x',z)\big)^{\nu}}{\nu!}\Bigg]$$

using the following notations for multi-indices and powers. For $\nu=(\nu_1,...,\nu_d)\in\mathbb{N}^d,\ x=(x_1,...,x_d)^*$ set $|\nu|=\nu_1+...+\nu_d,\ \nu!=\nu_1!...\nu_d!,\ (x)^{\nu}=x_1^{\nu_1}...\ x_d^{\nu_d},\ D_x^{\nu}=D_{x_1}^{\nu_1}...D_{x_d}^{\nu_d}$. Recalling the standard control

$$\exists c>0,C\geqslant 1,\ \forall\nu,|\nu|\leqslant 4,\forall 0\leqslant j<j'\leqslant N,\ (x,x')\in(\mathbb{R}^d)^2,$$
$$|D_x^{\nu}\widetilde{p}^{\Delta}(t_j,t_{j'},x,x')|\leqslant C(t_{j'}-t_j)^{-|\nu|/2}p_c(t_{j'}-t_j,x,x')\tag{A.3}$$

for the derivatives of Gaussian densities, we obtain:

$$\begin{aligned}|R^{\Delta}(t_j,t_{j'},x,x')|&\leqslant C\Delta|b|_{\infty}^2(t_{j'}-t_{j+1})^{-1}p_c(t_{j'}-t_{j+1},x,x')+3\Delta^{-1}\Bigg|\sum_{|\nu|=3}\int_{\mathbb{R}^d}\mathrm{d}zG(z)\int_0^1\mathrm{d}\delta(1-\delta)^2\times\\&\Bigg[D_x^{\nu}\widetilde{p}^{\Delta}(t_{j+1},t_{j'},x+\delta\,\mathrm{T}_{\Delta}(t_j,x,z),x')\left(\frac{(\mathrm{T}_{\Delta}(t_j,x,z))^{\nu}}{\nu!}-\frac{(\mathrm{T}_{\Delta}(t_j,x',z))^{\nu}}{\nu!}\right)\\&-\Big(D_x^{\nu}\widetilde{p}^{\Delta}(t_{j+1},t_{j'},x+\delta\,\mathrm{T}_{\Delta}(t_j,x',z),x')-D_x^{\nu}\widetilde{p}^{\Delta}(t_{j+1},t_{j'},x+\delta\,\mathrm{T}_{\Delta}(t_j,x,z),x')\Big)\times\frac{(\mathrm{T}_{\Delta}(t_j,x',z))^{\nu}}{\nu!}\Bigg]\Bigg|.\end{aligned}\tag{A.4}$$

Note now that,

$$\exists C>0,\ \forall \nu,\ |\nu|\leqslant 3,\ \forall (j,x',z)\in [\![0,N]\!]\times(\mathbb{R}^d)^2,\ |(\mathrm{T}_\Delta(t_j,x',z))^\nu|\leqslant C(\Delta^{|\nu|}+\Delta^{|\nu|/2}|z|^{|\nu|}),$$
$$\forall (j,x,x',z)\in [\![0,N]\!]\times(\mathbb{R}^d)^3,\ |\mathrm{T}_\Delta(t_j,x,z)-\mathrm{T}_\Delta(t_j,x',z)|\leqslant C(\Delta+\Delta^{1/2}|x-x'|^\eta|z|),$$

and that

$$\begin{aligned}\forall \nu, |\nu|=3,\ |(\mathrm{T}_\Delta(t_j,x',z))^\nu-(\mathrm{T}_\Delta(t_j,x,z))^\nu| &\leqslant C(\Delta+\Delta^{1/2}|x-x'|^\eta|z|)(\Delta^2+\Delta|z|^2)\\ &\leqslant C(\Delta^3+\Delta^2|z|^2+\Delta^{3/2}|x-x'|^\eta(\Delta|z|+|z|^3)).\end{aligned}$$

Hence, plugging these controls in (A.4) and using (A.3) we get:

$$\begin{aligned}|R^\Delta(t_j,t_{j'},x,x')|&\leqslant C\Bigg[p_c(t_{j'}-t_{j+1},x,x')\\ &+\int_{\mathbb{R}^d}\mathrm{d}zG(z)\int_0^1\mathrm{d}\delta(1-\delta)^2p_c(t_{j+1},t_{j'},x+\delta\,\mathrm{T}_\Delta(t_j,x,z),x')\frac{\left[\Delta^2+\Delta|z|^2+\Delta^{1/2}|x-x'|^\eta(\Delta|z|+|z|^3)\right]}{(t_{j'}-t_{j+1})^{3/2}}\\ &+\int_{\mathbb{R}^d}\mathrm{d}zG(z)\int_{[0,1]^2}\mathrm{d}\delta\mathrm{d}\gamma(1-\delta)^2p_c(t_{j+1},t_{j'},x+\delta\,\mathrm{T}_\Delta(t_j,x,z)+\gamma\delta(\mathrm{T}_\Delta(t_j,x',z)-\mathrm{T}_\Delta(t_j,x,z)),x')\\ &\times\frac{\left[(\Delta+\Delta^{1/2}|x-x'|^\eta|z|)(\Delta^2+\Delta^{1/2}|z|^3)\right]}{(t_{j'}-t_{j+1})^2}\Bigg].\end{aligned}$$

Now, using the inequality $\forall\varepsilon\in(0,1),\ |x-x'+\rho|^2\geqslant|x-x'|^2(1-\varepsilon)+|\rho|^2(1-\varepsilon^{-1}),\forall\rho\in\mathbb{R}^d$, taking $\rho=\delta\,\mathrm{T}_\Delta(t_j,x,z)$ and $\rho=\delta\,\mathrm{T}_\Delta(t_j,x,z)+\gamma\delta[\mathrm{T}_\Delta(t_j,x',z)-\mathrm{T}_\Delta(t_j,x,z)]$ respectively in the first and second integral we derive

$$\begin{aligned}|R^\Delta(t_j,t_{j'},x,x')|&\leqslant\frac{C}{(1-\varepsilon)^{d/2}}p_{(1-\varepsilon)c}(t_{j'}-t_{j+1},x,x')\left\{1+\int_{\mathbb{R}^d}\mathrm{d}zG(z)\exp\left(c\frac{|\sigma|_\infty^2|z|^2\Delta}{t_{j'}-t_{j+1}}(\varepsilon^{-1}-1)\right)\right.\\ &\times\left[\left(\Delta^{1/2}+\frac{|z|^2}{(t_{j'}-t_{j+1})^{1/2}}+|x'-x|^\eta\Big(\frac{|z|^3}{(t_{j'}-t_{j+1})}+|z|\Big)\right)\right.\\ &\left.\left.+\left(\Delta+\frac{|z|^3}{(t_{j'}-t_{j+1})^{1/2}}+|x'-x|^\eta\Big(\Delta^{1/2}|z|+\frac{|z|^4}{(t_{j'}-t_{j+1})}\Big)\right)\right]\right\}.\end{aligned}$$

Choosing ε sufficiently close to one the above integrals w.r.t. z are finite and therefore for different c,C depending on ε as well, we have

$$\begin{aligned}|R^\Delta(t_j,t_{j'},x,x')|&\leqslant Cp_c(t_{j'}-t_{j+1},x,x')\left(1+\frac{1}{(t_{j'}-t_{j+1})^{1/2}}+\frac{|x-x'|^\eta}{(t_{j'}-t_{j+1})}\right)\\ &\leqslant C(t_{j'}-t_{j+1})^{-1+\eta/2}p_c(t_{j'}-t_{j+1},x,x').\end{aligned}\tag{A.5}$$

Now with the definitions of (A.2) we also have from (A.3)

$$|H(t_j,t_{j'},x,x')|\leqslant C(t_{j'}-t_{j+1})^{-1+\eta/2}p_c(t_{j'}-t_{j+1},x,x').$$

Plugging this last estimate and (A.5) in (A.2) we derive

$$|H^\Delta(t_j,t_{j'},x,x')|\leqslant C(t_{j'}-t_{j+1})^{-1+\eta/2}p_c(t_{j'}-t_{j+1},x,x')\leqslant C(t_{j'}-t_j)^{-1+\eta/2}p_c(t_{j'}-t_j,x,x').$$

Case *(b)*. We first introduce some notations. Set $x_\Delta := x + \begin{pmatrix} \mathbf{0}_{d'\times 1} \\ x^{1,d'}\Delta \end{pmatrix}$, $(x')_{\Delta,j,j'} := x' - \begin{pmatrix} \mathbf{0}_{d'\times 1} \\ (x')^{1,d'} \end{pmatrix}(t_{j'} - t_j)$. Define $\forall y \in \mathbb{R}^d$, $B^\Delta(t_j, y) := \begin{pmatrix} b_1(t_j, y)\Delta \\ b_1(t_j, y)\Delta^2/2 \end{pmatrix}$, $\Sigma^\Delta(t_j, y) := \begin{pmatrix} \Delta^{1/2}\sigma(t_j, y) & 0 \\ \Delta^{3/2}\sigma(t_j, y)/2 & \Delta^{3/2}\sigma(t_j, y)/(2\sqrt{3}) \end{pmatrix}$. Introducing for all $(j, x, y, z) \in [\![0, N]\!] \times (\mathbb{R}^d)^3$ the transition

$$\mathrm{T}_\Delta(t_j, x, y, z) := B^\Delta(t_j, x) + \Sigma^\Delta(t_j, y)z,$$

we have

$$\begin{aligned} H^\Delta(t_j, t_{j'}, x, x') = \Delta^{-1}\int_{\mathbb{R}^d} G(z)\Big\{\Big(\widetilde{p}^\Delta(t_{j+1}, t_{j'}, x_\Delta + \mathrm{T}_\Delta(t_j, x, x, z), x') - \widetilde{p}^\Delta(t_{j+1}, t_{j'}, x_\Delta, x')\Big) \\ - \Big(\widetilde{p}^\Delta(t_{j+1}, t_{j'}, x_\Delta + \mathrm{T}_\Delta(t_j, x', (x')_{\Delta,j,j'}, z), x') - \widetilde{p}^\Delta(t_{j+1}, t_{j'}, x_\Delta, x')\Big)\Big\}\mathrm{d}z := T_1^{(b)} - T_2^{(b)}. \end{aligned}$$

The strategy now relies as in case (a) on Taylor expansions. Let us first perform an exact Taylor expansion of $T_1^{(b)}, T_2^{(b)}$ around the point x_Δ at order one, separating the components from 1 to d' and $d'+1$ to d. We obtain

$$\begin{aligned} &H^\Delta(t_j, t_{j'}, x, x') = \\ &\quad\Bigg\{\Delta^{-1}\int_{\mathbb{R}^d}\mathrm{d}z G(z)\int_0^1 \mathrm{d}\gamma\Big\{\Big\langle D_{x^{1,d'}}\widetilde{p}^\Delta(t_{j+1}, t_{j'}, x_\Delta + \gamma\,\mathrm{T}_\Delta(t_j, x, x, z), x'), \Big(\mathrm{T}_\Delta(t_j, x, x, z)\Big)^{1,d'}\Big\rangle \\ &\qquad - \Big\langle D_{x^{1,d'}}\widetilde{p}^\Delta(t_{j+1}, t_{j'}, x_\Delta + \gamma\,\mathrm{T}_\Delta(t_j, x', (x')_{\Delta,j,j'}, z), x'), \Big(\mathrm{T}_\Delta(t_j, x', (x')_{\Delta,j,j'}, z)\Big)^{1,d'}\Big\rangle\Big\}\Bigg\} \\ &+\Bigg\{\Delta^{-1}\int_{\mathbb{R}^d}\mathrm{d}z G(z)\int_0^1 \mathrm{d}\gamma\Big\{\Big\langle D_{x^{d'+1,d}}\widetilde{p}^\Delta(t_{j+1}, t_{j'}, x_\Delta + \gamma\,\mathrm{T}_\Delta(t_j, x, x, z), x'), \Big(\mathrm{T}_\Delta(t_j, x, x, z)\Big)^{d'+1,d}\Big\rangle \\ &\qquad - \Big\langle D_{x^{d'+1,d}}\widetilde{p}^\Delta(t_{j+1}, t_{j'}, x_\Delta + \gamma\,\mathrm{T}_\Delta(t_j, x', (x')_{\Delta,j,j'}, z), x'), \Big(\mathrm{T}_\Delta(t_j, x', (x')_{\Delta,j,j'}, z)\Big)^{d'+1,d}\Big\rangle\Big\}\Bigg\} \\ &\qquad\qquad := (M_1^\Delta + R_1^\Delta)(t_j, t_{j'}, x, x'), \end{aligned} \tag{A.6}$$

where $D_{x^{1,d'}}$ (resp. $D_{x^{d'+1,d}}$) denotes the differentiation w.r.t. the first d' (resp. the last d') components. Expanding the terms $D_{x^{1,d'}}\widetilde{p}^\Delta(t_{j+1}, t_{j'}, x_\Delta + \gamma\,\mathrm{T}_\Delta(t_j, x, x, z), x'), D_{x^{1,d'}}\widetilde{p}^\Delta(t_{j+1}, t_{j'}, x_\Delta + \gamma\,\mathrm{T}_\Delta(t_j, x', (x')_{\Delta,j,j'}, z), x')$ at order 2 around x_Δ in M_1^Δ, we get:

$$\begin{aligned} H^\Delta(t_j, t_{j'}, x, x') &= \langle b_1(t_j, x) - b_1(t_j, x'), D_{x^{1,d'}}\widetilde{p}^\Delta(t_{j+1}, t_{j'}, x_\Delta, x')\rangle \\ &\quad+\frac{1}{2}\mathrm{Tr}\Big\{\Big(a(t_j, x) - a\Big(t_j, (x')_{\Delta,j,j'}\Big)\Big)D^2_{x^{1,d'}}\widetilde{p}^\Delta(t_{j+1}, t_{j'}, x_\Delta, x')\Big\} \\ &\quad+(R_1^\Delta + R_2^\Delta)(t_j, t_{j'}, x, x') := H(t_j, t_{j'}, x_\Delta, x') + (R_1^\Delta + R_2^\Delta)(t_j, t_{j'}, x, x'). \end{aligned} \tag{A.7}$$

Here, similarly to (A.2), H is the difference of the generators at time t_j of the processes $(X_t)_{t\geqslant 0}$ satisfying (1.1) and the Gaussian process $\tilde{X}_t = x + \int_{t_j}^t \begin{pmatrix} b_1(s, x') \\ (\tilde{X}_s)^{1,d'} \end{pmatrix}\mathrm{d}s + \int_{t_j}^t B\sigma\left(s, x' - \begin{pmatrix} \mathbf{0}_{d'\times 1} \\ (x')^{1,d'} \end{pmatrix}(t_{j'} - s)\right)\mathrm{d}W_s$, $t \in [t_j, t_{j'}]$ (continuous version of $(\tilde{X}^\Delta_{t_i})_{i\in[\![j,j']\!]}$ introduced in (4.11)), applied to the Gaussian density $\widetilde{p}^{\Delta}(t_{j+1}, t_{j'}, \cdot, x')$ at point x_Δ. The second remainder term R_2^{Δ} writes $R_2^{\Delta}(t_j, t_{j'}, x, x') = (R_{21}^{\Delta} + R_{22}^{\Delta})(t_j, t_{j'}, x, x')$ where

$$
\begin{aligned}
R_{21}^{\Delta}(t_j, t_{j'}, x, x') &= \frac{\Delta}{2} \mathrm{Tr}\Big(\big(b_1 b_1^*(t_j, x) - b_1 b_1^*(t_j, x')\big) D^2_{x^{1,d'}} \widetilde{p}^{\Delta}(t_{j+1}, t_{j'}, x_\Delta, x') \Big) \\
&\quad + \frac{\Delta}{4}\Big(\mathrm{Tr}\Big(\big(b_1 b_1^*(t_j, x) - b_1 b_1^*(t_j, x')\big)\Delta + \big(a(t_j, x) - a(t_j, (x')_{\Delta,j,j'})\big) \Big) \\
&\qquad \times D^2_{x^{1,d'}, x^{d'+1,d}} \widetilde{p}^{\Delta}(t_{j+1}, t_{j'}, x_\Delta, x') \Big)
\end{aligned}
$$

and

$$
\begin{aligned}
R_{22}^{\Delta}(t_j, t_{j'}, x, x') &= 2\Delta^{-1} \sum_{|\theta|=2} \int_{\mathbb{R}^d} \mathrm{d}z G(z) \int_{[0,1]^2} \mathrm{d}\gamma \mathrm{d}\delta (1-\delta)\gamma^2 \\
&\times \Bigg[\Big\langle D^\theta D_{x^{1,d'}} \widetilde{p}^{\Delta}(t_{j+1}, t_{j'}, x_\Delta + \delta\gamma\, \mathrm{T}_\Delta(t_j, x, x, z), x') \frac{\big(\mathrm{T}_\Delta(t_j, x, x, z)\big)^\theta}{\theta!}, \mathrm{T}_\Delta(t_j, x, x, z)^{1,d'} \Big\rangle \\
&- \Big\langle D^\theta D^\nu_{x^{1,d'}} \widetilde{p}^{\Delta}(t_{j+1}, t_{j'}, x_\Delta + \delta\gamma\, \mathrm{T}_\Delta(t_j, x', (x')_{\Delta,j,j'}, z), x') \frac{\big(\mathrm{T}_\Delta(t_j, x', (x')_{\Delta,j,j'}, z)\big)^\theta}{\theta!}, \\
&\qquad (\mathrm{T}_\Delta(t_j, x', (x')_{\Delta,j,j'}, z)^{1,d'} \Big\rangle \Bigg]
\end{aligned}
$$

Let $\mu = (\mu_1, \cdots, \mu_{d'}) \in \mathbb{N}^{d'},\ \nu = (\nu_1, \cdots, \nu_{d'}) \in \mathbb{N}^{d'}$ be multi-indices. Similarly to (A.3) we have,

$$
\begin{aligned}
&\exists c > 0, C \geqslant 1,\ \forall (\mu, \nu),\ |\mu| \leqslant 3, |\nu| \leqslant 4, \forall 0 \leqslant j < j' \leqslant N,\ (x, x') \in \mathbb{R}^d \times \mathbb{R}^d, \\
&\qquad |D^\nu_{x^{1,d'}} D^\mu_{x^{d'+1,d}} \widetilde{p}^{\Delta}(t_j, t_{j'}, x, x')| \leqslant C(t_{j'} - t_j)^{-(|\nu|/2 + 3/2|\mu|)} p_c(t_{j'} - t_j, x, x'). \quad \text{(A.8)}
\end{aligned}
$$

Observe as well that there exists $C > 0$ s.t.

$$
|\,\mathrm{T}_\Delta(t_j, x, x, z)^{d'+1,d} - \mathrm{T}_\Delta(t_j, x, (x')_{\Delta,j,j'}, z)^{d'+1,d}| \leqslant C(\Delta^2 + |x - (x')_{\Delta,j,j'}|^\eta \Delta^{3/2}|z|).
$$

From (A.8), and the above equation, proceeding as in case (a), we get from (A.6) for ε sufficiently close to 1:

$$
\begin{aligned}
|R_1^{\Delta}(t_j, t_{j'}, x, x')| &\leqslant \frac{C}{(1-\varepsilon)^{d/2}} p_{(1-\varepsilon)c}(t_{j'} - t_{j+1}, x_\Delta, x') \Bigg\{ \int \mathrm{d}z G(z) \exp\Big(2c|\sigma|_\infty^2 |z|^2 (\varepsilon^{-1} - 1)\Big) \times \\
&\Bigg(\frac{1}{(t_{j'} - t_{j+1})^{1/2}} + \frac{|x - (x')_{\Delta,j,j'}|^\eta |z|}{(t_{j'} - t_{j+1})} \Bigg) + \Bigg(1 + \frac{|z|}{(t_{j'} - t_j)^{1/2}} + \frac{|x - (x')_{\Delta,j,j'}|^\eta |z|}{(t_{j'} - t_{j+1})^{1/2}} + \frac{|x - (x')_{\Delta,j,j'}|^\eta |z|^2}{t_{j'} - t_{j+1}} \Bigg) \Bigg\} \\
&\leqslant \frac{C}{(1-\varepsilon)^{d/2}} p_{(1-\varepsilon)c}(t_{j'} - t_{j+1}, x_\Delta, x') \Bigg(\frac{1}{(t_{j'} - t_{j+1})^{1/2}} + \frac{|x - (x')_{\Delta,j,j'}|^\eta}{t_{j'} - t_{j+1}} \Bigg). \quad \text{(A.9)}
\end{aligned}
$$

To conclude the proof, one has now to bound the terms $|x - (x')_{\Delta,j,j'}|^\eta$ deriving from the difference of the transitions in the above expressions with quantities that appear as well in the off-diagonal bounds of the density $p_{(1-\varepsilon)c}(t_{j'}-t_{j+1},x_\Delta,x')$. This allows to compensate the singularities in time of the remainders. We write:

$$
\begin{aligned}
|x-(x')_{\Delta,j,j'}|^\eta &\leqslant C(|x^{1,d'}-(x')^{1,d'}|^\eta+|x^{d'+1,d}-(x')^{d'+1,d}-(x')^{1,d}(t_{j'}-t_j)|^\eta)\\
&\leqslant C\left(|x^{1,d'}-(x')^{1,d'}|^\eta\left(1+\left(\frac{t_{j'}-t_j}{2}\right)^\eta\right)+|x^{d'+1,d}-(x')^{d'+1,d}-\frac{x^{1,d}+(x')^{1,d}}{2}(t_{j'}-t_j)|^\eta\right)\\
&\leqslant C\left(|x^{1,d'}-(x')^{1,d'}|^\eta+|x^{d'+1,d}-(x')^{d'+1,d}-\frac{x^{1,d}+(x')^{1,d}}{2}(t_{j'}-t_j)|^\eta\right). \qquad \text{(A.10)}
\end{aligned}
$$

On the other hand, the off-diagonal term of $p_{(1-\varepsilon)c}(t_{j'}-t_{j+1},x_\Delta,x')$ associated to components $d'+1,d$ writes:

$$
\begin{aligned}
&\frac{|(x')^{d'+1,d}-x_\Delta^{d'+1,d}-\frac{(x')^{1,d'}+x_\Delta^{1,d'}}{2}(t_{j'}-t_{j+1})|^2}{(t_{j'}-t_{j+1})^3}=\frac{|(x')^{d'+1,d}-x^{d'+1,d}-\frac{(x')^{1,d'}+x^{1,d'}}{2}(t_{j'}-t_{j+1})-x^{1,d'}\Delta|^2}{(t_{j'}-t_{j+1})^3}\\
&=\frac{|(x')^{d'+1,d}-x^{d'+1,d}-\frac{(x')^{1,d'}+x^{1,d'}}{2}(t_{j'}-t_j)-\frac{(x^{1,d'}-(x')^{1,d'})}{2}\Delta|^2}{(t_{j'}-t_{j+1})^3}\\
&\geqslant(1-\zeta)\frac{|(x')^{d'+1,d}-x^{d'+1,d}-\frac{(x')^{1,d'}+x^{1,d'}}{2}(t_{j'}-t_j)|^2}{(t_{j'}-t_j)^3}+(1-\zeta^{-1})\frac{|(x')^{1,d'}-x^{1,d'}|^2\Delta^2}{4(t_{j'}-t_j)^3},\zeta\in(0,1).
\end{aligned}
$$
(A.11)

Use now equations (A.10), (A.11) and the bound $\frac{1}{(t_{j'}-t_{j+1})}\leqslant\frac{2}{t_{j'}-t_j}$ (since $j'>j+1$) to obtain, for ζ sufficiently close to 1, from (A.9):

$$|R_1^\Delta(t_j,t_{j'},x,x')|\leqslant C(t_{j'}-t_j)^{-1+\frac{\eta}{2}}p_c(t_{j'}-t_j,x,x'),$$

up to modifications of C,c that both depend on ε,ζ.

The same line of reasoning yields $|R_{21}^\Delta(t_j,t_{j'},x,x')|\leqslant C\left(1+\frac{|x-(x')_{\Delta,j,j'}|^\eta}{(t_{j'}-t_{j+1})}\right)p_c(t_{j'}-t_{j+1},x_\Delta,x')\leqslant C(1+(t_{j'}-t_j)^{-1+\eta/2})p_c(t_{j'}-t_j,x,x')$. Proceeding as in case (a), using (A.10), (A.8) to handle $R_{22}^\Delta(t_j,t_{j'},x,x')$, we eventually derive that $|R_2^\Delta(t_j,t_{j'},x,x')|\leqslant C(t_{j'}-t_j)^{-1+\eta/2}p_c(t_{j'}-t_j,x,x')$. From the definition in equation (A.7) we also obtain $|H(t_j,t_{j'},x,x')|\leqslant C(t_{j'}-t_j)^{-1+\eta/2}p_c(t_{j'}-t_j,x,x')$ using (A.10), (A.8) and the statement (A.1) follows.

Remark A.1. *Note that the time dependence in the frozen dynamics* (4.11) *somehow corresponds to the backward transport of the terminal condition. It is crucial in order to allow from* (A.10) *the compensation of the exploding terms associated to derivatives in* $x^{1,d'}$ *of order greater than* 2 *and derivatives in* $x^{d'+1,d}$ *of order greater than* 1 *appearing in the kernel* H^Δ. *A similar construction was used in [15].*

A.2 Proof of Lemma 4.3

Let us first prove (4.31). We begin with $(x_i, x_{i+1}) \in B_i \times B_{i+1},\ i \in [\![1, \bar L - 2]\!]$. From Section 4.4, one can check that $|Q_{\varepsilon_i}^{-1/2}(R(s_{i+1}, s_i)x_i - x_{i+1})|^2 = 2d_{\varepsilon_i}^2(x_i, x_{i+1})$. Hence,

$$\begin{aligned}Q_i &:= d_{\varepsilon_i}(x_i, x_{i+1}) = \frac{1}{\sqrt{2}}|Q_{\varepsilon_i}^{-1/2}(R(s_{i+1}, s_i)x_i - x_{i+1})| \leqslant c|Q_{\varepsilon_i}^{-1/2}(x_i - R(s_i, s_{i+1})x_{i+1})|\\ &\leqslant c\{|Q_{\varepsilon_i}^{-1/2}(x_i - R(s_i, s_{i-1})y_{i-1})| + |Q_{\varepsilon_i}^{-1/2}(R(s_i, s_{i-1})y_{i-1} - y_i)| + |Q_{\varepsilon_i}^{-1/2}(y_i - R(s_i, s_{i+1})x_{i+1})|\}\\ &:= Q_i^1 + Q_i^2 + Q_i^3.\end{aligned}$$

One has

$$Q_i^1 \leqslant c\sum_{j=1}^{2}\varepsilon_i^{1/2-j}|(x_i - R(s_i, s_{i-1})y_{i-1})_j| \leqslant c\sum_{j=1}^{2}\left(\frac{\varepsilon_i}{K\rho^2}\right)^{1/2-j}(K^{1/2}\rho)^{1-2j}|(x_i - R(s_i, s_{i-1})y_{i-1})_j|,$$

denoting for all $z \in \mathbb{R}^d, z_1 := z^{1,d'}, z_2 := z^{d'+1,d}$ with a slight abuse of notation. Now, from (4.30), $\varepsilon_i/(K\rho^2) \geqslant c_1((t_{j'} - t_j)/L)/(Kd_{t_{j'}-t_j}^2(x, x')(t_{j'} - t_j)/L^2) = c_1\frac{L}{Kd_{t_{j'}-t_j}^2(x,x')}$. Thus, recalling $L = \lceil Kd_{t_{j'}-t_j}^2(x, x')\rceil$, $\exists c > 0,\ \forall j \in [\![1, 2]\!],\ \left(\frac{\varepsilon_i}{K\rho^2}\right)^{1/2-j} \leqslant c$ and

$$Q_i^1 \leqslant c\sum_{j=1}^{2}(K^{1/2}\rho)^{1-2j}|(x_i - R(s_i, s_{i-1})y_{i-1})_j| \leqslant c|Q_{K\rho^2}^{-1/2}(x_i - R(s_i, s_{i-1})y_{i-1})| \leqslant cR_0K^{-1/2},$$

exploiting $x_i \in B_i$ for the last identity. The term Q_i^3 could be handled in a similar way so that $Q_i^1 + Q_i^3 \leqslant cR_0K^{-1/2}$. Now $Q_i^2 := \sqrt{2}d_{\varepsilon_i}(y_i, y_{i+1}) \leqslant I(s_i, s_{i+1}, y_i, y_{i+1})^{1/2} \leqslant c\left(\int_{s_i}^{s_{i+1}}|\varphi_s|^2 \mathrm{d}s\right)^{1/2} \leqslant c\frac{d_{t_{j'}-t_j}(x,x')}{L^{1/2}} \leqslant \frac{c}{K^{1/2}}$. Hence, for all $i \in [\![1, \bar L - 2]\!]$, $Q_i \leqslant 2R_0$ for K large enough independent of $t_{j'} - t_j$. Eventually, for $x_1 \in B_1, x_{\bar L-1} \in B_{\bar L-1}$ the terms $Q_0 := |Q_{\varepsilon_0}^{-1/2}(R(s_1, 0)x - x_1)|$ and $Q_{\bar L-1} := |Q_{\varepsilon_{\bar L-1}}^{-1/2}(R(t_{j'} - t_j, s_{\bar L-1})x_{\bar L-1} - x')| \leqslant c|Q_{\varepsilon_{\bar L-1}}^{-1/2}(x_{\bar L-1} - R(s_{\bar L-1}, t_{j'} - t_j)x')|$ can be controlled as the previous $Q_i^1,\ i \in [\![1, \bar L - 2]\!]$ from the definitions of $B_1, B_{\bar L-1}$, so that $Q_i \leqslant 2R_0,\ i \in \{0, \bar L - 1\}$ as well. This proves (4.31).

It now remains to control the Lebesgue measure of the sets $(B_i)_{i\in[\![1,\bar L-1]\!]}$. Define for all $i \in [\![1, \bar L - 1]\!]$, $E_i := \{z \in \mathbb{R}^d : |Q_{K\rho^2}^{-1/2}(y_i - z)| \leqslant 2R_0(3K^{1/2})^{-1}\}$. One has $\exists \check c := \check c(d) > 0,\ |E_i| \geqslant \check c\rho^{2d}$. Let us now prove $E_i \subset B_i$. Write, for all $z \in E_i$,

$$\begin{aligned}R_i &:= |Q_{K\rho^2}^{-1/2}(R(s_i, s_{i-1})y_{i-1} - z)| + |Q_{K\rho^2}^{-1/2}(z - R(s_i, s_{i+1})y_{i+1})|\\ &\leqslant |Q_{K\rho^2}^{-1/2}(R(s_i, s_{i-1})y_{i-1} - y_i)| + 2|Q_{K\rho^2}^{-1/2}(y_i - z)| + |Q_{K\rho^2}^{-1/2}(y_i - R(s_i, s_{i+1})y_{i+1})|\\ &:= R_i^1 + R_i^2 + R_i^3.\end{aligned}$$

The previous definition of E_i gives $R_i^2 \leqslant \frac{4R_0}{3K^{1/2}}$. Now, arguments similar to those used to control the above $(Q_i^1, Q_i^2)_{i\in[\![1,M-2]\!]}$ yield

$$R_i^1 \leqslant c\sum_{j=1}^{2}\left(\frac{\varepsilon_i}{K\rho^2}\right)^{j-1/2}\varepsilon_i^{1/2-j}|(R(s_i, s_{i-1})y_{i-1} - y_i)_j| \leqslant c\frac{d_{t_{j'}-t_j}(x', x)}{L^{1/2}} \leqslant \frac{c}{K^{1/2}}.$$

Since the term R_i^3 could be handled in the same way we deduce that for K large enough and R_0 large enough w.r.t. the above c, $R_i \leqslant 2R_0K^{-1/2}$. Hence $E_i \subset B_i$ which completes the proof. □

References

[1] V. Bally and D. Talay, *The law of the Euler scheme for stochastic differential equations: I. Convergence rate of the distribution function*, Prob. Th. Rel. Fields **104-1** (1996), 43–60. MR1367666

[2] ______, *The law of the Euler scheme for stochastic differential equations, II. Convergence rate of the density*, Monte-Carlo methods and Appl. **2** (1996), 93–128. MR1401964

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

[4] Sergey G. Bobkov, Ivan Gentil, and Michel Ledoux, *Hypercontractivity of Hamilton-Jacobi equations*, J. Math. Pures Appl. (9) **80** (2001), no. 7, 669–696. MR1846020

[5] J. M. Coron, *Control and nonlinearity*, Mathematical Surveys and Monographs, 136, AMS, 2007. MR2302744

[6] François Delarue and Stéphane Menozzi, *Density estimates for a random noise propagating through a chain of differential equations*, Journal of Functional Analysis **259** (2010), no. 6, 1577 – 1630.

[7] A. Friedman, *Partial differential equations of parabolic type*, Prentice-Hall, 1964. MR0181836

[8] ______, *Stochastic differential equations*, Chapmann-Hall, 1975.

[9] E. Gobet and C. Labart, *Sharp estimates for the convergence of the density of the euler scheme in small time*, Elect. Comm. in Probab **13** (2008), 352–363. MR2415143

[10] A. M. Il'in, A. S. Kalashnikov, and O. A. Oleinik, *Second-order linear equations of parabolic type*, Uspehi Mat. Nauk **17–3(105)** (1962), 3–146. MR0138888

[11] A. Joulin and Y. Ollivier, *Curvature, concentration, and error estimates for Markov chain Monte Carlo*, http://arxiv.org/abs/0904.1312. To appear in Annals of Probability (2010).

[12] A. N. Kolmogorov, *Zufällige Bewegungen (zur Theorie der Brownschen Bewegung)*, Ann. of Math. **2-35** (1934), 116–117. MR1503147

[13] V. Konakov and E. Mammen, *Local limit theorems for transition densities of Markov chains converging to diffusions*, Prob. Th. Rel. Fields **117** (2000), 551–587. MR1777133

[14] ______, *Edgeworth type expansions for euler schemes for stochastic differential equations*, Monte Carlo Methods Appl. **8–3** (2002), 271–285. MR1931967

[15] V. Konakov, S. Menozzi, and S. Molchanov, *Explicit parametrix and local limit theorems for some degenerate diffusion processes*, To appear in Annales de l'Institut Henri Poincaré, Série B (2010).

[16] S. Kusuoka and D. Stroock, *Applications of the Malliavin calculus. III*, J. Fac. Sci. Univ. Tokyo Sect. IA Math **34** (1987), 391–442. MR0914028

[17] M. Ledoux, *Concentration of measure and logarithmic sobolev inequalities*, Séminaire de Probabilités XXXIII. LNM, 1709 (1999), 120–216. MR1767995

[18] F. Malrieu and D. Talay, *Concentration inequalities for euler schemes*, Monte Carlo and Quasi-Monte Carlo Methods 2004, H. Niederreiter and D. Talay (Eds.) (2006), 355–37. MR2208718

[19] H. P. McKean and I. M. Singer, *Curvature and the eigenvalues of the Laplacian.*, J. Differential Geometry **1** (1967), 43–69. MR0217739

[20] D. Nualart, *Malliavin calculus and related topics*, Springer Verlag, 1995. MR1344217

[21] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), no. 2, 361–400. MR1760620

[22] S. J. Sheu, *Some estimates of the transition density of a nondegenerate diffusion Markov process*, Ann. Probab. **19–2** (1991), 538–561. MR1106275

[23] D.W. Stroock and S.R.S. Varadhan, *Multidimensional diffusion processes*, Springer-Verlag Berlin Heidelberg New-York, 1979. MR0532498

[24] D. Talay and L. Tubaro, *Expansion of the global error for numerical schemes solving stochastic differential equations*, Stoch. Anal. and App. **8-4** (1990), 94–120. MR1091544