Evolution of Gaussian Concentration

J.-R Chazottes, P. Collet, F Redig

Evolution of Gaussian Concentration

J.-R. Chazottes *[1], P. Collet †[1], and F. Redig ‡[2]

[1]Centre de Physique Théorique, CNRS UMR 7644, F-91128 Palaiseau Cedex, France
[2]Delft Institute of Applied Mathematics, Delft University of Technology, van Mourik Broekmanweg 6, 2628 XE, Delft, The Netherlands

March 19, 2019

Abstract

We study the behavior of the Gaussian concentration bound (GCB) under stochastic time evolution. More precisely, in the context of a diffusion process on $\mathbb{R}^d$ we prove in various settings that if we start the process from an initial probability measure satisfying GCB, then at later times GCB holds, and estimates for the constant are provided. Under additional conditions, we show that GCB holds for the unique invariant measure. This yields a semigroup interpolation method to prove Gaussian concentration for measures which are not available in explicit form. We also consider diffusions "coming down from infinity" for which we show that, from any starting measure, at positive times, GCB holds.

Keywords: Gaussian concentration bound, diffusion processes, Ornstein-Uhlenbeck process, non Markovian diffusion, nonlinear semigroup, coupling, diffusion coming from infinity.

Contents

*Email: chazottes@cpht.polytechnique.fr
†Email: collet@cpht.polytechnique.fr
‡Email: f.h.j.redig@tudelft.nl

1 Introduction

Concentration inequalities are a well studied subject in probability and statistics and are very useful in the study of fluctuations of possibly complicated and indirectly defined functions of random variables, such as the Kantorovich distance between the empirical distribution and the true distribution, and various properties of random graphs. See, e.g., [2, 10] and references therein. Initially mostly studied in the i.i.d. context, many efforts have been done to extend concentration inequalities to the context of dependent random variables, and more generally dependent random fields. E.g. in the context of models of statistical mechanics, where the dependence is naturally encoded in the interaction potential, the relation between the Dobrushin uniqueness condition (high-temperature) and the Gaussian concentration inequality has been obtained in [9, 4, 3], whereas at low temperature weaker concentration inequalities are proved in [4].

In this paper we are interested in the behaviour of concentration inequalities under stochastic time-evolution. There are several motivations to be interested in this problem. First, in the context of non-equilibrium systems, non-equilibrium stationary states, or transient non-equilibrium states are usually characterized rather implicitly via an underlying dynamics. If we are interested in concentration properties of such measures, we are naturally lead to the question of time-evolution of measures satisfying a concentration inequality. It is also used in various contexts that a Markovian semigroup interpolates between different measures [1], [10, Section 2.3], and therefore it is of interest whether this interpolation conserves concentration properties. Notice that in the context of Gibbs measures, stochastic time-evolution (even high-temperature dynamics) can destroy the Gibbs property [6], therefore it is interesting to understand whether such measures -though not Gibbs- still enjoy concentration properties, or whether there can be phase transitions in the concentration behavior of a measure, e.g., from Gaussian concentration bound to weaker concentration bound in a dynamics leading from high to low-temperature regime.

As we will see later, in the study of these questions, an object popping up naturally is the so-called nonlinear semigroup $V_t f = \log S_t \,\mathrm{e}^f$ where S_t is the Markov semigroup of the process under consideration, and its associated nonlinear generator $\mathcal{H}f = \mathrm{e}^{-f}\,\mathcal{L}(\mathrm{e}^f)$ where $\mathcal{L}$ is the Markov generator.

In this paper, for the stochastic dynamics, we restrict to diffusion processes. In this setting, the nonlinear generator is a sum of a linear and a quadratic part, and this quadratic part coincides with the "carré du champ" operator. This implies that one can use general results on strong gradient bounds from [1]. Our paper is organized as follows. In section 2 we define the basic setting and define the problem of time-evolution of the Gaussian concentration bound. We also give a simple but enlightening example of the Ornstein-Uhlenbeck process, where starting from a normal distribution, we can explicitly see the time-evolution of the constant in the Gaussian concentration bound. In section 3 we use the method of the non-linear semigroup, which as we see in section 3.2, enters naturally in our context. The main problem is then to understand the evolution of the Lipschitz constant under the non-linear semigroup. In section 3, we control this via the method and framework of [1], using the strong gradient bound. This method applies in the reversible context. In section 4, we use a different approach based on coupling which can also be used in the non-reversible context. We give examples from non-equilibrium steady states, and non-gradient perturbations of reversible diffusions. In section 4, we use a third approach based on the exponential moment of the square distance function. With this technique, we give a class of examples where, starting from any initial measure, we have the Gaussian concentration bound at any positive time, and we also apply the technique for a time-dependent Markovian diffusion with confining drift condition. This applies for instance to the "noisy" Lorenz system. Finally, in section 6 we treat non-Markovian diffusions with linear drift, which can be studied using martingale moment inequalities. In the appendices we give a new proof of Gaussian concentration from the existence of an exponential moment of the square distance function, and provide a general approximation lemma, showing that in the context of a separable Banach space, the Gaussian concentration bound for smooth functions with bounded support implies the Gaussian concentration bound for general Lipschitz functions.

2 Setting and basic questions

2.1 Gaussian concentration bounds

We denote by $\mathscr{C}_b(\mathbb{R}^d, \mathbb{R})$ the space of bounded continuous functions from $\mathbb{R}^d$ to $\mathbb{R}$. For a probability measure μ on (the Borel σ-field of) $\mathbb{R}^d$ and $f \in \mathscr{C}_b(\mathbb{R}^d, \mathbb{R})$, we denote by $\mu(f) = \int f \, \mathrm{d}\mu$ the μ expectation of f. $\mathrm{Lip}(\mathbb{R}^d, \mathbb{R})$ denotes the set of real-valued Lipschitz functions. We further denote for $f \in \mathrm{Lip}(\mathbb{R}^d, \mathbb{R})$

$$\mathrm{lip}(f) := \sup_{x,y} \frac{|f(x) - f(y)|}{\|x - y\|}$$

the Lipschitz constant of f, where $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^d$. A Lipschitz function is almost surely differentiable by Rademacher's theorem [11, p. 101], and the supremum norm of the gradient coincides with the Lipschitz constant. For $f : \mathbb{R}^d \to \mathbb{R}$ we denote by ∇f the gradient of f, which we view as a column vector. We denote

$$\|\nabla f\|_\infty^2 := \operatorname*{ess\,sup}_{x \in \mathbb{R}^d} \|\nabla f(x)\|^2 .$$

We can now define the notion of Gaussian concentration bound.

Definition 2.1. *Let μ be a probability measure on (the Borel σ-field of) $\mathbb{R}^d$.*

a) We say that μ satisfies the smooth Gaussian concentration bound with constant D if we have

$$\log \mu\left(\mathrm{e}^{f - \mu(f)}\right) \le D \, \mathrm{lip}(f)^2 .$$

for all smooth compactly supported f. We abbreviate this property by GCBS(D).

b) We say that μ satisfies the Gaussian concentration bound with constant D if we have

$$\log \mu\left(\mathrm{e}^{f-\mu(f)}\right) \le D\,\mathrm{lip}(f)^2 .$$

for all Lipschitz functions $f \in \mathrm{Lip}(\mathbb{R}^d, \mathbb{R})$. We abbreviate this property by $\mathrm{GCB}(D)$.

In appendix B we prove in a much more general setting, i.e., in the context of a separable Banach space, that $\mathrm{GCBS}(\cdot)$ and $\mathrm{GCB}(\cdot)$ are equivalent. More precisely we prove that $\mathrm{GCBS}(D)$ implies $\mathrm{GCB}(D)$ (in general, we have to replace compact support by bounded support). Therefore, for the rest of the paper, we concentrate on the time evolution of $\mathrm{GCBS}(\cdot)$ rather than $\mathrm{GCB}(\cdot)$.

2.2 Time evolved Gaussian concentration bound

Let $\{X_t, t \ge 0\}$ denote a Markov diffusion process on $\mathbb{R}^d$, i.e., a process solving a SDE of the form

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sqrt{2a(X_t)}\,\mathrm{d}W_t \tag{1}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$, and $a : \mathbb{R}^d \to M_d^+$ where M_d^+ denotes the set of $d \times d$ symmetric positive definite matrices, and where $\{W_t, t \ge 0\}$ is standard Brownian motion on $\mathbb{R}^d$. The questions which we study in this paper are the following.

1. If μ satisfies $\mathrm{GCBS}(D)$ then does the same hold for μ_t, the distribution at time t of the process $\{X_t, t \ge 0\}$ when started initially from X_0 distributed according to μ.

2. Does the stationary measure (or stationary measures) of $\{X_t : t \ge 0\}$ satisfy $\mathrm{GCBS}(D_t)$ for some constant D_t? Can we estimate D_t?

2.3 Ornstein-Uhlenbeck process

A simple but inspiring example is given by the one-dimensional Ornstein-Uhlenbeck process, i.e., the process $\{X_t, t \ge 0\}$ solving the SDE

$$\mathrm{d}X_t = -\kappa X_t\,\mathrm{d}t + \sigma\,\mathrm{d}W_t \tag{2}$$

where $\sigma > 0$, and $\{W_t, t \ge 0\}$ is a standard Brownian motion. Let us denote X_t^x the solution starting from $X_0 = x$. Then we have

$$X_t^x = \mathrm{e}^{-\kappa t}\, x + \sigma \int_0^t \mathrm{e}^{-\kappa(t-s)}\,\mathrm{d}W_s\,.$$

If we start from X_0 normally distributed with expectation zero and variance θ^2 (notation $\mathcal{N}(0,\theta^2)$) then, at time $t > 0$, X_t is normally distributed with expectation zero and variance

$$\sigma_t^2 = \theta^2 \,\mathrm{e}^{-2\kappa t} + \frac{\sigma^2}{2\kappa}\big(1 - \mathrm{e}^{-2\kappa t}\big)\,.$$

Because the normal distribution $\mathcal{N}(0,a^2)$ satisfies GCBS(D) with $D = D_0 = \frac{1}{2}a^2$ we conclude that for this example, with $\mu = \mathcal{N}(0,\theta^2)$, μ_t satisfies GCBS(D_t) with

$$D_t = D_\infty + (D_0 - D_\infty)\,\mathrm{e}^{-2\kappa t}$$

with $D_\infty = \frac{\sigma^2}{2\kappa}$. Hence, μ_t satisfies GCBS(D_t) with a constant D_t interpolating smoothly between the initial constant D_0 and the constant D_∞ associated to the stationary normal distribution.

In case $\kappa = 0$ the process is σB_t, and we find

$$\sigma_t^2 = \theta^2 + \sigma^2 t$$

which implies that the constant of the Gaussian concentration bound evolves as

$$D_t = D_0 + \sigma^2 t.$$

3 Nonlinear semigroup approach

In this section we develop an abstract approach based on the so-called non-linear semigroup, combined with the Bakry-Emery Γ_2 criterion. We show that if the strong gradient bound is satisfied, then the Gaussian concentration bound is conserved in the course of the time-evolution, and in the limit $t \to \infty$.

3.1 The nonlinear semigroup

Let $\{X_t : t \geq 0\}$ be a Markov diffusion process on $\mathbb{R}^d$ as defined in (1) and denote by S_t its semigroup acting on $\mathscr{C}_b(\mathbb{R}^d, \mathbb{R})$. As usual, the generator is denoted by

$$\mathcal{L}f(x) = \lim_{t\to 0^+} \frac{S_t f(x) - f(x)}{t}\,.$$

on its domain $\mathcal{D}(\mathcal{L})$ of functions f such that $\frac{S_t f(x) - f(x)}{t}$ converges uniformly in x when $t \to 0^+$. The *non-linear* semigroup is denoted by

$$V_t(f) = \log S_t(\mathrm{e}^f)\,.$$

This is indeed a semigroup:

$$V_{t+s}f = \log S_{t+s}\, \mathrm{e}^f = \log(S_t(S_s(e^f)) = \log S_t(\log \mathrm{e}^{V_s f}) = V_t(V_s f)\,.$$

We denote by $\mathcal{H}$ its generator, i.e.,

$$\mathcal{H}f(x) = \lim_{t\to 0} \frac{V_t f(x) - f(x)}{t} = \left(\mathrm{e}^{-f}\, \mathcal{L}(\mathrm{e}^f)\right)(x) \tag{3}$$

defined on the domain $\mathcal{D}(\mathcal{H})$ where the defining limit in (3) converges uniformly. The relation between $\mathcal{H}$ and V_t is more subtle than the relation between $\mathcal{L}$ and S_t, i.e., we assume that we are in a context where there is a core where the right-hand side is well-defined and the closure of $\mathcal{H}$ defined on this core generates V_t. In the context of diffusion processes, e.g., this core consists of asymptotically constant smooth functions. In general, we call elements of this core "smooth functions". In general, whether $\mathcal{H}$ indeed generates the semigroup V_t is not obvious: to write down the expression $\left(\mathrm{e}^{-f}\, \mathcal{L}(\mathrm{e}^f)\right)$ one already needs some extra condition on the domain $\mathcal{D}(\mathcal{L})$. In all the cases below we will be in such a context where the domain of the contains a core that is closed under the map $f \mapsto \mathrm{e}^{af}$ for any $a \in \mathbb{R}$, and $\mathcal{H}$ is effectively the generator of V_t. We refer the reader to [12] for more details. In particular we have, for all f smooth (in the sense clarified above)

$$\frac{\mathrm{d}V_t f}{\mathrm{d}t} = \mathcal{H}V_t f\,.$$

Notice that, unlike in the case of the linear semigroup S_t, we *do not have commutation* of the semigroup with the generator, i.e., in general $\mathcal{H}V_t f \neq V_t\mathcal{H}f$.

3.2 Some preparatory computations

In order to start answering the questions of Section 2.2 we show here how the non-linear semigroup enters naturally into these questions. Indeed, for all $t \geq 0$, we have

$$\begin{aligned}\mu_t\left(\mathrm{e}^{f-\mu_t(f)}\right) &= \mu\left(S_t(\mathrm{e}^f)\right)\mathrm{e}^{-\mu(S_t f)}\\ &= \mu\left(\mathrm{e}^{V_t f - \mu(V_t f)}\right)\mathrm{e}^{\mu(V_t f - S_t f)}\,.\end{aligned} \tag{4}$$

Therefore, if μ satisfies GCBS(D), then we can estimate the first factor in the r.h.s. of (4)

$$\mu\left(\mathrm{e}^{V_t f - \mu(V_t f)}\right) \leq \mathrm{e}^{D\,\mathrm{lip}(V_t f)^2} \tag{5}$$

and so we have to estimate $\mathrm{lip}(V_t f)$, which in the case of diffusion processes will boil down to estimating $\nabla V_t f$. Concerning the second factor in (4) we define first the "truly non-linear" part of the non-linear generator as follows

$$\mathcal{H}_{\mathrm{nl}}(f) = \mathcal{H}(f) - \mathcal{L}(f)\,.$$

In the case of diffusion processes, this operator exactly contains the quadratic term of $\mathcal{H}$, which coincides in turn with the carré du champ operator (see section 3.3 below). We can then proceed as follows

$$\begin{aligned}\frac{\mathrm{d}(V_t f - S_t f)}{\mathrm{d}t} &= \mathcal{H}V_t f - \mathcal{L}S_t f = \mathcal{H}V_t f - \mathcal{L}V_t f + \mathcal{L}(V_t f - S_t f)\\ &= \mathcal{H}_{\mathrm{nl}}(V_t f) + \mathcal{L}(V_t f - S_t f)\,.\end{aligned}$$

As a consequence, we obtain, by the variation of constant method,

$$V_t f - S_t f = \int_0^t S_{t-s}\mathcal{H}_{\mathrm{nl}}(V_s f)\,\mathrm{d}s$$

and because S_t is a Markov semigroup, it is a contraction semigroup in the supremum norm and because μ is a probability measure, we obtain the inequality

$$\mu(|V_t f - S_t f|) \le \|V_t f - S_t f\|_\infty \le \int_0^t \|\mathcal{H}_{\mathrm{nl}}(V_s f)\|_\infty\,\mathrm{d}s\,. \tag{6}$$

As a consequence of (5) and (6), we first aim at obtaining estimates for $\mathrm{lip}(V_t f)$, or $\nabla V_t f$, and next use these estimates to further estimate the integral in the r.h.s. of (6). In particular, in the case of diffusion processes, $\mathcal{H}_{\mathrm{nl}} g$ is of the form $(\nabla g)^2$, and hence if we have a uniform estimate for $\nabla(V_t f)$, we can plug it in immediately. Summarizing, assuming that μ satisfies GCBS(D), when we combine (4), (5) and (6), we obtain, for all $t \ge 0$,

$$\mu_t\left(\mathrm{e}^{f-\mu_t(f)}\right) \le \exp\left(D\,\mathrm{lip}(V_t f)^2 + \int_0^t \|\mathcal{H}_{\mathrm{nl}}(V_s f)\|_\infty\,\mathrm{d}s\right)\,. \tag{7}$$

3.3 Abstract gradient bound approach

In this subsection we study the questions formulated in Section 2.2 in the context of Markovian diffusion triples, in the sense of [1], i.e., reversible diffusion processes for which we have the integration by parts formula relating the Dirichlet form and the carré du champ bilinear form. Let $\{X_t, t \ge 0\}$ be a Markov diffusion, i.e., a solution of the SDE of the form (1). Moreover, we will assume in this subsection that the covariance matrix $a(x)$ is not degenerate and bounded, uniformly in $x, v \in \mathbb{R}^d$, i.e., for some $C_1, C_2 > 0$,

$$C_1^{-2}\|v\|^2 \leq \langle v, a(x)v\rangle \leq C_2^2\|v\|^2 \tag{8}$$

where $\langle\cdot,\cdot\rangle$ denotes Euclidean inner product.

The generator of the process $\{X_t, t \geq 0\}$ solving the SDE (1), acting on a smooth compactly supported functions $f : \mathbb{R}^d \to \mathbb{R}$ is then given by

$$\mathcal{L}f(x) = \sum_{i=1} b_i(x)\,\partial_i f(x) + \sum_{i,j} a_{ij}(x)\,\partial_i\partial_j f(x) \tag{9}$$

where ∂_i denotes partial derivative w.r.t. x_i.

To the generator $\mathcal{L}$ is associated the carré du champ bilinear form

$$\Gamma(f, g) = \frac{1}{2}\left(\mathcal{L}(fg) - g\mathcal{L}(f) - f\mathcal{L}(g)\right) = \langle \nabla f, a \cdot \nabla g\rangle\,.$$

Notice that Γ satisfies the so-called diffusive condition, i.e., for all smooth functions $\psi : \mathbb{R} \to \mathbb{R}$

$$\Gamma(\psi(f), \psi(g))(x) = (\psi')^2(x)\,\Gamma(f, g)(x)\,.$$

We will further assume that there exists a reversible measure ν such that the integration by parts formula

$$\int f(-\mathcal{L}g)\,\mathrm{d}\nu = \int \Gamma(f, g)\,\mathrm{d}\nu$$

holds. The triple $(\mathbb{R}^d, \Gamma, \nu)$ is then a Markov diffusion triple in the sense of [1, section 3.1.7].

The second order carré du champ bilinear form is given by

$$\Gamma_2(f, g) = \frac{1}{2}\left(\mathcal{L}\Gamma(f, f) - \Gamma(\mathcal{L}f, g) - \Gamma(f, \mathcal{L}g)\right)\,.$$

In what follows, we abbreviate, as usual, $\Gamma(f, f) =: \Gamma(f)$, $\Gamma_2(f, f) = \Gamma_2(f)$. An important example is when $b = -\nabla W$ and $a = I$, in which case the second order the carré du champ bilinear form is given by

$$\Gamma_2(f, f) = \|\nabla\nabla f\|^2 + \langle \nabla f, \nabla\nabla W(\nabla f)\rangle$$

where $\nabla\nabla W$ denotes the Hessian of W, i.e., the matrix of the second derivatives. By the non-degeneracy and boundedness condition (8), we have, for all $x \in \mathbb{R}^d$

$$C_1^{-2}\|\nabla f(x)\|^2 \leq \Gamma(f)(x) \leq C_2^2\|\nabla f(x)\|^2\,.$$

Following [1] we say that the strong gradient bound is satisfied with constant $\rho \in \mathbb{R}$ if for all $t \in \mathbb{R}_+$

$$\sqrt{\Gamma(S_t f)} \leq \mathrm{e}^{-\rho t}\, S_t\big(\sqrt{\Gamma(f)}\,\big)\,. \tag{10}$$

This condition is fulfilled when, e.g., the Bakry-Emery curvature bound,

$$\Gamma_2(f) \geq \rho\,\Gamma(f)$$

is satisfied. We refer to [1, Chapter 3] for the proof and more background on this formalism. We then have the following general result.

Theorem 3.1. *Let $\{X_t, t \geq 0\}$ be a reversible diffusion process such that (10) is fulfilled. Assume that μ satisfies* GCBS(D). *Then, for every $t \geq 0$, μ_t satisfies* GCBS(D_t) *with*

$$D_t = D C_1^2 C_2^2\, \mathrm{e}^{-2\rho t} + \frac{C_1^2 C_2^4}{2\rho}\big(1 - \mathrm{e}^{-2\rho t}\big)\,. \tag{11}$$

In particular, if $\rho > 0$, then the unique reversible measure ν satisfies GCBS(D_∞) *with $D_\infty = \frac{C_2^4 C_1^2}{2\rho}$.*

Proof. Using (10) we start by estimating $\|\nabla V_t f\|$ for $f : \mathbb{R}^d \to \mathbb{R}$ smooth

$$\begin{aligned}
\|\nabla V_t f\| &= \frac{\|\nabla(S_t(\mathrm{e}^f))\|}{S_t(\mathrm{e}^f)} \leq C_1 \frac{\sqrt{\Gamma(S_t(\mathrm{e}^f))}}{S_t(\mathrm{e}^f)} \\
&\leq C_1\, \mathrm{e}^{-\rho t}\, \frac{S_t\big(\sqrt{(\Gamma(\mathrm{e}^f)}\,\big)}{S_t(\mathrm{e}^f)} = C_1\, \mathrm{e}^{-\rho t}\, \frac{S_t\big(\,\mathrm{e}^f \sqrt{\Gamma(f)}\,\big)}{S_t(\mathrm{e}^f)} \\
&\leq C_1\, \mathrm{e}^{-\rho t}\, \|\sqrt{\Gamma(f)}\|_\infty \leq C_1 C_2\, \mathrm{e}^{-\rho t}\, \|\nabla f\|_\infty\,.
\end{aligned}$$

As a consequence we obtain the estimate

$$\mathrm{lip}(V_t f) = \|\nabla V_t f\|_\infty \leq C_1 C_2\, \mathrm{e}^{-\rho t}\, \|\nabla f\|_\infty\,. \tag{12}$$

Now we recall that what we called the "truly non-linear part" of the non-linear generator $\mathcal{H}_{\mathrm{nl}}$ coincides here with the carré du champ bilinear form, i.e.,

$$\mathcal{H}_{\mathrm{nl}} f = \Gamma(f) \leq C_2^2\, \|\nabla f\|_\infty^2\,. \tag{13}$$

As a consequence, starting from (6), we further estimate

$$\|V_t f - S_t f\|_\infty \leq C_2^2 \int_0^t \|\nabla V_s f\|_\infty^2\, \mathrm{d}s \leq C_1^2 C_4^2\, \|\nabla f\|_\infty^2 \int_0^t \mathrm{e}^{-2\rho s}\, \mathrm{d}s\,. \tag{14}$$

Combining (12), (14) with (7) we obtain that μ_t satisfies GCBS(D_t) with

$$D_t = DC_1^2 C_2^2 \,\mathrm{e}^{-2\rho t} + C_1^2 C_2^4 \int_0^t \mathrm{e}^{-2\rho s}\,\mathrm{d}s$$

which is the claim of the theorem. □

Remark 3.1.

a) In case $\Gamma(f) = c^2\|\nabla f\|^2$, we have $C_1 = a^{-2}, C_2 = a^2$, so D_t in $t = 0$ equals D. In general, $C_1^2 C_2^2 > 1$, which means that at time $t = 0$ we do not recover the constant D, but a larger constant. This is an artefact of the method where we estimate the norm of the gradient via the carré du champ.

b) In case we have an exact commutation relation of the type

$$\nabla S_t f = \mathrm{e}^{-\rho t}\, S_t \nabla f$$

such as is the case for the Ornstein-Uhlenbeck process, we obtain directly

$$|\nabla V_t f\| \leq \mathrm{e}^{-\rho t}\, \|\nabla f\|_\infty$$

i.e., without using the bilinear form Γ.

4 Coupling approach

4.1 Coupling and the nonlinear semigroup

In the previous section, the essential input coming from the strong gradient bound is the estimate (12) which implies that for all $x, y \in \mathbb{R}^d$ and all $t \in \mathbb{R}_+$

$$\|V_t f(x) - V_t f(y)\| \leq C_t \, \|\nabla f\|_\infty \, \|x - y\| \, \mathrm{e}^{-\rho t} \; . \tag{15}$$

Once we have the bound (15), we can use it to further estimate the r.h.s. of (6), provided we have a control on $\mathcal{H}_{\mathrm{nl}}$. Instead of starting from the curvature bound, in this subsection we start from a coupling point of view. This has the advantage that reversibility is no longer necessary. We denote by X_t^x the process $\{X_t, t \geq 0\}$ started at $X_0 = x$.

As an important example to keep in mind, consider the Ornstein-Uhlenbeck process on $\mathbb{R}^d$, with generator

$$-\langle Ax, \nabla\rangle + \Delta$$

where Δ denotes the Laplacian in $\mathbb{R}^d$, and where A is a $d\times d$ matrix. In that case we have

$$X_t^x = \mathrm{e}^{-At}\, x + \int_0^t \mathrm{e}^{-2A(t-s)}\, \mathrm{d}W_s \tag{16}$$

which depends deterministically, and in fact linearly, on x.

Definition 4.1. *Let $\gamma : [0,\infty) \to [0,\infty)$ be a measurable function such that $\gamma(0) = 1$. We say that the process $\{X_t, t \geq 0\}$ can be coupled at rate γ if for all $x, y \in \mathbb{R}^d$ there exists a coupling of $\{X_t^x, t \geq 0\}$ and $\{X_t^y, t \geq 0\}$ such that almost surely in this coupling*

$$d(X_t^x, X_t^y) \leq d(x,y)\,\gamma(t)\,. \tag{17}$$

In the case of the Ornstein-Uhlenbeck process in $\mathbb{R}^d$, we have from (16) (which implicitly defines a coupling, because we use (16) for all x with the same Brownian realization)

$$\|X_t^x - X_t^y\| \leq \|\,\mathrm{e}^{-At}\,\|\, \|x-y\|$$

hence $\gamma(t) = \|\,\mathrm{e}^{-At}\,\|$. Notice that $\gamma(t)$ can be "expanding" or "contracting", depending on the spectrum of A. More precisely, γ will be eventually contracting if the numerical range of A lies in the half plane of complex numbers with non-positive real part.

We have the following result. Let $\mathcal{W}_1$ be the space of probability measures μ such that $\int d(0,x)\,\mathrm{d}\mu(x) < \infty$ equipped with the distance

$$\begin{aligned} d_{\mathcal{W}_1}(\mu,\nu) &= \sup\left\{\int f\,\mathrm{d}\mu - \int f\,\mathrm{d}\nu : \mathrm{lip}(f) \leq 1\right\} \\ &= \inf\left\{\int d(x,y)\,\mathrm{d}P : P \text{ coupling of } \mu,\nu\right\}. \end{aligned}$$

Theorem 4.1. *Assume that $\{X_t, t\geq 0\}$ can be coupled at rate γ. Assume that μ satisfies* GCBS(D), *then for all $t > 0$, for f smooth we have the estimate*

$$\log \mu_t(\mathrm{e}^{f-\mu_t(f)}) \leq D\,\mathrm{lip}(f)^2 \gamma(t)^2 + C_2^2\,\mathrm{lip}(f)^2 \int_0^t \gamma(s)^2\,\mathrm{d}s \tag{18}$$

where C_2 is defined in (8). As a consequence, μ_t satisfies GCBS(D_t) *with*

$$D_t = D\gamma(t)^2 + C_2^2 \int_0^t \gamma(s)^2\,\mathrm{d}s\,. \tag{19}$$

In particular, if $\int_0^\infty \gamma(s)^2 \, \mathrm{d}s < \infty$*, then every weak limit point of* $\{\mu_t, t \geq 0\}$ *satisfies* GCBS(D_∞) *with*

$$D_\infty = C_2^2 \int_0^\infty \gamma(s)^2 \, \mathrm{d}s \, .$$

Moreover, the unique invariant probability measure $\nu \in \mathcal{W}_1$ *satisfies* GCBS(D_∞)*.*

Proof. We start with a lemma which gives a general estimate on the variation of $V_t f$.

Lemma 4.1. *Let* f *be Lipschitz and assume that* $\{X_t, t \geq 0\}$ *can be coupled at rate* γ*. Then for all* $t \geq 0$ *and* $x, y \in \mathbb{R}^d$ *we have*

$$V_t f(x) - V_t f(y) \leq \mathrm{lip}(f) \, \gamma(t) \, d(x, y) \, .$$

As a consequence, for all $t \geq 0$*,*

$$\mathrm{lip}(V_t f) \leq \mathrm{lip}(f) \, \gamma(t) \, .$$

Proof. Let us denote by $\widehat{\mathbb{E}}$ expectation in the coupling of $\{X_t^x, t \geq 0\}$ and $\{X_t^y, t \geq 0\}$ for which (17) holds (which exists by assumption). Then we have

$$\begin{aligned}
\exp(V_t f(x) - V_t f(y)) &= \frac{\widehat{\mathbb{E}}\left(\mathrm{e}^{f(X_t^x)}\right)}{\widehat{\mathbb{E}}\left(\mathrm{e}^{f(X_t^y)}\right)} = \frac{\widehat{\mathbb{E}}\left(\mathrm{e}^{f(X_t^y)}(\mathrm{e}^{f(X_t^x) - f(X_t^y)})\right)}{\widehat{\mathbb{E}}\left(\mathrm{e}^{f(X_t^y)}\right)} \\
&\leq \frac{\widehat{\mathbb{E}}\left(\mathrm{e}^{f(X_t^y)} \, \mathrm{e}^{\mathrm{lip}(f) \, d(X_t^x, X_t^y)}\right)}{\widehat{\mathbb{E}}\left(\mathrm{e}^{f(X_t^y)}\right)} \\
&\leq \frac{\widehat{\mathbb{E}}\left(\mathrm{e}^{f(X_t^y)} \, \mathrm{e}^{\mathrm{lip}(f) \, d(x,y) \gamma(t)}\right)}{\widehat{\mathbb{E}}\left(\mathrm{e}^{f(X_t^y)}\right)} = \mathrm{e}^{\mathrm{lip}(f) d(x,y) \gamma(t)}
\end{aligned}$$

where in the last inequality we used (17). □

Notice that in lemma 4.1 it is not required that $\gamma(t) \to 0$ as $t \to \infty$, i.e., the coupling does not have to be successful. However if one wants to pass to the limit $t \to \infty$ then it is important that $\gamma(t) \to 0$ as $t \to \infty$. This in turn implies, as we see in the next lemma that among all probability measures in the Wasserstein space $\mathcal{W}_1$, there is a unique invariant probability measure ν, and for all $\mu \in \mathcal{W}_1$, $\mu_t \to \nu$ weakly as $t \to \infty$.

Lemma 4.2. *Assume that* $\{X_t, t \geq 0\}$ *can be coupled at rate* γ *and* $\gamma(t) \to 0$ *as* $t \to \infty$*. Then there exists a unique invariant probability measure* ν *in* $\mathcal{W}_1$*. Moreover, for all* $\mu \in \mathcal{W}_1$*,* $\mu_t \to \nu$ *as* $t \to \infty$*.*

Proof. Let μ, ν be elements of $\mathcal{W}_1$ and let f be a Lipschitz function with $\mathrm{lip}(f) \leq 1$. Because μ, ν are elements of $\mathcal{W}_1$, there exists a coupling P such that

$$\int d(x,y)\,\mathrm{d}P = d_{\mathcal{W}_1}(\mu,\nu) < \infty\,.$$

Then

$$\begin{aligned}\int f\,\mathrm{d}\mu_t - \int f\,\mathrm{d}\nu_t &= \int \widehat{\mathbb{E}}(f(X_t^x) - f(X_t^y))\,\mathrm{d}P(x,y) \\ &\leq \int \widehat{\mathbb{E}}(d(X_t^x, X_t^y))\,\mathrm{d}P(x,y) \\ &\leq \gamma(t)\,d_{\mathcal{W}_1}(\mu,\nu)\,.\end{aligned}$$

This shows that for all $\mu, \nu \in \mathcal{W}_1$, and for all $t \geq 0$,

$$d_{\mathcal{W}_1}(\mu_t, \nu_t) \leq \gamma(t)\,d_{\mathcal{W}_1}(\mu,\nu)\,. \tag{20}$$

Existence of an invariant measure $\nu \in \mathcal{W}_1$ now follows via a standard contraction argument. If $\mu, \nu \in \mathcal{W}_1$ are both invariant then (20) gives, after taking $t \to \infty$: $d_{\mathcal{W}_1}(\mu,\nu) = 0$, which shows uniqueness of the invariant measure $\nu \in \mathcal{W}_1$. The fact $\mu \in \mathcal{W}_1$, $\mu_t \to \nu$ as $t \to \infty$ then also follows from (20). □

To finish the proof of the theorem, we use (13)

$$\mathcal{H}_{\mathrm{nl}} f = \Gamma(f) \leq C_2^2\,\|\nabla f\|^2 \leq C_2^2\,\mathrm{lip}(f)^2\,.$$

Combining with (6) and (5) and lemma 4.2 this yields the result of the theorem.

□

As an application we have the following result on Markovian diffusions with covariance matrix a not depending on the location x.

Theorem 4.2. *Let X_t denote a diffusion process on $\mathbb{R}^d$ with generator of type (9), and where the covariance matrix a does not depend on location x, and is such that there exists $C_2 > 0$*

$$\langle u, au\rangle \leq C_2^2 \|u\|^2$$

for all $u \in \mathbb{R}^d$. Assume furthermore that the function $b : \mathbb{R}^d \to \mathbb{R}^d$ is continuously differentiable and the differential $D_x b$ satisfies the estimate

$$\langle D_x b(x)(u), u\rangle \leq -\kappa\,\|u\|^2 \tag{21}$$

for all $x, u \in \mathbb{R}^d$ and some $\kappa \in \mathbb{R}$. Let μ satisfy GCBS(D)*, then, for all $t > 0$, μ_t satisfies* GCBS(D_t) *with*

$$D_t = D\,\mathrm{e}^{-2\kappa t} + \frac{C_2^2}{2\kappa}\,(1 - \mathrm{e}^{-2\kappa t})\,. \tag{22}$$

Moreover, if $\kappa > 0$, then $\mu_t \to \nu$ as $t \to \infty$ where ν is the unique invariant probability measure, which satisfies $GCB(C_2^2/2\kappa)$. In particular, if $b = -\nabla W$, where the potential $W : \mathbb{R}^d \to \mathbb{R}$ is $\mathscr{C}^2$, then (21) reduces to the convexity condition

$$\langle \nabla\nabla W, u, u\rangle \geq \kappa \|u\|^2\,.$$

Proof. We have $\|\mathcal{H}_{\mathrm{nl}} f\| = \Gamma(f) \leq C_2^2 (\nabla f)^2$. Therefore by Theorem 4.1 it suffices to see that we have a coupling rate $\gamma(t) = \mathrm{e}^{-\kappa t}$. We couple X_t^x, X_t^y by using the same realization of the underlying Brownian motion $\{W_t, t \geq 0\}$, and as a consequence, because a does not depend on x, the difference $X_t^x - X_t^y$ is evolving according to

$$\frac{\mathrm{d}(X_t^x - X_t^y)}{\mathrm{d}t} = b(X_t^x) - b(X_t^y)\,.$$

By the mean-value theorem $b(X_t^x) - b(X_t^y) = D_x b(\xi)(X_t^x - X_t^y)$ for some $\xi \in \mathbb{R}^d$. As a consequence,

$$\frac{\mathrm{d}(\|X_t^x - X_t^y\|^2)}{\mathrm{d}t} = 2\langle X_t^x - X_t^y, D_x b(\xi)(X_t^x - X_t^y)\rangle \leq -2\kappa\,\|X_t^x - X_t^y\|^2$$

which gives

$$\|X_t^x - X_t^y\| \leq \mathrm{e}^{-\kappa t}\,\|x - y\|$$

for all $x, y \in \mathbb{R}^d$ and for all $t \in \mathbb{R}_+$. □

Remark 4.1.

a) Notice that in the approach based on the strong gradient bound, we needed non-degeneracy of the covariance matrix a in (1), cf. condition (8). In the coupling setting, we allow the matrix a to be degenerate, but not depending on x, and the condition is only on the drift b.

b) Unlike the time dependent constant D_t, given via the strong gradient bound (11), the bound (22) yields the correct constant D at time zero. Remark that the constant of the limiting stationary distribution, i.e., $C_2^2/2\kappa$ is invariant under linear rescaling of time, as it should. More precisely, if we multiply the generator with a factor α, C_2^2 is multiplied by this same factor α, and so is the constant κ.

c) With the same proof, we can cover the case where we have the condition

$$\langle b(x) - b(y), x - y\rangle \le -\kappa \|x-y\|^2$$

for all x, y.

4.2 Examples

Example 1: Ornstein-Uhlenbeck process and Brownian motion. Coming back to the simple example of the Ornstein-Uhlenbeck process (2), we have coupling rate

$$\gamma(t) = \mathrm{e}^{-\kappa t}$$

and we find (18), i.e., the time evolution of the constant in the Gaussian concentration bound is the same in general as for the special case of a Gaussian starting measure. If we have standard Brownian motion, then the coupling rate $\gamma(t) = 1$ and the formula (19) reads ($C_2 = 1$)

$$D_t = D + t$$

which is sharp if the starting measure is the normal law $\mu_0 = \mathcal{N}(0, \sigma^2)$, which at time t gives $\mu_t = \mathcal{N}(0, \sigma^2 + t)$.

Example 2: Ginzburg-Landau dynamics with boundary reservoirs. We consider the system process $\{X_t, t \ge 0\}$ on $\mathbb{R}^N$ with generator

$$\mathcal{L} = \sum_{i=1}^{N} (\partial_i - \partial_{i+1})^2 - (\varphi'(x_{i+1}) - \varphi'(x_i))(\partial_{i+1} - \partial_i) + L_1 + L_N$$

where ∂_i denotes partial derivative w.r.t. x_i, and where the extra operators L_1 and L_N model the reservoirs and are given by

$$L_1 = b_1(x_1)\,\partial_1 + \frac{1}{2}\sigma_1^2 \partial_1^2$$

$$L_N = b_N(x_N)\,\partial_N + \frac{1}{2}\sigma_N^2 \partial_N^2\,.$$

This models a non-equilibrium system with harmonic potential in the bulk, and driven by reservoirs with drift b_1, b_N. For the choice $b_1(x) = -\kappa_1 x, b_N(x) = -\kappa_N x$ this corresponds to a "non-equilibrium" Ornstein-Uhlenbeck process, for which it can be shown that the unique stationary measure μ is a Gaussian product measure, with an energy profile

$\mu(x_i^2) = \alpha + \beta i$ linearly interpolating between the left and right reservoirs.

The noise in the system is degenerate, but does not depend on x, which means that the coupling condition is satisfied. The covariance matrix a of (1) is given by $a_{ii} = -2, 2 \leq i \leq N-1, a_{11} = 1, a_{NN} = 1, a_{i,i+1} = 2, 1 \leq i \leq N-1$.

If the drifts associated to the reservoirs b_1, b_N are not linear, then the stationary non-equilibrium state is unknown and not Gaussian. In the following, direct application of Theorem 4.2 then gives the following.

Proposition 4.1. *If the reservoir drifts satisfy*

$$\langle u, -\Delta u\rangle - u_1^2 b_1'(x_1) - u_N^2 b_N'(x_N) \leq -\kappa_N \|u\|^2$$

with $-\Delta$ the discrete laplacian defined via $(\Delta u)_i = u_{i+1} + u_{i-1} - 2u_i$ for $2 < i < N-1$, and $(\Delta u)_1 = u_2 - u_1, (\Delta u)_N = u_{N-1} - u_N$, then the unique stationary measure of the process with generator $\mathcal{L}$ satisfies GCBS(D)*, with $D = C_N^2/2\kappa_N$, with $C_N = \|a\| \leq 4$.*

Example 3: Perturbation of the drift. Remark that if (21) is satisfied with $\kappa > 0$ for the drift b with constant κ and $\tilde{b}$ is such that $\langle D_x(\tilde{b} - b)(u), u\rangle \leq \epsilon \|u\|^2$, for some $0 < \epsilon < \kappa$, then obviously, (21) is satisfied for the drift $\tilde{b}$ with constant $\tilde{\kappa} = \kappa - \epsilon$. E.g., if $\tilde{b}(x) = -\nabla W(x) + \epsilon(x)$, where $W(x)$ is a strictly convex potential, then if $\|D_x \epsilon\|_\infty$ is sufficiently small, there is a unique invariant probability measure ν which satisfies GCB$(\cdot)$. However, ϵ is allowed to be of non-gradient form, which implies that ν is not known in explicit form. The same applies to systems where one adds sufficiently weak "boundary" reservoirs as long as the noise of these resevoirs does not depend on x.

5 Distance Gaussian moment approach

In this section, we start with a different approach, based on the equivalence between GCBS(D) and the existence of a Gaussian estimate of an exponential moment of the square of the distance (cf. Theorem 5.1 below).

5.1 A general equivalence

In this subsection, we work in a general separable metric space (Ω, d). We first generalize Definition 2.1.

Definition 5.1. *Let μ be a probability measure on (the Borel σ-field of) (Ω, d). We say that μ satisfies a Gaussian concentration bound with constant $D > 0$ on the metric space (Ω, d) if there exists $x_0 \in \Omega$ such that $\int d(x_0, x)\, \mathrm{d}\mu(x) < \infty$ and for all $f \in \mathrm{Lip}(\Omega, \mathbb{R})$, one has*

$$\int \mathrm{e}^{f-\mu(f)}\, \mathrm{d}\mu \leq \mathrm{e}^{D\, \mathrm{lip}(f)^2}\,.$$

For brevity we shall say that μ satisfies $\mathrm{GCB}(D)$ on (Ω, d).

Remark 5.1.

a) *Note that if there exists $x_0 \in \Omega$ such that $\int d(x_0, x)\, \mathrm{d}\mu(x) < \infty$ then, by the triangle inequality, $\int d(x_0, x)\, \mathrm{d}\mu(x) < \infty$ for all $x_0 \in \Omega$, and all Lipschitz functions on (Ω, d) are μ-integrable.*

b) *Note that one can find a topological space and a probability on the Borel sigma-algebra and two distances d_1 and d_2 s.t. μ satisfies GCB on the metric space with d_1 but it does not on the metric space with d_2. For example, take $\mathbb{R}$, μ to be the Gaussian measure, d_1 the Euclidean distance and $d_2(x, y) = \int_x^y (1 + |s|)\, \mathrm{d}s$.*

Theorem 5.1. *Let μ a probability measure on (Ω, d). Then μ satisifies a Gaussian concentration bound if and only it has a Gaussian moment. More precisely, we have the following:*

1. *If μ satisfies $\mathrm{GCB}(D)$, there exists $x_0 \in \Omega$ such that*

$$\int \mathrm{e}^{\frac{d(x_0,x)^2}{16D}}\, \mathrm{d}\mu(x) \leq 3\, \mathrm{e}^{\frac{\mu(d)^2}{D}} \tag{23}$$

where $\mu(d) := \int d(x, x_0)\, \mathrm{d}\mu(x)$.

2. *If there exist $x_0 \in \Omega$, $a > 0$ and $b \geq 1$ such that*

$$\int \mathrm{e}^{ad(x_0,x)^2}\, \mathrm{d}\mu(x) \leq b \tag{24}$$

then μ satisfies $\mathrm{GCB}(D)$. with

$$D = \frac{1}{2a}\left(1 \vee \frac{b^2\, \mathrm{e}}{2\sqrt{\pi}}\right)\,. \tag{25}$$

This result can be found in [7, Theorem 2.3] with less explicit constants. We provide a direct proof of the theorem in appendix A. Notice that, by the triangle inequality, if (23) holds for some x_0 then it holds for any x_0. Idem for (24).

5.2 Example 1: Diffusions coming down from infinity

As a first example of application, we consider diffusions “coming down from infinity” for which we show that from any starting measure, at positive times $t > 0$, GCBS(D) holds.

We consider a diffusion process on $\mathbb{R}^d$ which solves the SDE

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \mathrm{d}W_t\,.$$

We introduce the following condition on the drift.

Condition 5.1. *There exists an open subset $\mathcal{D} \subset \mathbb{R}^d$ (called “domain”) such that there exists a real, non-negative, non-decreasing and $\mathscr{C}^1$ function h and a constant $A > 0$ such that for all $x \in \mathcal{D}$*

$$\frac{\langle x, b(x)\rangle}{\|x\|} \leq A - h(\|x\|)\,. \tag{26}$$

Theorem 5.2. *Under condition H, if additionally we have the integrability condition*

$$\int_0^\infty \frac{\mathrm{d}u}{h(u)} < \infty \tag{27}$$

then there exists $t_* > 0$, a non-negative function $C(t)$ and a constant $\alpha > 0$ such that for all $0 \leq t \leq t_*$

$$\sup_{x\in\mathcal{D}} \mathbb{E}_x\left(\mathrm{e}^{\alpha\,\|X_t\|^2}\ \mathbb{1}_{\{T_\partial > t\}}\right) \leq C(t)$$

where T_∂ denotes the exit time of the domain $\mathcal{D}$.

We deduce the following result showing immediate Gaussian concentration in the course of diffusions coming down from infinity.

Theorem 5.3. *Assume that hypothesis (26) and (27) hold. Let μ be any probability measure on (the Borel field of) $\mathbb{R}^d$. Let t_* be as in Theorem 5.2. Then, for all $t > 0$, the probability measure $(\mu_t)_{t\geq 0}$ defined by*

$$\mu_t(f) = \mathbb{E}_\mu\big(f(X_t)\big|T_\partial > t\big),\ \forall f \in \mathscr{C}_b(\mathbb{R}^d)$$

satisfies GCBS(D_t) *where*

$$D_t = \frac{1}{2\alpha}\left(1 \vee \frac{C^2(t\wedge t_*)}{\mathbb{P}_\mu(T_\partial > t\wedge t_*)^2}\frac{\mathrm{e}}{2\sqrt{\pi}}\right)\,.$$

Proof. For $0 < t \leq t_*$, the result follows from Theorems 5.2 and 5.1. For $t > t_*$ the result follows from the semigroup property of $S_\partial(t)f(x) = \mathbb{E}_x\big(f(X_t)|T_\partial > t\big)$ and the result for $0 < t \leq t_*$. □

Proof of Theorem 5.2. Define

$$u(t,x) = \varphi(t)\,\mathrm{e}^{\alpha\,\|x\|^2}$$

where φ will be chosen later on. We have

$$\begin{aligned}
&\partial_t u(t,x) + \mathscr{L} u(t,x) \\
&= \mathrm{e}^{\alpha\,\|x\|^2}\left(\dot\varphi(t) + \varphi(t)\big[d\,\alpha/2 + 2\,\alpha^2\|x\|^2 + \langle x, b(x)\rangle\big]\right) \\
&\leq \mathrm{e}^{\alpha\,\|x\|^2}\ \left(\dot\varphi(t) + \varphi(t)\big[d\,\alpha + 4\,\alpha^2\|x\|^2 + A\,\|x\| - h(\|x\|)\,\|x\|\big]\right)\ .
\end{aligned}$$

Using integration by parts we get

$$\int_0^z \frac{\mathrm{d}u}{h(u)} = \frac{z}{h(z)} + \int_0^z \frac{u\,h'(u)}{h(u)^2}\,\mathrm{d}u$$

and using that h is non-decreasing we obtain

$$\liminf_{z\to\infty} \frac{h(z)}{z} \geq \frac{1}{\int_0^\infty \frac{\mathrm{d}u}{h(u)}} > 0\ .$$

Therefore, choosing $\alpha > 0$ sufficiently small and $y_* > 0$ sufficiently large, we have for $u \geq y_*$

$$h(u) - 2\alpha^2\,u - A - \frac{d\,\alpha}{2\,u} > \frac{h(u)}{2}\ .$$

We then define a non-increasing function $y(s)$ and the non-decreasing function $\varphi(s)$ via

$$\frac{\dot\varphi(s)}{\varphi(s)} = -\dot y(s)\,y(s) = y(s)\,\frac{h(s)}{2}\ .$$

We impose additionally $y(0) = \infty$ and obtain

$$\int_{y(s)}^\infty \frac{\mathrm{d}u}{h(u)} = \frac{s}{2}\,.$$

This define t_* via

$$\int_{y_*}^\infty \frac{\mathrm{d}u}{h(u)} = \frac{t_*}{2}$$

and

$$\varphi(s) = \mathrm{e}^{-y(s)^2/2}\ .$$

If $A > y_*$ using Ito's formula with T_A the hitting time of the boundary of the ball centered at x with radius A, (where $A > \|x\|$)

$$\mathbb{E}_x\big(u(t\wedge T_\partial \wedge T_A, X_{t\wedge T_\partial\wedge T_A})\big) = \mathbb{E}_x\left(\int_0^{t\wedge T_\partial\wedge T_A} \big(\partial_t u + \mathscr{L}\, u\big)(s, X_s)\,\mathrm{d}s\right).$$

If $X_s \geq y(s)\vee y_*$ we have

$$\big(\partial_t u + \mathscr{L}\, u\big)(s, X_s) \leq \mathrm{e}^{\alpha\,\|X_s\|^2}\left(\dot\varphi + \varphi\,\frac{h(y(s))}{2}\right) = 0$$

and if $X_s < y(s)\vee y_*$ we have

$$\big(\partial_t u + \mathscr{L}\, u\big)(s, X_s) \leq \mathrm{e}^{\alpha y(s)^2}\left(\dot\varphi(s) + \varphi(s)\,C\left(1+y(s)^2\right)\right)$$

for some (computable) constant $C > 0$ independent of s. Therefore

$$\begin{aligned}
&\mathbb{E}_x\left(\int_0^{t\wedge T_\partial\wedge T_A} \big(\partial_t u + \mathscr{L}\, u\big)(s, X_s)\,\mathrm{d}s\right)\\
&\leq \mathbb{E}_x\left(\int_0^{t\wedge T_\partial\wedge T_A} \mathrm{e}^{\alpha y(s)^2}\left(\dot\varphi(s) + \varphi(s)\,C\left(1+y(s)^2\right)\right)\mathrm{d}s\right)\\
&\leq \int_0^t \mathrm{e}^{\alpha y(s)^2}\left(\dot\varphi(s) + \varphi(s)\,C\left(1+y(s)^2\right)\right)\mathrm{d}s\\
&= -\int_0^t \mathrm{e}^{\alpha y(s)^2}\;\dot y(s)\,y(s)\,\mathrm{e}^{-y(s)^2/2}\,\mathrm{d}s + \int_0^t \mathrm{e}^{\alpha y(s)^2}\;\mathrm{e}^{-y(s)^2/2}\,C\left(1+y(s)^2\right)\mathrm{d}s
\end{aligned}$$

and if $\alpha < 1/2$

$$\leq \int_{y(t)}^\infty \mathrm{e}^{\alpha y^2}\;y\;\mathrm{e}^{-y^2/2}\;\mathrm{d}y + \mathcal{O}(1)\int_0^t \mathrm{d}s = \frac{1}{1-2\,\alpha}\;\mathrm{e}^{-(1-2\,\alpha)\,y(t)^2/2} + \mathcal{O}(1)\,t\,.$$

We now observe that since $u \geq 0$

$$\mathbb{E}_x\big(u(t\wedge T_\partial, X_{t\wedge T_\partial})\;\mathbb{1}_{\{T_A > t\wedge T_\partial\}}\big) \leq \mathbb{E}_x\big(u(t\wedge T_\partial\wedge T_A, X_{t\wedge T_\partial\wedge T_A})\big)$$

therefore by the monotone convergence theorem (let A tend to infinity)

$$\mathbb{E}_x\big(u(t\wedge T_\partial, X_{t\wedge T_\partial})\big) \leq \frac{1}{1-2\,\alpha}\;\mathrm{e}^{-(1-2\,\alpha)\,y(t)^2/2} + \mathcal{O}(1)\,t\;.$$

The result follows by observing that

$$\varphi(t)\;\mathbb{E}_x\left(\mathrm{e}^{\alpha\,\|X_t\|^2}\;\mathbb{1}_{\{T_\partial > t\}}\right) \leq \mathbb{E}_x\big(u(t\wedge T_\partial, X_{t\wedge T_\partial})\big)\;.$$

$\square$

5.3 Example 2: Markovian diffusion processes with space-time dependent drift and covariance

In this section, we consider stochastic differential equations on $\mathbb{R}^d$ given by

$$\mathrm{d}X_t = b(X_t, t)\,\mathrm{d}t + \sigma(X_t, t)\,\mathrm{d}W_t$$

where the vector field b and the matrix-valued σ are regular in x, t. We assume that, for any given initial condition x_0, the solution exists, is unique and defined for all times. This generalizes the coupling setting of Theorem 4.2, i.e., we impose a more general confining condition on the drift $b(x, t)$ and allow the covariance matrix $\sigma(x, t)$ to depend on time and location.

Theorem 5.4. *Assume that $\alpha > 0$, $\beta > 0$ and $\theta > 0$ such that, for all $x \in \mathbb{R}^d$ and $t \geq 0$*

$$\langle x\,, b(x, t)\rangle \leq \alpha\,\|x\| - \beta\|x\|^2 \tag{28}$$

and

$$\sigma^t(x, t)\,\sigma(x, t) \leq \theta\,\mathrm{Id}$$

where the second inequality is in the sense of the order on positive definite matrices. Then, for every initial probability measure μ_0 on $\mathbb{R}^d$ satisfying GCBS(D_0)*, the evolved probability measure μ_t satisfies* GCBS(D_t) *for all $t \geq 0$, where D_t is given by the formula (25), with*

$$\begin{aligned} a = a_0 &= \frac{\beta}{2\theta} \wedge \frac{1}{16 D_0} \\ b = b_t &= b_0\,\exp\left(-a_0\left(\theta d + \frac{2\alpha^2}{\beta}\right)t\right) \\ &\quad + 2\,\mathrm{e}^{\frac{4a_0}{\beta}\left(\theta d + \frac{2\alpha^2}{\beta}\right)}\left(1 - \exp\left(-a_0\left(\theta d + \frac{2\alpha^2}{\beta}\right)t\right)\right) \end{aligned}$$

and

$$b_0 = 3\,\mathrm{e}^{\mu_0(d)^2/8D}$$

where $\mu_0(d) = \int \|x\|\,\mathrm{d}\mu(x)$.

Proof. Let $a_0 = \frac{\beta}{2\theta} \wedge \frac{1}{16 D_0}$ and define $u(x) = \mathrm{e}^{a_0\|x\|^2}$. Using the assumptions we get

$$\begin{aligned} \mathcal{L}u(x) &\leq \left(2a_0^2\theta\|x\|^2 + a_0\theta d + 2a_0\alpha\|x\| - 2a_0\beta\|x\|^2\right)u(x) \\ &\leq a_0\left(\theta d + 2\alpha\|x\| - \beta\|x\|^2\right)u(x) \\ &\leq a_0\left(\theta d + \frac{2\alpha^2}{\beta} - \frac{\beta}{2}\|x\|^2\right)u(x)\,. \end{aligned}$$

For any $A > 0$, let $T_A = \inf\{t \geq 0 : \|X_t\| \geq A\}$. Using Dynkin's formula and Theorem 5.1, we get

$$\begin{aligned}
&\mathbb{E}_{\mu_0}\left(\mathrm{e}^{a_0\|X_{t\wedge T_A}\|^2}\right) \\
&\leq b_0 + a_0\,\mathbb{E}_{\mu_0}\left(\int_0^{t\wedge T_A} \mathrm{e}^{a_0\|X_s\|^2}\left(\theta d + \frac{2\alpha^2}{\beta} - \frac{\beta}{2}\|X_s\|^2\right)\mathrm{d}s\right) \qquad (29)
\end{aligned}$$

where, via (23)

$$b_0 = \int \mathrm{e}^{a_0\|x\|^2}\,\mathrm{d}\mu(x) \leq 3\,\mathrm{e}^{\frac{\mu(d)^2}{8D}}$$

where $\mu(d) = \int \|x\|\,\mathrm{d}\mu(x)$. We now estimate the expectation on the right-hand side of (29). Define, for $s > 0$, the event

$$\mathcal{E}_s = \left\{\|X_s\|^2 > \frac{4}{\beta}\theta d + \frac{2\alpha^2}{\beta}\right\}.$$

We have

$$\begin{aligned}
&\mathbb{E}_{\mu_0}\left(\mathrm{e}^{a_0\|X_{t\wedge T_A}\|^2}\right) \\
&\leq b_0 + 2a_0\left(\theta d + \frac{2\alpha^2}{\beta}\right)\mathbb{E}_{\mu_0}\left(\int_0^{t\wedge T_A} \mathrm{e}^{a_0\|X_s\|^2}\,\mathbb{1}_{\mathcal{E}_s^c}\,\mathrm{d}s\right) \\
&\qquad - a_0\left(\theta d + \frac{2\alpha^2}{\beta}\right)\mathbb{E}_{\mu_0}\left(\int_0^{t\wedge T_A} \mathrm{e}^{a_0\|X_s\|^2}\,\mathrm{d}s\right) \\
&\leq b_0 + 2a_0\left(\theta d + \frac{2\alpha^2}{\beta}\right)\mathrm{e}^{\frac{4a_0}{\beta}\left(\theta d + \frac{2D^2}{\beta}\right)}\,t \\
&\qquad - a_0\left(\theta d + \frac{2\alpha^2}{\beta}\right)\mathbb{E}_{\mu_0}\left(\int_0^{t\wedge T_A} \mathrm{e}^{a_0\|X_s\|^2}\,\mathrm{d}s\right).
\end{aligned}$$

By the Monotone Convergence Theorem, letting $A \uparrow \infty$, and Fubini's Theorem, we get

$$\begin{aligned}
&\mathbb{E}_{\mu_0}\left(\mathrm{e}^{a_0\|X_t\|^2}\right) \\
&\leq b_0 + 2a_0\left(\theta d + \frac{2\alpha^2}{\beta}\right)\mathrm{e}^{\frac{4a_0}{\beta}\left(\theta d + \frac{2\alpha^2}{\beta}\right)}\,t \\
&\qquad - a_0\left(\theta d + \frac{2\alpha^2}{\beta}\right)\int_0^t \mathbb{E}_{\mu_0}\left(\mathrm{e}^{a_0\|X_s\|^2}\right)\mathrm{d}s.
\end{aligned}$$

Using Grönwall's lemma, we obtain

$$\begin{aligned}
\mathbb{E}_{\mu_0}\left(\mathrm{e}^{a_0\|X_t\|^2}\right) &\leq b_0\,\exp\left(-a_0\left(\theta d + \frac{2\alpha^2}{\beta}\right)t\right) + \\
&\qquad 2\,\mathrm{e}^{\frac{4a_0}{\beta}\left(\theta d + \frac{2\alpha^2}{\beta}\right)}\left(1 - \exp\left(-a_0\left(\theta d + \frac{2\alpha^2}{\beta}\right)t\right)\right).
\end{aligned}$$

By Theorem 5.1, we deduce that μ_t satisfies GCBS(D_t) with the announced constant D_t. $\square$

As an application, we consider the famous Lorenz system

$$\begin{cases} \frac{\mathrm{d}x}{\mathrm{d}t} &= \sigma(y-x) \\ \frac{\mathrm{d}y}{\mathrm{d}t} &= rx - y - xz \\ \frac{\mathrm{d}z}{\mathrm{d}t} &= xy - bz\,. \end{cases}$$

which, for a certain range of (positive) parameters has a strange attractor [8, Chapter 14].

Adding a noise which satisfies the condition of Theorem 4.2, this leads to a unique invariant probability measure whose properties are largely unknown. However, this measure satisfies GCBS(). This can be proved observing that the Lorenz system translated by the vector $(0,0,-2r)$ satisfies (28) using the squared norm $\|(x,y,z)\|^2 = rx^2 + \sigma y^2 + \sigma z^2$ with

$$\beta = \inf \frac{rx^2 + y^2 + bz^2}{rx^2 + \sigma y^2 + \sigma z^2}$$

where the infimum is taken over x, y, z in such a way that $(x,y,z) \neq (0,0,0)$.

6 Non Markovian diffusions: Martingale moment approach

In this section we consider the simplest context beyond Markov, where we can no longer rely on methods based on generators.

We consider the stochastic differential equation on $\mathbb{R}$ given by

$$\mathrm{d}X_t = -\kappa X_t\,\mathrm{d}t + \sigma_t\,\mathrm{d}W_t \tag{30}$$

where we assume that the process σ_t is uniformly bounded and predictable. An example of this setting is

$$\begin{cases} \mathrm{d}Y_t &= -\theta Y_t + \mathrm{d}W_t \\ \mathrm{d}X_t &= -\kappa X_t + \sigma(Y_t)\,\mathrm{d}W_t\,. \end{cases}$$

Then the couple (X_t, Y_t) is Markov but X_t is not, and satisfies a SDE of the form (30).

Because the process $\{X_t, t \geq 0\}$ is no longer a Markov process (unless σ_t depends only on X_t) we can no longer use techniques based on the generator as we did before for processes of Ornstein-Uhlenbeck type. The main point is that as a consequence, X_t^x equals a *deterministic process* of bounded variation plus a martingale. As a consequence, the Gaussian concentration bound can be obtained from estimating the martingale, which can be done with the help of Burkholder's inequalities.

The assumption (30) allows us to write the solution in the form

$$X_t = X_0\,\mathrm{e}^{-\kappa t} + \int_0^t \mathrm{e}^{-\kappa(t-s)}\,\sigma_s\,\mathrm{d}W_s\,. \tag{31}$$

We have the following result.

Theorem 6.1. *Assume that there exists $M > 0$ such that*

$$\sup_{t\geq 0} \|\sigma_t\|_{L^\infty} \leq M\,.$$

Assume X_0 is distributed according to a probability measure μ satisfying GCBS(D). *Then we have that for all $t > 0$ there exists $C_t > 0$ such that X_t satisfies* GCBS(D_t). *Moreover, if $\kappa > 0$ then all weak limit points of $\{X_t, t \geq 0\}$ satisfy* GCBS(D_∞) *for some $D_\infty > 0$.*

Proof. We use Theorem 5.1, and will prove that there exist $a > 0, b > 0$ such that

$$\mathbb{E}\left(\mathrm{e}^{aX_t^2}\right) \leq b\,.$$

Then we can conclude via Theorem 5.1, that the distribution of X_t satisfies GCBS(C) with $C \leq \frac{1}{2a}(1 \vee \frac{b^2\,\mathrm{e}}{2\sqrt{\pi}})$. We start from (31) from which we derive the inequality

$$X_t^2 \leq 2X_0^2\,\mathrm{e}^{-2\kappa t} + 2\left(\int_0^t \mathrm{e}^{-\kappa(t-s)}\,\sigma_s\,\mathrm{d}W_s\right)^2\,. \tag{32}$$

We start by estimating, for $\gamma > 0$

$$\begin{aligned}
&\mathbb{E}\left(\exp\left(\gamma\left(\int_0^t \mathrm{e}^{-\kappa(t-s)}\,\sigma_s\,\mathrm{d}W_s\right)^2\right)\right)\\
&= \sum_{n=0}^\infty \frac{\gamma^n}{n!}\mathbb{E}\left(\left(\int_0^t \mathrm{e}^{-\kappa(t-s)}\,\sigma_s\,\mathrm{d}W_s\right)^{2n}\right)\,.
\end{aligned}$$

Next use Burkholder's inequality [5] which states that for a martingale $\{Z_t, t \geq 0\}$ w.r.t. Brownian filtration, with quadratic variation $[Z, Z]_t$, we have the estimate

$$\mathbb{E}(Z_t^{2n}) \leq A(2n)^n\mathbb{E}([Z, Z]_t^n)$$

with A an absolute constant. As a consequence,

$$\begin{aligned}\mathbb{E}\left(\left(\int_0^t \mathrm{e}^{-\kappa(t-s)}\,\sigma_s\,\mathrm{d}W_s\right)^{2n}\right) &= \mathrm{e}^{-2n\kappa t}\,\mathbb{E}\left(\left(\int_0^t \mathrm{e}^{\kappa s}\,\sigma_s\,\mathrm{d}W_s\right)^{2n}\right)\\ &\le \mathrm{e}^{-2n\kappa t}\,A(2n)^n\mathbb{E}\left(\left(\int_0^t \mathrm{e}^{2\kappa s}\,\sigma_s^2\,\mathrm{d}s\right)^{n}\right)\\ &\le \mathrm{e}^{-2n\kappa t}\,AM^{2n}(2n)^n\mathbb{E}\left(\left(\int_0^t \mathrm{e}^{2\kappa s}\,\mathrm{d}s\right)^{n}\right)\\ &\le AM^{2n}(2n)^n\left(\frac{1-\mathrm{e}^{-2\kappa t}}{2\kappa}\right)^n .\end{aligned}$$

As a consequence we obtain

$$\begin{aligned}&\mathbb{E}\left(\exp\left(\gamma\left(\int_0^t \mathrm{e}^{-\kappa(t-s)}\,\sigma_s\,\mathrm{d}W_s\right)^2\right)\right)\\ &\le A\sum_{n=0}^{\infty}\frac{1}{n!}\gamma^n M^{2n}(2n)^n\left(\frac{1-\mathrm{e}^{-2\kappa t}}{2\kappa}\right)^n .\end{aligned}$$

The r.h.s. of this inequality is a convergent series provided

$$\gamma < \left(2\,\mathrm{e}\,M^2\left(\frac{1-\mathrm{e}^{-2\kappa t}}{2\kappa}\right)\right)^{-1} .$$

We then estimate, using (32) and the Cauchy-Schwarz inequality

$$\mathbb{E}\left(\mathrm{e}^{aX_t^2}\right) \le \left(\mathbb{E}\left(\mathrm{e}^{4aX_0^2\,\mathrm{e}^{-2\kappa t}}\right)\right)^{1/2}\left(\mathbb{E}\left(\mathrm{e}^{4a\left(\int_0^t \mathrm{e}^{-\kappa(t-s)}\,\sigma_s\,\mathrm{d}W_s\right)^2}\right)\right)^{1/2} . \tag{33}$$

Because by assumption the distribution of X_0 satisfies GCBS(C), we have that the first factor in the r.h.s. in (33) is finite as soon as $4a\,\mathrm{e}^{-2\kappa t} < a_0$ where a_0 is such that $\mathbb{E}\left(\mathrm{e}^{a_0X_0^2}\right) < \infty$. The second factor is finite as soon as

$$a < \left(8\,\mathrm{e}\,M^2\left(\frac{1-\mathrm{e}^{-2\kappa t}}{2\kappa}\right)\right)^{-1} .$$

Therefore, $\mathbb{E}\left(\mathrm{e}^{aX_t^2}\right)$ is finite for

$$a < \left(8\,\mathrm{e}\,M^2\left(\frac{1-\mathrm{e}^{-2\kappa t}}{2\kappa}\right)\right)^{-1} \wedge a_0\,\mathrm{e}^{2\kappa t}$$

which, combined with Theorem 5.1 concludes the proof of the theorem. □

A Proof of Theorem 5.1

Statement 1. Choose $x_0 \in \Omega$ arbitrarily. Since $x \mapsto d(x_0, x)$ is 1-Lipschitz, GCBS(D) implies by the classical Chernoff bound that for all $r \geq 0$ we have

$$\mu\{x \in \Omega : d(x_0, x) > \mu(d) + r\} \leq \mathrm{e}^{-\frac{r^2}{4D}} \tag{34}$$

where

$$\mu(d) := \int d(x_0, x) \, \mathrm{d}\mu(x) \, .$$

We have

$$\begin{aligned} &\int \mathrm{e}^{ad(x_0,x)^2} \, \mathrm{d}\mu(x) \\ &= \int \mathrm{e}^{ad(x_0,x)^2} \, \mathbb{1}_{\{d(x,x_0)<\mu(d)\}} \, \mathrm{d}\mu(x) + \int \mathrm{e}^{ad(x_0,x)^2} \, \mathbb{1}_{\{d(x,x_0)\geq\mu(d)\}} \, \mathrm{d}\mu(x) \\ &\leq \mathrm{e}^{a\mu(d)^2} + \mathrm{e}^{2a\mu(d)^2} \int \mathrm{e}^{2a(d(x_0,x)-\mu(d))^2} \, \mathbb{1}_{\{d(x,x_0)\geq\mu(d)\}} \, \mathrm{d}\mu(x) \, . \end{aligned}$$

Now we use the fact that

$$\begin{aligned} &\int \mathrm{e}^{2a(d(x_0,x)-\mu(d))^2} \, \mathbb{1}_{\{d(x,x_0)\geq\mu(d)\}} \, \mathrm{d}\mu(x) \\ &= 1 + \int_1^\infty \mu\left(\left\{x : \mathrm{e}^{2a(d(x_0,x)-\mu(d))^2} \geq u\right\}\right) \mathrm{d}u \, . \end{aligned}$$

The result follows using (34) with $a = 1/(16D)$.
Statement 2. Since for all x and for all $a > 0$

$$d(x_0, x) \leq \frac{1}{\sqrt{a}} \, \mathrm{e}^{ad(x_0,x)^2}$$

it follows that $x \mapsto d(x_0, x)$ is μ-integrable. We also have that e^f is μ-integrable for any Lipschitz function. Now, using Jensen's inequality and then the triangle inequality, we obtain

$$\begin{aligned} &\int \mathrm{e}^{f-\mu(f)} \, \mathrm{d}\mu \\ &\leq \int\int \mathrm{e}^{f(x)-f(y)} \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y) \leq \left(\int \mathrm{e}^{\mathrm{lip}(f)d(x,x_0)} \, \mathrm{d}\mu(x)\right)^2 . \end{aligned} \tag{35}$$

Combining the elementary inequality

$$\mathrm{lip}(f) \, d(x, x_0) \leq \frac{\mathrm{lip}(f)^2}{4a} + ad(x, x_0)^2$$

with (24), we obtain

$$\int \mathrm{e}^{\mathrm{lip}(f) d(x,x_0)}\, \mathrm{d}\mu(x) \le b\, \mathrm{e}^{\frac{1}{4a}\,\mathrm{lip}(f)^2}\ . \tag{36}$$

We now show that the prefactor of the exponential can be changed to 1. We first establish the following lemma.

Lemma A.1. *Let Z be a symmetric random variable distributed according to a probability measure ν such that there exist $C_1 \ge 1$ and $C_2 > 0$ such that for all $\lambda \in \mathbb{R}$*

$$\mathbb{E}\left(\mathrm{e}^{\lambda Z}\right) \le C_1\, \mathrm{e}^{C_2 \lambda^2}\ .$$

Then for all $\lambda \in \mathbb{R}$ we have

$$\mathbb{E}\left(\mathrm{e}^{\lambda Z}\right) \le \mathrm{e}^{C_2\left(1 \vee \frac{C_1 \mathrm{e}}{2\sqrt{\pi}}\right)\lambda^2}\ .$$

Proof. We have for any $\lambda \in \mathbb{R}$

$$\mathbb{E}(Z^{2q}) = \mathbb{E}\big(Z^{2q}\, \mathrm{e}^{-\lambda Z}\, \mathrm{e}^{\lambda Z}\big) \le C_1\, (2q)^{2q} \lambda^{-2q}\, \mathrm{e}^{-2q}\, \mathrm{e}^{C_2 \lambda^2} \le C_1\, 4^q q^q\, \mathrm{e}^{-q}\, C_2^q$$

where the first inequality follows maximizing $x^{2q}\, \mathrm{e}^{-\lambda x}$ over x, while the second is obtained by minimizing over λ. Using the bound

$$\sqrt{2\pi}\, n^{n+\frac{1}{2}}\, \mathrm{e}^{-n} \le n! \le \mathrm{e}\, n^{n+\frac{1}{2}}\, \mathrm{e}^{-n}$$

which is valid for any $n \ge 1$, we get

$$\frac{C_1\, 4^q q^q\, \mathrm{e}^{-q}\, C_2^q}{(2q)!} \le \frac{\left(1 \vee \frac{C_1\, \mathrm{e}}{2\sqrt{\pi}}\right)^q C_2^q}{q!}\,,\ q \ge 1\,.$$

The result follows. □

We now combine the above lemma for $C_1 = b^2$ and $C_2 = \frac{\mathrm{lip}(f)^2}{2a}$ with (35) and (36). Theorem 5.1 follows.

B An approximation lemma

In this appendix, $(\Omega, \|\cdot\|)$ is a separable Banach space. We denote by $\mathrm{Lip}(\Omega, \mathbb{R})$ the space of real-valued Lipschitz functions on $(\Omega, \|\cdot\|)$, by $\mathrm{Lip}_s(\Omega, \mathbb{R})$ the space of real-valued Lipschitz functions with bounded support, and by $\mathrm{Lip}_b(\Omega, \mathbb{R})$ the space of real-valued bounded Lipschitz functions. If Ω is a Banach space, we denote by $\mathscr{C}^\infty(\Omega, \mathbb{R})$ the space of real-valued infinitely differentiable functions, and by $\mathscr{C}_s^\infty(\Omega,\mathbb{R})$ the space of real-valued infinitely differentiable functions with bounded support.

Let $\mathscr{C}$ be class of real-valued functions on Ω, we say that μ satisfies $\mathrm{GCB}(\mathscr{C};D)$ if there exists $D>0$ such that

$$\log\mu\left(\mathrm{e}^{f-\mu(f)}\right)\leq D\,\mathrm{lip}(f)^2\,.$$

for all $f\in\mathscr{C}$.

Lemma B.1. *Let μ be a probability measure on Ω. Then*

1. *If μ satisfies $\mathrm{GCB}(\mathscr{C}_s^\infty(\Omega,\mathbb{R});D)$ then it satisfies $\mathrm{GCB}(\mathrm{Lip}_s(\Omega,\mathbb{R});D)$.*
2. *If μ satisfies $\mathrm{GCB}(\mathrm{Lip}_s(\Omega,\mathbb{R});D)$ then it satisfies $\mathrm{GCB}(\mathrm{Lip}(\Omega,\mathbb{R});D)$.*

Proof. Let ν be a $\mathscr{C}^\infty$ (in the sense of distributions) probability measure on Ω with bounded support. For every $\lambda>0$ we define the rescaled measure ν_λ by

$$\nu_\lambda(f):=\nu(f_\lambda)$$

for any f continuous with bounded support, where $f_\lambda(x):=f(\lambda x)$. For $f\in\mathrm{Lip}_s(\Omega,\mathbb{R})$, we have $\nu_\lambda*f\in\mathscr{C}_s^\infty(\Omega,\mathbb{R})$ and $\mathrm{lip}(\nu_\lambda*f)\leq\mathrm{lip}(f)$. Since μ is assumed to satisfy $\mathrm{GCB}(\mathscr{C}_s^\infty(\Omega,\mathbb{R});D)$, it follows that

$$\mu\left(\mathrm{e}^{\nu_\lambda*f-\mu(\nu_\lambda*f)}\right)\leq\mathrm{e}^{D\,\mathrm{lip}(f)^2}\,.$$

The first statement then follows by dominated convergence.

For the second statement, as an intermediate step, we prove that if μ satisfies $\mathrm{GCB}(\mathrm{Lip}_s(\Omega,\mathbb{R});D)$ then it satisfies $\mathrm{GCB}(\mathrm{Lip}_b(\Omega,\mathbb{R});D)$. Let $\psi:\mathbb{R}^+\to\mathbb{R}^+$ be defined by

$$\psi(u)=\begin{cases}1 & \text{if}\quad u\leq 1\\ 2-u & \text{if}\quad 1\leq u\leq 2\\ 0 & \text{if}\quad u\geq 2.\end{cases}$$

For any $A>0$ define $\psi_A:\Omega\to\mathbb{R}^+$ by

$$\psi_A(x)=\psi\left(\frac{\|x\|}{A}\right).$$

We have $\psi_A\in\mathrm{Lip}_s(\Omega,\mathbb{R})$ and $\mathrm{lip}(\psi_A)\leq 1/A$. Take $f\in\mathrm{Lip}_b(\Omega,\mathbb{R})$ such that $f(0)=0$ (without loss of generality), define the function F_A by

$$F_A(x)=f(x)\psi_A(x).$$

We show that $F_A \in \mathrm{Lip}_s(\Omega, \mathbb{R})$. We have

$$F_A(x) - F_A(y) = f(x)\left[\psi_A(x) - \psi_A(y)\right] + \psi_A(y)\left[f(x) - f(y)\right].$$

Since $\|\psi_A\|_\infty \leq 1$ we get

$$\mathrm{lip}(F_A) \leq \frac{\|f\|_\infty}{A} + \mathrm{lip}(f).$$

Since μ is assumed to satisfy $\mathrm{GCB}(\mathrm{Lip}_s(\Omega, \mathbb{R}); D)$, we have

$$\mu\left(\mathrm{e}^{F_A - \mu(F_A)}\right) \leq \exp\left(D\left(\frac{\|f\|_\infty}{A} + \mathrm{lip}(f)\right)^2\right). \tag{37}$$

Using the Dominated Convergence Theorem, we take the limit $A \to +\infty$ and get

$$\mu\left(\mathrm{e}^{f - \mu(f)}\right) \leq \mathrm{e}^{D\,\mathrm{lip}(f)^2}.$$

Finally, let us prove that if μ satisfies $\mathrm{GCB}(\mathrm{Lip}_b(\Omega, \mathbb{R}); D)$ then it satisfies $\mathrm{GCB}(\mathrm{Lip}(\Omega, \mathbb{R}); D)$. Define for $M > 0$

$$f_M(x) = (f(x) \wedge M) \vee (-M).$$

By observing that $\mathrm{lip}(f_M) \leq \mathrm{lip}(f)$ and since μ satisfies $\mathrm{GCB}(\mathrm{Lip}_b(\Omega, \mathbb{R}); D)$ by assumption, we have

$$\mu\left(\mathrm{e}^{f_M - \mu(f_M)}\right) \leq \mathrm{e}^{D\,\mathrm{lip}(f)^2}. \tag{38}$$

We are going to take the limit $M \to +\infty$ and prove that the left-hand side converges to $\mu\left(\exp(f - \mu(f))\right)$. We first prove that $\sup_{M>0} |\mu(f_M)| < +\infty$. We start by proving that $\inf_{M>0} \mu(f_M) > -\infty$. Take a ball B such that $\mu(B) > 0$. Denote by x_B its center and by r_B its radius. Using (38) and the mean-value theorem, we deduce that there exists $y_M \in B$ such that

$$\mu(B)\,\mathrm{e}^{f_M(y_M) - \mu(f_M)} \leq \mathrm{e}^{D\,\mathrm{lip}(f)^2}.$$

Hence, using that $\mathrm{lip}(f_M) \leq \mathrm{lip}(f)$, we get

$$f_M(x_B) \leq \mu(f_M) + D\,\mathrm{lip}(f)^2 - \log \mu(B) + \mathrm{lip}(f) r_B.$$

Since $f_M(0) = 0$, we obtain $f_M(x_B) \geq -\mathrm{lip}(f)\|x_B\|$, which implies $\inf_{M>0} \mu(f_M) > -\infty$. A similar argument applies to $-f$, therefore

$$A_f := \sup_{M>0} |\mu(f_M)| < +\infty.$$

We now prove that e^f is integrable with respect to μ. We have

$$\mu\left(\mathrm{e}^{f_M}\right) = \mu\left(\mathbb{1}_{\{f\geq 0\}}\,\mathrm{e}^{f_M}\right) + \mu\left(\mathbb{1}_{\{f<0\}}\,\mathrm{e}^{f_M}\right). \tag{39}$$

If $x \in \Omega$ is such that $f(x) \geq 0$, then $f_M(x) \uparrow f(x)$ as $M \uparrow +\infty$, then

$$\mu\left(\mathbb{1}_{\{f\geq 0\}}\,\mathrm{e}^{f_M}\right) \leq \mu\left(\mathrm{e}^{f_M}\right) \leq \mathrm{e}^{D\,\mathrm{lip}(f)^2 + A_f}.$$

By the Monotone Convergence Theorem we thus get

$$\mu\left(\mathbb{1}_{\{f\geq 0\}}\,\mathrm{e}^{f}\right) = \lim_{M\to+\infty} \mu\left(\mathbb{1}_{\{f\geq 0\}}\,\mathrm{e}^{f_M}\right) \leq \mathrm{e}^{D\,\mathrm{lip}(f)^2 + A_f}. \tag{40}$$

Now we deal with the second term in the right-hand side of (39). Since the function $\mathbb{1}_{\{f<0\}}\,\mathrm{e}^{f_M}$ is nonnegative and bounded above by 1 and converges pointwise to $\mathbb{1}_{\{f<0\}}\,\mathrm{e}^{f}$ as M tends to $+\infty$, we apply the Dominated Convergence Theorem to get that

$$\lim_{M\to+\infty} \mu\left(\mathbb{1}_{\{f<0\}}\,\mathrm{e}^{f_M}\right) = \mu\left(\mathbb{1}_{\{f<0\}}\,\mathrm{e}^{f}\right).$$

Therefore, using this inequality, (40) and (39) we conclude that

$$\lim_{M\to+\infty} \mu\left(\mathrm{e}^{f_M}\right) = \mu\left(\mathrm{e}^{f}\right) < +\infty. \tag{41}$$

By a similar argument one shows that $\mu\left(\mathrm{e}^{-f}\right) < +\infty$.

We now prove that $\mu(f_M)$ converges to $\mu(f)$ as M tends to $+\infty$. We observe that $|f_M| \leq \mathrm{e}^f + \mathrm{e}^{-f}$. Hence by the Dominated Convergence Theorem we conclude that

$$\lim_{M\to+\infty} \mu(f_M) = \mu(f). \tag{42}$$

Using (42) and (41), we can take the limit $M \to +\infty$ in inequality (38) and obtain

$$\mu\left(\mathrm{e}^{f-\mu(f)}\right) \leq \mathrm{e}^{D\,\mathrm{lip}(f)^2}.$$

□

Acknowledgment

F. Redig thanks Ecole Polytechnique and the CNRS for financial support and hospitality. We thank R. Kraaij for useful discussions in an early stage of this work.

References

[1] D. Bakry, I. Gentil, M. Ledoux. *Analysis and geometry of Markov diffusion operators.* Grundlehren der Mathematischen Wissenschaften, 348. Springer, 2014.

[2] S. Boucheron, G. Lugosi, P. Massart. *Concentration inequalities. A nonasymptotic theory of independence.* Oxford University Press, 2013.

[3] J.-R. Chazottes, P. Collet, F. Redig. On concentration inequalities and their applications for Gibbs measures in lattice systems. J. Stat. Phys. 169 (2017), no. 3, 504–546.

[4] J.-R. Chazottes, P. Collet, C. Külske, F. Redig. Concentration inequalities for random fields via coupling. Probab. Theory Related Fields 137 (2007), no. 1-2, 201225.

[5] B. Davis. On the L_p norms of stochastic integrals and other martingales, Duke Math. J. 43, 697-704 (1976).

[6] A.C.D. van Enter, R. Fernández, F. den Hollander, F. Redig. Possible loss and recovery of Gibbsianness during the stochastic evolution of Gibbs measures. Comm. Math. Phys. 226 (2002), no. 1, 101130.

[7] H. Djellout, A. Guillin, and L. Wu. Transportation cost-information inequalities and applications to random dynamical systems and diffusions. Ann. Probab., Volume 32, Number 3B (2004), 2702-2732.

[8] M. Hirsch, S. Smale, R. Devaney. *Differential equations, dynamical systems, and an introduction to chaos.* Second edition. Pure and Applied Mathematics (Amsterdam), 60. Elsevier/Academic Press, Amsterdam, 2004.

[9] C. Külske. Concentration inequalities for functions of Gibbs fields with application to diffraction and random Gibbs measures. Comm. Math. Phys. 239 (2003), no. 1-2, 29-51.

[10] M. Ledoux. *The concentration of measure phenomenon.* Mathematical Surveys and Monographs, 89. American Mathematical Society, 2001.

[11] P. Mattila. *Geometry of Sets and Measures in Euclidean Spaces.* Cambridge University Press, 1995.

[12] I. Miyadera. *Nonlinear semigroups.* Translated from the 1977 Japanese original by Choong Yun Cho. Translations of Mathematical Monographs, 109. American Mathematical Society, Providence, RI, 1992.