HAL
open science

Measure concentration through non-Lipschitz observables and functional inequalities

Arnaud Guillin, Aldéric Joulin

To cite this version:

Arnaud Guillin, Aldéric Joulin. Measure concentration through non-Lipschitz observables and functional inequalities. Electronic Journal of Probability, 2013, 18. hal-00668800

HAL Id: hal-00668800

https://hal.science/hal-00668800v1

Submitted on 10 Feb 2012

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

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

MEASURE CONCENTRATION THROUGH NON-LIPSCHITZ OBSERVABLES AND FUNCTIONAL INEQUALITIES

ARNAUD GUILLIN AND ALDÉRIC JOULIN

Abstract. Non-Gaussian concentration estimates are obtained for invariant probability measures of reversible Markov processes. We show that the functional inequalities approach combined with a suitable Lyapunov condition allows us to circumvent the classical Lipschitz assumption of the observables. Our method is general and covers diffusions as well as pure-jump Markov processes on unbounded spaces.

1. Introduction

In the last few decades, the concentration of measure phenomenon has attracted a lot of attention. Given a metric probability space $(\mathcal{X}, d, \mu)$ and a sufficiently large class of functions defined on this space (we call them observables), the concentration of measure occurs when, observed through these functions, the space seems to be actually smaller than it is. In other words, there exists a non-decreasing continuous function $\alpha : [0, \infty) \to [0, \infty)$, null at the origin and tending to infinity at infinity, such that for a given (symmetric) class $\mathcal{C}$ of observables $f : \mathcal{X} \to \mathbb{R}$,

$$\mu\left(\{x \in \mathcal{X} : f(x) - \int_{\mathcal{X}} f \, d\mu > r\}\right) \quad \leq \quad e^{-\alpha(r)}, \quad r \geq 0.$$

The concentration is said to be Gaussian when α is quadratic-like. In connection with isoperimetry theory, the class $\mathcal{C}$ is usually taken to be the space of Lipschitz functions on $(\mathcal{X}, d, \mu)$, say $\mathrm{Lip}(\mathcal{X})$. A good review on the subject is the monograph [35] where the interested reader will find a clear introduction to the topic. One may mention also the recent progress in the area through mass transportation techniques, see the recent survey [24].

In this paper, we emphasize a dynamical point of view of concentration of measure. Given the invariant measure μ of an ergodic continuous-time Markov process $(X_t)_{t\geq 0}$ with *carré du champ* operator Γ (see below for the definition), we provide concentration properties of μ through observables which depend on the dynamics. As it is sometimes the case in previous studies, see for instance [36] for a state of the art, our starting point is to assume that the pair (μ, Γ) satisfies a convenient

2000 *Mathematics Subject Classification.* 46E35, 60E15, 60J27, 60J60, 60K35.

Key words and phrases. Concentration, invariant measure, reversible Markov process, Lyapunov condition, functional inequality, carré du champ, diffusion process, jump process.

functional inequality such as Poincaré or the entropic inequality. Such inequalities, which are verified by a wide variety of examples, are closely related to the long-time behaviour of the process. In particular, this approach allows to unify both continuous and discrete space settings even if, in essence, these two objects are rather different from each other. In the diffusion framework, the main ingredient to establish Gaussian concentration through functional inequalities is based on the chain rule derivation formula satisfied by the operator Γ. In the context of Markov jump processes, although this property is not verified - here Γ becomes a finite difference operator - this difficulty can be circumvented by using reversibility and Gaussian concentration properties are still available in this discrete situation. In both cases the carré du champ refers to a natural distance related to the dynamics and, within this notion of distance, the Lipschitz observables under which concentration estimates are obtained are the ones with a bounded carré du champ, that is to say the space $\mathrm{Lip}_\Gamma(\mathcal{X})$ of functions f such that $\Gamma(f,f)$ is μ-essentially bounded. Then a natural question arises: which type of measure concentration can we obtain beyond the space $\mathrm{Lip}_\Gamma(\mathcal{X})$? In particular in discrete space settings, such a study basically makes sense on an unbounded state space $\mathcal{X}$. Using the notion of Ricci curvature for Markov chains (a definition also available in continuous-time), a first result of this kind was given by Ollivier in [38], in which he obtains concentration bounds involving a mixed Gaussian-exponential regime, i.e. $\alpha(r)$ is quadratic/linear for small/large r. In our language, he requires that the carré du champ $\Gamma(f,f)$ belongs to the space $\mathrm{Lip}(\mathcal{X})$. Despite this interesting and new result, which is sufficiently robust to be extended to additive functionals, see e.g. [31] and [32], it seems to the authors that there is no satisfactory treatment yet to this question and we hope to give (the beginning of) an answer to this problem with the present article.

Our idea is to use a Lyapunov condition on the observables. This kind of criteria have been successfully used for proving various types of functional inequalities, cf. [2, 12, 13, 23] and for concentration estimates of additive functionals, see for instance [11, 22, 16]. Namely we will consider the class $\mathcal{L}_V(a,b)$ of observables f such that

$$\Gamma(f,f) \leq -a\,\frac{\mathcal{L}V}{V} + b,$$

where a, b are two positive constants and V is a convenient test function. When a vanishes the class $\mathcal{L}_V(a,b)$ reduces to the space $\mathrm{Lip}_\Gamma(\mathcal{X})$ and the classical concentration results apply [36]. In particular, the behaviour of the carré du champ $\Gamma(f,f)$ depends now on the growth of the term $-\mathcal{L}V/V$, which has no reason to be bounded. Of course there is a price to pay for such an improvement and it resides in the concentration property of the measure μ, which is no longer Gaussian but only of Gaussian-exponential type under this class of observables.
To fix the ideas, let us consider a simple example on the set $\mathbb{N} := \{0,1,2,\ldots\}$ endowed with the classical distance between integers. Denote μ_p the geometric

distribution on $\mathbb{N}$ of parameter $p \in (0,1)$, i.e. $\mu_p(\{x\}) := (1-p)p^x$, $x \in \mathbb{N}$. Then there exist infinitely many dynamics on $\mathbb{N}$ admitting μ_p as their reversible invariant measure, and among them let us consider the Markov process with carré du champ

$$\Gamma^{(n)}(f,f)(x) := \frac{1}{2}\left\{p(x+1)^n\left(f(x+1)-f(x)\right)^2 + x^n 1_{\{x\neq 0\}}\left(f(x-1)-f(x)\right)^2\right\},$$

where $x \in \mathbb{N}$ and $n \in \mathbb{N}$ is some fixed parameter. On the one hand we can prove that for every $n \in \mathbb{N}$ the dynamics $(\mu_p, \Gamma^{(n)})$ satisfies a Poincaré inequality and adapting then to the discrete case the method introduced by Aida and Stroock in [1] entails a mixed Gaussian-exponential regime for the measure μ_p, under observables $f \in \mathrm{Lip}_{\Gamma^{(n)}}(\mathbb{N})$. Actually, these observables correspond to:

- functions belonging to the space $\mathrm{Lip}(\mathbb{N})$ in the case $n = 0$. Such an estimate is sharp, cf. for instance [27, 29];
- functions of order $\sqrt{x}$ in the case $n = 1$. According to the previous line, the measure μ_p should concentrate at least like a Gaussian through these observables;
- functions of order $\log(x)$ in the case $n = 2$. Under these observables, $\alpha(r)$ should behave like e^{-e^r} for large r;
- bounded functions in the case $n \geq 3$.

On the other hand Ollivier's method cleverly applies, but only in the case $n = 1$. Hence one observes that these two methods does not lead to sharp results as soon as $n \geq 2$. Actually, thanks to the Lyapunov approach, the Gaussian-exponential concentration results we give in this paper apply for observables f with carré du champ $\Gamma^{(n)}(f,f)(x)$ of order x^n, i.e. of the expected order of magnitude.

The paper is organized as follows. In section 2, we recall some basic material on Markov processes and functional inequalities. Two types of processes are considered in our study: diffusions and pure-jump Markov processes. Next we state in section 3 our main results of the paper, theorems 3.6 and 3.11, in which some mixed Gaussian-exponential concentration properties of μ are obtained through observables satisfying the Lyapunov condition defined above and under a convenient functional inequality assumption satisfied by the dynamics (μ, Γ). As a result, such new concentration inequalities extend the classical estimates obtained when the observables belong to the space $\mathrm{Lip}_\Gamma(\mathcal{X})$, corresponding to the case where a vanishes. Finally, section 4 is devoted to numerous examples in continuous and discrete settings. First we investigate the diffusion case of Kolmogorov processes whose invariant measure has density proportional to e^{-U}, where U is some nice potential on $\mathbb{R}^d$, and for which we derive a concentration result when U is considered as an observable, in the spirit of the recent progress made by Bobkov and Madiman in [8]. Dealing with jump processes, the case of birth-death processes is addressed in detail. When we apply our concentration estimates to observables in $\mathrm{Lip}(\mathbb{N})$ equipped with the classical distance between integers, we are able to consider processes whose generator is unbounded, like the basic example investigated above. Finally, we focus our attention on an unbounded interacting particle

system, namely the Glauber dynamics associated to a Gibbs measure defined with respect to a Poisson reference measure.

2. Preliminaries

2.1. **Functional inequalities.** Throughout the paper, $(\mathcal{X}, d)$ is a Polish space endowed with the corresponding Borel σ-field $\mathscr{B}$ and $\mathrm{Lip}(\mathcal{X})$ is the space of Lipschitz functions on $\mathcal{X}$ with finite Lipschitz seminorm with respect to d, i.e.

$$\|f\|_{\mathrm{Lip}} \;:=\; \sup_{x\neq y} \frac{|f(x)-f(y)|}{d(x,y)} < \infty.$$

On a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$, let $\{(X_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in\mathcal{X}}\}$ be an $\mathcal{X}$-valued *càdlàg* ergodic Markov process with reversible invariant measure (or stationary distribution) μ and symmetric semigroup $(P_t)_{t\geq 0}$ on $L^2(\mu)$. In the sequel we denote $L^p(\mu) := L^p(\mathcal{X}, \mathscr{B}, \mu)$ for $p \in [1, \infty]$. Denote $\mathcal{L}$ the self-adjoint generator acting on its dense domain $\mathcal{D}_2(\mathcal{L})$ consisting of functions $f \in L^2(\mu)$ such that $t^{-1}(P_t f - f)$ admits a limit in $L^2(\mu)$ as $t \to 0$. One of the main protagonists of the present paper is the *carré du champ* Γ, which is a bilinear symmetric operator on $\mathcal{D}_2(\mathcal{L}) \times \mathcal{D}_2(\mathcal{L})$ defined by

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

As mentioned in the introduction, there is a natural pseudo-distance associated to the operator Γ which can be defined as

$$d_\Gamma(x,y) := \sup\Big\{|f(x)-f(y)| : \|\Gamma(f,f)\|_{L^\infty(\mu)} \leq 1\Big\}, \quad x, y \in \mathcal{X}.$$

Although this distance can be infinite, it is well-defined in the situations of interest and carries a lot of information about the structure of the underlying process. In the sequel, we denote $\mathrm{Lip}_\Gamma(\mathcal{X})$ the space of Lipschitz functions with respect to d_Γ. The associated Dirichlet form acts on $\mathcal{D}(\mathcal{E}_\mu) \times \mathcal{D}(\mathcal{E}_\mu)$ as

$$\mathcal{E}_\mu(f,g) \;=\; \int_{\mathcal{X}} \Gamma(f,g)\, d\mu,$$

where $\mathcal{D}(\mathcal{E}_\mu)$ is the subspace of all functions $f \in \mathcal{D}_2(\mathcal{L})$ such that $\mathcal{E}_\mu(f,f)$ is well-defined. In particular, the Donsker-Varadhan information of any probability measure ν on $\mathcal{X}$ with respect to the invariant measure μ is defined as

$$I(\nu|\mu) := \begin{cases} \mathcal{E}_\mu\left(\sqrt{f}, \sqrt{f}\right) & \text{if} \quad d\nu = f d\mu, \quad \sqrt{f} \in \mathcal{D}(\mathcal{E}_\mu); \\ \infty & \text{otherwise.} \end{cases}$$

A key point in our analysis is that the functional $\nu \mapsto I(\nu|\mu)$ is nothing but the rate function governing the Large Deviation Principle in large time of the empirical measure of $(X_t)_{t\geq 0}$. However in the non-reversible case, it is given by a contraction form of the Donsker-Varadhan entropy which is different from the Donsker-Varadhan information, so that our study will not extend to the non-symmetric case, unfortunately.

Now let us introduce the functional inequalities we will focus on in the paper. Given an integrable function $f \in L^1(\mu)$, we denote $\mu(f) := \int_{\mathcal{X}} f d\mu$. Let I be an open interval of $\mathbb{R}$ and for a convex function $\phi : I \to \mathbb{R}$ we define the ϕ-entropy of a function $f : \mathcal{X} \to I$ with $\phi(f) \in L^1(\mu)$ as

$$\mathrm{Ent}_\mu^\phi(f) := \mu\left(\phi(f)\right) - \phi\left(\mu(f)\right).$$

The dynamics (μ, Γ) satisfies a ϕ-entropy inequality with constant $C_\phi > 0$ if for any I-valued function $f \in \mathcal{D}(\mathcal{E}_\mu)$ such that $\phi'(f) \in \mathcal{D}(\mathcal{E}_\mu)$,

$$C_\phi \,\mathrm{Ent}_\mu^\phi(f) \leq \frac{1}{2}\,\mathcal{E}_\mu(f, \phi'(f)).$$

See for instance [15] for a careful investigation of the properties of ϕ-entropies. The latter inequality is satisfied if and only if the following entropy dissipation of the semigroup holds: for any I-valued function f such that $\phi(f) \in L^1(\mu)$,

$$\mathrm{Ent}_\mu^\phi(P_t f) \leq e^{-2C_\phi t}\,\mathrm{Ent}_\mu^\phi(f), \quad t \geq 0.$$

In this paper we will consider three cases:

(i) the Poincaré inequality: $\phi(u) = u^2$ with $I = \mathbb{R}$ and the ϕ-entropy inequality rewrites as

$$\lambda\,\mathrm{Var}_\mu(f) \leq \mathcal{E}_\mu\left(f, f\right), \tag{2.1}$$

where the variance of f under μ is given by

$$\mathrm{Var}_\mu(f) := \mu(f^2) - \mu(f)^2.$$

The optimal constant λ_1 (say) is nothing but the spectral gap in $L^2(\mu)$ of the operator $-\mathcal{L}$, i.e. its smallest non-zero eigenvalue. Estimating λ_1 allows us to obtain the optimal rate of convergence of the semigroup in $L^2(\mu)$.

(ii) the entropic inequality: $\phi(u) = u \log u$ with $I = (0, \infty)$ and the ϕ-entropy inequality is given by

$$\rho\,\mathrm{Ent}_\mu(f) \leq \mathcal{E}_\mu\left(f, \log f\right), \tag{2.2}$$

where the entropy under μ of the smooth positive function f is defined by

$$\mathrm{Ent}_\mu(f) := \mu(f\,\log f) - \mu(f)\,\log \mu(f).$$

We have skipped in the inequality the constant $1/2$ for convenience in future computations. Once again, the best constant ρ_0 in (2.2) gives the optimal exponential decay of the entropy along the semigroup.

(iii) the Beckner-type inequality: $\phi(u) = u^p$ with $p \in (1, 2]$ and $I = (0, \infty)$. We have in this case

$$\alpha_p\,\left(\mu(f^p) - \mu(f)^p\right) \leq \frac{p}{2}\,\mathcal{E}_\mu\left(f, f^{p-1}\right). \tag{2.3}$$

Estimating α_p gives the optimal rate of convergence of the semigroup in $L^p(\mu)$.

The entropic and Beckner-type inequalities are stronger than the Poincaré inequality (apply these inequalities to the function $1 + \varepsilon f$ and take the limit as $\varepsilon \to 0$). Moreover it reduces to the Poincaré inequality (2.1) if $p = 2$, whereas

dividing both sides by $p-1$ and taking the limit as $p \to 1$ we obtain the entropic inequality (2.2).

In this paper we will mainly consider two general classes of reversible Markov processes: diffusions and pure jump Markov processes to which we turn now.

2.2. **Diffusion processes.** A diffusion process on the Euclidean space $\mathcal{X} = \mathbb{R}^d$ corresponds to a path continuous Markov process on $\mathbb{R}^d$ whose generator $\mathcal{L}$ is a second order differential operator: for any sufficiently smooth function $f : \mathbb{R}^d \to \mathbb{R}$,

$$\mathcal{L}f(x) = \sum_{i,j=1}^{d} a_{i,j}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{d} b_i(x) \frac{\partial f}{\partial x_i}(x), \quad x \in \mathbb{R}^d.$$

Here $a := \sigma\sigma^*$ is a measurable and locally bounded function from $\mathbb{R}^d$ to the space of $d \times d$ symmetric positive definite matrices with smooth entries, σ^* being the transpose of the matrix σ, and the measurable drift $b : \mathbb{R}^d \to \mathbb{R}^d$ is also assumed to be smooth. In this case the carré du champ is given by

$$\begin{aligned} \Gamma(f,g) &= \sum_{i,j=1}^{d} a_{i,j} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \\ &= < \sigma^* \nabla f, \sigma^* \nabla g >, \end{aligned}$$

where $< \cdot, \cdot >$ stands for the Euclidean scalar product in $\mathbb{R}^d$ and ∇ is the usual gradient operator. In particular when σ is the identity matrix, the spaces $\mathrm{Lip}(\mathbb{R}^d)$ and $\mathrm{Lip}_\Gamma(\mathbb{R}^d)$ might be identified.
In contrast to the jump case introduced below, Γ is a differentiation, i.e. for all any smooth enough functions $(f_k)_{1 \le k \le n}$, $f : \mathbb{R}^d \to \mathbb{R}$ and any C^1 function $\phi : \mathbb{R}^n \to \mathbb{R}$,

$$\Gamma(\phi(f_1, \dots, f_n), f) = \sum_{i=1}^{n} \frac{\partial \phi}{\partial x_i}(f_1, \dots, f_n)\, \Gamma(f_i, f). \tag{2.4}$$

Due to this chain rule derivation formula, the entropic inequality (2.2) rewrites in the diffusion case as the famous log-Sobolev inequality

$$\rho\, \mathrm{Ent}_\mu(f^2) \le 4\, \mathcal{E}_\mu(f,f), \tag{2.5}$$

which is the original inequality (up to the extra factor 4) derived by Gross in [25] to study hypercontractivity of the underlying semigroup. When we will consider diffusion processes in the sequel, we will use the terminology "log-Sobolev inequality" instead of "entropic inequality".
On the other hand, letting $p = 2/q$ for $q \in [1,2)$ and $f = g^q$, the Beckner-type inequality (2.3) rewrites as the so-called standard Beckner inequality:

$$\alpha_{2/q} \left(\mu(g^2) - \mu(g^q)^{2/q} \right) \le (2-q)\, \mathcal{E}_\mu(g,g). \tag{2.6}$$

Such an inequality was introduced by Beckner in [3] for the Gaussian measure. In particular, the limiting case $q \to 2$ recovers the classical log-Sobolev inequality.

Note however that the inequality (2.6) is weaker than the log-Sobolev inequality, cf. [34].

2.3. **Markov jump processes.** Dealing with a pure-jump Markov process, the generator $\mathcal{L}$ is given for any function $f \in \mathcal{D}_2(\mathcal{L})$ by

$$\mathcal{L}f(x) = \int_{\mathcal{X}} (f(y) - f(x))\, Q_x(dy), \quad x \in \mathcal{X},$$

where the transition kernel $x \mapsto Q_x$ is a measurable mapping from $\mathcal{X}$ to the set of Radon measures on $\mathcal{X}$ endowed with the corresponding Borel σ-field. We assume that it satisfies the following stability assumption:

$$\int_{\mathcal{X}} Q_x(dy) < \infty, \quad x \in \mathcal{X}, \tag{2.7}$$

which entails that the process is piecewise constant. Here, reversibility means that the following detailed balance condition is satisfied:

$$Q_x(dy)\, \mu(dx) = Q_y(dx)\, \mu(dy). \tag{2.8}$$

The carré du champ operator Γ admits an explicit expression given for any $f, g \in \mathcal{D}_2(\mathcal{L})$ by

$$\Gamma(f,g)(x) = \frac{1}{2} \int_{\mathcal{X}} (f(y) - f(x))\, (g(y) - g(x))\; Q_x(dy),$$

and we have

$$\Gamma(f,f)(x) = \frac{1}{2} \int_{\mathcal{X}} (f(y) - f(x))^2\; Q_x(dy).$$

In particular, the spaces $\mathrm{Lip}(\mathcal{X})$ and $\mathrm{Lip}_\Gamma(\mathcal{X})$ have no reason to coincide since the kernel of the generator may be unbounded, i.e.

$$\sup_{x \in \mathcal{X}} \int_{\mathcal{X}} Q_x(dy) = \infty. \tag{2.9}$$

Finally the Dirichlet form rewrites for any $f, g \in \mathcal{D}(\mathcal{E}_\mu)$ as

$$\begin{aligned} \mathcal{E}_\mu(f,g) &= \frac{1}{2} \int_{\mathcal{X}} \int_{\mathcal{X}} (f(y) - f(x))\, (g(y) - g(x))\; Q_x(dy)\mu(dx) \\ &= \iint\limits_{f(x) > f(y)} (f(y) - f(x))\, (g(y) - g(x))\; Q_x(dy)\mu(dx), \end{aligned}$$

where in the last line the reversibility is used. In our jump framework, the entropic inequality (2.2) corresponds to one of the so-called modified log-Sobolev inequalities introduced by Bobkov and Ledoux in [6]. However, due to the lack of chain rule for discrete gradients, this inequality is different from the discrete version of the log-Sobolev inequality (2.5), and the same remark holds between the Beckner-type inequality (2.3) and the standard Beckner inequality (2.6). We refer to [21, 6, 9] for historical and tutorial references on these discrete functional inequalities, together with a hierarchy of the various modified log-Sobolev inequalities.

3. Main results

As announced, we obtain concentration properties of the invariant measure μ through observables which are not required to belong to the spaces $\mathrm{Lip}(\mathcal{X})$ nor $\mathrm{Lip}_\Gamma(\mathcal{X})$, but which satisfy a Lyapunov condition. In order to state this condition properly, let us introduce first the extended domain of the generator. Denote the probability measure $\mathbb{P}_\nu(\cdot) := \int_{\mathcal{X}} \mathbb{P}_x(\cdot)\,\nu(dx)$ where ν is an arbitrary initial probability distribution. A continuous function f is said to belong to the extended domain $\mathcal{D}_e(\mathcal{L})$ of the generator $\mathcal{L}$ if there exists some measurable function $g : \mathcal{X} \to \mathbb{R}$ such that for any $t \geq 0$, $\int_0^t |g(X_s)|\,ds < \infty$, $\mathbb{P}_\mu$-a.s. and the process

$$M_t^f = f(X_t) - f(X_0) - \int_0^t g(X_s)\,ds, \quad t \geq 0,$$

is a local $\mathbb{P}_\mu$-martingale. In this case we write $f \in \mathcal{D}_e(\mathcal{L})$ and $\mathcal{L}f = g$.

The first result on which our analysis is based is closely related to the theory of large deviations, see for instance [43].

Lemma 3.1. *The Donsker-Varadhan information $\nu \mapsto I(\nu|\mu)$ is the rate function governing the Large Deviation Principle of the empirical measure of the reversible process $(X_t)_{t\geq 0}$. In other words,*

$$I(\nu|\mu) = \sup_{0<V\in\mathcal{D}_e(\mathcal{L})} \int_{\mathcal{X}} -\frac{\mathcal{L}V}{V}\,d\nu. \tag{3.1}$$

Remark 3.2. Actually, we will only need the following inequality, available for any positive test function $V \in \mathcal{D}_e(\mathcal{L})$:

$$\int_{\mathcal{X}} -\frac{\mathcal{L}V}{V}\,d\nu \;\leq\; I(\nu|\mu).$$

For instance in the diffusion case, we can prove it as follows: assume that ν has density f^2 with respect to μ (trivial otherwise). Then by the chain rule formula satisfied by the carré du champ,

$$\begin{aligned}
\int_{\mathcal{X}} -\frac{\mathcal{L}V}{V}\,d\nu &= \int_{\mathcal{X}} \Gamma\left(V, \frac{f^2}{V}\right) d\mu \\
&= \int_{\mathcal{X}} \left(\frac{2f}{V}\,\Gamma(V,f) - \frac{f^2}{V^2}\,\Gamma(V,V)\right) d\mu \\
&\leq \int_{\mathcal{X}} \left(\frac{2f}{V}\sqrt{\Gamma(V,V)}\sqrt{\Gamma(f,f)} - \frac{f^2}{V^2}\,\Gamma(V,V)\right) d\mu \\
&\leq \int_{\mathcal{X}} \Gamma\left(f,f\right)\,d\mu \\
&= I(\nu|\mu),
\end{aligned}$$

where we used Cauchy-Schwarz' inequality to obtain the first inequality, and also the elementary bound $2xy \leq x^2 + y^2$, $x, y \in \mathbb{R}$, to get the final result.

Now we are able to state the Lyapunov condition we will focus on along this paper.

Definition 3.3. *Let a, b be two positive constants and let $V \in \mathcal{D}_e(\mathcal{L})$ be a positive test function. A function $f \in \mathcal{D}_2(\mathcal{L})$ belongs to the class $\mathcal{L}_V(a,b)$ if the following inequality is satisfied $\mu - a.s.$:*

$$\Gamma(f,f) \leq -a\,\frac{\mathcal{L}V}{V} + b. \tag{3.2}$$

Remark 3.4. The Poincaré inequality can be seen as a minimal assumption in our study of concentration by means of the Lyapunov condition (3.2). Indeed, if there exists a function $f \in \mathcal{D}_2(\mathcal{L})$ such that $\Gamma(f,f)$ is lower bounded by a positive constant at infinity, and this the case in the main examples of interest (except in the Cauchy-like case appearing in section 4), then the Poincaré inequality is satisfied, cf. [14]. Moreover, integrating with respect to μ entails that $f \in \mathcal{D}(\mathcal{E}_\mu)$ and by Poincaré inequality we have $\mathrm{Var}_\mu(f) \leq b/\lambda_1$. In other words the constant b/λ_1 can be interpreted in the sequel as a variance term of the observable f.

Before stating our first main result, let us provide a key lemma. In the remainder of this paper, we only give the proofs in the jump case since the diffusion framework requires no additional difficulties and is even simpler, according to the chain rule derivation formula (2.4) satisfied by the carré du champ.

Lemma 3.5. *Let f belong to the class $\mathcal{L}_V(a,b)$. Given $\lambda \in (0, 2/\sqrt{a})$, let μ_λ be the probability measure with density $f_\lambda := e^{\lambda f}/Z_\lambda$ with respect to μ, where Z_λ is the appropriate normalization constant. We assume that $\sqrt{f_\lambda} \in \mathcal{D}(\mathcal{E}_\mu)$. Then*

$$I(\mu_\lambda|\mu) \leq \frac{\lambda^2 b}{4 - \lambda^2 a}, \quad 0 < \lambda < \frac{2}{\sqrt{a}}.$$

Proof. Since $f \in \mathcal{L}_V(a,b)$, we have for any $\lambda \in (0, 2/\sqrt{a})$:

$$\begin{aligned}
I(\mu_\lambda|\mu) &= \frac{1}{Z_\lambda} \iint\limits_{f(x)>f(y)} \left(e^{\lambda f(x)/2} - e^{\lambda f(y)/2}\right)^2 Q_x(dy)\mu(dx) \\
&= \iint\limits_{f(x)>f(y)} \left(1 - e^{-\lambda(f(x)-f(y))/2}\right)^2 f_\lambda(x)\, Q_x(dy)\mu(dx) \\
&\leq \frac{\lambda^2}{4} \int_{\mathcal{X}} \Gamma(f,f)\, d\mu_\lambda \\
&\leq \frac{\lambda^2}{4} \int_{\mathcal{X}} \left(-a\,\frac{\mathcal{L}V}{V} + b\right) d\mu_\lambda \\
&\leq \frac{\lambda^2}{4} \left(aI(\mu_\lambda|\mu) + b\right),
\end{aligned}$$

where in the last line we used lemma 3.1. Finally rearranging the terms allows us to obtain the desired inequality. ∎

We turn now to our first main and new result which exhibits a non-Gaussian concentration estimate through observables belonging to the class $\mathcal{L}_V(a,b)$. Due to the approach we will use, the numerical constants in the estimates below have no reason to be sharp.

Theorem 3.6. *Assume that the pair (μ,Γ) satisfies the entropic inequality (2.2) with optimal constant ρ_0. Let $f\in\mathcal{L}_V(a,b)$ and let*

$$r_{\max}:=\frac{8b}{3\rho_0\sqrt{a}}$$

be the size of the Gaussian window. Then the invariant measure μ has the following concentration property: for any $0\le r\le r_{\max}$, the deviation is of Gaussian-type:

$$\mu\left(\{x\in\mathcal{X}: f(x)-\mu(f)>r\}\right)\le e^{-\frac{3\rho_0 r^2}{16b}}, \tag{3.3}$$

and for any $r\ge r_{\max}$, the decay is exponential:

$$\mu\left(\{x\in\mathcal{X}: f(x)-\mu(f)>r\}\right)\le e^{-\frac{r}{2\sqrt{a}}}, \tag{3.4}$$

Remark 3.7. In the sequel, a concentration property such as (3.3)-(3.4) will be called Gaussian-exponential concentration.

Proof. Denote $L_\lambda:=\lambda^{-1}\log Z_\lambda$, where $Z_\lambda:=\int_{\mathcal{X}}e^{\lambda f}\,d\mu$, with $\lambda\in(0,1/\sqrt{a})$, and let μ_λ be the absolutely continuous probability measure with density $f_\lambda:=e^{\lambda f}/Z_\lambda$ with respect to μ. Using a standard approximation procedure one may assume that the observable $f\in\mathcal{L}_V(a,b)$ is bounded so that $\sqrt{f_\lambda}\in\mathcal{D}(\mathcal{E}_\mu)$. The following proof is a modification of the famous Herbst method popularized by Ledoux. Using the entropic inequality (2.2),

$$\begin{aligned}
\frac{d}{d\lambda}L_\lambda &= \frac{1}{\lambda^2 Z_\lambda}\operatorname{Ent}_\mu(e^{\lambda f})\\
&\le \frac{1}{\rho_0\lambda^2 Z_\lambda}\,\mathcal{E}_\mu\left(\lambda f, e^{\lambda f}\right)\\
&= \frac{1}{\rho_0\lambda}\iint\limits_{f(x)>f(y)}(f(x)-f(y))\left(1-e^{-\lambda(f(x)-f(y))}\right)f_\lambda(x)\,Q_x(dy)\mu(dx)\\
&\le \frac{1}{\rho_0}\int_{\mathcal{X}}\int_{\mathcal{X}}\Gamma(f,f)\,d\mu_\lambda\\
&\le \frac{1}{\rho_0}\int_{\mathcal{X}}\left(-a\frac{\mathcal{L}V}{V}+b\right)d\mu_\lambda\\
&\le \frac{1}{\rho_0}\left(aI(\mu_\lambda|\mu)+b\right),
\end{aligned}$$

where we used that $f\in\mathcal{L}_V(a,b)$ and then lemma 3.1 in the two last lines. Thus lemma 3.5 entails the inequality

$$\frac{d}{d\lambda}L_\lambda \le \frac{4b}{3\rho_0},\quad 0<\lambda<\frac{1}{\sqrt{a}},$$

and therefore the following log-Laplace estimate is available for any $0 < \lambda < \frac{1}{\sqrt{a}}$:

$$\log \int_{\mathcal{X}} e^{\lambda f}\, d\mu \leq \lambda \mu(f) + \frac{4b\lambda^2}{3\rho_0}. \tag{3.5}$$

Finally using Chebyshev's inequality and optimizing in $\lambda \in (0, 1/\sqrt{a})$ yields the tail estimates (3.3) and (3.4). The proof of theorem 3.6 is thus complete. □

Remark 3.8. Two deviation regimes appear, Gaussian and exponential, with continuous transition from one to the other. In contrast to the classical Herbst method where the observables belong to $\mathrm{Lip}_\Gamma(\mathcal{X})$, i.e. $a = 0$ in the Lyapunov condition (3.2), our assumption allows us to go beyond this Lipschitz property. However the price to pay is to have a finite Gaussian window, i.e. $r_{\max} < \infty$.

Remark 3.9. By the Central Limit Theorem, the order of magnitude is correct in the Gaussian regime. Since the entropic inequality entails a Poincaré inequality, we have $\rho_0 \leq \lambda_1$ and thus for any observable $f \in \mathcal{L}_V(a, b)$, we get $\mathrm{Var}_\mu(f) \leq b/\rho_0$. Therefore, if $\mu = \nu^{\otimes d}$ is a product measure and $f(x) = \sum_{k=1}^d \phi(x_k)$, $x := (x_1, \ldots, x_d) \in \mathcal{X}$, then we obtain the following inequality, which is sharp for large d:

$$\mu\left(\left\{x \in \mathcal{X} : f(x) - \mu(f) > r\sqrt{d}\right\}\right) \leq e^{-\frac{3\rho_0 r^2}{16b}}, \quad 0 \leq r \leq \frac{8\tilde{b}\sqrt{d}}{3\rho_0\sqrt{a}}.$$

Here the important point is that the positive parameter $\tilde{b}$ depends on ϕ but is independent of d.

Remark 3.10. The method is sufficiently robust to get, for large deviation level r, other regimes than exponential under particular observables. For example, assume that we consider $f \in \mathcal{L}_V(a, b)$ such that $\Gamma(f, f) \ll -a\mathcal{L}V/V + b$ but that there exists two functions $\phi, \psi : (0, \infty) \to (0, \infty)$ such that for all $\varepsilon > 0$,

$$\Gamma(f, f) \leq \phi(\varepsilon)\left(-a\,\frac{\mathcal{L}V}{V} + b\right) + \psi(\varepsilon).$$

Then plugging this estimate in the previous proofs of lemma 3.5 and theorem 3.6, one has for all $\varepsilon > 0$,

$$I(\mu_\lambda|\mu) \leq \frac{\lambda^2}{4}\left(a\phi(\varepsilon)I(\mu_\lambda|\mu) + b\phi(\varepsilon) + \psi(\varepsilon)\right).$$

Optimizing in $\varepsilon > 0$ enables to get for some function $\Phi : [0, \infty) \to [0, \infty]$ and all $\lambda > 0$,

$$I(\mu_\lambda|\mu) \leq \Phi(\lambda),$$

leading then to super-exponential regime for large r. We will illustrate this on an example in section 4.

Inspired by Otto-Villani's method, cf. [39, 41] where the links between log-Sobolev and transportation inequalities are studied on continuous and finite state spaces respectively, let us recover theorem 3.6 by using a semigroup proof. Once

again we focus our attention on the jump case. Let h be a smooth density with respect to μ. Given $t > 0$, denote ν_t the probability measure with density $P_t h$ with respect to μ. We assume that the Donsker-Varadhan information $I(\nu_t|\mu)$ is well-defined, i.e. $\sqrt{P_t h} \in \mathcal{D}(\mathcal{E}_\mu)$. Using Cauchy-Schwarz's inequality and then reversibility,

$$\begin{aligned}
\mathcal{E}_\mu(P_t h, f) &= \frac{1}{2}\int_{\mathcal{X}}\int_{\mathcal{X}} \left(P_t h(x) - P_t h(y)\right)\left(f(x) - f(y)\right) Q_x(dy)\mu(dx) \\
&\leq \sqrt{I(\nu_t|\mu)}\sqrt{\frac{1}{2}\int_{\mathcal{X}}\int_{\mathcal{X}} \left(\sqrt{P_t h(x)} + \sqrt{P_t h(y)}\right)^2 \left(f(x) - f(y)\right)^2 Q_x(dy)\mu(dx)} \\
&\leq 2\sqrt{I(\nu_t|\mu)}\sqrt{\int_{\mathcal{X}} \Gamma(f,f)\, d\nu_t} \\
&\leq 2\sqrt{I(\nu_t|\mu)}\sqrt{\int_{\mathcal{X}} \left(-a\,\frac{\mathcal{L}V}{V} + b\right) d\nu_t} \\
&\leq 2\sqrt{I(\nu_t|\mu)}\sqrt{aI(\nu_t|\mu) + b},
\end{aligned}$$

where in the two last lines we used that $f \in \mathcal{L}_V(a,b)$ and then lemma 3.1. Using now the elementary inequality $2(a-b)^2 \leq (a^2 - b^2)\log(a/b)$ available for any $a, b > 0$ and then the entropic inequality (2.2), we get

$$\begin{aligned}
\mathcal{E}_\mu(P_t h, f) &\leq \mathcal{E}_\mu(P_t h, \log P_t h)\left(\frac{\sqrt{a}}{2} + \sqrt{\frac{b}{\mathcal{E}_\mu(P_t h, \log P_t h)}}\right) \\
&\leq \mathcal{E}_\mu(P_t h, \log P_t h)\left(\frac{\sqrt{a}}{2} + \sqrt{\frac{b}{\rho_0 \mathrm{Ent}_\mu(P_t h)}}\right).
\end{aligned}$$

Integrating time between 0 and infinity entails the covariance inequality:

$$\begin{aligned}
\mathrm{Cov}_\mu(f,h) &:= \mu(fh) - \mu(f)\,\mu(h) \\
&= \int_0^\infty \mathcal{E}_\mu(P_t h, f)\, dt \\
&\leq \int_0^\infty \mathcal{E}_\mu(P_t h, \log P_t h)\left(\frac{\sqrt{a}}{2} + \sqrt{\frac{b}{\rho_0 \mathrm{Ent}_\mu(P_t h)}}\right) dt \\
&= \frac{\sqrt{a}}{2}\,\mathrm{Ent}_\mu(h) + 2\sqrt{\frac{b\,\mathrm{Ent}_\mu(h)}{\rho_0}},
\end{aligned}$$

which in turn yields the inequality

$$\alpha\left(\mathrm{Cov}_\mu(f,h)\right) \leq \mathrm{Ent}_\mu(h),$$

where α is the function

$$\alpha(r) = \frac{\rho_0 r^2}{4b + 2\rho_0\sqrt{a}\, r}, \quad r > 0.$$

Finally using theorem 3.2 in [24], we obtain the following concentration estimate through the observable $f \in \mathcal{L}_V(a,b)$:

$$\mu\left(\{x \in \mathcal{X} : f(x) - \mu(f) > r\}\right) \;\leq\; e^{-\alpha(r)}, \quad r \geq 0.$$

One deduces that, up to numerical constants, this result is similar to that emphasized in theorem 3.6 since for small deviation level, $\alpha(r) = O(\rho_0 r^2/b)$ whereas $\alpha(r) = O(r/\sqrt{a})$ for large r.

As we have seen above, the entropic inequality (2.2) entails on the one hand a concentration property for the invariant measure μ through observables $f \in \mathcal{L}_V(a,b)$. On the other hand and as announced in remark 3.4, the Poincaré inequality can be seen as a minimal assumption in our study. Hence one may wonder if the Beckner-type inequality (2.3), which interpolates between both, provides qualitative concentration estimates through observables in $\mathcal{L}_V(a,b)$. Our second main result, theorem 3.11, goes in this way. However, although the estimates we provide below are somewhat similar to that of theorem 3.6 in regard of the mixed Gaussian-exponential behaviour, the price to pay is to lose the good order of magnitude for the Gaussian window since we obtain $r_{\max} = O(\sqrt{b})$ instead of $O(b)$.

Theorem 3.11. *Assume that there exists $p \in (1,2]$ such that the pair (μ,Γ) satisfies a Beckner-type inequality (2.3) and denote α_p the optimal constant. Moreover, assume that the observable $f \in \mathcal{L}_V(a,b)$ with the restriction $\alpha_p \leq 2b(p-1)/(3pa)$ and let*

$$r_{\max} := \sqrt{\frac{32bp}{27(p-1)\alpha_p}}$$

be the size of the Gaussian window. Then the following tail estimates hold: for any $0 \leq r \leq r_{\max}$,

$$\mu\left(\{x \in \mathcal{X} : f(x) - \mu(f) > r\}\right) \leq e^{-\frac{9\alpha_p\, r^2}{32b}}, \tag{3.6}$$

whereas for any $r \geq r_{\max}$,

$$\mu\left(\{x \in \mathcal{X} : f(x) - \mu(f) > r\}\right) \leq e^{-r\sqrt{\frac{3p\alpha_p}{32b(p-1)}}}. \tag{3.7}$$

Proof. The proof is adapted from the method of Aida and Stroock introduced in [1]. Assume without loss of generality that f is centered and bounded and for any $\lambda \in (0,\lambda_0)$, where

$$\lambda_0 := \sqrt{\frac{3p\alpha_p}{2b(p-1)}} \leq \frac{1}{\sqrt{a}},$$

denote once again $Z_\lambda := \int_{\mathcal{X}} e^{\lambda f}\, d\mu$ and μ_λ the probability measure with density $f_\lambda := e^{\lambda f}/Z_\lambda$ with respect to μ. We have by the Beckner-type inequality (2.3)

applied to the function $e^{\lambda f/p}$,

$$\begin{aligned}
Z_\lambda - Z_{\lambda/p}^p &\leq \frac{p}{2\alpha_p}\,\mathcal{E}_\mu(e^{\lambda f/p}, e^{\lambda(1-1/p)f}) \\
&= \frac{p}{2\alpha_p} \iint\limits_{f(x)>f(y)} e^{\lambda f(x)} \left(1 - e^{-\lambda(f(x)-f(y))/p}\right)\left(1 - e^{-\lambda(1-1/p)(f(x)-f(y))}\right) Q_x(dy)\mu(dx) \\
&\leq \frac{\lambda^2(p-1)Z_\lambda}{2p\alpha_p} \int_{\mathcal{X}} \Gamma(f,f)\, d\mu_\lambda \\
&\leq \frac{\lambda^2(p-1)Z_\lambda}{2p\alpha_p} \int_{\mathcal{X}} \left(-a\,\frac{\mathcal{L}V}{V} + b\right) d\mu_\lambda \\
&\leq \frac{\lambda^2(p-1)Z_\lambda}{2p\alpha_p}\,(aI(\mu_\lambda|\mu) + b) \\
&\leq \left(\frac{\lambda}{\lambda_0}\right)^2 Z_\lambda,
\end{aligned}$$

where we used that $f \in \mathcal{L}_V(a,b)$ and lemmas 3.1-3.5 in the three last lines. Hence rearranging the terms above and iterating the procedure yields for every $n \geq 1$,

$$Z_\lambda \leq \prod_{k=0}^{n-1} \left(\frac{\lambda_0^2}{\lambda_0^2 - \lambda^2/p^{2k}}\right)^{p^k} (Z_{\lambda/p^n})^{p^n}.$$

Since f is centered, the quantity $Z_{\lambda/p^n}^{p^n}$ goes to 1 as $n \to \infty$ and from the latter inequality we obtain after taking logarithm,

$$\begin{aligned}
\log Z_\lambda &\leq -\sum_{k=0}^{\infty} p^k \log\left(1 - \frac{(\lambda/\lambda_0)^2}{p^{2k}}\right) \\
&= \sum_{k=0}^{\infty} \frac{p^{2k+1}}{p^{2k+1}-1} \times \frac{(\lambda/\lambda_0)^{2(k+1)}}{k+1} \\
&\leq -\frac{p}{p-1} \log\left(1 - \left(\frac{\lambda}{\lambda_0}\right)^2\right).
\end{aligned}$$

In the last inequality we used the trivial inequality $p^{2k+1} \leq (\frac{p}{p-1})\,(p^{2k+1}-1)$ available for any integer k because $p \in (1,2]$. We thus obtain for any $0 < \lambda \leq \lambda_0/2$,

$$\begin{aligned}
Z_\lambda &\leq \left(1 + \frac{\lambda^2}{\lambda_0^2 - \lambda^2}\right)^{p/(p-1)} \\
&\leq \exp\left(\frac{p\lambda^2}{(p-1)(\lambda_0^2 - \lambda^2)}\right) \\
&\leq \exp\left(\frac{4p\lambda^2}{3(p-1)\lambda_0^2}\right).
\end{aligned}$$

Finally using the exponential Chebyshev inequality entails the desired result. ■

Remark 3.12. The assumption relying α_p to the parameters a and b is a technical detail but cannot be avoided. However it will be always satisfied as soon as b is taken sufficiently large (or a small enough).

Remark 3.13. As already mentioned, the Beckner-type inequality is stronger than the Poincaré inequality, i.e. $\alpha_p \leq \lambda_1$. However, theorem 3.11 does not entail a better concentration estimate than that obtained under the Poincaré inequality, except maybe when focusing on the constants depending on p (this is clearly not our interest here). The reason is due to the approach emphasized above which is exactly the same for any $p \in (1, 2]$, in contrast to theorem 3.6 where the Herbst method is used.

4. Examples

4.1. **Diffusion processes.** Let us apply now theorems 3.6 and 3.11 to diffusion processes. Below, the function U is a smooth potential such that e^{-U} is Lebesgue integrable, and denote μ the Boltzmann probability measure with density e^{-U}/Z with respect to the Lebesgue measure, where Z is the normalization factor.

The first example of interest is the so-called Kolmogorov process with generator given for any C^2 function $f : \mathbb{R}^d \to \mathbb{R}$ with bounded derivatives by

$$\mathcal{L}f = \Delta f - < \nabla U, \nabla f > .$$

One easily checks that μ is reversible for this process and the carré du champ is $\Gamma(f, f) = \|\nabla f\|^2$ where $\|\cdot\|$ stands for the Euclidean norm in $\mathbb{R}^d$. Hence by Rademacher's theorem, the spaces $\mathrm{Lip}(\mathcal{X})$ and $\mathrm{Lip}_\Gamma(\mathcal{X})$ coincide. Moreover the domain $\mathcal{D}(\mathcal{E}_\mu)$ of the Dirichlet form is $H^1(\mu)$.

4.1.1. *Ornstein-Uhlenbeck process and the standard Gaussian distribution.* Let us consider first the case of the Ornstein-Uhlenbeck process which has the standard Gaussian distribution as invariant measure. Here the potential is given by $U(x) = \|x\|^2/2$. By the famous Gross theorem [25], the pair (μ, Γ) satisfies the log-Sobolev inequality, i.e. the entropic inequality with (optimal) constant $\rho_0 = 2$. Hence theorem 3.6 will apply for observables in $\mathcal{L}_V(a, b)$ for some good test function V. For instance if $f(x) = \|x\|^2$, then choose the positive test function $V = e^{cU}$ with $c \in (0, 1)$, i.e. V is at the boundary of non-integrability. Then we have

$$-\frac{\mathcal{L}V(x)}{V(x)} \quad = \quad -cd + c(1-c)\,\|x\|^2.$$

Thus with the choice $c = 1/2$ we get $f \in \mathcal{L}_V(a, b)$ with $a = 16$ and $b = 8d$ and by theorem 3.6, for any $0 \leq r \leq 8d/3$,

$$\mu\left(\{x \in \mathbb{R}^d : \|x\|^2 > d + r\}\right) \quad \leq \quad e^{-\frac{3r^2}{64d}},$$

which is sharp up to a numerical constant since in this case $\mathrm{Var}_\mu(f) = 2d$. In the exponential regime we get for any $r \geq 8d/3$,

$$\mu\left(\{x \in \mathbb{R}^d : \|x\|^2 > d + r\}\right) \leq e^{-\frac{r}{8}}.$$

Actually, such a behaviour is expected since under μ, the variable $f(x)$ is $\chi^2(d)$-distributed, i.e. distributed as a chi-squared random variable with d degrees of freedom.

Let us come back to the example announced in remark 3.10. As observed above, our concentration result is convenient as soon as f is close to realize the equality in the Lyapunov condition (3.2). However what happens for a function g such that $\|\nabla g\| \ll \|\nabla f\|$ at infinity ? For instance if $f(x) = \|x\|^2$ as above then how concentrates the invariant measure through the observable $g(x) = \|x\|^{3/2}$? Let us investigate this point in detail now. Assume that there exists an observable g satisfying for any $\varepsilon > 0$,

$$\begin{aligned} \|\nabla g\|^2 &\leq \varepsilon \|\nabla f\|^2 + \frac{1}{\varepsilon} \\ &\leq -a_\varepsilon \frac{\mathcal{L}V}{V} + b_\varepsilon, \end{aligned}$$

where $a_\varepsilon := a\varepsilon$ and $b_\varepsilon := b\varepsilon + 1/\varepsilon$. Using the argument given in the proof of lemma 3.5, we have for any $\lambda > 0$,

$$\begin{aligned} I(\mu_\lambda|\mu) &\leq \inf_{\varepsilon>0} \frac{\lambda^2}{4} \left(a_\varepsilon I(\mu_\lambda|\mu) + b_\varepsilon\right) \\ &= \frac{\lambda^2}{2} \sqrt{aI(\mu_\lambda|\mu) + b}, \end{aligned}$$

where in the definition of $I(\mu_\lambda|\mu)$ we replaced f by g. Hence we obtain

$$I(\mu_\lambda|\mu) \leq \frac{\lambda^2}{2} \left(\frac{a\lambda^2}{2} + \sqrt{b}\right).$$

Now the same argument as in the proof of theorem 3.6 together with the latter inequality entail

$$\begin{aligned} \frac{d}{d\lambda} L_\lambda &\leq \inf_{\varepsilon>0} \frac{1}{\rho_0} \left(a_\varepsilon I(\mu_\lambda|\mu) + b_\varepsilon\right) \\ &= \frac{2}{\rho_0} \sqrt{aI(\mu_\lambda|\mu) + b} \\ &\leq \frac{2}{\rho_0} \left(\frac{a\lambda^2}{2} + \sqrt{b}\right), \end{aligned}$$

since $\sqrt{x(x+y) + y^2} \leq x + y$ for any $x, y \geq 0$. Hence we get for any $\lambda > 0$,

$$\log \int_{\mathcal{X}} e^{\lambda g}\, d\mu \leq \lambda\mu(g) + \frac{2}{\rho_0} \left(\frac{a\lambda^4}{6} + \lambda^2\sqrt{b}\right).$$

Finally, a bit more of analysis shows that the decay in the concentration estimate is of order $e^{-cr^{4/3}}$ for large deviation level r, which is of the good order of magnitude when choosing for instance $g(x)$ to be proportional to $\|x\|^{3/2}$.

Now take the observable f as a quadratic form on $\mathbb{R}^d$, i.e. there exists a positive definite symmetric matrix $A = (a_{i,j})_{i,j=1,\ldots,d}$ of size d such $f(x) =< Ax, x >$, $x \in \mathbb{R}^d$. Then in the Gaussian regime we should obtain the variance of f in the denominator,

$$\mathrm{Var}_\mu(f) \,=\, 2 \sum_{i,j=1}^{d} a_{i,j}^2.$$

However theorem 3.6 does not yield such a result, unfortunately. For instance using the same test function V as before would entail that $f \in \mathcal{L}_V(a,b)$ with $a = 16\,\|A\|_{\mathrm{op}}^2$ and $b = 8d\,\|A\|_{\mathrm{op}}^2$, where $\|A\|_{\mathrm{op}}$ is the (Euclidean) operator norm of the matrix A, i.e. its spectral radius. With this choice of parameters the inequality $\mathrm{Var}_\mu(f) \leq b/\rho_0$ is too weak to provide a reasonable variance estimate since b behaves badly in terms of dimension. To circumvent this difficulty, we choose $V = e^{cf}$ with $c = 1/(4\|A\|_{\mathrm{op}})$, so that V is positive and still integrable with respect to the Gaussian measure μ. We thus obtain that $f \in \mathcal{L}_V(a,b)$ with the same a as before, but now with the improved variance estimate

$$b = 2ac\,\mathrm{trace(A)} = 8\,\mathrm{trace(A)}\,\|\mathrm{A}\|_{\mathrm{op}},$$

where trace(A) is the trace of the matrix A. Therefore applying theorem 3.6 entails a tail estimate whose behaviour has been improved with respect to the dimension. This example emphasizes the inherent importance of the choice of the function V in the condition $\mathcal{L}_V(a,b)$. See also for instance [26, 33] for some nice studies on the concentration properties of Gaussian-like quadratic forms and Gaussian chaoses.

4.1.2. *Kolmogorov process and the Boltzmann invariant measure.* This class is a natural generalization of the Ornstein-Uhlenbeck process. To begin, assume that the measure μ is spherically log-concave, i.e. there exists a C^2 function $\phi : \mathbb{R} \to \mathbb{R}$ convex and non-decreasing such that $U(x) = \phi(\|x\|)$ for any $x \in \mathbb{R}^d$. By a famous result of Bobkov [5], the dynamics (μ, Γ) satisfy (at least) a Poincaré inequality. Let us consider the potential U as an observable and also the positive test function $V = e^{cU}$, which belongs to $L^1(\mu)$ for any $c \in (0,1)$ since ϕ is convex. Assume that there exists $r > 0$ and $M = M(r) \in (0, 1-c)$ such that

$$\frac{\Delta U(x)}{\|\nabla U(x)\|^2} \,=\, \frac{(d-1)\,\phi'(\|x\|) + \|x\|\,\phi''(\|x\|)}{\|x\|\,\phi'(\|x\|)^2} \,\leq\, M, \quad \|x\| \geq r.$$

Since we have

$$\begin{aligned} -\frac{\mathcal{L}V(x)}{V(x)} \;&=\; -c\,\Delta U(x) + c(1-c)\,\|\nabla U(x)\|^2 \\ &=\; -c(d-1)\,\frac{\phi'(\|x\|)}{\|x\|} - c\,\phi''(\|x\|) + c(1-c)\,\phi'(\|x\|)^2, \end{aligned}$$

one deduces that U belongs to the class $\mathcal{L}_V(a,b)$ with $a = 1/c(1-c-M)$, the parameter $b = b(r)$ being chosen conveniently on B_r, the centered ball of radius r in $\mathbb{R}^d$, i.e. b is the maximum between $3a\lambda_1$ and

$$\left\| \left(a\,\frac{\mathcal{L}V}{V} + \|\nabla U\|^2 \right) 1_{B_r} \right\|_{L^\infty(\mu)}.$$

As a result, one can apply theorem 3.11 to obtain a Gaussian-exponential concentration estimate through the observable U. See also the recent work of Bobkov and Madiman [8] where a somewhat similar tail estimate is established via a different approach.

Actually, spherically log-concave probability measures include the case of a potential U such that $U(x) = \|x\|^\beta$ with $\beta \geq 1$. Since the case $\beta = 2$ has already been considered, three different situations arise:

(i) the case $\beta = 1$, for which only the Poincaré inequality is satisfied.

(ii) the case $\beta \in (1,2)$: the standard Beckner inequality holds, cf. [34].

(iii) the case $\beta > 2$: using Wang's criterion [42], the log-Sobolev inequality is then verified.

In these three cases, one may choose the following parameters:

$$c := \frac{1}{2}, \quad M := \frac{1}{4}, \quad a := 8 \quad \text{and} \quad r := 4^{1/\beta} \left(\frac{d+\beta-2}{\beta} \right)^{1/\beta},$$

provided the restriction $d + \beta - 2 > 0$ holds. Finally, if $\beta > 2$ then the parameter b can be easily chosen, in contrast to the case $\beta \in [1,2)$ for which U is not C^2 at 0. Therefore, to obtain a convenient constant b in this non-smooth situation, one can choose a test function $V = e^{c\tilde{U}}$ where $\tilde{U}$ is C^2 on $\mathbb{R}^d$ and $U = \tilde{U}$ outside the ball B_r. Then the proof above remains valid with $\tilde{U}$ instead of U and an easy perturbation argument entails the standard Beckner inequality (or the Poincaré inequality in the case $\beta = 1$) for the Boltzmann probability measure defined with respect to the potential $\tilde{U}$.

One may also extend the result to the non symmetrically invariant case, when for example a logarithmic Sobolev inequality holds. Let us assume for example that the potential U is such that its Hessian matrix, denoted $\operatorname{Hess} U$, is lower bounded and that the following Lyapunov condition holds:

$$\mathcal{L}V(x) \leq \left(-c_1 \|x\|^2 + c_2 \right) V(x), \quad x \in \mathbb{R}^d,$$

where $c_1, c_2 > 0$ and V is a C^2 positive test function. Then a logarithmic Sobolev inequality holds, cf. [12], and one can apply theorem 3.6 for observables f such that the norm of $\nabla f(x)$ is at most $\|x\|$ at infinity, since in this case $f \in \mathcal{L}_V(a,b)$ with $a = 1/c_1$ and $b = c_2/c_1$. For instance, the Lyapunov condition above will be verified if at least one of the two conditions below is satisfied: there exist $\alpha \in (0,1)$

and $\beta > 0$ such that for sufficiently large x,

$$(1-\alpha)\,\|\nabla U(x)\|^2 - \Delta U(x) \;\geq\; \beta\|x\|^2 \quad \text{ or } \quad < x, \nabla U(x) > \geq \beta\|x\|^2.$$

4.1.3. *Log-Sobolev inequality for modified dynamics.* Our last example concerns the case where a log-Sobolev inequality holds for a slightly modified dynamics, but with the same Boltzmann invariant measure. In a sense, it corresponds to a modified (or weighted) log-Sobolev inequality. Consider the process with the following generator:

$$\mathcal{L}^{\sigma^2} f := \sigma^2\,\Delta f + < \nabla(\sigma^2) - \sigma^2 \nabla U, \nabla f >,$$

where σ is some measurable and locally bounded function from $\mathbb{R}^d$ to $\mathbb{R}$. Once again the measure μ is reversible for this process, but the notable difference relies on the weight σ^2 in the carré du champ, i.e. $\Gamma^{\sigma^2}(f,f) := \sigma^2\,\|\nabla f\|^2$, so that a Lipschitz function f may have an unbounded carré du champ $\Gamma^{\sigma^2}(f,f)$, in contrast to the Kolmogorov example studied above. In particular, the domain of the Dirichlet form is a weighted H^1 space, i.e.

$$\mathcal{D}(\mathcal{E}_\mu) := \left\{ f \in L^2(\mu) : \int_{\mathbb{R}^d} \sigma^2\,\|\nabla f\|^2\,d\mu < \infty \right\}.$$

We will focus mainly here on the simple case $U(x) := \|x\|^\alpha$ for $1 < \alpha < 2$, so that the standard Beckner inequality (thus the Poincaré inequality) holds for the classical dynamics (μ, Γ), but not a log-Sobolev inequality. On the one hand, according to a result of Latala and Oleszkiewicz [34], the measure μ concentrate like e^{-r^α} for large deviation level r through Lipschitz observables. On the other hand, the following weighted log-Sobolev inequality holds:

$$\mathrm{Ent}_\mu(f^2) \leq C \int_{\mathbb{R}^d} \left(1 + \|x\|^{2-\alpha}\right)\,\|\nabla f(x)\|^2\,\mu(dx),$$

where $C > 0$ is some constant depending on dimension d, cf. [13]. Letting the weight function $\sigma(x)^2 := 1 + \|x\|^{2-\alpha}$, one observes that this weighted inequality rewrites as the log-Sobolev inequality for the new dynamics (μ, Γ^{σ^2}). Choosing the positive test function $V(x) = e^{c\|x\|^\alpha}$, which belongs to $L^1(\mu)$ for any $c \in (0,1)$, we have for any x outside a neighborhood of 0,

$$\begin{aligned} -\frac{\mathcal{L}^{\sigma^2} V(x)}{V(x)} \;&=\; -c\alpha(d+\alpha-2)\,(1+\|x\|^{2-\alpha})\|x\|^{\alpha-2} + \alpha^2 c(1-c)(1+\|x\|^{2-\alpha})\|x\|^{2(\alpha-1)} \\ &\quad -c\alpha(2-\alpha), \end{aligned}$$

which behaves like $\alpha^2 c(1-c)\|x\|^\alpha$ at infinity. Hence using the same reasoning as in the case of the Kolmogorov process above, one deduces that observables f having a gradient $\|\nabla f(x)\|$ controlled by $\|x\|^{\alpha-1}$ for large x satisfy theorem 3.6 (the observable $f(x) = \|x\|^\alpha$ belongs to this class, as expected according to [34]). We point out that our results might be made more precise by following the approach provided in remark 3.10. To that aim, one has to consider the functional inequality

$I_\mu(a)$ involved in Latala and Oleszkiewicz's work [34], which is more general than the standard Beckner inequality emphasized above.

In fact using the modified dynamics (μ, Γ^{σ^2}), one can even consider interesting cases for which even the Poincaré inequality does not hold for the original dynamics (μ, Γ). For instance consider the generalized Cauchy measure μ with density proportional to $(1+\|x\|^2)^{-\beta}$, where the condition $\beta > d/2$ holds to ensure integrability at infinity. Such a measure satisfies both a weighted Poincaré inequality:

$$\mathrm{Var}_\mu(f) \leq C \int_{\mathbb{R}^d} \left(1+\|x\|^2\right) \|\nabla f(x)\|^2 \, \mu(dx), \tag{4.1}$$

with the slight restriction $\beta \geq d$, cf. [7], and also a weighted log-Sobolev inequality according to [13]:

$$\mathrm{Ent}_\mu(f^2) \leq \tilde{C} \int_{\mathbb{R}^d} \left(1+\|x\|^2\right) \log\left(e+\|x\|^2\right) \|\nabla f(x)\|^2 \, \mu(dx), \tag{4.2}$$

without further restriction on β. Here C and $\tilde{C}$ are some positive constants depending on β and d. Letting the weights

$$\sigma_1(x)^2 := 1+\|x\|^2 \quad \text{and} \quad \sigma_2(x)^2 := \left(1+\|x\|^2\right) \log\left(e+\|x\|^2\right),$$

then the weighted inequalities (4.1) and (4.2) rewrite as the Poincaré and the log-Sobolev inequality for the dynamics $(\mu, \Gamma^{\sigma_1^2})$ and $(\mu, \Gamma^{\sigma_2^2})$, respectively. Now let $V(x) = \|x\|^k$ for some $0 < k < 2\beta - d$, so that the positive test function V lies in $L^1(\mu)$. Then we have for any x outside a neighborhood of 0,

$$-\frac{\mathcal{L}^{\sigma_1^2} V(x)}{V(x)} = -k(d+k-2)\frac{1+\|x\|^2}{\|x\|^2} + 2k(\beta-1),$$

and also

$$\begin{aligned} -\frac{\mathcal{L}^{\sigma_2^2} V(x)}{V(x)} &= -k(d+k-2)\frac{1+\|x\|^2}{\|x\|^2} \log\left(e+\|x\|^2\right) - \frac{2k\left(1+\|x\|^2\right)}{e+\|x\|^2} \\ &\quad +2k(\beta-1)\log\left(e+\|x\|^2\right). \end{aligned}$$

The first quantity is constant at infinity whereas the second one is of order $k(2\beta - d - k)\log\left(e+\|x\|^2\right)$ for large $\|x\|$. Then we obtain by theorems 3.6 and 3.11 an exponential concentration estimate for large deviation level r through observables f having their gradient $\|\nabla f(x)\|$ dominated in both cases by $1/\|x\|$ for large x. Hence the previous example shows that weighted Poincaré and log-Sobolev inequalities (of course with a different weight) can lead to somewhat similar concentration estimates. Note that the function $f(x) = \log(\|x\|)$ belongs to this class of observables, leading to the well-known heavy tail phenomenon satisfied by Cauchy-type measures, cf. [7]. Finally, we mention that one can take profit of remark 3.10 to get intermediate concentration regime for observables not saturating the Lyapunov condition.

4.2. **Birth-death processes.** Let us begin the study of jump processes by considering a simple but however non trivial example, namely birth-death processes. Here $(X_t)_{t\geq 0}$ is a Markov process on the state space $\mathbb{N} := \{0, 1, 2, \ldots\}$ endowed with the classical metric $d(x, y) = |x - y|$, $x, y \in \mathbb{N}$. The transition probabilities are given by

$$\mathbb{P}_x(X_t = y) = \begin{cases} \lambda_x t + o(t) & \text{if } y = x + 1, \\ \nu_x t + o(t) & \text{if } y = x - 1, \\ 1 - (\lambda_x + \nu_x)t + o(t) & \text{if } y = x, \end{cases}$$

where $\lim_{t\to 0} t^{-1} o(t) = 0$. The transition rates λ and ν are respectively called the birth and death rates and satisfy to $\lambda > 0$ on $\mathbb{N}$ and $\nu > 0$ on $\mathbb{N}^* := \{1, 2, \ldots\}$ and $\nu_0 = 0$, so that the process is irreducible. Although we assume that the stability condition (2.7), which rewrites as

$$\lambda_x + \nu_x < \infty, \quad x \in \mathbb{N},$$

is satisfied, the generator might be unbounded in the sense of (2.9), i.e.

$$\sup_{x\in\mathbb{N}} \lambda_x + \nu_x = \infty.$$

The process is positive recurrent and non-explosive when the rates satisfy to

$$\sum_{x=1}^{\infty} \frac{\lambda_0\lambda_1 \cdots \lambda_{x-1}}{\nu_1\nu_2 \cdots \nu_x} < \infty \quad \text{and} \quad \sum_{x=1}^{\infty} \left(\frac{1}{\lambda_x} + \frac{\nu_x}{\lambda_x\lambda_{x-1}} + \cdots + \frac{\nu_x \cdots \nu_1}{\lambda_x \cdots \lambda_1\lambda_0} \right) = \infty,$$

respectively. In this case the detailed balance condition (2.8) rewrites as

$$\lambda_x \, \mu(\{x\}) = \nu_{x+1} \, \mu(\{x + 1\}), \quad x \in \mathbb{N},$$

where μ is the unique stationary distribution of the process given by

$$\mu(\{x\}) = \mu(\{0\}) \prod_{y=1}^{x} \frac{\lambda_{y-1}}{\nu_y}, \; x \in \mathbb{N}, \tag{4.3}$$

$\mu(\{0\})$ being the normalization constant. In the situations of interest, the death rate ν has to be bigger than λ to ensure such criteria.

For any function $f : \mathbb{N} \to \mathbb{R}$, the generator $\mathcal{L}$ of the process is given by

$$\mathcal{L}f(x) \;=\; \lambda_x \left(f(x + 1) - f(x)\right) + \nu_x \left(f(x - 1) - f(x)\right), \quad x \in \mathbb{N},$$

and the carré du champ is

$$\Gamma(f, f)(x) \;=\; \frac{1}{2} \left\{ \lambda_x \left(f(x + 1) - f(x)\right)^2 + \nu_x \left(f(x - 1) - f(x)\right)^2 \right\}, \quad x \in \mathbb{N}.$$

In particular, the Dirichlet form is given by

$$\mathcal{E}_\mu(f, g) := \sum_{x\in\mathbb{N}} \lambda_x \left(f(x + 1) - f(x)\right) \left(g(x + 1) - g(x)\right) \mu(\{x\}),$$

where f, g belong to the space $\mathcal{D}(\mathcal{E}_\mu)$ of functions $u : \mathbb{N} \to \mathbb{R}$ such that $\mathcal{E}_\mu(u, u)$ is finite.

On the one hand Joulin [30] gives, under some convenient ergodic assumptions, concentration estimates of Poisson-type through observables belonging to the space $\mathrm{Lip}(\mathbb{N})$. In particular, his proof requires the inclusion $\mathrm{Lip}(\mathbb{N}) \subset \mathrm{Lip}_\Gamma(\mathbb{N})$, enforcing the rates λ and ν to be bounded. On the other hand, when we apply Ollivier's result [38] to birth-death processes, his Gaussian-exponential concentration property is available for observables $f \in \mathrm{Lip}(\mathbb{N})$ such that $\Gamma(f,f) \in \mathrm{Lip}(\mathbb{N})$. It induces that $\lambda, \nu \in \mathrm{Lip}(\mathbb{N})$, extending Joulin's result from bounded to (at most) linear rates. As announced in the introduction through the study of a specific example, theorems 3.6 and 3.11 entail Gaussian-exponential concentration estimates beyond these cases since the carré du champ $\Gamma(f,f)$ is allowed to have a growth comparable to that of ν.

First let us provide some basic conditions which ensure an entropic or Poincaré inequality. The following necessary (but not sufficient) condition is due to Caputo, DaiPra and Posta [10] and has been recently recovered by Chafaï and Joulin [16] by using a semigroup approach: if λ is non-increasing and ν is non-decreasing and there exists $\alpha > 0$ such that

$$\inf_{x\in\mathbb{N}} \lambda_x - \lambda_{x+1} + \nu_{x+1} - \nu_x \geq \alpha, \tag{4.4}$$

then the entropic inequality (2.2) is satisfied with constant α, or equivalently $\rho_0 \geq \alpha$. Such an assumption exhibits very asymmetric rates. More precisely, it enforces the rates λ and ν to be bounded and super-linear, respectively, excluding some interesting cases which can be however considered for the Poincaré inequality. Indeed, Miclo [37] states that the spectral gap λ_1 is positive if and only if

$$\delta := \sup_{x\geq 1} \sum_{k=0}^{x-1} \frac{1}{\lambda_k \mu(\{k\})} \sum_{l\geq x} \mu(\{l\}) \quad < \quad \infty, \tag{4.5}$$

and in this case we have $1/\delta \geq \lambda_1 \geq 1/4\delta$, i.e. λ_1 is of order $1/\delta$. In particular, in contrast to the entropic inequality, one may find examples satisfying the Poincaré inequality with an unbounded birth rate λ. Now assume that the positive test function $V(x) := \kappa^x$ is in $L^1(\mu)$ for some constant $\kappa > 1$ depending on λ, ν. Then an observable f belongs to $\mathcal{L}_V(a,b)$ if and only if

$$\Gamma(f,f) \quad \leq \quad \frac{a(\kappa-1)}{\kappa}\,(\nu - \kappa\,\lambda) + b,$$

showing that on a large scale the behaviour of $\Gamma(f,f)$ is controlled by the growth of the death rate ν. To compare with the results of Joulin and Ollivier mentioned previously, assume that the observable $f \in \mathrm{Lip}(\mathbb{N})$. Then two extreme situations may appear when the death rate ν is unbounded:

(i) a small birth rate λ, i.e. λ is bounded. In this case one may choose the following parameters to ensure that $f \in \mathcal{L}_V(a,b)$:

$$a := \frac{\kappa}{2(\kappa-1)} \quad \text{and} \quad b := \frac{(1+\kappa)\|\lambda\|_{L^\infty(\mu)}}{2}.$$

(ii) a birth rate λ of the order of ν. Let $x_0 \in \mathbb{N}^*$ and assume that $\lambda_x \leq c\nu_x$ for all $x \geq x_0$, where $c \in (0,1)$ is some parameter. Then in order to get $f \in \mathcal{L}_V(a,b)$, one can choose for $\kappa \in (1, 1/c)$,

$$a := \frac{\kappa(1+c)}{2(1-c\kappa)(\kappa-1)},$$

$$\text{and} \qquad b := \left\| \left(\frac{\lambda+\nu}{2} + a\,\frac{\mathcal{L}V}{V} \right) 1_{[0,x_0]} \right\|_{L^\infty(\mu)} = \frac{1+\kappa}{2(1-c\kappa)} \left\| (\lambda - c\nu)\, 1_{[0,x_0]} \right\|_{L^\infty(\mu)}.$$

For instance the choice of $\kappa := 1/\sqrt{c}$ entails the integrability of V and then we obtain

$$a = \frac{1+\sqrt{c}}{2(1-\sqrt{c})} \qquad \text{and} \qquad b = \frac{1+\sqrt{c}}{2\sqrt{c}(1-\sqrt{c})} \left\| (\lambda - c\nu)\, 1_{[0,x_0]} \right\|_{L^\infty(\mu)}.$$

Of course b has also to be at least $3a\lambda_1$ if only the Poincaré inequality is satisfied. In both cases (i) and (ii) there exist plenty of examples satisfying Poincaré inequality and thus theorem 3.11, whereas only the case (i) may satisfy the entropic inequality and so theorem 3.6. In particular, the example emphasized in the introduction is a prototype of such a situation. Indeed denote μ_p the geometric distribution on $\mathbb{N}$ of parameter $p \in (0,1)$, i.e. $\mu_p(\{x\}) := (1-p)p^x$, $x \in \mathbb{N}$, and let us consider the carré du champ

$$\Gamma^{(n)}(f,f)(x) := \frac{1}{2} \left\{ p(x+1)^n \left(f(x+1) - f(x)\right)^2 + x^n 1_{\{x \neq 0\}} \left(f(x-1) - f(x)\right)^2 \right\},$$

where $x \in \mathbb{N}$ and $n \in \mathbb{N}$ is some fixed parameter. The measure μ_p is reversible with respect to these dynamics and by comparing the underlying Dirichlet form with that given in the case $n = 0$ (for which $\lambda_1^{(1)} = (1-\sqrt{p})^2$, cf. [17, 29]), the dynamics $(\mu_p, \Gamma^{(n)})$ satisfies a Poincaré inequality with optimal constant $\lambda_1^{(n)} \geq (1-\sqrt{p})^2$ (Miclo's result (4.5) would only entail $\lambda_1^{(n)} \geq (1-p)^2/4$). Hence for sufficiently large integer x_0 one can pick $c = (1+p)^2/4 \in (0,1)$, $\kappa = 2/(1+p)$ and for some integer $x_1 \in (0, x_0)$,

$$a = \frac{3+p}{2(1-p)} \qquad \text{and} \qquad b = \max\left\{ \frac{3+p}{4(1-p)^2} \left| 4p(x_1+1)^n - (1+p)^2\, x_1^n \right|, 3a\lambda_1^{(n)} \right\}.$$

To achieve the birth-death example, let us focus our attention on a model which mimics the diffusion case, namely ultra log-concave distributions on $\mathbb{N}$, see for instance [28, 10]. We say that a probability measure μ on $\mathbb{N}$ is ultra log-concave (resp. log-concave) if it satisfies for any $x \in \mathbb{N}^*$,

$$x\,\mu(\{x\})^2 \geq (x+1)\,\mu(\{x+1\})\,\mu(\{x-1\}) \quad (\text{resp.} \quad \mu(\{x\})^2 \geq \mu(\{x+1\})\,\mu(\{x-1\})).$$

For instance the Poisson distribution is ultra log concave whereas the geometric measure is only log-concave. Assume that the measure μ has density e^{-U}/Z with

respect to the counting measure on $\mathbb{N}$, where U is some nice function and Z is the normalization constant. Denote ΔU the discrete Laplacian of the potential U, i.e.

$$\Delta U(x) := U(x+1) - 2U(x) + U(x-1), \quad x \in \mathbb{N}^*.$$

Then μ is ultra log-concave (resp. log-concave) if and only if $\Delta U(x) \geq \log(1+1/x)$ for any integer $x \in \mathbb{N}^*$ (resp. ΔU is non-negative).

From a dynamical point of view, measure μ is the stationary distribution of the birth-death process with rates

$$\lambda_x = 1 \quad \text{and} \quad \nu_x = e^{U(x)-U(x-1)}\, 1_{\{x\neq 0\}}, \quad x \in \mathbb{N}.$$

Then under the ultra log-concavity assumption, we have for any integer $x \geq 2$,

$$\begin{aligned}
\lambda_x - \lambda_{x+1} + \nu_{x+1} - \nu_x &= \left(e^{\Delta U(x)} - 1\right) e^{\sum_{k=1}^{x-1} \Delta U(k) + U(1) - U(0)} \\
&\geq \prod_{k=1}^{x-1} \left(1 + \frac{1}{k}\right) \frac{e^{U(1)-U(0)}}{x} \\
&= e^{U(1)-U(0)} \\
&= \nu_1,
\end{aligned}$$

so that (4.4) is satisfied with $\alpha = \nu_1$ (the cases $x \in \{0,1\}$ being straightforward). Thus the entropic inequality holds with constant $\rho_0 \geq \nu_1$. In particular, the super-linearity of the death rate ν entails that the potential U has a growth at infinity at least $x\log(x)$, showing that the tail behaviour of μ can be compared to that of a Poisson distribution. Finally, note that the log-concavity assumption only is not sufficient to ensure an entropic inequality since one obtains in this case

$$\inf_{x\in\mathbb{N}} \lambda_x - \lambda_{x+1} + \nu_{x+1} - \nu_x \geq 0.$$

However, as in the diffusion case, one may find examples of log-concave distributions on $\mathbb{N}$ satisfying the Poincaré inequality by using Miclo's condition (4.5), which simply rewrites as

$$\sup_{x\in\mathbb{N}^*} \sum_{0\leq k\leq x-1<l} e^{U(k)-U(l)} < \infty.$$

4.3. **Glauber dynamics for unbounded particles.** We consider the situation where $\mathcal{X}$ is the unbounded configuration space $\mathbb{N}^\Lambda$, where Λ is a bounded subset of $\mathbb{Z}^d$. For each site $x \in \Lambda$, denote η_x the number of particles located at x. Given a bounded function $\lambda : \mathbb{Z}^d \to [0,\infty)$, let π be the Poisson measure on $\mathbb{N}^\Lambda$ with parameter λ, that is to say

$$\pi(\{\eta\}) = \prod_{x\in\Lambda} e^{-\lambda(x)} \frac{\lambda(x)^{\eta_x}}{\eta_x!}, \quad \eta \in \mathbb{N}^\Lambda.$$

We equip $\mathbb{N}^\Lambda$ with the total variation distance which counts the number of different particles. In other words, if η and $\bar{\eta}$ are two configurations in $\mathbb{N}^\Lambda$, then the total

variation distance is given by

$$d(\eta,\bar{\eta}) \;:=\; \sum_{x\in\Lambda} |\eta_x - \bar{\eta}_x|.$$

Our definition is a straightforward generalization of the classical notion of total variation distance between probability measures, since it coincides with the usual definition when the configurations are normalized by their total masses. For any $f:\mathbb{N}^\Lambda \to \mathbb{R}$, the discrete gradient operators are defined by

$$D_x^+ f(\eta) \;:=\; f(\eta+\delta_x) - f(\eta), \quad D_x^- f(\eta) \;:=\; f(\eta-\delta_x) - f(\eta), \quad \eta\in\mathbb{N}^\Lambda,$$

where δ_x is the Dirac mass at point $x\in\Lambda$ and by convention $D_x^- f(\emptyset) := 0$. Note that by [20], a given function f belongs to the space $\mathrm{Lip}(\mathbb{N}^\Lambda)$ if and only if

$$\sup_{(\eta,x)\in\mathbb{N}^\Lambda\times\Lambda} |D_x^+ f(\eta)| < \infty.$$

Now let $\phi:\mathbb{Z}^d\to[0,\infty)$ be an even function, null at the origin and summable on $\mathbb{Z}^d$, i.e. satisfying $\sum_{x\in\mathbb{Z}^d}\phi(x)<\infty$. We define the Hamiltonian $H:\mathbb{N}^\Lambda\to\mathbb{R}_+$ as

$$H(\eta) \;:=\; \frac{1}{2}\sum_{x,y\in\Lambda} \phi(x-y)\,\eta_x\,\eta_y.$$

Then the Gibbs measure μ at inverse temperature $\beta>0$ is the probability measure on $\mathbb{N}^\Lambda$ given by

$$\mu(\{\eta\}) \;=\; \frac{1}{Z}\,e^{-\beta\,H(\eta)}\,\pi(\{\eta\}),$$

where Z is the normalization constant. As observed below, our study is based on the configuration space $\mathbb{N}^\Lambda$ since our model exhibits free boundary condition, that is to say Λ is, in some sense, disconnected from the lattice $\mathbb{Z}^d$. However the aforementioned model might be extended outside Λ by introducing an appropriate boundary condition.

Now, let us introduce the Glauber dynamics associated to the Gibbs measure above, which can be seen as a spatial birth-death process, cf. [40]. If η is the configuration of the system at time t, then a particle appears or disappears at site $x\in\Lambda$ with rates $\lambda(x)\,e^{-\beta D_x^+ H(\eta)}dt$ and dt, respectively. In particular, the case $H=0$ corresponds to the non-interacting case. The generator $\mathcal{L}$ is thus of birth-death type and defined for any function $f:\mathbb{N}^\Lambda\to\mathbb{R}$ by

$$\mathcal{L}(\eta) \;:=\; \sum_{x\in\Lambda}\Big(c^-(\eta,x)\,D_x^- f(\eta) + c^+(\eta,x)\,D_x^+ f(\eta)\Big), \quad \eta\in\mathbb{N}^\Lambda,$$

where the rates of the dynamics c^+ and c^- are given by

$$\begin{cases} c^+(\eta,x) \;=\; \lambda(x)\,e^{-\beta D_x^+ H(\eta)} \;=\; \lambda(x)\,e^{-\beta\sum_{y\in\Lambda}\phi(x-y)\,\eta_y}; \\ c^-(\eta,x) \;=\; \eta_x. \end{cases}$$

In particular, the stability condition (2.7) is clearly satisfied since Λ is finite and moreover, according to the detailed balance condition (2.8) which in our context rewrites as

$$c^{\pm}(\eta,x)\,\mu(\{\eta\}) \;=\; c^{\mp}(\eta\pm\delta_x,x)\,\mu(\{\eta\pm\delta_x\}),\quad \eta_x>0,\quad (\eta,x)\in\mathbb{N}^{\Lambda}\times\Lambda,$$

the Gibbs measure μ is reversible for these dynamics. Finally, the carré du champ of an observable f is given by

$$\Gamma(f,f)(\eta) \;=\; \frac{1}{2}\sum_{x\in\Lambda}\left\{c^{-}(\eta,x)\,|D_x^{-}f(\eta)|^2+c^{+}(\eta,x)\,|D_x^{+}f(\eta)|^2\right\},\quad \eta\in\mathbb{N}^{\Lambda}.$$

Recently, the problem of finding the speed of convergence to equilibrium of this model has been addressed in several articles, cf. for instance [4] or [44] for a spectral method (i.e. related to Poincaré inequality) in the continuum $\mathbb{R}^d$, and also [18] for an approach through the entropic inequality. In all these papers, the objective is to find constants which are independent of Λ and of the boundary condition. In a recent work [19], Dai Pra and Posta established the entropic inequality with constant $\rho_0\geq 1-\|\lambda\|_\infty\,\varepsilon(\beta)$, under the following Dobrushin-type uniqueness condition:

$$\varepsilon(\beta)\;:=\;\sum_{x\in\mathbb{Z}^d}\left(1-e^{-\beta\phi(x)}\right)\;<\;\frac{1}{\|\lambda\|_\infty}. \tag{4.6}$$

Here $\|\cdot\|_\infty$ denotes the supremum norm of a bounded function on $\mathbb{Z}^d$. Note that assumption (4.6) will be verified as soon as β is small enough, i.e. the temperature of the system is sufficiently high. Therefore, if we choose for some $\kappa>1$ the test function $V(\eta):=\kappa^{\sum_{x\in\Lambda}\eta_x}$ which is in $L^1(\mu)$, then an observable f belongs to the class $\mathcal{L}_V(a,b)$ if and only if

$$\Gamma(f,f)(\eta)\;\leq\;\frac{a(\kappa-1)}{\kappa}\sum_{x\in\Lambda}\left(\eta_x-\kappa\,\lambda(x)\,e^{-\beta D_x^{+}H(\eta)}\right)+b,\quad \eta\in\mathbb{N}^{\Lambda},$$

as in the context of birth-death processes above. Thus the Gaussian-exponential concentration estimate of theorem 3.6 applies under these observables. Finally, we have $D_x^{+}H(\eta)\geq 0$ because ϕ is non-negative and if $f\in\mathrm{Lip}(\mathbb{N}^{\Lambda})$ then one may choose

$$a:=\frac{\kappa}{2(\kappa-1)}\quad\text{and}\quad b:=\frac{(1+\kappa)}{2}\sum_{x\in\Lambda}\lambda(x).$$

In particular if λ is assumed to be summable on $\mathbb{Z}^d$, then b no longer depends on the box Λ.

Acknowledgments. The authors are grateful to Paolo Dai Pra and G. Posta for useful discussions about the statistical mechanics part. They thank also the ANR Project EVOL for financial support.

References

[1] S. Aida and D. Stroock. Moment estimates derived from Poincaré and logarithmic Sobolev inequalities. *Math. Res. Lett.*, 1:75-86, 1994.

[2] F. Barthe, D. Bakry, P. Cattiaux and A. Guillin. Poincaré inequalities for log-concave probability measures: a Lyapunov function approach. *Elec. Comm. Probab.*, 13:60-66, 2008.

[3] W. Beckner. A generalized Poincaré inequality for Gaussian measures. *Proc. of the AMS*, 105:397-400, 1989.

[4] L. Bertini, N. Cancrini and F. Cesi. The spectral gap for a Glauber-type dynamics in a continuous gas. *Ann. Inst. H. Poincaré Probab. Statist.*, 38:91-108, 2002.

[5] S.G. Bobkov. Spectral gap and concentration for some spherically symmetric probability measures. *Geometric aspects of functional analysis*, 37-43, Lecture Notes in Math., 1807, Springer, Berlin, 2003.

[6] S.G. Bobkov and M. Ledoux. On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *J. Funct. Anal.*, 156:347-365, 1998.

[7] S.G. Bobkov and M. Ledoux. Weighted Poincaré-type inequalities for Cauchy and other convex measures. *Ann. Probab.*, 37(2):403-427, 2009.

[8] S.G. Bobkov and M. Madiman. Concentration of the information in data with log-concave distributions. *Ann. Probab.*, 39: 1528-1543, 2011.

[9] S.G. Bobkov and P. Tetali. Modified logarithmic Sobolev inequalities in discrete settings. *J. Theor. Probab.*, 19:289-336, 2006.

[10] P. Caputo, P. Dai Pra and G. Posta. Convex entropy decay via the Bochner-Bakry-Émery approach. *Ann. Inst. H. Poincaré Probab. Statist.*, 45, 734-753, 2009.

[11] P. Cattiaux and A. Guillin. Deviation bounds for additive functionals of Markov process. *ESAIM Probab. Statist.* 12:12-29, 2008.

[12] P. Cattiaux, A. Guillin and L. Wu. A note on Talagrand transportation inequality and logarithmic Sobolev inequality. *Probab. Theory Related Fields*, 148(1-2):285-304, 2010.

[13] P. Cattiaux, A. Guillin and L. Wu. Some remarks on weighted logarithmic Sobolev inequality. To appear in *Indiana University Math. J.*, 2011.

[14] P. Cattiaux, A. Guillin and P.A. Zitt. Poincaré inequalities and hitting times. To appear in *Ann. Inst. H. Poincaré Probab. Statist.*, 2012.

[15] D. Chafaï. Entropies, convexity, and functional inequalities. *J. Math. Kyoto Univ.*, 44:325-363, 2004.

[16] D. Chafaï and A. Joulin. Intertwining and commutation relations for birth-death processes. Preprint 2011. Available at http://arxiv.org/abs/1011.2331.

[17] M. F. Chen. Estimation of spectral gap for Markov chains. *Acta Math. Sin.*, 12:337- 360, 1996.

[18] P. Dai Pra, A.M. Paganoni and G. Posta. Entropy inequalities for unbounded spin systems. *Ann. Probab.*, 30:1969-1976, 2002.

[19] P. Dai Pra and G. Posta. Personal communication, september 2011.

[20] L. Decreusefond, A. Joulin and N. Savy. Upper bounds on Rubinstein distances on configuration spaces and applications. *Comm. Stoch. Anal.*, 4:377-399, 2010.

[21] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.*, 6:695-750, 1996.

[22] F. Gao, A. Guillin and L. Wu. Bernstein type's concentration inequalities for symmetric Markov processes. To appear in *SIAM Theor. Probab. Appl.* 2012.

[23] A. Guillin, C. Léonard, L. Wu and N. Yao. Transportation inequalities for Markov processes. *Probab. Theory Related Fields*, 144 (3-4):669-695, 2009.

[24] N. Gozlan and C. Léonard. Transport inequalities. A survey. *Markov Process. Related Fields*, 16:635-736, 2010.
[25] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97:1061-1083, 1975.
[26] D.L. Hanson and F.T. Wright. A bound on tail probabilities for quadratic forms in independent random variables. *Ann. Math. Statist.*, 42:1079-1083, 1971.
[27] C. Houdré. Remarks on deviation inequalities for functions of infinitely divisible random vectors. *Ann. Probab.*, 30(3):1223-1237, 2002.
[28] O. Johnson. Log-concavity and the maximum entropy property of the Poisson distribution. *Stochastic Process. Appl.*, 117:791-802, 2007.
[29] A. Joulin and N. Privault. Functional inequalities for discrete gradients and applications to the geometric distribution. *ESAIM Probab. Statist.*, 8:87-101, 2004.
[30] A. Joulin. Poisson-type deviation inequalities for curved continuous-time Markov chains. *Bernoulli*, 13:782-798, 2007.
[31] A. Joulin. A new Poisson-type deviation inequality for Markov jump processes with positive Wasserstein curvature. *Bernoulli*, 15(2):532-549, 2009.
[32] A. Joulin and Y. Ollivier. Curvature, concentration, and error estimates for Markov chain Monte Carlo. *Ann. Probab.*, 38(6): 2418-2442, 2010.
[33] R. Latala. Estimates of moments and tails of Gaussian chaoses. *Ann. Probab.*, 34:2315-2331, 2006.
[34] R. Latala and K. Oleszkiewicz. Between Sobolev and Poincaré. *Geometric aspects of functional analysis*, 147-168, Lecture Notes in Math., 1745, Springer, Berlin, 2000.
[35] M. Ledoux. *The concentration of measure phenomenon.* Mathematical Surveys and Monographs, 89. American Mathematical Society, Providence, RI, 2001.
[36] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. *Séminaire de Probabilités*, XXXIII, 120-216, Lecture Notes in Math., 1709, Springer, Berlin, 1999.
[37] L. Miclo. An example of application of discrete Hardy' s inequalities. *Markov Process. Related Fields*, 5: 319-330, 1999.
[38] Y. Ollivier. Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.*, 256:810-864, 2009.
[39] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173: 361-400, 2000.
[40] C. Preston. Spatial birth-and-death processes. Proceedings of the 40th session of the international statistical institute (Warsaw 1975). *Bull. Inst. Internat. Statist.*, 46:371-391, 1975.
[41] M.D. Sammer. Aspects of mass transportation in discrete concentration inequalities. Ph.D. Thesis, Georgia Institute of Technology. Available at http://smartech.gatech.edu/dspace/handle/1853/ 7006.
[42] F.Y. Wang. Logarithmic Sobolev inequalities on noncompact Riemannian manifolds. *Probab. Theory Related Fields*, 109:417-424, 1997.
[43] L. Wu. Uniformly integrable operators and large deviations for Markov processes. *J. Funct. Anal.*, 172:301-376, 2000.
[44] L. Wu. Estimate of spectral gap for continuous gas. *Ann. I. H. Poincaré*, 40:387-409, 2004.

(A. Guillin) UMR CNRS 6620, Université Blaise Pascal, Clermont-Ferrand II and Institut Universitaire de France (IUF), France
E-mail address: mailto:guillin(at)math.univ-bpclermont.fr
URL: http://math.univ-bpclermont.fr/~guillin/

(A. Joulin, corresponding author) UMR CNRS 5219, Institut de Mathématiques de Toulouse, Université de Toulouse, France
E-mail address: mailto:ajoulin(at)insa-toulouse.fr
URL: http://www-gmm.insa-toulouse.fr/~ajoulin/