Combinatorics, Probability and Computing (19XX) 00, 000–000. Printed in the United Kingdom

Path method for the logarithmic Sobolev constant

Cyril ROBERTO

roberto@cict.fr
Université Paul Sabatier,
L.S.P. and C.N.R.S
118 route de Narbonnes
31062 Toulouse, France.

7 March 2001.

abstract:

This paper is concerned with path techniques for quantitative analysis of the logarithmic Sobolev constant on a countable set. We present new upper bounds of the logarithmic Sobolev constant that generalize those given by Sinclair in [18] in the case of the spectral gap constant involving path combinatorics. Some examples of applications are given. We compare our bounds to the Hardy constant in the particular case of birth and death processes.

1. Introduction

Let $\mathcal{X}$ be a countable set and $K(x,y)$ a transition matrix for an *irreducible* chain. This means that for all $x, y \in \mathcal{X} \times \mathcal{X}$, $K(x,y) \geq 0$ and for $x \neq y$, there exists a sequence $x_0, \ldots, x_k$ such that $K(x_i, x_{i+1}) > 0$, $i = 0, \ldots, k-1$, $x_0 = x$ and $x_k = y$. We assume throughout that K is *reversible* with respect to the probability measure μ, that is, for all $x, y \in \mathcal{X} \times \mathcal{X}$, it satisfies the *detailed balance condition*

$$\mu(x)K(x,y) = \mu(y)K(y,x) .$$

Define the matrix L as $L(x,y) = K(x,y)$ if $x \neq y$ and $L(x,x) = -\sum_{y \in \mathcal{X}} K(x,y)$. Then, L is the generator of a process (in continuous time) that acts in particular on the functions f null excepted on a finite numbers of points in $\mathcal{X}$ as

$$L(f)(x) = \sum_{y \in \mathcal{X}} K(x,y)(f(y) - f(x)), \quad \text{for all} \quad x \in \mathcal{X} .$$

By irreductibility, the law of the position of the process converges to the equilibrium state μ as t goes to infinity.

By irredutibility, μ charges all the points, *i.e.* for all $x \in \mathcal{X}$, $\mu(x) > 0$.

Now, for any function $f \in \mathbf{L^2}(\mu)$, we define the DIRICHLET *form* of f associated to K and μ by

$$\mathcal{E}_{\mu,K}(f,f) := \frac{1}{2} \sum_{x,y \in \mathcal{X} \times \mathcal{X}} \mu(x) K(x,y) (f(y) - f(x))^2 \ .$$

We also define the *variance* of f,

$$\mathbf{Var}_\mu(f) = \mu((f - \mu(f))^2) \ ,$$

and the *entropy* of f^2,

$$\mathbf{Ent}_\mu(f^2) := \sum_{x \in \mathcal{X}} f^2(x) \ln \frac{f^2(x)}{\sum_{x \in \mathcal{X}} f^2(x) \mu(x)} \mu(x)$$

In order to study the behavior of the process, functional inequalities play a crucial role. We have in particular in mind the POINCARÉ and the logarithmic SOBOLEV inequalities: we say that μ satisfies a POINCARÉ *inequality* (or *spectral gap inequality*) if there exists a constant $\lambda > 0$ such that for all $f \in \mathbf{L^2}(\mu)$,

$$\mathbf{Var}_\mu(f) \leq \lambda \mathcal{E}_{\mu,K}(f,f) \ . \tag{1}$$

We say that μ satisfies a *logarithmic* SOBOLEV *inequality* if there exists a constant $\alpha > 0$ such that for all $f \in \mathbf{L^2}(\mu)$,

$$\mathbf{Ent}_\mu(f^2) \leq \alpha \mathcal{E}_{\mu,K}(f,f) \ . \tag{2}$$

The best ergodic constants λ^{-1} and α such that (1) and (2) hold are called respectively the spectral gap and the logarithmic SOBOLEV constant of μ. Both are of interest. They give in particular the speed of convergence of the chain to equilibrium and thus answer the question: starting from an arbitrary distribution, how many time must we wait to be near the equilibrium? We refer the reader principally to [5], [4], [17] and the references therein for details on this convergence. There are also applications to theorical computation science, see for instance [19] or [18].

In our context, we mention that it is not interesting to consider a MARKOV kernel, *i.e.* satisfying for all $x \in \mathcal{X}$, $\sum_{y \in \mathcal{X}} K(x,y) = 1$. First because the logarithmic SOBOLEV inequality and the POINCARÉ inequality considered above correspond to continuous time and not to discrete time (for discrete time, we would have considered the multiplicative symmetrized kernel KK^* instead of K). The second reason is that a bounded kernel is often (and always on an infinite set as we will see after) an obstacle to the existence of the logarithmic SOBOLEV constant.

The spectral gap and the logarithmic SOBOLEV constants give different bounds on the speed of convergence and it is often convenient to control both of them. To study the spectral gap, a lot of techniques can be used, from analytic tools to geometric tools like the CHEEGER inequality (see [3] and [9]), or also paths combinatorics. Those one have been introduced by JERRUM and SINCLAIR [7] in theorical computation science in their study of a stochastic algorithm that counts perfect matchings in a graph. One of the initial idea was to control the CHEEGER constant. The method was further developed by DIACONIS and STROOCK [5], FILL [6] and SINCLAIR [18] on finite sets. ROSENTHAL generalized them to countable sets (and even continuous sets) in [16]. In [11] the authors adapt the paths techniques on the inequalities of MATHIEU (see [10]).

This present paper develops the same kind of paths combinatorics to bound the logarithmic SOBOLEV constant. To that aim, we make use of an interesting idea of BOBKOV and GÖTZE [2] to reduce the study of the logarithmic SOBOLEV inequality (2) to an inequality of type POINCARÉ that allows us to generalize SINCLAIR's bounds.

We now present one of our results to illustrate our purpose. Look at $\mathcal{X}$ as a graph whose vertices are the points of $\mathcal{X}$ and edges are all couples $e = (x, y) \in \mathcal{X} \times \mathcal{X}$ such that $Q(e) = \mu(x)K(x, y) > 0$. Set $\mathbf{E}$ for the set of all edges. Then, by irreductibility, we can construct a path γ_{xy} from any x to any y, that is, a sequence of vertices $\gamma = (x_0, \dots, x_k)$ such that (x_i, x_{i+1}) is an edge, $i = 1 \dots k-1$, $x_0 = x$ and $x_k = y$. The length of such a path γ is $|\gamma| = k$. Then, we have the following result.

Theorem. *Let K be an irreductible chain with reversible probability measure μ on a countable set $\mathcal{X}$. For all $(x, y) \in \mathcal{X} \times \mathcal{X}$, $x \neq y$, choose one path γ_{xy} without repeated edges (i.e. for all $i \neq j$, $e_i \neq e_j$). Assume that for all $x \in \mathcal{X}$, $\mu(x) \leq 1/2$. Then, the logarithmic* SOBOLEV *constant α defined in (2) satisfies $\alpha \leq 20A$, where*

$$A := \sup_{e \in \mathbf{E}} \left\{ \frac{1}{Q(e)} \sum_{x,y : \gamma_{xy} \ni e} |\gamma_{xy}| \mu(x)\mu(y) \ln \frac{1}{\mu(y)} \right\} .$$

In the sequel, we will often write $\sup_{e \in \mathbf{E}}$ even when $|\mathbf{E}| < \infty$. The latter theorem corresponds to Theorem 5 of [18].

Next we extend the constant A through family of weight functions. A weight function is simply a positive function on the set of edges, $w : \mathbf{E} \to (0, \infty)$. The bound of Theorem 1 may then be improved (see Theorem 3.2 below) replacing A by

$$A^w := \sup_{e \in \mathbf{E}} \left\{ \frac{w(e)}{Q(e)} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}|_w \mu(x)\mu(y) \ln \frac{1}{\mu(y)} \right\}$$

where $|\gamma_{xy}|_w = \sum_{e \in \gamma_{xy}} 1/w(e)$. Note that if $w \equiv 1$ then $A^w = A$. One of the advantages of A and A^w is that they are easily computable. Several examples will be discussed.

The last generalization uses the notion of flow introduced by SINCLAIR in [18], see Theorem 3.3 at the end of section 3.

We finish our discussion by some comparison result between path techniques and HARDY's inequalities in the particular case of a birth and death process on $\mathbf{N}$. We exhibit a special weight function such that the SINCLAIR constant

$$A^w_{S.G} := \sup_{e \in \mathbf{E}} \left\{ \frac{w(e)}{Q(e)} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}|_w \mu(x)\mu(y) \right\}$$

corresponding to A^w introduced above is optimal with respect to the spectral gap. More precisely, we prove that for this special weight function w,

$$\frac{A^w_{S.G}}{8} \leq \lambda \leq A^w_{S.G} .$$

The same weight function gives also a result on the logarithmic SOBOLEV constant, but (unfortunately) not so sharp.

We must notice that on an infinite set $\mathcal{X}$, if the chain is given by a bounded transition matrix K, *i.e.* if there exists a constant, say 1, such that for all $x \in \mathcal{X}$, $\sum_{y \in \mathcal{X}} K(x,y) \leq 1$, then, the logarithmic SOBOLEV constant α is certainly infinite (see [12]). To see that claim, it is enough to consider a sequence of test functions $\mathbb{1}_{A_n}$ in the logarithmic SOBOLEV inequality (2) with $\mu(A_n) \to 0$ as n goes to infinity. This remark shows that to study the logarithmic SOBOLEV constant on an infinite set, we must consider an unbounded transition matrix. All our examples on infinite sets are built at the light of this remark.

At last, we mention that the results of this work hold true with a transition matrix for an irreductible chain with *invariant* (not necessarily reversible) measure μ. In that case, it is enough to consider $\frac{1}{2}(\mu(x)K(x,y) + \mu(y)K(y,x))$ instead of $\mu(x)K(x,y)$ in all what follow.

2. Logarithmic Sobolev inequality as Poincaré type inequality

This section reduces the study of the logarithmic SOBOLEV inequality to the study of an inequality of POINCARÉ type. This reduction is the starting point of our geometric approach of the logarithmic SOBOLEV constant in term of path combinatorics. The main idea comes from [2].

Let K be an irreductible chain with reversible probability measure μ on a countable set $\mathcal{X}$. In [2], the authors introduce the quantity $\mathcal{L}(f) := \sup_{t \in \mathbf{R}} \mathbf{Ent}_\mu(\mathbf{f}+\mathbf{t})^2$ for all functions $f \in \mathbf{L}^2(\mu)$. Then, by translation invariance of the DIRICHLET form, it is equivalent to consider, for all $f \in \mathbf{L}^2(\mu)$, the logarithmic SOBOLEV inequality (2)

$$\mathbf{Ent}_\mu(f^2) \leq \alpha \mathcal{E}_{\mu,K}(f,f)$$

or the inequality

$$\mathcal{L}(f) \leq \alpha \mathcal{E}_{\mu,K}(f,f) \ .$$

One of the features of the previous elementary remark is that for all functions $f \in \mathbf{L}^2(\mu)$, the quantity $\mathcal{L}(f)$ is linked to an ORLICZ norm of f. Let us present this link after some notations and useful facts on ORLICZ norms.

Let (Ω, μ) be a probability space and $\Theta : \mathbf{R} \to [0,\infty)$ a Young function, that is an even convex function with $\Theta(x) > 0$ for all $x > 0$ and $\Theta(0) = 0$. Note that $\Theta^{-1} : [0,\infty) \to [0,\infty)$ exists. The ORLICZ space associated to Θ is

$$L_\Theta(\Omega,\mu) := \left\{ f : \Omega \to \mathbf{R} \ \text{ measurable } \ : \exists \alpha \neq 0, \int_\Omega \Theta(\alpha \mathrm{f}) \mathrm{d}\mu < \infty \right\} \ .$$

On $L_\Theta(\Omega,\mu)$ we can define two equivalent norms defining a BANACH space structure, namely:

$$\|f\|_\Theta := \inf \left\{ \lambda > 0 : \int_\Omega \Theta\left(\frac{f}{\lambda}\right) d\mu \leq 1 \right\}$$

and

$$N_\Theta(f) := \sup_{g\in\mathcal{G}_\Theta}\left\{\int_\Omega |f|g d\mu\right\} .$$

Here, $\mathcal{G}_\Theta$ is a suitable family of non negative measurable functions g on Ω (see [15] or [1, chapter 6] for more details on this family). In particular, in [15] we learn that for all $f\in L_\Theta(\Omega,\mu)$, $N_\Theta(f)\le\|f\|_\Theta\le 2N_\Theta(f)$.

Define now $\Phi(x) = |x|\ln(1+|x|)$. Clearly Φ is a YOUNG function. MICLO in [14, proposition 11] gives the following result: for all functions $f^2\in L_\Phi(\mathcal{X},\mu)$,

$$\frac{2}{3}\inf_{t\in\mathbf{R}}\left\|(f-t)^2\right\|_\Phi \le \mathcal{L}(f) \le \frac{5}{2}\inf_{t\in\mathbf{R}}\left\|(f-t)^2\right\|_\Phi .$$

Hence, for all $x\in\mathcal{X}$, we have

$$\mathcal{L}(f) \le \frac{5}{2}\left\|(f-f(x))^2\right\|_\Phi .$$

Now observe that for all functions h such that $h^2\in L_\Phi(\mathcal{X},\mu)$, $\left\|h^2\right\|_\Phi \le 2N_\Phi(h^2)$. Thus, by definition of the norm N_Φ,

$$\begin{aligned}
\mathcal{L}(f) &\le 5\sup_{g\in\mathcal{G}_\Phi}\sum_{y\in\mathcal{X}}|f(y)-f(x)|^2\mu(y)g(y)\\
&\le 5\sum_{y\in\mathcal{X}}|f(y)-f(x)|^2\left(\sup_{g\in\mathcal{G}_\Phi}\mu(y)g(y)\right)\\
&\le 5\sum_{y\in\mathcal{X}}|f(y)-f(x)|^2 N_\Phi(\mathbb{1}_{\{y\}})
\end{aligned}$$

for a suitable family $\mathcal{G}_\Phi$ of non negative measurable functions g on $\mathcal{X}$. Then, an integration over all $x\in\mathcal{X}$ gives

$$\mathcal{L}(f) = \sum_{x\in\mathcal{X}}\mathcal{L}(f)\mu(x) \le 5\sum_{x,y\in\mathcal{X}\times\mathcal{X}}|f(y)-f(x)|^2\mu(x)N_\Phi(\mathbb{1}_{\{y\}}) .$$

The latter inequality will be the starting point of our study that we summarize in the next statement.

Lemma 2.1. *Let β be the best constant such that for all $f\in\mathbf{L^2}(\mu)$ on $\mathcal{X}$*

$$\sum_{x,y\in\mathcal{X}\times\mathcal{X}}|f(y)-f(x)|^2\mu(x)N_\Phi(\mathbb{1}_{\{y\}}) \le \beta\mathcal{E}_{\mu,\mathrm{K}}(\mathrm{f},\mathrm{f}) .$$

Then, if β is finite, the logarithmic SOBOLEV *constant α defined in (2) satisfies*

$$\alpha \le 5\beta .$$

Remark. The inequality in the lemma above is of POINCARÉ type. Indeed, the usual POINCARÉ inequality states that for all functions $f\in\mathbf{L^2}(\mu)$,

$$\mathbf{Var}_\mu(f) \le \lambda\mathcal{E}_{\mu,K}(f,f)$$

for some constant $\lambda > 0$. As it is classical, we can write

$$\mathbf{Var}_\mu(f) = \frac{1}{2} \sum_{x,y \in \mathcal{X} \times \mathcal{X}} |f(y) - f(x)|^2 \mu(x)\mu(y) .$$

Remark. It is well known (see [1, chapter 1] for instance) that the entropy has the following variational caracterization: for all $f \in \mathbf{L^2}(\mu)$,

$$\mathbf{Ent}_\mu(f^2) = \sup_{g:\mu(e^g)\leq 1} \mu(f^2 g) .$$

Here, the supremum is taken over a set of functions g possibly negative. On the contrary, in BOBKOV and GÖTZE approach $\mathcal{L}(f)$ is linked, up to universal constants, to $\sup_{g\in\mathcal{G}_\Phi} \mu(f^2 g)$ where g is non negative. The main feature of the non negativity of g is that it keeps the order of the inequalities.

3. Geometric bounds on the logarithmic Sobolev constant

From Lemma 2.1, we will be able to adapt the proof of SINCLAIR [18] (see also [17]) on geometric bounds for the spectral gap to the logarithmic SOBOLEV constant. To that aim we introduce some notations.

Our setting is that one defined in introduction: the state space $\mathcal{X}$ is viewed as a graph with vertices the points of $\mathcal{X}$ and with edges all couples $e = (x, y)$, $x \neq y$, such that $\mu(x)K(x, y) > 0$. We write $\mathbf{E}$ for the set of all edges. Then, a path from x to y is a sequence of vertices $\gamma = (x_0, \ldots, x_k)$ such that (x_i, x_{i+1}) is an edge, $i = 1 \ldots k-1$, $x_0 = x$ and $x_k = y$, and $|\gamma| = k$ is the length of the path. Let Γ be the set of all paths γ which have no repeated edges (*i.e.* for all $i \neq j$, $e_i \neq e_j$) and for all $(x, y) \in \mathcal{X} \times \mathcal{X}$, let Γ_{xy} be the set of all paths $\gamma_{xy} \in \Gamma$ starting at x and ending at y. By irreductibility of K, for all $(x, y) \in \mathcal{X} \times \mathcal{X}$, $\Gamma_{xy} \neq \emptyset$.

It is also convenient to introduce for any edge $e = (x, y) \in \mathbf{E}$ and any function f on $\mathcal{X}$, $df(e) = f(y) - f(x)$. Define now $Q(e) = \mu(x)K(x, y)$ and observe that

$$\mathcal{E}_{\mu,K}(f, f) = \frac{1}{2} \sum_{e\in\mathbf{E}} |df(e)|^2 Q(e) .$$

We start with the simplest result on path combinatorics.

Theorem 3.1. *Let K be an irreductible chain with reversible probability measure μ on a countable set $\mathcal{X}$. For all $(x, y) \in \mathcal{X} \times \mathcal{X}$, $x \neq y$, choose one path γ_{xy} in Γ_{xy}. Then, the logarithmic* SOBOLEV *constant α defined in (2) satisfies $\alpha \leq 10A_\Phi$, where*

$$A_\Phi := \sup_{e\in\mathbf{E}} \left\{ \frac{1}{Q(e)} \sum_{x,y:\gamma_{xy}\ni e} |\gamma_{xy}|\mu(x) N_\Phi(\mathbb{1}_{\{y\}}) \right\} .$$

Proof. We will make use of Lemma 2.1. For all $x, y \in \mathcal{X} \times \mathcal{X}$, write

$$f(y) - f(x) = \sum_{e\in\gamma_{xy}} df(e) .$$

By the CAUCHY-SCHWARZ inequality,

$$|f(y)-f(x)|^2 \le |\gamma_{xy}| \sum_{e\in\gamma_{xy}} |df(e)|^2 \ .$$

Thus,

$$\begin{aligned}
&\sum_{x,y\in\mathcal{X}\times\mathcal{X}} |f(y)-f(x)|^2\mu(x)N_\Phi(\mathbb{1}_{\{y\}}) \\
&\quad\le \sum_{x,y\in\mathcal{X}\times\mathcal{X}} |\gamma_{xy}| \sum_{e\in\gamma_{xy}} |df(e)|^2\mu(x)N_\Phi(\mathbb{1}_{\{y\}}) \\
&\quad\le \sum_{e\in\mathbf{E}} \left\{ \frac{1}{Q(e)} \sum_{x,y:\gamma_{xy}\ni e} |\gamma_{xy}|\mu(x)N_\Phi(\mathbb{1}_{\{y\}}) \right\} |df(e)|^2 Q(e) \\
&\quad\le 2A_\Phi\, \mathcal{E}_{\mu,K}(f,f) \ .
\end{aligned}$$

Applying Lemma 2.1 achieves the proof. □

The constant A_Φ of Theorem 3.1 must be compared to the one given by SINCLAIR in [18] in the case of the POINCARÉ inequality:

$$A_{S.G.} := \sup_{e\in\mathbf{E}} \left\{ \frac{1}{Q(e)} \sum_{x,y:\gamma_{xy}\ni e} |\gamma_{xy}|\mu(x)\mu(y) \right\} \ .$$

In our case, one can ask how to compute the spurious term $N_\Phi(\mathbb{1}_{\{y\}})$ (instead of $\mu(y)$)? Indeed, we have the following result:

Corollary 3.1. *Let K be an irreductible chain with reversible probability measure μ on a countable set $\mathcal{X}$. For all $(x,y)\in\mathcal{X}\times\mathcal{X}$, $x\ne y$, choose one path γ_{xy} in Γ_{xy}. Assume that for all $x\in\mathcal{X}$, $\mu(x)\le 1/2$. Then, the logarithmic* SOBOLEV *constant α defined in (2) satisfies $\alpha\le 20A$, where*

$$A := \sup_{e\in\mathbf{E}} \left\{ \frac{1}{Q(e)} \sum_{x,y:\gamma_{xy}\ni e} |\gamma_{xy}|\mu(x)\mu(y)\ln\frac{1}{\mu(y)} \right\} \ .$$

Proof. One of the BOBKOV and GÖTZE results [2, Lemma 5.4] states that for all $t\ge 2$,

$$\frac{1}{2}\frac{t}{\ln t} \le \Phi^{-1}(t) \le 2\frac{t}{\ln t} \ . \tag{3}$$

(where we recall that $\Phi(t)=|t|\ln(1+|t|)$). Thus, by definition of $\|\cdot\|_\Phi$, it is easy to check that for all $y\in\mathcal{X}$,

$$N_\Phi(\mathbb{1}_{\{y\}}) \le \left\|\mathbb{1}_{\{y\}}\right\|_\Phi = \frac{1}{\Phi^{-1}\left(\frac{1}{\mu(y)}\right)} \le 2\mu(y)\ln\frac{1}{\mu(y)} \ . \tag{4}$$

Applying Theorem 3.1 concludes the proof. □

The constant A of corollary 3.1 can now be compared to $A_{S.G.}$ introduced above. Obviously $A_{S.G.} \leq A/\ln 2$.

Remark. In practice, the hypothesis $\mu(x) \leq 1/2$ is often true and thus, Corollary 3.1 can be applied in many cases. However, we can be more precise. Indeed, Define $\mu^* := \sup_{x\in\mathcal{X}} \mu(x)$. It is not difficult to check that we can find a function Ψ such that for all $y \in \mathcal{X}$,

$$N_\Phi(\mathbb{1}_{\{\mathrm{y}\}}) \leq \frac{1}{\Phi^{-1}\left(\frac{1}{\mu(\mathrm{y})}\right)} \leq \Psi(\mu^*)\mu(\mathrm{y}) \ln \frac{1}{\mu(\mathrm{y})} .$$

Going back the proof of Corollary 3.1 yields $\alpha \leq 10\Psi(\mu^*)A$. Note that $\Psi(t)$ goes to infinity as t goes to 1.

There is an other way to compute $N_\Phi(\mathbb{1}_{\{\mathrm{y}\}})$. Indeed, there exists a universal constant c such that for any $0 \leq \mu(y) \leq 1$,

$$N_\Phi(\mathbb{1}_{\{\mathrm{y}\}}) \leq \frac{1}{\Phi^{-1}\left(\frac{1}{\mu(\mathrm{y})}\right)} \leq \mathrm{c}\mu(\mathrm{y}) \ln \left(2 \vee \frac{1}{\mu(\mathrm{y})}\right) .$$

Then, define $A := \sup_{e\in\mathbf{E}} \left\{ \frac{1}{Q(e)} \sum_{x,y:\gamma_{xy}\ni e} |\gamma_{xy}|\mu(x)\mu(y) \ln\left(2 \vee \frac{1}{\mu(y)}\right)\right\}$. Going back the proof of Corollary 3.1 yields $\alpha \leq 5cA$.

Example 1 (the simple random walk). We start with a simple example : let $\mathcal{X} = \{-N, \ldots, 0, \ldots, N\}$, $K(x+1,x) = K(x,x+1) = 1/2$ for all $-N \leq x < N$, $K(x,y) = 0$ if $|x-y| \neq 1$. K is reversible with respect to the uniform measure $\mu \equiv 1/(2N+1)$ on $\mathcal{X}$. The choice of a path γ_{xy} from any x to any y is imposed by the model. Let $e \in \mathbf{E}$ be an edge. Clearly, $e = (n, n+1)$ for some $n \in \mathcal{X}$ (or (n+1,n)), and $Q(e) = 1/2(2N+1)$. As $|\gamma_{xy}| \leq 2N+1$, we certainly have

$$\begin{aligned}
A &\leq \frac{2(2N+1)^2}{(2N+1)^2} \ln(2N+1) \max_e \left\{ \sum_{x,y:\gamma_{xy}\ni e} 1 \right\} \\
&\leq 2\ln(2N+1) \max_{-N\leq n<N} \left\{ \sum_{x\leq n<y} 1 \right\} \\
&\leq 2\ln(2N+1) \max_{-N\leq n<N} (n+N+1)(N-n) \\
&\leq 2N(N+1)\ln(2N+1) .
\end{aligned}$$

By Corollary 3.1 we can conclude that the logarithmic SOBOLEV constant satisfies $\alpha \leq 40N(N+1)\ln(2N+1)$. It is well known that the logarithmic SOBOLEV constant α is actually of order $O(N^2)$ (to see that claim, we can use HARDY's inequalities, see Section 4).

Example 2 (the hypercube $\mathbf{Z}_2^N$). We follow [17]. Let $\mathcal{X} = \{0,1\}^N$, $\mu \equiv 1/2^N$ and

$K(x,y)=0$ unless $|x-y|=1$ in which case $K(x,y)=1/N$. Define a path from x to y changing the coordinates of x to that one of y from the left to the right. Here, for all edges $e\in\mathbf{E}$, $Q(e)=1/(N2^N)$ and any path is at most of length N. Hence

$$\begin{aligned} A &\leq N^3\frac{1}{2^N}\ln 2\max_{e\in\mathbf{E}}\left\{\sum_{x,y:\gamma_{xy}\ni e}1\right\} \\ &\leq \frac{N^3\ln 2}{2^N}\max_{e\in\mathbf{E}}\#\left\{(x,y):\gamma_{xy}\ni e\right\}\ . \end{aligned}$$

Let $e=(u,v)$ be an edge of $\mathbf{E}$. By definition of the paths, there exists i such that $u_i\neq v_i$. Then we have

$$\begin{aligned} x &= (x_1,\ldots,x_{i-1},u_i,u_{i+1},\ldots,u_N) \\ y &= (v_1,\ldots,v_{i-1},v_i,y_{i+1},\ldots,y_N)\ . \end{aligned}$$

It follows that $i-1$ coordinates of x and $N-i$ coordinates of y are free, this yields that

$$\max_{e\in\mathbf{E}}\#\left\{(x,y):\gamma_{xy}\ni e\right\}=2^{N-1}\ .$$

Therefore

$$A\leq\frac{\ln 2}{2}N^3\ .$$

Corollary 3.1 allows us to conclude that the logarithmic SOBOLEV constant α is bounded from above by $10\ln 2N^3$. It is known that the right order of the logarithmic SOBOLEV constant α is $O(N)$. However, note the non trivial cancellation of the exponential.

Example 3 (a graph). It is interesting to study the order of magnitude of A on a graph. Let $\mathcal{X}=G$ be a graph and consider the random walk on this graph. We assume that G is connected and simple, that is, G has no loops or multiple edges. The random walk starts at one point and choose a neighbor vertex with uniform probability. Thus, if $d(x)$ is the degree of x (*i.e.* the number of neighbors), we have

$$K(x,y)=\begin{cases}1/d(x) & \text{if } x\sim y\\ 0 & \text{otherwise}\ .\end{cases}$$

Here, $x\sim y$ means that x and y are neighbors in G. The chain is reversible with respect to the measure $\mu(x)=d(x)/(2|\mathbf{E}|)$. Assume that for all $x\in G$, $\mu(x)\leq 1/2$. Since the graph is connected, the chain is irreductible. For any $e\in\mathbf{E}$, $Q(e)=1/(2|\mathbf{E}|)$. Then, consider any choice of geodesic path from any x to any y and define $d^*:=\sup d(x)$, $d_*:=\inf d(x)$, γ^* as the maximum number of edges in any path and $b:=\sup_e\#\{\gamma:e\in\gamma\}$. Corollary 3.1 gives

$$\alpha\leq 10\frac{(d^*)^2\gamma^*b}{|\mathbf{E}|}\ln\frac{2|\mathbf{E}|}{d_*}\ .$$

Note that b can be interpreted as a measure of bottlenecks and γ^* as an upper bound of the diameter of G.

As an application, consider the full binary tree of depth N (see [5, example 2.3]). DIACONIS and STROOCK compute that $d^*=3$, $d_*=2$, $\gamma^*=2N$, $|\mathbf{E}|=2^{N+1}-2$ and

$b = (2^N - 1)2^N$. Thus, $\alpha \leq 90 N^2 2^N$. In [5], we learn that the spectral gap (and so the logarithmic SOBOLEV constant) of the tree is of order greater that $O(2^N)$. Hence, the bound is reasonable thanks to the exponential term.

Example 4 (a birth and death process on N). Let $\mathcal{X} = \mathbf{N}$, the measure μ such that $\mu(n) = Z^{-1}\exp(-n(n+1)/2)$ with the normalization constant $Z = \sum_{x\in\mathbf{N}} \exp(-n(n+1)/2)$ (note that $Z \leq e$), $K(n,n+1) = 1$ for all $n \in \mathbf{N}$, $K(n,n-1) = e^n$ for all $n \in \mathbf{N}^*$ and $K(x,y) = 0$ if $|x-y| \neq 1$. The choice of a path from any x to any y is forced by the model.

Fix $n \in \mathbf{N}$ and consider $e = (n,n+1)$ and $e' = (n+1,n)$. Clearly

$$\sum_{x,y:\gamma_{xy}\ni e} |\gamma_{xy}|\mu(x)\mu(y)\ln\frac{1}{\mu(y)} \geq \sum_{x,y:\gamma_{xy}\ni e'} |\gamma_{xy}|\mu(x)\mu(y)\ln\frac{1}{\mu(y)} \; .$$

Thus, it is enough to study for each $n \in \mathbf{N}$

$$\begin{aligned}
&\sum_{x\leq n<y} |\gamma_{xy}|\mu(x)\mu(y)\ln\frac{1}{\mu(y)} \\
&\qquad\qquad \leq \sum_{x\leq n}\mu(x)\sum_{y\geq n+1} yZ^{-1}e^{-\frac{y(y+1)}{2}}\left(\frac{y(y+1)}{2} + \ln Z\right) \\
&\qquad\qquad \leq \sum_{y\geq n+1} y^2(y+1)Z^{-1}e^{-\frac{y(y+1)}{2}} \; .
\end{aligned}$$

Here, we simply used that $\sum_{x\leq n}\mu(x) \leq 1$ and $\ln Z \leq 1 \leq \frac{y^2(y+1)}{2}$. Now it is not difficult to check that

$$\sum_{y\geq n+1} y^2(y+1)Z^{-1}e^{-\frac{y(y+1)}{2}} \leq 2(n+1)^2(n+2)Z^{-1}e^{-\frac{(n+1)(n+2)}{2}} \; .$$

As $Q((n,n+1)) = \mu(n)$, it follows that

$$\begin{aligned}
A &\leq \sup_{n\in\mathbf{N}} 2(n+1)^2(n+2)e^{-\frac{1}{2}((n+1)(n+2)-n(n+1))} \\
&\leq \sup_{n\in\mathbf{N}} 2(n+1)^2(n+2)e^{-n-1} < \infty \; .
\end{aligned}$$

In the case of an infinite set, one of the important point is already to know if whether or not α is finite. Here, Corollary 3.1 yields $\alpha < \infty$. Note that a direct application of HARDY's inequality (see section 4) also gives this result.

Example 5 (an infinite star). Let $\mathcal{X} = \mathbf{N}$ and choose a sequence of non negative numbers $(w_i)_{i\in\mathbf{N}}$ such that $\sum_{i=1}^{\infty} w_i = 1/2$ and $S := (1/2) + \sum_{i=1}^{\infty} w_i \ln(1/w_i) < \infty$. Then, define the chain on $\mathcal{X}$ by its transition matrix K by $K(i,0) = 1/(2w_i)$, $K(0,i) = 1$ for all $i \geq 1$ and $K(i,j) = 0$ otherwise. This chain is reversible with respect to the probability measure $\mu(0) = 1/2$ and $\mu(i) = w_i$ for all $i \geq 1$.

This example is presented in [16] in the case of a MARKOV chain with $K(0,i) = w_i$, in such a way that K is a transition probability. But, as mentioned in the introduction, for a MARKOV chain on a infinite set, the logarithmic SOBOLEV constant is certainly infinite.

Note that a finite version of "the star" is discussed in [5, Remark 2.5] with equal weights.

The graph of $\mathcal{X}$ can be drawned as a "star" with 0 at the center and with all positive integers around, connected to 0.

The choice of path from i to j, $i \neq j$, $i,j \geq 1$ is forced: $\gamma_{ij} = ((i,0),(0,j))$, while $\gamma_{0i} = (0,i)$ and $\gamma_{i0} = (i,0)$. We certainly have $|\gamma_{xy}| \leq 2$ for all $x,y \in \mathcal{X} \times \mathcal{X}$. Now, let $e = (i,0)$ for some $i \geq 1$, as $Q(e) = 1/2$, it follows that

$$\begin{aligned}\frac{1}{Q(e)} \sum_{x,y:\gamma_{xy} \ni e} |\gamma_{xy}| \mu(x)\mu(y) \ln \frac{1}{\mu(y)} &\leq 4\mu(i) \sum_{j \neq i} \mu(j) \ln \frac{1}{\mu(j)} \\ &= 4w_i \left(S - w_i \ln \frac{1}{w_i} \right) \\ &\leq 2S \;.\end{aligned}$$

On the other hand, if $e = (0,i)$ for $i \geq 1$, $Q(e) = 1/2$ and

$$\begin{aligned}\frac{1}{Q(e)} \sum_{x,y:\gamma_{xy} \ni e} |\gamma_{xy}| \mu(x)\mu(y) \ln \frac{1}{\mu(y)} &\leq 4\mu(i) \ln \frac{1}{\mu(i)} \sum_{j \neq i} \mu(j) \\ &\leq 4w_i \ln \frac{1}{w_i}(1 - w_i) \\ &\leq 2 \;.\end{aligned}$$

Hence, Corollary 3.1 yields that the logarithmic Sobolev constant α defined in (2) satisfies $\alpha \leq 40(S \vee 1)$.

Now, we turn to more sophisticated geometric bounds. We introduce the notion of *weight function*, that is a positive function on the set of edges, $w : \mathbf{E} \to (0, \infty)$ and we define its associated w-length of a path $\gamma \in \Gamma$ by $|\gamma|_w := \sum_{e \in \gamma} \frac{1}{w(e)}$.

Theorem 3.2 (weight function). *Let K be an irreductible chain with reversible probability measure μ on a countable set $\mathcal{X}$. For all $(x,y) \in \mathcal{X} \times \mathcal{X}$, $x \neq y$, choose one path γ_{xy} in Γ_{xy}. Then, for any weight function w, the logarithmic* Sobolev *constant α defined in (2) satisfies $\alpha \leq 10A_\Phi^w$, where*

$$A_\Phi^w := \sup_{e \in \mathbf{E}} \left\{ \frac{w(e)}{Q(e)} \sum_{x,y:\gamma_{xy} \ni e} |\gamma_{xy}|_w \mu(x) N_\Phi(\mathbb{1}_{\{y\}}) \right\} .$$

Proof. Fix a weight function w. The proof starts as in the proof of Theorem 3.1 but introduce the weight function w in the Cauchy-Schwarz inequality to get

$$|f(y) - f(x)|^2 = \left(\sum_{e \in \gamma_{xy}} |df(e)| \right)^2 \leq \left(\sum_{e \in \gamma_{xy}} \frac{1}{w(e)} \right) \left(\sum_{e \in \gamma_{xy}} |df(e)|^2 w(e) \right) .$$

Then, follow step by step the proof of Theorem 3.1 to conclude. $\square$

Taking back the proof of Corollary 3.1 gives the following result:

Corollary 3.2 (weight function). *Let K be an irreductible chain with reversible probability measure μ on a countable set $\mathcal{X}$. For all $(x,y) \in \mathcal{X} \times \mathcal{X}$, $x \neq y$, choose one path γ_{xy} in Γ_{xy}. Assume that for all $x \in \mathcal{X}$, $\mu(x) \leq 1/2$. Then, for any weight function w, the logarithmic Sobolev constant α defined in (2) satisfies $\alpha \leq 20A^w$, where*

$$A^w := \sup_{e \in \mathbf{E}} \left\{ \frac{w(e)}{Q(e)} \sum_{x,y:\gamma_{xy} \ni e} |\gamma_{xy}|_w \mu(x)\mu(y) \ln \frac{1}{\mu(y)} \right\} .$$

Theorem 3.2 (resp. Corollary 3.2) is clearly a refinement of Theorem 3.1 (resp. Corollary 3.1). Indeed, it is enough to consider the trivial weight function $w \equiv 1$.

In [5], the authors study the spectral gap by means of the equivalent constant of A^w with the particular choice of weight function $w : e \mapsto Q(e)$.

We now apply Corollary 3.2 to an example where a nice choice of weight function improves considerably the bound A of Corollary 3.1.

Example 6 (birth and death process on $\{0,\dots,N\}$). On $\mathcal{X} = \{0,\dots,N\}$, define $\mu(n) = Z^{-1}e^{-n}$ with $Z = \sum_{n=0}^{N} e^{-n}$, $K(n,n+1) = 1$ and $K(n+1,n) = e$ for all $0 \leq n$ and $K(x,y) = 0$ if $|x-y| \neq 1$. The choice of a path from any x to any y in $\mathcal{X}$ is forced by the model.

Consider the following weight function $w : e = (n,n+1) \mapsto \sqrt{\mu(n)}$ and $e = (n+1,n) \mapsto \sqrt{\mu(n)}$. Clearly,

$$\begin{aligned} A^w \leq \quad & \sup_{0 \leq n < N} \left\{ \frac{w((n,n+1))}{Q((n,n+1))} \sum_{x \leq n < y} |\gamma_{xy}|_w \mu(x)\mu(y) \ln \frac{1}{\mu(y)} \right\} \vee \\ & \sup_{0 \leq n < N} \left\{ \frac{w((n+1,n))}{Q((n+1,n))} \sum_{y \leq n < x} |\gamma_{xy}|_w \mu(x)\mu(y) \ln \frac{1}{\mu(y)} \right\} . \end{aligned}$$

Fix $0 \leq n$, $x \leq n$ and $y \geq n+1$. For notational convenience, k denotes below a numerical constant possibly changing from line to line. First, remark that

$$|\gamma_{xy}|_w = \sum_{i=x}^{y} \frac{1}{\sqrt{\mu(i)}} = \sqrt{Z} \sum_{i=x}^{y} e^{i/2} \leq k\sqrt{Z}e^{y/2} = k\frac{1}{\sqrt{\mu(y)}} .$$

Then, an easy computation gives

$$\begin{aligned} \sum_{x \leq n < y} |\gamma_{xy}|_w \mu(x)\mu(y) \ln \frac{1}{\mu(y)} & \leq k \sum_{y \geq n+1} \sqrt{\mu(y)} \ln \frac{1}{\mu(y)} \\ & \leq k \frac{1}{\sqrt{Z}} \sum_{y \geq n+1} y e^{-y/2} \\ & \leq k \frac{1}{\sqrt{Z}} (n+1) e^{-(n+1)/2} \\ & = k(n+1)\sqrt{\mu(n+1)} . \end{aligned}$$

As $Q((n,n+1)) = \mu(n)$ and $w((n,n+1)) = \sqrt{\mu(n)}$, it follows that

$$\begin{aligned}\frac{w((n,n+1))}{Q((n,n+1))}\sum_{x\leq n<y}|\gamma_{xy}|_w\mu(x)\mu(y)\ln\frac{1}{\mu(y)} &\leq k(n+1)\sqrt{\frac{\mu(n+1)}{\mu(n)}}\\ &\leq k(n+1)\ .\end{aligned}$$

On the other hand, the same kind of calculus gives

$$\begin{aligned}&\frac{w((n+1,n))}{Q((n+1,n))}\sum_{y\leq n<x}|\gamma_{xy}|_w\mu(x)\mu(y)\ln\frac{1}{\mu(y)}\\ &\qquad\leq k\frac{1}{\sqrt{\mu(n)}}\sum_{y\leq n<x}\sqrt{\mu(x)}\mu(y)\ln\frac{1}{\mu(y)}\\ &\qquad\leq k\frac{(n+1)}{\sqrt{\mu(n)}}\sum_{x\geq n+1}\sqrt{\mu(x)}\\ &\qquad\leq k(n+1)\sqrt{\frac{\mu(n+1)}{\mu(n)}}\\ &\qquad\leq k(n+1)\ .\end{aligned}$$

Here, we used $|\gamma_{xy}|_w \leq k/\sqrt{\mu(x)}$ for all $x \geq y$ and the trivial bound $\mu(y)\ln\frac{1}{\mu(y)} \leq 1$.

We can now conclude that $A^w \leq kN$ and thus by Corollary 3.2 that $\alpha \leq kN$. It is known that the logarithmic Sobolev constant α is of order $O(N)$ (as we can see using Hardy's inequalities, see section 4). Note that the same kind of calculus as previously would have given $A \leq kN^2$ (where A is defined in Corollary 3.1). This example shows the effectiveness of the weight functions.

Our next step in complexifying geometric bounds is the notion of flow. In previous theorems, we used exactly one path from any x to any y in $\mathcal{X}$. We now consider more than one path. Introduce a *flow* Ψ on the set of paths Γ, that is, a non negative function $\Psi : \Gamma \to [0,\infty)$ satisfying for all $x, y \in \mathcal{X}\times\mathcal{X}$, $x \neq y$,

$$\sum_{\gamma\in\Gamma_{xy}}\Psi(\gamma) = \mu(x)\mu(y)\ln\frac{1}{\mu(y)}\ .$$

A flow can be viewed as a probability measure on the set of paths Γ_{xy} starting at x and ending at y, simply because $\sum_{\gamma\in\Gamma_{xy}}\Psi(\gamma)/\mu(x)\mu(y)\ln\frac{1}{\mu(y)} = 1$.

Theorem 3.3 (flow function). *Let K be an irreductible chain with reversible probability measure μ on a countable set $\mathcal{X}$. Assume that for all $x \in \mathcal{X}$, $\mu(x) \leq 1/2$. Then, for any flow function Ψ, the logarithmic Sobolev constant α defined in (2) satisfies $\alpha \leq 20A^\Psi$, where*

$$A^\Psi := \sup_{e\in\mathbf{E}}\left\{\frac{1}{Q(e)}\sum_{\gamma:\gamma\ni e}|\gamma|\Psi(\gamma)\right\}\ .$$

Proof. Fix a flow Ψ. For any $x,y \in \mathcal{X}\times\mathcal{X}$ and any $\gamma \in \Gamma_{xy}$, we have

$$|f(y)-f(x)|^2 \leq |\gamma| \sum_{e\in\gamma} |df(e)|^2 .$$

Thus, by definition of the flow,

$$|f(y)-f(x)|^2 \mu(x)\mu(y)\ln\frac{1}{\mu(y)} \leq \sum_{\gamma\in\Gamma_{xy}} |\gamma| \sum_{e\in\gamma} |df(e)|^2 \Psi(\gamma) .$$

Then, as in the proof of Theorem 3.1, we get

$$\begin{aligned}\sum_{x,y\in\mathcal{X}\times\mathcal{X}} |f(y)-f(x)|^2\mu(x)\mu(y)\ln\frac{1}{\mu(y)} &\leq \sum_{e\in\gamma}\left\{\frac{1}{Q(e)}\sum_{\gamma:\gamma\ni e}|\gamma|\Psi(\gamma)\right\}|df(e)|^2 Q(e) \\ &\leq 2A^{\Psi}\mathcal{E}_{\mu,K}(f,f) .\end{aligned}$$

Now, thanks to inequality (4),

$$\sum_{x,y\in\mathcal{X}\times\mathcal{X}} |f(y)-f(x)|^2\mu(x)N_{\Phi}(\mathbb{1}_{\{\mathrm{y}\}}) \leq 4 \sum_{\mathrm{x},\mathrm{y}\in\mathcal{X}\times\mathcal{X}} |\mathrm{f}(\mathrm{y})-\mathrm{f}(\mathrm{x})|^2\mu(\mathrm{x})\mu(\mathrm{y})\ln\frac{1}{\mu(\mathrm{y})}$$

Applying Lemma 2.1 achieves the proof. ☐

Remark. If we define a flow as a function satisfying for $x,y\in\mathcal{X}\times\mathcal{X}$,

$$\sum_{\gamma\in\Gamma_{xy}} \Psi(\gamma) = \mu(x)N_{\Phi}(\mathbb{1}_{\{\mathrm{y}\}}) ,$$

the proof above gives $\alpha \leq 10A^{\Psi}$.

Remark 1. If we choose one path γ_{xy} for all $(x,y)\in\mathcal{X}\times\mathcal{X}$, $x\neq y$ and let $\Psi(\gamma_{xy}) = \mu(x)\mu(y)\ln\frac{1}{\mu(y)}$ and $\Psi(\gamma)=0$ for all $\gamma\in\Gamma_{xy}\backslash\{\gamma_{xy}\}$, then Ψ is a trivial flow and Theorem 3.3 is Corollary 3.1. This is for example the case on a graph where the choice of any path is forced. In that case, the notion of flow does not play any role.

Example 7 (on groups). To illustrate the notion of flow, we give a result relative to the action of a group on a finite set $\mathcal{X}$. This is a direct adaptation of SALOFF-COSTE result [17, corollary 3.6].

Let G be a group that acts on a finite set $\mathcal{X}$ such that for all $g\in G$, all $x\in\mathcal{X}$,

$$\mu(gx)=\mu(x), \qquad Q(gx,gy)=Q(x,y) .$$

Hence, if $e=(u,v)\in\mathbf{E}$, then $ge=(gu,gv)\in\mathbf{E}$. Let $\mathbf{E}=\bigcup_1^k \mathbf{E}_i$ be the partition of $\mathbf{E}$ into transitive classes for the action of G (this means that for all $(e_i,e_i')\in\mathbf{E}_i\times\mathbf{E}_i$, there exists $g\in G$ such that $ge_i=e_i'$ and for all $(e_i,e_j)\in\mathbf{E}_i\times\mathbf{E}_j$, $i\neq j$, we cannot find $g\in G$ such that $e_i=ge_j$). Denote $Q(i)=Q(e_i)$ for $e_i\in\mathbf{E}_i$ and $d(x,y)$ the graph distance between x and y. Consider the set $\mathcal{G}_{xy}$ of all geodesic paths from x to y and define the flow Ψ by

$$\Psi(\gamma)=\begin{cases}\left(\mu(x)\mu(y)\ln\frac{1}{\mu(y)}\right)/\#\mathcal{G}_{xy} & \text{if } \gamma\in\mathcal{G}_{xy}\\ 0 & \text{otherwise .}\end{cases}$$

Then, a direct adaptation of the proof of Corollary 3.6 of [17] (we omit the proof) yields that $\alpha \leq 20A$, where

$$A := \max_{1\leq i\leq k}\left\{\frac{1}{|\mathbf{E}_i|Q(i)}\sum_{x,y} d(x,y)^2\mu(x)\mu(y)\ln\frac{1}{\mu(y)}\right\}.$$

If we look at the hypercube (see Example 2), $\mathbf{Z}_2^N$ acts on itself transitively, there are $N2^N$ oriented edges and $k=1$. Hence an easy calculus shows that $\alpha \leq 20\ln 2N^3$. Note that we loose a factor 2 with respect to Example 2 using Corollary 3.1.

We finish this section with a refinement on geometric bounds that have been suggested to us by Laurent MICLO. The bound mix the notions of flow and of weight functions and the weight functions depend not only of the edges but also of the paths. This bound is certainly interesting in order to study the optimality of paths method.

We give the result without any proof because it a direct easy adaptation of the proofs of Theorems 3.2 and 3.3.

We introduce the notion of path-weighted function, that is a set of positive functions on the set of edges, indexed by the family of paths Γ, $w_\gamma : \mathbf{E} \to (0,\infty)$. Also, we define its associated w_γ-length of the path γ by $|\gamma|_{w_\gamma} = \sum_{e\in\gamma}\frac{1}{w_\gamma(e)}$.

Then, we have the following result.

Theorem 3.4 (path-weighted and flow functions). *Let K be an irreductible chain with reversible probability measure μ on a countable set $\mathcal{X}$. Assume that for all $x\in\mathcal{X}$, $\mu(x)\leq 1/2$. Then, for any flow function Ψ and any path-weighted function $(w_\gamma)_\gamma$, the logarithmic* SOBOLEV *constant α defined in (2) satisfies $\alpha \leq 20A^{\Psi,(w_\gamma)_\gamma}$, where*

$$A^{\Psi,(w_\gamma)_\gamma} := \sup_{e\in\mathbf{E}}\left\{\frac{1}{Q(e)}\sum_{\gamma:\gamma\ni e}|\gamma|_{w_\gamma}w_\gamma(e)\Psi(\gamma)\right\}.$$

Remark. If we consider the trivial flow function introduced in Remark 1, then we have $\alpha \leq 20A^{(w_\gamma)_\gamma}$ where

$$A^{(w_\gamma)_\gamma} := \sup_{e\in\mathbf{E}}\left\{\frac{1}{Q(e)}\sum_{x,y:\gamma_{xy}\ni e}|\gamma_{xy}|_{w_{\gamma_{xy}}}w_{\gamma_{xy}}(e)\mu(x)\mu(y)\ln\frac{1}{\mu(y)}\right\}.$$

4. Comparison results

In this section, we compare different bounds on spectral gap and logarithmic SOBOLEV constant. We start by the comparison of the SINCLAIR path bound to the HARDY constant. Then, we turn to the same kind of comparison with A^w. In both cases, we restrict our study to the birth and death chains. At last, we compare our bound to a result of DIACONIS and SALOFF-COSTE result [4].

We mention that this section arised from various interesting discussions with Laurent MICLO that we warmly acknowledge for his hints.

4.1. On Poincaré inequality

Our setting is the following: let $\mathcal{X} = \mathbf{N}$ and K be a birth and death chain on $\mathcal{X}$, that is, $K(x,y) = 0$ unless $|x-y| = 1$. We assume that K is reversible with respect to some probability measure μ. For technical reasons, we make the assumption that $\mu(0) = 1/2$ in such a way that 0 is a median of μ. Note that there is only one way to construct a path from any x to any y in $\mathcal{X}$.

From [13], if

$$B_{S.G.} := \sup_{n\geq 1} \sum_{i=1}^{n-1} \frac{1}{\mu(i)K(i,i+1)} \sum_{j=n}^{\infty} \mu(j) ,$$

we know that the spectral gap λ^{-1} defined in (1) satisfies

$$\frac{1}{2} B_{S.G.} \leq \lambda \leq 4 B_{S.G.} . \tag{5}$$

On the other hand, SINCLAIR's result (see [18]) states that for any weight function w (see section 3 for notations),

$$\lambda \leq A^w_{S.G.} := \sup_{e\in\mathbf{E}} \left\{ \frac{w(e)}{Q(e)} \sum_{x,y:\gamma_{xy}\ni e} |\gamma_{xy}|_w \mu(x)\mu(y) \right\} . \tag{6}$$

As the previous inequality holds for any weight function, one can ask for the optimality of such a bound, *i.e.* can we find w such that $\lambda = A^w_{S.G.}$? Or at least, can we find w such that for some universal constant k, $\lambda \geq k A^w_{S.G.}$? KAHALE [8] answers such a question in various cases. Let us present an easy result on that topic by means of the constant of HARDY $B_{S.G.}$ defined above.

Define the weight function $w : e = (n,n+1) \mapsto \left(\sum_{i=0}^n 1/Q(i)\right)^{1/2} Q(n)$ with $Q(i) := \mu(i)K(i,i+1)$, $i \in \mathbf{N}$. Now, note that by concavity, for any $a,b \in \mathbf{R}_+$, $b^{1/2} - a^{1/2} \geq \frac{1}{2b^{1/2}}(b-a)$, thus, if $x \leq y$,

$$\begin{aligned}
|\gamma_{xy}|_w &= \sum_{j=x}^{y-1} \frac{1}{\left(\sum_{i=0}^j 1/Q(i)\right)^{1/2}} \frac{1}{Q(j)} \\
&\leq 2\left[\left(\sum_{i=0}^{y-1} 1/Q(i)\right)^{1/2} - \left(\sum_{i=0}^{x} 1/Q(i)\right)^{1/2}\right] \\
&\leq 2\left(\sum_{i=0}^{y-1} 1/Q(i)\right)^{1/2} \\
&\leq 2\sqrt{B_{S.G.}} \frac{1}{\sqrt{\sum_{j=y}^{\infty} \mu(j)}} .
\end{aligned}$$

The last inequality comes from the definition of $B_{S.G.}$. Then, for the edge $e = (n,n+1) \in \mathbf{E}$, we have

$$\frac{w(e)}{Q(e)} \sum_{x,y:\gamma_{xy}\ni e} |\gamma_{xy}|_w \mu(x)\mu(y)$$

$$\begin{aligned}
&\le 2\sqrt{B_{S.G.}}\left(\sum_{i=0}^{n} 1/Q(i)\right)^{1/2} \sum_{x\le n<y} \frac{\mu(x)\mu(y)}{\sqrt{\sum_{j=y}^{\infty}\mu(j)}} \\
&\le 2\sqrt{B_{S.G.}}\left(\sum_{i=0}^{n} 1/Q(i)\right)^{1/2} \sum_{y=n+1}^{\infty} \frac{\mu(y)}{\sqrt{\sum_{j=y}^{\infty}\mu(j)}} \qquad (7)\\
&\le 2\sqrt{B_{S.G.}}\left(\sum_{i=0}^{n} 1/Q(i)\right)^{1/2} 2\left(\sum_{j=n+1}^{\infty}\mu(j)\right)^{1/2} \\
&\le 4B_{S.G.}\ .
\end{aligned}$$

Here, we used the trivial bound $\sum_{x\le n}\mu(x)\le 1$, the concave inequality $b^{1/2}-a^{1/2}\ge \frac{1}{2b^{1/2}}(b-a)$ and the definition of $B_{S.G.}$. The latter inequality holds for any edge $e\in\mathbf{E}$. Hence, $A^w_{S.G.}\le 4B_{S.G.}$. From (5) and (6) we thus deduce the following caracterization:

Proposition 4.1. *Let $\mathcal{X}=\mathbf{N}$ and K be such that $K(x,y)=0$ unless $|x-y|=1$. Assume that K is reversible with respect to a probability measure μ with $\mu(0)=1/2$. Then, if $w: e=(n,n+1)\mapsto \left(\sum_{i=0}^{n}1/Q(i)\right)^{1/2}Q(n)$, with $Q(i):=\mu(i)K(i,i+1)$, $i\in\mathbf{N}$, we have*

$$\frac{1}{8}A^w_{S.G.}\le\lambda\le A^w_{S.G.}\ .$$

The proposition above says that, up to universal constants, a particular choice of weight function always gives the right order of magnitude for $A^w_{S.G.}$. Exactly the same result applies to birth and death chains on $\{0,\dots,N\}$.

4.2. On logarithmic Sobolev inequality

Consider the same setting as in Section 4.1 with $\mathcal{X}=\{0,\dots,N\}$. From [13], if

$$B_{L.S.}:=\max_{n\ge 1}\sum_{i=1}^{n-1}\frac{1}{\mu(i)K(i,i+1)}\left(\sum_{j=n}^{N}\mu(j)\right)\ln\frac{1}{\sum_{j=n}^{N}\mu(j)}\ ,$$

we know that the logarithmic Sobolev constant α defined in (2) satisfies

$$\frac{1}{31}B_{L.S.}\le\alpha\le 20B_{L.S.}\ . \qquad (8)$$

We now want to compare $B_{L.S.}$ to the constant A^w_Φ defined in Theorem 3.2. To that aim introduce as in Section 4.1 the weight function

$$w:(n,n+1),(n+1,n)\mapsto\left(\sum_{i=0}^{n}1/Q(i)\right)^{1/2}Q(n)\ ,$$

with $Q(i) := \mu(i)K(i,i+1)$, $i \in \mathcal{X} \setminus \{N\}$. Define for any $g \in \mathcal{G}_\Phi$ (see section 2 for the definitions of $\mathcal{G}_\Phi$ and Φ),

$$A_g^{w,+} := \max_{n\in\mathcal{X}\setminus\{N\}} \left\{ \frac{w(n,n+1)}{Q(n)} \sum_{x,y:\gamma_{xy}\ni(n,n+1)} |\gamma_{xy}|_w \mu(x)\mu(y)g(y) \right\} ,$$

$$A_g^{w,-} := \max_{n\in\mathcal{X}\setminus\{N\}} \left\{ \frac{w(n+1,n)}{Q(n)} \sum_{x,y:\gamma_{xy}\ni(n+1,n)} |\gamma_{xy}|_w \mu(x)\mu(y)g(y) \right\} ,$$

and

$$B_g := \max_{n\geq 1} \sum_{i=1}^{n-1} \frac{1}{\mu(i)K(i,i+1)} \sum_{j=n}^{N} \mu(j)g(j) \ . \tag{9}$$

We recall that in the proof of Theorem 3.2, we proved that the constant A_Φ^w satisfies $A_\Phi^w \geq \sup_{g\in\mathcal{G}_\Phi}\left(A_g^{w,+} \vee A_g^{w,-}\right)$. On the other hand, using a comparison between a sum and a supremum, it is not difficult to check that $A_\Phi^w \leq |\mathcal{X}|^2 \sup_{g\in\mathcal{G}_\Phi}\left(A_g^{w,+} \vee A_g^{w,-}\right)$. Thus,

$$\sup_{g\in\mathcal{G}_\Phi}\left(A_g^{w,+} \vee A_g^{w,-}\right) \leq A_\Phi^w \leq |\mathcal{X}|^2 \sup_{g\in\mathcal{G}_\Phi}\left(A_g^{w,+} \vee A_g^{w,-}\right) . \tag{10}$$

Now, the same proof as in Section 4.1 yields that for any $g \in \mathcal{G}_\Phi$, $A_g^{w,+} \leq 4B_g$.

A similar inequality holds for $e = (n+1,n)$. Indeed, let $e = (n+1,n)$ for $n \in \mathcal{X} \setminus \{N\}$ and $g \in \mathcal{G}_\Phi$. Start the proof as in section 4.1 until inequality (7). Here, instead of $\sum_{x\leq n}\mu(x) \leq 1$, use

$$\sum_{y\leq n}\mu(y)g(y) \leq N_\Phi(\mathbb{1}_\mathcal{X}) \leq \|\mathbb{1}_\mathcal{X}\|_\Phi = \frac{1}{\Phi^{-1}(1)} \leq 1 \ .$$

Recall that $\Phi(t) = |t|\ln(1+|t|)$. Then, complete the proof as in Section 4.1 to get that for any $g \in \mathcal{G}_\Phi$, $A_g^{w,-} \leq 4B_{\mathbb{1}}$. Here $B_{\mathbb{1}}$ is defined in (9) with $g \equiv 1$ (even if $\mathbb{1} \notin \mathcal{G}_\Phi$).

By definition of $\|\cdot\|_\Phi$ and from (3), we have

$$\begin{aligned}\sup_{g\in\mathcal{G}_\Phi}\sum_{j=n}^{N}\mu(j)g(j) &\leq N_\Phi(\mathbb{1}_{\{\mathrm{n},\ldots,\mathrm{N}\}}) \leq \left\|\mathbb{1}_{\{\mathrm{n},\ldots,\mathrm{N}\}}\right\|_\Phi = \frac{1}{\Phi^{-1}(1/\mu(\{\mathrm{n},\ldots,\mathrm{N}\}))} \\ &\leq 2\mu(\{n,\ldots,N\})\ln\frac{1}{\mu(\{n,\ldots,N\})} \ .\end{aligned}$$

Hence, $\sup_{g\in\mathcal{G}_\Phi} B_g \leq 2B_{L.S.}$ and thus $\sup_{g\in\mathcal{G}_\Phi} A_g^{w,+} \leq 8B_{L.S.}$.

On the other hand, there exists a universal constant $c > 0$ such that $\mathbb{1}/\mathrm{c} \in \mathcal{G}_\Phi$. It follows that for any $g \in \mathcal{G}_\Phi$, $A_g^{w,-} \leq 8cB_{L.S.}$. Putting all together the previous results gives

$$\sup_{g\in\mathcal{G}_\Phi}\left(A_g^{w,+} \vee A_g^{w,-}\right) \leq 8(c\vee 1)B_{L.S.} \ .$$

Now, remark that there is a way to compute the constant c. Indeed, define $\Psi(x) := \int_0^x (\Phi')^{-1}(t)dt$ for $x \geq 0$, then $c \geq 1/\Psi^{-1}(1)$ (see [1, chapter 6]). Thus, a rought computation gives $c \geq 2$.

Now, inequalities (8) and (10) and Theorem 3.2 allow us to conclude to the following result.

Proposition 4.2. *Let $\mathcal{X} = \{0, \ldots, N\}$ and K be such that $K(x,y) = 0$ unless $|x-y| = 1$. Assume that K is reversible with respect to a probability measure μ with $\mu(0) = 1/2$. Then, if $w : e = (n, n+1), (n+1, n) \mapsto \left(\sum_{i=0}^{n} 1/Q(i)\right)^{1/2} Q(n)$, with $Q(i) := \mu(i)K(i, i+1)$, $i \in \mathcal{X} \setminus \{N\}$, we have*

$$\frac{1}{496|\mathcal{X}|^2} A_\Phi^w \leq \alpha \leq 5A_\Phi^w \ .$$

Remark 2. From the inequalities $N_\Phi(f) \leq \|f\|_\Phi \leq 2N_\Phi(f)$ (see [15]) and (3), we have $A^w/4 \leq A_\Phi^w \leq 4A^w$. Thus, the previous proposition gives also a comparison between A^w and α: $\frac{1}{1984|\mathcal{X}|^2} A^w \leq \alpha \leq 10A^w$.

Remark. Note that in example 6, if we take the weight function introduced above, we find again that $\alpha \leq kN$ for some constant k, *i.e.* the right order for α. This is true simply because for any n, $\left(\sum_{i=0}^{n} 1/Q(i)\right)^{1/2} Q(n)$ is of the same order, up to universal constants, as $\sqrt{Q(n)}$.

The latter proposition partially answers the question of the optimality of the constant A_Φ^w. We will now see how it can help us to compare the bound A^w to the following result of Diaconis and Saloff-Coste [4]: for any finite set $\mathcal{X}$, if $\mu_* := \min_{x\in\mathcal{X}} \mu(x)$, they proved that

$$\alpha \leq \lambda \frac{\ln(1/\mu_*) - 1}{1 - 2\mu_*} \ . \tag{11}$$

Here, λ^{-1} is the spectral gap defined in (1).

Assume that one can compute λ (for example using path techniques). Then, the latter inequality gives a bound for the logarithmic Sobolev constant α. One can ask whether or not this bound is better than A^w ? Indeed, in examples 1 and 6, both bounds are of the same order. In example 2, if one knows that λ is of order $O(N)$, then, Diaconis and Saloff-Coste's bound is better than ours. In this case, λ must have been computed in a different way than path techniques. Actually, as we know, path techniques give nothing better than $O(N^2)$, see [17]. Tensorization techniques give the right order for λ (and even, exactly the right constant) in the case of the hypercube.

More generally, if μ is the uniform distribution over a finite set $\mathcal{X}$, then, even using path techniques to compute λ, the Diaconis and Saloff-Coste bound is certainly better than A^w (just use inequality (6)).

On the other hand a first remark is that inequality (11) is no more valid when $|\mathcal{X}| = \infty$ while A^w still has a meaning. This is the case of Examples 4 and 5 for instance.

Moreover, we have the following example suggested to us by Laurent Miclo: on $\mathcal{X} = \{0, 1, 2\}$, let the birth and death process $K(0,1) = \varepsilon^2$ and $K(2,1) = 1$ reversible with respect to $\mu(0) = 1/2$, $\mu(1) = (1/2) - \varepsilon$ and $\mu(2) = \varepsilon$. Here $\varepsilon \ll 1$ is a positive constant. Proposition 4.1 and 4.2 yield that the spectral gap and the logarithmic Sobolev constant are of order $O(1/\varepsilon^2)$, while inequality (11) gives $\alpha \leq \frac{1}{\varepsilon^2} \ln \frac{1}{\varepsilon}$. In that case, A^w is certainly better than inequality (11). Note that the HARDY inequalities (5) and (8) are an other way to compute the spectral gap and the logarithmic SOBOLEV constant (the maximum is given by the edge $(0,1)$).

If one computes the spectral gap using path techniques, then, we must remark that it is certainly better to compute directly A or A^w instead of computing first $A_{S.G.}$ or $A^w_{S.G.}$ and then using inequality (11). Indeed, in that case, note that in the definition of A or A^w we average $\ln(1/\mu(y))$ over $\mathcal{X}$ while inequality (11) considers only $\ln(1/\mu_*)$. As $A_{S.G.}$ and A are both easy (or difficult) to compute, it is better to compute directly A (or A^w).

At last, if one looks at the examples presented in [5] as the circle $\mathbf{Z}_N$ [5, Example 2.1], or the "star" (the finite version of Example 5), or also the random walk on graphs (see our Example 3), then, DIACONIS and SALOFF-COSTE's bound is at least better than A^w.

As a conclusion, inequality (11) and A^w are not universally comparable. For finite sets, DIACONIS and SALOFF-COSTE's bound seems to be more convenient to compute in a lot of cases, while for infinite sets, only A^w makes sense.

Acknowledgement. I warmly acknowledge Laurent MICLO for useful hints and discussions. I am also grateful to Michel LEDOUX and Laurent SALOFF-COSTE for their helpful comments and to Sébastien BLACHÈRE for a careful reading of the manuscript.

References

[1] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques*. 2000. to appear in Panorama et Synthèse, S.M.F.

[2] S. G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163(1):1–28, 1999.

[3] Jeff Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. In *Problems in analysis (Papers dedicated to Salomon Bochner, 1969)*, pages 195–199. Princeton Univ. Press, Princeton, N. J., 1970.

[4] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.*, 6(3):695–750, 1996.

[5] P. Diaconis and D. Stroock. Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Probab.*, 1(1):36–61, 1991.

[6] J.A. Fill. Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains, with an application to the exclusion process. *Ann. Appl. Probab.*, 1(1):62–87, 1991.

[7] M. Jerrum and A. Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM J. Comput.*, 22(5):1087–1116, 1993.

[8] N. Kahale. A semidefinite bound for mixing rates of Markov chains. *Random Structures Algorithms*, 11(4):299–313, 1997.

[9] G. F. Lawler and A. D. Sokal. Bounds on the L^2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality. *Trans. Amer. Math. Soc.*, 309(2):557–580, 1988.

[10] P. Mathieu. Hitting times and spectral gap inequalities. *Ann. Inst. H. Poincaré Probab. Statist.*, 33(4):437–465, 1997.

[11] P. Mathieu and P. Picco. Metastability and convergence to equilibrium for the random field Curie-Weiss model. *J. Statist. Phys.*, 91(3-4):679–732, 1998.

[12] L. Miclo. Remarques sur l'hypercontractivité et l'évolution de l'entropie pour des chaînes de Markov finies. In *Séminaire de Probabilités XXXI*, volume 1655 of *Lecture Notes in Math.*, pages 136–167. Springer, Berlin, 1997.

[13] L. Miclo. An example of application of discrete Hardy's inequalities. *Markov Process. Related Fields*, 5(3):319–330, 1999.

[14] L. Miclo. Relations entre isopérimetrie et trou spectral pour les chaînes de Markov finies. *Probab. Theor. Relat. Fields*, 114:431–485, 1999.

[15] M. M. Rao and Z. D. Ren. *Theory of Orlicz spaces.* Marcel Dekker Inc., New York, 1991.

[16] J. S. Rosenthal. Markov chain convergence: from finite to infinite. *Stochastic Process. Appl.*, 62(1):55–72, 1996.

[17] L. Saloff-Coste. Lectures on finite Markov chains. In *Lectures on probability theory and statistics. École d'été de probabilités de St-Flour 1996*, volume 1665 of *Lecture Notes in Math.*, pages 301–413. Springer, Berlin, 1997.

[18] A. Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combin. Probab. Comput.*, 1(4):351–370, 1992.

[19] A. Sinclair. *Algorithms for random generation and counting.* Birkhäuser Boston Inc., Boston, MA, 1993. A Markov chain approach.