Combinatorics, Probability and Computing (2003) 12, 431–455. © 2003 Cambridge University Press
DOI: 10.1017/S0963548302005515 Printed in the United Kingdom

A Path Method for the Logarithmic Sobolev Constant

CYRIL ROBERTO

Université Paul Sabatier, L.S.P., 118 route de Narbonnes, 31062 Toulouse, France
(e-mail: roberto@cict.fr)

Received 7 March 2001; revised 22 May 2002

This paper is concerned with path techniques for quantitative analysis of the logarithmic Sobolev constant on a countable set. We present new upper bounds on the logarithmic Sobolev constant, which generalize those given by Sinclair [20], in the case of the spectral gap constant involving path combinatorics. Some examples of applications are given. Then, we compare our bounds to the Hardy constant in the particular case of birth and death processes. Finally, following the approach of Rosenthal in [18], we generalize our bounds to continuous sets.

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 $x,y \in \mathcal{X}\times\mathcal{X}$, $K(x,y)\geqslant 0$ 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 by $L(x,y)=K(x,y)$, 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) and, in particular, it acts on those functions f that are null, except at finitely many points in $\mathcal{X}$, by

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

By irreducibility, the law of the position of the process converges to the equilibrium state μ, as t tends to infinity. By irreducibility, μ *charges* all the points, that is, for all $x\in\mathcal{X}$, $\mu(x)>0$.

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

Now, for any function $f \in \mathbf{L}^2(\mu)$, we define the *Dirichlet form* of f, associated with 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,

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

Functional inequalities play a crucial role in the analysis of the behaviour of the process. In particular, we have 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 $f \in \mathbf{L}^2(\mu)$,

$$\mathbf{Var}_\mu(f) \leqslant \lambda \mathcal{E}_{\mu,K}(f,f). \tag{1.1}$$

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

$$\mathbf{Ent}_\mu(f^2) \leqslant \alpha \mathcal{E}_{\mu,K}(f,f). \tag{1.2}$$

The best ergodic constants λ^{-1} and α such that (1.1) and (1.2) 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 long must we wait to be near equilibrium? We refer the reader principally to [5], [4], [19] and the references therein for details on this convergence. There are also applications to theoretical computer science; see, for instance, [20] or [21].

In our context, we mention that it is not interesting to consider a Markov kernel, *i.e.*, satisfying, for all $x \in \mathcal{X}$, $\sum_{y \in \mathcal{X}} K(x,y) = 1$. This is firstly 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 hereafter) 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 path combinatoric methods. These were introduced by Jerrum and Sinclair [7] in theoretical computer science, in their study of a stochastic algorithm that counts perfect matchings in a graph. One of the initial ideas was to control the Cheeger constant. The method was further developed on finite sets by Diaconis and Stroock [5], Fill [6] and Sinclair [20]. Rosenthal generalized them to

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

countable sets (and even continua) in [18]. In [11] the authors adapt path techniques to the inequalities of Mathieu (see [10]).

The present paper develops the same kind of path combinatorics to bound the logarithmic Sobolev constant. To that end, we make use of an interesting idea of Bobkov and Götze [2], to reduce the study of the logarithmic Sobolev inequality (1.2) to an inequality of Poincaré type that allows us to generalize Sinclair's bounds.

We now present one of our results to illustrate our purpose. Regard $\mathcal{X}$ 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 irreducibility, we can construct a path γ_{xy} from any x to any y, that 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 $x_k = y$. The length of such a path γ is $|\gamma| = k$. Then we have the following result.

Theorem 1.1. *Let K be an irreducible 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 $x \in \mathcal{X}$, $\mu(x) \leqslant 1/2$. Then the logarithmic Sobolev constant α defined in* (1.2) *satisfies $\alpha \leqslant 20A$, where*

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

From here on, we will often write $\sup_{e \in \mathbf{E}}$ even when $|\mathbf{E}| < \infty$. The latter theorem corresponds to Theorem 5 of [20].

Next we extend Theorem 1.1 to a family of weight functions. A weight function is simply a positive function$w : \mathbf{E} \to (0, \infty)$ on the set of edges. The bound of Theorem 1.1 may then be improved (see Theorem 3.3 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 [20]: see Theorem 3.5 below.

We finish our discussion with a comparison of path techniques and Hardy's inequalities in the particular case of a birth and death process on $\mathbb{N}$. 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,

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

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

The same weight function also gives a result on the logarithmic Sobolev constant, but (unfortunately) it is less 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 $x \in \mathcal{X}$, $\sum_{y\in\mathcal{X}} K(x,y) \leqslant 1$, then the logarithmic Sobolev constant α is certainly infinite (see [12]). To see why, it is sufficient to consider a sequence of test functions $\mathbf{1}_{A_n}$ in the logarithmic Sobolev inequality (1.2) with $\mu(A_n) \to 0$ as n tends 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 constructed in the light of this remark.

In the last section, we generalize path methods for the logarithmic Sobolev constant to uncountable sets, following the approach of Rosenthal in [18].

Finally, we mention that the results of this work hold true with a transition matrix for an irreducible 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 that follows.

2. Logarithmic Sobolev inequality via 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 to the logarithmic Sobolev constant in terms of path combinatorics. The main idea comes from [2].

Let K be an irreducible chain with reversible probability measure μ on a countable set $\mathcal{X}$. In [2], the authors introduce the quantity $\mathcal{L}(f) := \sup_{t\in\mathbb{R}} \mathbf{Ent}_\mu(f+t)^2$ for all functions $f \in \mathbf{L}^2(\mu)$. Then, by translation invariance of the Dirichlet form, it is equivalent to considering, for all $f \in \mathbf{L}^2(\mu)$, the logarithmic Sobolev inequality (1.2)

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

or the inequality

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

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 notation and useful facts on Orlicz norms.

Let (Ω,μ) be a probability space and $\Theta : \mathbb{R} \to [0,\infty)$ 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 with Θ is

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

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

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

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

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 nonnegative measurable functions g on Ω (see [17] or [1, Chapter 6] for more details on this family). In particular, in [17] we learn that, for all $f \in L_\Theta(\Omega,\mu)$, $N_\Theta(f) \leqslant \|f\|_\Theta \leqslant 2N_\Theta(f)$.

Now define $\Phi(x) = |x|\ln(1+|x|)$. Clearly Φ is a Young function. Miclo in [14, Proposition 11] gives the following result. For all functions f with $f^2 \in L_\Phi(\mathcal{X},\mu)$,

$$\frac{2}{3}\inf_{t\in\mathbb{R}}\|(f-t)^2\|_\Phi \leqslant \mathcal{L}(f) \leqslant \frac{5}{2}\inf_{t\in\mathbb{R}}\|(f-t)^2\|_\Phi.$$

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

$$\mathcal{L}(f) \leqslant \tfrac{5}{2}\|(f-f(x))^2\|_\Phi.$$

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

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

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

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

The latter inequality will be the starting point of our study, which 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 $\mathcal{X}$,*

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

Then, if β is finite, the logarithmic Sobolev constant α defined in (1.2) *satisfies*

$$\alpha \leqslant 5\beta.$$

Remarks. **(1)** The usual Poincaré inequality states that, for all functions $f \in \mathbf{L}^2(\mu)$,

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

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

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

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

Thus, for the quadratic expression, the inequality in the lemma above is of Poincaré type. However, as in general $\mu(y) \neq N_\Phi(\mathbf{1}_{\{y\}})$, there is some asymmetry.

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

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

Here, the supremum is taken over a set of functions g possibly negative. On the contrary, in the 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 nonnegative. The main feature of the nonnegativity 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 [20] (see also [19]) on geometric bounds for the spectral gap to the logarithmic Sobolev constant. To that end, we introduce some notation.

Our setting is that defined in the Introduction: the state space $\mathcal{X}$ is viewed as a graph whose vertices are the points of $\mathcal{X}$ and whose edges are 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 the irreducibility 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 $\mathcal{X}$, $df(e) = f(y) - f(x)$. Now define $Q(e) = \mu(x)K(x, y)$, 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 irreducible 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* (1.2) *satisfies $\alpha \leqslant 10 A_\Phi$, where*

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

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

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,

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

Thus,

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

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 [20] 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 $N_\Phi(\mathbf{1}_{\{y\}})$ (instead of $\mu(y)$). Indeed, we have the following result.

Corollary 3.2. *Let K be an irreducible 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 $x \in \mathcal{X}$, $\mu(x) \leqslant 1/2$. Then, the logarithmic Sobolev constant α defined in* (1.2) *satisfies $\alpha \leqslant 20A$, 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 $t \geqslant 2$,

$$\frac{1}{2} \frac{t}{\ln t} \leqslant \Phi^{-1}(t) \leqslant 2 \frac{t}{\ln t} \tag{3.1}$$

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

$$N_\Phi\left(\mathbf{1}_{\{y\}}\right) \leqslant \left\|\mathbf{1}_{\{y\}}\right\|_\Phi = \frac{1}{\Phi^{-1}\left(\frac{1}{\mu(y)}\right)} \leqslant 2\mu(y) \ln \frac{1}{\mu(y)}. \tag{3.2}$$

Applying Theorem 3.1 concludes the proof. □

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

The constant A of Corollary 3.2 can now be compared to $A_{S.G.}$ introduced above. Obviously $A_{S.G.} \leqslant A/\ln 2$.

Remark. In practice, the hypothesis $\mu(x) \leqslant 1/2$ is often true, and thus Corollary 3.2 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 $y \in \mathcal{X}$,

$$N_\Phi\left(\mathbf{1}_{\{y\}}\right) \leqslant \frac{1}{\Phi^{-1}\left(\dfrac{1}{\mu(y)}\right)} \leqslant \Psi(\mu^*)\mu(y)\ln\frac{1}{\mu(y)}.$$

Going back to the proof of Corollary 3.2 yields $\alpha \leqslant 10\Psi(\mu^*)A$. Note that $\Psi(t)$ tends to infinity as t tends to 1.

There is another way to compute $N_\Phi(\mathbf{1}_{\{y\}})$. Indeed, there exists a universal constant c such that, for any $0 \leqslant \mu(y) \leqslant 1$,

$$N_\Phi(\mathbf{1}_{\{y\}}) \leqslant \frac{1}{\Phi^{-1}\left(\dfrac{1}{\mu(y)}\right)} \leqslant c\mu(y)\ln\left(2 \vee \frac{1}{\mu(y)}\right).$$

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 to the proof of Corollary 3.2 yields $\alpha \leqslant 5cA$.

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 $-N \leqslant x < N$, $K(x,y) = 0$ 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 $|\gamma_{xy}| \leqslant 2N+1$, we certainly have

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

By Corollary 3.2 we can conclude that the logarithmic Sobolev constant satisfies $\alpha \leqslant 40N(N+1)\ln(2N+1)$. 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).

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

Example 2. (The hypercube $\mathbb{Z}_2^N$) We follow [19]. 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 those of y from left to right. Here, for all edges $e \in \mathbf{E}$, $Q(e) = 1/(N2^N)$ and any path is at most of length N. Hence

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

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

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

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

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

Therefore

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

Corollary 3.2 allows us to conclude that the logarithmic Sobolev constant α is bounded from above by $10 \ln 2 N^3$. It is known that the correct order of the logarithmic Sobolev constant α is $O(N)$. However, note the nontrivial 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 neighbour vertex with uniform probability. Thus, if $d(x)$ is the degree of x (*i.e.*, the number of neighbours), 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 neighbours in G. The chain is reversible with respect to the measure $\mu(x) = d(x)/(2|\mathbf{E}|)$. Assume that, for all $x \in G$, $\mu(x) \leqslant 1/2$. Since the graph is connected, the chain is irreducible. 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 3.2 gives

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

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

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

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, $\alpha \leqslant 90N^2 2^N$. 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. (Birth and death process on $\mathbb{N}$) Let $\mathcal{X} = \mathbb{N}$, the measure μ such that $\mu(n) = Z^{-1}\exp(-n(n+1)/2)$ with the normalization constant $Z = \sum_{x\in\mathbb{N}} \exp(-n(n+1)/2)$ (note that $Z \leqslant e$), $K(n, n+1) = 1$ for all $n \in \mathbb{N}$, $K(n, n-1) = e^n$ for all $n \in \mathbb{N} \setminus \{0\}$ 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 $n \in \mathbb{N}$ and consider $e = (n, n+1)$ and $e' = (n+1, n)$. Clearly

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

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

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

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

$$\sum_{y\geqslant n+1} y^2(y+1)Z^{-1}e^{-\frac{y(y+1)}{2}} \leqslant 2(n+1)^2(n+2)Z^{-1}e^{-\frac{(n+1)(n+2)}{2}}.$$

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

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

In the case of an infinite set, one of the important points is to know whether or not α is finite. Here, Corollary 3.2 yields $\alpha < \infty$. Note that a direct application of Hardy's inequality (see Section 4) also gives this result.

Example 5. (Infinite star) Let $\mathcal{X} = \mathbb{N}$ and choose a sequence of nonnegative numbers $(w_i)_{i\in\mathbb{N}}$ 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 $i \geqslant 1$, 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 $i \geqslant 1$.

This example is presented in [18] 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.

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

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 drawn as a 'star' with 0 at the centre and with all positive integers around, connected to 0.

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

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

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

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

Hence, Corollary 3.2 yields that the logarithmic Sobolev constant α defined in (1.2) satisfies $\alpha \leqslant 40(S \vee 1)$.

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 3.3. (Weight function) *Let K be an irreducible 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* (1.2) *satisfies $\alpha \leqslant 10 A_\Phi^w$, where*

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

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

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

Then, we follow the proof of Theorem 3.1 step by step. □

Going back to the proof of Corollary 3.2 gives the following result.

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

Corollary 3.4. (Weight function) *Let K be an irreducible 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 $x \in \mathcal{X}$, $\mu(x) \leqslant 1/2$. Then, for any weight function w, the logarithmic Sobolev constant α defined in* (1.2) *satisfies $\alpha \leqslant 20A^w$, 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 3.3 (resp. Corollary 3.4) is clearly a refinement of Theorem 3.1 (resp. Corollary 3.2). 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 3.4 to an example where a nice choice of weight function considerably improves the bound A of Corollary 3.2.

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 $0 \leqslant n$ 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 weight function $w : e = (n,n+1) \mapsto \sqrt{\mu(n)}$ and $e = (n+1,n) \mapsto \sqrt{\mu(n)}$. Clearly,

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

Fix $0 \leqslant n$, $x \leqslant n$ and $y \geqslant n+1$. For notational convenience, k will denote a numerical constant, possibly changing from line to line. First, note that

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

Then, easy computation gives

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

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

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

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

On the other hand, the same kind of calculation gives

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

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

We can now conclude that $A^w \leqslant kN$ and thus, by Corollary 3.4, that $\alpha \leqslant kN$. 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 calculation as before would have given $A \leqslant kN^2$ (where A is defined in Corollary 3.2). 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 nonnegative 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 3.5. (Flow function) *Let K be an irreducible chain with reversible probability measure μ on a countable set $\mathcal{X}$. Assume that, for all $x \in \mathcal{X}$, $\mu(x) \leqslant 1/2$. Then, for any flow function Ψ, the logarithmic Sobolev constant α defined in* (1.2) *satisfies $\alpha \leqslant 20A^\Psi$, where*

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

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

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

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

Thus, by definition of the flow,

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

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

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

Now, thanks to inequality (3.2),

$$\sum_{x,y \in \mathcal{X} \times \mathcal{X}} |f(y) - f(x)|^2 \mu(x) n_{\Phi}(\mathbf{1}_{\{y\}}) \leqslant 4 \sum_{x,y \in \mathcal{X} \times \mathcal{X}} |f(y) - f(x)|^2 \mu(x)\mu(y) \ln \frac{1}{\mu(y)}.$$

Applying Lemma 2.1 achieves the proof. □

Remarks. **(1)** If we define a flow as a function satisfying, for $x, y \in \mathcal{X} \times \mathcal{X}$,

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

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

(2) 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 $\gamma \in \Gamma_{xy} \setminus \{\gamma_{xy}\}$, then Ψ is a trivial flow and Theorem 3.5 becomes Corollary 3.2. 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's result [19, Corollary 3.6].

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

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

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 a $g \in G$ 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 let $d(x, y)$ be 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

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

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

Then, a direct adaptation of the proof of Corollary 3.6 of [19] (we omit the proof) yields that $\alpha \leqslant 20A$, where

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

If we look at the hypercube (see Example 2), $\mathbb{Z}_2^N$ acts on itself transitively, there are $N2^N$ oriented edges and $k = 1$. Hence an easy calculation shows that $\alpha \leqslant 20 \ln 2 N^3$. Note that we lose a factor of 2 with respect to Example 2, using Corollary 3.2.

Although the notion of flow seems difficult to apply, in [15] we find a nontrivial computation of a spectral gap using flow functions. This shows the effectiveness of the notion.

We finish this section with a refinement on geometric bounds suggested to us by Laurent Miclo. The bound mixes the notions of flow and weight functions, and the weight functions depend not only on the edges but also on the paths. This bound is certainly interesting for studying the optimality of the path method.

We give the result without any proof because it is a direct, easy adaptation of the proofs of Theorems 3.3 and 3.5.

We introduce the notion of path-weighted functions, 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 the 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 3.6. (Path-weighted and flow functions) *Let K be an irreducible chain with reversible probability measure μ on a countable set $\mathcal{X}$. Assume that, for all $x \in \mathcal{X}$, $\mu(x) \leqslant 1/2$. Then, for any flow function Ψ and any path-weighted function $(w_\gamma)_\gamma$, the logarithmic Sobolev constant α defined in* (1.2) *satisfies $\alpha \leqslant 20A^{\Psi,(w_\gamma)_\gamma}$, 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 (2) on page 444, then we have $\alpha \leqslant 20A^{(w_\gamma)_\gamma}$, 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. Finally, we compare our bound to a result of Diaconis and Saloff-Coste [4].

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

4.1. On Poincaré inequality

Our setting is the following: let $\mathcal{X} = \mathbb{N}$ 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

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

we know that the spectral gap λ^{-1} defined in (1.1) satisfies

$$\tfrac{1}{2} B_{S.G.} \leqslant \lambda \leqslant 4 B_{S.G.}. \tag{4.1}$$

The latter inequality comes from Hardy's constant. We mention that Muckenhoupt (see [16]) proved a similar result on the line in the continuous setting.

On the other hand, Sinclair's result (see [20]) states that, for any weight function w (see Section 3 for notation),

$$\lambda \leqslant A^w_{S.G.} := \sup_{e \in \mathbf{E}} \left\{ \frac{w(e)}{Q(e)} \sum_{x,y : \gamma_{xy} \ni e} |\gamma_{xy}|_w \mu(x)\mu(y) \right\}. \tag{4.2}$$

As the previous inequality holds for any weight function, we consider the optimality of such a bound, that is, 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, $\lambda \geqslant k A^w_{S.G.}$? Kahale [8] answers such a question in various cases. Let us present an easy result on that topic by means of Hardy's constant $B_{S.G.}$, 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 $Q(i) := \mu(i)K(i, i+1)$, $i \in \mathbb{N}$. Now note that, by concavity, for any $a, b \in \mathbb{R}_+$, $b^{1/2} - a^{1/2} \geqslant \frac{1}{2b^{1/2}}(b - a)$; thus, if $x \leqslant y$,

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

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

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

$$\begin{aligned}
&\frac{w(e)}{Q(e)} \sum_{x,y:\gamma_{xy} \ni e} |\gamma_{xy}|_w \mu(x)\mu(y) \\
&\quad \leqslant 2\sqrt{B_{S.G.}} \left(\sum_{i=0}^{n} 1/Q(i) \right)^{1/2} \sum_{x \leqslant n < y} \frac{\mu(x)\mu(y)}{\sqrt{\sum_{j=y}^{\infty} \mu(j)}} \\
&\quad \leqslant 2\sqrt{B_{S.G.}} \left(\sum_{i=0}^{n} 1/Q(i) \right)^{1/2} \sum_{y=n+1}^{\infty} \frac{\mu(y)}{\sqrt{\sum_{j=y}^{\infty} \mu(j)}} \qquad (4.3) \\
&\quad \leqslant 2\sqrt{B_{S.G.}} \left(\sum_{i=0}^{n} 1/Q(i) \right)^{1/2} 2 \left(\sum_{j=n+1}^{\infty} \mu(j) \right)^{1/2} \\
&\quad \leqslant 4B_{S.G.}.
\end{aligned}$$

Here, we used the trivial bound $\sum_{x \leqslant n} \mu(x) \leqslant 1$, the concave inequality $b^{1/2} - a^{1/2} \geqslant \frac{1}{2b^{1/2}}(b-a)$ and the definition of $B_{S.G.}$. The latter inequality holds for any edge $e \in \mathbf{E}$. Hence, $A^w_{S.G.} \leqslant 4B_{S.G.}$. From (4.1) and (4.2) we thus deduce the following characterization.

Proposition 4.1. *Let $\mathcal{X} = \mathbb{N}$ 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 $Q(i) := \mu(i)K(i, i+1)$, $i \in \mathbb{N}$, we have

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

The proposition above states 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 $\{0, \ldots, N\}$.

4.2. On logarithmic Sobolev inequality

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

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

we know that the logarithmic Sobolev constant α defined in (1.2) satisfies

$$\frac{1}{31} B_{L.S.} \leqslant \alpha \leqslant 20 B_{L.S.}. \qquad (4.4)$$

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

We now want to compare $B_{L.S.}$ to the constant A_Φ^w defined in Theorem 3.3. To that end, 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

$$B_g := \max_{n \geqslant 1} \sum_{i=1}^{n-1} \frac{1}{\mu(i)K(i, i+1)} \sum_{j=n}^{N} \mu(j)g(j). \tag{4.5}$$

We recall that in the proof of Theorem 3.3 we proved that the constant A_Φ^w satisfies $A_\Phi^w \geqslant \sup_{g \in \mathcal{G}_\Phi} (A_g^{w,+} \vee A_g^{w,-})$. On the other hand, using a comparison between a sum and a supremum, it is not difficult to check that $A_\Phi^w \leqslant |\mathcal{X}|^2 \sup_{g \in \mathcal{G}_\Phi} (A_g^{w,+} \vee A_g^{w,-})$. Thus

$$\sup_{g \in \mathcal{G}_\Phi} \left(A_g^{w,+} \vee A_g^{w,-} \right) \leqslant A_\Phi^w \leqslant |\mathcal{X}|^2 \sup_{g \in \mathcal{G}_\Phi} (A_g^{w,+} \vee A_g^{w,-}). \tag{4.6}$$

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

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 (4.3). Here, instead of $\sum_{x \leqslant n} \mu(x) \leqslant 1$, use

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

Recall that $\Phi(t) = |t| \ln(1 + |t|)$. Then, complete the proof as in Section 4.1 to get that, for any $g \in \mathcal{G}_\Phi$, $A_g^{w,-} \leqslant 4B_\mathbf{1}$. Here $B_\mathbf{1}$ is defined in (4.5) with $g \equiv 1$ (even if $\mathbf{1} \notin \mathcal{G}_\Phi$).

Then, by definition of $\|\cdot\|_\Phi$ and from (3.1), we have

$$\sup_{g \in \mathcal{G}_\Phi} \sum_{j=n}^{N} \mu(j)g(j) \leqslant N_\Phi\left(\mathbf{1}_{\{n,\dots,N\}}\right) \leqslant \left\|\mathbf{1}_{\{n,\dots,N\}}\right\|_\Phi = \frac{1}{\Phi^{-1}(1/\mu(\{n, \dots, N\}))}$$

$$\leqslant 2\mu(\{n, \dots, N\}) \ln \frac{1}{\mu(\{n, \dots, N\})}.$$

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

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

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

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

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

Now, inequalities (4.4) and (4.6) and Theorem 3.3 allow us to conclude with 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 $Q(i) := \mu(i)K(i, i+1)$, $i \in \mathcal{X} \backslash \{N\}$, we have

$$\frac{1}{496|\mathcal{X}|^2} A_\Phi^w \leqslant \alpha \leqslant 10 A_\Phi^w.$$

Remarks. **(1)** From the inequalities $N_\Phi(f) \leqslant \|f\|_\Phi \leqslant 2N_\Phi(f)$ (see [17]) and (3.1), we have $A^w/4 \leqslant A_\Phi^w \leqslant 4A^w$. Thus, the previous proposition also gives a comparison between A^w and α: $\frac{1}{1984|\mathcal{X}|^2} A^w \leqslant \alpha \leqslant 20A^w$.

(2) Note that, in Example 6, if we take the weight function introduced above, we find again that $\alpha \leqslant kN$ for some constant k, *i.e.*, the right order for α. This is true simply because, for any n, $(\sum_{i=0}^{n} 1/Q(i))^{1/2} Q(n)$ 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

$$\alpha \leqslant \lambda \frac{\ln(1/\mu_*) - 1}{1 - 2\mu_*}. \tag{4.7}$$

Here, λ^{-1} is the spectral gap defined in (1.1).

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 we know 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 to path techniques. Actually, as we know, path techniques give nothing better than $O(N^2)$ (see [19]). 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 (4.2)).

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

On the other hand we remark that inequality (4.7) is no longer valid when $|\mathcal{X}| = \infty$, while A^w is still meaningful. 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. Propositions 4.1 and 4.2 yield that the spectral gap and the logarithmic Sobolev constant are of order $O(1/\varepsilon^2)$, while inequality (4.7) gives $\alpha \leqslant \frac{1}{\varepsilon^2} \ln \frac{1}{\varepsilon}$. In that case, A^w is certainly better than inequality (4.7). Note that the Hardy inequalities (4.1) and (4.4) are another way to compute the spectral gap and the logarithmic Sobolev constant (the maximum is given by the edge $(0, 1)$).

If we compute the spectral gap using path techniques, then we must remark that it is certainly better to compute A or A^w directly, rather than first compute $A_{S.G.}$ or $A^w_{S.G.}$ and then use inequality (4.7). 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 (4.7) considers only $\ln(1/\mu_*)$. As $A_{S.G.}$ and A are both easy (or difficult) to compute, it is better to compute A (or A^w) directly.

Finally, if we look at the examples presented in [5] such as the circle $\mathbb{Z}_N$ [5, Example 2.1], or the 'star' (the finite version of Example 5), or 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 (4.7) and A^w are not universally comparable. For finite sets, Diaconis and Saloff-Coste's bound seems more convenient to compute in many of cases, while for infinite sets only A^w makes sense.

5. Path method on $\mathbb{R}^n$

In this section, we consider extensions of path methods to uncountable state spaces. Our result is a simple generalization of Rosenthal's [18] approach to the spectral gap. We follow [18] for notation and definitions. We conclude with an example of an application.

Let us start with some basic definitions.

Throughout this section, we assume that $\mathcal{X}$ is an open subset of $\mathbb{R}^n$ with $\mathcal{C}^1$ boundary. The transition matrix is now a transition kernel $K(x, \cdot)$, reversible with respect to a stationary measure μ, and irreducible with respect to the Lebesgue measure v on $\mathbb{R}^n$. Let $\mathcal{C}_c^\infty(\mathcal{X})$ be the set of $\mathcal{C}^\infty$ functions on $\mathbb{R}^n$ with compact support in $\mathcal{X}$.

Then the Dirichlet form of $f \in \mathcal{C}_c^\infty(\mathcal{X})$ associated with K and μ is defined by

$$\mathcal{E}_{\mu,K}(f,f) := \mu(f(\mathrm{Id} - K)f) = \frac{1}{2} \int\!\!\int_{\mathcal{X}\times\mathcal{X}} (f(y) - f(x))^2 K(x, dy)\mu(dx),$$

where Id is the operator identity, while the entropy of f^2 is

$$\mathbf{Ent}_\mu(f^2) := \mu\left(f^2 \ln \frac{f^2}{\mu(f^2)} \right).$$

With this setting, we say that μ satisfies a logarithmic Sobolev inequality if there exists $\alpha > 0$ such that, for all $f \in \mathcal{C}_c^\infty(\mathcal{X})$,

$$\mathbf{Ent}_\mu(f^2) \leqslant \alpha \mathcal{E}_{\mu,K}(f,f). \tag{5.1}$$

An optimal value of α is called the logarithmic Sobolev constant of μ.

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

As for countable spaces, our aim is to bound α from above, using path techniques.

We need some regularity conditions because we use differentiability properties. We follow [18].

We call a subset of $\mathbb{R}^n$ *gentle* if it is contained in some finite union of $\mathcal{C}^1$ submanifolds of $\mathbb{R}^n$ of codimension 1 ($\mathcal{C}^1$ hypersurfaces). This means that a gentle subset is small, and in particular that its Lebesgue measure is null.

We assume that μ has density h with respect to the Lebesgue measure, and that $h > 0$ on $\mathcal{X}$. Furthermore, we assume that we can write $K(x, dy) = a_x \delta_x(dy) + f_x(y)v(dy)$, for all $x \in \mathcal{X}$, with $a_x = 1 - \int_{\mathcal{X}} f_x(y)v(dy)$. We assume that $h(\cdot)$ and $\int_A f.dv$ are uniformly continuous functions of $x \in \mathcal{X}$ and that $f_x(y)$ is a uniformly continuous function of $(x, y) \in \mathcal{X} \times \mathcal{X}$, all of them defined on some specific gentle subset of $\mathbb{R}^n(\mathbb{R}^{2n})$.

We now recall how Rosenthal [18] chooses a collection of paths. For each $x \neq y$, let $\gamma_{xy} = (x_0, x_1, \ldots, x_k)$ for some $k = |\gamma_{xy}| < \infty$ be a sequence of points such that $x_i \in \mathcal{X}$, $i = 0, \ldots, k$, $x_0 = x$ and $x_k = y$, with $\{x_i\}_{0 \leqslant i \leqslant k}$ distinct, and with $f_{x_i}(x_{i+1}) > 0$ for $0 \leqslant i \leqslant k-1$. For notational convenience we set $\gamma_{xy}(l) = x_l$ for $l \leqslant |\gamma_{xy}|$. We assume that, for each l, γ_{xy} is a $\mathcal{C}^1$ function of (x, y) on a subset of $\mathcal{X} \times \mathcal{X}$ with $\mathcal{C}^1$ boundaries, except on some gentle subset of $\mathbb{R}^n$. Then we assume that there exists a sequence of bounded subsets $S_1 \subseteq S_2 \subseteq \ldots$ with $\mathcal{C}^1$ boundaries and with $\mathcal{X} = \cup_n S_n$, such that, if $x, y \in S_j$ for some j, then $\gamma_{xy}(l) \in S_j$ for all $0 \leqslant l \leqslant |\gamma_{xy}|$: we say that $\{\gamma_{xy}\}$ is *unfolding*. Note that if $\mathcal{X}$ is bounded itself, then this condition is satisfied.

To deal with discontinuities, Rosenthal [18] introduces the following notation. For any function f (possibly undefined at some points or with discontinuities), we let $\lceil f \rceil(x)$ (resp. $\lfloor f \rfloor(x)$) be the lim sup (resp. lim inf) of f. More precisely,

$$\lfloor f \rfloor_z(t) = \lim_{\varepsilon \downarrow 0} \inf_{\|z'-z\| \leqslant \varepsilon; \|t'-t\| \leqslant \varepsilon} f_{z'}(t').$$

When f is continuous, $\lfloor f \rfloor = \lceil f \rceil = f$.

Finally, for $l \leqslant |\gamma_{xy}|$, we define $J_{xy}(l)$ to be the Jacobian of $(x, y) \mapsto (\gamma_{xy}(l), \gamma_{xy}(l+1))$:

$$J_{xy}(l) = \det \begin{pmatrix} \dfrac{\partial}{\partial x_i}(\gamma_{xy}(l))_j & \dfrac{\partial}{\partial x_i}(\gamma_{xy}(l+1))_j \\ \dfrac{\partial}{\partial y_i}(\gamma_{xy}(l))_j & \dfrac{\partial}{\partial y_i}(\gamma_{xy}(l+1))_j \end{pmatrix}.$$

Note that, at some discontinuity points, $\lfloor J_{xy}(l) \rfloor$ and $\lceil J_{xy}(l) \rceil$ are still defined.

The last assumption is that, for any $x, y \in \mathcal{X} \times \mathcal{X}$, we assume that $\lfloor J_{xy}(l) \rfloor > 0$.

The result is the following.

Theorem 5.1. *Let $K(\cdot, \cdot)$ be a transition kernel on an open subset $\mathcal{X}$ of $\mathbb{R}^n$, reversible with respect to the probability measure μ and irreducible with respect to the Lebesgue measure v on $\mathbb{R}^n$. We assume that μ and K satisfy the regularity condition mentioned above. For all $(x, y) \in \mathcal{X} \times \mathcal{X}$, $x \neq y$, choose one path γ_{xy}. Assume that the collection of paths $\{\gamma_{xy}\}$ is unfolding, and satisfies the regularity conditions as above.*

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

Then, the logarithmic Sobolev constant α *defined in* (5.1) *satisfies* $\alpha \leqslant 20A$, *where*

$$A = \sup_{z,t\in\mathcal{X}} (\lfloor h \rfloor(z) \lfloor f \rfloor_z(t))^{-1} \sum_{\substack{x,y,l \\ \gamma_{xy}(l)=z \\ \gamma_{xy}(l+1)=t}} |\gamma_{xy}| \frac{\lceil h \rceil(x) \lceil h \rceil(y) \lceil \ln \frac{1}{h} \rceil(y)}{\lfloor J_{xy}(l) \rfloor}.$$

Proof. We omit the proof because it is almost identical to that given by Rosenthal [18, Theorem 5]. □

Now we give an example of an application.

Example 8. (Metropolis chain on $\mathbb{R}$) We let $\mathcal{X} = \mathbb{R}$ and consider a transition probability $Q(x, dy) = Z(s)e^{-|y-x|^s}dy$ with $Z(s)^{-1} = \int_{\mathbb{R}} e^{-|t|^s}dt$, $s > 1$, reversible with respect to the Lebesgue measure on $\mathbb{R}$.

Given $M > 0$, by a Metropolis algorithm we build a transition kernel

$$K(x, dy) = a_x \delta_x(dy) + f_x(y)v(dy),$$

with

$$f_x(y) = \begin{cases} Z(s)e^{-(|y|^s-|x|^s)_+-|y-x|^s} & \text{for } |y| \leqslant M, \\ 0 & \text{otherwise,} \end{cases}$$

and $a_x = 1 - \int_{-M}^{M} f_x(y)v(dy)$, reversible with respect to the probability measure $\mu(dy) = h(y)v(dy)$, with

$$h(y) = \begin{cases} Z e^{-|y|^s} & \text{for } |y| \leqslant M, \\ 0 & \text{otherwise,} \end{cases}$$

and $Z^{-1} = \int_{-M}^{M} e^{-|y|^s}v(dy)$ (truncated measure). Here $(x)_+ = \max(x, 0)$.

The case $s = 2$ is considered by Rosenthal [18].

We choose paths linearly as follows (see [18]). Given $x, y \in \mathcal{X}$, $x \neq y$, let $L_{xy} = \max(1, \lfloor |y-x| \rfloor)$ be the length of the path (where $\lfloor \cdot \rfloor$ is the entire part). Let $A_{xy} = \frac{y-x}{L_{xy}}$ be the jump rate, and for $0 \leqslant l \leqslant L_{xy} = |\gamma_{xy}|$, let $\gamma_{xy}(l) = x + lA_{xy}$.

It is easy to verify that all the regularity conditions and unfolding assumptions of Theorem 5.1 are satisfied. Moreover,

$$J_{xy}(l) = \det \begin{pmatrix} 1 - \frac{l}{L_{xy}} & 1 - \frac{l+1}{L_{xy}} \\ \frac{l}{L_{xy}} & \frac{l+1}{L_{xy}} \end{pmatrix} = \frac{1}{L_{xy}}.$$

Let $e = (z, t)$ be an edge. Assume that $z > |t|$ (the other case is almost identical, so we omit the computation).

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

First, if $z-t<1$, then there is only one path using e, namely the path γ_{zt}. Notice that $J_{zt}(0)=1$. Therefore we have

$$(h(z)f_z(t))^{-1}\sum_{x,y,l}|\gamma_{xy}|\frac{h(x)h(y)\ln\frac{1}{h(y)}}{J_{xy}(l)}=\frac{h(y)\ln\frac{1}{h(y)}}{f_z(t)}$$
$$=\frac{Z}{Z(s)}\left(\ln\frac{1}{Z}+|t|^s\right)e^{-|t|^s+|t-z|^s}$$
$$\leqslant\frac{Z}{Z(s)}e\left(\ln\frac{1}{Z}+\frac{1}{e}\right).$$

Above, the first sum is over all x,y,l such that $\gamma_{xy}(l)=z$ and $\gamma_{xy}(l+1)=t$.

If $1\leqslant z-t<2$, then the path γ_{xy} goes through the edge e if $x=z+m(z-t)$ and $y=t-n(z-t)$ for some $m,n\in\mathbb{N}$. Hence, if we consider the sum appearing in the definition of the constant A of Theorem 5.1, we have

$$\sum_{x,y,l}|\gamma_{xy}|\frac{h(x)h(y)\ln\frac{1}{h(y)}}{J_{xy}(l)}\leqslant(z-t)^2Z^2\sum_{n,m\in\mathbb{N}}(m+n+1)^2\left(\ln\frac{1}{Z}+|t-n(z-t)|^s\right)$$
$$\times e^{-|z+m(z-t)|^s}e^{-|t-n(z-t)|^s}$$
$$\leqslant k\left(\sum_{n=0}^{\infty}(n+1)^2\left(\ln\frac{1}{Z}+|t-n(z-t)|^s\right)e^{-|t-n(z-t)|^s}\right)$$
$$\times\left(\sum_{m=0}^{\infty}(m+1)^2e^{-|z+m(z-t)|^s}\right)$$

for some numerical constant k.

As $|z+m(z-t)|^s\geqslant|z|^s+m|z-t|^s$, the second sum is bounded by

$$\sum_{m=0}^{\infty}(m+1)^2e^{-|z|^s-m|z-t|^s}\leqslant e^{-|z|^s}\sum_{m=0}^{\infty}(m+1)(m+2)e^{-m|z-t|^s}$$
$$=e^{-|z|^s}\frac{2}{(1-e^{-|z-t|^s})^3}.$$

The first sum is more difficult, since $t-n(z-t)$ may be small for a large value of n. The change of variables $j=n-\lfloor\frac{t}{z-t}\rfloor$ and some manipulation imply that this sum is bounded by

$$k'\left(\left(\frac{z}{z-t}\right)^2+1\right)$$

for some numerical constant k'.

In conclusion,

$$(h(z)f_z(t))^{-1}\sum_{x,y,l}|\gamma_{xy}|\frac{h(x)h(y)\ln\frac{1}{h(y)}}{J_{xy}(l)}\leqslant k''\left(\left(\frac{z}{z-t}\right)^2+1\right)e^{-|z|^s}e^{|z|^s+|z-t|^s}$$
$$\leqslant k'''M^2$$

for some suitable numerical constants k'' and k'''.

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

A similar bound holds for the case $|z| \leqslant t$ and a similar discussion gives the same result for the edge (t, z) instead of (z, t).

Then, Theorem 5.1 gives that the logarithmic Sobolev constant associated with μ satisfies

$$\alpha \leqslant 20k''' M^2.$$

Note that this example successfully bounds the constant A of Theorem 5.1 for an uncountable space. But it is not difficult to see that this bound tends to ∞ as M tends to ∞ (see [18]). Thus, it does not give a bound for unbounded space.

Remark. Note that Theorem 5.1 can be generalized as in the previous sections, introducing, for instance, the notion of weight function. But it seems to us that bound A of Theorem 5.1 is already hard to compute, and as mentioned above, it is not clear that it gives an optimal bound.

Acknowledgement

I warmly thank Laurent Miclo for useful hints and discussions, particularly relating to Section 4. 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] Ané, C., Blachère, S., Chafaï, D., Fougères, P., Gentil, I., Malrieu, F., Roberto, C. and Scheffer, G. (2000) *Sur les Inégalités de Sobolev Logarithmiques*, Société Mathématique de France, Paris.

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

[3] Cheeger, J. (1970) A lower bound for the smallest eigenvalue of the Laplacian. In *Problems in Analysis: Papers Dedicated to Salomon Bochner, 1969*, Princeton University Press, Princeton, NJ, pp. 195–199.

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

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

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

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

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

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

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

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

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press

[12] Miclo, L. (1997) Remarques sur l'hypercontractivité et l'évolution de l'entropie pour des chaînes de Markov finies. In *Séminaire de Probabilités XXXI*, Vol. 1655 of *Lecture Notes in Mathematics*, Springer, Berlin, pp. 136–167.
[13] Miclo, L. (1999) An example of application of discrete Hardy's inequalities. *Markov Process. Rel. Fields* **5** 319–330.
[14] Miclo, L. (1999) Relations entre isopérimetrie et trou spectral pour les chaînes de Markov finies. *Probab. Theor. Rel. Fields* **114** 431–485.
[15] Miclo, L. and Roberto, C. (2000) Trous spectraux pour certains algorithmes de Métropolis sur **R**. In *Séminaire de Probabilités, XXXIV*, Springer, Berlin, pp. 336–352.
[16] Muckenhoupt, B. (1972) Hardy's inequality with weights. *Studia Math.* **44** 31–38.
[17] Rao, M. M. and Ren, Z. D. (1991) *Theory of Orlicz Spaces*, Marcel Dekker Inc., New York.
[18] Rosenthal, J. S. (1996) Markov chain convergence: From finite to infinite. *Stochastic Process. Appl.* **62** 55–72.
[19] Saloff-Coste, L. (1997) Lectures on finite Markov chains. In *Lectures on Probability Theory and Statistics: École d'Été de Probabilités de St-Flour 1996*, Vol. 1665 of *Lecture Notes in Mathematics*, Springer, Berlin, pp. 301–413.
[20] Sinclair, A. (1992) Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combin. Probab. Comput.* **1** 351–370.
[21] Sinclair, A. (1993) *Algorithms for Random Generation and Counting: A Markov Chain Approach*, Birkhäuser, Boston, MA.

https://doi.org/10.1017/S0963548302005515 Published online by Cambridge University Press