

Some drawbacks of finite modified logarithmic Sobolev inequalities

Laurent Miclo

Institut de Mathématiques de Toulouse, UMR 5219
Université de Toulouse and CNRS, France

Abstract

Classically, finite modified logarithmic Sobolev inequalities are used to deduce a differential inequality for the evolution of the relative entropy with respect to the invariant measure. We will check that these inequalities are ill-behaved with respect, on one hand, to the symmetrization procedure, and on the other hand, to the umbrella sampling procedure for Poincaré's inequalities. A short spectral proof of the latter method is given to estimate the spectral gap of a finite reversible Markov generator L in terms of the spectral gap of the restrictions of L on two subsets whose union is the whole state space and whose intersection is not empty.

Keywords: finite irreducible Markov generator, modified logarithmic Sobolev inequality, symmetrization, Poincaré's inequality, umbrella sampling.

MSC2010: primary: 46E39, secondary: 60J27, 94A17, 60E15, 15A18.

1 Introduction

The resort to the study of the evolution of the relative entropy is a traditional technique in the investigation of the convergence of Markov processes to equilibrium. By differentiation with respect to time, one ends up with an entropy dissipation. To compare this term with the relative entropy, one is led to introduce modified logarithmic Sobolev inequalities. The goal of this note is to present some examples of bad behaviors of these inequalities.

More precisely, the setting is as follows: consider V a finite state space endowed with an **Markov generator** matrix $L := (L(x, y))_{x, y \in V}$, namely satisfying

$$\begin{aligned} \forall x \neq y \in V, \quad L(x, y) &\geq 0 \\ \forall x \in V, \quad L(x, x) &= - \sum_{y \in V \setminus \{x\}} L(x, y) \end{aligned}$$

We assume that L is **irreducible**:

$$\forall x, y \in V, \exists n \in \mathbb{Z}_+ : L^n(x, y) > 0$$

Let $\mu := (\mu(x))_{x \in V}$ be the unique **invariant probability measure** for L , i.e. satisfying

$$\forall y \in V, \quad \sum_{x \in V} \mu(x) L(x, y) = 0$$

It charges all the points of V : $\mu(x) > 0$ for all $x \in V$. On $\mathcal{P}(V)$, standing for the **set of probability measures on V** , define the **relative entropy** with respect to μ via

$$\forall m := (m(x))_{x \in V} \in \mathcal{P}(V), \quad \text{Ent}(m|\mu) := \sum_{x \in V} \ln \left(\frac{m}{\mu}(x) \right) m(x)$$

It is a way to measure the discrepancy between m and μ , in particular Pinsker's inequality asserts that the total variation between m and μ is bounded by the square root of twice the relative entropy, cf. e.g. the book of Ané, Blachère, Chafaï, Fougères, Gentil, Malrieu, Roberto, and Scheffer [1].

The **Markov semigroup** $(P_t)_{t \geq 0}$ associated to L is given by

$$\forall t \geq 0, \quad P_t := \exp(tL)$$

and to any initial law $m_0 \in \mathcal{P}(V)$, the corresponding distribution m_t at time $t \geq 0$ is

$$m_t := m_0 P_t$$

It is the law of the position at time $t \geq 0$ of a Markov process generated by L and whose initial state is sampled according to m_0 .

The irreducibility of L (equivalent to the positivity of m_t for any $t > 0$ and any initial distribution m_0) implies that for any given $m_0 \in \mathcal{P}(V)$, m_t converges to μ for large $t \geq 0$. One way to quantify this convergence is to study the evolution of $\text{Ent}(m_t|\mu)$ by differentiating it with respect to time:

$$\forall m_0 \in \mathcal{P}(V), \forall t \geq 0, \quad \partial_t \text{Ent}(m_t|\mu) = -F(f_t) \tag{1}$$

where f_t is the density m_t/μ and where

$$\forall f \in \mathcal{F}_+(V), \quad F(f) := \sum_{x, y \in V} \mu(x) L(x, y) f(x) (\ln(f(x)) - \ln(f(y)))$$

with $\mathcal{F}_+(V)$ standing for the cone of non-negative functions on V . The quantity $F(f)$ is called the **entropy dissipation** and is non-negative (it takes the value $+\infty$ if there exist $x, y \in V$ with $L(x, y) > 0$, $f(x) > 0$ and $f(y) = 0$). To adopt functional notation, define

$$\forall f \in \mathcal{F}_+(V) \setminus \{0\}, \quad E(f) := \text{Ent}(f \cdot \mu / \mu[f] | \mu)$$

where $f \cdot \mu / \mu[f]$ is the probability measure admitting $f / \mu[f]$ as density with respect to μ . To transform (1) into a differential inequality for the evolution of the relative entropy, one introduces the following **modified logarithmic Sobolev inequality**:

$$\forall f \in \mathcal{F}_+(V) \setminus \{0\}, \quad \alpha E(f) \leq F(f)$$

where $\alpha \geq 0$ is the best constant such that this bound holds (hereafter called the **modified logarithmic Sobolev constant** and denoted $\alpha(L)$ when we need to emphasize the underlying generator).

The **symmetrization** $G(f)$ of the entropy dissipation $F(f)$ is defined by

$$\forall f \in \mathcal{F}_+(V), \quad G(f) := \frac{1}{2} \sum_{x, y \in V} \mu(x) L(x, y) (f(x) - f(y)) (\ln(f(x)) - \ln(f(y)))$$

and its interest relies on the **symmetric modified logarithmic Sobolev inequality**

$$\forall f \in \mathcal{F}_+(V) \setminus \{0\}, \quad \beta E(f) \leq G(f)$$

where $\beta \geq 0$ is the best constant such that this bound holds (called the **symmetric modified logarithmic Sobolev constant**). This bound corresponds to the previous modified logarithmic Sobolev inequality, but with L replaced by its **additive symmetrization** $(L + L^*)/2$ in $\mathbb{L}^2(\mu)$: L^* is the adjoint operator of L in $\mathbb{L}^2(\mu)$, which is a Markovian generator, because μ is invariant for L . More explicitly, one computes that

$$\forall x, y \in V, \quad L^*(x, y) = \frac{\mu(y)}{\mu(x)} L(y, x)$$

The invariant probability μ is said to be **reversible** with respect to L when $L^* = L$, i.e.

$$\forall x, y \in V, \quad \mu(x) L(x, y) = \mu(y) L(y, x)$$

In this case, we have

$$\forall f \in \mathcal{F}_+(V), \quad F(f) = G(f) \tag{2}$$

(in Lemma 4 below, we will check that conversely (2) implies that μ is reversible with respect to L). When μ is not reversible with respect to L , the introduction of G and β is an attempt to come back to the reversible situation, since μ is reversible for the Markov generator $(L + L^*)/2$. It is then natural to wonder if it would not be possible to compare the functionals F and G . An easy relation is

$$\forall f \in \mathcal{F}_+(V), \quad F(f) \leq 2G(f)$$

This is a consequence of the non-negativeness of the entropy dissipation F and of the fact that for any $f \in \mathcal{F}_+(V)$, the quantity

$$2G(f) - F(f) = \sum_{x, y \in V} \mu(x) L(x, y) f(y) (\ln(f(y)) - \ln(f(x)))$$

is also non-negative, since it can be viewed as an entropy dissipation for the time-reversed generator L^* . As a consequence, we deduce that $\alpha \leq 2\beta$.

Nevertheless, in view of the above considerations, a reverse bound would be more desirable, unfortunately there is no such relation in general:

Proposition 1 *As soon as $\text{card}(V) \geq 3$, there exists an irreducible Markov generator L on V such that*

$$\inf_{f \in \mathcal{F}_>(V)} \frac{F(f)}{G(f)} = 0$$

where $\mathcal{F}_>(V)$ is set of positive functions on V (considered instead of $\mathcal{F}_+(V)$, just to avoid the value $+\infty$ for F and G).

When $\text{card}(V) = 2$, μ is necessarily reversible with respect to L , so that (2) applies.

It is possible to avoid the comparison of the functionals F and G , in particular through the resort to **logarithmic Sobolev inequalities**. Due to the inequality

$$\forall a, b \in \mathbb{R}_+, \quad 4(\sqrt{b} - \sqrt{a})^2 \leq (b - a)(\ln(b) - \ln(a))$$

we have $\beta \geq 4\gamma$, where γ is the (symmetric) **logarithmic Sobolev constant**, namely the best constant $\gamma \geq 0$ such that

$$\forall f \in \mathcal{F}_+(V) \setminus \{0\}, \quad \gamma E(f) \leq H(\sqrt{f})$$

where H is the **energy** associated to L :

$$\forall f \in \mathcal{F}(V), \quad H(f) := \frac{1}{2} \sum_{x, y \in V} \mu(x) L(x, y) (f(x) - f(y))^2$$

($\mathcal{F}(V)$ is the space of all real functions defined on V). The energy is automatically symmetric and there is no need to consider a non-symmetric version. Indeed, we compute that for any $f \in \mathcal{F}(V)$,

$$\begin{aligned} \sum_{x, y \in V} \mu(x) L(x, y) f(x) (f(x) - f(y)) &= -\mu[fL[f]] \\ &= -\mu[fL^*[f]] \\ &= -\mu\left[f \frac{L + L^*}{2}[f]\right] \\ &= H(f) \end{aligned}$$

In our finite setting, it is well-known that $\gamma > 0$ if and only if L is irreducible, see e.g. the lecture notes of Saloff-Coste [4]. Nevertheless, β can be more convenient than γ , as there are natural examples on denumerable state spaces with $\beta > 0$ while $\gamma = 0$ (see for instance Wu [5]). In fact, if one intends to use the logarithmic Sobolev constant γ , it is pointless to consider the functional G , since it follows from [3] that F and H are easy to compare directly: we always have

$$\forall f \in \mathcal{F}_+(V), \quad F(f) \geq H(f) \tag{3}$$

In particular, this bound implies $2\beta \geq \alpha \geq \gamma$ and so the (symmetric) modified logarithmic Sobolev inequality constants α and β are positive for the irreducible Markov generator L .

In fact, (3) and (2) imply that it would only have been interesting to bound below F in terms of G in the non-reversible situations where it is possible to estimate the modified logarithmic Sobolev constant β without going through the logarithmic Sobolev constant γ .

Another usual way to measure the discrepancy between two probability measures m and μ is the **chi-2 distance** defined by

$$\chi_2(m, \mu) := \sqrt{\sum_{x \in V} \left(\frac{m(x)}{\mu(x)} - 1\right)^2 \mu(x)}$$

The corresponding functional is

$$\begin{aligned} \forall f \in \mathcal{F}(V), \quad D(f) &:= \sqrt{\mu[(f - \mu[f])^2]} \\ &= \sqrt{\frac{1}{2} \sum_{x,y \in V} \mu(x)\mu(y)(f(y) - f(x))^2} \end{aligned}$$

Considering the evolution of the chi-2 distance of the time-marginal laws to equilibrium instead of the relative entropy, one is led to the **Poincaré's inequality**

$$\forall f \in \mathcal{F}(V), \quad \lambda D(f) \leq H(f)$$

where the **spectral gap** $\lambda \geq 0$ is the best possible constant in this bound. There is a general comparison between the logarithmic Sobolev constant and the spectral gap: $2\gamma \leq \lambda$, cf. e.g. the book of Ané, Blachère, Chafaï, Fougères, Gentil, Malrieu, Roberto, and Scheffer [1].

For any positive $\mu \in \mathcal{P}(V)$, consider the generator L_μ defined by

$$\forall x \neq y, \quad L_\mu(x, y) := \mu(y)$$

It is irreducible and μ is its unique invariant measure which is furthermore reversible. It is the generator of the Markov process which, from any initial distribution, wait an exponential time of parameter 1 and then choose as next position a point sampled according to μ . In some sense, this process jumps directly to equilibrium. In this situation, we have $D(f) = H(f)$, for any $f \in \mathcal{F}(V)$ and by consequence $\lambda = 1$.

This observation is the key point in the **intersection method**. It amounts to the following procedure for the Poincaré's inequality. Assume that $V = \tilde{V} \cup \hat{V}$ and that $\tilde{V} \cap \hat{V}$ is the singleton $\{x_0\}$. Let \tilde{L} (respectively \hat{L}) be an irreducible and reversible Markov generator on \tilde{V} (resp. \hat{V}). Denote by $\tilde{\mu}$ and $\tilde{\lambda} > 0$ (resp. $\hat{\mu}$ and $\hat{\lambda} > 0$) the reversible probability measure and the spectral gap of \tilde{L} (resp. \hat{L}). Define $\tilde{\chi} := \tilde{\lambda}\tilde{\mu}(x_0)$ and $\hat{\chi} := \hat{\lambda}\hat{\mu}(x_0)$. Consider $L = \tilde{L} + \hat{L}$, namely the Markovian generator given by

$$\forall x, y \in V, \quad L(x, y) := \begin{cases} \tilde{L}(x, y) & , \text{ if } x, y \in \tilde{V} \\ \hat{L}(x, y) & , \text{ if } x, y \in \hat{V} \end{cases}$$

Theorem 2 *The Markov generator L is irreducible and reversible and its spectral gap λ satisfies*

$$\lambda \geq \min(\tilde{\lambda}, \hat{\lambda}, \tilde{\chi} + \hat{\chi} - \sqrt{\tilde{\chi}^2 + \hat{\chi}^2 - \tilde{\chi}\hat{\chi}}) > 0$$

The hypothesis that $\tilde{V} \cap \hat{V}$ is a singleton can be relaxed, for instance the proof of Theorem 2 can be extended immediately to the situation where the restrictions of $\tilde{\mu}$ and $\hat{\mu}$ on $\tilde{V} \cap \hat{V}$ are proportional. Then in the definition of $\tilde{\chi}$ and $\hat{\chi}$, $\tilde{\mu}(x_0)$ and $\hat{\mu}(x_0)$ have to be replaced respectively by $\tilde{\mu}(\tilde{V} \cap \hat{V})$ and $\hat{\mu}(\tilde{V} \cap \hat{V})$. More generally, the intersection method is a particular case of the procedure of umbrella sampling described in Madras and Randall [2]. There, the authors start with a reversible Markov transition kernel P and relate its spectral gap to the spectral gaps of the restriction of P to several subsets and to the spectral gap of another transition kernel standing for the motions between the subsets. Theorem 2 is more precise than Theorem 1.1 from Madras and Randall [2], because we will encapsulate the spectral gaps $\tilde{\lambda}$ and $\hat{\lambda}$ of the subsets \tilde{V} and \hat{V} into the definition of the generator describing the motions between them. This slight improvement could be extended to the setting of Madras and Randall [2].

However, our goal here is to give a straightforward spectral proof of Theorem 2 and to show that the relative entropy does not follow the same pattern: when $L = L_\mu$, we compute that

$$\begin{aligned} \forall f \in \mathcal{F}_+(V), \quad F(f) &= E(f) - \mu[\ln(f)] \\ &\geq E(f) \end{aligned}$$

where we used the Jensen's inequality with respect to the convex function $(0, +\infty) \ni u \mapsto -\ln(u)$ to deduce the last inequality. In particular we get $\alpha(L_\mu) \geq 1$.

By analogy with the Poincaré's inequality, one can wonder if there exists a constant $\kappa > 0$ such that the following **bi-modified logarithmic Sobolev inequality** holds:

$$\forall f \in \mathcal{F}_+(V), \quad \kappa E_m(f) \leq F(f) \quad (4)$$

where the **modified relative entropy** $E_m(f)$ is the quantity $E(f) - \mu[\ln(f)]$ (since $E \leq E_m$, of course we have $\kappa \leq \alpha$).

Unfortunately, it is often not possible:

Proposition 3 *The best constant κ in (4) satisfies*

$$\kappa \leq \min_{x \neq y \in V} \frac{L(y, x)}{\mu(x)}$$

In particular it is non-null if and only if the transition graph of L is the complete graph on V .

The latter drawback prevents the obtention of (bi-)modified logarithmic Sobolev inequalities by an intersection method, whose crucial observation was that $D = H$ for $L = L_\mu$. Of course (4) would have not been satisfactory in itself, since $F(f)$ should be replaced by the dissipation functional associated to $E_m(f)$, namely

$$\forall f \in \mathcal{F}_>(V), \quad \sum_{x, y \in V} \mu(x) L(x, y) f(x) \left(\ln(f(x)) - \ln(f(y)) + \frac{1}{f(y)} - \frac{1}{f(x)} \right)$$

leading to a new type of modified logarithmic Sobolev inequality. But we will not push further in this direction here.

2 Proofs and examples

Here we check the results presented in the introduction, via the exhibit of appropriate examples for Propositions 1 and 3.

Let us begin with the assertion made after (2):

Lemma 4 *The identity (2) is satisfied if and only if μ is reversible with respect to L .*

Proof

Fix $x_0 \in V$ and consider a function $U \in \mathcal{F}(V)$ such that $U(x_0) = 0 > \max\{U(x) : x \in V \setminus \{x_0\}\}$. For $r \geq 0$, define $f_r \in \mathcal{F}(V)$ via

$$\forall x \in V, \quad f_r(x) := \exp(rU(x))$$

Letting r go to $+\infty$ in $F(f_r) = G(f_r)$, we get

$$-\sum_{y \in V} \mu(x_0) L(x_0, y) U(y) = -\frac{1}{2} \left(\sum_{y \in V} \mu(x_0) L(x_0, y) U(y) + \sum_{y \in V} \mu(y) L(y, x_0) U(y) \right)$$

i.e.

$$\sum_{y \in V} \mu(x_0) L(x_0, y) U(y) = \sum_{y \in V} \mu(y) L(y, x_0) U(y)$$

Fix another point $x_1 \in V \setminus \{x_0\}$ and let $U(x_1)$ go to $-\infty$, while letting the other other values of U fixed: it follows that $\mu(x_0)L(x_0, x_1) = \mu(x_1)L(x_1, x_0)$. Since this is true for all $x_0 \neq x_1 \in V$, we get that μ is reversible with respect to L . ■

Next we find an example leading to the assertion of Proposition 1.

Proof of Proposition 1

We begin with the case where $V := \mathbb{Z}_3$. Consider the irreducible and Markov generator L given by

$$\forall x \neq y \in \mathbb{Z}_3, \quad L(x, y) := \begin{cases} 1 & , \text{ if } y = x + 1 \\ 0 & , \text{ otherwise} \end{cases}$$

Its invariant probability measure is the uniform measure μ on \mathbb{Z}_3 .

For $r > 1$, let $f_r \in \mathcal{F}_>(\mathbb{Z}_3)$ defined by

$$f_r(0) = 1, \quad f_r(1) = r, \quad f_r(2) = r/\ln(r)$$

We compute that

$$\begin{aligned} F(f_r) &= \sum_{x \in \mathbb{Z}_3} \mu(x)L(x, x+1)f_r(x)(\ln(f_r(x)) - \ln(f_r(x+1))) \\ &= \frac{1}{3} \sum_{x \in \mathbb{Z}_3} f_r(x)(\ln(f_r(x)) - \ln(f_r(x+1))) \\ &= \frac{1}{3} \left(-\ln(r) + r \left(\ln(r) - \ln\left(\frac{r}{\ln(r)}\right) \right) + \frac{r}{\ln(r)} \ln\left(\frac{r}{\ln(r)}\right) \right) \\ &= \frac{1}{3} \left(-\ln(r) + r \ln(\ln(r)) + \frac{r}{\ln(r)}(\ln(r) - \ln(\ln(r))) \right) \\ &\sim r \ln(\ln(r))/3 \end{aligned}$$

as $r > 1$ goes to $+\infty$.

Similarly, we have

$$\begin{aligned} G(f_r) &= \sum_{x \in \mathbb{Z}_3} \mu(x)L(x, x+1)(f_r(x) - f_r(x+1))(\ln(f_r(x)) - \ln(f_r(x+1))) \\ &= \frac{1}{3} \sum_{x \in \mathbb{Z}_3} (f_r(x) - f_r(x+1))(\ln(f_r(x)) - \ln(f_r(x+1))) \\ &\geq \frac{1}{3}(f_r(0) - f_r(1))(\ln(f_r(0)) - \ln(f_r(1))) \\ &= (1 - r)(0 - \ln(r))/3 \\ &\sim r \ln(r)/3 \end{aligned}$$

as $r > 1$ goes to $+\infty$. In particular, we get

$$\lim_{r \rightarrow +\infty} \frac{F(f_r)}{G(f_r)} = 0$$

To get the same result on any finite set V with $\text{card}(V) \geq 4$, choose two points $x_0 \neq x_1$ in V and consider the irreducible and Markov generator L given by

$$\forall x \neq y \in V, \quad L(x, y) := \begin{cases} 1/(\text{card}(V) - 2) & , \text{ if } x = x_0 \text{ and } y \notin \{x_0, x_1\} \\ 1 & , \text{ if } x \notin \{x_0, x_1\} \text{ and } y = x_1 \\ 1 & , \text{ if } x = x_1 \text{ and } y = x_0 \\ 0 & , \text{ otherwise} \end{cases}$$

By considering functions which are constant on $V \setminus \{x_0, x_1\}$, we are brought back to the previous situation on \mathbb{Z}_3 . ■

We now come to the positive result about Poincaré's inequality.

Proof of Theorem 2

The Markov generator L is clearly irreducible: from x_0 , the transitions of \tilde{L} (respectively \hat{L}) enable to rejoin any point of \tilde{V} (resp. \hat{V}) and to come back to x_0 . Furthermore, from the reversibility of $\tilde{\mu}$ and $\hat{\mu}$ with respect to \tilde{L} and \hat{L} , it appears that if we can find a probability measure μ on V which is proportional to $\tilde{\mu}$ (resp. $\hat{\mu}$) on \tilde{V} (resp. \hat{V}), then μ is reversible with respect to L (and in particular is invariant for L). Indeed, this is a consequence of the fact that if $L(x, y) > 0$ then x, y are both belonging to \tilde{V} or to \hat{V} . We are thus looking for three positive constants \tilde{a} , \hat{a} and b such that

$$\mu := \tilde{a}\tilde{\mu} + \hat{a}\hat{\mu} - b\delta_{x_0}$$

is a probability measure satisfying

$$\mu(x_0) = \tilde{a}\tilde{\mu}(x_0) \quad \text{and} \quad \mu(x_0) = \hat{a}\hat{\mu}(x_0)$$

These equalities lead to

$$\tilde{a}\tilde{\mu}(x_0) = b = \hat{a}\hat{\mu}(x_0) \tag{5}$$

and since we must also have $\tilde{a} + \hat{a} - b = 1$, we deduce that the solution to this problem is

$$\begin{aligned} \tilde{a} &= \frac{\hat{\mu}(x_0)}{\tilde{\mu}(x_0) + \hat{\mu}(x_0) - \tilde{\mu}(x_0)\hat{\mu}(x_0)} \\ \hat{a} &= \frac{\tilde{\mu}(x_0)}{\tilde{\mu}(x_0) + \hat{\mu}(x_0) - \tilde{\mu}(x_0)\hat{\mu}(x_0)} \end{aligned}$$

(and b given by (5)).

By definition of the energy H associated to L , we have for any $f \in \mathcal{F}(V)$,

$$\begin{aligned} H(f) &:= \frac{1}{2} \sum_{x \neq y \in V} \mu(x)L(x, y)(f(x) - f(y))^2 \\ &= \frac{1}{2} \sum_{x \neq y \in \tilde{V}} \mu(x)L(x, y)(f(x) - f(y))^2 + \frac{1}{2} \sum_{x \neq y \in \hat{V}} \mu(x)L(x, y)(f(x) - f(y))^2 \\ &= \tilde{a}\tilde{H}(f) + \hat{a}\hat{H}(f) \end{aligned}$$

where \tilde{H} (resp. \hat{H}) is the energy associated to \tilde{L} (resp. \hat{L}) and $\tilde{H}(f)$ is standing for \tilde{H} applied to the restriction of f to \tilde{V} .

By definition of the spectral gaps $\tilde{\lambda}$ and $\hat{\lambda}$, we have

$$\tilde{\lambda}\tilde{D}(f) \leq \tilde{H}(f) \quad \text{and} \quad \hat{\lambda}\hat{D}(f) \leq \hat{H}(f)$$

so that

$$\begin{aligned} \forall f \in \mathcal{F}(V), \quad H(f) &\geq \tilde{a}\tilde{\lambda}\tilde{D}(f) + \hat{a}\hat{\lambda}\hat{D}(f) \\ &= H_K(f) \end{aligned}$$

where H_K is the energy associated to the Markov generator K defined by

$$\forall x \neq y \in V, \quad K(x, y) := \begin{cases} \tilde{\lambda}\tilde{\mu}(y) & , \text{ if } x \in \tilde{V} \text{ and } y \in \tilde{V} \\ \hat{\lambda}\hat{\mu}(y) & , \text{ if } x \in \hat{V} \text{ and } y \in \hat{V} \end{cases}$$

It is immediate to check that μ is also reversible for K . Let θ be the spectral gap of K . From the above considerations, we have

$$\forall f \in \mathcal{F}(V), \quad H(f) \geq \theta D(f)$$

namely $\lambda \geq \theta$. To prove Theorem 2, it remains to show that

$$\theta \geq \min(\tilde{\lambda}, \hat{\lambda}, \tilde{\chi} + \hat{\chi} - \sqrt{\tilde{\chi}^2 + \hat{\chi}^2 - \tilde{\chi}\hat{\chi}}) \quad (6)$$

To go in this direction, we compute that for any $f \in \mathcal{F}(V)$,

$$\forall x \in V, \quad K[f](x) = \begin{cases} \tilde{\lambda}\tilde{\mu}[f] - \tilde{\lambda}f(x) & , \text{ if } x \in \tilde{V} \setminus \{x_0\} \\ \hat{\lambda}\hat{\mu}[f] - \hat{\lambda}f(x) & , \text{ if } x \in \hat{V} \setminus \{x_0\} \\ \tilde{\lambda}\tilde{\mu}[f] + \hat{\lambda}\hat{\mu}[f] - (\tilde{\lambda} + \hat{\lambda})f(x_0) & , \text{ if } x = x_0 \end{cases}$$

Let W be the vector subspace of \mathcal{F} consisting of the functions which are constant on $\tilde{V} \setminus \{x_0\}$ and constant on $\hat{V} \setminus \{x_0\}$. From the above expression, W is left stable by K , so by reversibility, the same is true for W^\perp , its orthogonal complement in $\mathbb{L}^2(\mu)$. A function $f \in \mathcal{F}(V)$ belongs to W^\perp if and only if $\mu[f\mathbb{1}_{V \setminus \{x_0\}}] = \mu[f\mathbb{1}_{V \setminus \{x_0\}}] = f(x_0) = 0$. In particular, for $f \in W^\perp$, we have

$$\forall x \in V, \quad K[f](x) = \begin{cases} -\tilde{\lambda}f(x) & , \text{ if } x \in \tilde{V} \setminus \{x_0\} \\ -\hat{\lambda}f(x) & , \text{ if } x \in \hat{V} \setminus \{x_0\} \\ 0 & , \text{ if } x = x_0 \end{cases}$$

It follows that $\tilde{\lambda}$ and $\hat{\lambda}$ are the only possible eigenvalues of the restriction of $-K$ to W^\perp (and they are indeed eigenvalues, as soon as $\text{card}(\tilde{V}) \geq 3$ and $\text{card}(\hat{V}) \geq 3$ respectively).

In the basis $(\mathbb{1}_{\{x_0\}}, \mathbb{1}_{\tilde{V} \setminus \{x_0\}}, \mathbb{1}_{\hat{V} \setminus \{x_0\}})$ of W , the matrix associated to K has the form

$$\begin{pmatrix} -\tilde{\chi} - \hat{\chi} & \tilde{\chi} & \hat{\chi} \\ \tilde{\chi} & -\tilde{\chi} & 0 \\ \hat{\chi} & 0 & -\hat{\chi} \end{pmatrix}$$

Thus the eigenvalues of the restriction of K to W are exactly the eigenvalues of this matrix. The characteristic polynomial of the opposite of this matrix is

$$X^3 - 2(\tilde{\chi} + \hat{\chi})X^2 + 3\tilde{\chi}\hat{\chi}X$$

whose roots are 0 and $X_\pm := \tilde{\chi} + \hat{\chi} \pm \sqrt{\tilde{\chi}^2 + \hat{\chi}^2 - \tilde{\chi}\hat{\chi}}$. As a consequence, we get, when $\tilde{\lambda}$ and $\hat{\lambda}$ are eigenvalues of $-K|_{W^\perp}$,

$$\begin{aligned} \theta &= \min\{\tilde{\lambda}, \hat{\lambda}, X_-, X_+\} \\ &= \min\{\tilde{\lambda}, \hat{\lambda}, X_-\} \end{aligned}$$

When $\tilde{\lambda}$ or $\hat{\lambda}$ is not an eigenvalue of $-K|_{W^\perp}$, we only end up with the lower bound (6), thus always valid as announced. ■

Remark 5 Assume for instance that $\tilde{\chi} \leq \hat{\chi}$. Using the bound $\tilde{\chi}^2 \leq \tilde{\chi}\hat{\chi} \leq \hat{\chi}^2$ in the definition of X_- , we find that

$$\tilde{\chi} \leq X_- \leq \hat{\chi}$$

with strict inequalities when $\tilde{\chi} < \hat{\chi}$. Thus if furthermore $\tilde{\mu}(x_0)$ is sufficiently close to 1, we can end up with $\tilde{\lambda} < X_-$. In more ‘‘typical’’ situations where $\tilde{\mu}(x_0)$ and $\hat{\mu}(x_0)$ are quite small, we will get that $X_- < \min(\tilde{\lambda}, \hat{\lambda})$.

Finally, we come to the last assertion of the introduction.

Proof of Proposition 3

Fix $x \in V$ and $f \in \mathcal{F}(V \setminus \{x\})$ such that $\mu[f \mathbf{1}_{V \setminus \{x\}}] = 1$. For any $\epsilon > 0$, define $f_\epsilon \in \mathcal{F}_+(V)$ via

$$\forall y \in V, \quad f_\epsilon(y) := \begin{cases} f(y)/Z_\epsilon & , \text{ if } y \neq x \\ \epsilon/Z_\epsilon & , \text{ if } y = x \end{cases}$$

where $Z_\epsilon := 1 + \mu(x)\epsilon$ is the normalization such that $\mu[f_\epsilon] = 1$. Letting ϵ go to 0_+ , we get

$$E_m[f_\epsilon] \sim \mu[\ln[1/f_\epsilon]] \sim \mu(x) \ln(1/\epsilon)$$

and

$$F[f_\epsilon] \sim \sum_{y \neq x} \mu(y) L(y, x) f[y] \ln(1/\epsilon)$$

It follows that

$$\begin{aligned} \kappa &\leq \inf \left\{ \frac{\sum_{y \neq x} \mu(y) L(y, x) f[y]}{\mu(x)} : f \in \mathcal{F}(V \setminus \{x\}) \text{ with } \mu[f \mathbf{1}_{V \setminus \{x\}}] = 1 \right\} \\ &= \min_{y \in V \setminus \{x\}} \frac{L(y, x) f[y]}{\mu(x)} \end{aligned}$$

whence the announced result, since x was arbitrary chosen. ■

Acknowledgments:

I'm grateful to José Alfredo Cañizo for discussions about functional inequalities, which motivated this paper, and Arnaud Guillin for pointing out the paper of Madras and Randall [2]. I'm also thankful to the ANR STAB (Stabilité du comportement asymptotique d'EDP, de processus stochastiques et de leurs discrétisations : 12-BS01-0019) for its support, as well as to the hospitality of the Institut Mittag-Leffler where this work was carried out.

References

- [1] Cécile Ané, Sébastien Blachère, Djalil Chafaï, Pierre Fougères, Ivan Gentil, Florent Malrieu, Cyril Roberto, and Grégory Scheffer. *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses [Panoramas and Syntheses]*. Société Mathématique de France, Paris, 2000. With a preface by Dominique Bakry and Michel Ledoux.
- [2] Neal Madras and Dana Randall. Markov chain decomposition for convergence rate analysis. *Ann. Appl. Probab.*, 12(2):581–606, 2002.
- [3] Laurent 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.
- [4] Laurent Saloff-Coste. Lectures on finite Markov chains. In *Lectures on probability theory and statistics (Saint-Flour, 1996)*, volume 1665 of *Lecture Notes in Math.*, pages 301–413. Springer, Berlin, 1997.

- [5] Liming Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. *Probab. Theory Related Fields*, 118(3):427–438, 2000.

miclo@math.univ-toulouse.fr
Institut de Mathématiques de Toulouse
Université Paul Sabatier
118, route de Narbonne
31062 Toulouse Cedex 9, France