Combinatorics, Probability and Computing (2018) 27, 334–357.
doi:10.1017/S0963548318000044

On Quantitative Noise Stability and Influences for Discrete and Continuous Models

RAPHAËL BOUYRIE

Institut de Mathématiques de Toulouse, UMR 5219 du CNRS Université de Toulouse, 31062, Toulouse, France
(e-mail: raphael.bouyrie@gmail.com)

Received 23 April 2013; revised 30 November 2017; first published online 22 March 2018

Keller and Kindler recently established a quantitative version of the famous Benjamini–Kalai–Schramm theorem on the noise sensitivity of Boolean functions. Their result was extended to the continuous Gaussian setting by Keller, Mossel and Sen by means of a Central Limit Theorem argument. In this work we present a unified approach to these results, in both discrete and continuous settings. The proof relies on semigroup decompositions together with a suitable cut-off argument, allowing for the efficient use of the classical hypercontractivity tool behind these results. It extends to further models of interest such as families of log-concave measures and Cayley and Schreier graphs. In particular we obtain a quantitative version of the Benjamini–Kalai–Schramm theorem for the slices of the Boolean cube.

2010 *Mathematics subject classification*: Primary 60C05
Secondary 05D40

1. Introduction

The notion of influences of variables on Boolean functions has been extensively studied over the last twenty years, with applications in various areas such as combinatorics, statistical physics and theoretical computer science, in particular cryptography and computational lower bounds (see *e.g.* the survey [13]). Similarly, the noise sensitivity of a Boolean function is a measure of how its values are likely to change under a slightly perturbed input. Noise sensitivity has become an important concept which is useful in many fields, for instance percolation, in which field it was originally defined in [6]. For an overview of the topic, see *e.g.* the book [11], in which the noise sensitivity concept and the results of [6] are presented in the context of percolation theory.

Noise sensitivity and influences are closely related. In this work, we will be concerned with recent connections between influences and asymptotic noise sensitivity. To start with, let us recall these two important concepts on the discrete cube $\{-1,1\}^n$. Rather than noise sensitivity, we describe the related notion of noise stability. Throughout this paper, let ν_p denote the product measure $(p\delta_{-1}+(1-p)\delta_1)^{\otimes n}$ on $\{-1,1\}^n$, where $p\in(0,1)$.

Let $A\subset\{-1,1\}^n$. For $x=(x_1,\dots,x_n)\in\{-1,1\}^n$ and $i=1,\dots,n$, let $\tau_i x\in\{-1,1\}^n$ be the vector obtained from x by changing x_i in $-x_i$ and leaving the other coordinates unchanged. The *influence* of the ith coordinate on a function $f:\{-1,1\}^n \to \mathbb{R}$ is given by

$$I_i(f) = \|f(x) - f(\tau_i x)\|_{L^1(\nu_p)}.$$

Similarly, for sets $A \subset \{-1,1\}^n$, the influence of the ith coordinate is defined using characteristic functions by $I_i(A) = I_i(\mathbf{1}_A)$. Notice then that in the uniform case ($p = 1/2$), if we denote the shifted set in the ith direction by $\tau_i(A) = \{x \in \{-1,1\}^n,\ \tau_i x \in A\}$, we obtain

$$I_i(A) = \frac{1}{2^n}|A\Delta\tau_i(A)|,$$

where Δ denotes the symmetric difference between two sets.

Turning to noise stability, let $\eta \in (0,1)$ and let $X = (X_1,\dots,X_n)$ be distributed according to ν on $\{-1,1\}^n$. Let $X^\eta = (X_1^\eta,\dots,X_n^\eta)$ be a $(1-\eta)$-correlated copy of X, that is, $X_j^\eta = X_j$ with probability $1-\eta$ and $X_j^\eta = X_j'$ with probability η where X' is an independent copy of X. For a function $f:\{-1,1\}^n \to \mathbb{R}$, following the notations of [16], define its *noise stability* as

$$\mathcal{S}_\eta^c(f) = \mathbb{E}_\nu[f(X)f(X^\eta)] - \mathbb{E}_\nu[f(X)]^2.$$

Similarly, for a subset $B \subset \{-1,1\}^n$, define its noise stability by $\mathcal{S}_\eta^c(B) := \mathcal{S}_\eta^c(\mathbf{1}_B)$. Thus, the noise stability of a function $\mathcal{S}_\eta^c(f)$ is a measure of the sensitivity of f to a small noise η in its input. A sequence of functions/sets $(f_{n_\ell})_{n_\ell}, (B_{n_\ell})_{n_\ell}$ over $\{-1,1\}^{n_\ell}$, $\ell \in \mathbb{N}$, where $n_\ell \nearrow \infty$, is said to be *(asymptotically) noise sensitive* if its noise stability tends to 0 as ℓ tends to ∞, that is,

$$\lim_{\ell\to\infty} \mathcal{S}_\eta^c(f_{n_\ell}) = 0$$

or

$$\lim_{\ell\to\infty} \mathcal{S}_\eta^c(B_{n_\ell}) = 0,$$

for each $\eta \in (0,1)$ (keeping in mind η small but fixed).

The first connection between noise sensitivity and influences was established by Benjamini, Kalai and Schramm [6]. They gave a criterion for a sequence of sets to be noise sensitive in terms of the sum of squares of the influences. More precisely, one of the main results of [6] is the following theorem.

Theorem 1.1. *Let $B_\ell \subset \{-1,1\}^{n_\ell}$, $\ell \in \mathbb{N}$. If*

$$\lim_{\ell\to\infty} \sum_{i=1}^{n_\ell} I_i(B_\ell)^2 = 0, \tag{1.1}$$

then $(B_\ell)_{\ell\in\mathbb{N}}$ is asymptotically noise sensitive.

In fact, Benjamini, Kalai and Schramm only proved this theorem for the uniform measure (*i.e.* in the case $p = 1/2$). The general result, including the case of biased measure (assuming the bias p does not depend on n), was proved recently by Keller and Kindler [14], who moreover established a quantitative version of this result. The proof of this quantitative bound uses the Fourier–Walsh expansion of functions of the discrete cube – and the expression of the noise stability in the Fourier–Walsh basis (see Section 2.2). Then, hypercontractivity is used to control the Fourier weights of a function in terms of the sum of the square of its influences (hypercontractivity is also at the root of the asymptotic Theorem 1.1). The result is stated below as (1.2), but first we present the analogous continuous version which has a similar form.

The quantitative version of [14] was extended to the continuous Gaussian setting by Keller, Mossel and Sen [16]. To state the result, we need to introduce the corresponding definitions. Let

$$d\mu(x) = e^{-|x|^2/2} \frac{dx}{(2\pi)^{n/2}}$$

be the canonical Gaussian measure on $\mathbb{R}^n$. For W, W' independent with distribution μ, and $\eta > 0$, set

$$W^\eta = \sqrt{1-\eta^2}\, W + \eta W',$$

so that (W, W^η) is a $\sqrt{1-\eta^2}$-correlated Gaussian vector. For $f : \mathbb{R}^n \to \mathbb{R}$, if $\|f(W)\|_{L^2(\mu)} < \infty$, define the Gaussian noise stability of f as

$$\mathcal{S}^{\mathcal{G}}_\eta(f) = \mathbb{E}_\mu[f(W)f(W^\eta)] - \mathbb{E}_\mu[f(W)]^2.$$

Similarly, for a (Borel measurable) subset $A \subset \mathbb{R}^n$, set

$$\mathcal{S}^{\mathcal{G}}_\eta(A) = \mathcal{S}^{\mathcal{G}}_\eta(\mathbf{1}_A).$$

In order to discuss the analogue of Theorem 1.1 in the Gaussian case, it is necessary to define the influence of the ith coordinate on a subset $A \subset \mathbb{R}^n$ in this context. To this end, in a continuous setting, Keller, Mossel and Sen [16] introduce the notion of geometric influence defined by

$$I^{\mathcal{G}}_i(A) = \mathbb{E}_x[\mu^+(A^x_i)].$$

In the latter expression, $A^x_i \subset \mathbb{R}$ is the restriction of A along the fibre of $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, that is,

$$A^x_i = \{y \in \mathbb{R}, (x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) \in A\},$$

and μ^+ denotes the lower Minkowski content, that is, for any Borel measurable set $D \subset \mathbb{R}$,

$$\mu^+(D) = \liminf_{r\to 0} \frac{\mu(D + [-r, r]) - \mu(D)}{r}$$

(where μ is the standard Gaussian distribution on $\mathbb{R}$).

From a more intuitive point of view, for each $i \in \{1, \ldots, n\}$, $I^{\mathcal{G}}_i(A)$ is obtained as a limit of $\|\partial_i f_{\varepsilon_k}\|_{L^1(\mu)}$ for a sequence of smooth functions $(f_{\varepsilon_k})_{\varepsilon_k \geqslant 0}$ such that $\lim_{\varepsilon_k\to 0} f_{\varepsilon_k} = \mathbf{1}_A$.

We refer the reader to [15, 16] for further developments on geometric influences and their applications.

In both the cube (with bias p) and Gaussian settings, the quantitative noise sensitivity bounds of [14] and [16] may then be expressed in the following form: for any sets $A \subset \{-1, 1\}^n$, and any $\eta \in (0, 1)$,

$$\mathcal{S}^c_\eta(A) \leqslant 20 \left(\sum_{i=1}^n I_i(A)^2 \right)^{C(p)\eta} \tag{1.2}$$

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

for the discrete cube where $C(p) > 0$ depends only on p, and for any sets $A \subset \mathbb{R}^n$ and any $\eta \in (0,1)$,

$$\mathcal{S}_\eta^{\mathcal{G}}(A) \leqslant 80\left(\sum_{i=1}^{n}(I_i^{\mathcal{G}}(A))^2\right)^{\eta^2/10} \tag{1.3}$$

for the Gaussian space (the numerical constants are not sharp). These results are actually proved in their functional forms, that is, for any bounded (by 1) function f defined respectively in the discrete cube and in the Gaussian space (C^1 smooth in the latter case), we have

$$\mathcal{S}_\eta^{c}(f) \leqslant 20\left(\sum_{i=1}^{n} I_i(f)^2\right)^{C(p)\eta} \quad \text{and} \quad \mathcal{S}_\eta^{\mathcal{G}}(f) \leqslant 80\left(\sum_{i=1}^{n}\|\partial_i f\|_1^2\right)^{\eta^2/10}.$$

The results in [14] and [16] in fact extend to functions in $L^2(\mu)$. Indeed, the proof of the Gaussian result in [16] relies on the result proved in [14] for the discrete cube together with an appropriate Central Limit Theorem argument. Close inspection of the arguments from [14] reveals that the Fourier–Walsh decomposition approach may be adapted to a Fourier–Hermite decomposition in the Gaussian case to yield the same conclusion. Therefore, the boundedness of the functions can be weakened to obtain the same result for functions in $L^2(\mu)$.

The Fourier decomposition approach, however, is somewhat limited to the examples of the discrete cube and the Gaussian space. Indeed, a key property is the fact that the elements of the orthogonal basis of the underlying space are eigenvectors of the corresponding semigroup. Spaces satisfying such a property have been characterized in [19].

In this paper we develop a new, simpler proof of the quantitative relationships (1.2) and (1.3), which moreover applies to more general examples. Once again, the main ingredient will be hypercontractivity, as in [6, 14, 16]. The starting point of the proof follows closely the work of [8], which generalizes Talagrand's inequality of [22] for more general models. The framework of [8] applies in a rather general context where hypercontractivity holds together with specific commutation properties. It covers the product of (strictly) log-concave probability measures and discrete examples as well, including the biased discrete cube or, more generally, Schreier or Cayley graphs.

In this framework, we will establish quantitative relationships between noise stability and influences, thereby strengthening the results of [6, 14, 16], under furthermore weaker assumptions. The proposed simpler and more efficient proof relies on semigroup decompositions and cut-off arguments.

The general setting contains two main illustrations: probability measures on finite state spaces that are invariant for some Markov kernel, and continuous product probability measures on $\mathbb{R}^n$ where each measure is of the form $d\mu(x) = e^{-V(x)}dx$ for V a smooth potential. The following is a sample illustration of the results of this work (in the continuous setting). More complete statements will be presented during the course of the paper.

Theorem 1.2. *Let $(\mathbb{R}^n, \mu^{\otimes n})$, with μ a probability measure on $\mathbb{R}$ of the form $d\mu(x) = e^{-v(x)}dx$, where $v'' \geqslant c$ uniformly for some $c > 0$. Let $\mathcal{S}_\eta$ denote the noise stability with respect to η in this context. Then, for any $\eta \in [0,1)$, and any Borel measurable set $A \in \mathbb{R}^n$, there exist positive*

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

constants C, c_1 depending only on c such that

$$\mathcal{S}_\eta(A) \leqslant C\left(\sum_{i=1}^n (I_i^{\mathcal{G}}(A))^2\right)^{c_1\eta^2} \mu^{\otimes n}(A)^{2-2c_1\eta^2}.$$

The generalized definition of the noise stability in this context will be given in the next section. In particular, for the standard Gaussian measure, $\mathcal{S}_\eta = \mathcal{S}_\eta^{\mathcal{G}}$, so we recover the quantitative estimate (1.3) established by Keller, Mossel and Sen.

With a common scheme of proof, an analogous statement is established for discrete models covering in particular the discrete cube endowed with any biased measure $(p\delta_{-1} + (1-p)\delta_1)^{\otimes n}$. The precise setting and the corresponding notion of influence will be presented in Section 2.3.

Theorem 1.3. *Let $(\Omega,\mu) = (\Omega_1 \times \cdots \times \Omega_n, \mu_1 \otimes \cdots \otimes \mu_n)$ be a Cartesian product of finite probability spaces. Let $\mathcal{S}_\eta$ denote the noise stability with respect to η in this context. Then, for any $\eta \in [0,1)$, and any set $A \subset \Omega$, and some absolute constants $C, c_1 > 0$,*

$$\mathcal{S}_\eta(A) \leqslant C\left(\sum_{i=1}^n (I_i(A))^2\right)^{c_1(\rho/\lambda)\eta} \mu(A)^{2-2(\rho/\lambda)c_1\eta},$$

where λ and ρ are respectively the spectral gap and the Sobolev logarithmic constants of the product graph Ω.

The definitions of λ and ρ as well as $\mathcal{S}_\eta$ in this general context will be given in Sections 2.4 and 2.5. In particular, if $\Omega = \{-1,1\}^n$, $\mu = \nu_p$, $\mathcal{S}_\eta = \mathcal{S}_\eta^c$ and we will see that $\lambda = 1$ and

$$\rho = 2\frac{\log p - \log(1-p)}{2p-1}.$$

Therefore, the above theorem extends the quantitative estimate (1.2) of Keller and Kindler.

This paper is organized as follows. In Section 2, we first describe a convenient abstract framework and recall some basic facts about Markov semigroups that will be used in the proofs of our results. Then, in Sections 2.1, 2.2 and 2.3, after developing the examples of the Gaussian space and the discrete cube, we present the general setting into which our results fall. Then, we present our generalized definition of noise stability in Section 2.4, and in Section 2.5 we describe further tools required in the proofs of our results, in particular the hypercontractive property. In Section 3 we establish our main result on $(\mathbb{R}^n, d\mu(x) = \bigotimes_{i=1}^m e^{-V_i(x)}dx)$ when the potentials V_i are convex, which represents a generalization of Theorem 1.2. Section 4 is devoted to the discrete case. Firstly, we focus on product spaces, proving Theorem 1.3 once again in a slightly more general form, and then we turn to the case of Cayley or Schreier graphs. In particular, we prove that the analogous result holds for the slices of the Boolean cube. We then briefly conclude in the last section with a similar inequality on the Euclidean spheres.

2. A general framework

This section presents the framework and the main tools that will be required in the proofs. The setting emphasized here is quite general, but for the sake of clarity, we discuss in Sections 2.1

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

and 2.2 the two main cases that formed the starting point of this investigation, namely the Gaussian space and the discrete cube endowed with measures ν_p, $p \in (0,1)$.

Let $(\Omega, \mathcal{A}, \mu)$ be a probability space. For a function $f : \Omega \to \mathbb{R}$ in $L^2(\mu)$, denote its variance with respect to μ by

$$\mathrm{Var}_\mu(f) = \int_\Omega f^2\, d\mu - \left(\int_\Omega f\, d\mu\right)^2.$$

In the same way, if $f \geqslant 0$, provided it is well defined, we denote its entropy with respect to μ by

$$\mathrm{Ent}_\mu(f) = \int_\Omega f \log f\, d\mu - \int_\Omega f\, d\mu \log\left(\int_\Omega f\, d\mu\right).$$

The main argument of the proof will be based on interpolation along a Markov semigroup with invariant measure μ. We refer to the general references [1, 2, 4] for background on Markov semigroups. For the reader's convenience, we briefly recall a few basic aspects, illustrated next for the two basic model examples.

A family $(P_t)_{t \geqslant 0}$ of operators acting on a domain $\mathcal{D}$ of functions on Ω is said to be a semigroup if $P_0 = \mathrm{Id}$ and, for all $s, t \geqslant 0$, $P_{t+s} = P_t \circ P_s$. The semigroup $(P_t)_{t \geqslant 0}$ is said to be Markov if, for all $t \geqslant 0$, $P_t \mathbf{1} = \mathbf{1}$. The infinitesimal generator L of $(P_t)_{t \geqslant 0}$ is defined by

$$\forall f \in \mathcal{D}_2(L), Lf := \lim_{t \to 0} \frac{P_t f - f}{t},$$

where the Dirichlet domain $\mathcal{D}_2(L) \subset \mathcal{D}$ is the set of all functions f in $L^2(\mu)$ for which the above limit exists. Conversely, L and $\mathcal{D}_2(L)$ completely determine $(P_t)_{t \geqslant 0}$. From the definition and the semigroup property, it follows that $\partial_t P_t f = L P_t f$ and $P_0 f = f$, justifying the intuitive notation $P_t = e^{tL}$.

Given such a Markov semigroup $(P_t)_{t \geqslant 0}$, the measure μ is said to be reversible with respect to $(P_t)_{t \geqslant 0}$ if

$$\forall f, g \in L^2(\mu), \int_\Omega f L g\, d\mu = \int_\Omega g L f\, d\mu,$$

and invariant with respect to $(P_t)_{t \geqslant 0}$ if

$$\forall f \in L^1(\mu), \int_\Omega P_t f\, d\mu = \int_\Omega f\, d\mu.$$

The Dirichlet form associated with (L, μ) is the bilinear symmetric operator

$$\mathcal{E}(f, g) = \int_\Omega f(-Lg)\, d\mu$$

on suitable real-valued functions f, g in the Dirichlet domain.

Finally, it will be assumed moreover that $(P_t)_{t \geqslant 0}$ is ergodic with respect to μ, which means that μ a.e., $P_t f \to \int_\Omega f\, d\mu$ as $t \to \infty$. We then notice – as a basic starting point of the future investigation – that the variance of function f with respect to μ can be represented via the semigroup as

$$\mathrm{Var}_\mu(f) = \int_\Omega f^2\, d\mu - \left(\int_\Omega f\, d\mu\right)^2 = \lim_{t \to \infty}\left(\int_\Omega (P_0 f)^2\, d\mu - \int_\Omega (P_t f)^2\, d\mu\right).$$

The following paragraphs aim to illustrate this general set-up by examples of interest. The first one discusses the Gaussian model, and its extension to log-concave measures. The next ones deal with the discrete cube and more general discrete models attached to Markov chains.

2.1. The Gaussian space and continuous setting

Let

$$d\mu(x) = e^{-|x|^2/2} \frac{dx}{(2\pi)^{n/2}}$$

be the standard Gaussian measure on $\Omega = \mathbb{R}^n$, and consider the Ornstein–Uhlenbeck semigroup acting of suitable functions $f : \mathbb{R}^n \to \mathbb{R}$ as

$$U_t f(x) = \int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1-e^{-2t}}y)\, d\mu(y), \quad t \geqslant 0,\ x \in \mathbb{R}^n.$$

As is classical (see *e.g.* [2, p. 4]), the generator of the Ornstein–Uhlenbeck semigroup $(P_t)_{t\geqslant 0}$ is given by $L = \Delta - x \cdot \nabla$. The Ornstein–Uhlenbeck semigroup $(U_t)_{t\geqslant 0}$ is invariant and symmetric with respect to μ, and ergodic (as is easily checked on the previous integral representation). The associated Dirichlet domain contains $L^2(\mu) \cap C^2(\mathbb{R}^n)$, and it follows from integration by parts that for C^2 functions f, g on $\mathbb{R}^n$,

$$\mathcal{E}(f,g) = \int_{\mathbb{R}^n} f(-Lg)\, d\mu = \int_{\mathbb{R}^n} \nabla f \cdot \nabla g\, d\mu.$$

In particular, we have the following decomposition of the Dirichlet form along directions:

$$\mathcal{E}(f,f) = \int_{\mathbb{R}^n} |\nabla f|^2\, d\mu = \sum_{i=1}^{n} \int_{\mathbb{R}^n} (\partial_i f)^2\, d\mu. \tag{2.1}$$

According to these properties, it is immediately checked that the (Gaussian) noise stability $\mathcal{S}_\eta^{\mathcal{G}}(f)$ of a function $f : \mathbb{R}^n \to \mathbb{R}$ as described in the Introduction may be reinterpreted in terms of the semigroup in the following way.

Lemma 2.1. *For $f : \mathbb{R}^n \to \mathbb{R}$ and $\eta > 0$,*

$$\mathcal{S}_\eta^{\mathcal{G}}(f) = \mathrm{Var}_\mu(U_{t/2} f),$$

with $e^{-t} = \sqrt{1-\eta^2}$.

The preceding Gaussian example may be amplified along the same lines to cover families of log-concave measures on $\mathbb{R}^n$. Indeed, let $d\mu(x) = e^{-V(x)}dx$ be a probability measure on the Borel sets of $\mathbb{R}^n$ where $V : \mathbb{R}^n \to \mathbb{R}$ is a smooth potential, invariant and symmetric with respect to the second-order diffusion operator $L = \Delta - \nabla V \cdot \nabla$ with associated semigroup $P_t = e^{tL}$, $t \geqslant 0$. As in the Gaussian case, integration by parts yields, for smooth functions f, g on $\mathbb{R}^n$,

$$\mathcal{E}(f,g) = \int_{\mathbb{R}^n} f(-Lg)\, d\mu = \int_{\mathbb{R}^n} \nabla f \cdot \nabla g\, d\mu,$$

and we therefore obtain a similar decomposition of the Dirichlet form $\mathcal{E}(f,f)$.

We will be concerned more generally with products of such measures, namely $\mu = \bigotimes_{i=1}^m \mu_i$ on $\mathbb{R}^n = \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$ where, for $i = 1, \ldots, m$, $d\mu_i(x)$ is of the form $e^{-V_i(x)}dx$ with $V_i : \mathbb{R}^{n_i} \to \mathbb{R}$

some smooth potential. The product generator L of the L_i is given by

$$L = \sum_{i=1}^{m} \mathrm{Id}_{\mathbb{R}^{n_1}} \otimes \cdots \otimes \mathrm{Id}_{\mathbb{R}^{n_{i-1}}} \otimes L_i \otimes \mathrm{Id}_{\mathbb{R}^{n_{i+1}}} \otimes \cdots \otimes \mathrm{Id}_{\mathbb{R}^{n_m}}$$

with associated (product) semigroup $(P_t)_{t \geqslant 0}$. Setting ∇_i for the gradient in the direction $\mathbb{R}^{n_i}$, the Dirichlet form is decomposed into

$$\mathcal{E}(f,f) = \int_{\mathbb{R}^n} |\nabla f|^2 \, d\mu = \sum_{i=1}^{m} \int_{\mathbb{R}^n} |\nabla_i f|^2 \, d\mu.$$

In this context, we may then state the classical decomposition of the variance (and accordingly of noise stability) along the semigroup which will be the starting point of our investigation.

Lemma 2.2. *For every smooth $f : \mathbb{R}^n \to \mathbb{R}$ and every $t \geqslant 0$,*

$$\mathrm{Var}_\mu(P_t f) = 2 \int_t^\infty \sum_{i=1}^{m} \int_{\mathbb{R}^n} |\nabla_i P_s f|^2 \, d\mu \, ds. \tag{2.2}$$

Proof. By ergodicity and the fundamental theorem of calculus,

$$\begin{aligned} \mathrm{Var}_\mu(P_t f) &= \int_{\mathbb{R}^n} (P_t f)^2 \, d\mu - \left(\int_{\mathbb{R}^n} f \, d\mu \right)^2 \\ &= - \int_t^\infty \frac{d}{ds} \int_{\mathbb{R}^n} (P_s f)^2 \, d\mu \, ds \\ &= -2 \int_t^\infty \int_{\mathbb{R}^n} P_s f L(P_s f) \, d\mu \, ds. \end{aligned}$$

Then, after integration by parts,

$$\mathrm{Var}_\mu(P_t f)1 = 2 \int_t^\infty \mathcal{E}(P_s f, P_s f) \, ds = 2 \int_t^\infty \sum_{i=1}^{m} \int_{\mathbb{R}^n} |\nabla_i P_s f|^2 \, d\mu \, ds$$

from which the lemma follows. □

2.2. The discrete cube

Denote the discrete cube $\{-1,1\}^n$ by C_n, endowed with the measure $\nu_p = (p\delta_{-1} + q\delta_1)^{\otimes n}$, where $p + q = 1$. $L^2(C_n, \nu_p)$ is a Euclidean space with respect to the standard scalar product $\langle \cdot, \cdot \rangle_{L^2(\nu_p)}$. We will use the standard notation $[n]$ for $\{1, \ldots, n\}$.

As described in [22], there is an orthonormal basis, called the Fourier–Walsh basis $(\omega_S)_{S \subset [n]}$, given by $(\prod_{i \in S} \omega_i)_{S \subset [n]}$, where

$$\omega_i = \frac{x_i - (q-p)}{2\sqrt{pq}}.$$

Notice that $\omega_i = -\sqrt{q/p}$ if $x_i = -1$ and $\sqrt{p/q}$ if $x_i = 1$. Each function defined on C_n can then be decomposed into multilinear polynomials as

$$f = \sum_{S \subset [n]} \hat{f}(S) \omega_S, \quad \text{with } \hat{f}(S) = \langle f, \omega_S \rangle_{L^2(\nu_p)}.$$

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

Define the operator L by

$$Lf = \sum_{i=1}^{n} L_i f = \sum_{i=1}^{n} \left(\int_{\{-1,1\}} f \nu_p^{(i)} - f \right),$$

where for all $i \in [n]$, $\nu_p^{(i)}$ denotes the integration with respect to the ith coordinate. It is classical that L acts diagonally on the Fourier–Walsh basis. Indeed, it is immediate that if $i \notin S$, $L_i \omega_S = 0$. If $i \in S$, then

$$L_i \omega_S = \left(\prod_{j \in S,\, j \neq i} \omega_j \right) \left(-p\sqrt{\frac{q}{p}} + q\sqrt{\frac{p}{q}} - \omega_i \right) = -\omega_S,$$

so that for all $S \subset [n]$, $L\omega_S = -|S|\omega_S$.

Now we define the family of semigroups $(T_t^p)_{t \geqslant 0}$ generated by L, which acts on functions $f : C_n \to \mathbb{R}$ via

$$T_t^p f = \sum_{S \subset [n]} e^{-t|S|} \hat{f}(S) \omega_S, \quad t \geqslant 0.$$

In the uniform case $p = 1/2$, this semigroup is classically referred to as the Bonami–Beckner semigroup. For all $p \in (0,1)$ it follows from the definition that $(T_t^p)_{t \geqslant 0}$ is a Markov semigroup, and by orthogonality of the basis $(\omega_S)_{S \subset [n]}$ it is immediately checked that ν_p is its invariant and reversible probability measure. Further, $(T_t^p)_{t \geqslant 0}$ is ergodic with respect to ν_p.

With this background, as in the Gaussian space, the noise stability may be expressed in terms of the semigroup $(T_t^p)_{t \geqslant 0}$.

Lemma 2.3. *For $f : C_n \to \mathbb{R}$ and $\eta > 0$,*

$$S_\eta^c(f) = \mathrm{Var}_\nu(T_{t/2}^p f),$$

with $e^{-t} = 1 - \eta$.

Proof. Expressed in the Fourier–Walsh basis, it follows from the above properties that

$$\mathrm{Var}_\nu(T_{t/2}^p f) = \sum_{S \subset [n], S \neq \emptyset} e^{-t|S|} (\hat{f}(S))^2 = \sum_{S \subset [n], S \neq \emptyset} (1-\eta)^{|S|} (\hat{f}(S))^2.$$

Since the basis is orthonormal, it suffices to check that for any non-empty subset $S \subset [n]$,

$$S_\eta^c(\omega_S) = \mathbb{E}_{\nu_p}(\omega_S(X)\omega_S(X^\eta)) - (\mathbb{E}_{\nu_p}\omega_S)^2 = \mathbb{E}_{\nu_p}(\omega_S(X)\omega_S(X^\eta)) = (1-\eta)^{|S|}.$$

This is immediate since (X, X^η) is $(1-\eta)$-correlated. □

In addition, as an important fact for further purposes, the Dirichlet form $\mathcal{E}$ takes the form

$$\mathcal{E}(f,f) = \sum_{i=1}^{n} \int_{C_n} |L_i(f)|^2 d\nu_p.$$

Indeed, since for each $i \in [n]$, $\int_{\{-1,1\}} (L_i f) \nu_p^{(i)} = 0$, by the product structure and Fubini's theorem,

$$\int_{C_n} \left(\int_{\{-1,1\}} f \nu_p^{(i)} \right) (-L_i f) d\nu_p = \int_{C_n^{(i)}} \left(\int_{\{-1,1\}} f \nu_p^{(i)} \right) \left(\int_{\{-1,1\}} (-L_i f) \nu_p^{(i)} \right) d\nu_p^{([n]\setminus i)} = 0,$$

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

so

$$\int_{C_n} f(-L_i f) d\nu_p = \int_{C_n} |L_i(f)|^2 d\nu_p.$$

To make the connection with influences, let D_i be the ith derivative of $f : C_n \to \mathbb{R}$ defined by $D_i(f)(x) = f(\tau_i x) - f(x)$, with $x = (x_1, \dots, x_n)$ and $\tau_i x$ as in the Introduction.

It is easily seen that for each $r \geqslant 1$,

$$\int_{C_n} |L_i f|^r d\nu_p = (pq^r + p^r q) \int_{C_n} |D_i(f)|^r d\nu_p,$$

so

$$\|L_i f\|_1 = 2pq\|D_i f\|_1 = 2pqI_i(f)$$

according to the definition given in the Introduction.

Thus, for both the uniform and biased measure on the cube, the decomposition of the variance along the semigroup is similar to the one emphasized in the continuous setting in Lemma 2.2.

Lemma 2.4. *For every $f : C_n \to \mathbb{R}$, every $t \geqslant 0$, and every $p \in (0,1)$,*

$$\mathrm{Var}_{\nu_p}(T_t^p f) = 2\int_t^\infty \sum_{i=1}^n \int_{C^n} |L_i T_s^p f|^2 \, d\nu \, ds = 2pq \int_t^\infty \sum_{i=1}^n \int_{C^n} |D_i T_s^p f|^2 \, d\nu \, ds. \tag{2.3}$$

2.3. General discrete case

In this subsection we discuss extensions of the discrete cube model to general discrete spaces, which we assume to be finite (as will be the case in all of our applications).

Let Ω be a finite space with probability measure μ on which there is a Markov kernel K, invariant and reversible with respect to μ, *i.e.* such that

$$\forall (x,y) \in \Omega^2, \quad \sum_{x\in\Omega} K(x,y)\mu(x) = \mu(y) \quad \text{and} \quad K(x,y)\mu(x) = K(y,x)\mu(y).$$

Define L by $L = K - \mathrm{Id}$, generator of the semigroup $P_t = e^{tL}$, $t \geqslant 0$. The associated Dirichlet form is given by

$$\mathcal{E}(f,g) = \int_\Omega f(-Lg)\, d\mu = \frac{1}{2} \sum_{x,y\in\Omega} (f(x) - f(y))(g(x) - g(y))K(x,y)\mu(y)$$

for functions f, g on Ω.

The discrete cube model enters this setting by a suitable choice of the kernel K. Consider the operator given by $Lf = \int_\Omega f\, d\mu - f$, that is, $Kf = \int_\Omega f\, d\mu$, or $K = \mathrm{diag}(\mu(x))_{x\in\Omega}$. In particular, a simple computation shows that

$$\mathrm{Var}_\mu(f) = \int_\Omega f(-Lf)\, d\mu = \mathcal{E}(f,f).$$

An interesting instance of the preceding setting, extending the case of the Boolean cube, is given by product spaces with product measures,

$$\Omega = \Omega_1 \times \cdots \times \Omega_n \quad \text{with } \mu = \mu_1 \otimes \cdots \otimes \mu_n,$$

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

when we take the product of the Markov operators. That is, for each $i = 1, \dots, n$, and $f : \Omega \to \mathbb{R}$, set $L_i f = \int_{\Omega_i} f \, d\mu_i - f$ and consider the generator on the product space given by

$$Lf = \sum_{i=1}^{n} L_i f.$$

In this case, by an argument similar to that described in the previous subsection, the Dirichlet form $\mathcal{E}$ may be decomposed as

$$\mathcal{E}(f,f) = \sum_{i=1}^{n} \int_{\Omega} f(-L_i f) \, d\mu = \sum_{i=1}^{n} \int_{\Omega} L_i(f)^2 \, d\mu. \tag{2.4}$$

We will refer to this setting as a discrete product structure, which covers the cube with $\Omega_1 = \cdots = \Omega_n = \{-1, 1\}$, where each is equipped with the measure $p\delta_{-1} + q\delta_1$. Among other relevant examples, one can take $\Omega_i = \mathbb{Z}/q\mathbb{Z}$, for any $q \geqslant 3$, endowed with the uniform measure. In this context, the influence of the ith coordinate is naturally defined as $\|L_i f\|_1$ (notice that both definitions over the Boolean cube agree up to a constant depending on the bias p).

The above discussion can be extended to more general Cayley or Schreier graphs (see [8, 21]), therefore covering non-product examples. Let G be a (finite) group for which there is a finite set of generators S that is symmetric, *i.e.* $S^{-1} = S$, and stable by conjugacy. Assume that G is acting transitively on a finite set Ω and let x^g denote the action of g on x, for each $g \in G$, $x \in \Omega$. The associated Schreier graph is the graph of vertices Ω and edges (x, y) if and only if there exists an $s \in S$ such that $y = x^s$. A Cayley graph corresponds to the particular case $G = \Omega$, consisting therefore of the set of vertices $(g)_{g \in G}$ with edges $(g, gs)_{g \in G, s \in S}$. One basic example is the symmetric group $\mathfrak{S}_n$ with generating set transpositions $\mathcal{T}_n$.

Given a Cayley or Schreier graph G, consider the transition kernel K given by

$$K(g_1, g_2) = \frac{1}{|S|} 1_S(g_1 g_2^{-1}), \quad g_1, g_2 \in G,$$

corresponding to the random walk to nearest neighbour and the uniform measure μ on G. Again, it generates the family of semigroups $(P_t = e^{tL})_{t \geqslant 0}$, with $L = K - \mathrm{Id}$. The associated Dirichlet forms $\mathcal{E}$ can be written as

$$\mathcal{E}(f,f) = \frac{1}{2|S|} \sum_{g \in G} \sum_{s \in S} [f(gs) - f(g)]^2 \mu(g) = \frac{1}{2|S|} \sum_{s \in S} \|D_s f\|_{L^2(G)}^2$$

where $D_s f : g \mapsto f(gs) - f(g)$ in the Cayley graph case, and

$$\mathcal{E}(f,f) = \frac{1}{2|S|} \sum_{x \in \Omega} \sum_{s \in S} [f(x^s) - f(x)]^2 \mu(x) = \frac{1}{2|S|} \sum_{s \in S} \|D_s f\|_{L^2(\Omega)}^2$$

where $D_s f : x \mapsto f(x^s) - f(x)$ in the Schreier graph case.

Note finally that the influence of a generator element $s \in S$ on a function f is then naturally defined in this context by $\|D_s f\|_1$.

2.4. The generalized version of noise stability

In this subsection we extend the definition of noise stability to the various models presented in the preceding subsections.

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

Consider therefore the preceding setting of a probability space $(\Omega, \mathcal{A}, \mu)$ equipped with a semigroup $(P_t)_{t\geqslant 0}$ with generator L and Dirichlet form $\mathcal{E}$, invariant and symmetric with respect to μ.

We say that the pair (L, μ) satisfies a spectral gap (or Poincaré) inequality whenever there exists $\lambda > 0$ such that

$$\lambda \operatorname{Var}_\mu(f) \leqslant \mathcal{E}(f, f) \tag{2.5}$$

for every function f of the Dirichlet domain. The spectral gap constant is the largest λ such that (2.5) holds. It is equivalent to the fact that for every function f in $L^2(\mu)$, and every $t > 0$,

$$\operatorname{Var}_\mu(P_t f) \leqslant e^{-2\lambda t} \operatorname{Var}_\mu(f). \tag{2.6}$$

This follows immediately from Gronwall's lemma since

$$\frac{d}{dt} \operatorname{Var}_\mu(P_t f) = \frac{d}{dt} \|P_t f\|_2^2 = -2\mathcal{E}(P_t f, P_t f).$$

For further purposes, it is then not hard to check (see [8]), that the spectral gap inequality with constant λ is equivalent to the fact that for every centred function f (*i.e.* $\int_\Omega f \, d\mu = 0$),

$$\forall t > 0, \operatorname{Var}_\mu(f) \leqslant \frac{1}{1 - e^{-2\lambda t}} (\|f\|_2^2 - \|P_t f\|_2^2). \tag{2.7}$$

As standard examples, $\lambda = 1$ for the standard Gaussian measure on $\mathbb{R}^n$ and similarly for the discrete cube equipped with any biased measure ν. As a result, from Lemmas 2.1 and 2.3, the spectral gap inequality in its formulation (2.6) implies

$$\mathcal{S}_\eta^c(f) \leqslant (1-\eta)\|f\|_2^2 \quad (\text{resp. } \mathcal{S}_\eta^{\mathcal{G}}(f) \leqslant \sqrt{1-\eta^2}\|f\|_2^2)$$

for (centred) functions of the discrete cube (resp. the Gaussian space). Thus, for a sequence of (centred) functions $(f_n)_{n\geqslant 0}$ defined on the discrete cube or on the Gaussian space such that $\inf_{n\geqslant 0} \|f_n\|_2 > 0$, the spectral gap inequality does not determine whether or not it is a noise sensitive sequence.

We then extend the definition of noise stability of sets/functions in our both discrete and continuous setting in order to preserve this property.

Turning to log-concave measures, we give the following definition.

Definition 1. Let $(\mathbb{R}^n, \mu)$, where μ is a (product) log-concave probability measures, with spectral gap constant λ and underlying semigroup $(P_t)_{t\geqslant 0}$. For a measurable function f in $L^2(\mu)$, we define its noise stability with parameter $\eta \in [0,1)$ by

$$\mathcal{S}_\eta(f) = \operatorname{Var}_\mu(P_t f),$$

where $e^{-2\lambda t} = \sqrt{1-\eta^2}$, and similarly for Borel sets $A \subset \mathbb{R}^n$, $\mathcal{S}_\eta(A) = \mathcal{S}_\eta(\mathbf{1}_A)$.

In the context of Section 2.3, extending the case of the Boolean cube, the definition of noise stability is as follows (once again λ is the spectral gap constant and $(P_t)_{t\leqslant 0}$ the underlying semigroup).

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

Definition 2. Let Ω be a finite Schreier graph. For a function $f : \Omega \to \mathbb{R}$, we define its noise stability with parameter $\eta \in [0,1)$ by

$$\mathcal{S}_\eta(f) = \mathrm{Var}_\mu(P_t f),$$

where $e^{-2\lambda t} = 1-\eta$, and similarly for sets $A \subset \Omega$, $\mathcal{S}_\eta(A) = \mathcal{S}_\eta(\mathbf{1}_A)$.

In the case of Schreier graphs, the noise stability can be interpreted in more probabilistic terms. Indeed, since $(P_t)_{t\geqslant 0} = (e^{-t}e^{tK})_{t\geqslant 0}$ and by symmetry of $(P_t)_{t\geqslant 0}$ with respect to μ,

$$\mathcal{S}_\eta(f) = \mathbb{E}_{\mu\otimes\mu}(f(x)f(y)) - (\mathbb{E}_\mu f)^2.$$

In the above equation, y is obtained from x by acting randomly on m elements of the generating set S, where m is a Poisson random variable with mean $2t$, and $t = t(\eta)$ satisfies $e^{-2\lambda t} = 1-\eta$.

From Section 2.2, the last definition agrees with $\mathcal{S}^c_\eta$ in the case of the Boolean cube. Moreover, in this general case the spectral gap inequality (2.6) can be similarly rewritten as

$$\mathcal{S}_\eta(f) \leqslant (1-\eta)\mathrm{Var}_\mu(f).$$

Thus, with this normalization, for each fixed noise η, the spectral gap inequality does not bring any information about the sensitivity of a sequence of functions $(f_n)_{n\geqslant 0}$ of fixed variance. A quantitative relationship given by the generalization of Theorem 1.3, as developed in Section 4, therefore represents an improvement upon the spectral gap inequality similar to that in previous work [14].

2.5. The hypercontractive tool and the decomposition along 'directions'

In this subsection we present the main tool used in the proof of the results of this work, that is, hypercontractivity of the underlying semigroups. First proposed by Nelson [20], it has turned out to be a very useful property, known to be equivalent to the so-called logarithmic Sobolev inequalities. We also present, in a fairly abstract setting, the common decomposition of the Dirichlet form along 'directions' with the associated influences with respect to these directions.

Say that (L,μ) satisfies a logarithmic Sobolev inequality whenever there exists $\rho > 0$ such that

$$\rho\,\mathrm{Ent}_\mu(f^2) \leqslant 2\mathcal{E}(f,f) \tag{2.8}$$

for every function f on the Dirichlet domain. The logarithmic Sobolev constant is the largest $\rho > 0$ such that (2.8) holds. Since the work of Gross [12] in the continuous setting (see *e.g.* [1, 2, 4]) and Diaconis and Saloff-Coste [9] in the discrete setting, it is known that a logarithmic Sobolev inequality is equivalent to hypercontractivity of the semigroup $(P_t)_{t\geqslant 0}$ in the sense that for every $f \in L^p(\mu)$, $t > 0$ and $1 < p < q < \infty$ with $p \geqslant 1 + (q-1)e^{-2\rho t}$,

$$\|P_t f\|_q \leqslant \|f\|_p. \tag{2.9}$$

As a result, the Sobolev logarithmic constant is often also referred to as the hypercontractive constant. It is classical that ρ is smaller than the spectral gap constant λ. It is also a main feature of these inequalities that both spectral gap and logarithmic Sobolev constants of a product space are the minimum of the spectral gap and logarithmic Sobolev constants of each factor in the product. Recall that $\lambda = \rho = 1$ for the standard Gaussian measure on $\mathbb{R}^n$ and similarly for the uniform measure on the discrete cube. For future purposes, it is known that if $d\mu(x) = e^{-V(x)}dx$

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

on $\mathbb{R}^n$ is such that the Hessian of the potential V satisfies $\mathrm{Hess}(V) \geqslant c > 0$ (uniformly, as symmetric matrices), then $\lambda \geqslant \rho \geqslant c$. This very classical fact follows from the pioneering work of Bakry and Émery [3] on hypercontractive diffusions.

In each class of examples of the above subsections, a key property is the decomposition of the Dirichlet form along 'directions'. We will be interested in situations in which these directions commute in an appropriate sense with derivation. This may be expressed in the following abstract formulation, which immediately applies to the various examples of interest. Namely, assume that there is a decomposition of the Dirichlet form $\mathcal{E}$ along 'directions' i for some operators Γ_i such that, for any suitable f on Ω,

$$\mathcal{E}(f,f) = \sum_{i=1}^{m} \int_{\Omega} \Gamma_i(f)^2 \, d\mu, \tag{2.10}$$

where the operators Γ_i commute with the semigroup in a sense that there exists a real constant κ such that, for every $i = 1, \dots, m$ and $t \geqslant 0$,

$$\Gamma_i(P_t f) \leqslant e^{\kappa t} P_t(\Gamma_i(f)). \tag{2.11}$$

To illustrate this rather formal property, consider the example of the Gaussian space $(\mathbb{R}^n, \mu)$ and its Ornstein–Uhlenbeck semigroup $(U_t)_{t \geqslant 0}$. Recall the decomposition (2.1), and observe that $\partial_i U_t f = e^{-t} U_t(\partial_i f)$ so that (2.11) holds with $m = n$, $\Gamma_i(f) = |\partial_i f|$ and $\kappa = -1$. In the framework of a log-concave measure $d\mu(x) = e^{-V(x)} dx$, it is known (see *e.g.* [2, 4]) that whenever the Hessian of the potential V satisfies $\mathrm{Hess}(V) \geqslant c$ where $c \in \mathbb{R}$, for every smooth $f : \mathbb{R}^n \to \mathbb{R}$ and every $t \geqslant 0$,

$$|\nabla P_t f| \leqslant e^{-ct} P_t(|\nabla f|). \tag{2.12}$$

Therefore, a product of log-concave measures $d\mu_i(x) = e^{-V_i(x)} dx$, $i = 1, \dots, m$, for which each potential V_i satisfies $\mathrm{Hess}(V_i) \geqslant c_i$ with $c_i \in \mathbb{R}$, is another instance of the decomposition (2.10) with $\Gamma_i(f) = |\nabla_i f|$ and $\kappa = -\min c_i$.

In the discrete product setting, we may take $\Gamma_i = L_i$. Then it follows immediately from the definition that for each $i = 1, \dots, m$, $L_i L = L L_i$ so that $L_i P_t = P_t L_i$ for every $t \geqslant 0$. These equalities ensure the suitable commutation (2.11) with $\kappa = 0$.

Such general decompositions were emphasized in [8] in connection with the study of influences and the extension of Talagrand's inequality. The starting point of [8] is the variance representation along the semigroup. We proceed in the same manner. Indeed, by ergodicity and the decomposition (2.10), we may express both Lemma 2.2 and 2.4 in a general form. Namely, the noise stability can be decomposed as

$$\mathcal{S}_\eta(f) = \mathrm{Var}_\mu(P_t f) = 2 \sum_{i=1}^{m} \int_t^\infty \int_\Omega \Gamma_i(P_s f)^2 \, d\mu \, ds,$$

where $t = t(\eta)$ is given as in Definition 1 in the continuous case or Definition 2 in the discrete case. Together with hypercontractivity, this decomposition will be at the root of our main conclusions. As well as the above illustrations, another example of interest given by Euclidean spheres will be presented in Section 5 below.

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

To conclude this section, it is worthwhile mentioning that the content of our statements are invariant under translation of the functions by a constant. Therefore, for the remainder of the paper, it will be implicitly assumed that all functions are centred.

3. The Gaussian and log concave settings

This section will be devoted to the proof of Theorem 1.2, in fact in a more general formulation.

Theorem 3.1. *Let μ be a probability measure on $\mathbb{R}^n = \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$ of the form $d\mu(x) = \bigotimes_{i=1}^m e^{-V_i(x)}dx$ with $\mathrm{Hess}(V_i) \geqslant c > 0$ for every $i = 1, \ldots, m$. Let also $f : \mathbb{R}^n \to \mathbb{R}$ be C^1-smooth and in $L^2(\mu)$. Then, for every $\eta \in (0,1)$,*

$$\mathcal{S}_\eta(f) \leqslant \max\left(4, \frac{4}{c}\right)(1-\eta^2)^{c/(4\lambda)} \left(\sum_{i=1}^m \|\nabla_i f\|_1^2\right)^{\alpha(\eta)} \|f\|_2^{2-2\alpha(\eta)}$$

where

$$\alpha(\eta) = \frac{1-(1-\eta^2)^{\rho/(4\lambda)}}{2}.$$

Recall that ρ denotes the hypercontractive constant and λ the Poincaré constant, and that in this setting we have $\lambda \geqslant \rho \geqslant c$.

To compare with the results of [16] in the Gaussian case (corresponding therefore to $m = n$, $n_1 = \cdots = n_m = 1$, and V_i quadratic) for which $\lambda = \rho = c = 1$, Theorem 3.1 implies that for all $f : \mathbb{R}^n \to \mathbb{R}$ and every $\eta \in (0,1)$,

$$\mathcal{S}_\eta^{\mathcal{G}}(f) \leqslant 4(1-\eta^2)^{1/4} \left(\sum_{i=1}^n \|\partial_i f\|_{L^1(\mu)}^2\right)^{(1-(1-\eta^2)^{1/4})/2} \|f\|_2^{1+(1-\eta^2)^{1/4}}.$$

It may always be assumed that

$$\sum_{i=1}^n \|\partial_i f\|_{L^1(\mu)}^2 \leqslant \|f\|_2^2,$$

otherwise the above inequality is implied by the spectral gap inequality (in its formulation $\mathcal{S}_\eta^{\mathcal{G}}(f) \leqslant \sqrt{1-\eta^2}\|f\|_2^2$). As

$$1-(1-\eta^2)^{1/4} \geqslant \frac{\eta^2}{4},$$

it thus yields the inequality of [16] with $C_1 = 4$ and $C_2 = 1/8$.

Proof of Theorem 3.1. We make the assumption that f is such that

$$\sum_{i=1}^m \|\nabla_i f\|_1^2 \leqslant \|f\|_2^2,$$

otherwise there is nothing to prove (again by the spectral gap inequality). We set

$$t = -\frac{\log(1-\eta^2)}{4\lambda},$$

so that $\mathcal{S}_\eta(f) = \mathrm{Var}_\mu(P_t f)$.

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

We start with Lemma 2.2, from which

$$\mathrm{Var}_\mu(P_t f) = 2\int_t^\infty \sum_{i=1}^m \int_{\mathbb{R}^n} |\nabla_i P_s f|^2 \, d\mu \, ds.$$

The main step of the proof consists of the following cut-off argument. Namely, for $i = 1, \dots, m$ and $M > 0$,

$$\int_{\mathbb{R}^n} |\nabla_i P_s f|^2 \, d\mu = \int_{\{|\nabla_i P_s f| \leqslant M \|\nabla_i f\|_1\}} |\nabla_i P_s f|^2 \, d\mu + \int_{\{|\nabla_i P_s f| > M \|\nabla_i f\|_1\}} |\nabla_i P_s f|^2 \, d\mu. \tag{3.1}$$

The first integral is bounded from above as

$$\begin{aligned}\int_{\{|\nabla_i P_s f| \leqslant M \|\nabla_i f\|_1\}} |\nabla_i P_s f|^2 \, d\mu &\leqslant M \|\nabla_i f\|_1 \int_{\mathbb{R}^n} |\nabla_i P_s f| \, d\mu \\ &\leqslant M \|\nabla_i f\|_1 e^{-cs} \int_{\mathbb{R}^n} P_s(|\nabla_i f|) \, d\mu,\end{aligned}$$

where in the last inequality we use the commutation property (2.12), which ensures that $|\nabla_i P_s f| \leqslant e^{-cs} P_s(|\nabla_i f|)$. Since the measure μ is invariant with respect to $(P_s)_{s \geqslant 0}$, it follows that

$$\int_{\{|\nabla_i P_s f| \leqslant M \|\nabla_i f\|_1\}} |\nabla_i P_s f|^2 \, d\mu \leqslant e^{-cs} M \|\nabla_i f\|_1^2.$$

After integrating over time and summing over $i = 1, \dots, m$, we reach a first bound

$$\sum_{i=1}^m \int_t^\infty \int_{\{|\nabla_i P_s f| \leqslant M \|\nabla_i f\|_1\}} |\nabla_i P_s f|^2 \, d\mu \, ds \leqslant \frac{1}{c} e^{-ct} M \sum_{i=1}^m \|\nabla_i f\|_1^2. \tag{3.2}$$

We now focus on bounding from above

$$\sum_{i=1}^m \int_{\{|\nabla_i P_s f| > M \|\nabla_i f\|_1\}} |\nabla_i P_s f|^2 \, d\mu.$$

To this end, for every $i = 1, \dots, m$, Hölder's inequality applied to $|\nabla_i P_s f|^2$ and $\mathbf{1}_{\{|\nabla_i P_s f| > M \|\nabla_i f\|_1\}}$ yields

$$\int_{\{|\nabla_i P_s f| > M \|\nabla_i f\|_1\}} |\nabla_i P_s f|^2 \, d\mu \leqslant \mu\big\{ |\nabla_i P_s f| > M \|\nabla_i f\|_1 \big\}^{1/q} \|\nabla_i P_s f\|_{2p}^2 \tag{3.3}$$

for every $p, q \geqslant 1$ such that $1/p + 1/q = 1$. However, since $P_s f = P_{s/2}(P_{s/2} f)$, using (2.12) once again,

$$|\nabla_i P_s f|^{2p} \leqslant e^{-cps} [P_{s/2}(|\nabla_i P_{s/2} f|)]^{2p},$$

so that by integration

$$\|\nabla_i P_s f\|_{2p}^2 \leqslant e^{-cs} \big\| P_{s/2}(|\nabla_i P_{s/2} f|) \big\|_{2p}^2.$$

Now, the hypercontractive property (2.9) ensures that, if ρ denotes the hypercontractive constant, for $p = p(s)$ with $2p(s) - 1 \; = \; e^{\rho s}$,

$$\big\| P_{s/2}(|\nabla_i P_{s/2} f|) \big\|_{2p}^2 \leqslant \|\nabla_i P_{s/2} f\|_2^2.$$

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

Putting (3.3) and the last two inequalities together, we infer that

$$\int_{\{|\nabla_i P_s f| > M\|\nabla_i f\|_1\}} |\nabla_i P_s f|^2 \, d\mu \leqslant e^{-cs} \mu\big\{|\nabla_i P_s f| > M\|\nabla_i f\|_1\big\}^{1/q(s)} \|\nabla_i P_{s/2} f\|_2^2, \tag{3.4}$$

with

$$q(s) = \frac{p(s)}{p(s)-1} = \frac{e^{\rho s}+1}{e^{\rho s}-1}.$$

By Markov's inequality, we further obtain, using again the commutation (2.12) and the invariant property of the semigroup, that

$$\mu\big\{|\nabla_i P_s f| > M\|\nabla_i f\|_1\big\} \leqslant \frac{1}{M\|\nabla_i f\|_1} \int_{\mathbb{R}^n} |\nabla_i P_s f| \, d\mu \leqslant \frac{e^{-cs}}{M}. \tag{3.5}$$

The above bounds (3.4) and (3.5) therefore imply

$$\begin{aligned} \int_t^\infty \int_{\{|\nabla_i P_s f| > M\|\nabla_i f\|_1\}} |\nabla_i P_s f|^2 \, d\mu \, ds &\leqslant \int_t^\infty \left(\frac{e^{-cs}}{M}\right)^{1/q(s)} e^{-cs} \|\nabla_i P_{s/2} f\|_2^2 \, ds \\ &\leqslant e^{-ct} \int_t^\infty \frac{1}{M^{1/q(s)}} \|\nabla_i P_{s/2} f\|_2^2 \, ds. \end{aligned}$$

We then notice that the function

$$s \mapsto \frac{1}{q(s)} = \tanh(\rho s/2)$$

is increasing. Hence, for every $M \geqslant 1$,

$$e^{-ct} \int_t^\infty \frac{1}{M^{1/q(s)}} \|\nabla_i P_{s/2} f\|_2^2 \, ds \leqslant \frac{e^{-ct}}{M^{1/q(t)}} \int_t^\infty \|\nabla_i P_{s/2} f\|_2^2 \, ds.$$

Summing over $i = 1, \ldots, m$,

$$\sum_{i=1}^m \int_t^\infty \int_{\{|\nabla_i P_s f| > M\|\nabla_i f\|_1\}} |\nabla_i P_s f|^2 \, d\mu \, ds \leqslant \frac{e^{-ct}}{M^{1/q(t)}} \sum_{i=1}^m \int_t^\infty \|\nabla_i P_{s/2} f\|_2^2 ds.$$

By Lemma 2.2 again,

$$\sum_{i=1}^m \int_t^\infty \|\nabla_i P_{s/2} f\|_2^2 ds = \mathrm{Var}_\mu(P_{t/2} f) \leqslant \|f\|_2^2$$

(recall that f is centred), so

$$\sum_{i=1}^m \int_t^\infty \int_{\{|\nabla_i P_s f| > M\|\nabla_i f\|_1\}} |\nabla_i P_s f|^2 \, d\mu \, ds \leqslant \frac{e^{-ct}}{M^{1/q(t)}} \|f\|_2^2. \tag{3.6}$$

By the decomposition (3.1), the two bounds (3.2) and (3.6) therefore yield

$$\mathrm{Var}_\mu(P_t f) \leqslant \max\left(2, \frac{2}{c}\right) e^{-ct} \left(M \sum_{i=1}^m \|\nabla_i f\|_1^2 + \frac{1}{M^{1/q(t)}} \|f\|_2^2\right).$$

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

Here, we recall that $M \geqslant 1$. Given the assumption $\|f\|_2^2 \geqslant \sum_{i=1}^m \|\nabla_i f\|_1^2$, we can choose M such that

$$M \sum_{i=1}^m \|\nabla_i f\|_1^2 = \frac{1}{M^{1/q(t)}} \|f\|_2^2.$$

We therefore get

$$M^{1+1/q(t)} = \frac{\|f\|_2^2}{\sum_{i=1}^m \|\nabla_i f\|_1^2},$$

so that finally

$$\mathrm{Var}_\mu(P_t f) \leqslant \max\left(4, \frac{4}{c}\right) e^{-ct} \left(\sum_{i=1}^n \|\nabla_i f\|_1^2\right)^{1/(1+q(t))} \|f\|_2^{2q(t)/(1+q(t))}.$$

Replacing $q(t)$ with its explicit form, and recalling that $t = -\log(1-\eta^2)/(4\lambda)$, we obtain the stated claim. Theorem 3.1 is established. □

4. The discrete setting

4.1. The case of the Boolean cube and other discrete product spaces

This section develops the corresponding analysis for the cube and discrete product models.

To deal with the Boolean cube $\{-1,1\}^n$, consider the discrete product structure as emphasized in Section 2.3 consisting of a product space $\Omega = \Omega_1 \times \cdots \times \Omega_n$ with product probability measure $\mu = \mu_1 \otimes \cdots \otimes \mu_n$, each factor (Ω_i, μ_i) being endowed with the Markov operator $L_i f = \int_{\Omega_i} f\, d\mu_i - f$. We recall that the Dirichlet form $\mathcal{E}$ admits the decomposition

$$\mathcal{E}(f,f) = \sum_{i=1}^n \int_\Omega L_i(f)^2\, d\mu,$$

and that the equality $\mathrm{Var}_\mu f = \mathcal{E}(f,f)$ implies that the spectral gap constant is equal to 1. The underlying (product) semigroup $(P_t)_{t \geqslant 0}$ will be assumed to be hypercontractive with constant ρ (equivalently, each (L_i, μ_i) is hypercontractive with constant ρ).

Theorem 4.1. *Let $\Omega = \Omega_1 \times \cdots \times \Omega_n$ equipped with a product probability measure $\mu = \mu_1 \otimes \cdots \otimes \mu_n$, and let $f : \Omega \to \mathbb{R}$. Then, if $\eta \in [0,1)$,*

$$\mathcal{S}_\eta(f) \leqslant 5 \left(\sum_{i=1}^n \|L_i f\|_1^2\right)^{\alpha(\eta)} \|f\|_2^{2-2\alpha(\eta)},$$

where

$$\alpha(\eta) = \frac{1-(1-\eta)^{\rho/2}}{2} \geqslant \frac{\rho}{4}\eta.$$

Theorem 4.1 contains the result of [14] for the discrete cube $\Omega = \{-1,1\}^n$ with $\mu = \nu_p = (p\delta_{-1} + q\delta_1)^{\otimes n}$. To make the connection, note that we can assume that

$$\sum_{i=1}^n \|L_i f\|_1^2 \leqslant \|f\|_2^2,$$

otherwise there is nothing to prove. Recall that

$$\|L_i f\|_1 = 2pq\|D_i f\|_1 = 2pqI_i(f).$$

Thereby, we get

$$\mathcal{S}^c_\eta(f) \leqslant 5(2pq)^{\rho\eta/2}\left(\sum_{i=1}^n (I_i f)^2\right)^{\rho\eta/4}\|f\|_2^{2-\rho\eta/2}.$$

so that, since $4pq \leqslant 1$,

$$\mathcal{S}^c_\eta(f) \leqslant 5\left(\sum_{i=1}^n (I_i f)^2\right)^{\rho\eta/4}\|f\|_2^{2-\rho\eta/2}.$$

On the other hand, the logarithmic Sobolev constant is equal to

$$\rho = \frac{2(p-q)}{\log p - \log q} \quad (= 1 \text{ if } p = q)$$

(see [9]), which implies the main result of [14] as emphasized in the Introduction. In the uniform case (*i.e.* $p = q$), we have $\rho = 1$ and the above result is similar to Theorem 4 of [14] with weaker assumptions on f. In the biased case, the constants are somewhat weaker for small p or q.

It is worth pointing out that the quantitative relationship of Theorem 4.1 yields empty results when $\log p$ is of order $\log n$. In this range, indeed, the Benjamini–Kalai–Schramm relationship does not hold even qualitatively, since there exist sequences of noise stable functions $(f_n)_{n\geqslant 0}$ such that $\sum_{i=1}^n I_i(f_n)^2$ goes to 0 (see [14]).

Proof of Theorem 4.1. The scheme of the proof is almost identical to that developed in the continuous setting for Theorem 3.1. Nevertheless, since (2.11) holds with $\kappa = 0$, there is no longer an exponential decay in time. Thus, in the decomposition of the variance with respect to the semigroup of Lemma 2.4, we need an integration over a finite domain $[0,T]$ for some $T > 0$. This is possible thanks to inequality (2.7):

$$\mathrm{Var}_\mu(f) \leqslant \frac{1}{1-e^{-2T}}(\|f\|_2^2 - \|P_T f\|_2^2).$$

Since

$$\begin{aligned}\|P_t f\|_2^2 - \|P_{t+T} f\|_2^2 &= -\int_t^{T+t} \frac{d}{ds}\|P_s f\|_2^2\,ds \\ &= 2\int_t^{t+T} \mathcal{E}(P_s f, P_s f)\,ds \\ &= 2\int_t^{t+T} \sum_{i=1}^n \int_\Omega |L_i P_s f|^2\,d\mu\,ds\end{aligned}$$

and $\mathcal{S}_\eta(f) = \mathrm{Var}_\mu(P_t f)$ (with $e^{-2t} = 1-\eta$), for every $T > 0$,

$$\mathcal{S}_\eta(f) \leqslant \frac{2}{1-e^{-2T}}\int_t^{t+T} \sum_{i=1}^n \int_\Omega |L_i P_s f|^2\,d\mu\,ds.$$

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

For each $i = 1, \dots, n$, we again cut the integral into two parts with $M \geqslant 1$. The same commutation and contraction argument yields

$$\int_{\{|L_i P_s f| \leqslant M\|L_i f\|_1\}} |L_i P_s f|^2 \, d\mu \leqslant M \|L_i f\|_1 \int_\Omega |L_i P_s f| \, d\mu \leqslant M \|L_i f\|_1^2.$$

Therefore

$$\mathrm{Var}_\mu(P_t f) \leqslant \frac{2T}{1-e^{-2T}} \left(M \sum_{i=1}^n \|L_i f\|_1^2 + \frac{1}{T} \int_t^{t+T} \sum_{i=1}^n \int_{\{|L_i P_s f| > M\|L_i f\|_1\}} |L_i P_s f|^2 \, d\mu \, ds \right).$$

Above the truncation level, for each $i = 1, \dots, n$, the same argument based on the Hölder and hypercontractivity inequalities yields

$$\begin{aligned} \int_{\{|L_i P_s f| > M\|L_i f\|_1\}} |L_i P_s f|^2 \, d\mu &\leqslant \mu\big\{|L_i P_s f| > M\|L_i f\|_1\big\}^{1/q} \|L_i P_s f\|_{2p}^2 \\ &\leqslant \|L_i P_{s/2} f\|_2^2 \, \mu\big\{|L_i P_s f| > M\|L_i f\|_1\big\}^{1/q(s)} \\ &\leqslant \frac{1}{M^{1/q(s)}} \|L_i P_{s/2} f\|_2^2, \end{aligned}$$

with

$$q(s) = \frac{e^{\rho s} + 1}{e^{\rho s} - 1}.$$

Therefore,

$$\int_t^{t+T} \sum_{i=1}^n \int_{\{|L_i P_s f| > M\|L_i f\|_1\}} |L_i P_s f|^2 \, d\mu \, ds \leqslant \frac{1}{M^{1/q(t)}} \sum_{i=1}^n \int_t^{t+T} \|L_i P_{s/2} f\|_2^2 \, ds.$$

Since

$$\sum_{i=1}^n \int_t^{t+T} \|L_i P_{s/2} f\|_2^2 \, ds \leqslant \sum_{i=1}^n \int_0^\infty \|L_i P_{s/2} f\|_2^2 \, ds = \|f\|_2^2,$$

we thus get that

$$\mathrm{Var}_\mu(P_t f) \leqslant \frac{2\max(1,T)}{1-e^{-2T}} \left(M \left(\sum_{i=1}^n \|L_i f\|_1^2 \right) + \|f\|_2^2 \frac{1}{M^{1/q(t)}} \right).$$

The theorem then follows as in the conclusion of the proof of Theorem 3.1, using moreover the fact that

$$\inf_{T>0} \frac{4\max(1,T)}{1-e^{-2T}} = \frac{4}{1-e^{-2}} \leqslant 5.$$

□

4.2. The case of more general Schreier graphs and non-product examples

This subsection briefly discusses further examples of interest, basically non-product models, for which the preceding approach may be developed similarly. The basic ingredients for such extension are the decomposition of the variance into directional derivatives and hypercontractivity.

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

Among discrete examples, the recent work of O'Donnell and Wimmer [21] investigates the examples of general Schreier or Cayley graphs. In the context of Section 2.2, recall that the Dirichlet form $\mathcal{E}$ takes the form

$$\mathcal{E}(f,f) = \frac{1}{2|S|}\sum_{x\in\Omega}\sum_{s\in S}[f(x^s)-f(x)]^2\mu(x) = \frac{1}{2|S|}\sum_{s\in S}\|D_s f\|^2_{L^2(\Omega)}.$$

Therefore, the 'directions' are given by the elements of the generating set, and moreover it is shown in [21] that the commutation (2.11) holds with $\kappa = 0$ (see also [8]). As usual, let $(P_t)_{t\geqslant 0}$ denote the underlying semigroup attached to this Dirichlet form, and let λ be the spectral gap constant and ρ the logarithmic Sobolev constant. Noting that

$$\inf_{T>0}\frac{\max(1,T)}{1-e^{-\lambda T}} \leqslant \frac{2}{\lambda},$$

we get the following theorem from the general proof scheme developed in the preceding subsection.

Theorem 4.2. *Let Ω be a Schreier or Cayley graph and $f:\Omega\to\mathbb{R}$. Then, for any $\eta\in(0,1)$,*

$$\mathcal{S}_\eta(f) \leqslant \frac{4}{\lambda|S|}\left(\sum_{s\in\mathcal{S}} I_s(f)^2\right)^{\alpha(\eta)}\|f\|_2^{2-2\alpha(\eta)},$$

with

$$\alpha(\eta) = \frac{1-(1-\eta)^{\rho/(2\lambda)}}{2} \geqslant \frac{\rho}{4\lambda}\eta.$$

Thus, for a sequence of graphs $(\Omega_n)_{n\geqslant 0}$, we see from Theorem 4.2 that the original Benjamini–Kalai–Schramm criterion holds whenever $\inf_{n\in\mathbb{N}}\rho_n/\lambda_n > 0$, since in this case there exists a universal constant $c>0$ such that $\alpha_n(\eta)\geqslant c\eta$. Theorem 4.2 then represents a quantitative form similar to (1.2).

Explicit examples with known respective spectral gap and logarithmic Sobolev constants are given in [21]. Among these, we can cite the discrete tori $(\mathbb{Z}/m\mathbb{Z})^n$, $m\geqslant 2$. Then, given the product structure, both constants λ_n and ρ_n are of the same order, so Theorem 4.2 provides a quantitative Benjamini–Kalai–Schramm relationship. The case of the Boolean cube C_n with uniform measure can be seen as the Cayley graph $(\mathbb{Z}/2\mathbb{Z})^n$ generated by its canonical basis $S=(e_i)_{1\leqslant n}$. Then the Dirichlet form is

$$\mathcal{E}(f,f) = \frac{1}{2n}\sum_{i=1}^n\|D_i f\|^2_{L^2(C_n)} = \frac{2}{n}\sum_{i=1}^n\|L_i f\|^2_{L^2(C_n)},$$

so the spectral gap constant and the hypercontractive constant are both equal to $2/n$ with this normalization. The above statement is therefore another generalization of the results of [14].

The main novelty with respect to the previous subsection is that it also covers two non-product models of graphs, namely the symmetric group and the slices of the Boolean cube. In the case of the symmetric group $\mathfrak{S}_n$, recall its Cayley graph structure with the generating set given by the subset of the transpositions $\mathcal{T}_n$. The spectral gap λ_n is equal to $2/(n-1)$, and the logarithmic

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

Sobolev constant ρ_n is smaller than $a/(n\log n)$ for some $a > 0$ (see [9]). Hence, Theorem 4.2 does not improve upon the spectral gap inequality since ρ_n/λ_n goes to 0 as n goes to infinity. Notice that this conclusion also holds for the inequalities established in [8] and [21] in the case of the symmetric group.

However, as pointed out in [21], in the case of the slices of the Boolean cube, both the spectral gap and Sobolev logarithmic constants may be of the same order. For $1 \leqslant k < n$, the slices (of order k) of the Boolean cube are defined by $\binom{[n]}{k} = \{x \in \{0,1\}^n, \sum_{i=1}^n x_i = k\}$. The symmetric group is acting on $\binom{[n]}{k}$ by $x^\sigma = (x_{\sigma(i)})_{1\leqslant i\leqslant n}$, so it has a Schreier graph structure. The generators are given by the transpositions τ_{ij}, $1 \leqslant i < j \leqslant n$, with $x^{\tau_{ij}}$ being obtained from x by switching i and j. Then, it is a result of Lee and Yau [18] that the spectral gap constant is $\lambda = 1/n$ and the logarithmic Sobolev constant ρ satisfies $\rho^{-1} \sim n\log n^2/(k(n-k))$. In particular when k/n is bounded away from 0 and 1, both constants are of the same order. We rescale the Dirichlet form by multiplying by n so that

$$\mathcal{E}'(f,f) = \frac{1}{n-1}\sum_{1\leqslant i<j\leqslant n}\|D_{\tau_{ij}}f\|_2^2.$$

Thus, the spectral gap constant associated with $\mathcal{E}'$ is equal to 1 and the corresponding Sobolev logarithmic constant ρ' satisfies $\rho'^{-1} \sim \log n^2/(k(n-k))$. In particular, whenever k/n is bounded away from 0 and 1, ρ' is bounded away from 0. Letting $\mathcal{S}_\eta^s = \mathcal{S}_\eta^{s_k}$ denote the noise stability in this context, Theorem 4.2 therefore implies that

$$\mathcal{S}_\eta^s f \leqslant 5\left(\frac{1}{n-1}\sum_{1\leqslant i<j\leqslant n}(I_{\tau_{ij}}f)^2\right)^{c\eta}\|f\|_2^{2-2c\eta}, \tag{4.1}$$

for some positive constant c. We recall that $I_{\tau_{ij}}f$ is the influence of the transposition τ_{ij} on f. Therefore, the original [6] criterion holds over slices of the Boolean cube of order $k \in [c'n,(1-c')n]$, for each positive constant c'. That is, if the sequence

$$\left(\frac{1}{n}\sum_{1\leqslant i<j\leqslant n}(I_{\tau_{ij}}f_n)^2\right)_{n\geqslant 0}$$

goes to 0 when n goes to infinity, then the sequence $(f_n)_{n\geqslant 0}$ is (asymptotically) noise sensitive. It is worth mentioning that a variant of this qualitative result over the slices has recently been established by Forsström [10], who uses this result in connection with the notion of exclusion sensitivity. We point out that (4.1) represents a quantitative version similar to the results of [14, 16]. It would be interesting to provide some application of this quantitative result.

5. The case of the Euclidean spheres

To conclude, we present the continuous model given by the Euclidean spheres in which the preceding proof scheme applies. That is, a decomposition (2.10) holds with commutation (2.11) as well as hypercontractivity (2.9) and spectral gap (2.5) (see [8]). Although we could give a similar definition of noise stability, we will not do so as it does not have a clear meaning. We will express the results only in terms of the heat semigroup associated with the spherical Laplacian.

https://doi.org/10.1017/S0963548318000044 Published online by Cambridge University Press

Let $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ $(n \geqslant 2)$ denote the $(n-1)$-dimensional Euclidean sphere equipped with its normalized surface measure μ. For $i, j \in \{1, \ldots, n\}$, consider $D_{i,j} = x_j \partial_i - x_i \partial_j$. The Dirichlet form associated with the spherical Laplacian

$$\Delta = \frac{1}{2} \sum_{i,j=1}^{n} D_{i,j}^2$$

takes the form

$$\mathcal{E}(f, f) = \int_{\mathbb{S}^{n-1}} f(-\Delta f)\, d\mu = \frac{1}{2} \sum_{i,j=1}^{n} \int_{\mathbb{S}^{n-1}} (D_{i,j} f)^2\, d\mu.$$

Since we clearly have $\Delta D_{i,j} = D_{i,j}\Delta$, (2.11) holds with $\kappa = 0$. Consider again the heat semigroup $(P_t f)_{t \geqslant 0} = (e^{t\Delta} f)_{t \geqslant 0}$ generated by the Laplacian Δ. It is known (see *e.g.* [2]) that the spectral gap constant and the logarithmic Sobolev constant are both equal to $n-1$. Our result can then be stated as follows.

Theorem 5.1. *Let $f : \mathbb{S}^{n-1} \to \mathbb{R}$, where $n \geqslant 2$, be C^1-smooth and in $L^2(\mu)$. Then, for any $t \geqslant 0$,*

$$\mathrm{Var}_\mu(P_t f) \leqslant 5 \left(\sum_{i,j=1}^{n} \|D_{i,j} f\|_1^2 \right)^{\alpha(t)} \|f\|_2^{2-2\alpha(t)},$$

where

$$\alpha(t) = \frac{1 - e^{-(n-1)t}}{2}.$$

Sketch of the proof. The proof follows that of Theorem 3.1, but with the twist emphasized in the discrete framework of Section 4, that is, restricting to a finite domain of integration in time $[0, T]$ for some $T \geqslant 0$ using (2.7) and replacing the logarithmic Sobolev constant by its value $n-1$. Then, we simply notice that for every $\lambda \geqslant 1$

$$\inf_{T>0} \frac{4 \max(1, T)}{1 - e^{-2\lambda T}} = \frac{4}{1 - e^{-2\lambda}} \leqslant 5. \qquad \square$$

Recall that the Poincaré inequality implies that $\mathrm{Var}_\mu(P_t f) \leqslant e^{-(n-1)t} \mathrm{Var}_\mu(f)$. Since the logarithmic Sobolev constant and the spectral gap constant are equal in this case, Theorem 5.1 provides some non-trivial bound on $\mathrm{Var}_\mu(P_t f)$ in specific ranges of t (*i.e.* for t of order $1/n$). To the best of our knowledge, this inequality is new. It would therefore be of interest to exhibit functions $(f_n)_{n \geqslant 0}$ over the spheres such that the sequence $(\sum_{i,j=1}^{n} \|D_{i,j} f_n\|_1^2)_{n \geqslant 2}$ goes to 0.

Acknowledgement

This work was completed during my doctoral research at the University of Toulouse. I warmly thank my PhD advisor Michel Ledoux for introducing this problem to me, and for fruitful discussions. I also thank the anonymous referee for helpful comments in improving the exposition.

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*, Vol. 10 of Panoramas et Synthèses, Société Mathématique de France.

[2] Bakry, D. (1994) L'hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on Probability Theory: École d'Été de Probabilités de Saint-Flour XXII*, Vol. 1581 of Lecture Notes in Mathematics, Springer, pp. 1–114.

[3] Bakry, D. and Émery, M. (1985) Diffusions hypercontractives. In *Séminaire de Probabilités XIX*, Vol. 1123 of Lecture Notes in Mathematics, Springer, pp. 177–206.

[4] Bakry, D., Gentil, I. and M. Ledoux, M. (2014) *Analysis and Geometry of Markov Diffusion Operators*, Vol. 348 of Grundlehren der Mathematischen Wissenschaften, Springer.

[5] Beckner, W. (1975) Inequalities in Fourier analysis. *Ann. of Math.* **102** 159–182.

[6] Benjamini, I., Kalai, G. and Schramm, O. (1999) Noise sensitivity of Boolean functions and applications to percolation. *Publ. Math. Inst. Hautes Etudes Sci.* **90** 5–43.

[7] Bonami, A. (1970) Étude des coefficients de Fourier des fonctions de $L^p(G)$. *Ann. Inst. Fourier (Grenoble)* **20** 335–402.

[8] Cordero-Erausquin, D. and Ledoux, M. (2012) Hypercontractive measures, Talagrand's inequality, and influences. In *Geometric Aspects of Functional Analysis: Israel Seminar 2006–2010*, Vol. 2050 of Lecture Notes in Mathematics, Springer, pp. 169–189.

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

[10] Forsström, M. P. (2015) A noise sensitivity theorem for Schreier graphs. `arXiv:1501.01828`

[11] Garban, C. and Steif, J. E. (2014) *Lectures on Noise Sensitivity and Percolation*, Cambridge University Press.

[12] Gross, L. (1975) Logarithmic Sobolev inequalities. *Amer. J. Math.* **97** 1061–1083.

[13] Kalai, G. and Safra, S. (2006) Threshold phenomena and influence: Perspectives from mathematics, computer science, and economics. In *Computational Complexity and Statistical Physics* (A. G. Percus *et al.*, eds), Oxford University Press, pp. 25–60.

[14] Keller, N. and Kindler, G. (2013) Quantitative relationship between noise sensitivity and influences. *Combinatorica* **33** 45–71.

[15] Keller, N., Mossel, E. and Sen, A. (2012) Geometric influences. *Ann. Probab.* **40** 1135–1166.

[16] Keller, N., Mossel, E. and Sen, A. (2014) Geometric influences II: Correlation inequalities and noise sensitivity. *Ann. Inst. H. Poincaré* **50** 1121–1139.

[17] Ledoux, M. (2000) The geometry of Markov diffusion generators. *Ann. Fac. Sci. Toulouse Math.* **9** 305–366.

[18] Lee, T. Y. and Yau, H. T. (1998) Logarithmic Sobolev inequality for some models of random walks. *Ann. Probab.* **26** 1855–1873.

[19] Mazet, O. (1997) Classification des semi-groupes de diffusion sur $\mathbb{R}$ associés à une famille de polynômes orthogonaux. In *Séminaire de Probabilités XXXI*, Vol. 1655 of Lecture Notes in Mathematics, Springer, pp. 40–53.

[20] Nelson, E. (1973) The free Markov field. *J. Funct. Anal.* **12** 211–227.

[21] O'Donnell, R. and Wimmer, K. (2009) KKL, Kruskal–Katona, and monotone nets. *SIAM J. Comput.* **42** 2375–2399 (*50th Annual IEEE Symposium on Foundations of Computer Science: FOCS 2009*).

[22] Talagrand, M. (1994) On Russo's approximate zero-one law. *Ann. Probab.* **22** 1576–1587.

[23] Talagrand, M. (1996) How much are increasing sets positively correlated? *Combinatorica* **16** 243–258.

[24] Talagrand, M. (1997) On boundary and influences. *Combinatorica* **17** 275–285.