An unified approach to the Junta theorem for discrete and continuous models.

RAPHAËL BOUYRIE

University of Marne-la-Vallée, France

Abstract

In a recent paper, T. Austin has proved an analogous theorem for the continuous torus of the original Junta theorem proved by Friedgut in the case of the Boolean cube. Analogous statements have been established recently in discrete cases such as the discrete Tori by Ellis et.al., and in the case of slices of the Boolean cube by Wimmer and Filmus. In the continuous case, through the notion of geometric influences, a statement has also be made by Keller, Mossel and Sen for Boltzmann probability measures. In this article, we broaden the scope of the arguments of T. Austin to get an unified proof of these results, removing the restriction to Boolean functions. Indeed, the proof of T. Austin relies on semigroup arguments and can be performed in a general framework that covers both Cayley or Schreier graphs or product of log-concave probability measures.

MSC: 60C05; 05D40
Keywords: Juntas, Influences, Schreier graphs, Geometric influences, Uniform enlargement.

1 Introduction

Analysis of Boolean functions is an area at the intersection of theoretical computer science, functional analysis and probability theory, which originally studies Boolean functions - i.e. functions mapping to \{0, 1\} - on the Boolean cube. A central concept in this field in the concept of influence. The first lightening result with respect to influences is probably the KKL theorem of Kahn-Kalai and Linial [K-K-L], which provides a non trivial lower bound on the maximal influence of any Boolean function. This theorem has numerous important applications in areas of computer science and mathematics. Since then, several results related to influences has been established for function defined on the Boolean cube. For an complete overview over analysis of Boolean functions and its recent developments, we refer the reader to the monograph [OD].

In this paper, we will be concern with Friedgut’s Junta Theorem [Fri]. It states that a Boolean function over the discrete Boolean cube with a bounded total influence essentially depends on few coordinates. The original application of Friedgut’s result was related to threshold phenomenons in randoms graphs. Recently, a lot of effort has been made to extend to other discrete spaces many of the existing results on the discrete Boolean cube. We will present some extensions of the Junta theorem of Friedgut, in discrete but also continuous cases. To name a few, [S-T] and [Beal] generalized this work in a discrete setting respectively to Cartesian product of Graphs and discrete tori. Non-product examples has also been investigated and an analogous theorem has been proven for the slices of the cube (also called Bernoulli–Laplace model) independently by Wimmer [Wim].
and by Filmus [Fil]. Another direction, at the root of the present paper, is the generalization of the Junta theorem to the continuous tori of Austin [Aus].

The main ingredients in the proof of [Aus] are semigroups interpolation together with hypercontractive tool combined with a reverse martingale argument. Our purpose is to generalize Austin’s arguments to obtain a Junta theorem both for Cartesian product of graphs and for continuous models considered in the works [CE-L], [Bou]. Typically, in the latter case, the setting is \((\mathbb{R}^n, \mu^{\otimes n})\) such that \((\mathbb{R}, \mu)\) is hypercontractive, and the appropriate notion of influence is the one introduced in [K-M-S1], that will be recalled below.

Before starting, we state some basic definitions. Let a function \(f : \{-1,1\}^n \to \mathbb{R}\) and let \(\nu_p\) to be the probability measure on \(\{-1,1\}^n\) defined by \((p\delta_{-1} + (1-p)\delta_1)^{\otimes n}\), \(p \in (0,1)\). The influence of the \(i\)-th coordinate on \(f\) is given by

\[
I_i(f) = \|f(x) - f(\tau_i x)\|_{L^1(\nu_p)},
\]

where \(\tau_i x = (x_1, \ldots, x_{i-1}, -x_i, x_{i+1}, \ldots, x_n)\). Friedgut’s theorem deals with the total influence \(\sum_{i=1}^n I_i(f)\) of a function \(f\) denoted by \(I(f)\).

A Boolean function \(f : \{-1,1\}^n \to \{0,1\}\) is called a \(k\)-junta, or simply a junta, if it depends only on \(k\) coordinates, where \(k\) does not depend on \(n\). Notice that when \(k = 1\), such a function is called “dictatorship”.

If \(f\) is a junta, it is an immediate consequence that the total influence does not depend on \(n\), i.e. \(I(f) = O(1)\). The Junta theorem of Friedgut [Fri] is a kind of converse statement of this fact. Namely, if \(p\) does not depend on \(n\), the following holds.

**Theorem 1.1.** Set \(f : \{-1,1\}^n \to \{0,1\}\) with a bounded total influence \(I(f)\). Then there exists an \(e^{O(I(f)/\eps)}\)-junta function \(g\) such that

\[
\|f - g\|_{L^1(\nu_p)} \leq \eps.
\]

In order to make an analogous statement in \(\mathbb{R}^n\), we recall the concept of geometric influences for a Borel probability measure \(\mu\) defined by Keller, Mossel and Sen in [K-M-S1] (see also [K-M-S2]). For a (Borel measurable) subset \(A\) of \(\mathbb{R}^n\), the geometric influence of the \(i\)-th coordinate by

\[
I_i^\mu(A) = \mathbb{E}_x[\mu^+(A_i^x)].
\]

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

\[
A_i^x = \{y \in \mathbb{R}, (x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) \in A\}
\]

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

\[
\mu^+(D) = \lim_{r \to 0} \inf \frac{\mu(D + [-r, r]) - \mu(D)}{r}.
\]

For a \((C^1,\cdot)\)-smooth function \(f\), the geometric influences correspond to the \(L^1\)-norm of its partial derivatives, that is for all \(i \in \{1, \cdots, n\}\), \(I_i^\mu(f) = \|\partial_i f\|_{L^1(\mu)}\). Its total influence \(I(f)\) is thus simply \(\sum_{i=1}^n \|\partial_i f\|_{L^1(\mu)}\).

In the case of Boltzmann probability measures on \(\mathbb{R}^n\) of the form \(\frac{1}{Z_p} e^{-|x|^p} \, dx\), \(p \geq 1\), the authors have proved an analogous result of the original Friedgut’s theorem.

The recent work [Aus] on the continuous torus \([0,1]^n\) combines a reverse martingale argument relying on the Cartesian product structure of \([0,1]^n\) and the hypercontractive property of the heat
semigroup on $[0, 1]^n$. It is by now classical that such hypercontractive tool can be used in a setting that covers both discrete and continuous models, as for example in the papers [CE-L], [Bou]. Therefore, in the Cartesian product setting, we are able to deduce a Junta theorem by a rather simple generalization of [Aus]. Such results are not necessary new, and besides we obtain somewhat weaker constants. However, a main novelty in our results is that we remove the restriction to Boolean functions and we consider real-valued functions. The other advantage of our proof (actually Austin’s argument [Aus]) is its simplicity. Indeed, a function with small total influence is such that a large number of its coordinates have either no or few influences. This remark strongly suggests that the function must remain close to its average over such coordinates and therefore should be essentially determined by a small number of them.

As a sample illustration in the continuous case, we prove the following theorem.

**Theorem 1.2.** Let $(\mathbb{R}^n, \mu \otimes n)$ with $d\mu(x) = e^{-v(x)}dx$ a probability measure such that $v'' \geq c > 0$ (uniformly). Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ in $L^2(\mu)$ with a total influence $I(f)$ independent of $n$. Then, there exist a function $g$ and a positive constant $C(\varepsilon, c)$ independent of $n$ such that $g$ depends on at most $C(\varepsilon, c)$ coordinates and

$$\|f - g\|_{L^1(\mu \otimes n)} \leq \varepsilon.$$  

The paper is organized as follow. In the next section, we describe a convenient framework subsequent to this work, both in discrete and continuous setting. In Section 3 we present our generalization of Austin’s proof in the case of Cartesian product of graphs. In the next section, we make use of the recent work of Filmus [Fil] over the slices of the Boolean cube to conclude similarly to a Junta theorem in this space. Finally, in the last section, we discuss the case of product of log-concave measures, proving Theorem 1.2 and its applications to geometric influences for sets.

## 2 Framework

This section aims at presenting the framework and the main tools that will be required in the proofs. This is completely similar to the framework presented in the previous works [CE-L] and [Bou]. In its discrete version, this also recover the setting of [O-W] of particular Schreier graphs and it is slightly more restrictive than the setting of [S-T].

### 2.1 Discrete setting

Let $\Omega$ be a finite space with probability measure $\mu$ on which there is a Markov kernel $K$, invariant and reversible with respect to $\mu$, 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 $L = K - Id$. 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 $\Omega$. Among examples of such spaces, we will discuss the ones of Cayley or Schreier graphs.

Let $G$ be a finite group acting transitively on a finite set $X$; we write $x^g$ for the action of $g \in G$ on $x \in X$. Assume that there exists a generating set $S$ for $G$ which is symmetric: $S = S^{-1}$. The
associated Schreier graph $X = (X, G, S)$ has vertex set $X$ and an edge $(x; y)$ whenever $x^s = y$ for some $s \in S$. A Cayley graph corresponds to the special case $X = G$. In what follows let $X$ be a Schreier or Cayley graph endowed with uniform probability measure $\mu$. Given a Cayley of Schreier graph $X$, consider the transition kernel $K$ given by $K(x_1, x_2) = \frac{1}{|S|} 1_S(x_1 x_2^{-1})$, $x_1, x_2 \in X$.

Such kernel generates the following family of continuous time semigroups $(P_t := e^{tL})_{t \geq 0}$, that is, with the property $P_0 = \text{Id}$, and $\forall t, s \geq 0$, $P_{t+s} = P_t \circ P_s$, where we recall $L = K - \text{Id}$. Thus, given the definition of $K$, in a more probabilistic point of view, $P_t f(x) = f(y)$ where $y$ is obtained from $x$ by taking $m$ random transpositions and $m \sim P(t)$, the Poisson law of parameter $t$.

The associate 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. The condition $S = S^{-1}$ implies moreover the commutation $D_s P_t = P_t D_s$ for every $s \in S, t \geq 0$ (see [O-W], [CE-L], [Bou]).

In this context, define the influence $I_s(f)$ of an element $s \in S$ on a real-valued function $f$ by $\|D_s f\|_1$.

Relevant examples of Cayley or Schreier graphs are given by the discrete tori $(\mathbb{Z}/m\mathbb{Z})^n, m \geq 2$, with generating set $S = (e_i)_{1 \leq i \leq n}$ where $e_i = \{0, 1\}$ with 1 at the $i$-th place. In these particular cases, the Dirichlet form takes the following explicit expression

$$\mathcal{E}(f, f) = \frac{1}{2n} \sum_{i=1}^n \sum_{x \in (\mathbb{Z}/m\mathbb{Z})^n} |f(x + e_i) - f(x)|^2 = \frac{1}{2n} \sum_{i=1}^n \|D_i f\|_{L^2((\mathbb{Z}/m\mathbb{Z})^n)}^2.$$

The Boolean cube (with uniform measure) can be seen as the case $m = 2$.

Another instances are given by the symmetric group $S_n$, $n \geq 2$ or the slices of the Boolean cube 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 $x^\sigma = (x_{\sigma(i)})_{1 \leq i \leq n}$, so that it has a Schreier graph structure. The generators in both case are given by the transpositions $t_{ij}, 1 \leq i < j \leq n$.

Another direction is considering as [CE-L] the operator given by $Lf = \int f d\mu - f$, i.e $K f = \int f d\mu$, or $K = \text{diag}(\mu(x))_{x \in \Omega}$. Extending the case of the Boolean cube, we can consider such product spaces with product measures

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

when we take product of the above Markov operators. That is, set, for $i = 1, \ldots, n$, and $f : \Omega \to \mathbb{R}$

$$L_i f = \int_{\Omega_i} f d\mu_i - f$$

and consider the generator on the product space given by

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

In this case the Dirichlet form $\mathcal{E}$ may be decomposed as

$$\mathcal{E}(f, f) = \sum_{i=1}^n \int_{\Omega_i} L_i(f)^2 d\mu_i.$$
In the original case of the Boolean cube endowed with the measure \( \nu_p \), it corresponds to \( \Omega_1 = \cdots = \Omega_n = \{-1, 1\} \), \( K(x, y) = \nu_p(y) \) and \( L_i f = \int_{\{-1,1\}} f \, d\nu_p - f \). Thus, the Dirichlet form is given by
\[
\mathcal{E}(f, f) = \sum_{i=1}^{n} \int_{\{-1,1\}^n} L_i(f)^2 \, d\nu_p = 2p(1-p) \sum_{i=1}^{n} \int_{\{-1,1\}^n} D_i(f)^2 \, d\nu_p,
\]
where \( D_i f : x \mapsto f(\tau_i x) - f(x) \) and \( \tau_i \) defined as in the introduction.

In the general context, we will define the influence of the \( i \)-th coordinate for a function \( f \) by \( I_i(f) = \|L_i f\|_1 \), although on the discrete cube with measure \( \nu_p \), it agrees with the previous definition only up to a constant depending on \( p \). Since we are interested in functions such that their total influences are independent of \( n \), this slight abuse of notation does not change the content of our results.

In the preceding context, define the spectral gap constant \( \lambda \) as the largest \( \lambda \) such that
\[
\lambda \text{Var}_\mu(f) \leq \mathcal{E}(f, f),
\]
holds for all functions \( f \), where
\[
\text{Var}_\mu(f) = \int_{\Omega} f^2 \, d\mu - \left( \int_{\Omega} f \, d\mu \right)^2.
\]
stands for the variance of a function \( f \in L^2(\Omega) \).

Similarly, the Sobolev logarithmic constant \( \rho \) is the largest \( \rho \) such that
\[
\rho \text{Ent}_\mu(f^2) = \rho \leq 2\mathcal{E}(f, f)
\]
holds for all functions \( f \), where
\[
\text{Ent}_\mu(f) = \int_{\Omega} f \log f \, d\mu - \int_{\Omega} f \, d\mu \log \left( \int_{\Omega} f \, d\mu \right)
\]
stands for the entropy of a positive function \( f \). We recall (see [D-SC]) that it always holds \( \rho \leq \lambda \).

A basic - but nonetheless crucial - property of these inequalities is their stability by products. Namely, if \((\Omega_1, \mu_1)\) has spectral gap constant \( \lambda_1 \) (resp. Sobolev logarithmic constant \( \rho_1 \)) and \((\Omega_2, \mu_2)\) has spectral gap constant \( \lambda_2 \) (resp. Sobolev logarithmic constant \( \rho_2 \)), then the Cartesian product space \((\Omega_1 \times \Omega_2, \mu_1 \otimes \mu_2)\) has spectral gap constant \( \min(\lambda_1, \lambda_2) \) (resp. Sobolev logarithmic constant \( \min(\rho_1, \rho_2) \)).

It is a classical result, proven by Gross [Gro] in a continuous case and adapted in the discrete cases by Diaconis and Saloff-Coste [D-SC], that a Sobolev logarithmic inequality is equivalent to hypercontractivity of the underlying semigroup \((P_t)_{t \geq 0}\). More precisely, if \( \rho \) designs the Sobolev logarithmic constant, for all \( f \in L^p(\mu) \) and all \( t > 0, 1 < p < q < \infty \) with \( p \geq 1 + (q-1)e^{-2\rho t} \),
\[
\|P_t f\|_q \leq \|f\|_p.
\]
(4)
The hypercontractive tool is at the root of many results about Boolean functions. It is important to point out that in the normalization (1), for the discrete cube with uniform measure, both spectral gap and Sobolev logarithmic constants depends on \( n \) and are equal to \( \frac{2}{n} \), whereas in the normalization (3), both constants are equal to 1. In its classical formulation, both constants are equal to 1 in the case of the Boolean cube. For the discrete Tori \((\mathbb{Z}/m\mathbb{Z})^n\) we will therefore rescale the Dirichlet form.
by multiplying by \( \frac{2}{m} \), so that these constants agree for \( m = 2 \). In implies that we will consider the following Dirichlet form \( \mathcal{E}'(f, f) = \frac{1}{4} \sum_{i=1}^{n} \| D_i f \|^2_2 \). With this normalization, it is known (see e.g. [O-W]) that the spectral gap constant \( \lambda \) attached to \( \mathcal{E}' \) is equal to \( 1 - \cos \left( \frac{\pi m}{2} \right) \) and that the Sobolev logarithmic constant \( \rho \) is such that \( \rho \geq \frac{c}{m^2} \) for some positive constant \( c \).

In the statement of ours results in the case of Cartesian product, we will use the normalization \( (3) \). In this case simple computation shows that

\[
\text{Var}_\mu(f) = \int_\Omega f(-L f) d\mu = \mathcal{E}(f, f),
\]

so that in this case the spectral gap constant is always equal to 1.

### 2.2 The continuous setting

Such abstract Markov framework contains continuous examples. For a complete account, we refer the (patient) reader to the monograph [B-G-L]. In this paper, we restrict ourselves to the Euclidean space \( \mathbb{R}^n \) although it may be considered in a broader setting of Riemannian manifolds and we will recall some basic properties that will be used in Section 5.

Let \( (\mathbb{R}, \mu) \) the real line equipped with a probability measure. Assume that \( \mu \) has a (smooth) density, so that we can write \( d\mu(x) = e^{-v(x)} dx \). Then, it follows from integration by parts that the operator \( L \) acting on \( C^2 \)-smooth functions \( f \) such that \( L f(x) = f''(x) - v'(x) f'(x) \) is reversible for \( \mu \), that is

\[
\forall f, g \in L^2(\mu) \cap C^2(\mathbb{R}), \quad \int_{\mathbb{R}} f L g d\mu = \int_{\mathbb{R}} g L f d\mu.
\]

Similarly, define the Dirichlet form as the positive bilinear symmetric form by

\[
\mathcal{E}(f, g) = \int_{\mathbb{R}} f(-L g) d\mu = \int_{\mathbb{R}} f' g' d\mu,
\]

for each functions \( f, g \) in the Dirichlet domain, i.e. functions such that the above quantity is well defined. The spectral gap constant is the largest constant \( \lambda \geq 0 \) such that

\[
\lambda \text{Var}_\mu(f) \leq \mathcal{E}(f, f),
\]

and the Sobolev logarithmic constant \( \rho \) as the largest \( \rho \) such that

\[
\rho \text{Ent}_\mu(f^2) \leq 2 \mathcal{E}(f, f)
\]

(again for all functions of the Dirichlet domain). The operator \( L \) generates a semigroup \( (P_t)_{t \geq 0} \). For a (smooth) function \( f \) of the Dirichlet domain, \( P_t f \) is the unique solution of

\[
\frac{\partial}{\partial t} P_t f = LP_t f \quad \text{with} \quad P_0 f = f.
\]

Say that \( (\mathbb{R}, \mu) \) is hypercontractive with constant \( \rho \) for all \( f \in L^p(\mu) \) and all \( t > 0 \), \( 1 < p < q < \infty \) with \( p \geq 1 + (q - 1) e^{-2 \rho t} \), \( (4) \) holds. Equivalently (by Gross’ argument), the Sobolev logarithmic constant of \( (\mathbb{R}, \mu) \) is equal to \( \rho \).

We will be concern with Cartesian product of such measures \( (\mathbb{R}^n, \mu = \mu_1 \otimes \cdots \otimes \mu_n) \). The product generator \( \bar{L} \) of the \( L_i \) is given by

\[
L = \sum_{i=1}^{n} \text{Id}_{\mathbb{R}^{n-1}} \otimes L_i \otimes \text{Id}_{\mathbb{R}^{n-i}}
\]
with associated (product) semigroup \((P_t)_{t \geq 0}\). The Dirichlet form is decomposed into

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

The spectral gap constant is then given by \(\lambda = \min_{1 \leq i \leq n} \lambda_i\) and the hypercontractive constant is given by \(\rho = \min_{1 \leq i \leq n} \rho_i\).

One basic example is the case of strictly log-concave measures that is \((\mathbb{R}^n, \mu^{\otimes n})\) with \(d\mu(x) = e^{-v(x)} dx\), \(v'' \geq c > 0\) for \(n \geq 1\) (then it is well known (see [Bak], [B-G-L]) that \(\lambda \geq \rho \geq c\)). In particular for the Gaussian space, \(\rho = \lambda = 1\). In [K-M-S1] the authors deal with the family of Boltzmann probability measures given by \(\mu^{\otimes n}_p (p \geq 1)\) where

\[
d\mu_p(x) = \frac{1}{Z_p} e^{-|x|^p} dx,
\]

and \(Z_p\) is the normalizing constant. Convexity of the one dimensional potentials \(x^p\) is not strict anymore (unless \(p = 2\) corresponding to the Gaussian space), nonetheless these measures are hypercontractive for \(p > 2\). Therefore, such measures fall within our framework, and results such that Talagrand’s inequality or quantitative Benjamini–Kalai–Scramm criterion has been established respectively in [CE-L] and [Bon].

To conclude this section, let us mention another important property of these semigroups. If \(v'' \geq \kappa \in \mathbb{R}\) (uniformly), it is well-known (see e.g. [B-G-L]), that \((P_t)_{t \geq 0}\) commutes with the gradient operator \(\nabla\), that is it holds, for all smooth function \(f\),

\[
|\nabla P_t f| \leq e^{rt} P_t(|\nabla f|).
\]

Since we have restrict ourselves to case of one-dimensional products, it implies by the product structure that for each \(i \in \{1, \ldots, n\}\),

\[
|\partial_i P_t f| \leq e^{rt} P_t(|\partial_i f|).
\]

### 3 The case of Cartesian products graphs

In what follows let \(G\) be a Schreier or Cayley graph with uniform probability measure \(\mu\) and generating set \(S\). Let \(f : G \to \mathbb{R}\). For \(\{s_1, \ldots, s_k\} \subset S\), denote \(T = S \setminus \{s_1, \ldots, s_k\}\), \(\mathcal{C}\) the set generated by \((s_j)_{j \in T}\) and

\[
\Pi_T f(x) = \sum_{s \in \mathcal{C}} f(x^s)\mu(s),
\]

the function obtained from \(f\) by averaging over \(\mathcal{C}\). Clearly, \(\Pi_T f\) depends only on \(s_1, \ldots, s_k\), i.e. \(\sum_{s \in T} I_s(\Pi_T f) = 0\). The main idea is to show that when the total influence does not depend on \(n\), \(f\) is close to the \(k\)-junta \(\Pi_T f\), where \(s_1, \ldots, s_k\) are coordinates of “large” influences.

In the case of product structure \(\Omega^n\), the analogous of the operator \(\Pi_T\) is more explicit and consists in integrating with respect to coordinates of “small” influences. Namely for a subset \(\{j_1, \ldots, j_k\}\) of \([n] := \{1, \ldots, n\}\) (that will correspond the coordinates of “large influence”), denote \(T = \{t_1, \ldots, t_{n-k}\}\) so that \([n]\) is the disjoint union of \(\{j_1, \ldots, j_k\}\) and \(\{t_1, \ldots, t_{n-k}\}\). Recall that the (product) Markov chain \(K\) is defined as \(K = \text{diag}(\mu(x))_{x \in \Omega^n} = \text{diag}(K_i)_{1 \leq i \leq n}\) with \(K_i = \text{diag}(\mu(x))_{x \in \Omega}\). Thus \(\Pi_T\) corresponds to

\[
\Pi_T = K_1 \circ \cdots \circ K_{n-k},
\]

that is the integration operator with respect to the coordinates of \(T\).
3.1 The case of product structures.

We use the normalization \( \rho \). Recall that in this context the spectral gap constant is \( \lambda = 1 \) and that \( \rho \) designs the hypercontractive constant.

In this context, we prove the following junta theorem.

**Theorem 3.1.** Let \( (\Omega^n, \mu) \) be a Cayley or Schreier graph product with Sobolev logarithmic constant \( \rho \). Let \( f : \Omega^n \rightarrow \mathbb{R} \) and set \( I(f) = \sum_{i=1}^{n} I_i(f) \). Then there exists a \( \exp\left(\frac{\|f\|_1 \log(\varepsilon)}{\rho^2}\right) \)-junta function \( g : \Omega^n \rightarrow \mathbb{R} \) such that \( \|f - g\|_{L^2(\mu)} \leq \varepsilon \).

The proof of this theorem relies essentially of the following lemma, due to Austin [Aus].

Before stating this, let \( f : \Omega \rightarrow \mathbb{R} \) and define \( 1 \leq i \leq n - k \), \( f_i = K_i(f_{i-1}) \) with \( f_0 = f \), that is integrate \( f \) successively with respect to the coordinates of \( T \). Then, as used in [Aus], the sequence \((f_i)_{1 \leq i \leq n - k}\) is a reverse martingale and by definition of \( \Pi_T \), \( f_{n-k} = \Pi_T f \).

Up to a scaling factor, we can assume that \( \max_{1 \leq i \leq n} \|I_i f\|_\infty = 1 \). Without loss of generality, we can assume \( I(f) \geq 1 \) - otherwise Theorem 3.1 is still true. Then, the following lemma holds (cf Lemma 2.5 in [Aus]).

**Lemma 3.2.** Let \( f : \Omega^n \rightarrow \mathbb{R} \). Denote by \( I(f) = \sum_{i=1}^{n} \|L_i f\|_1 \) to total influence of \( f \). If \( \eta > 0 \) and \( t > 0 \) are fixed, then

\[
\|P_t f - \Pi_T P_t f\|^2_{L^2(\Omega^n)} \leq I(f) \eta \frac{1 - e^{-2t\rho}}{1 + e^{-2t\rho}},
\]

where \( T \) is such that \( \forall i \in T, I_i f \leq \eta \).

To establish this lemma, we mimic the arguments of [Aus]. It combines the spectral gap inequality, the hypercontractive property of \((P_t)_{t \geq 0}\) and log-convexity of the \( L^p \)-norms.

**Proof of lemma 3.2.** By the fact that \((P_t f)_i\) is a reverse martingale, we can write

\[
\|P_t f - \Pi_T P_t f\|^2_{L^2(\Omega^n)} = \sum_{i \in T} \| (P_t f)_{i-1} - \Pi_{S \setminus \{i\}} (P_t f)_{i-1} \|^2_{L^2(\Omega^n)} = \sum_{i \in T} \| (P_t f)_{i-1} - (P_t f)_i \|^2_{L^2(\Omega^n)}. \tag{7}
\]

Besides, the spectral gap inequality applied to the one dimensional functions \( x_i \mapsto (P_t f)_{i-1}(x) \) implies, since the operators \( \Pi_{S \setminus \{i\}} \) are projections, the following inequality (cf [Aus])

\[
\| (P_t f)_{i-1} - (P_t f)_i \|^2_{L^2(\Omega^n)} \leq \|L_i (P_t f)\|^2_2.
\]

Since \( P_t \) commutes with \( L_i \), we can apply the hypercontractive inequality \( 4 \) for each \( i \in T \):

\[
\|L_i (P_t f)\|^2_2 = \|P_t (L_i f)\|^2_2 \leq \|L_i f\|^2_{1 + e^{-2\rho t}}. \tag{8}
\]

By log-convexity of the \( L^p \) norms, it follows that

\[
\|L_i f\|_{1 + e^{-2\rho t}} \leq \|L_i f\|_1^{\alpha(t)} \|L_i f\|_2^{1 - \alpha(t)},
\]

where \( \alpha(t) = \frac{1 - e^{-2\rho t}}{1 + e^{-2\rho t}} \) (notice that \( \alpha(t) = 2\rho t + o(t) \) as \( t \) goes to 0).

Therefore, starting from (7) and using the previous three inequalities all together,

\[
\|P_t f - \Pi_T P_t f\|^2_{L^2(\Omega^n)} = \sum_{i \in T} \| (P_t f)_{i-1} - (P_t f)_i \|^2_{L^2(\Omega^n)} \leq \sum_{i \in T} \|L_i f\|_{2}^{2\alpha(t)} \|L_i f\|_2^{2 - 2\alpha(t)}. \tag{9}
\]
Applying Hölder inequality with exponents \( \left( \frac{1}{\alpha(t)}, \frac{1}{1-\alpha(t)} \right) \) yields

\[
\sum_{i \in T} \| L_i f \|_1^{2\alpha(t)} \| L_i f \|_2^{2-2\alpha(t)} \leq \left( \sum_{i \in T} \| L_i f \|_1^2 \right)^{\alpha(t)} \left( \sum_{i \in T} \| L_i f \|_2^2 \right)^{1-\alpha(t)}. \tag{10}
\]

Using the assumption \( \max_{1 \leq i \leq n} \| L_i f \|_\infty = 1 \), it follows that

\[
\sum_{i \in T} \| L_i f \|_1 \leq \sum_{i \in T} \| L_i f \|_1 \| L_i f \|_\infty \leq \sum_{i \in T} \| L_i f \|_1.
\]

Besides, by definition of \( T \), for all \( i \in T \), \( \| L_i f \|_1 = I_i(f) \leq \eta \), so that we also get

\[
\sum_{i \in T} \| L_i f \|_1^2 \leq \eta \left( \sum_{i \in T} \| L_i f \|_1 \right).
\]

Thus, putting together the above two inequalities,

\[
\left( \sum_{i \in T} \| L_i f \|_1^2 \right)^{\alpha(t)} \left( \sum_{i \in T} \| L_i f \|_2^2 \right)^{1-\alpha(t)} \leq \eta^{\alpha(t)} \sum_{i \in T} \| L_i f \|_1 \leq \eta^{\alpha(t)} I(f).
\]

Recalling (9), and (10), we get

\[
\| P_t f - \Pi_T P_t f \|^2_{L^2(\Omega^n)} \leq I(f) \eta^{\alpha(t)}. \tag{11}
\]

Replacing \( \alpha(t) \) by its explicit expression, it ends the proof of the lemma.

To conclude to amount to Theorem 3.1, we now use the following lemma, that is due to Bakry in continuous setting.

**Lemma 3.3.** For every function \( f : \Omega^n \to \mathbb{R} \), and every \( t \geq 0 \),

\[
\| f - P_t f \|^2_{L^2(\mu)} \leq t E(f, f).
\]

**Proof.** Recall that \( -L \) is a (semi) positive operator on \( L^2(\Omega^n) \). Let \( \{0, \lambda_1 = \lambda, \lambda_2, \ldots\} \) to be its spectrum and \( (\varphi_k)_{k \geq 0} \) an orthonormal basis of eigenvectors with corresponding eigenvalues \( \lambda_k \). Every function \( f \) in \( L^2(\mu) \) can be written as \( \sum_{k \geq 0} f_k \varphi_k \) with \( f_k = \langle f, \varphi_k \rangle_{L^2(\mu)} \). Hence the Dirichlet form is equal to

\[
E(f, f) = \langle f, -Lf \rangle_{L^2(\mu)} = \sum_{k \geq 1} \lambda_k f_k^2
\]

and similarly

\[
\| f - P_t f \|^2_{L^2(\mu)} = \| f - e^{tL} f \|^2_{L^2(\mu)} = \sum_{k \geq 1} (1 - e^{-t\lambda_k}) f_k^2.
\]

The lemma follows then from the inequality \( 1 - e^{-x} \leq x \) valid for all \( x > 0 \). 

We can now finish the proof of Theorem 3.1.
Proof of Theorem \[3.1\] Recall \(I(f) \geq 1\). Since \(\Pi_T\) is a projection, it follows
\[
\|\Pi_T f - \Pi_T P_t f\|_2^2 \leq \|f - P_t f\|_2^2.
\] (12)
Besides \(E(f, f) = \sum^n_i \|L_i f\|_2^2 \leq \sum^n_i \|L_i f\|_1 \|L_i f\|_\infty \leq I(f)\) so that
\[
\|f - P_t f\|_2^2 \leq \sqrt{I(f)}
\] (13)
It follows, from (11), (12) and (13), by the triangle inequality that for all \(t \geq 0\) (since \(\sqrt{I(f)} \leq I(f)\)),
\[
\|f - \Pi_T f\|_2 \leq (2\sqrt{t} + \eta^2)I(f).
\]
But \(\Pi_T f\) depends only on \(k\) coordinates, and since for each \(i \in \{1, \ldots, k\}\), \(I_{\ast_i} f \geq \eta\), one have \(k \leq I(f)\eta^{-1}\). Choose now \(t = \frac{\varepsilon^2}{16I(f)}\), and \(\eta\) such that \(\eta^{-1}(I(f)\eta^{-1}) = \frac{\varepsilon}{2}\). Then \(\eta = \exp(-O(\frac{I(f)}{\log(\varepsilon\rho)})\)) and \(\|f - \Pi_T f\|_2 \leq \varepsilon\). Therefore, every function \(f\) satisfying the assumption \(I(f) = O(1)\) is a \(O(\varepsilon, \exp(O(\frac{I(f)}{\log(\varepsilon\rho)})))\)-junta.

Remark : Notice that, although the estimate is however weaker (due to the extra \(\log \varepsilon\) term) than the ones following from the previous works, the functions can be real-valued. In the case of Boolean functions, the proof can be substantially simplified. Indeed, from (8), one case use in several concrete cases that for Boolean valued function \(L_i f\) takes values in \(\{-1, 0, 1\}\). Therefore, \(\|L_i f\|_p^p\) is a constant with respect to \(p\) (i.e. is equal to \(I_i(f)\)). Thus, \(\mathbf{[K]}\) implies that
\[
\|L_i P_t f\|_2^2 \leq I_i(f)^{\beta(t)},
\]
where \(\beta(t) = \frac{2}{1 + e^{-rt}} > 1\). From then, one can adapt Lemma \[3.2\] to reach the desired conclusion more directly.

To emphasis its interest, Theorem \[3.1\] contains numerous of known results. The simplest case is \(V = (\{-1, 1\}, \nu_p)\). Then recall that he hypercontractive constant is \(\rho = 2\frac{p - (1 - p)}{\log p - \log(1 - p)}\) and that \(\|L_i f\|_1 = p(1 - p)I_i(f)\), where \(I_i(f)\) is define as in the introduction. If \(p\) is independent of \(n\), the following result holds (the case \(p = 1/2\) is the original Friedgut’s junta theorem for Boolean functions):

\textbf{Corollary 3.4.} Let \(f : (\{0, 1\}^n, \nu_p) \mapsto \{0, 1\}\) with total influence \(I(f)\). Then there exists \(g : (\{0, 1\}^n, \nu_p^n) \mapsto \{0, 1\}\) such that \(g\) is a \(e^{O_p(\frac{I(f)\log(\varepsilon)}{\varepsilon^2})}\)-junta and \(\|f - g\|_1 \leq \varepsilon\).

\textbf{Proof.} Indeed, according to our previous results \(\|f - \Pi_T f\|_2 \leq \varepsilon\). Setting \(g = \frac{\text{sgn}(\Pi_T f) + 1}{2}\), we then have
\[
\mathbb{P}_{\nu_p}(f \neq g) \leq \|f - \Pi_T f\|_2^2 \leq \varepsilon^2,
\]
and \(g\) depends on coordinates of \(S\), with \(|S| = e^{O_p(\frac{I(f)}{\log(\varepsilon)}\frac{1}{\varepsilon^2})}\). Substituting \(\varepsilon^2\) by \(\varepsilon\) yields the result.

Another interesting instance is given by the discrete tori \((\mathbb{Z}/m\mathbb{Z})^n\) for \(m \geq 2\) as in \[Beal\]. Recall that \(\lambda = \frac{1 - \cos(\frac{2\pi}{m})}{m}\) and \(\rho \geq \frac{\lambda}{m^2}\) for some numerical constant \(c\). Since it is a Cartesian product, by similar arguments, the following corollary holds (states for Boolean functions).
Corollary 3.5. Let \( m \geq 2 \). Let \( f : (\mathbb{Z}/m\mathbb{Z})^n \to \{0,1\} \) with total influence \( I(f) \) defined as

\[
I(f) = \frac{1}{m^n} \sum_{i=1}^{n} \sum_{x \in (\mathbb{Z}/m\mathbb{Z})^n} |f(x + e_i) - f(x)|.
\]

Then there exists a function \( g \) depending on at most \( \exp\left( O\left( \frac{I(f)m^2 \log(1/\varepsilon)}{\varepsilon} \right) \right) \) coordinates such that

\[
\|f - g\|_{L^1((\mathbb{Z}/m\mathbb{Z})^n)} = \frac{1}{m^n} \sum_{x \in (\mathbb{Z}/m\mathbb{Z})^n} |f(x) - g(x)| \leq \varepsilon.
\]

This is a weak form of Theorem 5 of [Beal] (both in the dependance on \( \varepsilon \) and \( m \)), but with a somewhat simpler proof.

More generally, Theorem 3.1 is a particular case of the recent work of Sachdeva and Tulsiani [S-T]. Namely, if \((G, V)\) is a graph, and \( V^n \) is its \( n \)-th power, the main result of [S-T] ensures that any boolean function with total influence \( I(f) \) on \( V^n \) is a \( O(\varepsilon, \exp(O(I(f)/\rho))) \)-junta (that is, Theorem 3.1 is a particular case of this result up to the logarithmic factor on \( \varepsilon \)).

The proof of [S-T] relies on an appropriate control on the entropy in the spirit of a work of Rossignol [Ros]. It is mentioned that the results can be extend for more general Markov chains \( K \) that the one attached with the standard random walk to the nearest neighbour.

We notice that the proof of lemma 3.2 heavily relies on the Cartesian product structure. Indeed, in the non product setting, the reverse martingale argument fails. For a general graph, the lack of structure impedes to bound efficiently \( \|f - \Pi_T f\|_{L^1(G)} \). In the next section, we make use of a construction by Filmus [Fil] to obtain a similar conclusion over the slices of the Boolean cube with this scheme of proof.

4 The case of the slice of the Boolean cube.

The simplest - and the most popular - case of a non-product Schreier graph is probably the slices of the Boolean cube \( \left( \left[ \begin{array}{c} n \end{array} \right] \right) \), for which a Junta theorem has been established in recent papers by Wimmer [Wim] and Filmus [Fil]. In the last few years, other results of harmonic analysis have been extended over \( \left( \left[ \begin{array}{c} n \end{array} \right] \right) \), and also over the symmetric group \( \mathfrak{S}_n \), such as the KKL theorem [O-W], Talagrand’s inequality [CE-L] or the quantitative Benjamini–Kalai–Schramm relationship between noise stability and influences [Don]. All of these above results rely on the hypercontractivity of the underlying semi-group. In the case of the symmetric group, these results are not improving upon the spectral gap inequality. The reason for it is that the hypercontractive constant, of order \( 1/(n \log n) \), is too small with respect to the spectral gap equal to \( 1/n \). This is however not the case for the slices of the Boolean cube. Indeed, the spectral gap constant is equal to \( 1/n \) and the hypercontractive constant \( \rho \) has been computed by Lee and Yau [L-Y] and is of order \( (n \log \omega(k,n))^{-1} \) with \( \omega(k,n) = \frac{n^2}{k(n-k)} \). Therefore, if \( k \) is of order \( n \), both spectral gap and hypercontractive constants are of the same order, leading to an improvement over the spectral gap inequality.

Recall that the symmetric group is acting on \( \left( \left[ \begin{array}{c} n \end{array} \right] \right)_k \) by \( x^\sigma = (x_{\sigma(i)})_{1 \leq i \leq n} \). Denote

\[
D_{n_{ij}} f : x \mapsto f(x^{\tau_{ij}}) - f(x),
\]
so that \( x^{\tau_{ij}} \) is obtained from \( x \) by switching the coordinates \( i \) and \( j \).
In this context the total influence is defined by

$$\text{Inf}(f) = \frac{1}{n} \sum_{1 \leq i < j \leq n} I_{\tau_{ij}}(f)$$

and similarly, the total influence up to \( k \) coordinates is

$$\text{Inf}^{(k)}(f) = \frac{1}{k} \sum_{1 \leq i < j \leq k} I_{\tau_{ij}}(f).$$

Notice that, if \( \max_{\tau_{ij} \in \mathcal{T}_n} \|D_{\tau_{ij}}f\|_\infty = 1 \) then \( \|D_{\tau_{ij}}f\|_2^2 \leq I_{\tau_{ij}}(f) \) (and equality holds for Boolean functions). Therefore the Dirichlet form is then related to the influences by

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

Wimmer’s original proof \([\text{Wim}]\) of the Junta theorem is done on the symmetric group and uses the properties of Young’s orthogonal representation. It is pointed out that the Junta theorem is false for the symmetric group, but the author is able to deduce it by reduction for Boolean valued the properties of Young’s orthogonal representation. It is pointed out that the Junta theorem is bound on the preceding quantity \( \|f\|_2 \). Without loss of generality (that is, up to a composition with an appropriate permutation), we can choose a threshold \( \eta > 0 \) such that \( I_{\tau_{ij}}f \leq \eta \) whenever \( i, j \notin S \) where \( S = \{n - m + 1, \ldots, n\} \).

Without loss of generality (that is, up to a composition with an appropriate permutation), we can choose a threshold \( \eta > 0 \) such that \( I_{\tau_{ij}}f \leq \eta \) whenever \( i, j \notin S \) where \( S = \{n - m + 1, \ldots, n\} \).

In the above notations the set \( S \) is therefore generated by the transpositions \( (\tau_{ij})_{1 < i < j \leq n-m} \), that is all the permutations of \( [n-m] \) identify as \( \mathcal{G}_{n-m} \) (seen as a subset of \( \mathcal{G}_n \)). The operator given by \([\text{Fil}]\) is thus

$$\Pi_{T}f(x) = \frac{1}{(n-m)!} \sum_{\sigma \in \mathcal{G}_{n-m}} f(x^\sigma)$$

and depends only on the last \( m \) coordinates. Here we say that \( f \) depends on a coordinate of a subset \( S \) of \( [n] \) if \( I_{\tau_{ij}}(f) = 0 \) whenever \( i, j \notin S \).

The Fourier orthogonal basis build in \([\text{Fil}]\) consists of multilinear polynomials \((\chi_B)_{B \in \mathcal{B}_n}\) where \( \mathcal{B}_n \) are subsets of \( [n] \) called “top sets”. Each function \( f \) defined on the slices can then be decomposed into

$$f = \sum_{B \in \mathcal{B}_n} \hat{f}(B) \chi_B,$$

where as usual \( \hat{f}(B) = \frac{\mathbb{E}_x f \chi_B}{\|\chi_B\|_2} = \frac{\langle f \chi_B \rangle_2}{\|\chi_B\|_2^2}. \) Moreover, we have the property that

$$f - \Pi_T f = \sum_{B \in \mathcal{B}_n, B \cap [n-m] \neq \emptyset} \hat{f}(B) \chi_B,$$

and thus

$$\|f - \Pi_T f\|_2^2 = \sum_{B \in \mathcal{B}_n, B \cap [n-m] \neq \emptyset} \hat{f}(B)^2 \|\chi_B\|_2^2.$$
Thus, (14) and Lemma 4.2 enable us to upper bound the quantity \( \|H_t f - \Pi_T H_t f\|_2^2 \) in terms of the total influence of \( H_t f \). Indeed,

\[
\text{Inf}^k(H_t f) = \sum_{B \in \mathcal{B}_n} \exp \left( -t \frac{|B|(n+1-|B|)}{n} \right) |B \cap [k]|(k+1-|B \cap [k]|) \hat{f}(B)^2 \|\chi_B\|_2^2.
\]

Since for each \( k \geq 1 \) and \( |B \cap [k]| > 0 \) one have \( 1 \leq \frac{|B \cap [k]|(k+1-|B \cap [k]|)}{k} \),

\[
\|H_t f - \Pi_T H_t f\|_2^2 = \sum_{B \in \mathcal{B}_n, B \cap [n-m]\neq\emptyset} \exp \left( -t \frac{|B|(n+1-|B|)}{n} \right) \hat{f}(B)^2 \|\chi_B\|_2^2 \leq \text{Inf}^{(n-m)}(H_t f). \quad (15)
\]

We now show how it suffices to conclude to a Junta theorem, in the same manner as in the preceding section. Since the derivatives operators commute with \((H_t)_{t \geq 0}\) in the sense that for any function \( f \), \( H_t(D_{\tau_{ij}} f) = D_{\tau_{ij}}(H_t f) \), the hypercontractive inequality expresses that

\[
\|D_{\tau_{ij}}(H_t f)\|_2^2 = \|H_t(D_{\tau_{ij}} f)\|_2^2 \leq \|D_{\tau_{ij}} f\|_1^2 e^{-\rho t}. \quad (16)
\]
The log-convexity of the $L^p$-norms and Hölder inequality yields this time

$$\inf^{(n-m)}(H_t f) \leq \frac{1}{n-m} \left( \sum_{1 \leq i < j \leq n-m} \|D_{\tau_{ij}} f\|_2^2 \right)^{\alpha(t)} \left( \sum_{1 \leq i < j \leq n-m} \|D_{\tau_{ij}} f\|_2^2 \right)^{1-\alpha(t)}$$

$$\leq \frac{1}{n-m} \left( \sum_{1 \leq i < j \leq n-m} \left( I_{\tau_{ij}} f \right)^2 \right)^{\alpha(t)} \left( \sum_{1 \leq i < j \leq n-m} I_{\tau_{ij}} f \right)^{1-\alpha(t)}.$$  

Recall that $\alpha(t) = \frac{1-2e^{-2t}}{1+e^{-2t}}$ and $\rho$ is the hypercontractive constant of $(H_t)_{t \geq 0}$. The Lee–Yau’s result implies that if $k/n$ is bounded away from 0 and 1, $\rho = O(1)$, and therefore, $\alpha(t)$ does not depend on $n$ if $t$ does not depend on $n$. Recall that, by assumption, for each $1 \leq i < j \leq n-m$, $I_{\tau_{ij}} (f) \leq \eta$.

We therefore get, similarly as in Section 3

$$\|H_t f - \Pi_T H_t f\|_2^2 \leq \eta \alpha(t) \frac{1}{n-m} \sum_{1 \leq i < j \leq n-m} I_{\tau_{ij}} (f) \leq \frac{n}{n-m} \eta \alpha(t) \inf(f).$$

Now, Lemma 4.3 ensures that

$$\|f - H_t f\|_{L^2(\mu)}^2 \leq \frac{(n-1)t}{2} E(f, f) \leq \frac{t}{2} \inf(f).$$

Since $\Pi_T$ is a projection, the triangular inequality yields this time,

$$\|f - \Pi_T f\|_2^2 \leq (t + \frac{n}{n-m} \eta \alpha(t)) \inf(f).$$

Taking $t = \frac{\varepsilon}{2 \inf(f)}$, there exists a constant $c$ such that $\eta \alpha(t) \leq \eta^c$. We choose now $\eta$ such that $\eta^c \leq \frac{\varepsilon}{2 \inf(f)}$. By assumption $\inf(f) = O(1)$ so that $\eta = \varepsilon O(\frac{1}{n})$. In order to conclude, we use Lemma 4.2 of [Fil] (see also [Wim]):

**Lemma 4.3.** For every function on the slice $f$, every $\eta > 0$, there exists a set $S$ of cardinality at most $O(\frac{\inf(f)}{\eta})$ such that for every $i, j \notin S$, $I_{\tau_{ij}} (f) < \eta$.

It implies that one can take $m = O(\frac{\inf(f)}{\eta})$ and it yields - we refer to [Fil] for the details -the Junta conclusion for slices of the hypercube in the form of the following Theorem.

**Theorem 4.4.** Let $f : \binom{[n]}{k} \to \mathbb{R}$ with $\inf(f) = O(1)$, and denote $\mu$ the uniform measure on $\binom{[n]}{k}$. Then there exists $g : \binom{[n]}{k} \to \mathbb{R}$ such that $g$ is an $e^{O(\log(1/\varepsilon))}$-junta and $\|f - g\|_{L^1(\mu)} \leq \varepsilon$.

## 5 The continuous case.

We can extend the preceding results in continuous setting using the definition of geometric influences of [K-M-SI], as considered in [Bon] [CE-L]. In this section, the setting consists of the product space $\mathbb{R}^n$ equipped with a product measure $\mu = \mu_1 \otimes \cdots \otimes \mu_n$ so that each $(\mathbb{R}, \mu_i)$ is hypercontractive with constant $\rho > 0$. For sake of simplicity, we will take $\mu_1 = \cdots = \mu_n$. As already mentioned earlier, in [K-M-SI] the authors are able to deal with family of Boltzmann probability measures $\mu^{2n}_p$ ($p \geq 1$) given by

$$d\mu_p (x) = \frac{1}{Z_p} e^{-|x|^p} \, dx,$$
where $Z_p$ is the normalizing constant. This is a family of log-concave probability measures. Besides, these measures are hypercontractive for $p \geq 2$ ($p = 2$ is the case of the standard Gaussian space), but this is not the case for $p \in [1, 2]$.

Let then $(\mathbb{R}^n, \mu^{\otimes n})$ be such $d\mu(x) = e^{-v(x)}dx$, with $v'' \geq 0$ and such that $(\mathbb{R}, \mu)$ is hypercontractive with constant $\rho$. Actually, as in the work [CE-L], we shall need only an assumption $v'' \geq \kappa$ with $\kappa \in \mathbb{R}$ to have the commutation property \cite{[5]}. These includes potentials of the form $ax^4 - bx^2$, $a, b > 0$. The arguments below can be adapted in this general case, however for sake of clarity and since it is the case for the concrete example of Boltzmann measures, we will consider only $\kappa = 0$.

Then, from the same arguments as in Section \cite{[3]} (keeping the same notations), one can reach the following inequality

$$\|P_tf - \Pi_T P_tf\|_2 \leq \frac{1}{\lambda} \left( \sum_{i \in T} \|\partial_i f\|_2^2 \right)^{\alpha(t)} \left( \sum_{i \in T} \|\partial_i f\|_2^2 \right)^{1-\alpha(t)}, \quad (17)$$

where $\lambda$ is the spectral gap constant. Since we recall $\rho \leq \lambda$, we can replace $\lambda$ by $\rho$.

However, in connection to geometric influences, the $L^2$-norm of the partial derivatives are not well suited. We will therefore get rid of them, using arguments already developed in \cite{[A]}.

Applying (17) in $t/2$ for $P_{t/2} f$ yields, using the semigroup property,

$$\|P_{t/2} f - \Pi_T P_{t/2} f\|_2 \leq \frac{1}{\rho} \left( \sum_{i \in T} \|\partial_i P_{t/2} f\|_2^2 \right)^{\alpha(t/2)} \left( \sum_{i \in T} \|\partial_i P_{t/2} f\|_2^2 \right)^{1-\alpha(t/2)}. \quad (18)$$

Then, it is well known (see e.g. [B-G-L]) that under convexity assumption of the potential $v$ (or the so-called $CD(0, \infty)$ condition), for a fixed $t > 0$, $\varphi_t : s \mapsto P_s((P_{t-s} f)^2)$ is a convex function. Therefore, the convexity of $\varphi_t$ implies the point-wise upper point

$$\varphi_t'(0) = 2|\nabla P_t f|^2 \leq \frac{\varphi_t(t) - \varphi_t(0)}{t} = \frac{P_t f^2 - (P_t f)^2}{t} \leq \frac{P_t f^2}{t}. \quad (19)$$

Integrating in space and using invariance of $(P_t)_{t \geq 0}$ with respect to $\mu$, it implies a reverse spectral gap inequality of the following form :

$$\|\nabla P_t f\|_2^2 \leq \frac{\|f\|_2^2}{2t}.$$  

One therefore have, for each $T \subset [n]$,

$$\left( \sum_{i \in T} \|\partial_i P_{t/2} f\|_2^2 \right)^{1-\alpha(t/2)} \leq \left( \|\nabla P_{t/2} f\|_2^2 \right)^{1-\alpha(t/2)} \leq \left( \frac{\|f\|_2^2}{t} \right)^{1-\alpha(t/2)}. \quad (20)$$

Besides, by convexity of the potential $v$, recall that commutation \cite{[5]} holds with $\kappa = 0$. Using the product structure, for each $i \in [n]$, the point-wise upper bound $\|\partial_i P_{t/2} f\|_2 \leq P_{t/2} (|\partial_i f|)$ holds. Integrating this upper bound in space and using then the invariance of $(P_t)_{t \geq 0}$ with respect to $\mu$, it yields

$$\|\partial_i P_{t/2} f\|_1^2 \leq P_{t/2} (|\partial_i f|) L^2 = |\partial_i f| L^2. \quad (21)$$

Thus, putting the three inequalities \cite{[18], [20]} and (21) together

$$\|P_t f - \Pi_T P_t f\|_2 \leq \frac{1}{\rho} \left( \sum_{i \in T} \|\partial_i f\|_2^2 \right)^{\alpha(t/2)} \left( \frac{\|f\|_2^2}{2t} \right)^{1-\alpha(t/2)}. \quad (22)$$
In view of application to geometric influences, one needs to replace Lemma 3.3 by a \( L^1 \)-version. This was done by Ledoux [Led], who showed (actually in a more general form) that, under convexity of \( v \),

\[
\| f - P_t f \|_1 \leq 2\sqrt{t} \| \nabla f \|_1 \leq 2\sqrt{t} \sum_{i=1}^{n} \| \partial_i f \|_1.
\]

For a fixed \( \eta > 0 \), \( 0 \leq t \leq 1 \), define \( T \) such that for all \( i \in T \), \( \| \partial_i f \|_1 \leq \eta \) and assume \( \sum_{i=1}^{n} \| \partial_i f \|_1 \geq 1 \). Thus using that \( \| \cdot \|_1 \leq \| \cdot \|_2 \), by the triangular inequality (similarly as in Section 4), we get

\[
\| f - \Pi_T f \|_1 \leq \left( 4\sqrt{t} + \frac{1}{\sqrt{\rho t}} \eta^{\alpha(t/2)} \| f \|_2^{1-\alpha(t/2)} \right) \max \left( \sum_{i=1}^{n} \| \partial_i f \|_1, \left( \sum_{i=1}^{n} \| \partial_i f \|_1 \right)^{\alpha(t/2)} \right)
\]

\[
\leq \left( 4\sqrt{t} + \frac{1}{\sqrt{\rho t}} \eta^{\alpha(t/2)} \| f \|_2^{1-\alpha(t/2)} \right) \sum_{i=1}^{n} \| \partial_i f \|_1.
\]

Choosing again \( t \) and \( \eta \) appropriately so that \( \| f - \Pi_T f \|_1 \leq \varepsilon \), we get the following generalization of Theorem 1.2.

**Theorem 5.1.** Let \( f : \mathbb{R}^n \to \mathbb{R} \) with \( \sum_{i=1}^{n} \| \partial_i f \|_1 = I(f) \) and \( \| f \|_{L^2(\mu \otimes \nu)} < \infty \). Then, there exists a function \( g \) such that \( g \) depends of at most \( \exp \left( O \left( \frac{I(f)}{\varepsilon^2 \rho} \right) \log \frac{\varepsilon^2 \rho}{I(f)} \| f \|_2^2 \right) \) coordinates and \( \| f - g \|_{L^1(\mu \otimes \nu)} \leq \varepsilon \).

**Remark:** If we assume \( f \) bounded (say by 1 - as for characteristics functions), (19) implies \( \| \nabla P_t f \|_\infty \leq \frac{1}{\sqrt{2t}} \), and thus from (18) we get that, for all \( t \leq 1 \),

\[
\| P_t f - \Pi_T P_t f \|_2 \leq \frac{\eta^{\alpha(t/2)}}{\rho} \sum_{i=1}^{n} \| \partial_i f \|_1.
\]

This leads more directly to a somewhat improved estimate over \( \eta \).

Applying the above theorem to (smooth approximations of) characteristics functions of sets, we get a condition over the sum of the geometric influences. Such quantity can be interpreted geometrically.

Say that a set \( A \) is increasing if whenever \( x = (x_1, \ldots, x_n) \in A \), \( y = (y_1, \ldots, y_n) \in A \) as soon as for each \( i \in 1, \ldots, n \), \( x_i \leq y_i \) or decreasing if whenever \( x = (x_1, \ldots, x_n) \in A \), \( y = (y_1, \ldots, y_n) \in A \) when for each \( i \in 1, \ldots, n \), \( x_i \geq y_i \). For monotone (either increasing or decreasing) sets, the total influence \( \sum_{i=1}^{n} I_i(A) \) is the measure of the boundary under uniform enlargement \( \mu_\infty^+(A) \) defined by

\[
\mu_\infty^+(A) = \liminf_{r \to 0} \frac{\mu(A + [-r, r]^n) - \mu(A)}{r}
\]

(see [K-M-S1] and also [B-H] for a more complete account on isoperimetric inequalities for the uniform enlargement). Notice that it follows immediately from the definition that for every Borel measurable subset \( A \subset \mathbb{R}^n \), \( \mu^+(A) \leq \mu_\infty^+(A) \) where \( \mu^+(A) \) stands for the usual boundary measure defined by

\[
\mu^+(A) = \liminf_{r \to 0} \frac{\mu(A + B_r^n) - \mu(A)}{r},
\]

with \( B_r^n \) the Euclidean ball centered in 0 of radius \( r \). Theorem 5.1 implies therefore following corollary.
Corollary 5.2. Take \((\mathbb{R}^n, \mu^\otimes n)\) with \(d\mu(x) = e^{-v(x)}dx\) a log-concave measure on the real line with hypercontractive constant \(\rho\). For any monotone set \(A\) with boundary \(\mu_\infty^+(A)\) there exist a constant \(C(\varepsilon, \rho, \mu_\infty^+(A))\) and a set \(B\) such that \(1_B\) is determined by at most \(C(\varepsilon, \rho, \mu_\infty^+(A))\)-coordinates and \(\mu(A\Delta B) = \|1_A - 1_B\|_1 \leq \varepsilon\).

This corollary expresses that any monotone set whose boundary measure under uniform enlargement - and therefore usual boundary measure - does not depend on the dimension can be essentially written as \(A_1 \times \mathbb{R}^{n-m}\), where \(A_1 \subset \mathbb{R}^m\) lies on a subspace of fixed dimension. We refer to Theorem 3.13 of [K-M-S1] for similar results in that direction. We note that in [K-M-S1] the authors are able to deal with family of Boltzmann probability measures \(\mu_p^\otimes n\) even for \(p \in (1, 2)\).

Acknowledgment. The main part of this work has been completed when I made my Ph.D at the University of Toulouse. I thank my Ph.D advisor Michel Ledoux for drawing my attention to [Aus] and for fruitful discussions. I also thank Matthieu Fradelizi for his careful reading and helpful comments.

References

[Aus] T. Austin, On the failure of concentration for the \(l^\infty\)-ball. Israel J. Math, Vol. 211, Issue 1, 221-238 (2016)

[Bak] D. Bakry L’hypercontractivité et son utilisation en théorie des semigroupes., Ecole d’été de Probabilités de Saint-Flour. Lecture Notes in Math. 1581, 1-114 1994. Springer (1994)

[B-G-L] D. Bakry, I. Gentil, M. Ledoux. Analysis and geometry of Markov diffusion operators. Grundlehren der mathematischen Wissenschaften 348. Springer (2014).

[B-H] F. Barthe, B. Huou Perimeters, uniform enlargement and high dimensions. To appear in Bernoulli Journal.

[Beal] I. Benjamini, D. Ellis, E. Friedgut, N. Keller, A. Sen Junta in the \(l^1\)-grid and Lipschitz maps between discrete tori, available online at http://arxiv.org/pdf/1311.6958v3.pdf

[Bou] R. Bouyrie, On quantitative noise stability and influences for discrete and continuous models. To appear in Combinatorics, probability and computing.

[CE-L] D. Cordero-Erausquin, M. Ledoux. Hypercontractives measures, Talagrand’s inequality, and influences, Geometric Aspects of Functional Analysis, Springer, 2012, p. 169-189 (2012)

[D-SC] P. Diaconis, L. Saloff-Coste Logarithmic Sobolev inequalities for finite Markov chains, Ann. Appl. Prob. 6, p. 695-750 (1996)

[Fri] E. Friedgut Boolean functions with low average sensitivity depend on few coordinates. Combinatorica, 27-35 (1998)

[Fil] Y. Filmus Orthogonal basis for functions over a slice of the Boolean hypercube Electronic Journal of Combinatorics, Vol. 23, Issue 1, 1 - 23 (2016)

[Gro] L. Gross Logarithmic Sobolev inequalities, Amer. J. Math., 97: 1061 - 1083 (1975)
[K-K-L] J. Kahn, G. Kalai, N. Linial The influence of variables on Boolean functions, Proc. 29th Ann. Symp. on Foundations of Comp. Sci., Computer Society Press, 68-80 (1988)

[K-M-S1] N. Keller, E. Mossel, A. Sen Geometric influences, Ann. of Probab. 2012, Vol.40 : 1135 - 1166 (2012)

[K-M-S2] N. Keller, E. Mossel, A. Sen Geometric influences II : Correlation inequalities and noise sensitivity, Ann. Inst. H. Poincaré, Vol.50, No4 : 1121 - 1139 (2014)

[Led] M. Ledoux A simple analytic proof of an inequality by P. Buser., Proc. Amer. Math. Soc. 121, p.951-959 , (1994)

[L-Y] . Lee, H.T. Yau Logarithmic Sobolev inequality for some models of random walks Ann. Of Probab. 1998; Vol.26 : 1855-1873 (1998)

[OD] R.O'Donnell, Analysis of Boolean functions. Cambridge University Press (2014)

[O-W] R. O'Donnell, K. Wimmer KKL, Kruskal-Katona, and monotone nets. 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2009), 725-734, IEEE Computer Soc., Los Alamitos, CA, (2009)

[Ros] R. Rossignol. Threshold for monotone symmetric properties through a logarithmic Sobolev inequality. The Annals of Probability, 34(5) : 1707-1725 (2006)

[S-T] S. Sachdeva M. Tulsiani Cuts in Cartesian Products of Graphs Available online at http://arxiv.org/abs/1105.3383

[Wim] K. Wimmer Low influence functions over slices of the Boolean hypercube depend on few coordinates Conference on Computational Complexity (CCC), 120-131 (2014)

Raphaël Bouyrie,
Laboratoire d’Analyse de Mathématiques Appliquées, UMR 8050 du CNRS, Université Paris-Est Marne-la-Vallée, 5 Bd Descartes, Champs-sur-Marne, 77454 Marne-la-Vallée Cedex, France
E-mail address: raphael.bouyrie@upem.fr