Influences in Mixing Measures

Frederic Koehler* Noam Lifshitz† Dor Minzer‡ Elchanan Mossel§

August 25, 2025

Abstract

The theory of influences in product measures has profound applications in theoretical computer science, combinatorics, and discrete probability. This deep theory is intimately connected to functional inequalities and to the Fourier analysis of discrete groups. Originally, influences of functions were motivated by the study of social choice theory, wherein a Boolean function represents a voting scheme, its inputs represent the votes, and its output represents the outcome of the elections. Thus, product measures represent a scenario in which the votes of the parties are randomly and independently distributed, which is often far from the truth in real-life scenarios.

We begin to develop the theory of influences for more general measures under mixing or spectral independence conditions. More specifically, we prove analogues of the KKL and Talagrand influence theorems for Markov Random Fields on bounded degree graphs when the Glauber dynamics mix rapidly. We thus resolve a long standing challenge, stated for example by Kalai and Safra (2005). We show how some of the original applications of the theory of in terms of voting and coalitions extend to these general dependent measures. Our results thus shed light both on voting with correlated voters and on the behavior of general functions of Markov Random Fields (also called "spin-systems") where the Glauber dynamics mixes rapidly.

*fkoehler@uchicago.edu, University of Chicago
†noam.lifshitz@mail.huji.ac.il, Hebrew University.
‡dminzer@mit.edu, MIT.
§elmos@mit,edu, MIT.

1 Introduction

Starting with the works of Ben-Or and Linial [6] and Kahn, Kalai, and Linial [29], Analysis of Boolean functions became a major area of research in combinatorics, probability and theoretical computer science. It has deep and interesting connections to functional and isoperimetric inequalities, and other important areas in probability and combinatorics. It has deep impact in property testing, hardness of approximation, the theory of voting and the theory of percolation, see e.g. [44, 21, 41].

At the technical level this theory crucially relies on:

- Hyper-contractive inequalities that hold for product measures that are not too biased, and
- Explicit representations of functions in explicit bases, which correspond to Fourier bases and their generalizations.

Major recent effort has been devoted to extend the theory to space for which hyper-contractive inequalities do not hold. Notably it was shown that a notion of *global hypercontraction* holds for such spaces and that this in turn implies many interesting applications [34, 32, 33, 24, 2, 31]. In the other direction, extending the theory to spaces that are not highly symmetric and do not have explicit bases remained a major challenge.

Our main contribution in this paper is to prove very general versions of two major theorems of analysis of Boolean functions, the KKL and the Talagrand theorem in the setting of general Gibbs measures on bounded degree graphs with correlation decay. The study of such measures is fundamental in statistical physics, graphical models, and in the analysis of Markov chains and spectral independence, see e.g. [42, 43, 16, 1, 38, 53, 8, 40, 14]. Such measures are known to satisfy the log-Sobolev inequality (equivalently they are hyper-contractive) but do not posses explicit orthogonal bases. The study of such measures is fundamental in statistical physics, graphical models, and in the analysis of Markov chains and spectral independence, see e.g. [42, 43, 16, 1, 38, 53, 8, 40, 14]. Such measures are known to satisfy the log-Sobolev inequality (equivalently they are hyper-contractive) but do not posses explicit orthogonal bases.

Our results provide an answers a major challenge that is open for about 20 years. Indeed in their survey, Kalai and Safra [30] state that: "One of the major research challenges is to extend the results described in this chapter to models where the probability distribution is not a product distribution. Important cases are the Ising and the more general Potts and random cluster models, as well as models based on random walks of various types." Results of Graham and Grimmett [23] proved one version of these theorem that is useful to prove phase transitions for such models, in the special case that they are ferromagnetic/"monotone". But their version of the KKL theorem uses "effects" instead of the true "influences", which means they are not useful for other applications including the original motivations in theoretical computer science to collective coin-tossing and voting. Our results are the first to prove a variant of the KKL theorem that is applicable to these applications — see Section 2.5 for more explanation. In particular, we show how some of the original applications of the theory of influences extend to the new setup: for general voting functions on n voters there exist a voter who influence is $\Omega(\log n/n)$ times the variance. For monotone voting functions there exist a coalition of $O(n/\log n)$ voters who by flipping their votes can control the elections with probability arbitrary close to 1. See also the discussion below for comparison between our results and those of [23].

Our results also have interesting interpretation for the theory of Markov Chain. Informally our results show that functions of low influences are unstable with respect to the natural Markov chains.

See subsection 2.4 for a more detailed discussion.

Other related work. Many works have studied versions of the KKL theorem in algebraic settings like Cayley graphs, e.g. [46, 47, 52, 12], and repeatedly posed the challenge of proving the KKL theorem in other settings with dependent coordinates. The KKL theorem has deep connections and applications to many different areas of research which continue to be investigated. See e.g. [27, 13, 19, 50, 28, 10, 39, 51] for some recent works studying connections between the KKL/Talagrand inequalities and metric geometry, quantum physics, and distributed computing.

2 Definitions and Main Results

We recall the definition of the Glauber dynamics, log-Sobolev constant, etc. See e.g., [3, 40, 54] for references.

Glauber dynamics. Let ν be a probability distribution on the space Σ^n where Σ is an arbitrary finite set of size $q = |\Sigma|$. Let P_i be the Markov operator that resamples coordinate i from stationary distribution ν conditioned on all other coordinates, so that

$$(P_i f)(x) = \mathbb{E}_\nu[f(X) \mid X_{\sim i} = x_{\sim i}],$$

where $x_{\sim i}$ is the vector of all coordinates other than i. We will consider the continuous time Glauber dynamics, where every coordinate (a.k.a. site) is equipped with an independent Poisson clock. Whenever the clock at coordinate i ticks, the spin at this coordinate is updated according to P_i. It is well known that this defines a semigroup H_t where H_t is the transition matrix of the configuration from time 0 to time t. We recall that H being a *semigroup* means that it satisfies that $H_{s+t} = H_s H_t = H_t H_s$ for all s and t. Moreover, we can write $H_t = e^{tL}$, where L, called the *generator*, is given by $L = \sum_i L_i$ and $L_i f = P_i f - f$ so that $L_i^2 = (P_i - I)^2 = -P_i + I = -L_i$. With this notation, the *Dirichlet form* of the Glauber dynamics is defined to be

$$\mathcal{E}_\nu(f, f) = -\mathbb{E}_{X\sim\nu}[f(X)(Lf)(X)] = \sum_i \mathbb{E}_\nu(L_i f)^2.$$

Each L_i can be thought of as a generalized notion of partial derivative with respect to coordinate i, so the Dirichlet form can be viewed as a natural measure of the size of the gradient of the function f (from the perspective of the chosen semigroup).

Log-Sobolev inequality. We say the Glauber dynamics for ν satisfy the log-Sobolev inequality with constant $\rho > 0$ if

$$\rho \operatorname{Ent}_\nu[f] \le 2\mathcal{E}_\nu(\sqrt{f}, \sqrt{f})$$

for all functions $f : \Sigma^n \to \mathbb{R}_{\ge 0}$, where $\operatorname{Ent}_\nu[f] = \mathbb{E}_\nu[f \log f] - \mathbb{E}_\nu[f] \log \mathbb{E}_\nu[f]$ is the relative entropy functional. This is equivalent to the hypercontractivity statement that for all functions f, $t \ge 0$, and $p \ge 1 + e^{-2\rho t}$,

$$\|H_t f\|_2 \le \|f\|_p$$

where $\|\cdot\|_p$ denotes the $L_p(\nu)$ norm $\|f\|_p = (\mathbb{E}_\nu |f|^p)^{1/p}$.

The log-Sobolev inequality implies that the Poincaré inequality

$$\lambda \mathrm{Var}_\nu(f) \le \mathcal{E}_\nu(f, f)$$

holds with some constant $\lambda \ge \rho$ and for all functions $f : \Sigma^n \to \mathbb{R}$. This is equivalent to the statement that $\mathrm{Var}(H_t f) \le e^{-\lambda t}\mathrm{Var}(f)$ for all such f.

Markov property. We say ν is a *Markov random field* with respect to a graph G if it satisfies the *Markov property*: for any vertex i with neighbors $\mathcal{N}(i)$ in G and for $X \sim \nu$, X_i is conditionally independent of $X_{\sim i}$ given $X_{\mathcal{N}(i)}$. Such a distribution is also referred to as an *undirected graphical model*, see [36]. Given a graph G, we let $d_G(i, j)$ denote the graph distance between i and j.

Other notation. Given square matrices X, Y we write $[X, Y] = XY - YX$ for the usual commutator. We write $[X, \cdot]$ to denote the adjoint map $Y \mapsto [X, Y]$. We now come to the important definition of influences for our setting.

Definition 2.1. *Given a function $f : \Sigma^n \to \{0, 1\}$, we define the influence of coordinate i to be*

$$I_i(f) = \Pr_{X\sim\nu}[\exists x_i', f(X) \ne f(X_1, \ldots, X_{i-1}, x_i', X_{i+1}, \ldots, X_n)].$$

We write $d_H(x, y) = \#\{i : x_i \ne y_i\}$ to denote the usual Hamming metric on Σ^n. Given a vector $x \in \Sigma^n$ and $i \in [n]$, $x_{\sim i} \in \Sigma^{n-1}$ denotes the same vector with coordinate i removed. More generally, the notation $x_{\sim S}$ denotes the vector with the coordinates in set S removed.

2.1 Main Results

Our results hold in a very general setting: they apply to all undirected graphical models with bounded marginals, bounded degree, and which satisfy the log-Sobolev inequality. These assumptions are formally laid out below. In Section 4, we illustrate some of the special cases where the log-Sobolev inequality is known to hold and give references to others.

Assumption 1. *Let Σ be a finite alphabet of size q. The probability measure ν on Σ^n for some $n \ge 1$ satisfies that:*

1. *There exists a constant $b \ge 1$ such that*

$$\nu(x)/\nu(y) \in [1/b, b] \tag{1}$$

   for any $x, y \in \Sigma^n$ with Hamming distance one. In other words, ν has bounded marginals under pinning.

2. *The Glauber dynamics for ν satisfy the log-Sobolev inequality with constant $\rho \in (0, 1]$.*

3. *The distribution ν is a Markov random field with respect to a graph G, and every vertex of G has degree at most $\Delta \ge 1$.*

In our key contribution, we show that these assumptions suffice to prove general versions of Talagrand's theorem and the KKL inequality:

Theorem 2.2 (Theorem 3.7 below). *For any $n \geq 1$, ν satisfying Assumption 1, and any $f : \Sigma^n \to \mathbb{R}$, we have*

$$\mathrm{Var}_\nu(f) \leq \frac{Cq^4 b^4 \Delta^2}{\rho} \sum_j \frac{\|L_j f\|_2^2}{1 + \log(\|L_j f\|_2 / \|L_j f\|_1)} \tag{2}$$

for some absolute constant $C > 0$.

Theorem 2.3 (Theorem 3.8 below). *There exists $\alpha_{b,\rho,\Delta,q} > 0$ such that the following is true. For any $n \geq 1$, ν satisfying Assumption 1, and any $f : \Sigma^n \to \{0, 1\}$, there exists a coordinate $k \in [n]$ such that*

$$I_k(f) \geq \alpha_{b,\rho,\Delta,q} \mathrm{Var}(f) \log(n)/n.$$

2.2 Proof Ideas

Our main results are derived as consequences of a new comparison inequality between the variance and derivatives of a function f (Theorem 3.1 below). Our statements in (2) vastly generalize results of Cordero-Erasquin and Ledoux [11]. In [11] a statement similar to (2) was proven under the strong assumption that the operators L_i and semigroup H_t " weakly commute" (equation (15) there). This is valid for product measures and a few other interesting examples in [11] such as the symmetric group, the sphere etc. However, in our setting it fails very badly — an update at one site affects all of its neighbors, which affects their neighbors, and so on. As in [11] and following [3] the proof begins by writing the variance as an "integral over the heat semi-group" (equation (6) below).

Then, in our main contribution we provide a new analysis for this noncommutative setting which controls the commutators corresponding to all of these interactions. A key observation is that even though the operators L_i and H_t do not commute, we can derive an exact formula for their commutator in terms of an expansion into (appropriately labeled) connected subgraphs:

Lemma 2.4 (Lemma 3.3 below). *For any $T \geq 0$ and $i \in [n]$ we have $L_i H_T = H_T M_{T,i}$ where*

$$M_{T,i} := \sum_{k=0}^{\infty} \frac{T^k}{k!} \sum_{(j_1,\ldots,j_k) \in \mathcal{S}_{k,i}} [\cdots[[[L_i, P_{j_1}], P_{j_2}], P_{j_3}] \cdots P_{j_k}]. \tag{3}$$

Here

$$\mathcal{S}_{k,i} := \{(j_1, \ldots, j_k) : j_i \in \mathcal{N}^+(\{i, j_1, \ldots, j_{i-1}\})\} \tag{4}$$

and $\mathcal{N}^+(U)$ denotes the union of U and the neighbors of nodes U in the graph.

The reason that only connected subgraphs contribute is the Markov property (which implies $[P_i, P_j] = 0$ unless i neighbors j). The next stage of the argument is to very carefully bound the terms in this expansion, and relate them to the commutator-free terms $\|L_j f\|_p$ which appear on the right hand side of the desired inequality. Ultimately, we are able to show that at small times T, the total contribution of higher-order diagrams is damped exponentially in terms of the size k, which ultimately lets us show the effect of an update at site i decays exponentially in the graph-theoretic distance (equation (10) below). The bounded degree assumption is used here to control the number of diagrams, and rapid mixing of the dynamics (more precisely, the log-Sobolev inequality) lets us argue that analyzing the behavior at small times T is sufficient to prove the comparison inequality.

2.3 Applications to voting

There is a long history of using Markov random fields/statistical physics models to model the correlated preferences of voters in elections, for example to estimate the probability of a Condorcet paradox (e.g. [49, 9, 20, 22, 35]). Our results have a natural interpretation in the voting context. If each entry $X \sim \nu$ corresponds to the preference of an individual, and $f : \Sigma^n \to \{0,1\}$ is an election rule which takes as input these preferences and aggregates them into a choice between two candidates, then our generalized KKL theorem says that one voter has influence $\Omega(\log(n)/n)$ provided both candidates have a non-negligible chance of winning *a priori*.

What about larger coalitions? Before stating our result, it is natural in the context of elections to assume that voters preferences are also binary valued (i.e. $\Sigma = \{\pm 1\}$) and that the function f is *monotone*, i.e. if $x \le y$ then $f(x) \le f(y)$. Under these assumptions, the following corollary shows in particular that a coalition of size $\omega(n/\log(n))$ has influence $1 - o(1)$ on a fair election. It follows by iteratively applying our generalization of the KKL theorem, and generalizes Corollary 3.5 of [29] where the case of the uniform measure was considered.

Corollary 2.5 (Corollary 3.10 below)**.** *For any $n \ge 1$ and ν satisfying Assumption 1, the following is true. For any $\epsilon > 0$ and and monotone function $f : \{\pm 1\}^n \to \{0,1\}$ satisfying $\mathbb{E}_\nu[f] \ge \epsilon$, there exists a set of coordinates $S \subset [n]$ such that*

$$\mathbb{E}_{X \sim \nu}[f(X_{\sim S}, X_S \to 1)] \ge 1 - \epsilon$$

and

$$|S| \le \frac{4(1+b)\log(1/2\epsilon)}{\alpha_{b,\rho,\Delta}} \cdot \frac{n}{\log(n)}$$

where $\alpha_{b,\rho,\Delta} > 0$ is the constant (independent of n) from Theorem 3.8.

Here the notation $\mathbb{E}_{X \sim \nu}[f(X_{\sim S}, X_S \to 1)]$ refers to the expectation of $f(Y)$ where X is drawn from μ and $Y_i = 1$ for $i \in S$ while $Y_i = X_i$ for $i \notin S$.

2.4 Graph Partitioning, Spectral Gap, and Cheeger's Inequality

Our main results have a natural interpretation in terms of the theory of Markov chains and in terms of graph partitioning. We briefly explain what these are.

In the context of the theory of Markov chains, it is well known that in the setting of Theorem 2.2 the spectral gap of the continuous time Glauber dynamics and the Cheeger constant are both $\Theta(1)$. By this we mean that the Poincaré inequality

$$\mathrm{Var}(f) \lesssim \sum_j \|L_j f\|_2^2$$

holds up to a constant independent of the dimension (more precisely, the constant is $O(1/\rho)$). This inequality is tight up to constants, even for boolean functions, since it is saturated up to constants for any dictator function $f(x) = x_i$. It is natural to ask if there are many other partitions (Boolean functions) for which the ratio of the right hand side to the left hand side (i.e. the Rayleigh quotient up to constants) is $\Theta(1)$.

Our main result (Theorem 2.2) implies that Boolean functions all of whose influences are small (say less than δ) have a Rayleigh quotient that is $\Omega(\log(1/\delta))$, i.e.

$$\mathrm{Var}(f) \lesssim \frac{1}{\log(1/\delta)} \sum_j \|L_j f\|_2^2.$$

The term $\|L_j f\|_2^2$ is the same, up to constant factors, as the influence (see Lemma 3.9 below). Thus, for the Poincaré inequality to be a constant factor close to tight, it is necessary for the function to have at least one variable with high influence. This also has a dynamical interpretation: while for dictators, there will be a $\Theta(1/n)$ chance of a single step of the Glauber dynamics flipping the value of f at steady state, for balanced functions with small influences the probability will necessarily be larger! More precisely, starting from the stationary distribution, a single step of Glauber will have a chance that is $\Omega(\log(1/\delta)/n)$ of flipping the value of f. Likewise, if we run the dynamics for a sufficiently long time, the proportion of time steps at which the function value flips will be $\Omega(\log(1/\delta)/n)$. So functions without influential variables are more sensitive/less stable under the natural dynamics.

In a different language, if we look for a almost balanced partition of the state space that has minimal bisection (in the sense that the two parts have proportional measures), then our main theorem states that such a partition mush have a high influence variable. It is natural to ask if the statements above can be strengthened to state that any function that is close to an optimal bisection is close to a junta.

2.5 Comparison to the Results on Phase Transitions for Monotone Measures

We next compare our results to work by Graham and Grimmett [23] who proved a version of the KKL theorem and sharp thresholds for "monotonic" measures. Consider a monotone function $f : \{0,1\}^n \to \{0,1\}$ and a measure μ on $\{0,1\}^n$. Recall the definition of influence, Definition 2.1. We now define the *effect* $e_i(f,\mu)$ of a variable i on f under μ as $\mathrm{Cov}_\mu[f, x_i] = \mathbb{E}_\mu[f x_i] - \mathbb{E}_\mu[f] E_\mu[x_i]$ (note that this is $p(1-p)$ times the effect as defined in [25]). We note that

1. If μ is the uniform measure and f is monotone then the effect and the influence are the same up to a constant factor. If μ is a monotone measure in the sense of [23] and f is monotone, the size of the effect can be lower bounded by the influence using the FKG inequality (see [23]).
2. The paper [23] proves sharp phase transitions based on the effects. In [23], they do so by proving a version of KKL Theorem based on a reduction to the KKL theorem in the standard case of the uniform measure. We also mention the work of [15] who prove sharp thresholds by generalizing the results of [45] using effects. Interestingly, these results do not require any correlation decay of the measure, so unlike our results they do not require the log-Sobolev inequality. They also require monotonicity of the measure which our results do not.

There are very important differences between the interpretations of effects and influences. (The importance of this difference was also discussed by Graham and Grimmett [23] where they called effects and influences the "conditional influences" and "absolute influences" respectively.)

To compare influences and effects in a concrete setting, we consider the finite-volume Ising model with parameter β on the square lattice in dimension $d \geq 2$. In this (classical) setting, the vertices of our graph correspond to the integer elements of $[-L/2, L/2]^d$ where $L \geq 1$ is the sidelength of the box, and the edges E of the graph connect vertices which are neighbors in the square lattice, i.e. which are Euclidean distance 1 from each other. Note that there are $n = (L+1)^d$ many vertices in total. Given this graph, the ferromagnetic Ising model is the distribution on $\{\pm 1\}^n$ of the form:

$$\nu(x) \propto \exp\left(\beta \sum_{(i,j)\in E} x_i x_j\right).$$

Let $\beta_c(d)$ be the critical inverse temperature of the lattice Ising model in dimension d (see e.g. [18, 48]). Below β_c is the high-temperature/subcritical regime and above β_c is the low-temperature/supercritical regime of the model. Informally speaking, in the low temperature phase, the model exhibit symmetry breaking, a typical sample from the model lies either in a mostly $+$ phase or in a mostly $-$ phase, and because of this the Glauber dynamics mix torpidly.

Let f be a monotone function from $\{\pm 1\}^n \to \{0,1\}$ with variance $\Omega(1)$. The results of [23] imply that for all $\beta \geq 0$:

1. There exists a variable whose effect is at least $\Omega(\log n/n)$.
2. There exists a set S consisting of $O(n \,/\log n)$ many variables such that $E[f|X_S = +] = 1 - o(1)$.

As we will now illustrate, the analogous results with influences replaces by effects will fail badly due to the aforementioned phase transition in the Ising model.

The log-Sobolev inequality for this measure, see e.g. [37] allows us to apply our results to deduce that for $\beta < \beta_c$, i.e. in the *subcritical regime* of the model, we have that:

- There exists a variable whose influence is at least $\Omega(\log n/n)$.
- There exists a set S consisting of $O(n \,/\log n)$ many variables such that $E[f(X_{-S}, X_S \to 1)] = 1 - o(1)$.

On the other hand, when $\beta > \beta_c$, i.e. in the *supercritical regime*, it immediately follows from rigorous results on the large deviations of the magnetization in the Ising model [48, 7] that when f is the majority function:

- For every i, the effect of X_i is $\Theta(1)$.
- For every i, the influence of X_i is $\exp(-\Theta(L^{d-1}))$.
- For a uniformly random set S with $|S| = \omega(1)$ it holds that $E[f|X_S = +] = 1 - o(1)$.
- For every set S with $|S| = o(n)$ it holds that $E[f(X_{-S}, X_S \to 1)] = 0.5 + \exp(-\Theta(L^{d-1}))$.

This shows that our results cannot be proven without assuming rapid mixing of the dynamics.

Intuitively, for non-product measures there is a dramatic difference between fixing a variable and conditioning on a variable, as conditioning on a variable changes the measure and therefore changes all other variables. This shows that our results and the results of GC and DCRT and incomparable. In the setting where both our results and theirs apply (monotone measures which satisfy Assumption 1), our versions of Talagrand and KKL are stronger since the influences lower bound the effects.

3 Proof of Main Results

In this section, we prove all of our results. It was observed by Cordero-Erasquin and Ledoux [11] that Talagrand's inequality (and then KKL) can be deduced from an estimate of the form (5) below. The most important contribution of our work is to prove this estimate (Theorem 3.1) in our very general setting, which we do in Section 3.1 below. Given this estimate, we derive the generalized Talagrand's inequality and KKL in Section 3.2, and then show how to obtain the consequences for coalitions in Section 3.3.

3.1 Main functional inequality

The following is the main technical claim which implies Talagrand's inequality and KKL.

Theorem 3.1. *There exists absolute constants $c, c' > 0$ such that the following is true. For any ν satisfying Assumption 1, $f : \Sigma^n \to \mathbb{R}$, and for any positive $T \le c/b^2q^2\Delta^2$,*

$$\mathrm{Var}_\nu(f) \le \frac{c'q^2b^2}{1-e^{-\rho T}} \int_0^T \sum_{j=1}^n \|L_j f\|^2_{1+e^{-2\rho t}}\, dt. \tag{5}$$

Proof. Since the log-Sobolev inequality implies the Poincare inequality, we have that for any $T \ge 0$

$$\mathrm{Var}(f) = \mathrm{Var}(f) - \mathrm{Var}(H_T f) + \mathrm{Var}(H_T f) \le \mathrm{Var}(f) - \mathrm{Var}(H_T f) + e^{-\rho T}\mathrm{Var}(F)$$

and so

$$\mathrm{Var}(f) \le \frac{1}{1-e^{-\rho T}}[\mathrm{Var}(f) - \mathrm{Var}(H_T f)].$$

To upper bound $\mathrm{Var}(f)$, it thereby suffices to upper bound for some $T > 0$ the quantity

$$\mathrm{Var}(f) - \mathrm{Var}(H_T f) = 2\int_0^T \mathcal{E}(H_t f, H_t f)dt = \sum_i \int_0^T \mathbb{E}(L_i H_t f)^2 dt. \tag{6}$$

The first equality in the equation above holds for any Markov semigroup as proven in [11].

We recall the following fact, sometimes called the Hadamard or Baker-Hausdorff Lemma:

Lemma 3.2 (Proposition 3.35 of [26])**.** *For square matrices X, Y, we have $e^X Y e^{-X} = e^{[X,\cdot]}Y$.*

The following lemma computes the effect of commuting L_i and H_T.

Lemma 3.3. *For any $T \ge 0$ and $i \in [n]$ we have $L_i H_T = H_T M_{T,i}$ where*

$$M_{T,i} := \sum_{k=0}^\infty \frac{T^k}{k!} \sum_{(j_1,\dots,j_k)\in\mathcal{S}_{k,i}} [\cdots[[[L_i, P_{j_1}], P_{j_2}], P_{j_3}]\cdots P_{j_k}]. \tag{7}$$

Here

$$\mathcal{S}_{k,i} := \{(j_1,\dots,j_k) : j_i \in \mathcal{N}^+(\{i, j_1, \dots, j_{i-1}\})\} \tag{8}$$

and $\mathcal{N}^+(U)$ denotes the union of U and the neighbors of nodes U in the graph.

Proof. Note that by applying Lemma 3.2 to a negated matrix X, we have the identity for square matrices X, Y

$$e^{-X} Y e^{X} = e^{[\cdot, X]} Y.$$

Since $H_t = e^{tL}$, we therefore get

$$H_T^{-1} L_i H_T = \sum_{k=0}^{\infty} \frac{T^k}{k!} [L_i, L]^{(k)}$$

where $[L_i, L]^{(k)}$ denotes the iterated commutator of the following form: $[L_i, L]^{(0)} = L_i$ and $[L_i, L]^{(k)} = [[L_i, L]^{(k-1)}, L]$.

To compute the commutator, first observe

$$[L_i, L] = \sum_{j : i \sim j} [L_i, P_j]$$

since $L_j = P_j - I$ and P_i commutes with P_j when $i \not\sim j$. For the same reason, we have more generally that

$$[L_i, L]^{(k)} = \sum_{(j_1, \ldots, j_k) \in \mathcal{S}_{k,i}} [\cdots [[[L_i, P_{j_1}], P_{j_2}], P_{j_3}], \cdots P_{j_k}]$$

which proves the result. □

Lemma 3.4. *With the notation of* (8), $|\mathcal{S}_{k,i}| \leq (\Delta + 1)^k k^k$ *for any* i, k.

Proof. Observe that we can encode j_k as an element of $[k] \times [\Delta + 1]$ by choosing one of its predecessors $i, \ldots, j_{k-1}$ and specifying whether j_k is equal to that node or one of that node's Δ neighbors. Performing this encoding recursively proves the result. □

Therefore recalling the definition of $M_{t,i}$ in (7) to get the first equality and applying hypercontractivity to get the following inequality we have

$$
\begin{aligned}
&\int_0^T \|L_i H_t f\|_2^2 dt \\
&\quad = \int_0^T \|H_t M_{t,i} f\|_2^2 dt \\
&\quad \le \int_0^T \|M_{t,i} f\|_{1+e^{-2\rho t}}^2 dt \\
&\quad = \int_0^T \left\| \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{(j_1,\dots,j_k)\in\mathcal{S}_{i,k}} [\cdots[[[L_i, P_{j_1}], P_{j_2}], P_{j_3}], \cdots, P_{j_k}] f \right\|_{1+e^{-2\rho t}}^2 dt \\
&\quad \le \int_0^T \left(\sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{(j_1,\dots,j_k)\in\mathcal{S}_{i,k}} \|[\cdots[[[L_i, P_{j_1}], P_{j_2}], P_{j_3}], \cdots, P_{j_k}] f\|_{1+e^{-2\rho t}} \right)^2 dt \\
&\quad \le \left(\sum_{k=0}^{\infty} \frac{T^k}{k!} (\Delta+1)^k k^k \right) \int_0^T \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{(j_1,\dots,j_k)\in\mathcal{S}_{i,k}} \|[\cdots[[[L_i, P_{j_1}], P_{j_2}], P_{j_3}], \cdots, P_{j_k}] f\|_{1+e^{-2\rho t}}^2 dt \\
&\quad \le 2 \int_0^T \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{(j_1,\dots,j_k)\in\mathcal{S}_{i,k}} \|[\cdots[[[L_i, P_{j_1}], P_{j_2}], P_{j_3}], \cdots, P_{j_k}] f\|_{1+e^{-2\rho t}}^2 dt \qquad (9)
\end{aligned}
$$

where we used the triangle inequality, in the second-to-last step we applied the Cauchy-Schwarz inequality and Lemma 3.4, and in the last step we used the assumption that T is small compared to $1/\Delta^2$.

Lemma 3.5. *For any $p \ge 1$, $i \in [n]$, $k \ge 0$ and for $\mathcal{S}_{i,k}$ as defined in* (8)*, we have*

$$\sum_{(j_1,\dots,j_k)\in\mathcal{S}_{i,k}} \|[\cdots[[[L_i, P_{j_1}], P_{j_2}], P_{j_3}], \cdots P_{j_k}] f\|_p^2 \le 2(\Delta+1)^k (k+1)^{k+4} (2qb)^{2k+2} \max_{j: d_G(j,i)\le k} \|L_j f\|_p^2$$

Proof. For notational convenience, define $j_0 = i$. The result is trivial if $k = 0$. Otherwise, we can replace $L_i = P_i - I$ by P_i since the identity commutes with everything. Now observe that

$$
\begin{aligned}
&[\cdots[[[P_i, P_{j_1}], P_{j_2}], P_{j_3}], \cdots, P_{j_k}] \\
&\quad = P_{j_k}[\cdots[[[P_i, P_{j_1}], P_{j_2}], P_{j_3}], \cdots, P_{j_{k-1}}] - [\cdots[[[P_i, P_{j_1}], P_{j_2}], P_{j_3}], \cdots, P_{j_{k-1}}] P_{j_k} \\
&\quad = \sum_{\alpha} (-1)^{r(\alpha)} \left(P_{j_k} P_\alpha - P_\alpha P_{j_k} \right)
\end{aligned}
$$

where α ranges over a subset of permutations of $(j_0, \dots, j_{k-1})$ of size at most 2^k that arise when expanding out the iterated commutator, and $r(\alpha) \in \{0, 1\}$ encodes the corresponding sign of this term. Let

$$K_{j_0,\dots,j_k}(x) = \{y : y_{\sim\{j_0,\dots,j_k\}} = x_{\sim\{j_0,\dots,j_k\}}\}$$

denote the set of spin configurations which disagree with x only within $\{i, j_1, \dots, j_k\}$. Using that the dynamics only update sites $j_0, \dots, j_k$ and using the triangle inequality we have that

$$
\begin{aligned}
|([P_{j_k} P_\alpha - P_\alpha P_{j_k}] f)(x)| &\le \max_{y,y' \in K_{i,j_1,\dots,j_k}(x)} |f(y) - f(y')| \\
&\le (k+1) \max_{z,z' \in K_{i,j_1,\dots,j_k}(x): d_H(z,z')=1} |f(z) - f(z')|.
\end{aligned}
$$

Hence taking the average over x, we find

$$\begin{aligned}
&\sum_x \nu(x)|([P_{j_k}P_\alpha - P_\alpha P_{j_k}]f)(x)|^p \\
&\quad\le (k+1)^p \sum_x \nu(x) \max_{z,z'\in K_{j_0,\dots,j_k}(x):d_H(z,z')=1} |f(z)-f(z')|^p \\
&\quad\le 2^p(k+1)^p \sum_x \nu(x) \max_{z\in K_{j_0,\dots,j_k}(x),\ell\in\{j_0,\dots,j_k\}} |(L_\ell f)(z)|^p \\
&\quad\le 2^p(k+1)^p \sum_x \nu(x) \max_{z\in K_{j_0,\dots,j_k}(x)} (|(L_{j_0}f)(z)|+\cdots+|(L_{j_k}f)(z)|)^p \\
&\quad\le 2^p(k+1)^p(qb)^{k+1} \sum_z \nu(z)(|(L_{j_0}f)(z)|+\cdots+|(L_{j_k}f)(z)|)^p
\end{aligned}$$

where in the second to last step we used Lemma 3.6, and we arrived at the last step by considering the z which achieves the inner maximum, and used the fact that $\nu(x)\le b^{k+1}\nu(z)$ and that there are at most q^{k+1} such x for each z. Hence by the L_p triangle inequality, $p\ge 1$, and $1\le b$,

$$\|[P_{j_k}P_\alpha - P_\alpha P_{j_k}]f\|_p \le 2(k+1)(qb)^{k+1}\sum_{r=0}^k \|L_{j_r}f\|_p \le 2(k+1)^2(qb)^{k+1}\max_r \|L_{j_r}f\|_p.$$

Using that α ranges over a set of size at most 2^k, we find by the L_p triangle inequality

$$\|[\cdots[[[P_i,P_{j_1}],P_{j_2}],P_{j_3}],\cdots,P_{j_k}]f\|_p \le \sum_\alpha \|P_{j_k}P_\alpha - P_\alpha P_{j_k}]\|_p \le 2(k+1)^2(2qb)^{k+1}\max_{0\le r\le k}\|L_{j_r}f\|_p$$

and using Lemma 3.4 we have

$$\sum_{(j_1,\dots,j_k)\in\mathcal{S}_{i,k}} \|[\cdots[[[P_i,P_{j_1}],P_{j_2}],P_{j_3}],\cdots P_{j_k}]f\|_p^2 \le 2(\Delta+1)^k(k+1)^{k+4}(2qb)^{2k+2}\max_{j:d_G(j,i)\le k}\|L_jf\|_p^2$$

as desired. □

Lemma 3.6. *Suppose ν is a distribution on Σ^n. For any function $f:\Sigma^n\to\mathbb{R}$, and $y,z\in\Sigma^n$ differing only at site i we have*

$$\frac{1}{2}|f(y)-f(z)| \le \max\{|(L_if)(y)|,|(L_if)(z)|\}.$$

For any $x\in\Sigma^n$ we have

$$|(L_if)(x)| \le \max_{y,z:x_{\sim i}=y_{\sim i}=z_{\sim i}} |f(y)-f(z)|.$$

Proof. Expanding the definition, we have

$$(L_if)(x) = (P_if)(x) - f(x) = \mathbb{E}[f(X)\mid X_{\sim i}=x_{\sim i}] - f(x)$$

so the latter bound follows immediately, and the former bound follows from the triangle inequality as

$$|f(y)-f(z)| \le |(L_jf)(y)| + |(L_jf)(z)| \le 2\max\{|(L_jf)(y)|,|(L_jf)(z)|\}.$$

□

Using Lemma 3.5, if $T \le c/q^2b^2\Delta^2$ for some absolute constant $c > 0$, we have for all $t \le T$ that for some constant $c' > 0$,

$$\sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{(j_1,\dots,j_k)\in\mathcal{S}_{i,k}} \|[\cdots[[[L_i, P_{j_1}], P_{j_2}], P_{j_3}], \cdots, P_{j_k}]f\|^2_{1+e^{-2\rho t}}$$
$$\le q^2b^2 \sum_{j=1}^{n} (c'/\Delta)^{d_G(j,i)} \|L_j f\|^2_{1+e^{-2\rho t}} \tag{10}$$

and summing over i and using that the number of nodes at exactly distance k from node j is at most Δ^k, this gives

$$2\sum_i \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{(j_1,\dots,j_k)\in\mathcal{S}_{i,k}} \|[\cdots[[[L_i, P_{j_1}], P_{j_2}], P_{j_3}], \cdots, P_{j_k}]f\|^2_{1+e^{-2\rho t}} \le c'q^2b^2 \sum_{j=1}^{n} \|L_j f\|^2_{1+e^{-2\rho t}}.$$

Hence, recalling (9), we have for $T \le c/q^2b^2\Delta^2$ that

$$\sum_i \int_0^T \|L_i H_t f\|_2^2 dt \le c'q^2b^2 \int_0^T \sum_{j=1}^{n} \|L_j f\|^2_{1+e^{-2\rho t}}\, dt$$

which, recalling (6), gives the desired bound

$$\mathrm{Var}(f) \le \frac{c'q^2b^2}{1-e^{-\rho T}} \int_0^T \sum_{j=1}^{n} \|L_j f\|^2_{1+e^{-2\rho t}}\, dt.$$

□

3.2 Generalized Talagrand and KKL Inequalities

We now show how to deduce the Talagrand and KKL inequalities from Theorem 3.1. The proof of these implications follows from the work of Cordero-Erausquin and Ledoux [12] and is reproduced for convenience. The first result generalizes Talagrand's inequality:

Theorem 3.7. *For any $n \ge 1$, ν satisfying Assumption 1, and any $f : \Sigma^n \to \mathbb{R}$, we have*

$$\mathrm{Var}_\nu(f) \le \frac{Cq^4b^4\Delta^2}{\rho} \sum_j \frac{\|L_j f\|_2^2}{1 + \log(\|L_j f\|_2/\|L_j f\|_1)} \tag{11}$$

for some absolute constant $C > 0$.

Proof. Making the change of variables $p = 1 + e^{-2\rho t}$, $dp = -2\rho e^{-2\rho t} dt$ and assuming $T \le 1/2\rho$ we have by Holder's inequality

$$\int_0^T \|L_j f\|^2_{1+e^{-2\rho t}} dt \le \frac{2}{\rho} \int_1^2 \|L_j f\|_p^2 dp \le \frac{2}{\rho} \|L_j f\|_2^2 \int_1^2 d_j^{2\theta(p)} dp$$

where $1/p = \theta + (1-\theta)/2 = (1+\theta)/2$ and

$$d_j := \|L_j f\|_1/\|L_j f\|_2 \le 1.$$

Note that

$$\frac{d\theta}{dp} = -2/p^2$$

so making the change of variables $s = 2\theta(p)$, $ds = (-4/p^2)dp$ we have

$$\begin{aligned}\int_1^2 d_j^{2\theta(p)} dp &\le \int_0^2 d_j^s (p(s)^2/4) ds \\ &\le \int_0^2 d_j^s ds = \frac{1 - d_j^2}{\log(1/d_j)} = \frac{(1-d_j^2)(1 + 1/\log(1/d_j))}{1 + \log(1/d_j)} \le \frac{2}{1 + \log(1/d_j)}.\end{aligned}$$

hence

$$\int_0^T \|L_j f\|^2_{1+e^{-2\rho t}} dt \le \frac{4}{\rho(1 + \log(1/d_j))}.$$

Combining with Theorem 3.1, we have for $T = c/q^2 b^2 \Delta^2$ that for some absolute constant $C > 0$

$$\mathrm{Var}(f) \le \frac{Cq^4 b^4 \Delta^2}{\rho} \sum_j \frac{\|L_j f\|_2^2}{1 + \log(\|L_j f\|_2 / \|L_j f\|_1)}$$

which proves the analogue of Talagrand's inequality. □

Now we generalize KKL:

Theorem 3.8. *There exists $\alpha_{b,\rho,\Delta,q} > 0$ such that the following is true. For any $n \ge 1$, ν satisfying Assumption 1, and any $f : \Sigma^n \to \{0,1\}$, there exists a coordinate $k \in [n]$ such that*

$$I_k(f) \ge \alpha_{b,\rho,\Delta,q} \mathrm{Var}(f) \log(n)/n.$$

Proof. By combining Lemma 3.9 with Theorem 3.7 we have that

$$\mathrm{Var}(f) \le C \sum_j \frac{I_j(f)}{1 - \log(bq\sqrt{I_j(f)})} \tag{12}$$

where $C = C_{b,\rho,\Delta,q} > 0$. Fix b, ρ, Δ, q and suppose for contradiction that the conclusion of the theorem is false. The conclusion of the theorem is trivially true if $n = 1$, so it must be that for any $\alpha \in [0,1]$ there exists $n \ge 2$, ν satisfying Assumption 1, and $f : \Sigma^n \to \{0,1\}$ so that

$$I_k(f) \le \alpha \mathrm{Var}(f) \log(n)/n$$

for all $k \in [n]$. In particular $I_k(f) \le \alpha \log(n)/n$ since $\mathrm{Var}(f) \le 1$. Combining with (12) and dividing through by $\mathrm{Var}(f)$, we have

$$\begin{aligned}1 &\le \frac{C\alpha \log(n)}{1 - \log(bq\sqrt{\alpha \log(n)/n})} \\ &= \frac{C\alpha \log n}{1 - \log(bq\alpha^{1/2}) + (1/2)[\log(n) - \log\log(n)]} \\ &= \frac{C\alpha}{1/\log(n) - \log(bq\alpha^{1/2})/\log(n) + (1/2)[1 - [\log\log(n)]/\log(n)]}.\end{aligned}$$

which is a contradiction for any

$$\alpha < \min\left\{\frac{1}{b^2q^2}, \frac{1}{C}\inf_{n\geq 2}[1/\log(n) + (1/2)[1 - [\log\log(n)]/\log(n)]]\right\}.$$

□

Lemma 3.9. *For $f : \Sigma^n \to \{0,1\}$ and any ν satisfying Assumption 1, we have for any $p \geq 1$*

$$I_i(f) \geq \mathbb{E}|L_i f|^p \geq \frac{1}{(qb)^p} I_i(f)$$

Proof. Recall that

$$I_i(f) = \Pr_{X\sim\nu}[\exists x_i' \in \Sigma, f(X) \neq f(X_1, \ldots, X_{i-1}, x_i', X_{i+1}, \ldots, X_n)].$$

Given $x \in \Sigma^n$, if $f(x) = f(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n)$ for all $x_i' \in \Sigma$ then this means that $(L_i f)(x) = 0$. Since $|L_i f| \leq 1$, this implies that

$$|(L_i f)(x)| \leq \mathbb{1}(\exists x_i' \in \Sigma, f(x) \neq f(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n)).)$$

On the other hand, if there exists some x_i' such that $f(x) \neq f(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n)$ then this implies that

$$|(L_i f)(x)| = |f(x) - \mathbb{E}[f(X) \mid X_{\sim i} = x_{\sim i}]| \geq \Pr(X_i = x_i' \mid X_{\sim i} = x_{\sim i}) \geq 1/qb$$

by (1). Therefore

$$\frac{1}{qb}\mathbb{1}(\exists x_i' \in \Sigma, f(x) \neq f(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n)) \leq |(L_i f)(x)|$$

Hence taking expectation over X we have for any $p \geq 1$

$$I_i(f) \geq \mathbb{E}|L_i f|^p \geq \frac{1}{(qb)^p} I_i(f)$$

as claimed. □

3.3 Application to coalitions

We now discuss the application of our result to the existence of coalitions for monotone voting rules. In this section, we restrict to the case of $\Sigma = \{\pm 1\}$ and recall that a function $f : \{\pm 1\}^n \to \mathbb{R}$ is *monotone* if

$$f(x) \leq f(y)$$

for any pair such that $x \leq y$ coordinatewise.

The following corollary shows in particular that a coalition of size $\omega(n/\log(n))$ has influence $1 - o(1)$ on a fair election. It follows by iteratively applying our generalization of the KKL theorem, and generalizes Corollary 3.5 of [29] where the case of the uniform measure was considered.

Corollary 3.10. *For any $n \geq 1$ and ν satisfying Assumption 1, the following is true. For any $\epsilon > 0$ and and monotone function $f : \{\pm 1\}^n \to \{0,1\}$ satisfying $\mathbb{E}_\nu[f] \geq \epsilon$, there exists a set of coordinates $S \subset [n]$ such that*

$$\mathbb{E}_{X\sim\nu}[f(X_{\sim S}, X_S \to 1)] \geq 1 - \epsilon$$

and

$$|S| \leq \frac{4(1+b)\log(1/2\epsilon)}{\alpha_{b,\rho,\Delta}} \cdot \frac{n}{\log(n)}$$

where $\alpha_{b,\rho,\Delta} > 0$ is the constant (independent of n) from Theorem 3.8.

Proof. We construct a sequence of sets $S_0, S_1, \ldots$ iteratively. Let $S_0 = \{\}$. For each $t \geq 0$, define $f_t(x) = f(x_{\sim S_t}, 1_{S_t})$, i.e. f_t is the same as f except that it ignores the input x_{S_t} and replaces it by all-ones. Either

$$\Pr_{X\sim\nu}[f_t = 1] \geq 1 - \epsilon$$

or we define a set S_{t+1} in the following way. By Theorem 3.8, there exists some $k_t \in [n]$ such that

$$I_{k_t}(f_t) \geq \alpha \mathrm{Var}(f_t)\frac{\log(n)}{n}$$

where $\alpha = \alpha_{b,\rho,\Delta} > 0$ does not depend on n, and we let $S_{t+1} = S_t \cup k_t$. Now defining $f_{t+1}(x) = f(x_{\sim S_{t+1}}, 1_{S_{t+1}})$, we have by monotonicity that

$$\Pr_\nu(f_{t+1} = 1) = \Pr_\nu(f_t = 1) + \Pr_\nu(f_{t+1} > f_t).$$

Furthermore,

$$\begin{aligned}
\Pr_\nu(f_{t+1} > f_t) &= \mathbb{E}_{X\sim\nu}[1(f_t(X_{\sim k_t}, 1) > f_t(X))] \\
&= \mathbb{E}_{X\sim\nu}[1(f_t(X_{\sim k_t}, 1) > f_t(X)) \cdot 1(X_{k_t} = -1)] \\
&= \mathbb{E}_{X\sim\nu}[1(f_t(X_{\sim k_t}, 1) > f_t(X_{\sim k_t}, -1)) \cdot 1(X_{k_t} = -1)] \\
&= \mathbb{E}_{X\sim\nu}[1(f_t(X_{\sim k_t}, 1) > f_t(X_{\sim k_t}, -1)) \cdot \Pr(X_{k_t} = -1 \mid X_{\sim k_t})] \\
&\geq \frac{I_{k_t}(f_t)}{1+b}
\end{aligned}$$

where in the last equality we applied the law of total expectation, and in the final step we used that

$$\Pr_{X\sim\nu}(X_{k_t} = -1 \mid X_{\sim k_t}) \geq \frac{1}{1+b}$$

by Assumption 1. Therefore, if $p_t = \Pr(f_t = 1)$ we have that

$$p_{t+1} \geq p_t + \frac{\alpha}{1+b} p_t(1-p_t)\frac{\log(n)}{n}.$$

It follows that if $p_t < 1/2$, $p_{t+1} \geq (1 + \frac{\alpha\log(n)}{(1+b)n})p_t \geq \exp\left(\frac{\alpha\log(n)}{2(1+b)n}\right)p_t$, so $p_t > 1/2$ for any $t > \frac{2(1+b)n}{\alpha\log(n)}\log(1/2\epsilon)$. By a symmetrical argument, we have that $p_t \geq 1-\epsilon$ for $t > \frac{4(1+b)n}{\alpha\log(n)}\log(1/2\epsilon)$. $\square$

4 Some examples

There is a vast literature establishing log-Sobolev inequalities for spin systems on the hypercube. For concreteness, we give a few examples of settings where the log-Sobolev constant is known to be bounded, and as a consequence our results can be applied.

Sparse Markov random field under ℓ_2-Dobrushin uniqueness condition. Suppose that ν is a Markov random field on a graph of maximum degree Δ with n vertices, and define the Dobrushin matrix $A \in \mathbb{R}^{n \times n}$ to have zero diagonal and off-diagonal entries

$$A_{ij} = \max_{y \in \Sigma^n, z} d_{TV}(\Pr_\nu[X_i = \cdot \mid X_{\sim i} = y_{\sim i}], \Pr_\nu[X_i = \cdot \mid X_{\sim i,j} = y_{\sim i,j}, X_j = z]).$$

Suppose also that ν satisfies the b-bounded marginal assumption from Assumption 1. Then if $\|A\|_{OP} < 1$, it was shown by Marton [38] that ν satisfies the log-Sobolev inequality with log-Sobolev constant polynomial in b and q.

Special case: Ising under Dobrushin's uniqueness threshold. As a special case of the above, suppose that

$$\nu(x) \propto \exp\left(\sum_{(i,j) \in E} J_{ij} x_i x_j + \sum_i h_i x_i\right)$$

is a probability measure on the hypercube $\{\pm 1\}^n$ parameterized by J, h where E is the edge set of a sparse graph of maximum degree Δ. If $\sum_j |J_{ij}| < 1 - \delta$ for all i, and $\sum_i |h_i| < H$, one can directly show from the definition of the model that it is marginally bounded with $b = \exp(O(1 + H))$ and satisfies Dobrushin's uniqueness condition (by applying Gershgorin's disk theorem), hence our result applies. Note that we do not need any assumption on the sign of the interactions J_{ij} or external field h_i.

Additional references. There are many settings outside of Dobrushin's uniqueness condition where the log-Sobolev inequality is known. For example, the case of the lattice Ising model we discussed earlier is not contained in this regime. See e.g. [53, 37, 8, 4, 17, 5] for a few relevant references. In particular, by the result of Chen, Liu, and Vigoda [8], the log-Sobolev constant can be bounded purely as a function of b, Δ and the "spectral independence" constant of the distribution ν — so our assumption that the log-Sobolev constant is bounded can be replaced by the assumption of spectral independence.

Acknowledgment F.K. was supported in part by NSF award CCF-1704417, NSF award IIS-1908774, and N. Anari's Sloan Research Fellowship. D.M. was supported by a Sloan Research Fellowship, NSF CCF award 2227876 and NSF CAREER award 2239160. E.M. is partially supported by Vannevar Bush Faculty Fellowship award ONR-N00014-20-1-2826, ARO MURI N000142412742 and a Simons Investigator Award in Mathematics. N. L. was supported by the Israel Science Foundation (grant no. 1980/22).

References

[1] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. *SIAM Journal on Computing*, (0):FOCS2021, 2021.

[2] Mitali Bafna, Max Hopkins, Tali Kaufman, and Shachar Lovett. Hypercontractivity on high dimensional expanders. In *STOC*, volume 22, page 54th, 2022.

[3] Dominique Bakry, Ivan Gentil, Michel Ledoux, et al. *Analysis and geometry of Markov diffusion operators*, volume 103. Springer, 2014.

[4] Roland Bauerschmidt and Thierry Bodineau. A very simple proof of the lsi for high temperature spin systems. *Journal of Functional Analysis*, 276(8):2582–2588, 2019.

[5] Roland Bauerschmidt and Benoit Dagallier. Log-sobolev inequality for near critical ising models. *arXiv preprint arXiv:2202.02301*, 2022.

[6] M. Ben-Or and N. Linial. Collective coin flipping. In S. Micali, editor, *Randomness and Computation.* Academic Press, New York, 1990.

[7] Thierry Bodineau. Slab percolation for the ising model. *arXiv preprint math/0309300*, 2003.

[8] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Optimal mixing of glauber dynamics: Entropy factorization via high-dimensional expansion. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1537–1550, 2021.

[9] GL Columbu, A De Martino, and A Giansanti. Nature and statistics of majority rankings in a dynamical model of preference aggregation. *Physica A: Statistical Mechanics and its Applications*, 387(5-6):1338–1344, 2008.

[10] Dario Cordero-Erausquin and Alexandros Eskenazis. Talagrand's influence inequality revisited. *Analysis & PDE*, 16(2):571–612, 2023.

[11] Dario Cordero-Erausquin and Michel Ledoux. Hypercontractive measures, talagrand's inequality, and influences. In *Geometric Aspects of Functional Analysis*, volume 2050 of *Lecture Notes in Mathematics.* 2012.

[12] Dario Cordero-Erausquin and Michel Ledoux. Hypercontractive measures, talagrand's inequality, and influences. In *Geometric Aspects of Functional Analysis: Israel Seminar 2006–2010*, pages 169–189. Springer, 2012.

[13] Nikhil R Devanur, Subhash A Khot, Rishi Saket, and Nisheeth K Vishnoi. Integrality gaps for sparsest cut and minimum linear arrangement problems. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 537–546, 2006.

[14] R. L. Dobrushin and S. B. Shlosman. Constructive criterion for uniqueness of a Gibbs field. In J. Fritz, A. Jaffe, and D. Szasz, editors, *Statistical Mechanics and dynamical systems*, volume 10, pages 347–370. 1985.

[15] Hugo Duminil-Copin, Aran Raoufi, and Vincent Tassion. Sharp phase transition for the random-cluster and potts models via decision trees. *Annals of Mathematics*, 189(1):75–99, 2019.

[16] Charilaos Efthymiou, Thomas P Hayes, Daniel Stefankovic, Eric Vigoda, and Yitong Yin. Convergence of mcmc and loopy bp in the tree uniqueness region for the hard-core model. *SIAM Journal on Computing*, 48(2):581–643, 2019.

[17] Ronen Eldan, Frederic Koehler, and Ofer Zeitouni. A spectral condition for spectral gap: fast mixing in high-temperature ising models. *Probability theory and related fields*, 182(3-4):1035–1051, 2022.

[18] Richard S Ellis. *Entropy, large deviations, and statistical mechanics*, volume 1431. Taylor & Francis, 2006.

[19] Yuval Filmus, Lianna Hambardzumyan, Hamed Hatami, Pooya Hatami, and David Zuckerman. Biasing boolean functions and collective coin-flipping protocols over arbitrary product distributions. *arXiv preprint arXiv:1902.07426*, 2019.

[20] Serge Galam. Rational group decision making: A random field ising model at t= 0. *Physica A: Statistical Mechanics and its Applications*, 238(1-4):66–80, 1997.

[21] Christophe Garban and Jeffrey E Steif. *Noise sensitivity of Boolean functions and percolation*, volume 5. Cambridge University Press, 2014.

[22] William V Gehrlein. *Condorcet's paradox*. Springer, 2006.

[23] Benjamin T Graham and Geoffrey R Grimmett. Influence and sharp-threshold theorems for monotonic measures. *Annals of probability*, 34(5):1726–1745, 2006.

[24] Tom Gur, Noam Lifshitz, and Siqi Liu. Hypercontractivity on high dimensional expanders. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 176–184, 2022.

[25] O. Häggström, G. Kalai, and E. Mossel. A law of large numbers for weighted majority. *Advances in Applied Mathematics*, 37(1):112–123, 2006.

[26] Brian C Hall. *Lie groups, Lie algebras, and representations*. Springer, 2013.

[27] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.

[28] Paata Ivanisvili and Yonathan Stone. The kkl inequality and rademacher type 2. *arXiv preprint arXiv:2204.05195*, 2022.

[29] J. Kahn, G. Kalai, and N. Linial. The influence of variables on boolean functions. In *Proceedings of the 29th Annual Symposium on Foundations of Computer Science*, pages 68–80, 1988.

[30] Gil Kalai and Shmuel Safra. Threshold phenomena and influence with some perspectives from mathematics. *Computer Science and Economics*, 2004.

[31] Tali Kaufman and Dor Minzer. Optimal testing of generalized reed-muller codes in fewer queries. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, 2023.

[32] Peter Keevash, Noam Lifshitz, Eoin Long, and Dor Minzer. Global hypercontractivity and its applications. *arXiv preprint arXiv:2103.04604*, 2021.

[33] Peter Keevash, Noam Lifshitz, and Dor Minzer. On the largest product-free subsets of the alternating groups. *arXiv preprint arXiv:2205.15191*, 2022.

[34] Subhash Khot, Dor Minzer, and Muli Safra. Pseudorandom sets in grassmann graph have near-perfect expansion. *Annals of Mathematics*, 198(1):1–92, 2023.

[35] Frederic Koehler and Elchanan Mossel. A phase transition in arrow's theorem. *arXiv preprint arXiv:2004.12580*, 2020.

[36] Steffen L Lauritzen. *Graphical models*, volume 17. Clarendon Press, 1996.

[37] Fabio Martinelli. Lectures on glauber dynamics for discrete spin models. *Lectures on probability theory and statistics (Saint-Flour, 1997)*, 1717:93–191, 1999.

[38] Katalin Marton. Logarithmic sobolev inequalities in discrete product spaces. *Combinatorics, Probability and Computing*, 28(6):919–935, 2019.

[39] Ashley Montanaro and Tobias J Osborne. Quantum boolean functions. *arXiv preprint arXiv:0810.2435*, 2008.

[40] Ravi Montenegro, Prasad Tetali, et al. Mathematical aspects of mixing times in markov chains. *Foundations and Trends® in Theoretical Computer Science*, 1(3):237–354, 2006.

[41] Elchanan Mossel. Probabilistic view of voting, paradoxes, and manipulation. *BULLETIN OF THE AMERICAN MATHEMATICAL SOCIETY*, 59(3):297–330, 2022.

[42] Elchanan Mossel and Allan Sly. Rapid mixing of gibbs sampling on graphs that are sparse on average. *Random Structures & Algorithms*, 35(2):250–270, 2009.

[43] Elchanan Mossel, Dror Weitz, and Nicholas Wormald. On the hardness of sampling independent sets beyond the tree threshold. *Probability Theory and Related Fields*, 143(3-4):401–439, 2009.

[44] Ryan O'Donnell. *Analysis of boolean functions*. Cambridge University Press, 2014.

[45] Ryan O'Donnell, Michael E. Saks, Oded Schramm, and Rocco A. Servedio. Every decision tree has an influential variable. In *FOCS*, pages 31–39, 2005.

[46] Ryan O'Donnell and Karl Wimmer. Kkl, kruskal–katona, and monotone nets. *SIAM Journal on Computing*, 42(6):2375–2399, 2013.

[47] Ryan O'Donnell and Karl Wimmer. Sharpness of kkl on schreier graphs. 2013.

[48] Agoston Pisztora. Surface order large deviations for ising, potts and percolation models. *Probability Theory and Related Fields*, 104:427–466, 1996.

[49] Giacomo Raffaelli and Matteo Marsili. Statistical mechanics model for the emergence of consensus. *Physical Review E*, 72(1):016114, 2005.

[50] Cambyse Rouzé, Melchior Wirth, and Haonan Zhang. Quantum talagrand, kkl and friedgut's theorems and the learnability of quantum boolean functions. *arXiv preprint arXiv:2209.07279*, 2022.

[51] Alexander Russell, Michael Saks, and David Zuckerman. Lower bounds for leader election and collective coin-flipping in the perfect information model. In *Proceedings of the thirty-first annual ACM symposium on Theory of Computing*, pages 339–347, 1999.

[52] Sushant Sachdeva and Madhur Tulsiani. Cuts in cartesian products of graphs. *arXiv preprint arXiv:1105.3383*, 2011.

[53] Daniel W Stroock and Boguslaw Zegarlinski. The logarithmic sobolev inequality for discrete spin systems on a lattice. *Communications in Mathematical Physics*, 149:175–193, 1992.

[54] Ramon Van Handel. Probability in high dimension. *Lecture Notes (Princeton University)*, 2014.