ON REVERSE HYPERCONTRACTIVITY

ELCHANAN MOSSEL, KRZYSZTOF OLESZKIEWICZ, AND ARNAB SEN

ABSTRACT. We study the notion of reverse hypercontractivity. We show that reverse hypercontractive inequalities are implied by standard hypercontractive inequalities as well as by the modified log-Sobolev inequality. Our proof is based on a new comparison lemma for Dirichlet forms and an extension of the Strook-Varapolos inequality.

A consequence of our analysis is that *all* simple operators $L = Id - \mathbb{E}$ as well as their tensors satisfy uniform reverse hypercontractive inequalities. That is, for all $q < p < 1$ and every positive valued function f for $t \geq \log \frac{1-q}{1-p}$ we have $\|e^{-tL} f\|_q \geq \|f\|_p$. This should be contrasted with the case of hypercontractive inequalities for simple operators where t is known to depend not only on p and q but also on the underlying space.

The new reverse hypercontractive inequalities established here imply new mixing and isoperimetric results for short random walks in product spaces, for certain card-shufflings, for Glauber dynamics in high-temperatures spin systems as well as for queueing processes. The inequalities further imply a quantitative Arrow impossibility theorem for general product distributions and inverse polynomial bounds in the number of players for the non-interactive correlation distillation problem with m-sided dice.

1. Introduction

1.1. **Background.** Log Sobolev and hypercontractive inequalities play a fundamental role in a number of areas in analysis including the study of Gaussian processes (see. e.g. [Gro78, Jan97]), analysis of Markov chains (see e.g. [SC97]) and discrete Fourier analysis starting in [KKL88, Tal94].

One of the first and most useful hypercontractive inequalities due to Bonami-Nelson-Gross-Beckner [Bon70, Nel73, Gro75, Bec75] states that if $(\Omega, \mu) = (\{0,1\}, \frac{1}{2}(\delta_0 + \delta_1))$ then the operator $T_t = e^{t(Id - \mathbb{E})}$ satisfies

$$\|T_t f\|_p \leq \|f\|_q, \quad \forall f : \Omega \to \mathbb{R},\, p > q > 1,\, t \geq \frac{1}{2}\frac{q-1}{p-1}. \tag{1.1}$$

The fact that the inequality tensorises allows to apply it in discrete Fourier analysis (starting with [KKL88]) and even earlier in the study of Gaussian processes. Extending (1.1) to other spaces turned out to be a non-trivial task.

2000 *Mathematics Subject Classification.* Primary: 60E15, Secondary: 60J27.

The first author is supported by DMS 0548249 (CAREER) award, by DOD ONR grant N000141110140, by ISF grant 1300/08 and by a Minerva Grant. Most of this work was conducted when the author was at the Weizmann institute.

The second author is partially supported by Polish MNiSzW Grant N N201 397437.

For the case of the spaces $(\Omega,\mu) = (\{0,1\}, \alpha\delta_0 + (1-\alpha)\delta_1)$, with $\alpha \le 1/2$, the first bounds were obtained by Talagrand [Tal94] - these bounds had an extra $\log(1/\alpha)$ factor. Exact formula were obtained by Oleszkiewicz [Ole03] in the cases where either $p > 2 = q$ or $p = 2 > q > 1$. The analysis of [Ole03] implies in particular that there exist constants $C_1, C_2, C_3, C_4 \in (0,\infty)$ such that for all $\alpha \in (0,1/2)$:

(1.2)
$$t \ge C_2(\frac{1}{q}-\frac{1}{2})\log(\frac{1}{\alpha}) \implies \forall f : \Omega \to \mathbb{R}, \|T_t f\|_2 \le \|f\|_q \implies t \ge C_1(\frac{1}{q}-\frac{1}{2})\log(\frac{1}{\alpha})$$

and

(1.3)
$$t \ge C_3(\frac{1}{2}-\frac{1}{p})\log(\frac{1}{\alpha}) \implies \forall f : \Omega \to \mathbb{R}, \|T_t f\|_p \le \|f\|_2 \implies t \ge C_4(\frac{1}{2}-\frac{1}{p})\log(\frac{1}{\alpha})$$

Wolff then generalized the results above [Wol07] to general discrete spaces. Letting (Ω,μ) with $\alpha = \min_{\omega\in\Omega}\mu\{\omega\}$ and $p > q > 1$ he proved that there exist constants $C_5, C_6 \in (0,\infty)$ such that for all $p > q > 1$,

(1.4)
$$t \ge C_5(\frac{1}{q}-\frac{1}{p})\log(\frac{1}{\alpha}) \implies \forall f : \Omega \to \mathbb{R}, \|T_t f\|_p \le \|f\|_q \implies t \ge C_6(\frac{1}{q}-\frac{1}{p})\log(\frac{1}{\alpha})$$

Note that (1.2), (1.3) and (1.4) all have a factor of $\log(1/\alpha)$ depending on the smallest atom in space. This dependency is present in many applications of hypercontractivity starting with [Tal94]. We note further that the same dependency is arrived at using the exact calculation of the log-Sobolev constant of simple operators by Diaconis and Saloff-Coste [DSC96].

A seemingly obscure 'reverse' hypercontractivity is shortly proved and discussed in a paper by Borell [Bor82] in the 80's. This result, proven for the measure $(\Omega,\mu) = (\{0,1\}, \frac{1}{2}(\delta_0+\delta_1))$, states that

$$\|T_t f\|_q \ge \|f\|_p, \quad \forall f : \Omega \to \mathbb{R}_+,\ 1 > p > q,\ t \ge \frac{1}{2}\frac{q-1}{p-1}. \tag{1.5}$$

This inequality which also tensorizes is indeed 'reverse' in many ways. Not only the inequality goes 'the other way' and $p > q$, it is also the case that p and q are less than 1 (indeed they may be negative(!); note however the function f has to take positive values).

As far as we know, Borell's result was first used in a paper published more than 20 years later [MOR+06], where it is used to analyze mixing of short random walks on the discrete cube $\{0,1\}^n$ as well as to provide tight bounds on the non-interactive correlation distillation problem (NICD).

Motivated by generalization of applications in [MOR+06] as well as by other applications that will be discussed later, we wish to extend Borell's results to other discrete probability spaces. Noting the similarity of the inequalities (1.1) and (1.5) it is tempting to conjecture (as the first named author have done) that the formula for hypercontracitivity and reverse hypercontractivity are 'the same'. In particular, for discrete spaces there is a

$\log(1/\alpha)$ dependency on the size of the smallest atom in space as in (1.2), (1.3) and (1.4).

This turns out to be far from true. In fact in our results show that for every discrete probability space (Ω, μ):

$$\|T_t f\|_q \le \|f\|_p, \quad \forall f : \Omega \to \mathbb{R}_+,\ 1 > p > q,\ t \ge \frac{q-1}{p-1}. \tag{1.6}$$

In particular, reverse hypercontractive inequalities hold uniformly for all probability spaces.

It is well known that hypercontractive inequalities are intimately related to log Sobolev inequalities and our proof of (1.6) is based on extension of this connection to 'norms' $p < 1$ (such extensions were noted before, see e.g. Bakry's lecture notes [BGM94]). At the heart of the proof is a new monotonicity result showing that under the appropriate normalization log-Sobolev inequalities are monotone in the norm p for all $p \in [0, 2]$. This result in turn is based on an extension of the Strook-Varapolous inequality to general norms. The result allows to show how reverse hypercontractive inequalities follow directly from standard hypercontractive inequalities and furthermore from standard log-Sobolev and modified log-Sobolev inequalities.

After we develop the theory of reverse hypercontractive inequalities, we derive a number of novel results regarding mixing of Markov chains run for short time from large sets, in general cubes, the symmetric group, Ising configurations (via Glauber dynamics). We further derive a quantitative Arrow's Theorem for general distributions and inverse polynomial bounds for the NICD problem for general m-sided dice. We proceed with formal definitions and statements of the main results.

1.2. **General setup.** We now turn to the general mathematical setup of the paper. Let (Ω, μ) be a finite probability space (with a natural σ-field of all subsets of Ω). We assume $\mu\{\omega\} > 0$ for $\omega \in \Omega$. Let $\mathbb{E}$ denote the expectation operator: $\mathbb{E}f = \int_\Omega f\, d\mu$. We also use the standard notation for the *variance* of f, $\mathrm{Var}(f) = \mathbb{E}f^2 - (\mathbb{E}f)^2$, and the *entropy* of $f > 0$, $Ent(f) = \mathbb{E}(f \log f) - \mathbb{E}f \cdot \log \mathbb{E}f$. Let $\mathcal{H}$ be the space of real-valued functions on Ω. Let $\mathcal{H}_{(0,\infty)}$ denote the strictly positive functions.

Definition 1.1. Let $L : \mathcal{H} \to \mathcal{H}$ be a linear operator such that

- $L1 = 0$ and
- L is self-adjoint with respect to $L^2(\Omega, \mu)$ structure, i.e. $\mathbb{E}fLg = \mathbb{E}gLf$ for any $f, g \in \mathcal{H}$, and
- L is positive semidefinite, i.e. $\mathbb{E}fLf \ge 0$ for all $f \in H$.

The *Markov semigroup* of operators $(T_t)_{t\ge0} : \mathcal{H} \to \mathcal{H}$ *generated by* L is given by

$$T_t f = e^{-tL} f,$$

with $T_0 f = f$ and $\frac{d}{dt}T_t f = -LT_t f = -T_t L f$. The *Dirichlet form* $\mathcal{E} : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ *associated with* L is given by

$$\mathcal{E}(f,g) = \mathbb{E}(fLg) = \mathbb{E}(gLf) = \mathcal{E}(g,f) = -\frac{d}{dt}\mathbb{E}fT_t g\Big|_{t=0}.$$

Recall that in the setup of the definition above: $T_t 1 = 1$ for $t \geq 0$ and $\mathcal{E}(f,1) = 0$ for $f \in \mathcal{H}$. The operators T_t are symmetric linear contractions in L^p-norm for every $p \in [1,\infty)$ and $t \geq 0$. They are mean-preserving, i.e.

$$\mathbb{E}T_t f = \mathbb{E}1T_t f = \mathbb{E}fT_t 1 = \mathbb{E}f,$$

and positivity preserving, i.e. $T_t f \geq 0$ for $f \in H_{[0,\infty)}$, thus also order preserving ($f \geq g$ implies $T_t f \geq T_t g$). In fact, they preserve also strict positivity: $f > 0$ implies $T_t f > 0$ for $t \geq 0$. The positivity preserving property implies that $\mathcal{E}(|f|,|f|) \leq \mathcal{E}(f,f)$ for any $f \in \mathcal{H}$.

Definition 1.2. For $f \in \mathcal{H}$ and $p > 0$ we denote by $\|f\|_p$ the *p-th norm* of f: $(\mathbb{E}|f|^p)^{1/p}$. We extend the definition to $p \in \mathbb{R}$ and $f \in \mathcal{H}_{(0,\infty)}$ by setting $\|f\|_p = (\mathbb{E}f^p)^{1/p}$ for $p \neq 0$, and $\|f\|_0 = \exp(\mathbb{E}\log f)$.

Recall that $\|\cdot\|_p$ is a norm for $p \geq 1$ and it is a semi-norm only for $p \in (0,1)$ (unless $|\Omega| = 1$). It is an easy and well known fact that for any $f \in \mathcal{H}_{(0,\infty)}$ the map $p \mapsto \|f\|_p$ is continuous and non-decreasing.

Following Borell [Bor82] we extend the definition of duality to $p \in \mathbb{R}$:

Definition 1.3. For a number $p \in \mathbb{R}\setminus\{0,1\}$ we define its (Hölder) conjugate $p' = p/(p-1)$, so that $\frac{1}{p} + \frac{1}{p'} = 1$. We also set $0' = 0$.

Note that the map $p \mapsto p'$ is a continuous order-reversing involution on $(-\infty,1)$ and $(1,\infty)$ with fixed points 0 and 2. It is worth observing that $(2-p)' = 2 - p'$ for $p \neq 1$, even though we will not make use of this fact.

1.3. **Log Sobolev inequalities.** We now recall the definition of log-Sobolev inequalities.

Definition 1.4. For $p \in \mathbb{R} \setminus \{0,1\}$ we say that p-logSob is satisfied with constant $C > 0$ if

$$Ent(f^p) \leq \frac{Cp^2}{4(p-1)}\mathcal{E}(f^{p-1},f) \tag{1.7}$$

for every $f \in \mathcal{H}_{(0,\infty)}$. We will say that 1-logSob is satisfied with constant $C > 0$ if

$$Ent(f) \leq \frac{C}{4}\mathcal{E}(f,\log f)$$

for $f \in \mathcal{H}_{(0,\infty)}$. Finally, we will say that 0-logSob is satisfied with constant $C > 0$ if

$$\mathrm{Var}(\log f) \leq -\frac{C}{2}\mathcal{E}(f,1/f)$$

for every $f \in \mathcal{H}_{(0,\infty)}$.

Remark 1.5. Obviously, the cases $p = 0$ and $p = 1$ of the p-logSob inequality are limit cases of the p-logSob for $p \in \mathbb{R} \setminus \{0, 1\}$ (with the same C).

Remark 1.6. Logarithmic Sobolev inequalities were introduced by Gross in his seminal paper [Gro75]. Gross defined logarithmic Sobolev inequality for $p > 1$. The definition was later extended by Bakry [BGM94] to any real p (including 1-log Sobolev inequality). However, Gross's definition is more general than (1.7) as it involves an extra term of the form $\gamma \mathbb{E}[f^p]$, for some nonnegative constant γ, in the right hand side of the inequality (Gross's definition reduces to (1.7) when $\gamma = 0$ and $p > 1$). Finally, we remark that 1-logSob inequality is also known in the literature as modified log-Sobolev inequality (see e.g. [Wu00, BT06]).

Remark 1.7. Our definition uses a novel and non-standard normalization factor $\frac{p^2}{4(p-1)}$ in (1.7). The choice of this normalization makes our p-logSob constants invariant under Hölder conjugation (see Lemma 3.2). Moreover, this normalization is crucial to prove the main result of the paper - the monotonicity of the inequality for $p \in [0, 2]$ (see Theorem 1.8).

In our main result we prove a very general result relating p-logSob inequalities for different values of p.

Theorem 1.8. *Let $0 \leq q \leq p \leq 2$. Assume that p-logSob holds with a constant $C > 0$. Then also q-logSob holds true with the same constant C.*

Remark 1.9. It has been proved in [BGM94, Proposition 3.1] that if 2-logSob holds with constant C, then any p-logSob also holds with the same constant and the converse is true in case of diffusions with invariant measure μ.

Using the fact that simple operators satisfy 1-logSob with the constant 4 (proved in [BT06]; the proof is reproduced in our Lemma 3.5 below) we obtain the following corollary:

Corollary 1.10. *Assume that the semigroup $(T_t)_{t \geq 0}$ is generated by $L = Id - \mathbb{E}$ or by a tensor of simple operators. Then it satisfies the r-logSob inequality with constant 4 for all $r \leq 1$.*

1.4. **Reverse Hyper Contractive Estimates.** Using theorem 1.8 we derive the following general reverse hypercontractive bounds:

Theorem 1.11. *If a symmetric Markov semigroup $(T_t)_{t \geq 0}$ satisfies r-logSob with constant C and $r \geq 1$ then for all $q < p < 1$ and every $f \in \mathcal{H}_{(0,\infty)}$ for all $t \geq \frac{C}{4} \log \frac{1-q}{1-p}$ we have $\|T_t f\|_q \geq \|f\|_p$.*

Using corollary 1.10 this implies in turn that:

Corollary 1.12. *If a symmetric Markov semigroup $(T_t)_{t \geq 0}$ has a simple generator $L = Id - \mathbb{E}$ or it is a tensor product of such simple semigroups, then for all $q < p < 1$ and every $f \in \mathcal{H}_{(0,\infty)}$ for $t \geq \log \frac{1-q}{1-p}$ we have $\|T_t f\|_q \geq \|f\|_p$.*

In fact, for simple operators we derive the following stronger result:

Theorem 1.13. *Assume that a symmetric Markov semigroup $(T_t)_{t\geq 0}$ has a simple generator $L = Id - \mathbb{E}$ or it is a tensor product of such simple semigroups. Let $f \in \mathcal{H}$ be strictly positive. Then for all $q < p \leq 0$ and $t \geq \log \frac{2-q}{2-p}$, and also for all $0 \leq q < p < 1$ and $t \geq \log \frac{(1-q)(2-p)}{(1-p)(2-q)}$ we have $\|T_t f\|_q \geq \|f\|_p$.*

Remark 1.14. As far as we know the only reverse hypercontractive inequality derived prior to ours is (1.5) for a simple operator and the space $\{-1, 1\}$ with the uniform measure. This result is tight.

1.5. **Application 1: Mixing of Large Sets in Markov Chains.** Our first application of the new inequality is to mixing of Markov chains from on large sets. The statement and proof of the theorem below are a generalization of the main result of [MOR+06] where it was proven for the random walk on the discrete cube $\{0, 1\}^n$.

Theorem 1.15. *Let $(X_t)_{t\geq 0}$ be a a continuous time Markov chain on a finite state space Ω which is reversible with respect to the invariant probability measure π. Let $(T_t)_{t\geq 0}$ be the semigroup defined by $T_t f(x) = \mathbb{E}^x f(X_t)$ for $f \in \mathcal{H}$. Assume $(T_t)_{t\geq 0}$ satisfies 1-logSob with constant C. Let $A, B \subseteq \Omega$ with $\pi(A) = \exp(-a^2/2)$ and $\pi(B) = \exp(-b^2/2)$. Let X_0 be distributed according to π. Then*

$$\mathbb{P}\{X_0 \in A, X_t \in B\} \geq \exp\left(-\frac{1}{2}\frac{a^2 + 2e^{-2t/C}ab + b^2}{1 - e^{-4t/C}}\right). \tag{1.8}$$

This theorem should be compared to the two main techniques for proving lower bounds on $\mathbb{P}\{X_0 \in A, X_t \in B\}$.

- First, the *Expander Mixing Lemma* (see e.g. [AS08, Chapter 9]) implies that if Poincaré inequality holds with constant D then:

$$\mathbb{P}\{X_0 \in A, X_t \in B\} \geq \pi(A)\pi(B) - \sqrt{\pi(A)\pi(B)}e^{-t/D}. \tag{1.9}$$

  The inequality 1.9 will be better (up to constants) than our inequality (1.8) in the case where the sets A and B are large, say $\pi(A)\pi(B) \geq 2\sqrt{\pi(A)\pi(B)}e^{-t/D}$, since in this case we obtain the lower bound of $\frac{1}{2}\pi(A)\pi(B)$ which is (except for the factor 2) the best that one can hope for. However, in the case where the sets A and B are small, say $\pi(A)\pi(B) \leq \sqrt{\pi(A)\pi(B)}e^{-t/D}$, the expander mixing lemma gives nothing while (1.8) gives a lower bound that is a power of the measures of the original sets.
- The second technique uses total variation mixing times. Indeed, if the worst total variation distance at time t is at most ϵ, then we have:

$$\mathbb{P}\{X_0 \in A, X_t \in B\} \geq \pi(A)(\pi(B) - \epsilon). \tag{1.10}$$

Again - applying this bound requires that one of the sets A or B is large (of measure at least ϵ). Moreover, in many examples the time t when the total variation distance is at most $1/e$ is much larger than the 1-logSob constant C. Therefore if t is of order C, then the mixing time bound (1.10) gives nothing while our result (1.8) gives an efficient lower bound.

In section 9 we demonstrate this point by proving new mixing bounds from large sets for various classical Markov chains including:

- Random walks on general product spaces.
- Glauber dynamics on Ising model on finite boxes.
- The random transposition card shuffle on the symmetric group.
- The top to random card shuffle on the symmetric group.
- A natural random walk on the spanning trees of a graph considered in [JS02].
- A natural random walk for the Bernoulli-Laplace model, i.e., the natural random walk on subsets of size r of a ground set $[n]$.
- A natural Markovian queueing process - the $q/q/\infty$ Markov process. The last example is interesting since it has infinite 2-logSob constant and an infinite mixing time.

1.6. **Application 2: A General Quantitative Arrow Theorem.** Arrow's Impossibility theorem [Arr50, Arr63] is a fundamental result in social choice theory. Arrow considers functions $F : S_k^n \to \{-1,1\}^{\binom{k}{2}}$ aggregating individual rankings (elements of the permutation group S_k) to result in a preference between every pair of the k alternatives. Arrow showed that if the following desired properties hold simultaneously when $k \geq 3$:

- Transitivity - $F(\sigma)$ induces a transitive ranking for all $\sigma \in S_k^n$.
- Unanimity - for every pair of alternatives a and b, if all voters rank a above b then F also ranks a above b.
- IIA - for every pair of alternatives, the resulting outcome regarding the preference between a and b is determined by the individual preferences between a and b.

Then F is a dictator function, i.e, it is determined by a single voter.

It is natural to ask how robust is the result when considering natural distributions over S_k^n. This questions was analyzed by Kalai [Kal02] who studied it for the case of the uniform distribution over S_3^n and showed that for every $\epsilon > 0$, there exist a $\delta > 0$ such that if F satisfies:

- δ-Transitivity: $\mathbb{P}\{F(\sigma) \text{ is transitive}\} \geq 1 - \delta$.
- Fairness: for every a, b, $\mathbb{P}\{F \text{ ranks } a \text{ above } b\} = 1/2$.
- IIA.

Then there exists a dictator function G such that $\mathbb{P}\{F(\sigma) \neq G(\sigma)\} \leq \epsilon$.

Following a challenge by Kalai, Mossel [Mos11] proved a stronger result for any number of alternatives and without the assumption that F is fair.

His result shows that for $k \geq 3$ and every $\epsilon > 0$, there exist a $\delta > 0$ such that if F satisfies

- δ-Transitivity: $\mathbb{P}\{F(\sigma) \text{ is transitive}\} \geq 1 - \delta$.
- IIA.

Then there exists a function G which is transitive and satisfies the IIA property such that $\mathbb{P}\{F(\sigma) \neq G(\sigma)\} \leq \epsilon$. A complete characterization of all functions G that are IIA and transitive is given by Wilson [Wil72] - these functions include dictators, functions taking two values etc.

A key ingredient of the proof in [Mos11] is the use of reverse hypercontractive inequalities. It is further noted in [Mos11] that it should be possible to extend the proof to general product distributions on S_k^n given reverse hypercontractive bounds for general two point spaces. Here we prove the following extension:

Theorem 1.16 (Quantitative Arrow's theorem for general distribution)**.** *Let ϱ be general distribution on S_k with ϱ assigning positive probability to each element of S_k. Let $\mathbb{P}$ denote the distribution $\varrho^{\otimes n}$ on S_k^n. Then for any number of alternatives $k \geq 3$ and $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$, such that for every n, if $F : S_k^n \to \{-1, 1\}^{\binom{k}{2}}$ satisfies*

- *IIA and*
- $\mathbb{P}\{F(\sigma)$ *is transitive*$\} \geq 1 - \epsilon$.

Then there exists a function G which is transitive and satisfies the IIA property and $\mathbb{P}\{F(\sigma) \neq G(\sigma)\} \leq \epsilon$

We note that considering general product distributions gives a more realistic model of actual voting (though the independence assumption in this line of work is still problematic in real voting scenarios).

1.7. **Application 3: Non-interactive correlation distillation from dice source.** The problem of non-interactive correlation distillation deals with players who receive correlated random strings and whose collective goal is to agree with the highest possible probability on a random variable with a given distribution. Suppose there are $k \geq 2$ players and a 'cosmic source'. Assume first that the source generates a string x of n i.i.d. bits. Each player gets to receive an independent noisy copy of x. Each player then produces a single random bit based on her input. The players wish to have unanimous agreement on their outputs but are not allowed to communicate. The problem is to understand to what extent the players can successfully 'distill' the correlations in their strings into a shared random bit.

This problem has been considered in [AMW91, MO05, Yan04, MOR$^+$06] when the input strings consists of i.i.d. fair coin flips. Here we consider a more general the problem where each bit is an outcome of a throw of a fair dice with m-faces, $m \geq 2$. Let us introduce some notations. Let $\Omega = \{1, 2, \ldots, m\}$ denote set of the possible outcomes of a dice. Let $x = (x_1, x_2, \ldots, x_n)$ be a random vector consisting of n i.i.d. random variables, each being uniformly distributed over Ω. Fix $\rho \in [0,1)$. Let y be a ρ-correlated copy of x and let $(y^i)_{1\le i\le k}$ be conditionally independent copies of y given x. Let player i use the function $F_i : \Omega^n \to \Omega$ to produce her output $F_i(y^i)$. The functions F_i are all assumed to be balanced, that is, $\mathbb{P}\{F_i(x) = j\} = m^{-1}$ for all i, j.

Define

$$\mathcal{M}_\rho(k,n) = \sup_{(F_i)_{1\le i\le k}} \mathbb{P}\{\text{all players output the same bit}\},$$

where the supremum is taken over all choices of balanced functions $(F_i)_{1\le i\le k}$. Since a balanced function defined on n variables can be also thought of a balanced function of $(n+1)$ variables, for a fixed k and ρ, $\mathcal{M}_\rho(k,n)$ is a non-decreasing function of n and so, $\lim_{n\to\infty} \mathcal{M}_\rho(k,n)$ exists. One of the main results of [MOR$^+$06] says that when $m = 2$, we have

$$\lim_{n\to\infty} \mathcal{M}_\rho(k,n) = k^{-\frac{1}{\rho^2}+1+o(1)} \quad \text{as } k \to \infty.$$

The upper bound of the above result uses an application of reverse hypercontractivity for simple semigroup on symmetric two-point space. Here we generalize this bound and give an inverse polynomial bounds (in k) on the agreement probability for general m.

Theorem 1.17. *Fix $\rho \in (0,1]$. Then there exist positive constants $\gamma_1 = \gamma_1(\rho), \gamma_2 = \gamma_2(\rho), c_1 = c_1(m,\rho)$ and $c_2 = c_2(m,\rho)$ such that for all $k \ge 2$,*

$$c_2 k^{-\gamma_2} \le \lim_{n\to\infty} \mathcal{M}_\rho(k,n) \le c_1 k^{-\gamma_1}.$$

Acknowledgements. The first named author enjoyed the hospitality of Isaac Newton Institute, Cambridge while completing part of this research. The second named author enjoyed hospitality of University of California, Berkeley and Isaac Newton Institute, Cambridge while doing this research. We thank Dominique Bakry, Franck Barthe, Nick Crawford, Michel Ledoux and Cyril Roberto for helpful comments and discussions.

2. Comparison of Dirichlet form

The following Theorem extends the classical Stroock-Varopoulos inequality (covering the case $p = 2$, $q \in (1,2]$ of the present result). Theorem 2.1 is the main tool in proving Theorem 1.8. Note that some terms in the statement below may take negative values.

Theorem 2.1. *Let $p, q \in (0,2] \setminus \{1\}$ and $p > q$. Then*

$$qq'\mathcal{E}(g^{1/q}, g^{1/q'}) \ge pp'\mathcal{E}(g^{1/p}, g^{1/p'})$$

for every positive $g \in \mathcal{H}$.

Remark 2.2. The above result has a natural extension to the case $p = 1$ or $q = 1$, with $\mathcal{E}(\log g, g)$ replacing the right (resp. left) hand side of the asserted inequality; then it suffices to use functions $\varphi_1(x) = \log x$ and $\varphi_2(x) = x$ (or $\psi_1(x) = \log x$ and $\psi_2(x) = x$, respectively) in the proof.

The proof of the theorem will use the following lemmas.

Lemma 2.3. *Let I be a non-empty convex subset of $\mathbb{R}$. Assume that some functions $\varphi_1, \varphi_2, \psi_1, \psi_2 : I \to \mathbb{R}$ satisfy*

$$(\varphi_1(a) - \varphi_1(b))(\varphi_2(a) - \varphi_2(b)) \le (\psi_1(a) - \psi_1(b))(\psi_2(a) - \psi_2(b))$$

for all $a, b \in I$. Then for every $f : \Omega \to I$ there is

$$\mathcal{E}(\varphi_1(f), \varphi_2(f)) \le \mathcal{E}(\psi_1(f), \psi_2(f)),$$

where by $\varphi_1(f)$ we denote $\varphi_1 \circ f \in \mathcal{H}$, etc.

Proof. By applying order preserving and linear T_t to a pointwise inequality

$$(\varphi_1(f) - \varphi_1(b))(\varphi_2(f) - \varphi_2(b)) \le (\psi_1(f) - \psi_1(b))(\psi_2(f) - \psi_2(b))$$

we obtain the inequality

$$T_t(\varphi_1(f)\varphi_2(f)) - \varphi_1(b)T_t(\varphi_2(f)) - \varphi_2(b)T_t(\varphi_1(f)) + \varphi_1(b)\varphi_2(b) \le$$
$$T_t(\psi_1(f)\psi_2(f)) - \psi_1(b)T_t(\psi_2(f)) - \psi_2(b)T_t(\psi_1(f)) + \psi_1(b)\psi_2(b)$$

which also holds pointwise, so that it yields yet another pointwise inequality:

$$T_t(\varphi_1(f)\varphi_2(f)) - \varphi_1(f)T_t(\varphi_2(f)) - \varphi_2(f)T_t(\varphi_1(f)) + \varphi_1(f)\varphi_2(f) \le$$
$$T_t(\psi_1(f)\psi_2(f)) - \psi_1(f)T_t(\psi_2(f)) - \psi_2(f)T_t(\psi_1(f)) + \psi_1(f)\psi_2(f).$$

By averaging both sides of it and using the fact that T_t is symmetric and mean-preserving we arrive at

$$\mathbb{E}\varphi_1(f)\varphi_2(f) - \mathbb{E}\varphi_1(f)T_t(\varphi_2(f)) \le \mathbb{E}\psi_1(f)\psi_2(f) - \mathbb{E}\psi_1(f)T_t(\psi_2(f)). \tag{2.1}$$

Let $\alpha : [0, \infty) \to \mathbb{R}$ be given by

$$\alpha(t) = \mathbb{E}\varphi_1(f)T_t(\varphi_2(f)) - \mathbb{E}\psi_1(f)T_t(\psi_2(f)).$$

By (2.1) the function α has a local minimum at zero, so that

$$\mathcal{E}(\psi_1(f), \psi_2(f)) - \mathcal{E}(\varphi_1(f), \varphi_2(f)) = \alpha'_+(0) \ge 0.$$

$\square$

Lemma 2.4. *Assume that I is a convex non-empty subset of $\mathbb{R}$ and functions φ_1, φ_2, ψ_1, $\psi_2 : I \to \mathbb{R}$ are differentiable and such that*

$$\varphi_1'(a)\varphi_2'(b) + \varphi_1'(b)\varphi_2'(a) \le \psi_1'(a)\psi_2'(b) + \psi_1'(b)\psi_2'(a)$$

for all $a, b \in I$. Then for every $f : \Omega \to I$ there is

$$\mathcal{E}(\varphi_1(f), \varphi_2(f)) \le \mathcal{E}(\psi_1(f), \psi_2(f)).$$

Proof. By Lemma 2.3 it suffices to check whether $\Phi : I \times I \to \mathbb{R}$ given by

$$\Phi(a, b) = (\psi_1(a) - \psi_1(b))(\psi_2(a) - \psi_2(b)) - (\varphi_1(a) - \varphi_1(b))(\varphi_2(a) - \varphi_2(b))$$

is nonnegative. Clearly, $\Phi(x, x) = 0$ and $\frac{\partial \Phi}{\partial a}(x, x) = 0$ for all $x \in I$. Now it is enough to notice that the assumptions of Lemma 2.4 yield $\frac{\partial}{\partial b}\frac{\partial}{\partial a}\Phi \ge 0$ which implies that $\Phi(\cdot, x)$ is non-decreasing on $[x, \infty) \cap I$ and non-increasing on $(-\infty, x] \cap I$. $\square$

We are now ready to prove Theorem 2.1.

Proof. It suffices to use Lemma 2.4 with $I = (0, \infty)$, $\varphi_1(x) = px^{1/p}$, $\varphi_2(x) = p'x^{1/p'}$, $\psi_1(x) = qx^{1/q}$, and $\psi_2(x) = q'x^{1/q'}$. Indeed, to verify the assumptions of Lemma 2.4 one needs to check whether

$$(a/b)^{1/p} + (a/b)^{-1/p} \le (a/b)^{1/q} + (a/b)^{-1/q}$$

for all $a, b > 0$. This is, however, obvious since the function $w \mapsto s^w + s^{-w}$ is even and convex for every $s > 0$ and thus it is non-decreasing on $(0, \infty)$. Choosing $s = a/b$ and recalling that $0 < 1/p \le 1/q$ ends the proof. $\square$

Remark 2.5. Theorem 2.1 has a natural extension to the case $p = 1$ or $q = 1$, with $\mathcal{E}(\log g, g)$ replacing the right (resp. left) hand side of the asserted inequality; then it suffices to use functions $\varphi_1(x) = \log x$ and $\varphi_2(x) = x$ (or $\psi_1(x) = \log x$ and $\psi_2(x) = x$, respectively) in the proof.

We further obtain the following.

Corollary 2.6. *For any positive $f \in \mathcal{H}$ there is*

$$\mathcal{E}(\log f, \log f) \le -\mathcal{E}(f, 1/f).$$

Proof. It follows immediately from Lemma 2.4 applied to $I = (0, \infty)$ with $\varphi_1(x) = \log x$, $\varphi_2(x) = \log x$, $\psi_1(x) = x$, and $\psi_2(x) = -1/x$. $\square$

3. Logarithmic Sobolev inequalities

In this section we prove various properties of Log-Sobolev inequalities and in particular Theorem 1.8. We begin with a simple claim relating 0-logSob to a Poincaré inequality. We suspect that both Lemma 3.1 and Lemma 3.2 below are previously known in the literature but we did not find any explicit reference.

Lemma 3.1. *0-logSob holds with constant C if and only if the standard Poincaré inequality holds with constant $C/2$, i.e.*

$$\mathrm{Var}(g) \le \frac{C}{2}\mathcal{E}(g, g)$$

for every $g \in \mathcal{H}$.

Proof. For $g \in \mathcal{H}$ and $\delta > 0$ set $f = e^{\delta g}$. Assuming that f satisfies 0-logSob with constant C we obtain

$$\mathrm{Var}(\delta g) \le -\frac{C}{2}\mathcal{E}(e^{\delta g}, e^{-\delta g}).$$

By the homogeneity of variance and bilinearity of $\mathcal{E}$ we get

$$\mathrm{Var}(g) \le \frac{C}{2}\mathcal{E}\Big(\frac{e^{\delta g} - 1}{\delta}, \frac{1 - e^{-\delta g}}{\delta}\Big) \xrightarrow{\delta \to 0^+} \frac{C}{2}\mathcal{E}(g, g).$$

On the other hand, let $f \in \mathcal{H}$ be positive and assume that the Poincaré inequality holds with constant $C/2$. By using it for $g = \log f$ we arrive at

$$\mathrm{Var}(\log f) \le \frac{C}{2}\mathcal{E}(\log f, \log f) \le -\frac{C}{2}\mathcal{E}(f, 1/f),$$

where we have used Corollary 2.6. □

The following easy observation allows us to restrict study of the p-logSob inequalities to the case $p \in [0, 2]$ (also, it reveals that 1-logSob is, in a sense, a replacement for $\pm\infty$-logSob).

Lemma 3.2. *If $p \in \mathbb{R} \setminus \{1\}$ and p-logSob is satisfied with a constant C then also p'-logSob holds, with the same constant.*

Proof. For $p = 0$ there is nothing to prove, whereas for $p \neq 0$ it suffices to notice that by setting $g = f^p$ we obtain an equivalent 'self-dual' version of p-logSob:

$$Ent(g) \leq \frac{Cpp'}{4} \mathcal{E}(g^{1/p}, g^{1/p'}) \tag{3.1}$$

for all positive $g \in \mathcal{H}$. □

We now prove Theorem 1.8 using the extension of classical Stroock-Varopoulos inequality proven in Theorem 2.1.

Proof. It is a direct consequence of the 'self-dual' reformulation (3.1) of p-logSob, Theorem 2.1, and Remark 2.5. The fact that p-logSob with constant C implies 0-logSob (with the same constant) for every $p \neq 0$ may be proved in two natural ways. One can deduce the Poincaré inequality with constant $C/2$ from p-logSob by setting $f = e^{\delta g}$ and letting δ tend to zero (as in the first part of proof of Proposition 3.1, and use Proposition 3.1 to finish the argument. Alternatively, one can first deduce from p-logSob the q-logSob inequalities (with the same constant) for q arbitrarily close to zero, and then simply apply limit transition $q \to 0$. □

In view of Proposition 3.1 and Theorem 1.8, the p-logSob inequalities, $p \in [0, 2]$, can be treated as a family interpolating in a continuous and monotone way between the classical Poincaré and logaritmic Sobolev inequalities. Another approach to the interpolation problem may be found in [LO00]. Relation between the two approaches seems unclear to the present authors and perhaps it deserves some further investigation.

However, it is well known that all the p-logSob inequalties for $p \in (1, 2]$ are in a sense equivalent, at least if we do not care too much about constants (we do not know whether all p-logSob inequalities for $p \in (0, 1)$ are equivalent in a similar sense).

Proposition 3.3. *Let $1 < q \leq p \leq 2$. Assume that q-logSob holds true with a constant $C > 0$. Then also p-logSob holds true, with constant $\frac{(p-1)q^2}{(q-1)p^2} C$.*

Remark 3.4. It follows from [BGM94, Proposition 3.1] that in case of diffusions with invariant measure μ any q-logSob implies any p-logSob with same constant. However, we did not find in the literature any reference to the results similar to Proposition 3.3 regarding reversible Markov chains.

Proof. Indeed, since $\frac{(p-1)q^2}{(q-1)p^2} = \frac{qq'}{pp'}$ it suffices to prove that $\mathcal{E}(g^{1/q}, g^{1/q'}) \leq \mathcal{E}(g^{1/p}, g^{1/p'})$ for every positive $g \in \mathcal{H}$, which follows easily from Lemma 2.3 applied to $I = (0, \infty)$, $\varphi_1(x) = x^{1/p}$, $\varphi_2(x) = x^{1/p'}$, $\psi_1(x) = x^{1/q}$, and $\psi_2(x) = x^{1/q'}$. The inequality

$$(\varphi_1(a) - \varphi_1(b))(\varphi_2(a) - \varphi_2(b)) \leq (\psi_1(a) - \psi_1(b))(\psi_2(a) - \psi_2(b))$$

is equivalent to

$$(a/b)^{\frac{1}{p}-\frac{1}{2}} + (a/b)^{\frac{1}{2}-\frac{1}{p}} \leq (a/b)^{\frac{1}{q}-\frac{1}{2}} + (a/b)^{\frac{1}{2}-\frac{1}{q}}.$$

Since for every $s > 0$ the function $w \mapsto s^w + s^{-w}$ is non-decreasing on $[0, \infty)$, we finish the proof by setting $s = a/b$ and noting that $\frac{1}{q} - \frac{1}{2} \geq \frac{1}{p} - \frac{1}{2} \geq 0$. □

Usually it is not easy to prove the classical logarithmic Sobolev inequality (2-logSob in our notation). On the other hand, the following lemma provides a modified logarithmic Sobolev inequality (1-logSob in our notation) for a large class of simple semigroups.

Lemma 3.5. *Assume that the semigroup $(T_t)_{t\geq 0}$ is generated by $L = Id - \mathbb{E}$. Then it satisfies 1-logSob with constant 4, i.e.*

$$Ent(f) \leq \mathcal{E}(f, \log f)$$

for all positive $f \in \mathcal{H}$.

Proof. Indeed, it suffices to note that the logarithm function is concave on $(0, \infty)$, so that $\mathbb{E} \log f \leq \log \mathbb{E} f$. Thus

$$Ent(f) = \mathbb{E} f \log f - \mathbb{E} f \cdot \log \mathbb{E} f \leq \mathbb{E} f \log f - \mathbb{E} f \cdot \mathbb{E} \log f =$$

$$\mathbb{E} f(\log f - \mathbb{E} \log f) = \mathbb{E} f L \log f = \mathcal{E}(f, \log f).$$

□

Remark 3.6. Lemma 3.5 was proved in [BT06]. We remark that the constant 4 is not always the optimal 1-logSob constants for the semigroups generated by $L = Id - \mathbb{E}$. For example, for the two point space $(\{0,1\}, \alpha\delta_0 + (1-\alpha)\delta_1)$ the best 1-logSob constant C is known [BT06] to satisfy $C \leq \frac{4}{1+2\sqrt{\alpha(1-\alpha)}}$ which is strictly less than 4 unless $\alpha = \frac{1}{2}$. For $\alpha = \frac{1}{2}$, the best 1-logSob constant is 2 [BT06]. When $\alpha \neq \frac{1}{2}$, the best 1-logSob constant is not known (though the best 2-logSob constant is already known [DSC96]). It's not difficult to show that the best 1-logSob constant is strictly greater than 2 when $\alpha \neq \frac{1}{2}$.

3.1. **Tensorization.** The p-logSob inequalities obviously share the tensorization property of the classical logarithmic Sobolev and Poincaré inequalities. This is a standard observation but we include it here for reader's convenience. For $i = 1, 2, \ldots, n$ assume that (Ω_i, μ_i) is a finite (this assumption may be relaxed) probability space with an associated space $\mathcal{H}_i$ of real functions on Ω_i, and a Markov semigroup $(T_t^{(i)})_{t\geq 0} : \mathcal{H}_i \to \mathcal{H}_i$ generated by a self-adjoint positive semi-definite operator L_i, and a Dirichlet form $\mathcal{E}_i$ (all of them enjoying properties described in the Preliminaries section). Now let us consider a new semigroup $(T_t)_{t\geq 0}$ of operators acting on a space $\mathcal{H} = \mathcal{H}_1 \times \mathcal{H}_2 \times \ldots \times \mathcal{H}_n$ of real-valued functions on a product probability space

$$(\Omega, \mu) = (\Omega_1 \times \Omega_2 \times \ldots \times \Omega_n, \mu_1 \otimes \mu_2 \otimes \ldots \otimes \mu_n).$$

We obtain it by defining its generator $L : \mathcal{H} \to \mathcal{H}$ as

$$L = \sum_{i=1}^{n} Id_{\mathcal{H}_1} \otimes \ldots \otimes Id_{\mathcal{H}_{i-1}} \otimes L_i \otimes Id_{\mathcal{H}_{i+1}} \ldots \otimes Id_{\mathcal{H}_n}.$$

Equivalently, we may define it by setting, for $t \geq 0$,

$$T_t = T_t^{(1)} \otimes T_t^{(2)} \otimes \ldots \otimes T_t^{(n)}.$$

Proposition 3.7. *Let $p \in \mathbb{R}$. In the setting as above, assume that there exist positive constants $C_1, C_2, \ldots, C_n$ such that the semigroup $(T_t^{(i)})_{t\geq 0}$ satisfies p-logSob with constant C_i for $i = 1, 2, \ldots, n$. Then the semigroup $(T_t)_{t\geq 0}$ satisfies p-logSob with the constant $C = \max(C_1, C_2, \ldots, C_n)$.*

Proof. We skip the proof, referring the reader to the classical tensorization argument: subadditivity of entropy (or variance, if $p = 0$). □

4. Hypercontractivity

4.1. **Control of moments under semigroup action.** Let $t = t(p)$ be a differentiable nonnegative function defined on a convex subset of $\mathbb{R} \setminus \{0, 1\}$, and let $f \in \mathcal{H}$ be strictly positive. We will study behavior of the moments of functions $f_{t(p)} := T_{t(p)} f$. An elementary though tedious standard calculation shows that

$$\frac{d}{dp} \log \|T_{t(p)} f\|_p = \frac{Ent(f_{t(p)}^p) - p^2 t'(p) \mathcal{E}(f_{t(p)}^{p-1}, f_{t(p)})}{p^2 \mathbb{E} f_{t(p)}^p}. \tag{4.1}$$

Since, as explained in the Preliminaries, $f_{t(p)}$ is also strictly positive, we may apply to it the p-logSob inequality, which will yield monotonicity of the map $p \mapsto \|T_{t(p)} f\|_p$ upon appropriate choice of the function $t(p)$.

4.2. **Hypercontractivity estimate.**

Proposition 4.1. *Let $r \in (1, 2]$ and let $(T_t)_{t\geq 0}$ be a symmetric Markov semigroup.*

Assume that $(T_t)_{t\geq 0}$ satisfies r-logSob with constant C. Let $r' \leq q \leq p$ or $1 < q \leq p \leq r$. Then for every $t \geq \frac{C}{4} \log \frac{p-1}{q-1}$ and every $f \in \mathcal{H}$ there is $\|T_t f\|_p \leq \|f\|_q$. In other words, T_t is a linear contraction from $L^q(\Omega, \mu)$ to $L^p(\Omega, \mu)$.

Conversely, if there exists $C > 0$ such that

$$\|T_{\frac{C}{4} \log \frac{p-1}{q-1}} f\|_p \leq \|f\|_q \tag{4.2}$$

for all p and q such that $1 < q < p \le r$, and for all positive $f \in \mathcal{H}$ then $(T_t)_{t\ge 0}$ sastisfies r-logSob with the constant C.

Remark 4.2. That the concepts of logarithmic Sobolev inequality and hypercontractivity are intimately connected goes back to Gross [Gro75]. In fact, the converse part of the above proposition follows from Theorem 1.2 of [Gro75] though we add a short proof here for the sake of completeness. But the hypothesis of the forward direction (r-logSob implies hypecontractivity) of our proposition is weaker than that of [Gro75] since [Gro75] assumes r-logSob holds for a nonempty open interval - see Theorem 1.1 of [Gro75] for more details. We also comment that for $r = 2$ we recover the part (i) and (ii) of Theorem 3.5 of Diaconis and Saloff-Coste [DSC96] on the classical equivalence of the logarithmic Sobolev inequality and hypercontractivity for the reversible Markov chains (with essentially the same proof).

Proof. In the proof of the first assertion without loss of generality we can assume that $f \ge 0$ - indeed, since T_t is order preserving, the pointwise inequality $-|f| \le f \le |f|$ implies that $|T_t f| \le T_t|f|$ pointwise, and thus $\|T_t f\|_p \le \|T_t|f|\,\|_p$ whereas f and $|f|$ have the same q-th norm. Furthermore, without loss of generality we may assume that f is strictly positive (which follows by considering functions $f + \varepsilon$ instead of f and then letting $\varepsilon \to 0^+$).

Theorem 1.8 and Lemma 3.2 imply that $(T_t)_{t\ge 0}$ satisfies s-logSob with constant C for all $s \in (1, r] \cup [r', \infty)$. Let $t(s) = \frac{C}{4}\log\frac{s-1}{q-1}$, so that $t(q) = 0$. Then $s^2 t'(s) = \frac{Cs^2}{4(s-1)}$ and (4.1) together with the s-logSob imply that the map $s \mapsto \|T_{t(s)} f\|_s$ in non-increasing on $[q, p]$. Comparing its values at the ends of the interval we arrive at $\|T_{t_{p,q}} f\|_p \le \|f\|_q$ for $f > 0$, where $t_{p,q} = \frac{C}{4}\log\frac{p-1}{q-1}$. To finish the proof of the first assertion for $t > t_{p,q}$ it suffices to express T_t as $T_{t-t_{p,q}} \circ T_{t_{p,q}}$, and use the fact that the semigroup is contractive in L^p-norm ($p > 1$).

To prove the second assertion let us fix some $q \in (1, r)$ and some positive $f \in \mathcal{H}$. For $p \in [q, r)$ let $t(p) = \frac{C}{4}\log\frac{p-1}{q-1}$, so that $t(q) = 0$. Since the map $p \mapsto \|T_{t(p)} f\|_p$ is non-increasing on $[q, r)$ by using (4.1) at $p = q$ we infer that q-logSob holds true with the constant C. Passing to the limit $q \to r^-$ ends the proof. □

5. Reverse hypercontractivity - preliminary results

In this section we prove some preliminary results regarding reverse hypercontractivity.

5.1. **Reverse Contraction.** We begin with the following standard fact showing 'reverse contraction':

Lemma 5.1. *Let I be a non-empty convex open subset of $\mathbb{R}$ and let $(T_t)_{t\ge 0}$ be a symmetric Markov semigroup. Then for every $t > 0$ and every concave $\Phi : I \to \mathbb{R}$ there is $\mathbb{E}\Phi(T_t f) \ge \mathbb{E}\Phi(f)$ for all $f \in \mathcal{H}$ with values in I. In particular, for every $q < 1$ and positive $f \in \mathcal{H}$ we have $\|T_t f\|_q \ge \|f\|_q$.*

Proof. Indeed, Φ may be expressed as infimum of a family $\mathcal{C}_\Phi$ of affine functions:

$$\Phi(x) = \inf\{\phi(x); \phi \in \mathcal{C}_\Phi\}$$

for $x \in I$. Thus from the pointwise inequality $\Phi(f) \le \phi(f)$ and positivity preserving by T_t we deduce $T_t\Phi(f) \le T_t\phi(f) = \phi(T_t f)$ for all $\phi \in \mathcal{C}_\Phi$ and hence $T_t\Phi(f) \le \inf\{\phi(T_t f); \phi \in \mathcal{C}_\Phi\} = \Phi(T_t f)$, pointwise, again. So $\mathbb{E}\Phi(f) = \mathbb{E}T_t(\Phi(f)) \le \mathbb{E}\Phi(T_t f)$, and we are done. The fact that also $T_t f$ has values in I is a consequence of the order preservation. $\square$

Note that the lemma used for $I = \mathbb{R}$ and $\Phi(x) = -|x|^p$, $p \ge 1$ implies the contractivity of $(T_t)_{t\ge 0}$ in L^p-norm.

5.2. **Duality and Tensorization.** The standard statement of the duality of L^p-norms is that for $p > 1$ and $f \in \mathcal{H}$ we have

$$\|f\|_p = \sup\{\mathbb{E}fg;\ \|g\|_{p'} \le 1\}.$$

A slightly less known observation can be found in [Bor82]:

Lemma 5.2. *Let $p \in (-\infty, 1)$. Then for any positive $f \in \mathcal{H}$ there is*

$$\|f\|_p = \inf\{\mathbb{E}fg;\ g > 0, \|g\|_{p'} \ge 1\}.$$

We skip its proof since it is an easy exercise.

The standard duality of the L^p-norms implies that $L^{p'}(\Omega,\mu)$ is Banach space dual to $L^p(\Omega,\mu)$ for any $p > 1$, and from the symmetry of the semigroup $(T_t)_{t\ge 0}$ we deduce that

$$\|T_t\|_{L^p(\Omega,\mu)\to L^q(\Omega,\mu)} = \|T_t\|_{L^{q'}(\Omega,\mu)\to L^{p'}(\Omega,\mu)}$$

for any $p, q > 1$ and $t \ge 0$.

The case $p, q \in (-\infty, 1)$ is less standard and a bit more delicate (in particular, note that this is no longer the Banach space setting). We will need the following auxiliary result which was previously used by Borell [Bor82].

Proposition 5.3. *Let $p, q \in (-\infty, 1)$ and $t \ge 0$. Assume that $\|T_t f\|_q \ge \|f\|_p$ for every positive $f \in \mathcal{H}$. Then also $\|T_t f\|_{p'} \ge \|f\|_{q'}$ for every positive $f \in \mathcal{H}$.*

Proof. Indeed,

$$\|T_t f\|_{p'} = \inf\{\mathbb{E}gT_t f;\ g > 0, \|g\|_p \ge 1\} =$$

$$\inf\{\mathbb{E}fT_t g;\ g > 0, \|g\|_p \ge 1\} \ge \inf\{\mathbb{E}fh;\ h > 0, \|h\|_q \ge 1\} = \|f\|_{q'},$$

where we have used Lemma 5.2, the symmetry of T_t, assumptions of the proposition, and again Lemma 5.2. $\square$

Lemma 5.4. *Assume the set-up of Subsection 3.1. Let $-\infty < q < p < 1$. If for each $1 \le i \le n$, $\|T_t^{(i)} f\|_q \ge \|f\|_p$ for all positive functions $f \in \mathcal{H}_i$, then $\|T_t f\|_q \ge \|f\|_p$ for all positive functions $f \in \mathcal{H}$.*

Proof. The proof is an easy modification of the standard argument for showing the usual hypercontractive inequalities tensorize where Minkowski inequality is to be replaced by the reverse Minkowski inequality (Lemma 5.1). We omit details. $\square$

6. Reverse Hypercontractivity - General Results

We establish an analogue of Proposition 4.1 for p and q below 1, extending results of Borell, [Bor82]. Now we restrict our considerations to positive functions.

Proposition 6.1. *Let $r \in (0,1)$ and let $(T_t)_{t\geq 0}$ be a symmetric Markov semigroup.*

Assume that $(T_t)_{t\geq 0}$ satisfies r-logSob with some constant $C > 0$. Let $r' \leq q \leq p \leq r$. Then for every $t \geq \frac{C}{4}\log\frac{1-q}{1-p}$ and every positive $f \in \mathcal{H}$ there is $\|T_t f\|_q \geq \|f\|_p$.

Conversely, if there exists $C > 0$ such that

$$\|T_{\frac{C}{4}\log\frac{1-q}{1-p}} f\|_q \geq \|f\|_p \tag{6.1}$$

for all p and q such that $0 < q < p \leq r$, and for all positive $f \in \mathcal{H}$ then $(T_t)_{t\geq 0}$ satisfies r-logSob with the constant C.

Remark 6.2. Theorem 3.3 of Bakry's lecture note [BGM94] established similar equivalence between reverse hypercontractivity and r-logSob when $r < 1$. Indeed the converse part of Proposition 6.1 follows from that. But the forward direction, which turns out to be more useful in practice, Theorem 3.3 of [BGM94] assumes that r-logSob holds for some nonempty open interval instead of a single point.

Proof. Let us divide the proof of the first assertion into two basic cases: $0 < q \leq p \leq r$ and $r' \leq q \leq p < 0$ (and in fact we will need to prove only first of them since the second follows then by Proposition 5.3). Once they are proved, the assertion for $0 \leq q \leq p \leq r$ and $r' \leq q \leq p \leq 0$ will follow by passing to a limit ($q \to 0^+$ and $p \to 0^-$, respectively), while the case $q < 0 < p$ will follow from

$$\|T_t f\|_q = \|T_{t-\frac{C}{4}\log\frac{1}{1-p}}(T_{\frac{C}{4}\log\frac{1}{1-p}} f)\|_q \geq \|T_{\frac{C}{4}\log\frac{1}{1-p}} f\|_0 \geq \|f\|_p$$

since $t - \frac{C}{4}\log\frac{1}{1-p} \geq \frac{C}{4}\log(1-q)$ for $t \geq \frac{C}{4}\log\frac{1-q}{1-p}$ (we "glue" the two cases together at zero).

Let us assume $0 < q \leq p \leq r$, then. Consider a function $t(q) = \frac{C}{4}\log\frac{1-q}{1-p}$ defined on $(0,p]$. Then $t(p) = 0$ and $q^2 t'(q) = \frac{Cq^2}{4(q-1)}$, so that by (4.1) the map $q \mapsto \|T_{t(q)} f\|_q$ is non-increasing on $(0,p]$ because r-logSob implies q-logSob, with the same constant C, by Proposition 1.8. At the right end of the interval the map takes on the value $\|f\|_p$, so that $\|T_{t_{p,q}} f\|_q \geq \|f\|_p$ for $t_{p,q} = \frac{C}{4}\log\frac{1-q}{1-p}$. For $t > t_{p,q}$ we simply express $T_t f$ as $T_{t-t_{p,q}}(T_{t_{p,q}} f)$ and use Lemma 5.1.

To prove the converse assertion, let us fix some $p \in (0,r)$ and a positive $f \in \mathcal{H}$. For $q \in (0,p]$ let $t(q) = \frac{C}{4}\log\frac{1-q}{1-p}$, so that $t(p) = 0$. Since the map $q \mapsto \|T_{t(q)} f\|_q$ is non-decreasing on $(0,p]$ formula (4.1) used at $q = p$ yields p-logSob with the constant C. Passing to the limit $p \to r^-$ ends the proof. $\square$

We can now prove Theorem 1.11. In fact we will prove the following result which includes an inverse.

Corollary 6.3. *If a symmetric Markov semigroup $(T_t)_{t\geq 0}$ satisfies r-logSob with constant C and $r \geq 1$ then for all $q < p < 1$ and every positive $f \in \mathcal{H}$ for all $t \geq \frac{C}{4}\log\frac{1-q}{1-p}$ we have $\|T_t f\|_q \geq \|f\|_p$.*

Conversely, if for some $C > 0$ a symmetric Markov semigroup $(T_t)_{t\geq 0}$ satisfies $\|T_{\frac{C}{4}\log\frac{1-q}{1-p}} f\|_q \geq \|f\|_p$ for all $0 < q < p < 1$ and all positive $f \in \mathcal{H}$ then it also satisfies 1-logSob with the constant C.

Proof. By Proposition 1.8 for all $r \in (0,1)$ also r-logSob holds, with the same constant C. The assertion follows immediately from Proposition 6.1.

The converse assertion is easy - Proposition 6.1 implies that $(T_t)_{t\geq 0}$ satisfies r-logSob with the same constant C for all $r \in (0,1)$, and it suffices to pass to the limit $(r \to 1^-)$. ☐

As in [Bor82, MOR$^+$06] we can now obtain the two function version corollary.

Corollary 6.4. *If a symmetric Markov semigroup $(T_t)_{t\geq 0}$ satisfies 1-logSob with constant C then for all $0 < p, q < 1$ and every nonnegative $f, g \in \mathcal{H}$ for all $t \geq -\frac{C}{4}\log[(1-p)(1-q)]$ we have $\mathbb{E}[fT_t g] \geq \|f\|_p\|g\|_q$.*

Proof. Fix $0 < p, q < 1$ and $t \geq \frac{C}{4}\log[(1-p)(1-q)]$. Approximating the nonnegative functions f and g by positive functions $f+\epsilon$ and $g+\epsilon$ and then in the end letting $\epsilon \downarrow 0$, we can assume, without loss of generality, that the functions $f, g \in \mathcal{H}$ are positive. Applying the reverse Hölder's inequality (Lemma 5.2), we have $\mathbb{E}[fT_t g] \geq \|f\|_p\|T_t g\|_{p'}$. It remains to show that $\|T_t g\|_{p'} \geq \|g\|_q$, where immediately follows from Corollary 6.3 once we note that $(1-p') = (1-p)^{-1}$. ☐

We conclude this section by proving Corollary 1.12.

Proof. Indeed, Lemma 3.5 and Proposition 3.7 (in the product case) imply that $(T_t)_{t\geq 0}$ satisfies 1-logSob with constant 4, so that it suffices to use Corollary 6.3. ☐

7. Improved Reverse Bounds for Simple Semigroups

Actually, we can significantly weaken the condition $t \geq \log\frac{1-q}{1-p}$ in Corollary 1.12 for simple operators and prove Theorem 1.13.

Proof. It suffices to prove the claim in the case $q < p \leq 0$ - Proposition 5.3 together with an observation that for $0 \leq q < p < 1$ there is $p' < q' \leq 0$ and

$$\frac{2-p'}{2-q'} = \frac{(1-q)(2-p)}{(1-p)(2-q)}$$

will do the rest. Also, we can restrict to the case $L = Id - \mathbb{E}$ - the product case will follow by Lemma 5.4.

Let $q < p < 0$ and $L = Id - \mathbb{E}$, then, so that

$$T_t f = e^{-tL} f = \mathbb{E}f + e^{-t}(f - \mathbb{E}f) = \mathbb{E}f + \tilde{\theta}(f - \mathbb{E}f),$$

where $\tilde{\theta} := e^{-t} \leq \theta := \frac{2-p}{2-q} \in (0,1)$. For $s < 0$, define $\Psi_s : (-1, \infty) \to [0, \infty)$,

$$\Psi_s(x) = \frac{1 + sx - (1+x)^s}{s}.$$

It is easy to check that Ψ_s is a convex function with $\Psi_s(0) = \Psi_s'(0) = 0$ and $\Psi_s''(x) = (1-s)(1+x)^{s-2}$ (actually, the same properties hold true in the case $s > 0$, and in the case $s = 0$ with $\Psi_0(x) = x - \log(1+x)$, but we will not need those). The inequality

$$\Psi_q(\theta x) \leq \Psi_p(x) \tag{7.1}$$

holds true for every $x \in (-1, \infty)$. Indeed, due to the properties of Ψ_s listed above it suffices to prove that $\theta^2 \Psi_q''(\theta x) \leq \Psi_p''(x)$. This is equivalent to the inequality

$$\frac{(1+x)^\theta}{(1+\theta x)} \leq \Big(\frac{(2-q)^2(1-p)}{(2-p)^2(1-q)}\Big)^{\frac{1}{2-q}} \tag{7.2}$$

which immediately follows from the elementary inequality $(1+x)^\theta \leq 1 + \theta x$, and from the fact that the map

$$s \mapsto (2-s)^{-2}(1-s) = \frac{1}{2-s} - \Big(\frac{1}{2-s}\Big)^2$$

is positive and non-decreasing on $(-\infty, 0)$. We are to prove $\|T_t f\|_q \geq \|f\|_p$ for every positive $f \in \mathcal{H}$. By the homogeneity, we may and will assume that $\mathbb{E}f = 1$, so that $g = f - 1$ and $\tilde{g} = \tilde{\theta}\theta^{-1} g$ are zero-mean and take values in $(-1, \infty)$. Then we have

$$\|T_t f\|_q^p = \|1 + \tilde{\theta} g\|_q^p = \|1 + \theta \tilde{g}\|_q^p = (\mathbb{E}(1 + \theta\tilde{g})^q)^{p/q} = (1 - q\mathbb{E}\Psi_q(\theta\tilde{g}))^{|p|/|q|} \leq$$

$$1 - \frac{|p|}{|q|} q\mathbb{E}\Psi_q(\theta\tilde{g}) = 1 + |p|\mathbb{E}\Psi_q(\theta\tilde{g}) \overset{(7.1)}{\leq} 1 + |p|\mathbb{E}\Psi_p(\tilde{g}) = 1 + |p|\mathbb{E}\Psi_p(\tilde{\theta}\theta^{-1} g) =$$

$$1 + |p|\mathbb{E}\Psi_p\Big(\tilde{\theta}\theta^{-1} g + (1 - \tilde{\theta}\theta^{-1}) \cdot 0\Big) \leq 1 + |p|\mathbb{E}\tilde{\theta}\theta^{-1}\Psi_p(g) + (1 - \tilde{\theta}\theta^{-1})\Psi_p(0)$$

$$\leq 1 + |p|\mathbb{E}\Psi_p(g) = 1 - p\mathbb{E}\Psi_p(g) = \mathbb{E}(1+g)^p = \mathbb{E}f^p = \|f\|_p^p$$

which ends the proof - recall that the exponent p is negative. The first inequality above was just an application of the elementary $(1+x)^a \leq 1 + ax$, with $a = |p|/|q| \in (0,1)$, $x > -1$.

The case $p = 0$ follows by an obvious limit transition. □

Remark 7.1. Note that we could obtain a better reverse hypercontractivity constant than those given in Proposition 1.13 by first maximizing the function $\Psi_q(\theta x)/\Psi_p(x)$ over $x \in (-1, \infty)$ and then by trying to solve for θ in terms of p and q such that (7.1) holds. This would lead to an equation of the form

$$\left(\frac{1-\theta}{1-r}\right)^{-(p-q)} = \frac{(1-p)\theta^p}{(1-q)r^{2+p}} \quad \text{where } r = (2-p)/(2-q).$$

But unfortunately, in general, θ can not be recovered explicitly from the above equation. However, in the special case when $-\infty < q < p = 0$, θ can explicitly be solved as

$$\theta = 1 + q(2-q)^{-1+2/q}[4(1-q)]^{-1/q}.$$

Corollary 7.2. *If a symmetric Markov semigroup $(T_t)_{t\geq 0}$ has a simple generator $L = Id - \mathbb{E}$ or it is a tensor product of such simple semigroups, then for all $0 < p, q < 1$ and all nonnegative $f, g \in \mathcal{H}$, we have*

$$\mathbb{E}[fT_t g] \geq \|f\|_p \|g\|_q,$$

for all $t \geq \log \frac{(2-p)(2-q)}{4(1-p)(1-q)}$.

Proof. Fix $0 < p, q < 1$ and $t \geq \log \frac{(2-p)(2-q)}{4(1-p)(1-q)}$. Approximating the nonnegative functions by postive functions if necessary, we can assume, without loss of generality, that the functions $f, g \in \mathcal{H}$ are positive. Applying the reverse Hölder's inequality (Lemma 5.2), we have $\mathbb{E}[fT_t g] \geq \|f\|_p \|T_t g\|_{p'}$. Now suppose $t_1, t_2 > 0$ such that $t_1 = \log \frac{2-p'}{2} = \log \frac{(2-p)}{2(1-p)}$ and $t_2 = \log \frac{(2-q)}{2(1-q)}$. Note that $t \geq t_1 + t_2$. Thus, using the semigroup property, we can write $T_t = T_{t_1} \circ T_{t_2} \circ T_{t-(t_1+t_2)}$. Therefore we conclude that

$$\|T_t g\|_{p'} \geq \|T_{t_2} \circ T_{t-(t_1+t_2)} g\|_0 \geq \|T_{t-(t_1+t_2)} g\|_q \geq \|g\|_q,$$

where we used Proposition 1.13 in the first and the second inequality and Lemma 5.1 in the third inequality. □

We now obtain the following corollary regarding *ρ-correlation*.

Definition 7.3. Consider a product space $(\Omega, \mu) = (\prod_{i=1}^n \Omega_i, \prod_{i=1}^n \mu_i)$ where (Ω_i, μ_i) are finite probability spaces. We say that $(x, y) \in \Omega^2$ are *ρ-correlated* if x is distributed according to μ and for each i independently, the conditional distribution of y given x is: with probability ρ, $y_i = x_i$ and with probability $1-\rho$, y_i is sampled independently from μ_i.

Lemma 7.4. *Let (Ω, μ) be the product probability space in definition 7.3. Let $A, B \subseteq \Omega$ be two sets such that $\mu(A) \geq \epsilon, \mu(B) \geq \epsilon$. Let x be distributed according to the product measure μ and y is a ρ-correlated copy of x for some $0 \leq \rho < 1$. Then*

$$\mathbb{P}\{x \in A, y \in B\} \geq \epsilon^{\frac{2-\sqrt{\rho}}{1-\sqrt{\rho}}}.$$

Proof. Let f and g be the characteristic function of the sets A and B respectively. Note that

$$\mathbb{P}\{x \in A, y \in B\} = \mathbb{E}[f(x)g(y)] = \mathbb{E}\big[f(x)\mathbb{E}[g(y)|x]\big] = \mathbb{E}[fT_t g],$$

where $t = \log(1/\rho)$ and $T_t = \otimes_{i=1}^n T_t^i$ where $T_t^i = e^{-t(Id-\mathbb{E})}$. So, by Corollary 7.2, we have

$$\mathbb{E}[f T_t g] \geq \|f\|_p \|g\|_q, \tag{7.3}$$

for all $0 < p, q < 1$ such that $\rho = \frac{4(1-p)(1-q)}{(2-p)(2-q)}$. We now take $p = q = \frac{2(1-\sqrt{\rho})}{2-\sqrt{\rho}}$ in (7.3) to conclude the proof. □

Remark 7.5.
(1) The lower bound in Lemma 7.4 does not depend on the marginal measures $\mu_i, 1 \leq i \leq n$.
(2) We can also use Corollary 6.4 which deals with general symmetric Markov semigroups to get a lower bound $\epsilon^{\frac{2}{1-\sqrt{\rho}}}$. But this bound is worse than what we have achieved by using Corollary 7.2 that improves on the bounds provided by Corollary 6.4 in the case of simple semigroups.

8. Reverse hypercontractivity for some non-simple operators

For some of the applications below we will be interested in operators that are not necessarily simple but are obtained by composing a simple operator with a non-simple operator. In this section we extend some of the reverse hypercontractive results to this setup.

Proposition 8.1. *Assume that (Ω, μ) is a finite probability space and K is Markov kernel on Ω. Let $\nu = \mu K$ and*

$$\alpha := \min_{x,y:\nu(y)>0} \frac{K(x,y)}{\nu(y)} > 0 \tag{8.1}$$

Let $\alpha^\star = -\log(1-\alpha)$. Let the operator $K^{\otimes n}$ be the n-fold tensor product of the kernel K on the product space $(\Omega^n, \mu^{\otimes n})$. Then for all $f : \Omega^n \to \mathbb{R}_+$, for all $q < p \leq 0$ and $\alpha^\star \geq \log \frac{2-q}{2-p}$, and also for all $0 \leq q < p < 1$ and $\alpha^\star \geq \log \frac{(1-q)(2-p)}{(1-p)(2-q)}$ we have $\|K^{\otimes n} f\|_{L^q(\mu^{\otimes n})} \geq \|f\|_{L^p(\nu^{\otimes n})}$.

Proof. Let $T_t = e^{-t} I + (1 - e^{-t})\mathbb{E}_\mu, t \geq 0$ be the simple Markov semigroup on (Ω, μ). We can extend the definition of T_t for $t < 0$ as well, though it may no longer be a Markov operator. However, it is easy to check that $T_{-t}K = e^t K - (e^t - 1)\mathbb{E}_\nu$ is a Markov operator for small enough $t > 0$. First, for all t, $T_{-t}K1 = 1$ so it remains to check the positivity of $T_{-t}K$. Second, to ensure positivity, we need to show that $e^t K(x, y) - (e^t - 1)\nu(y) \geq 0$ for all $x, y \in \Omega$ which holds if $0 \leq t \leq -\log(1-\alpha) = \alpha^\star$.

The above discussion ensures that the kernel K can be always written as composition of two Markov kernels in the following way:

$$K = T_{\alpha^\star} \circ S, \tag{8.2}$$

where $S := T_{-\alpha^\star} \circ K$. Note that (8.2) implies that

$$\mu S = \mu T_{\alpha^\star} \circ S = \mu K = \nu.$$

Using the decomposition (8.2), we can follow the same argument given in the proof of Proposition 1.13 to obtain

$$\|Kf\|_{L^q(\mu)} \geq \|Sf\|_{L^q(\mu)} = \|f\|_{L^q(\nu)},$$

for p and q as in the hypothesis, where the last inequality follows from the fact that $\mu K = \nu$. The proof can now be completed by standard tensorization argument. $\square$

Corollary 8.2. *Consider the set-up of Proposition 8.1. Then for all* $0 < p, q < 1$ *and all nonnegative* f, g*, we have*

$$\mathbb{E}[fK^{\otimes n}g] \geq \|f\|_{L^q(\mu^{\otimes n})}\|g\|_{L^p(\nu^{\otimes n})},$$

for all $\alpha^\star \geq \log \frac{(2-p)(2-q)}{4(1-p)(1-q)}$.

Proof. Same as Corollary 7.2. $\square$

Lemma 8.3. *Let* $(x_i, y_i)_{1 \leq i \leq n}$ *be i.i.d. on* Ω^2*. Let* μ *and* γ *be the marginal distributions of* x_i *and* y_i *respectively and let* K *denote the conditional probability kernel* $K(a, b) = \mathbb{P}\{y_i = b | x_i = a\}$*. Assume that* $\alpha > 0$ *where* α *is given in* (8.1)*. Define* $x = (x_1, x_2, \ldots, x_n)$ *and* $y = (y_1, y_2, \ldots, y_n)$ *to be the random vectors in* Ω^n*. Let* $A, B \subset \Omega^n$ *be any two sets such that* $\mu^{\otimes n}\{A\} \geq \epsilon$ *and* $\nu^{\otimes n}\{B\} \geq \epsilon$*, then:*

$$\mathbb{P}\{x \in A, y \in B\} \geq \epsilon^{\frac{2-\sqrt{1-\alpha}}{1-\sqrt{1-\alpha}}}. \tag{8.3}$$

Proof. Let f and g be the characteristic function of the sets A and B respectively. Note that

$$\mathbb{P}\{x \in A, y \in B\} = \mathbb{E}[f(x)g(y)] = \mathbb{E}_{\mu^{\otimes n}}\big[f(x)\mathbb{E}[g(y)|x]\big] = \mathbb{E}_{\mu^{\otimes n}}[fK^{\otimes n}g].$$

Now by Corollary 8.2,

$$\mathbb{E}_{\mu^{\otimes n}}[fK^{\otimes n}g] \geq \|f\|_{L^p(\mu^{\otimes n})}\|g\|_{L^q(\nu^{\otimes n})},$$

for all $0 < p, q < 1$ such that $1-\alpha = \frac{4(1-p)(1-q)}{(2-p)(2-q)}$. We take $p = q = \frac{2(1-\sqrt{1-\alpha})}{2-\sqrt{1-\alpha}}$ to conclude the proof. $\square$

Remark 8.4. The following example shows that the condition $\alpha > 0$ cannot be dropped in general. Take $\Omega = \{0, 1\}$ and μ to be the unbiased bernoulli measure on Ω. Let the kernel K be as follows:

$$K = \begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix}$$

so that $\nu = (1/4)\delta_0 + (3/4)\delta_1$. Now take $A = \{x_1 = 0\}$ and $B = \{y_1 = 0\}$. Then $\mu^{\otimes n}\{A\} = 1/2$ and $\nu^{\otimes n}\{B\} = 1/4$ but $\mathbb{P}\{x \in A, y \in B\} = 0$.

9. Mixing of Markov Chains for big sets

In this section we prove Theorem 1.15 which establishes mixing for Markov chains satisfying 1-LogSob. We then give a number of examples where the theorem can be applied to yield new results on mixing of Markov chains starting from big sets. We begin with a proof of the theorem:

Proof. Take f and g to be the characteristic functions of A and B, respectively. Then by Corollary 6.4, for any choice of $0 < p, q < 1$ with $(1-p)(1-q) = e^{-4t/C}$, we get

$$\mathbb{P}\{X_0 \in A, X_t \in B\} = \mathbb{E}[fT_tg] \geq \|f\|_p\|g\|_q = \exp(-a^2/2p)\exp(-b^2/2q). \tag{9.1}$$

Write $p = 1 - e^{-2t/C}r$, $q = 1 - e^{-2t/C}/r$ with $r > 0$. Maximizing the right-hand side of (9.1) as a function of r the best choice is $r = ((b/a) + e^{-2t/C})/(1 + e^{-2t/C}(b/a))$ which yields in turn

$$p = 1 - e^{-2t/C}r = \frac{1-e^{-4t/C}}{1+e^{-2t/C}(b/a)}, \quad q = 1 - \rho/r = \frac{b}{a}\frac{1-e^{-4t/C}}{e^{-2t/C}+(b/a)}.$$

We conclude the proof by substituting this choice of p and q in (9.1). $\square$

The above result says that as long as $\pi(A), \pi(B) \geq c$ and $t \geq \tau C$, we have a uniform lower bound on the hitting probabilities which only depends on c and τ and is independent of the chain. As was discussed at the introduction the bound in the theorem is often stronger than expander mixing lemma bounds which require the sets to be large relative to $e^{-t/C'}$ where C' is the 0-logSob constant. Further the bound is relevant at times t smaller than the mixing time since once mixing is reached within ϵ we have $\mathbb{P}\{X_0 \in A, X_t \in B\} \geq \pi(A)(\pi(B) - \epsilon)$.

We follow below with a list of examples where the new bound is of natural interest. We highlight examples where the mixing time is much larger than the 1-logSob constant since in these examples our bound is better than the mixing bound even if one of the sets A or B has large measure.

9.1. **Short walks in general hypercube.**

The first example is simply a random walk on a product space. Let (Ω, μ) be a finite probability space and $n \geq 1$. Consider the hypercube $(\Omega^n, \mu^{\otimes n})$. The *continuous-time random walk* on this space corresponds to selecting one coordinate uniformly at random with each ring of a Poisson clock and updating that coordinate according to the distribution μ. This is a reversible Markov chain with invariant distribution $\mu^{\otimes n}$. When $\Omega = \{0, 1\}$ and $\mu\{0\} = \mu\{1\} = 1/2$, we have the standard continuous-time random walk on the hypercube. If $L = Id - \mathbb{E}$ is the generator of the simple Markov semigroup on (Ω, μ) and $(T_t)_{t\geq 0}$ be the corresponding semigroup, then the generator of the random walk on the general hypercube is given by

$$L^{\text{prod}} = \frac{1}{n}\sum_{i=1}^{n} Id \otimes Id \otimes \cdots \otimes \underbrace{L}_{i} \otimes \cdots \otimes Id,$$

and the corresponding Markov semigroup can be expressed as

$$T_t^{\mathrm{prod}} = T_{t/n} \otimes T_{t/n} \otimes \cdots \otimes T_{t/n}, \quad \text{for } t \geq 0.$$

From Lemma 3.5, Proposition 3.7 and Theorem 1.15 it follows that:

Proposition 9.1. *Let X_t be the continuous-time random walk on the general hypercube $(\Omega^n, \mu^{\otimes n})$ with X_0 distributed according to the product measure $\mu^{\otimes n}$. Then for any $A, B \subseteq \Omega^n$ with $\mu^{\otimes n}\{A\} = e^{-a^2/2}$ and $\mu^{\otimes n}\{B\} = e^{-b^2/2}$, and for $t \geq \tau n$, we have*

$$\mathbb{P}\{X_0 \in A, X_t \in B\} \geq \exp\left(-\frac{1}{2}\frac{a^2 + 2e^{-\tau/2}ab + b^2}{1 - e^{-\tau}}\right).$$

Remark 9.2. In fact a better bound can be obtained by repeating the proof of Theorem 1.15 with the two function bound in Corollary 6.4 which applies to simple operators and their tensors. We omit the details.

We note that the mixing time of the walk is of order $n \log n$. Therefore using the mixing time it is impossible to obtain effective bounds even when one of the sets A or B has a large measure and t is of order n (as noted before if both sets are of large measure, the expander mixing lemma applies).

9.2. **An example from queueing theory.** In this subsection, we will give an example where we will show reverse hypercontractivity for Markov semigroup arising from a standard q/q/∞ process (defined below) as a Poissonian limit of the reverse hypercontractive estimate for n dimensional discrete cube with product Bernoulli measure ($p = \lambda/n$) without any knowledge about p-logSob constant of the semigroup. The example is of interest for a number of reasons:

- It deals with a Markov chain defined on an infinite state space.
- It is an example where the 2-logSob and the mixing time are both infinite, yet it is possible to obtain reverse hypercontractive and mixing estimates.
- It is a natural example for queueing theory.

.

Let μ_p be the Bernoulli measure $(1-p)\delta_0 + p\delta_1$ on $\{0,1\}$. Let $(X_t^{(n)})_{t\geq 0}$ be the Markov process on state space $\{0,1\}$ corresponding to the simple semigroup generated by $I - \mathbb{E}$ w.r.t. the measure μ_p with $p = \lambda/n, \lambda > 0$ fixed. Let $X^{n,1}, X^{n,2}, \ldots, X^{n,n}$ be i.i.d. copies of $X^{(n)}$. The process $Y_t^{(n)} := X_t^{n,1} + X_t^{n,2} + \ldots + X_t^{n,n}$ is again Markov whose generator $L^{(n)}$ satisfies

$$L^{(n)} f(x_1 + x_2 + \ldots + x_n) = (I - \mathbb{E}_{\mu_{\lambda/n}})^{\otimes n} \hat{f}(x_1, x_2, \ldots, x_n), \quad x_i \in \{0,1\},$$

where $\hat{f} : \{0,1\}^n \to \mathbb{R}$ is given by the relation $\hat{f}(x_1, x_2, \ldots, x_n) = f(x_1 + x_2 + \ldots + x_n)$. Clearly, $\nu_n = \mu_{\lambda/n}^{*n}$, the n-fold convolution of $\mu_{\lambda/n}$, is the reversible measure of $Y^{(n)}$. A simple calculation yields

$$L^{(n)} f(k) = \big(1 - \frac{\lambda}{n}\big)k\big(f(k) - f(k-1)\big) + \frac{\lambda}{n}(n-k)\big(f(k) - f(k+1)\big).$$

So, as $n \to \infty$, the sequence of generators $L^{(n)}$ converges to the generator L which is given by

$$Lf(k) = -(k+\lambda)f(k) - kf(k-1) - \lambda f(k+1),$$

for all $f : \mathbb{N} \to \mathbb{R}$. One can easily recognize the above generator as the generator for the well-known q/q/∞ process which we denote by $(Y_t)_{t\geq 0}$. $(Y_t)_{t\geq 0}$ is a continuous time Markov process taking values in non-negative integers where $Y_t =$ the number of customer in the queue at time t in the following set-up. There are infinite number of servers, the customers arrive according to a Poisson process with rate λ and the service time of each customer follows an independent exponential with mean 1. This process is reversible w.r.t. $\lim_n \nu_n = \text{Poisson}(\lambda)$. Since convergence of the generator implies the convergence of the process, we have, for each $t \geq 0$,

$$Y_t^{(n)} \xrightarrow{d} Y_t, \quad \text{when } Y_0^{(n)} = Y_0. \tag{9.2}$$

Here $\xrightarrow{d}$ means the convergence in distribution.

Let $(T_t^{(n)})_{t\geq 0}$ (resp. $(T_t)_{t\geq 0}$) be the semigroup corresponding to the Markov process $(Y_t^{(n)})_{t\geq 0}$ (resp. $(Y_t)_{t\geq 0}$).

By Theorem 1.13, for any bounded $f : \mathbb{N} \to \mathbb{R}_+$, $0 \leq q < p < 1$ and $n \geq 1$,

$$\|T_t^{(n)} f\|_{L^q(\nu_n)} \geq \|f\|_{L^p(\nu_n)}, \quad \text{for } t \geq \log \frac{(1-q)(2-p)}{(1-p)(2-q)}.$$

Letting $n \to \infty$, by (9.2), we conclude that

$$\|T_t f\|_{L^q(\text{Poi}(\lambda))} \geq \|f\|_{L^p(\text{Poi}(\lambda))}, \quad \text{for } t \geq \log \frac{(1-q)(2-p)}{(1-p)(2-q)}.$$

Similarly by approximating the process $(Y_t)_{t\geq 0}$ by the process $Y^{(n)}$ and applying Theorem 1.15 we obtain that for any two sets A, B of integers of (Poisson λ) measure $e^{-a^2/2}$ and $e^{-b^2/2}$ respectively it holds that

$$\mathbb{P}\{X_0 \in A, X_t \in B\} \geq \exp\left(-\frac{1}{2}\frac{a^2 + 2e^{-t/2}ab + b^2}{1-e^{-t}}\right).$$

Note again that this result holds in an example where the mixing time and 2-logSob constant are infinite (see [BL98] where it is shown that the 1-logSob is finite).

9.3. **Glauber dynamics on Ising model on finite boxes of $\mathbb{Z}^d$.** The Ising model on a finite graph (V, E) has the state space $\Omega = \{-1, +1\}^V$. The probability of a spin configuration $\sigma \in \Omega$ is given by the Gibbs distribution,

$$\mu(\sigma) = \frac{1}{Z(\beta, h)} \exp\left(\beta \sum_{uv\in E} \sigma(u)\sigma(v) + h \sum_{u\in V} \sigma(u)\right),$$

where $Z(\beta, h)$ is the normalizing constant. The parameters $\beta \geq 0$ and h are called the inverse temperature and the external field respectively. These definitions extend to infinite locally finite graphs like $\mathbb{Z}^d$.

The Glauber dynamics for the Ising model is a family of continuous time Markov chains on the state space Ω, reversible with respect to Gibbs distribution, given by the generator

$$(Lf)(\sigma) = \sum_{u \in V} c(u, \sigma)(f(\sigma^u) - f(\sigma)),$$

where σ^u is the configuration σ with the spin at u flipped. We consider the two examples of transition rates $c(u, \sigma)$:

(1) Metropolis: $c(u, \sigma) = \exp\big(2h\sigma(u) + 2\beta\sigma(u)\sum_{v\sim u}\sigma(u)\big) \wedge 1$.
(2) Heat-bath: $c(u, \sigma) = \big[1 + \exp\big(-2h\sigma(u) - 2\beta\sigma(u)\sum_{v\sim u}\sigma(u)\big)\big]^{-1}$.

Let $\Lambda := [-n, n]^d \subseteq \mathbb{Z}^d$ be a finite box in the d-dimensional lattice. Let $\partial_+\Lambda \subseteq \Lambda^c$ be the vertex boundary of Λ in $\mathbb{Z}^d$. Let μ be the Gibbs distribution on $\mathbb{Z}^d$. Given a boundary condition $\tau \in \{-1, +1\}^{\partial_+\Lambda}$, we define a Gibbs distribution on Λ as a conditional measure:

$$\mu_\Lambda^\tau = \mu(\cdot | \sigma_{\partial_+\Lambda} = \tau).$$

Suppose that the inverse-temperature β and external field h are such that the Ising model on $\mathbb{Z}^d$ has strong spatial mixing. Then there exists a constant K, independent of n, such that given any boundary condition, the 2-logSob constant for the Glauber dynamics for the Ising model on the finite box Λ is bounded above by K, independent of n (see [MO94a, MO94b] which succeed [SZ92b, SZ92a, Zeg92] where uniform bound for 2-logSob constant was established under a more stronger Dobrushin-Shlosman mixing conditions). It is also known that in the regime of strong spatial mixing, the mixing time of the Glauber dynamics is $t_{\text{mix}} = \Theta(\log n)$.

The example above can be easily extended to other spin systems and other graphs as long as 1-logSob inequality is established.

9.4. **Random transposition walk on Symmetric group.** The random transposition walk on the group S_n of permutations of n elements is the walk generated by the set of all transpositions $\mathcal{C}_n = \{(i, j) : 1 \leq i < j \leq n\}$. The Markov transition from any $\sigma \in S_n$ is described by picking a transposition τ uniformly at random from $\mathcal{C}_n$ and compose it with σ to get a new permutation $\tau \circ \sigma \in S_n$. It was shown in [GQ03, Goe04] that the 1-logSob constant C of this chain is of order n. More precisely,

$$\frac{n-1}{2} \leq C \leq 2(n-1).$$

On the other hand, it is well known (see [DS81]) that the mixing time $t_{\text{mix}} = \Theta(n \log n)$. It's worth mentioning that the 2-logSob constant of the random transposition walk was determined in [LY98a] to satisfy $C' = \Theta(n \log n)$.

9.5. **Top-to-random transposition walk on symmetric group.** This is a random walk on S_n generated by the set of transpositions $\mathcal{D}_n = \{(1,j) : 2 \le j \le n\}$. Again the 1-logSob constant C of this chain satisfies [Goe04]

$$\frac{n-1}{2} \le C \le 2(n-1),$$

whereas the mixing time is $t_{\text{mix}} = \Theta(n \log n)$ (see [DFP92]).

9.6. **Random walk on spanning trees.** This is a natural random walk on the space of the space of all spanning trees of a graph $G = (V, E)$. Suppose T be our current spanning tree. We choose an edge $e \in E$ and another edge $f \in T$ uniformly at random. If $T' = T \cup \{e\} \setminus \{f\}$ is a spanning tree of G, we update T to T', otherwise we remain at T. It was shown in [JS02] that the 2-logSob constant of this walk satisfies

$$C \le |V||E|,$$

and consequently, $t_{\text{mix}} = O(|V||E| \log |V|)$. In general, the upper bound for the mixing time is tight. For example, consider a line of length n and replace each edge by a double edge. Thus the new graph has $|v| = n+1$ and $|E| = 2n$. The mixing time for the random walk on the spanning trees of this graph is same as the coupon collector problem with a delay of $\Theta(n)$ between successive moves.

9.7. **Bernoulli-Laplace model.** This is natural random walk on the subsets of size r of the ground set $\{1, 2, \dots, n\}$, $1 \le r < n$. So, the state space has size $\binom{n}{r}$. If the current state of Markov chain is a r set A, we pick an element i uniformly at random from A and pick an element j uniformly at random from $\{1, 2, \dots, n\} \setminus A$ and switch the elements to obtain a new r set $A' = A \cup \{j\} \setminus \{i\}$. This is also known as simple exclusion process on the complete graph on n vertices. The 1-logSob constant of this chain satisfies [GQ03, Goe04]

$$\frac{r(n-r)}{2n} \le C \le \frac{2r(n-r)}{n}.$$

The mixing time for Bernoulli-Laplace model is $t_{\text{mix}} = O(\frac{r(n-r)}{n} \log\log \binom{n}{r})$.

10. A Quantitative Arrow's theorem for general ranking distributions

Our goal in this section is to prove Theorem 10.1. We begin by briefly introducing some additional notation. Let $A = \{a, b, \dots, \}$ be a set of $k \ge 3$ alternatives. A *transitive preference* over A is a ranking of the alternatives from top to bottom where ties are not allowed. Such a ranking naturally corresponds to a *permutation* σ of the elements $1, \dots, k$. The group of all rankings will be denoted by S_k. A *constitution* is a function F that associates to every n-tuple $\sigma = (\sigma(1), \dots, \sigma(n))$ of transitive preferences, and every pair of alternatives $a, b \in A$, a (strict) preference between a and b. Some key properties of constitutions include:

- *Transitivity.* The constitution F is *transitive* if $F(\sigma)$ is transitive for all σ. In other words, for all σ and for all three alternatives a, b and c, if $F(\sigma)$ prefers a to b, and prefers b to c, it also prefers a to c. Thus F is transitive if and only if its image is a subset of the permutations on k elements.
- *Independence of Irrelevant Alternatives (IIA).* The constitution F satisfies the IIA property if for every pair of alternatives a and b, the social ranking of a vs. b (higher or lower) depends only on their relative rankings by all voters.
- *Unanimity.* The constitution F satisfies *Unanimity* if the social outcome ranks a above b whenever all individuals rank a above b.
- The constitution F is a *dictator* on voter j, if $F(\sigma) = \sigma(j)$, for all σ, or $F(\sigma) = \sigma(j)^{-1}$, for all σ, where $\sigma(j)^{-1}$ is the inverse of the permutation $\sigma(j)$.

We will assume each voter chooses one ranking from S_k according to some fixed distribution ϱ, independently of others. We will write $\mathbb{P}$ for the underlying probability measure and $\mathbb{E}$ for the corresponding expected value. Given two constitutions F, G on n voters, we denote the statistical distance between F and G by $D(F, G)$, so that:

$$D(F, G) := \mathbb{P}\{F(\sigma) \neq G(\sigma)\}.$$

Our goal is to prove the following theorem which is a more explicit version of Theorem 1.16

Theorem 10.1 (Quantitative Arrow's theorem for general distribution)**.** *Let ϱ be general distribution on S_k which assigns positive probability at least α to each element of S_k. For any number of alternatives $k \geq 3$ and $\epsilon > 0$ sufficiently small, there exists $\delta = \delta(\epsilon) > 0$, such that for every n, if F is a constitution on n voters and k alternatives satisfying:*

- *IIA and*
- $\mathbb{P}\{F(\sigma)$ *is transitive*$\} \geq 1 - \delta$

Then there exist a function G which is IIA and transitive and such that $D(F, G) < \epsilon$. Moreover, one can take

$$\delta = \exp\left(-\frac{C\alpha^{-7}2^{\alpha^{-2}}(\log(1/\epsilon))^2}{\epsilon^{2+\frac{1}{2\alpha^2}}}\right), \tag{10.1}$$

for some suitable absolute constant $C > 0$.

Our proof is an adaptation of the proof in [Mos11]. We begin with some notation and definitions from [Mos11].

Given $\sigma = \sigma(1), \ldots, \sigma(n)$ and for each pair of alternatives $a, b \in A$, we define binary vectors $x^{a>b} = x^{a>b}(\sigma)$ in the following manner:

$$x^{a>b}(j) = 1, \quad \text{if voter } j \text{ ranks } a \text{ above } b;$$

and

$$x^{a>b}(j) = -1, \quad \text{if voter } j \text{ ranks } a \text{ above } b.$$

Thus, if F satisfies the IIA property then there exists functions $f^{a>b}$ for every pair of candidates a and b such that

$$F(\sigma) = ((f^{a>b}(x^{a>b}) : \{a, b\}, a \neq b \in A)$$

where $f^{a>b} : \{-1, 1\}^n \to \{-1, 1\}$ is such that $f = +1$ if F ranks a over b and $f = -1$ otherwise.

We define $\mathrm{PX}(f_1, f_2, f_3)$ (PX stands for paradox) for three function $f_1, f_2, f_3 : \{-1, 1\}^n \to [-1, 1]$ by letting

$$\begin{aligned}\mathrm{PX}(f_1, f_2, f_3) = \frac{1}{4}\Big(1 + &\mathbb{E}[f_1(x^{a>b})f_2(x^{b>c})] + \mathbb{E}[f_2(x^{b>c})f_3(x^{c>a})] \\ &+ \mathbb{E}[f_3(x^{c>a})f_1(x^{a>b})]\Big).\end{aligned}$$

Note that for $k = 3$, the probability of non-transitive outcome is given by

$$\begin{aligned}P(F) &:= \mathbb{P}\{(f^{a>b}, f^{b>c}, f^{c>a}) \in \{(1, 1, 1), (-1, -1, -1)\}\} \\ &= \mathrm{PX}(f^{a>b}, f^{b>c}, f^{c>a}).\end{aligned}$$

In the rest of the subsection, we denote by α, the probability mass of smallest atom of the distributions of the random vectors $(x^{a>b}(1), x^{b>c}(1), x^{c>a}(1))$ on $\{-1, 1\}^3$ for triplets of distinct alternatives $a, b, c \in A$.

We now quickly discuss some definitions of influences that are needed in the proof. For a function $f : \{-1, 1\}^n \to \mathbb{R}$ where $\{-1, 1\}^n$ equipped with n-fold product of some biased measure $\mu_p = (1 - p)\delta_{-1} + p\delta_1$, define the *influence* of variable i on f by

$$\mathrm{Inf}_i(f) := \mu_p^{\otimes n}\{f(x_1, \ldots, \underbrace{-1}_{i}, \ldots, x_n) \neq f(x_1, \ldots, \underbrace{+1}_{i}, \ldots, x_n)\}.$$

Let $\{\psi_0 \equiv 1, \psi_1\}$ form a basis of $L^2(\{-1, 1\}, \mu_p)$. Then f we can express f in its Fourier basis as follows:

$$f(x) = \sum_{S \subseteq [n]} \hat{f}(S) \prod_{i \in S} \psi_1(x_i).$$

We define variance-influence of variable i on f as

$$I_i(f) := \sum_{S : i \in S} \hat{f}(S)^2.$$

When f is ± 1-valued, it can be easily checked that the above two notions of influences are equivalent up to a multiplicative factor (independent of n) as follows:

$$I_i(f) \leq \mathrm{Inf}_i(f) \leq \frac{1}{4p(1 - p)} I_i(f).$$

We also need the notion of low-degree variance-influences. For $d > 0$, this is defined as follows:

$$I_i^{\leq d}(f) := \sum_{S : i \in S, |S| \leq d} \hat{f}(S)^2.$$

Under the above assumption on ϱ, it's not difficult to show that for any three distinct alternatives $a, b, c \in A$ and any voter i, we have

$$|\mathrm{Corr}(x^{a>b}(i), x^{b>c}(i))| \le 1 - 4\alpha.$$

The following lemma is a consequence of the reverse hypercontractivity in the biased space. It is the key ingredient needed to extend the argument of [Mos11] to the nonuniform case.

Lemma 10.2. *Consider a social choice function on* 3 *candidates* a, b *and* c *and* n *voters denoted* $1, \ldots, n$. *Assume that the social choice function satisfies that IIA condition and that voters vote independently according to* ϱ *whose atoms are bounded below by a constant* $\beta > 0$. *Assume further that* $\mathrm{Inf}_1(f^{a>b}) > \epsilon$ *and* $\mathrm{Inf}_2(f^{b>c}) > \epsilon$. *Let*

$$A = \{\sigma : 1 \text{ is pivotal for } f^{a>b}\}, \quad B = \{\sigma : 2 \text{ is pivotal for } f^{b>c}\}.$$

Then

$$\mathbb{P}\{A \cap B\} \ge \epsilon^{\frac{2-\sqrt{1-\alpha}}{1-\sqrt{1-\alpha}}}.$$

Here voter j is called 'pivotal' for $f^{a>b}$ at σ if $f^{a>b}$ is a non-constant function of the j^{th} variable when we freeze the other $(n-1)$ variables at $x^{a<b}(\sigma)$.

Proof. Clearly, $(x^{a>b}(i), x^{b>c}(i))_{1\le i\le n}$ are i.i.d. with a joint distribution on $\{-1,1\}^2$ determined by ϱ. Let μ and ν be the marginal distributions of $x^{a>b}(i)$ and $x^{b>c}(i)$ respectively. Note that the event A is determined by $x^{a>b}$ and the event B is determined by $x^{b>c}$ and the their intersection probability is determined by the joint probability distribution of the random vectors $x^{a<b}$ and $x^{b<c}$. Let A_0 and B_0 be the subsets of $\{-1,1\}^n$ defined by:

$$A_0 = \{x^{a<b} \in \{-1,1\}^n : f^{a<b}(x^{a<b}) \ne f^{a<b}(x^{a<b} \oplus e_1)\},$$

and

$$B_0 = \{x^{b<c} \in \{-1,1\}^n : f^{b<c}(x^{b<c}) \ne f^{b<c}(x^{b<c} \oplus e_2)\}.$$

Observe that $\mu^{\otimes n}\{A_0\} = \mathrm{Inf}_1(f^{a>b}) > \epsilon$ and $\nu^{\otimes n}\{B_0\} = \mathrm{Inf}_2(f^{b>c}) > \epsilon$, and our goal is to obtain a bound on $\mathbb{P}\{A_0 \cap B_0\}$. But now we are exactly in the set-up of Lemma 8.3. If K denotes the conditional distribution of $x^{b<c}(i)$ given $x^{a<b}(i)$, then we have the following lower bound on

$$\min_{u,v\in\{-1,1\}} \frac{K(u,v)}{\nu(v)} = \min_{u,v\in\{-1,1\}} \frac{\mathbb{P}\{x^{a<b}(i) = u, x^{b<c}(i) = v\}}{\mu(u)\nu(v)} \ge \alpha.$$

The proof now follows from Lemma 8.3. □

The reminder of the proof is a straightforward (though somewhat tedious) generalization of the proof given in [Mos11] that does not use reverse hypercontractivity. A sketch of the modifications needed is given in Appendix A.

11. Non-interactive correlation distillation for dice

The proof of theorem 1.17 is a generalization of the proof given in [MOR+06]. The proof of the upper bound uses reverse hypercontractivity while the lower bound is based on the analysis of a simple protocol that is based on the plurality function and the analysis relies the on normal approximation. Here we give the proof of the upper bound while the proof of the lower bound is given in Appendix B.

Proof of upper bound. Note that the probability of all players output $j \in \Omega$ is

$$\mathbb{E}\left[\prod_{i=1}^{k} \mathbb{P}\{F_i(y) = j|x\}\right], \tag{11.1}$$

where y is a ρ-correlated copy of x. Let $f_{i,j}(x) := \mathbf{1}_{\{F_i(x)=j\}}$. Thus if $t = \log(1/\rho)$ and T_t is the simple semigroup then

$$(11.1) = \mathbb{E}\left[\prod_{i=1}^{k} \mathbb{E}[f_{i,j}(y)|x]\right] = \mathbb{E}\left[\prod_{i=1}^{k} T_t f_{i,j}(x)\right] \le \prod_{i=1}^{k} \|T_t f_{i,j}\|_k,$$

where the last step from Hölder's inequality. Since $\mathbb{E} f_{i,j} = m^{-1}$ for all i, j, we conclude, by Lemma 11.1 below, that the probability of total agreement among k players is bounded above by $\sum_{j\in\Omega} \prod_{i=1}^{k} \|T_t f_{i,j}\|_k \le C m k^{-\gamma_1}$ for some $\gamma_1 > 0$ depending on ρ. □

Lemma 11.1. *Let $\Omega = \{1, 2, \dots, m\}$ and μ be the uniform measure on Ω. Fix any $\rho \in (0, 1]$. Then there exist $C = C(\rho) > 0, \beta = \beta(\rho) > 0$ such that for any $f : \Omega^n \to [0, 1]$ and for any $k \ge 1$ such that $\mathbb{E} f \le 1/2$,*

$$\|T_t f\|_k^k \le C k^{-\beta}.$$

Proof of Lemma 11.1. Suppose $\|T_t f\|_k^k \ge 2\delta$. Define $S = \{x \in \Omega^n : [T_t f(x)]^k \ge \delta\}$. Since $T_t f$ is bounded between 0 and 1, it follows that $\mathbb{E}[1_S] \ge \delta$. If we write $\bar{f}$ for $1 - f$, then the set S has the following equivalent description

$$S = \{x \in \{0, 1\}^n : T_t \bar{f}(x) \le 1 - \delta^{1/k}\}.$$

Thus clearly we have

$$\mathbb{E}[1_S T_t \bar{f}] \le (1 - \delta^{1/k})\mathbb{P}\{S\}. \tag{11.2}$$

On the other hand, Corollary 7.2 gives us that

$$\mathbb{E}[1_S T_t \bar{f}] \ge \|1_S\|_p \|\bar{f}\|_q \quad \text{for any } 0 < p, q < 1 \text{ satisfying } \rho \le \frac{4(1-p)(1-q)}{(2-p)(2-q)}.$$

If we take $p = q = \frac{2(1-\sqrt{\rho})}{2-\sqrt{\rho}}$ in the above inequality, we have

$$\mathbb{E}[1_S T_t \bar{f}] \ge \mathbb{P}\{S\}^{\frac{2-\sqrt{\rho}}{2(1-\sqrt{\rho})}} (\mathbb{E}\bar{f})^{\frac{2-\sqrt{\rho}}{2(1-\sqrt{\rho})}}, \tag{11.3}$$

where we have used the fact that $\mathbb{E}[\bar{f}^q] \geq \mathbb{E}\bar{f}$ for any $q \in (0,1)$. Now, comparing (11.2) and (11.3), we have

$$\mathbb{P}\{S\}^{\frac{\sqrt{\rho}}{2(1-\sqrt{\rho})}} (\mathbb{E}\bar{f})^{\frac{2-\sqrt{\rho}}{2(1-\sqrt{\rho})}} \leq 1 - \delta^{1/k}.$$

Since, $\mathbb{P}\{S\} \geq \delta$ and $\mathbb{E}\bar{f} \geq 1/2$, we have

$$\delta^{\frac{\sqrt{\rho}}{2(1-\sqrt{\rho})}} 2^{-\frac{2-\sqrt{\rho}}{2(1-\sqrt{\rho})}} \leq 1 - \delta^{1/k},$$

which implies that $\delta \leq k^{-\beta}$ for any $\beta < \frac{2(1-\sqrt{\rho})}{\sqrt{\rho}}$ and k sufficiently large. $\square$

Remark 11.2. It is an interesting problem to find the exact exponent γ in Theorem 1.17 for which $\lim_{n\to\infty} \mathcal{M}_\rho(k,n) = k^{-\gamma+o(1)}$ as $k \to \infty$. A priori such an exponent might depend on m.

12. Observations and open problems

Our main result showing the monotonicity of r-LogSob inequalities implies that the Poincaré inequality is the weakest among them.

However several open problems regarding monotonicity:

I. Are there intervals I such that r-logSob inequalities are equivalent for all operators and all $r \in I$. In other words, for which intervals I, there exist constants $c(I)$ such that for all $r, r' \in I$, r-logSob with constant C implies r'-logSob with constant $c(I)C$? Note that Proposition 3.3 implies a positive answer to this question with the interval $(1+\epsilon, 2]$ for any $\epsilon > 0$. Note that this interval can not be extended to $[1,2]$. This follows for example from the fact that for the random transposition card shuffling on the symmetric group S_n, 2-logSob constant is $\Theta(n \log n)$ [LY98b] whereas 1-logSob constant is known to be $\Theta(n)$ [GQ03, Goe04].

II. Can one establish similar monotonicity property for hypercontractive inequalities?

12.1. **Reverse hypercontractivity implies spectral gap.**

Here we show that the Poincaré inequality may be deduced from reverse hypercontractivity for *fixed* $q < p < 1$. This provides a partial answer to question II. above.

Lemma 12.1. *Let $q < p < 1$ and $t > 0$. Assume that a symmetric Markov semigroup satisfies the hypercontractivity estimate $\|T_t f\|_q \geq \|f\|_p$ for every positive $f \in \mathcal{H}$. Then it also satisfies the Poincaré inequality*

$$\mathrm{Var}(g) \leq \frac{2t}{\log(1-q) - \log(1-p)} \cdot \mathcal{E}(g,g)$$

for every $g \in \mathcal{H}$.

Proof. Let $\lambda = \inf \sigma(L|_{1^\perp})$, where $1^\perp$ denotes a subspace of $\mathcal{H}$ consisting of all functions orthogonal (in the standard $L^2(\Omega,\mu)$ setting) to the constant function 1, i.e. zero-mean functions. For a zero-mean $g \in \mathcal{H}$ choose $\varepsilon > 0$ small enough to make $f = 1 + \varepsilon g > 0$. The inequality

$$(\|T_t f\|_q - 1)/\varepsilon^2 \geq (\|f\|_p - 1)/\varepsilon^2$$

upon passing to the limit $\varepsilon \to 0^+$ yields $\mathbb{E}[ge^{-2tL}g] = \mathbb{E}[(T_t g)^2] \leq \frac{1-p}{1-q}\mathbb{E}[g^2]$. Since this bound holds for all $g \in 1^\perp$ we infer that $e^{-2t\lambda} \leq (1-p)/(1-q)$ which ends the proof. □

12.2. **Spectral gap does not imply 1-logSob.** Here we show that the 0-logSob inequality does not imply the 1-logSob inequality. In particular it gives a partial answer to question I. above by showing that the r-logSob inequalities in the interval $[0, 1]$ are not all equivalent. Recall that a family graphs $\mathcal{G} = \{G_1, G_2, \ldots\}$ is called a d-regular (spectral) expander if

(1) $G_n = (V_n, E_n)$ is a d-regular graph on n vertices.
(2) The random walk on G_n satisfies a Poincaré inequality with constant C_0 that does not depend on n.

Suppose, if possible, assume that there exists a constant C_1 such that for each n, 1-logSob constant for the random walks on G_n is bounded above by C_1. Let $(X_t)_{t\geq 0}$ be the continuous time random walk on G_n. Since the underlying graph is d-regular, the stationary distribution π is the uniform measure on G_n. So, if we take $A = \{u\}$ and $B = \{v\}$ for $u, v \in V_n$ in (1.8), then we have

$$\mathbb{P}\{X_0 = u, X_1 = v\} \geq n^{-\alpha} \quad \forall u, v \in V_n,$$

where $\alpha > 0$ is a constant that depends on C_1. This implies that

$$\mathbb{P}^u\{X_1 = v\} \geq n^{-\alpha+1} \quad \forall u, v \in V_n. \tag{12.1}$$

By considering volume growth around a fixed point for a bounded degree graph, it is easy to convince ourselves that the diameter of the graph G_n has to be at least $cd \log n$ for some constant $c > 0$. We choose $u, v \in V_n$ so that their graph distance is at least $cd \log n$. So, starting from u, a discrete time random walk needs at least $cd \log n$ many jumps before it can reach the vertex v. Note that the number of jumps made by the continuous time walk X_t during the time interval $[0, 1]$ is distributed according to Poisson random variable with mean 1. Hence, from the tail bound for the Poisson distribution,

$$\mathbb{P}\{X_0 = u, X_1 = v\} \leq \mathbb{P}\{\text{Poisson}(1) \geq cd \log n\} \leq \exp(-c'd \log n \log\log n),$$

for some constant $c' > 0$. Since the right hand side of the above inequality decays faster than any polynomial, it contradicts (12.1). This proves that 1-logSob constant for the random walk on G_n tends to infinity as $n \to \infty$.

Remark 12.2. Explicit lower bounds on 1-logSob constants for connected d-regular graphs on n vertices can be found in [Goe04, BT06]. But our proof is different in the sense that it relies on the new mixing bounds implied by reverse hypercontractivity.

12.3. **Generalizations to infinite spaces.** It is straightforward to generalize the result of sections 1-9 of the paper to infinite probability spaces. The only point which requires some care is to work with the appropriate classes of functions. Since most of the applications in the current paper deal with finite spaces we omit this straightforward extension.

References

[AMW91] N. Alon, U. Maurer, and A. Wigderson, Unpublished results, 1991.

[Arr50] K.J. Arrow, *A difficulty in the concept of social welfare*, The Journal of Political Economy **58** (1950), no. 4, 328–346.

[Arr63] ______, *Social choice and individual values*, John Wiley and Sons, 1963.

[AS08] Noga Alon and Joel H. Spencer, *The probabilistic method*, third ed., Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons Inc., Hoboken, NJ, 2008, With an appendix on the life and work of Paul Erdős. MR 2437651 (2009j:60004)

[Bar80] S. Barbera, *Pivotal voters:: A new proof of arrow's theorem*, Economics Letters **6** (1980), no. 1, 13–16.

[Bec75] William Beckner, *Inequalities in Fourier analysis*, Ann. of Math. (2) **102** (1975), no. 1, 159–182. MR MR0385456 (52 #6317)

[BGM94] D. Bakry, R. D. Gill, and S. A. Molchanov, *Lectures on probability theory*, Lecture Notes in Mathematics, vol. 1581, Springer-Verlag, Berlin, 1994, Lectures from the Twenty-second Saint-Flour Summer School held July 9–25, 1992, edited by P. Bernard.

[BL98] S. G. Bobkov and M. Ledoux, *On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures*, J. Funct. Anal. **156** (1998), no. 2, 347–365. MR 1636948 (99e:60051)

[Bon70] Aline Bonami, *Étude des coefficients de Fourier des fonctions de $L^p(G)$*, Ann. Inst. Fourier (Grenoble) **20** (1970), no. fasc. 2, 335–402 (1971).

[Bor82] Christer Borell, *Positivity improving operators and hypercontractivity*, Math. Z. **180** (1982), no. 2, 225–234. MR 661699 (84b:47029)

[BT06] Sergey G. Bobkov and Prasad Tetali, *Modified logarithmic Sobolev inequalities in discrete settings*, J. Theoret. Probab. **19** (2006), no. 2, 289–336. MR 2283379 (2007m:60215)

[DFP92] Persi Diaconis, James Allen Fill, and Jim Pitman, *Analysis of top to random shuffles*, Combin. Probab. Comput. **1** (1992), no. 2, 135–155. MR 1179244 (93f:60011)

[DS81] Persi Diaconis and Mehrdad Shahshahani, *Generating a random permutation with random transpositions*, Z. Wahrsch. Verw. Gebiete **57** (1981), no. 2, 159–179. MR 626813 (82h:60024)

[DSC96] P. Diaconis and L. Saloff-Coste, *Logarithmic Sobolev inequalities for finite Markov chains*, Ann. Appl. Probab. **6** (1996), no. 3, 695–750. MR 1410112 (97k:60176)

[Goe04] Sharad Goel, *Modified logarithmic Sobolev inequalities for some models of random walk*, Stochastic Process. Appl. **114** (2004), no. 1, 51–79. MR 2094147 (2006h:60122)

[GQ03] Fuqing Gao and Jeremy Quastel, *Exponential decay of entropy in the random transposition and Bernoulli-Laplace models*, Ann. Appl. Probab. **13** (2003), no. 4, 1591–1600. MR 2023890 (2005a:60155)

[Gro75] Leonard Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97** (1975), no. 4, 1061–1083. MR 0420249 (54 #8263)

[Gro78] ———, *Logarithmic Sobolev inequalities—a survey*, Vector space measures and applications (Proc. Conf., Univ. Dublin, Dublin, 1977), I, Lecture Notes in Math., vol. 644, Springer, Berlin, 1978, pp. 196–203. MR 502406 (80f:46035)

[Jan97] S. Janson, *Gaussian Hilbert Spaces*, Cambridge Tracts in Mathematics, vol. 129, Cambridge University Press, Cambridge, 1997. MR 99f:60082

[JS02] Mark Jerrum and Jung-Bae Son, *Spectral gap and log-sobolev constant for balanced matroids*, 721–729.

[Kal02] G. Kalai, *A Fourier-theoretic perspective on the Concordet paradox and Arrow's theorem*, Adv. in Appl. Math. **29** (2002), no. 3, 412–426.

[Kel11] N. Keller, *A tight quantitative version of arrow's impossibility theorem*, Journal of the European Mathematical Society (to appear), 2011.

[KKL88] J. Kahn, G. Kalai, and N. Linial, *The influence of variables on boolean functions*, Proceedings of the 29th Annual Symposium on Foundations of Computer Science, 1988, pp. 68–80.

[LO00] R. Latała and K. Oleszkiewicz, *Between Sobolev and Poincaré*, Geometric aspects of functional analysis, Lecture Notes in Math., vol. 1745, Springer, Berlin, 2000, pp. 147–168. MR 1796718 (2002b:60025)

[LY98a] Tzong-Yow Lee and Horng-Tzer Yau, *Logarithmic Sobolev inequality for some models of random walks*, Ann. Probab. **26** (1998), no. 4, 1855–1873. MR 1675008 (2001b:60090)

[LY98b] ———, *Logarithmic Sobolev inequality for some models of random walks*, Ann. Probab. **26** (1998), no. 4, 1855–1873. MR 1675008 (2001b:60090)

[MO94a] F. Martinelli and E. Olivieri, *Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case*, Comm. Math. Phys. **161** (1994), no. 3, 447–486. MR 1269387 (96c:82040)

[MO94b] ———, *Approach to equilibrium of Glauber dynamics in the one phase region. II. The general case*, Comm. Math. Phys. **161** (1994), no. 3, 487–514. MR 1269388 (96c:82041)

[MO05] Elchanan Mossel and Ryan O'Donnell, *Coin flipping from a cosmic source: on error correction of truly random bits*, Random Structures Algorithms **26** (2005), no. 4, 418–436. MR 2139875 (2006c:60012)

[MOO10] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz, *Noise stability of functions with low influences: invariance and optimality*, Ann. of Math. (2) **171** (2010), no. 1, 295–341. MR 2630040

[MOR+06] Elchanan Mossel, Ryan O'Donnell, Oded Regev, Jeffrey E. Steif, and Benny Sudakov, *Non-interactive correlation distillation, inhomogeneous Markov chains, and the reverse Bonami-Beckner inequality*, Israel J. Math. **154** (2006), 299–336. MR 2254545 (2008b:60018)

[Mos11] E. Mossel, *A quantitative arrow theorem*, To appear in PTRF. Arxiv preprint arXiv:0903.2574, 2011.

[Nel73] Edward Nelson, *The free Markoff field*, J. Functional Analysis **12** (1973), 211–227. MR 0343816 (49 #8556)

[Ole03] Krzysztof Oleszkiewicz, *On a nonsymmetric version of the Khinchine-Kahane inequality*, Stochastic inequalities and applications, Progr. Probab., vol. 56, Birkhäuser, Basel, 2003, pp. 157–168. MR 2073432 (2005f:60049)

[Saz81] Vjačeslav V. Sazonov, *Normal approximation—some recent advances*, Lecture Notes in Mathematics, vol. 879, Springer-Verlag, Berlin, 1981. MR 643968 (83g:60035)

[SC97] Laurent Saloff-Coste, *Lectures on finite Markov chains*, Lectures on probability theory and statistics (Saint-Flour, 1996), Lecture Notes in Math., vol. 1665, Springer, Berlin, 1997, pp. 301–413. MR 1490046 (99b:60119)

[SZ92a] Daniel W. Stroock and Bogusław Zegarliński, *The equivalence of the logarithmic Sobolev inequality and the Dobrushin-Shlosman mixing condition*, Comm. Math. Phys. **144** (1992), no. 2, 303–323. MR 1152374 (93b:82005)

[SZ92b] ———, *The logarithmic Sobolev inequality for discrete spin systems on a lattice*, Comm. Math. Phys. **149** (1992), no. 1, 175–193. MR 1182416 (93j:82013)

[Tal94] M. Talagrand, *On Russo's approximate 0-1 law*, Annals of Probability **22** (1994), 1576–1587.

[Wil72] R. Wilson, *Social choice theory without the pareto principle*, Journal of Economic Theory **5** (1972), no. 3, 478–486.

[Wol07] P. Wolff, *Hypercontractivity of simple random variables*, Studia Mathematica (2007), 219–326.

[Wu00] Liming Wu, *A new modified logarithmic Sobolev inequality for Poisson point processes and several applications*, Probab. Theory Related Fields **118** (2000), no. 3, 427–438. MR 1800540 (2002f:60109)

[Yan04] Ke Yang, *On the (im)possibility of non-interactive correlation distillation*, LATIN 2004: Theoretical informatics, Lecture Notes in Comput. Sci., vol. 2976, Springer, Berlin, 2004, pp. 222–231. MR 2095197 (2005f:94045)

[Zeg92] Bogusław Zegarliński, *Dobrushin uniqueness theorem and logarithmic Sobolev inequalities*, J. Funct. Anal. **105** (1992), no. 1, 77–111. MR 1156671 (93d:58190)

Appendix A. Proof of Theorem 10.1

We continue in the proof of a general quantitative Arrow theorem following [Mos11].

The next step is to replace Theorem 7.1 and Theorem 11.11 in [Mos11] by the following two lemmas respectively.

Lemma A.1. *For every $\epsilon > 0$ there exist $\delta(\epsilon) > 0$ and $\tau(\delta) > 0$ such that the following hold. Let $f_1, f_2, f_3 : \{-1,1\}^n \to \{-1,1\}$ and let F be the social choice function defined by letting $f^{a>b} = f_1$, $f^{b>c} = f_2$ and $f^{c>a} = f_3$. Assume that for all $1 \le i \le 3$ and $j > 1$ it holds that*

$$I_j^{\le(\log(1/\tau)+1}(f_i) < \alpha\tau/2. \tag{A.1}$$

Then either

$$\mathbb{P}(f_1, f_2, f_3) \ge \alpha\delta, \tag{A.2}$$

or there exists a function G which is either a dictator or always ranks one candidate at top/bottom such that $D(F,G) \le 9\epsilon$. Moreover, one can take:

$$\delta = \frac{1}{4}(\epsilon/2)^{2+1/(2\alpha^2)}, \quad \tau = \delta^{C\frac{\log(2/\alpha)}{\alpha}\frac{\log(1/\delta)}{\delta}}.$$

Lemma A.2. *For every $\epsilon > 0$, there exist $\delta(\epsilon) > 0$ and $\tau(\delta) > 0$ such that the following hold. Let $f_1, f_2, f_3 : \{-1,1\}^n \to [-1,1]$. Assume that for all $1 \le i \le 3$ and all $u \in \{-1,1\}$ it holds that*

$$\min(u\mathbb{E}[f_i], -u\mathbb{E}[f_{i+1}]) \le 1 - 3\epsilon \quad (\textit{with convention } f_4 = f_1) \tag{A.3}$$

and for all $1 \le j \le n$ it holds that

$$|\{1 \le i \le 3 : I_j^{\le(\log(1/\tau))^2}(f_i) > \tau\}| \le 1. \tag{A.4}$$

Then we have

$$\mathrm{PX}(f_1, f_2, f_3) \geq \delta.$$

Moreover, one can take:

$$\delta = \frac{1}{8}(\epsilon/2)^{2+1/(2\alpha^2)}, \quad \tau = \delta^{C\frac{\log(2/\alpha)}{\alpha}\frac{\log(1/\delta)}{\delta}}.$$

The proofs of the above two lemmas are almost identical to those given in [Mos11]. The only difference is that instead of Theorem 11.10 of [Mos11] we now use its modified version as follows.

Lemma A.3. *For every $\epsilon > 0$, there exist $\delta(\epsilon) > 0$ and $\tau(\delta) > 0$ such that the following hold. Let $f_1, f_2, f_3 : \{-1,1\}^n \to [-1,1]$. Assume that for all $1 \leq i \leq 3$ and all $u \in \{-1,1\}$ it holds that*

$$\min(u\mathbb{E}[f_i], -u\mathbb{E}[f_{i+1}]) \leq 1 - 3\epsilon \quad (\textit{with convention } f_4 = f_1) \tag{A.5}$$

and for all $1 \leq i \leq 3$ and $1 \leq j \leq n$ it holds that

$$I_j^{\log(1/\tau)}(f_i) < \tau,$$

Then we have

$$\mathrm{PX}(f_1, f_2, f_3) > \delta.$$

Moreover, one can take:

$$\delta = \frac{1}{4}(\epsilon/2)^{2+1/(2\alpha^2)}, \quad \tau = \delta^{C\frac{\log(2/\alpha)}{\alpha}\frac{\log(1/\delta)}{\delta}}.$$

The proof of Lemma A.3 depends on Gaussian Arrow's theorem (see Theorem 11.7 of [Mos11]) and the following generalization of some Gaussian invariance result proved in [Mos11] (see Theorem 11.9). The latter may be of independent interest.

Lemma A.4 (Invariance)**.** *Let $\epsilon > 0, -1 < \rho < 1$. Then for every measurable function $f : \{-1,1\}^n \to [-1,1]$ there exists a measurable function $\widetilde{f} : \mathbb{R}^n \to [-1,1]$ such that the following hold for any $n \geq 1$. Let (X,Y) be distributed on $\{-1,1\}^n \times \{-1,1\}^n$ where $(X_i, Y_i)_{1 \leq i \leq n}$ are i.i.d. with $\mathrm{Corr}(X_i, Y_i) = \rho$. Let $\gamma > 0$ be a lower bound for the smallest atoms of the random variables X_i and Y_i on $\{-1,1\}$. Consider (N,M) jointly Gaussian and distributed in $\mathbb{R}^n \times \mathbb{R}^n$ with $(N_i, M_i)_{1 \leq i \leq n}$ are i.i.d. with*

$$\mathbb{E}[N_i] = \mathbb{E}[M_i] = 0, \quad \mathbb{E}[N_i^2] = \mathbb{E}[M_i^2] = 1, \quad \mathbb{E}[N_i M_i] = \rho.$$

Then

- *For the constant functions 1 and -1 it holds that $\widetilde{1} = 1$ and $\widetilde{-1} = -1$.*
- *If f and g are two functions such that for all $1 \leq i \leq n$, it holds that*

$$\max(I_i^{\leq \log(1/\tau)}(f), I_i^{\leq \log(1/\tau)}(g)) < \tau,$$

then

$$\left|\mathbb{E}[f(X)g(Y)] - \mathbb{E}[\widetilde{f}(N)\widetilde{g}(M)]\right| \leq \epsilon, \tag{A.6}$$

whenever

$$\tau \leq \epsilon^{C \frac{\log(2/\gamma)}{(1-|\rho|)} \cdot \frac{\log(1/\epsilon)}{\epsilon}}, \tag{A.7}$$

for some absolute constant $C > 0$.

Proof of Lemma A.4. The proof is same as Theorem 11.9 of [Mos11]. The only difference is that we now need to apply the version of Theorem 3.20 in [MOO10] under hypothesis H3 instead of hypothesis H4. □

Proof of Theorem 10.1. We will only give a brief sketch the proof of theorem for $k = 3$. The proof for $k > 3$ follows from a general argument given in [Mos11].

Take $\tau = \tau(\epsilon) = \delta_0^{C \log(2/\alpha) \frac{\log(1/\delta_0)}{\alpha\delta_0}}$, $\delta_0 = \frac{1}{8}(\epsilon/2)^{2+1/(2\alpha^2)}$ as in Lemma A.2 and $\eta = \alpha\tau/2$.

Let $f^{a>b}, f^{b>c}, f^{c>a} : \{-1,1\}^n \to \{-1,1\}$ be the three pairwise preference functions. Let $\eta = \delta$ (where the values of C will be determined later). We will consider three cases:

- There exist two voters $i \neq j \in [n]$ and two functions $f \neq g \in \{f^{a>b}, f^{b>c}, f^{c>a}\}$ such that

$$I_i^{\leq(\log(1/\tau))^2}(f) > \eta, \quad I_j^{\leq(\log(1/\tau))^2}(g) > \eta. \tag{A.8}$$

- For every two functions $f \neq g \in \{f^{a>b}, f^{b>c}, f^{c>a}\}$ and every $i \in [n]$, it holds that

$$\min(I_i^{\leq(\log(1/\tau))^2}(f), I_i^{\leq(\log(1/\tau))^2}(g)) < \eta. \tag{A.9}$$

- There exists a voter j' such that for all $j \neq j'$

$$\max(I_j^{\leq(\log(1/\tau))^2}(f^{a>b}), I_j^{\leq(\log(1/\tau))^2}(f^{b>c}), I_j^{\leq(\log(1/\tau))^2}(f^{c>a})) < \eta. \tag{A.10}$$

First note that each F satisfies at least one of the three conditions (A.8), (A.9) or (A.10). Thus it suffices to prove the theorem for each of the three cases.

In (A.8), we have $\mathrm{Inf}_i(f) > \eta$ and $\mathrm{Inf}_j(g) > \eta$. By Lemma 10.2 combined with Barbera's Lemma [Bar80] (see Proposition 3.1 of [Mos11]), we obtian

$$P(F) > \alpha^2 \eta^{\frac{2-\sqrt{1-\alpha}}{1-\sqrt{1-\alpha}}} \geq \alpha^2 \eta^{\frac{4}{\alpha}}.$$

We thus obtain that $P(F) > \delta$ where δ is given in (10.1) by taking large C.

In case (A.9), by Lemma A.2, it follows that Either (if (A.3) does not hold) there exists a function G which always puts a candidate at top/bottom and $D(F,G) < 3\epsilon$, Or, $P(F) > \frac{1}{8}(\epsilon/2)^{2+1/(2\alpha^2)} \gg \delta$.

Similarly in the remaining case (A.10), we have by Lemma A.1 that Either $D(F,G) < 9\epsilon$ Or $P(F) > \frac{1}{4}(\epsilon/2)^{2+1/(2\alpha^2)} \gg \delta$. The proof follows. □

Remark A.5. Keller [Kel11] proved that one may take $\delta = C\epsilon^3$ in the special case when ϱ is *uniform*. It's an interesting open question to see whether such polynomial dependence of δ on ϵ holds for general distribution ϱ.

Appendix B. A lower bound for the NICD problem using a plurality function

Proof of the lower bound for the NICD problem. We will analyze the protocol where all players use some balanced plurality function $\texttt{PLU}_n$ which we are going to described below. Define $n_j = \#\{i : x_i = j\}$ to be the number of times j is present in the string x and set $R = \{j \in \Omega : n_j = \max_{l\in\Omega} n_l\}$. Then we define our pluraity function as

$$\texttt{PLU}_n(x) = x_{i*} \quad \text{where } i_* = \min\{i : x_i \in R\}.$$

Note that if j is the unique value in Ω which occurs most frequently in string x, that is, $R = \{j\}$, then $\texttt{PLU}_n(x) = j$. Also, note that if σ is any permutation of Ω then

$$\texttt{PLU}_n(\sigma(x_1), \sigma(x_2), \cdots, \sigma(x_n)) = \sigma(\texttt{PLU}_n(x)),$$

which implies that $\texttt{PLU}_n$ is balanced.

Define $W_j = W_j^{(n)} := n^{-1/2}\sum_{i=1}^n(\mathbf{1}_{\{x_i=j\}} - m^{-1})$ and $W'_j = W_j'^{(n)} := n^{-1/2}\sum_{i=1}^n(\mathbf{1}_{\{y_i=j\}} - m^{-1})$ where y is a ρ-correlated of x.

The probability of total agreement among k players is bounded below by the probability event that they all output 1 which is at least

$$\mathbb{E}\left[\mathbb{P}\{W'_1 > \max_{j\neq 1} W'_j | W_j, 1 \le j \le m\}^k\right]. \tag{B.1}$$

Next we proceed to bound $\liminf_n \mathbb{P}\{W'_1 > \max_{j\neq 1} W'_j | A\}$ where $A = \{W_1 \ge \frac{2}{\rho}\sqrt{\log k}$ and $W_j \le 0$ for all $2 \le j \le m\}$. Note that

$$\mathbb{P}\{W'_1 > \max_{j\neq 1} W'_j | A\} \ge \mathbb{P}\{W'_1 > \sqrt{\log k} \ge | A\} - \sum_{j=2}^m \mathbb{P}\{W'_j \ge \sqrt{\log k} | A\}$$

$$= \mathbb{P}\{W'_1 > \sqrt{\log k} | W_1 \ge \frac{2}{\rho}\sqrt{\log k}\} - \sum_{j=2}^m \mathbb{P}\{W'_j \ge \sqrt{\log k} | W_j \le 0\}.$$

The last step is justified by the fact that W_j is a sufficient statistics for the conditional distribution of W'_j given x.

Note that for all $1 \le j \le m$

$$\mathbb{E}\mathbf{1}_{\{x_i=j\}} = \mathbb{E}(\mathbf{1}_{\{y_i=j\}} = m^{-1}, \ \mathrm{Var}(\mathbf{1}_{\{x_i=j\}}) = \mathrm{Var}(\mathbf{1}_{\{y_i=j\}}) = m^{-1}(1-m^{-1})$$

and for all $1 \le j \ne j' \le m$

$$\mathrm{Cov}(\mathbf{1}_{\{x_i=j\}}, \mathbf{1}_{\{y_i=j\}}) = \rho m^{-1}(1-m^{-1}), \mathrm{Cov}(\mathbf{1}_{\{x_i=j\}}, \mathbf{1}_{\{x_i=j'\}}) = -m^{-2}.$$

It now follows from multidimensional Central Limit Theorem that

$$(W_j, W'_j) \xrightarrow{d} N_2(0, \Sigma)$$

and

$$(W_j, 1 \le j \le m) \xrightarrow{d} N_m(0, \Gamma)$$

as $n \to \infty$, where $N_2(0, \Sigma)$ (resp. $N_m(0, \Gamma)$) is the two-dimensional (resp. m-dimensional) normal distribution with mean zero and covariance matrix Σ (resp. Γ) given by

$$\Sigma = m^{-1}(1 - m^{-1}) \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \text{ and } \Gamma = m^{-1} I_m - m^{-2}\mathbf{11'}.$$

Moreover, for any any convex regions $R_1 \subseteq \mathbb{R}^2$ and $R_2 \subseteq \mathbb{R}^m$, we have the Berry-Esséen-type error bound [Saz81] as the following:

$$\left|\mathbb{P}\{(W_j, W_j') \in R_1\} - \mathbb{P}\{(Z_1, Z_2) \in R_1\}\right| = O(n^{-1/2}), \tag{B.2}$$

and

$$|\mathbb{P}\{(W_j, 1 \le j \le m) \in R_2\} - \mathbb{P}\{(X_j, 1 \le j \le m) \in R_2\}| = O(n^{-1/2}), \tag{B.3}$$

where $(Z_1, Z_2) \sim N_2(0, \Sigma)$ and $(X_j, 1 \le j \le m) \sim N_m(0, \Gamma)$. From (B.2), it follows that as $n \to \infty$,

$$\mathbb{P}\{W_1' > \sqrt{\log k} | W_1 \ge \frac{2}{\rho}\sqrt{\log k}\} \to \mathbb{P}\{Z_2 > \sqrt{\log k} | Z_1 \ge \frac{2}{\rho}\sqrt{\log k}\}$$

and

$$\mathbb{P}\{W_j' \ge \sqrt{\log k} | W_j \le 0\} \to \mathbb{P}\{Z_2 \ge \sqrt{\log k} | Z_1 \le 0\}.$$

Recall that the conditional distribution Z_2 given Z_1 is $N(\rho Z_1, \sigma_{2.1})$ where $\sigma_{2.1} = (1 - \rho^2)m^{-1}(1 - m^{-1}) \le m^{-1}$. Also recall that if N is a standard normal random variable, then

$$\mathbb{P}\{N > x\} \le x^{-1} e^{-x^2} \qquad \text{for } x > 0,$$

and this bound is sharp in the asymptotic sense

$$\mathbb{P}\{N > x\} = \Theta(x^{-1} e^{-x^2/2}) \text{ as } x \to \infty.$$

Now for any $m \ge 2$,

$$\begin{aligned} \mathbb{P}\left\{Z_2 > \sqrt{\log k} \middle| Z_1 \ge \frac{2}{\rho}\sqrt{\log k}\right\} &\ge \mathbb{P}\left\{\frac{Z_2 - \rho Z_1}{\sigma_{2.1}} > -\frac{\sqrt{\log k}}{\sigma_{2.1}} \middle| Z_1 > \frac{2}{\rho}\sqrt{\log k}\right\} \\ &\ge \mathbb{P}\left\{N > -m\sqrt{\log k}\right\} \ge 1 - \frac{1}{mk}. \end{aligned}$$

Similarly,

$$\mathbb{P}\left\{Z_2 \ge \sqrt{\log k} \middle| Z_1 \le 0\right\} \le \frac{1}{k}.$$

Therefore,

$$\lim_n \mathbb{P}\{W_1' \ge \max_{j \ne 1} W_j' | A\} \ge 1 - \frac{1}{k}.$$

Consequently, for $k \ge 2$,

$$\begin{aligned} \liminf_n \mathcal{M}_\rho(k, n) &\ge \left(1 - \frac{1}{k}\right)^k \liminf_n \mathbb{P}\{A\} \\ &\ge \frac{1}{2}\mathbb{P}\{X_1 \ge \frac{2}{\rho}\sqrt{\log k} \text{ and } X_j \le 0 \text{ for all } 2 \le j \le m\} \text{ [by (B.3)]}. \end{aligned}$$

The proof of the lower bound is now complete by Lemma B.1. □

Lemma B.1. *Let* $(X_j, 1 \leq j \leq m) \sim N_m(0, m^{-1}I_m - m^{-2}\mathbf{1}_m\mathbf{1}_m')$. *Fix* $d > 0$. *Then there exists* $\gamma_2 = \gamma_2(d) > 0$ *such that for* $k \geq 2$,

$$\mathbb{P}\{X_1 \geq d\sqrt{\log k} \text{ and } X_j \leq 0 \text{ for all } 2 \leq j \leq m\} \geq c_2(m)k^{-\gamma_2},$$

where $c_2(m) \to 0$ *as* $m \to 0$.

Proof of Lemma B.1. Note that $X_1 + X_2 + \ldots + X_m = 0$ with probability one. Therefore,

$$\mathbb{P}\{X_1 \geq d\sqrt{\log k}, X_j \leq 0\ \forall j \geq 2\} = \mathbb{P}\{\sum_{j=2}^{m} X_j \leq -d\sqrt{\log k}, X_j \leq 0\ \forall j \geq 2\}$$

$$\geq \mathbb{P}\{X_2 \leq -d\sqrt{\log k}, -1 \leq X_j \leq 0\ \forall j \geq 3\}. \tag{B.4}$$

The conditional distribution of X_2 given $X_j, j \geq 3$ is

$$N\left(-(m-1)m^{-1}\sum_{j=3}^{m} X_j, 2m^{-2}(1-m^{-1})\right).$$

Hence, it can be easily seen that

$$\begin{aligned}(\text{B.4}) &\geq \mathbb{P}\{X_2 + (m-1)m^{-1}\sum_{j=3}^{m} X_j \leq -d\sqrt{\log k} - m\}\mathbb{P}\{-1 \leq X_j \leq 0, j \geq 3\} \\ &\geq \mathbb{P}\{N \leq -d\sqrt{\log k} - m\}\mathbb{P}\{-1 \leq X_j \leq 0, j \geq 3\}\end{aligned}$$

where $N \sim N(0,1)$. The lemma now follows from the normal tail estimate. □

Elchanan Mossel, 367 Evans Hall, University of California, Berkeley, CA, USA
E-mail address: mossel@stat.berkeley.edu

Elchanan Mossel, Faculty of Mathematics and Computer Science, Weizmann Institute, Israel
E-mail address: elchanan.mossel@weizmann.ac.il

Krzysztof Oleszkiewicz, Institute of Mathematics, University of Warsaw, ul. Banacha 2, 02-097, Warsaw, Poland
E-mail address: koles@mimuw.edu.pl

Statistical Laboratory, DPMMS, Wilberforce Road, Cambridge, CB3 0WB, United Kingdom
E-mail address: a.sen@statslab.cam.ac.uk