

# A Query Efficient Non-Adaptive Long Code Test with Perfect Completeness

Suguru Tamaki      Yuichi Yoshida

School of Informatics, Kyoto University, Kyoto 606-8501, Japan.  
{tamak@, yyoshida@lab2.}kuis.kyoto-u.ac.jp

## Abstract

Long Code testing is a fundamental problem in the area of property testing and hardness of approximation. Long Code is a function of the form  $f(x) = x_i$  for some index  $i$ . In the Long Code testing, the problem is, given oracle access to a collection of Boolean functions, to decide whether all the functions are the same Long Code, or cross-influences of any two functions are small. In this paper, we study the following problem: How small the soundness  $s$  of the Long Code test with perfect completeness can be by using non-adaptive  $q$  queries? We give a Long Code test with  $s = (2q + 3)/2^q$ , where  $q$  is of the form  $2^k - 1$  for any integer  $k > 2$ . Our test is a “noisy” version of Samorodnitsky-Trevisan’s Hyper Graph linearity test with suitably chosen noise distribution. To bound the soundness, we use Invariance-Principle style analysis in the spirit of O’Donnell and Wu (STOC 2009).

Previously, Håstad and Khot (Theory of Computing, 2005) have shown  $s = 2^{4\sqrt{q}}/2^q$  for infinitely many  $q$ . Chen (RANDOM 2009) has improved this to  $s = q^3/2^q$  for infinitely many  $q$  with “adaptive” queries. As for the Long Code test with “almost” perfect completeness, Trevisan and Samorodnitsky (SICOMP, 2009) have shown  $s = 2q/2^q$  (or even  $(q + 1)/2^q$  for infinitely many  $q$ ). Austrin and Mossel (Computational Complexity) have improved this to  $s = (q + o(q))/2^q$  (or even  $(q + 4)/2^q$  assuming the famous Hadamard Conjecture) for any  $q$ .

**Key words:** Long Code test, dictatorship test, influence of variables, amortized query complexity

# 1 Introduction

Testing basic properties of Boolean functions is a well studied subject in the area of property testing, complexity theory and learning theory [18, 38]. Examples of such properties include: Linear functions, singleton functions, juntas, low-degree polynomials, and several concise representation [12, 8, 2, 35, 3, 16, 10, 11]. In property test setting, our goal is, given oracle access to Boolean function(s) of the form  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ , to decide whether the functions satisfy a certain property, or they are *far* from the property, where the definition of farness varies according to the property of interest and intended applications of the test. There are several parameters that characterize testers: *Completeness*  $c$  is the probability such that the tester accepts functions that satisfy the property with probability at least  $c$ . *Soundness*  $s$  is the probability such that the tester accepts functions that are far from the property with probability at most  $s$ . *Query complexity*  $q$  is the total number of oracle accesses in one execution of the tester. An *adaptive* tester means that the tester can make queries that depend on the answers of previous queries. Otherwise the tester is said *non-adaptive*.

The property of our interest in this paper is the Long Code. *Long Code* is a function of the form  $f(x) = x_i$  for some index  $i$ . Multi-function Long Code testing problem is, given oracle access to a collection of Boolean functions, to decide whether they are all the same Long Code for some  $i$ , or any two functions are not close to some Long Code at the same time. The notion of Long Code was introduced by Bellare et al. [9] in the context of Probabilistically Checkable Proofs (PCPs) (also studied by Parnas et al. [35] in name of “dictatorship function”). We quickly review the use of (multi-function) Long Code test in PCPs and hardness of approximation. From now on, when we say just PCP, it means a PCP for some NP-complete language. Recall that most efficient PCP constructions today such as the celebrated work of Håstad [22] follow the paradigm of composing “outer verifier” and “inner verifier”.

An outer verifier is a certain kind of two prover one round games (or equivalently label cover problems); e.g., the combination of the PCP theorem [6, 4] and Raz’s parallel repetition theorem [37] (we often call the combination Raz verifier), or Unique Games, or  $d$ -to-1 Games [26]. An inner verifier is essentially a multi-function Long Code tester. By composing outer and inner verifiers, we obtain PCPs, which often has the same parameters of the underlying Long Code tester. Therefore improving the quality of Long Code testers yields better constructions of PCPs. Here we note that what kind of criterion for soundness is used in a Long Code tester/inner verifier may restrict the type of an outer verifier to ensure that they can be composed. We elaborate on this issue later.

The use of PCP constructions in hardness of approximation is quite standard. The acceptance criteria of Long Code testers/PCPs are corresponded to specific optimization problems and the ratio of completeness and soundness  $s/c$  gives the inapproximability ratio of the corresponding problems. For more details on the relation between Long Code tests, PCPs and hardness of approximation, see [27] for general framework and [33] for the emphasis on CSPs.

The connection between Long Code testers and PCPs in mind, we sometimes compare the parameters of them at the same time without distinction. There are several interesting trade-offs between parameters as follows:

**Criteria for soundness** The definition of farness is quite an important issue in the composition paradigm. Several possible criteria for rejection are; distance from any Long Codes or any juntas, quasi-randomness, and low-degree influences. Which criterion is used in Long Code test poses some restriction on an outer verifier, e.g., if we use quasi-randomness, the tester can be composed with Raz verifier, but if we use low-degree influences, it is hard to combine the tester with Raz verifier though it may be composed with Unique Games. It may also affect the lowest possible value of the

soundness with other parameters fixed.

**Perfect completeness vs. almost perfect completeness** It is known that the soundness of PCPs with perfect completeness ( $c = 1$ ) can be worse than that of PCPs with almost perfect completeness ( $c = 1 - \epsilon$ ). Håstad [22] showed that non-adaptive 3-query PCPs with almost perfect completeness and  $s = 1/2$ . Zwick [43] showed that any non-adaptive 3-query PCPs with perfect completeness must have soundness  $s \geq 5/8$ . He conjectured  $s = 5/8$  is optimal. Håstad [22] showed  $s = 3/4$  is possible and Khot and Saket [28] improved this to  $s = 20/27$ . Finally, O’Donnell and Wu [34] showed  $s = 5/8$  under Khot’s  $d$ -to-1 Games conjecture. We remark that perfect completeness is important for several reasons: First, it is natural to require that testers should always accept correct objects. Second, it is required to show the inapproximability results for certain problems such as satisfiable CSPs and approximate coloring problems. Third, it is robust when we compose testers with other protocols.

**Adaptive vs. non-adaptive queries** It is also known that adaptive queries are powerful than non-adaptive queries. Guruswami et al. [19] showed adaptive 3-query PCPs with  $c = 1, s = 1/2$ . As mentioned previously, Zwick showed non-adaptive 3-query PCPs with  $c = 1$  have  $s \geq 5/8$ . Non-adaptivity is crucial in applications for inapproximability; we need queries to be non-adaptive to relate the parameters of PCPs to the inapproximability ratio of optimization problems.

**Query complexity vs. soundness** Trade-offs between query complexity and soundness is also formalized as amortized query complexity, which is defined as  $q/\log_2(1/s)$ . Observe that repeating the tester  $t$  times independently, we can reduce the soundness to  $s^t$  while query complexity grows as  $t \cdot q$ . In this case, the amortized query complexity of the repeated tester remains the same as that of the original one. Hence this can be thought as a good and non-trivial measure for such trade-offs. Constructions of testers that achieve better amortized query complexity is interesting problems.

In this paper, we study the trade-offs between query complexity and soundness in the Long Code testing problem. Our focus is on non-adaptive testers with perfect completeness ( $c = 1$ ).

## 1.1 Previous Work

The study of amortized query complexity in PCPs was initiated by Trevisan [42] and many works followed it [39, 40, 25, 24, 17]. One of the most notable earlier work is due to Samorodnitsky and Trevisan: They constructed  $q$ -query non-adaptive PCPs with almost perfect completeness and  $s = 2^{2\sqrt{q}}/2^q$  for infinitely many  $q$ . Later Engebretsen and Holmerin [17] improved this to  $s = 2^{\sqrt{2q}}/2^q$ . Håstad and Khot constructed PCPs with perfect completeness and  $s = 2^{4\sqrt{q}}/2^q$ . All these constructions use Raz verifier as an outer verifier and thus are unconditional.

Samorodnitsky and Trevisan [41] observed that the trade-offs achieved by these works are essentially tight if the soundness criterion is with respect to the distance from juntas. (Similar limitations were shown by [14, 29].) To break the  $s = 2^{\Theta(\sqrt{q})}/2^q$  barrier, they reworked the problem under the relaxed soundness criterion, i.e., low-degree influences. They constructed so called the Hyper Graph Long Code test, that achieves almost perfect completeness and  $s = 2q/2^q$  (or even  $(q + 1)/2^q$  for infinitely many  $q$ ), and combined it with Unique Games to construct conditional PCPs with the same parameters. They use a novel technique from additive combinatorics, namely, Gowers uniformity norm, in the analysis. Later Austrin and Mossel [5] improved the soundness to  $s = (q + o(q))/2^q$  (or even  $(q + 4)/2^q$  assuming the famous Hadamard Conjecture) for any  $q$ . Their technique is completely different from [41] and based on Invariance Principle, which was developed

by Mossel et al. [30, 31]. Similar constructions were also given in [23, 1, 20] where the emphasis is on the hardness of CSPs over “random predicates” or predicates of non-Boolean domains. However, all these works are affected by the loss of perfect completeness. Recently, Chen [15], building on the work of Samorodnitsky and Trevisan [41] and Håstad and Khot [24], showed  $q$ -query “adaptive” Long Code testers with perfect completeness and  $s = q^3/2^q$  for infinitely many  $q$ .

## 1.2 Our Contribution

In this paper, we prove the following theorem:

**Theorem 1.1.** *For every  $k > 2$ , there exists a non-adaptive Long Code tester that makes  $q = 2^k - 1$  queries, has completeness 1 and soundness  $(2q + 3)/2^q$ .*

## 1.3 Our Method

Our tester and its analysis is based on natural but non-trivial extension of O’Donnell and Wu’s analysis for hardness of satisfiable 3-CSPs over Not-Two (NTW) predicates [34]. Most technical part of O’Donnell-Wu’s work is to bound the term of the form  $\mathbf{E}[f(\mathbf{x})g(\mathbf{y})g(\mathbf{z})]$  where  $\mathbf{x}, \mathbf{y}, \mathbf{z}$  are drawn from “NTW”-distribution. We extend this to the case for the product of large number of functions with more complex distribution.

First we need to define suitable test distributions to make queries. To achieve good query complexity-soundness trade-offs, we adopt the Hyper Graph test introduced by Samorodnitsky and Trevisan [41]. The Hyper Graph test can only test linearity; for testing Long Code, we need to perturb the Hyper Graph test with some “noise” distribution.

One often used noise distribution is uniform distribution. A mixture of the distribution for linearity test and uniform distribution with sufficiently small probability often yields a desired Long Code test. However, this forces the loss of perfect completeness because possible query-answers supports over all the domain, and “yes-instance” can be rejected with small probability. To avoid this, Chen [15] introduced a noise distribution based on Håstad and Khot’s query-efficient PCPs with perfect completeness [24]. His analysis is based on Fourier analysis and Gowers norms as Samorodnitsky and Trevisan did. Although his test achieves good query-soundness trade-offs, one drawback is its adaptivity. In this paper, we use a different noise distribution, modification of the All-Equal noise. Previously, the All-Equal noise was used by [33] combined with 3-bit linearity test to achieve  $c = 1, s = 5/8$ , non-adaptive Long Code test based on NTW predicate. The analysis is based on Fourier analysis and hypercontractivity inequality. Later the same authors [34] managed to construct PCPs with the same parameters  $c = 1$  and  $s = 5/8$  under Khot’s  $d$ -to-1 games conjecture [26]. The analysis is based on Invariance Principle style argument [30, 31]. The crucial point is bounding the correlation of probability spaces.

We follow the approach of [34]. Most part of our analysis becomes complicated to bound terms of the form  $\mathbf{E}[f_1 f_2 \cdots f_q]$  rather than a “cubic”-term, but some analysis becomes easier because we only work with functions over the same domain (or functions with “unique” constraints) rather than functions with  $d$ -to-1 constraints as in the case of [34].

## 1.4 Organization of Paper

In Section 2, we introduce notations and definitions needed throughout this paper. We include some backgrounds on correlated spaces in Appendix A. In Section 3, we define our multi-function Long Code testing problem formally, then we describe our test distribution and Long Code test. In

Section 4, we present completeness and soundness analysis of our tester. Finally, we notice some future work in Section 5.

## 2 Preliminaries

Let  $[k] = \{1, 2, \dots, k\}$ ,  $2_{\geq 1}^{[k]} = \{S \subseteq [k] : |S| \geq 1\}$  and  $2_{\geq 2}^{[k]} = \{S \subseteq [k] : |S| \geq 2\}$ . For a vector  $x \in \{-1, 1\}^n$ ,  $x_i$  denotes the  $i$ th element of  $x$ . Write an unit vector as  $\mathbf{1} = (1, 1, \dots, 1)$  and denote by  $\mathbf{1}_{-i}$  a vector obtained by flipping the  $i$ -th coordinate of  $\mathbf{1}$  to  $-1$ . We denote by  $|\mathbf{S}|$  the cardinality of a set  $\mathbf{S}$ . A *Boolean* function is a function  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ .

### 2.1 Fourier Expansion

**Definition 2.1** (Fourier expansion). *For a function  $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ , the Fourier expansion of  $f$  is*

$$f(x) = \sum_{S \subseteq [n]} \widehat{f}(S) \chi_S(x),$$

where

$$\chi_S(x) = \prod_{i \in S} x_i, \quad \widehat{f}(S) = \mathbf{E}_x[f(x) \chi_S(x)],$$

and expectation is with respect to uniform distribution.

It is easy to see that for any  $S, T \subseteq [n]$ ,  $\mathbf{E}_x[\chi_S(x) \chi_T(x)] = 1$  if  $S = T$  and 0 otherwise. For more on Fourier expansion and Fourier analysis see, e.g., [32].

### 2.2 Influences

In this section we recall basic notions from Fourier analysis, influence and the Bonami-Beckner operator. We first define a notion of the influence of a set of coordinates on a function  $f$ :

**Definition 2.2.** *For a function  $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ , we define the influence of  $i \in [n]$  on  $f$  to be*

$$\text{Inf}_i(f) = \sum_{S \ni i} \widehat{f}(S)^2.$$

We also define the degree- $w$  influence of  $i \in [n]$  on  $f$  to be

$$\text{Inf}_i^{\leq w}(f) = \sum_{S \ni i: |S| \leq w} \widehat{f}(S)^2.$$

We next recall the Bonami-Beckner operator  $T_\rho$  acting on Boolean functions:

**Definition 2.3.** *Let  $0 \leq \rho \leq 1$ . The Bonami-Beckner operator  $T_\rho$  is a linear operator mapping functions  $f : \{-1, 1\}^n \rightarrow \mathbb{R}$  into functions  $T_\rho f : \{-1, 1\}^n \rightarrow \mathbb{R}$  via*

$$(T_\rho f)(x) = \mathbf{E}[f(\mathbf{y})],$$

where in the expectation,  $\mathbf{y}$  is formed from  $x$  by setting  $\mathbf{y}_i = x_i$  with probability  $\rho$  and setting  $\mathbf{y}_i$  to be a uniformly random bit with probability  $1 - \rho$ .

The operator  $T_\rho$  can alternately be defined by the following formula:

**Proposition 2.4.**

$$T_\rho f = \sum_{S \subseteq [n]} \rho^{|S|} \widehat{f}(S) \chi_S$$

□

**Definition 2.5.** Let  $f_1, \dots, f_t$  be a collection of Boolean functions. Then the cross-influence of  $i$  for them is defined as

$$\text{XInf}_i(f_1, \dots, f_t) = \max_{j \neq k} \min(\text{Inf}_i(f_j), \text{Inf}_i(f_k)).$$

Similarly,

$$\text{XInf}_i^{\leq w}(f_1, \dots, f_t) = \max_{j \neq k} \min(\text{Inf}_i^{\leq w}(f_j), \text{Inf}_i^{\leq w}(f_k)).$$

**Proposition 2.6.** Let  $f$  be a Boolean function and  $\gamma > 0$ . Suppose that  $\text{Inf}_i(T_{1-\gamma/2}f) \geq \tau$ , then there exists  $w$  such that  $\text{Inf}_i^{\leq w}(f) \geq \frac{\tau}{2}$

*Proof.* Set  $w$  such that  $(1 - \frac{\gamma}{2})^w \leq \frac{\tau}{2}$ , then

$$\begin{aligned} \tau &\leq \text{Inf}_i(T_{1-\gamma/2}f) = \sum_{S \subseteq [n]} (1 - \frac{\gamma}{2})^{|S|} \widehat{f}(S)^2 \\ &= \sum_{S \subseteq [n], |S| \leq w} (1 - \frac{\gamma}{2})^{|S|} \widehat{f}(S)^2 + \sum_{S \subseteq [n], |S| > w} (1 - \frac{\gamma}{2})^{|S|} \widehat{f}(S)^2 \\ &\leq \sum_{S \subseteq [n], |S| \leq w} \widehat{f}(S)^2 + (1 - \frac{\gamma}{2})^w = \text{Inf}_i^{\leq w}(f) + \frac{\tau}{2}. \end{aligned}$$

□

### 2.3 Correlations

We now recall the definition of correlation for correlated probability spaces, as introduced by Mossel [31].

**Definition 2.7.** Let  $(\Omega, \Psi, \mu)$  be a (finite) correlated probability space, meaning that  $\mu$  is a distribution on the finite product set  $\Omega \times \Psi$  and that the marginals of  $\mu$  on  $\Omega$  and  $\Psi$  have full support. Define the “correlation” between  $\Omega$  and  $\Psi$  to be

$$\rho(\Omega, \Psi; \mu) = \max \left\{ \mathbf{Cov}_{(\omega, \psi) \sim \mu} [f(\omega)g(\psi)] \mid f : \Omega \rightarrow \mathbb{R}, g : \Psi \rightarrow \mathbb{R}, \mathbf{Var}_{(\omega, \psi) \sim \mu} [f(\omega)] = \mathbf{Var}_{(\omega, \psi) \sim \mu} [g(\psi)] = 1 \right\}$$

It is clear that in the definition of  $\rho(\Omega, \Psi; \mu)$ , we can equivalently maximize  $|\mathbf{E}[fg]|$  over  $f$  restricted to have  $\mathbf{E}[f] = 0, \mathbf{E}[f^2] \leq 1$  under  $\mu$ 's marginal on  $\Omega$ ; or, over similarly restricted  $g$  (or both).

## 3 Long Code Test

**Definition 3.1.** For  $i \in [n]$ , the  $i$ -th Long Code is the function  $f(x) = x_i$

Now let us define a  $t$ -function Long Code Test. Suppose we are given oracle access to a collection of Boolean functions  $f_1, f_2, \dots, f_t$ . We want to make as few queries as possible into these functions to decide if all the functions are the same Long Code, or no two functions have some common structure. More precisely, we have the following definition:

**Definition 3.2.** We say that a test  $\mathbf{T} = \mathbf{T}^{f_1, f_2, \dots, f_t}$  is a  $t$ -functions Long Code test with completeness  $c$  and soundness  $s$  if  $\mathbf{T}$  is given oracle access to a family of  $t$  functions  $f_1, f_2, \dots, f_t : \{-1, 1\}^n \rightarrow \{-1, 1\}$ , such that

- if there exists some  $i \in [n]$  such that for all  $a \in [t]$ ,  $f_a(x) = x_i$ , then  $\mathbf{T}$  accepts with probability at least  $c$ , and
- for every  $\varepsilon > 0$ , there exist a constant  $\tau > 0$  and a fixed positive integer  $w$  such that if  $\mathbf{T}$  accepts with probability at least  $s + \varepsilon$ , then there exists some  $i \in [n]$  such that  $\text{XInf}_i^{\leq w}(f_1, \dots, f_t) \geq \tau$

### 3.1 Folding

As introduced in [9], we assume that the functions are folded. We do so by requiring our Long Code test to make queries in a special manner. Suppose the test wants to query  $f$  at point  $x \in \{-1, 1\}^n$ . If  $x_1 = 1$ , then the test queries  $f(x)$  as usual. If  $x_1 = -1$ , then the test queries at the point  $-x$  and use  $-f(-x)$  as the evaluation of  $x$ . Folding ensures that  $f(-x) = -f(x)$  and  $f(\emptyset) = 0$ .

### 3.2 The Test Distribution

We now view  $f_S : \{-1, 1\}^n \rightarrow \{-1, 1\}$  as a function over an  $n$ -fold product set  $\mathcal{X}_S^1 \times \mathcal{X}_S^2 \times \dots \times \mathcal{X}_S^n$  where each  $\mathcal{X}_S^i = \{-1, 1\}^{\{i\}}$ .

For each integer  $k > 2$ , we construct  $\mathcal{T}_k$  as a product probability distribution over the  $n$ -fold product set

$$\prod_{i=1}^n \prod_{S \in 2_{\geq 1}^{[k]}} \mathcal{X}_S^i = \prod_{i=1}^n \mathcal{X}_1^i \times \prod_{i=1}^n \mathcal{X}_2^i \times \dots \times \prod_{i=1}^n \mathcal{X}_n^i \times \prod_{i=1}^n \mathcal{X}_{\{1,2\}}^i \times \dots \times \prod_{i=1}^n \mathcal{X}_{\{1,2,\dots,k\}}^i$$

For each  $i$  we have a distribution  $\mathcal{T}_k^i$  on  $\prod_{S \in 2_{\geq 1}^{[k]}} \mathcal{X}_S^i$  and we think of this as a “correlated space”.

The first distribution is the “Hyper Graph Test” distribution which is a slight modification of the one used by Samorodnitsky and Trevisan [41].

**Definition 3.3.** Define distribution  $\mathcal{H}_k$  generating  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{\{1,2,\dots,k\}}) \in \prod_{S \in 2_{\geq 1}^{[k]}} \mathcal{X}_S^i$  as follows: The bits  $\{\mathbf{x}_j\}_{j \in [k]}$  are independent and uniformly at random; then for each  $S \in 2_{\geq 2}^{[k]}$ ,  $\mathbf{x}_S$  is set to be  $-\prod_{j \in S} \mathbf{x}_j$ .

Note that under  $\mathcal{H}_k$ , the marginal distribution on  $\mathbf{x}_S$  for each  $S \in 2_{\geq 2}^{[k]}$  is also uniformly random. The second distribution is the “Noise” distribution.

**Definition 3.4.** Let  $p, q, r > 0$  be such that  $p + q + (2^k - 1)r = 1$ ,  $p + r = q + (2^k - 2)r = 1/2$ ,  $q = r$  ( $p = 2(2^k - 2)/(2^{k+1} - 2)$ ,  $q = r = 2/(2^{k+1} - 2)$ ). Define distribution  $\mathcal{N}_k$  generating  $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{\{1,2,\dots,k\}}) \in \prod_{S \in 2_{\geq 1}^{[k]}} \mathcal{X}_S^i$  as follows: (i) With probability  $p$ , set  $\mathbf{x} = \mathbf{1}$ , (ii) with probability  $q$ , set  $\mathbf{x} = -\mathbf{1}$ , and (iii) with probability  $(2^k - 1)r$ , first pick  $S \in 2_{\geq 1}^{[k]}$  uniformly at random, then set  $\mathbf{x} = -\mathbf{1}_{-S}$ . Additionally, for  $0 < \delta < 1$ , define distribution  $\mathcal{H}_{k,\delta}$  to be the mixture distribution  $\mathcal{H}_{k,\delta} = (1 - \delta)\mathcal{H}_k + \delta\mathcal{N}_k$ ; i.e., one draws from  $\mathcal{H}_k$  with probability  $1 - \delta$  and from  $\mathcal{N}_k$  with probability  $\delta$ .

Again note that under  $\mathcal{N}_k$  (also  $\mathcal{H}_{k,\delta}$ ), the marginal distribution on  $\mathbf{x}_S$  for each  $S \in 2_{\geq 1}^{[k]}$  is uniformly random.

We are now ready to define the test distribution  $\mathcal{T}_k^i$ :

**Definition 3.5.** For each  $i \in [n]$  we define  $\mathcal{T}_k^i$  to be  $\mathcal{H}_{k,\delta}$ , with

$$\delta = (\varepsilon/9)^2,$$

where the domain of  $\mathcal{H}_{k,\delta}$  is appropriately identified with the domain  $\prod_{S \in 2_{\geq 1}^{[k]}} \mathcal{X}_S^i$  of  $\mathcal{T}_k^i$ .  $\mathcal{T}_k$  is the product distribution of these distributions  $\mathcal{T}_k = \otimes_{i=1}^n \mathcal{T}_k^i$

**Definition 3.6.** For any positive integer  $k$  and  $S_0 \in 2_{\geq 1}^{[k]}$ , we define distribution  $\mathcal{I}_{k,S_0}$  generating  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{\{1,2,\dots,k\}}) \in \prod_{S \in 2_{\geq 1}^{[k]}} \mathcal{X}_S^i$  as follows: First draw from  $\mathcal{H}_k$ ; then uniformly rerandomize  $\mathbf{x}_{S_0}$ .

**Definition 3.7.** For  $0 < \delta < 1$ , define distribution  $\mathcal{I}_{k,S_0,\delta}$  to be the mixture distribution  $\mathcal{I}_{k,S_0,\delta} = (1 - \delta)\mathcal{I}_{k,S_0} + \delta\mathcal{N}_k$ .

### 3.3 A Query Efficient Long Code Test

Let  $k > 2$  be an integer and  $\{f_S\}_{S \in 2_{\geq 1}^{[k]}}$  be a collection of Boolean functions. We define the following Long Code test based on distribution  $\mathcal{T}_k$ :

**Long Code Test  $\mathbf{T}_k$ :** with oracle access to  $\{f_S\}_{S \in 2_{\geq 1}^{[k]}}$

1. Pick  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{\{1,2,\dots,k\}}) \in \prod_{S \in 2_{\geq 1}^{[k]}} \prod_{i=1}^n \mathcal{X}_S^i$  from  $\mathcal{T}_k$ .
2. For each  $S \in 2_{\geq 1}^{[k]}$ , query  $f_S(\mathbf{x}_S)$ .
3. Accept iff either of the followings holds:
  - (a) (Hyper Graph) For every  $S \in 2_{\geq 2}^{[k]}$ ,  $\prod_{s \in S} f_s(\mathbf{x}_s) \neq f_S(\mathbf{x}_S)$ , or
  - (b) (Noise)  $(f_S(\mathbf{x}_1), f_S(\mathbf{x}_2), \dots, f_{\{1,2,\dots,k\}}(\mathbf{x}_{\{1,2,\dots,k\}}))$  is either  $\mathbf{1}$  or  $-\mathbf{1}$  or  $-\mathbf{1}_{-S}$  for some  $S \in 2_{\geq 1}^{[k]}$ .

**Theorem 3.8** (main theorem restated). For every integer  $k > 2$ , there exists a non-adaptive Long Code test  $\mathbf{T}_k$  that makes  $q = 2^k - 1$  queries, has completeness 1 and soundness  $(2q + 3)/2^q$ .

## 4 Analysis of the Test

In this section, we prove Theorem 3.8.

**Completeness Analysis:** The completeness analysis is entirely standard.



**Soundness Analysis:** We arithmetize the probability that  $\mathbf{T}_k$  accepts. Let  $\mathbf{Z}_S = -f_S(\mathbf{x}_S) \prod_{s \in S} f_s(\mathbf{x}_s)$ . Then

$$\Pr[\mathbf{T}_k \text{ accepts}] = \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \prod_{S \in 2_{\geq 2}^{[k]}} \left( \frac{1 + \mathbf{Z}_S}{2} \right) + \prod_{S \in 2_{\geq 1}^{[k]}} \left( \frac{1 - f_S(\mathbf{x}_S)}{2} \right) + \prod_{S \in 2_{\geq 1}^{[k]}} \left( \frac{1 + f_S(\mathbf{x}_S)}{2} \right) + \sum_{S' \in 2_{\geq 1}^{[k]}} \left( \frac{1 - f_{S'}(\mathbf{x}_{S'})}{2} \right) \prod_{S \in 2_{\geq 1}^{[k]} \setminus \{S'\}} \left( \frac{1 + f_S(\mathbf{x}_S)}{2} \right) \right].$$

This can be written as

$$\begin{aligned} &= \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \frac{1}{2^{|2_{\geq 2}^{[k]}|}} \sum_{\mathbf{S} \subseteq 2_{\geq 2}^{[k]}} \prod_{S \in \mathbf{S}} \mathbf{Z}_S + \frac{1}{2^{|2_{\geq 1}^{[k]}|}} \sum_{\mathbf{S} \subseteq 2_{\geq 1}^{[k]}} (-1)^{|\mathbf{S}|} \prod_{S \in \mathbf{S}} f_S(\mathbf{x}_S) + \frac{1}{2^{|2_{\geq 1}^{[k]}|}} \sum_{\mathbf{S} \subseteq 2_{\geq 1}^{[k]} \setminus \{S'\}} \prod_{S \in \mathbf{S}} f_S(\mathbf{x}_S) + \sum_{S' \in 2_{\geq 1}^{[k]}} \left( \frac{1 - f_{S'}(\mathbf{x}_{S'})}{2^{|2_{\geq 1}^{[k]}|}} \right) \sum_{\mathbf{S} \subseteq 2_{\geq 1}^{[k]} \setminus \{S'\}} \prod_{S \in \mathbf{S}} f_S(\mathbf{x}_S) \right] \\ &\leq \frac{1}{2^{|2_{\geq 2}^{[k]}|}} \sum_{\mathbf{S} \subseteq 2_{\geq 2}^{[k]}} \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \prod_{S \in \mathbf{S}} \mathbf{Z}_S \right] + \frac{2 + |2_{\geq 1}^{[k]}|}{2^{|2_{\geq 1}^{[k]}|}} \sum_{\mathbf{S} \subseteq 2_{\geq 1}^{[k]}} \left| \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \prod_{S \in \mathbf{S}} f_S(\mathbf{x}_S) \right] \right|. \end{aligned}$$

We would like to prove

$$\begin{aligned} \frac{1}{2^{|2_{\geq 2}^{[k]}|}} \sum_{\mathbf{S} \subseteq 2_{\geq 2}^{[k]}} \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \prod_{S \in \mathbf{S}} \mathbf{Z}_S \right] &= \frac{1}{2^{|2_{\geq 2}^{[k]}|}} + \frac{1}{2^{|2_{\geq 2}^{[k]}|}} \sum_{\mathbf{S} \subseteq 2_{\geq 2}^{[k]}, \mathbf{S} \neq \emptyset} \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \prod_{S \in \mathbf{S}} \mathbf{Z}_S \right] \\ &\leq \frac{1}{2^{|2_{\geq 2}^{[k]}|}} + \varepsilon/3 \end{aligned}$$

and

$$\begin{aligned} \frac{2 + |2_{\geq 1}^{[k]}|}{2^{|2_{\geq 1}^{[k]}|}} \sum_{\mathbf{S} \subseteq 2_{\geq 1}^{[k]}} \left| \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \prod_{S \in \mathbf{S}} f_S(\mathbf{x}_S) \right] \right| &= \frac{2 + |2_{\geq 1}^{[k]}|}{2^{|2_{\geq 1}^{[k]}|}} + \frac{2 + |2_{\geq 1}^{[k]}|}{2^{|2_{\geq 1}^{[k]}|}} \sum_{\mathbf{S} \subseteq 2_{\geq 1}^{[k]}, \mathbf{S} \neq \emptyset} \left| \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \prod_{S \in \mathbf{S}} f_S(\mathbf{x}_S) \right] \right| \\ &\leq \frac{2 + |2_{\geq 1}^{[k]}|}{2^{|2_{\geq 1}^{[k]}|}} + 2\varepsilon/3 \end{aligned}$$

for some  $\varepsilon > 0$ . To do so, the following lemma is enough.

**Lemma 4.1.** *For any  $\varepsilon > 0$ , there exists  $\tau, \delta, w$  such that if  $\text{XInf}_i^{\leq w}(f_1, \dots, f_{[k]}) < \tau$  for any  $i$  then, for any  $\mathbf{S} \subseteq 2_{\geq 1}^{[k]}$*

$$\left| \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \prod_{S \in \mathbf{S}} f_S(\mathbf{x}_S) \right] \right| \leq \varepsilon/3.$$

#### 4.1 Proof of Lemma 4.1

First note that for any  $\mathbf{S} \subseteq 2_{\geq 1}^{[k]}$ ,  $|\mathbf{S}| \leq 2$ ,

$$\left| \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \prod_{S \in \mathbf{S}} f_S(\mathbf{x}_S) \right] \right| = 0,$$

since the marginal distribution of  $\mathcal{H}_{k,\delta}^n$  on  $\prod_{S \in \mathbf{S}} \prod_{i=1}^n \mathcal{X}_S^i$  is uniformly random and functions are folded. We need to prove Lemma 4.1 for  $\mathbf{S} \subseteq 2_{\geq 1}^{[k]}$ ,  $|\mathbf{S}| \geq 3$  and actually we only prove Lemma 4.1 for  $\mathbf{S} = 2_{\geq 1}^{[k]}$ ; any choice of  $\mathbf{S}$  can be handled in almost the same way. It is easy to see that following lemmas with the triangle inequality concludes the proof of Lemma 4.1.

**Lemma 4.2.** *By taking  $\gamma > 0$  small enough as a function of  $\delta, k$ ,*

$$\left| \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \prod_{S \in \mathbf{S}} f_S(\mathbf{x}_S) - \prod_{S \in \mathbf{S}} T_{1-\gamma} f_S(\mathbf{x}_S) \right] \right| \leq \sqrt{\delta}.$$

**Lemma 4.3.** *For any  $S_0 \in \mathbf{S}$ ,  $|\mathbf{S}| \geq 3$ ,  $\gamma, \delta$ , there exist  $\tau, w$  such that if  $\text{XInf}_i^{\leq w}(f_1, \dots, f_{[k]}) < \tau$  for any  $i$  then,*

$$\left| \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \prod_{S \in \mathbf{S}} T_{1-\gamma} f_S(\mathbf{x}_S) \right] - \mathbf{E}_{\mathcal{I}_{k,S_0,\delta}^n} \left[ \prod_{S \in \mathbf{S}} T_{1-\gamma} f_S(\mathbf{x}_S) \right] \right| \leq \sqrt{\delta}.$$

**Lemma 4.4.** *For any  $S_0 \in \mathbf{S}$ ,  $\gamma, \delta$ , it holds that*

$$\left| \mathbf{E}_{\mathcal{I}_{k,S_0,\delta}^n} \left[ \prod_{S \in \mathbf{S}} T_{1-\gamma} f_S(\mathbf{x}_S) \right] \right| \leq \sqrt{\delta}.$$

#### 4.2 Proof of Lemma 4.2

We define a lexicographical order  $<_{\text{lex}}$  between elements of  $2_{\geq 1}^{[k]}$  as

$$1 <_{\text{lex}} 2 <_{\text{lex}} \dots <_{\text{lex}} n <_{\text{lex}} \{1, 2\} <_{\text{lex}} \{1, 3\} <_{\text{lex}} \dots <_{\text{lex}} \{1, 2, \dots, k\},$$

and let  $\mathbf{S}_l = \{S \in 2_{\geq 1}^{[k]} : S \text{ appears within } l\text{-th by } <_{\text{lex}}\}$ . It is easy to see that the following lemma with the triangle inequality completes the proof of Lemma 4.2.

**Lemma 4.5.** *By taking  $\gamma > 0$  small enough as a function of  $\delta, k$ , we ensure, for any  $l \in [2_{\geq 1}^{[k]}]$ ,*

$$\left| \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \left( \prod_{S \in \mathbf{S}_l} f_S(\mathbf{x}_S) \right) \left( \prod_{S \in \mathbf{S} \setminus \mathbf{S}_l} T_{1-\gamma} f_S(\mathbf{x}_S) \right) - \left( \prod_{S \in \mathbf{S}_{l-1}} f_S(\mathbf{x}_S) \right) \left( \prod_{S \in \mathbf{S} \setminus \mathbf{S}_{l-1}} T_{1-\gamma} f_S(\mathbf{x}_S) \right) \right] \right| \leq \sqrt{\delta} / 2^{|2_{\geq 1}^{[k]}|}.$$

For the proof of Lemma 4.5, see Appendix B.

#### 4.3 Proof of Lemma 4.3

From Proposition 2.6, by setting  $w$  as a function of  $\gamma$  and  $\tau$ , we can assume  $\text{XInf}_i(T_{1-\gamma/2} f_1, \dots, T_{1-\gamma/2} f_{[k]}) < 2\tau$  for all  $i$  where  $\gamma$  is the value chosen in Appendix B. To prove Lemma 4.3 for the case  $|\mathbf{S}| \geq 3$ , we use the inductive proof. We replace the distribution from  $\mathcal{H}_{k,\delta}^n = \otimes_{i=1}^n \mathcal{H}_{k,\delta}$  to  $\mathcal{I}_{k,\delta}^n = \otimes_{i=1}^n \mathcal{I}_{k,S_0,\delta}$  one component at a time. We will show the following lemma.

**Lemma 4.6.** For any  $l \in [n]$ ,

$$\left| \mathbf{E}_{\otimes_{i=1}^{l-1} \mathcal{I}_{k,S_0,\delta} \times \otimes_{i=l}^n \mathcal{H}_{k,\delta}} \left[ \prod_{S \in \mathbf{S}} T_{1-\gamma} f_S(\mathbf{x}_S) \right] - \mathbf{E}_{\otimes_{i=1}^l \mathcal{I}_{k,S_0,\delta} \times \otimes_{i=l+1}^n \mathcal{H}_{k,\delta}} \left[ \prod_{S \in \mathbf{S}} T_{1-\gamma} f_S(\mathbf{x}_S) \right] \right| \leq \Delta_l.$$

where

$$\Delta_l = \tau^{\gamma/2|\mathbf{S}|} 2^{|\mathbf{S}|} \sum_{S \in \mathbf{S}} \text{Inf}_l(T_{1-\gamma/2} f_S)$$

For the proof of Lemma 4.6, see Appendix C.

*Proof of Lemma 4.3.* From Lemma 4.6,

$$\begin{aligned} & \left| \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \prod_{S \in \mathbf{S}} T_{1-\gamma} f_S(\mathbf{x}_S) \right] - \mathbf{E}_{\mathcal{I}_{k,\delta}^n} \left[ \prod_{S \in \mathbf{S}} T_{1-\gamma} f_S(\mathbf{x}_S) \right] \right| \\ & \leq (2\tau)^{\gamma/2|\mathbf{S}|} 2^{|\mathbf{S}|} \sum_{i \in [n]} \sum_{S \in \mathbf{S}} \text{Inf}_i(T_{1-\gamma/2} f_S) \\ & \leq |\mathbf{S}| (2\tau)^{\gamma/2|\mathbf{S}|} 2^{|\mathbf{S}|} / \gamma. \end{aligned}$$

By choosing  $\tau = O\left((\gamma\sqrt{\delta}/2|\mathbf{S}|)^{\gamma/|\mathbf{S}|}\right)$ , the lemma follows.  $\square$

#### 4.4 Proof of Lemma 4.4

**Lemma 4.7.** For any  $i \in [n]$  and  $S_0 \in \mathbf{S}$ , it holds that

$$\rho(\mathcal{X}_{S_0}^i, \prod_{S \in \mathbf{S} \setminus \{S_0\}} \mathcal{X}_S^i; \mathcal{I}_{k,S_0,\delta}) \leq \sqrt{\delta}.$$

For the proof of Lemma 4.7, see Appendix D.

*Proof of Lemma 4.4.* From Lemma 4.7 and Proposition 2.12 of [31],

$$\rho(\mathcal{X}_{S_0}, \prod_{S \in \mathbf{S} \setminus \{S_0\}} \mathcal{X}_S; \mathcal{I}_{k,S_0,\delta}^n) \leq \sqrt{\delta}.$$

Since  $\mathbf{E}_{\mathcal{I}_{k,S_0,\delta}^n} [T_{1-\gamma} f_S]$  has mean 0 and variance at most 1 for any  $S$ , the lemma immediately follows.  $\square$

## 5 Discussion

Unfortunately, our Long Code test does not immediately imply a new PCP characterization of NP. A Long Code test without consistency checks is most easily composed with Unique Games [26] as an outer verifier in PCP constructions. However, since Unique Games cannot have perfect completeness, an obvious approach that combines our tester and Unique Games does not imply a new PCP construction. A variant of Unique Games, which is called  $d$ -to-1 Games [26], is conjectured to have perfect completeness. Though it seems hard to combine our tester and  $d$ -to-1 Games, we hope that it is possible and we obtain a new conditional PCP construction.

## References

- [1] Per Austrin, Johan Håstad. Randomly supported independence and resistance. In *Proc. of STOC 2009*, pages 483–492, 2009.
- [2] Yonatan Aumann, Johan Håstad, Michael O. Rabin, Madhu Sudan. Linear-Consistency Testing. *J. Comput. Syst. Sci.*, 62(4):589–607, 2001.
- [3] N. Alon, T. Kaufman, M. Krivelevich, S. Litsyn, and D. Ron. Testing low-degree polynomials over  $\text{GF}(2)$ . In *Proc. of RANDOM-APPROX 2003*, pages 188–199, 2003.
- [4] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, Mario Szegedy. Proof Verification and the Hardness of Approximation Problems. *J. ACM*, 45(3):501–555, 1998.
- [5] Per Austrin, Elchanan Mossel. Approximation Resistant Predicates from Pairwise Independence. *Computational Complexity* (Online First), 23 pages, 2009.
- [6] Sanjeev Arora, Shmuel Safra. Probabilistic Checking of Proofs: A New Characterization of NP. *J. ACM*, 45(1):70–122, 1998.
- [7] William Beckner. Inequalities in Fourier analysis. *Annals of Mathematics*, 102(1):159–182, 1975.
- [8] Mihir Bellare, Don Coppersmith, Johan Håstad, Marcos A. Kiwi, Madhu Sudan. Linearity testing in characteristic two. *IEEE Transactions on Information Theory*, 42(6):1781–1795, 1996.
- [9] Mihir Bellare, Oded Goldreich, Madhu Sudan. Free Bits, PCPs, and Nonapproximability-Towards Tight Results. *SIAM J. Comput.*, 27(3):804–915, 1998.
- [10] Eric Blais. Improved Bounds for Testing Juntas. In *Proc. of APPROX-RANDOM 2008*, pages 317–330, 2008.
- [11] Eric Blais. Testing juntas nearly optimally. In *Proc. of STOC 2009*, pages 151–158, 2009.
- [12] Manuel Blum, Michael Luby, Ronitt Rubinfeld. Self-Testing/Correcting with Applications to Numerical Problems. *J. Comput. Syst. Sci.*, 47(3):549–595, 1993.
- [13] Aline Bonami. Étude des coefficients de Fourier des fonctions de  $L^p(G)$ . *Annales de l’Institut Fourier*, 20(2):335.402, 1970.
- [14] Eli Ben-Sasson, Prahladh Harsha, Sofya Raskhodnikova. Some 3CNF Properties Are Hard to Test. *SIAM J. Comput.*, 35(1):1–21, 2005.
- [15] Victor Chen. A hypergraph dictatorship test with perfect completeness In *Proc. of RANDOM 2009*, to appear.
- [16] Ilias Diakonikolas, Homin K. Lee, Kevin Matulef, Krzysztof Onak, Ronitt Rubinfeld, Rocco A. Servedio, Andrew Wan. Testing for Concise Representations. In *Proc. of FOCS 2007*, pages 549–558, 2007.
- [17] Lars Engebretsen, Jonas Holmerin. More efficient queries in PCPs for NP and improved approximation hardness of maximum CSP. *Random Struct. Algorithms*, 33(4):497–514, 2008.

- [18] Eldar Fischer. The art of uninformed decisions: A primer to property testing. *The Bulletin of the European Association for Theoretical Computer Science*, 75:97–126, 2001.
- [19] Venkatesan Guruswami, Daniel Lewin, Madhu Sudan, Luca Trevisan. A Tight Characterization of NP with 3 Query PCPs. In *Proc. of FOCS 1998*, pages 8–17, 1998.
- [20] Venkatesan Guruswami, Prasad Raghavendra. Constraint Satisfaction over a Non-Boolean Domain: Approximation Algorithms and Unique-Games Hardness. In *Proc. of APPROX-RANDOM 2008*, pages 77–90, 2008.
- [21] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97:1061–1083, 1975.
- [22] Johan Håstad. Some optimal inapproximability results. *J. ACM*. 48(4):798–859, 2001.
- [23] Johan Håstad. On the Approximation Resistance of a Random Predicate. In *Proc. of APPROX-RANDOM 2007*, pages 149–163, 2007.
- [24] Johan Håstad, Subhash Khot. Query Efficient PCPs with Perfect Completeness. *Theory of Computing*, 1(1):119–148, 2005.
- [25] Johan Håstad, Avi Wigderson. Simple analysis of graph tests for linearity and PCP. *Random Struct. Algorithms*, 22(2):139–160, 2003.
- [26] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proc. of STOC 2002*, pages 767–775, 2002.
- [27] Subhash Khot. Inapproximability results via Long Code based PCPs. *SIGACT News*, 36(2):25–42, 2005.
- [28] Subhash Khot, Rishi Saket. A 3-Query Non-Adaptive PCP with Perfect Completeness. In *Proc. of IEEE Conference on Computational Complexity 2006*, pages 159–169, 2006.
- [29] Shachar Lovett. Lower bounds for adaptive linearity tests. In *Proc. of STACS 2008*, pages 515–526, 2008.
- [30] Elchanan Mossel, Ryan O’Donnell and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. *Ann. Math.*, to appear.
- [31] Elchanan Mossel. Gaussian Bounds for Noise Correlation of Functions. *GAF*, to appear.
- [32] Ryan O’Donnell. Analysis of Boolean functions lecture notes, 2007. <http://www.cs.cmu.edu/~odonnell/boolean-analysis/>.
- [33] Ryan O’Donnell, Yi Wu. 3-bit dictator testing: 1 vs. 5/8. In *Proc. of SODA 2009*, pages 365–373, 2009.
- [34] Ryan O’Donnell, Yi Wu. Conditional hardness for satisfiable 3-CSPs. In *Proc. of STOC 2009*, pages 493–502, 2009.
- [35] Michal Parnas, Dana Ron, Alex Samorodnitsky. Testing Basic Boolean Formulae. *SIAM J. Discrete Math.*, 16(1):20–46, 2002.
- [36] Prasad Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proc. of STOC 2008*, pages 245–254, 2008.

- [37] Ran Raz. A Parallel Repetition Theorem. *SIAM J. Comput.*, 27(3):763–803, 1998.
- [38] Dana Ron. Property Testing: A Learning Theory Perspective. Now Publishers Inc., 112 pages, 2008.
- [39] Madhu Sudan, Luca Trevisan. Probabilistically Checkable Proofs with Low Amortized Query Complexity. In *FOCS 1998*, pages 18–27, 1998.
- [40] Alex Samorodnitsky, Luca Trevisan. A PCP characterization of NP with optimal amortized query complexity. In *Proc. of STOC 2000*, pages 191–199, 2000.
- [41] Alex Samorodnitsky, Luca Trevisan. Gowers Uniformity, Influence of Variables, and PCPs. *SIAM J. Comput.*, 39(1):323–360, 2009.
- [42] Luca Trevisan. Recycling Queries in PCPs and in Linearity Tests (Extended Abstract). In *Proc. of STOC 1998*, pages 299–308, 1998.
- [43] Uri Zwick. Approximation Algorithms for Constraint Satisfaction Problems Involving at Most Three Variables per Constraint. In *Proc. of SODA 1998*, pages 201–210, 1998.

## Appendix

### A Correlated Spaces and Noise

**Definition A.1** (Markov Operator). *Let  $(\Omega, \Psi, \mu)$  be a correlated probability space. Markov Operator associated with the probability space is the operator mapping  $f \in L^p(\Psi, \mu)$  to  $Uf \in L^p(\Omega, \mu)$  by*

$$(Uf)(x) = \mathbf{E}[f(Y)|X = x]$$

for  $x \in \Omega$  where  $(X, Y) \in (\Omega, \Psi)$  is distributed according to  $\mu$

**Lemma A.2** (Lemma 2.9 of [31]). *Let  $(\Omega, \Psi, \mu)$  be two correlated spaces such that the probability of the smallest atom in  $\Omega \times \Psi$  is at least  $\alpha > 0$ . Define a bipartite graph  $G = (\Omega, \Psi, E)$  where  $(a, b) \in \Omega \times \Psi$  satisfies  $(a, b) \in E$  if  $\mu(a, b) > 0$ . If  $G$  is connected then*

$$\rho(\Omega, \Psi; \mu) \leq 1 - \alpha^2/2.$$

Efron-Stein Decomposition is an analogous to Fourier expansion for (non-Boolean) functions over general domains. Given a vector  $x$  in an  $n$  dimensional product space and  $S \subseteq [n]$ , we write  $x_S$  for the vector  $(x_i)_{i \in S}$ .

**Definition A.3** (Efron-Stein Decomposition). *Let  $(\Omega_1, \mu_1), \dots, (\Omega_n, \mu_n)$  be discrete probability spaces  $(\Omega, \mu) = \prod_{i=1}^n (\Omega_i, \mu_i)$ . The Efron-Stein decomposition of  $f : \Omega \rightarrow \mathbb{R}$  is given by*

$$f(x) = \sum_{S \subseteq [n]} f_S(x),$$

where

$$f_S(x) = \sum_{S' \subseteq S} (-1)^{|S \setminus S'|} \mathbf{E}[f(X)|X_{S'} = x_{S'}]$$

From the definition of  $f_S(x)$ ,  $f_S$  depends only on  $x_S$ , and it is known that for all  $S \not\subseteq S'$  and  $x_{S'}$  it holds that  $\mathbf{E}[f_S|X_{S'} = x_{S'}] = 0$  [31].

**Proposition A.4** (Proposition 2.11 of [31]). *Let  $(\Omega_i, \Psi_i, \mu_i)$  be correlated probability spaces and let  $U_i$  be the Markov operator associated with  $\Omega_i$  and  $\Psi_i$  for  $1 \leq i \leq n$ . Suppose  $f \in L^2(\Psi)$  has Efron-Stein decomposition. Let*

$$\Omega = \prod_{i=1}^n \Omega_i, \Psi = \prod_{i=1}^n \Psi_i, \mu = \prod_{i=1}^n \mu_i, U = \bigotimes_{i=1}^n U_i$$

Then the Efron-Stein decomposition of  $Uf$  satisfies

$$(Uf)_S = U(f_S)$$

**Proposition A.5** (Proposition 2.12 of [31]). *Assume the setting of Lemma A.4 and that further for all  $i$  it holds that  $\rho(\Omega_i, \Psi_i, \mu_i) \leq \rho_i$ , Then for all  $f$  it holds that*

$$\|U(f_S)\|_2 \leq \left( \prod_{i \in S} \rho_i \right) \|f_S\|_2.$$

**Proposition A.6** (Proposition 2.13 of [31]). *Assume the setting of Lemma A.4. Then*

$$\rho(\Omega, \Psi; \mu) = \max \rho(\Omega_i, \Psi_i, \mu_i).$$

**Lemma A.7** (Hypercontractive inequality [7, 13]). *Let  $p, q, \gamma$  be real numbers such that  $1 \leq p \leq q$  and  $1 - \gamma \leq \sqrt{\frac{p-1}{q-1}}$ , Then it holds that for all  $f : \{-1, 1\}^n \rightarrow \mathbb{R}$*

$$\|T_{1-\gamma} f\|_q \leq \|f\|_p$$

## B Proof of Lemma 4.5

**Lemma B.1** (Lemma 4.5 restated). *By taking  $\gamma > 0$  small enough as a function of  $\delta, k$ , we ensure, for any  $l \in [|2_{\geq 1}^{[k]}|]$ ,*

$$\left| \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \left( \prod_{S \in \mathbf{S}_l} f_S(\mathbf{x}_S) \right) \left( \prod_{S \in \mathbf{S} \setminus \mathbf{S}_l} T_{1-\gamma} f_S(\mathbf{x}_S) \right) - \left( \prod_{S \in \mathbf{S}_{l-1}} f_S(\mathbf{x}_S) \right) \left( \prod_{S \in \mathbf{S} \setminus \mathbf{S}_{l-1}} T_{1-\gamma} f_S(\mathbf{x}_S) \right) \right] \right| \leq \sqrt{\delta} / 2^{|2_{\geq 1}^{[k]}|}.$$

*Proof.* We only prove the case for  $l = |2_{\geq 1}^{[k]}|$ . Other cases can be handled in almost the same way by changing the role of  $f_{[k]}$  and  $\prod_{i=1}^n \mathcal{X}_{[k]}^i$  to  $f_S$  and  $\prod_{i=1}^n \mathcal{X}_S^i$  for each  $S \in 2_{\geq 1}^{[k]}$ . (Notice that we can do so by Lemma E.1 that ensures the correlation is strictly less than 1. This is not the case in the proof of Theorem C.1 in [34]; they need some additional analysis for some choice of  $l$ , i.e., the proof of Lemme C.5.) We have

$$\begin{aligned} & \left| \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \prod_{S \in 2_{\geq 1}^{[k]}} f_S(\mathbf{x}_S) - \left( \prod_{S \in 2_{\geq 1}^{[k]} \setminus [k]} f_S(\mathbf{x}_S) \right) T_{1-\gamma} f_{[k]}(\mathbf{x}_{[k]}) \right] \right| \\ &= \left| \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \left( \prod_{S \in 2_{\geq 1}^{[k]} \setminus [k]} f_S(\mathbf{x}_S) \right) (id - T_{1-\gamma}) f_{[k]}(\mathbf{x}_{[k]}) \right] \right| \\ &= \left| \mathbf{E}_{\mathcal{H}_{k,\delta}^n} \left[ \left( \prod_{S \in 2_{\geq 1}^{[k]} \setminus [k]} f_S(\mathbf{x}_S) \right) U(id - T_{1-\gamma}) f_{[k]}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{\{2,3,\dots,k\}}) \right] \right|, \end{aligned} \quad (1)$$

where  $U$  is the Markov operator for the correlated probability space  $(\prod_{S \in 2_{\geq 1}^{[k]} \setminus [k]} (\prod_{i=1}^n \mathcal{X}_S^i), \prod_{i=1}^n \mathcal{X}_{[k]}^i; \mathcal{H}_{k,\delta}^n)$ , which maps functions on the latter space  $\prod_{i=1}^n \mathcal{X}_{[k]}^i$  into functions on the former space  $\prod_{S \in 2_{\geq 1}^{[k]} \setminus [k]} (\prod_{i=1}^n \mathcal{X}_S^i)$ .

Now we consider the quantity inside the expectation in (1) to be a product of two functions on  $\prod_{S \in 2_{\geq 1}^{[k]} \setminus [k]} (\prod_{i=1}^n \mathcal{X}_S^i)$ , namely,  $F = \prod_{S \in 2_{\geq 1}^{[k]} \setminus [k]} f_S$  and  $G = U(id - T_{1-\gamma})f_{[k]}$ . We take the Efron-Stein decomposition of these two functions with respect to the (product) distribution  $\prod_{S \in 2_{\geq 1}^{[k]} \setminus [k]} (\prod_{i=1}^n \mathcal{X}_S^i)$ . Then by the orthogonality of the Efron-Stein decomposition and the Cauchy-Schwartz inequality,

$$(1) = \left| \sum_{V \subseteq [n]} \mathbf{E}_{\mathcal{H}_{k,\delta}^n} [F_V(\mathbf{x}_1, \dots, \mathbf{x}_{\{2, \dots, k\}}) \cdot G_V(\mathbf{x}_1, \dots, \mathbf{x}_{\{2, \dots, k\}})] \right| \leq \sqrt{\sum_{V \subseteq [n]} \|F_V\|_2^2} \sqrt{\sum_{V \subseteq [n]} \|G_V\|_2^2}. \quad (2)$$

where the 2-norms  $\|\cdot\|_2$  are with respect to  $\mathcal{H}_{k,\delta}^n$ 's marginal distribution on  $\prod_{S \in 2_{\geq 1}^{[k]} \setminus [k]} (\prod_{i=1}^n \mathcal{X}_S^i)$ .

By orthogonality again, the quantity  $\sqrt{\sum_{V \subseteq [n]} \|F_V\|_2^2}$  is just  $\|F\|_2^2$ , which is precisely 1 because  $F$ 's range is  $\{-1, 1\}$ . Hence we have

$$(2) \leq \sqrt{\sum_{V \subseteq [n]} \|G_V\|_2^2}. \quad (3)$$

From Proposition 2.11 of [31],  $G_V = UG'_V$  where  $G' = (id - T_{1-\gamma})f_{[k]}$ . Here the Efron-Stein decomposition is with respect to  $\mathcal{H}_{k,\delta}^n$ 's marginal distribution on  $\prod_{i=1}^n \mathcal{X}_{[k]}^i$ , viz., the uniform distribution. It is also easy to check that this Efron-Stein decomposition of  $f_{[k]}$  has

$$(f_{[k]})_V = \widehat{f}_{[k]}(V)\chi_V.$$

It follows that applying the Bonami-Beckner operator  $T_{1-\gamma}$  to  $f_{[k]}$  also commutes with taking the Efron-Stein decomposition (technically, this uses  $\gamma < 1$ ). Hence we have  $G_V = UG'_V = U(id - T_{1-\gamma})(f_{[k]})_V$ . Substituting this into (3) yields

$$(3) = \sqrt{\sum_{V \subseteq [n]} \|U(id - T_{1-\gamma})(f_{[k]})_V\|_2^2}. \quad (4)$$

From Lemma E.1, which is proved later, the correlation

$$\rho\left(\prod_{S \in 2_{\geq 1}^{[k]} \setminus [k]} \left(\prod_{i=1}^n \mathcal{X}_S^i\right), \prod_{i=1}^n \mathcal{X}_{[k]}^i; \mathcal{H}_{k,\delta}^n\right) \leq \rho_0 = 1 - \alpha^2/2$$

where  $\alpha = 2\delta/(2^{k+1} - 2)$ ; Applying Lemma 2.9 of [31], we conclude that for each  $V \subseteq [n]$ ,

$$\|U(id - T_{1-\gamma})(f_{[k]})_V\|_2 \leq \rho_0^{|V|} \|(id - T_{1-\gamma})(f_{[k]})_V\|_2, \quad (5)$$

where the 2-norm on the right is with respect to the uniform distribution on  $\{-1, 1\}^n$ .

Next, observe that

$$\|(id - T_{1-\gamma})(f_{[k]})_V\|_2^2 = (1 - (1 - \gamma)^{2|V|}) \widehat{f}_{[k]}(V)^2.$$



Substituting (5) and then into (4), we determine

$$(4) \leq \sqrt{\sum_{V \subseteq [n]} \rho_0^{2|V|} (1 - (1 - \gamma)^{2|V|}) \widehat{f}_{[k]}(V)^2}. \quad (6)$$

We now bound

$$\rho_0^{2|V|} (1 - (1 - \gamma)^{2|V|}) \leq \exp(-\alpha^2|V|) \cdot (2|V|\gamma);$$

simple calculation shows that the maximum of this, over  $|S|$ , is at most

$$\frac{2}{e\alpha^2}\gamma. \quad (7)$$

By choosing  $\gamma > 0$  small enough we can upper-bound (7) by  $(\sqrt{\delta}/2^{\lfloor 2^{\lfloor k/2 \rfloor} \rfloor})^2$ ; specifically, we need to choose

$$\gamma = O(\alpha^2 \delta / 2^{\lfloor 2^{\lfloor k/2 \rfloor} \rfloor}).$$

Doing so, we get

$$(6) \leq \sqrt{\sum_{V \subseteq [n]} (\sqrt{\delta}/2^{\lfloor 2^{\lfloor k/2 \rfloor} \rfloor})^2 \widehat{f}_{[k]}(V)^2} = (\sqrt{\delta}/2^{\lfloor 2^{\lfloor k/2 \rfloor} \rfloor}) \sqrt{\mathbf{E}[f_{[k]}^2]} = \sqrt{\delta}/2^{\lfloor 2^{\lfloor k/2 \rfloor} \rfloor}.$$

□

## C Proof of Lemma 4.6

**Lemma C.1** (Lemma 4.6 restated). *For any  $l \in [n]$ ,*

$$\left| \mathbf{E}_{\otimes_{i=1}^{l-1} \mathcal{I}_{k, S_0, \delta} \times \otimes_{i=l}^n \mathcal{H}_{k, \delta}} \left[ \prod_{S \in \mathbf{S}} T_{1-\gamma} f_S(\mathbf{x}_S) \right] - \mathbf{E}_{\otimes_{i=1}^l \mathcal{I}_{k, S_0, \delta} \times \otimes_{i=l+1}^n \mathcal{H}_{k, \delta}} \left[ \prod_{S \in \mathbf{S}} T_{1-\gamma} f_S(\mathbf{x}_S) \right] \right| \leq \Delta_l.$$

where

$$\Delta_l = \tau^{\gamma/2} 2^{|\mathbf{S}|} \sum_{S \in \mathbf{S}} \text{Inf}_l(T_{1-\gamma/2} f_S)$$

We only show the proof of Lemma 4.6 for  $l = 1$ . The case for  $l = 2, \dots, n$  is the same. Actually, we only use the fact that the unchanged product distribution  $\mathcal{H}'_{k, \delta} = \otimes_{i=2}^n \mathcal{H}_{k, \delta}$  (when  $l = 1$ ) have the property that its marginals on  $\mathbf{x}_S$  for  $S \in \mathbf{S}$  are uniform. Here, We identify a set  $S \in [k]$  with a vector over  $\mathbb{F}_2^k$ , i.e., the  $i$ th element of the vector corresponding to  $S$  is 1 iff  $S$  contains  $i$ . Let  $\mathfrak{X}_{\mathbf{S}}$  be a set of  $\mathbf{T} \subseteq \mathbf{S}$  such that  $S_0$  and the sum of all elements of  $\mathbf{T}$  equals the zero vector. The following lemma holds.

**Lemma C.2.** *For any functions  $f_S : \{-1, 1\} \rightarrow \mathbb{R}$ ,*

$$\mathbf{E}_{\mathcal{H}_{k, \delta}} \left[ \prod_{S \in \mathbf{S}} f_S(\mathbf{x}_S) \right] - \mathbf{E}_{\mathcal{I}_{k, S_0, \delta}} \left[ \prod_{S \in \mathbf{S}} f_S(\mathbf{x}_S) \right] = -(1 - \delta) \sum_{\mathbf{T} \in \mathfrak{X}_{\mathbf{S}}} \prod_{S \in \mathbf{T}} \widehat{f}_S(\{1\})$$

*Proof.* Taking Fourier expansion of each function in the left hand side, we get

$$\sum_{\{U_S\}_{S \in \mathbf{S}}} \prod_{S \in \mathbf{S}} \widehat{f}_S(U_S) \left( \mathbf{E}_{\mathcal{H}_{k,\delta}} \left[ \prod_{S \in \mathbf{S}} \chi_{U_S}(\mathbf{x}_S) \right] - \mathbf{E}_{\mathcal{I}_{k,S_0,\delta}} \left[ \prod_{S \in \mathbf{S}} \chi_{U_S}(\mathbf{x}_S) \right] \right).$$

From  $\mathcal{H}_{k,\delta} = (1 - \delta)\mathcal{H}_k + \delta\mathcal{N}_k$  and  $\mathcal{I}_{k,\delta} = (1 - \delta)\mathcal{I}_k + \delta\mathcal{N}_k$ , it follows that

$$\mathbf{E}_{\mathcal{H}_{k,\delta}} \left[ \prod_{S \in \mathbf{S}} \chi_{U_S}(\mathbf{x}_S) \right] - \mathbf{E}_{\mathcal{I}_{k,S_0,\delta}} \left[ \prod_{S \in \mathbf{S}} \chi_{U_S}(\mathbf{x}_S) \right] = (1 - \delta) \left( \mathbf{E}_{\mathcal{H}_k} \left[ \prod_{S \in \mathbf{S}} \chi_{U_S}(\mathbf{x}_S) \right] - \mathbf{E}_{\mathcal{I}_{k,S_0}} \left[ \prod_{S \in \mathbf{S}} \chi_{U_S}(\mathbf{x}_S) \right] \right).$$

Notice that  $\mathcal{H}_k$  and  $\mathcal{I}_{k,S_0}$  have the same marginal distribution except  $\mathcal{X}_{S_0}$ . Thus, to make the expression above non zero,  $\mathbf{S}$  must contain  $S_0$  and  $U_{S_0}$  must be  $\{1\}$ . In such a case,

$$\mathbf{E}_{\mathcal{I}_{k,S_0}} \left[ \chi_{\{1\}}(\mathbf{x}_{S_0}) \prod_{S \in \mathbf{S} \setminus S_0} \chi_{U_S}(\mathbf{x}_S) \right] = \mathbf{E}_{\mathcal{I}_{k,S_0}} [\chi_{\{1\}}(\mathbf{x}_{S_0})] \mathbf{E}_{\mathcal{I}_{k,S_0}} \left[ \prod_{S \in \mathbf{S} \setminus S_0} \chi_{U_S}(\mathbf{x}_S) \right] = 0.$$

It is not hard to see that  $\mathbf{E}_{\mathcal{H}_k} [\prod_{S \in \mathbf{S}} \chi_{U_S}(\mathbf{x}_S)]$  becomes non zero iff there exists  $\mathbf{T} \in \mathfrak{X}_{\mathbf{S}}$  such that  $U_S = \{1\}$  for  $S \in \mathbf{T}$  and  $U_S = \emptyset$  for  $S \in \mathbf{S} \setminus \mathbf{T}$ , and then the value is  $-1$ .  $\square$

**Lemma C.3.** For any function  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ ,  $p \geq 3$  and  $\gamma > 0$ ,

$$\|T_{1-\gamma}f\|_p \leq \|T_{1-\gamma/2}f\|_2^{(2+\gamma)/p}.$$

*Proof.* It holds that

$$\|T_{1-\gamma}f\|_p = \mathbf{E}[|T_{1-\gamma}f|^p]^{1/p} \leq \mathbf{E}[|T_{1-\gamma}f|^{2+2\gamma}]^{1/p} = \|T_{1-\gamma}f\|_{2+2\gamma}^{(2+2\gamma)/p}.$$

Here, we use  $|T_{1-\gamma}f| \leq 1$ . Since  $1 - \gamma \leq \frac{1}{\sqrt{1+2\gamma}}$ , applying the Hypercontractive inequality, which appears in Appendix A

$$\|T_{1-\gamma}f\|_p \leq \|f\|_2^{(2+2\gamma)/p}.$$

Using this,

$$\|T_{1-\gamma}f\|_p \leq \|T_{1-\gamma/2}T_{1-\gamma/2}f\|_p \leq \|T_{1-\gamma/2}f\|_2^{(2+\gamma)/p}.$$

$\square$

*Proof of Lemma 4.6.* Let  $x = (x_1, x')$  where  $x' = (x_2, \dots, x_n)$ . We divide the Fourier expansion of  $f_S(x)$  according to its dependence to  $x_1$ , i.e.,  $f_S(x) = F_{S,\emptyset}(x') + x_1 F_{S,1}(x')$ . Then, we note that  $T_{1-\gamma}f_S(x) = T_{1-\gamma}F_{S,\emptyset}(x') + (1 - \gamma)x_1 T_{1-\gamma}F_{S,1}(x')$ .

Using Lemma C.2, the left hand side is equal to

$$\begin{aligned} \left[ \mathbf{E}_{\mathcal{H}'_{k,\delta}} \left[ \mathbf{E}_{\mathcal{H}_{k,\delta}} \left[ \prod_{S \in \mathbf{S}} T_{1-\gamma}f_S(\mathbf{x}) \right] - \mathbf{E}_{\mathcal{I}_{k,S_0,\delta}} \left[ \prod_{S \in \mathbf{S}} T_{1-\gamma}f_S(\mathbf{x}) \right] \right] \right] &= \left| (1 - \delta) \sum_{\mathbf{T} \in \mathfrak{X}_{\mathbf{S}}} (1 - \gamma)^{|\mathbf{T}|} \mathbf{E}_{\mathcal{H}'_{k,\delta}} \left[ \prod_{S \in \mathbf{T}} T_{1-\gamma}F_{S,1}(\mathbf{x}'_S) \right] \right| \\ &\leq (1 - \delta) \sum_{\mathbf{T} \in \mathfrak{X}_{\mathbf{S}}} (1 - \gamma)^{|\mathbf{T}|} \mathbf{E}_{\mathcal{H}'_{k,\delta}} \left[ \left| \prod_{S \in \mathbf{T}} T_{1-\gamma}F_{S,1}(\mathbf{x}'_S) \right| \right] \\ &\leq (1 - \delta) \sum_{\mathbf{T} \in \mathfrak{X}_{\mathbf{S}}} (1 - \gamma)^{|\mathbf{T}|} \prod_{S \in \mathbf{T}} \|T_{1-\gamma}F_{S,1}\|_{|\mathbf{T}|}. \quad (8) \end{aligned}$$

The last step uses the Hölder's inequality and the fact that the marginal of  $\mathcal{H}'_{k,\delta}$  on  $\mathbf{x}'_S$  is uniform for each  $S$ . Using Lemma C.3,

$$\prod_{S \in \mathbf{T}} \|T_{1-\gamma} F_{S,1}\|_{|\mathbf{T}|} \leq \prod_{S \in \mathbf{T}} \|T_{1-\gamma/2} F_{S,1}\|_2^{(2+\gamma)/|\mathbf{T}|}$$

Here,

$$\|T_{1-\gamma/2} F_{S,1}\|_2^2 = \sum_{Q \ni 1} (1-\gamma/2)^{2|Q|} \widehat{f}_S(Q)^2 \leq \text{Inf}_1(T_{1-\gamma/2} f_S)$$

Thus, (8) is upper bounded by

$$\sum_{\mathbf{T} \in \mathfrak{X}_{\mathbf{S}}} \prod_{S \in \mathbf{T}} \text{Inf}_1(T_{1-\gamma/2} f_S)^{(2+\gamma)/2|\mathbf{T}|}. \quad (9)$$

Since  $|\mathbf{T}| \geq 2$  for all  $\mathbf{T} \in \mathfrak{X}_{\mathbf{S}}$ ,  $\min_{S \in \mathbf{T}} \text{Inf}_1(T_{1-\gamma/2} f_S) \leq 2\tau$ . Then,

$$\begin{aligned} (9) &\leq \sum_{\mathbf{T} \in \mathfrak{X}_{\mathbf{S}}} \prod_{S \in \mathbf{T}} (2\tau)^{\gamma/2|\mathbf{T}|} \text{Inf}_1(T_{1-\gamma/2} f_S)^{1/|\mathbf{T}|} \\ &\leq (2\tau)^{\gamma/2|\mathbf{S}|} \sum_{\mathbf{T} \in \mathfrak{X}_{\mathbf{S}}} \sum_{S \in \mathbf{T}} \text{Inf}_1(T_{1-\gamma/2} f_S) \quad (\text{using the AM-GM inequality}) \\ &\leq (2\tau)^{\gamma/2|\mathbf{S}|} 2^{|\mathbf{S}|} \sum_{S \in \mathbf{S}} \text{Inf}_1(T_{1-\gamma/2} f_S) \end{aligned}$$

□

## D Proof of Lemma 4.7

**Lemma D.1** (Lemma 4.7 restated). *For any  $i \in [n]$  and  $S_0 \in \mathbf{S}$ , it holds that*

$$\rho(\mathcal{X}_{S_0}^i, \prod_{S \in \mathbf{S} \setminus \{S_0\}} \mathcal{X}_S^i; \mathcal{I}_{k,S_0,\delta}) \leq \sqrt{\delta}.$$

*Proof.* Let  $f : \mathcal{X}_{S_0}^i \rightarrow \mathbb{R}$  be any function with  $\mathbf{E}_{\mathcal{I}_{k,S_0,\delta}}[f] = 0, \mathbf{E}_{\mathcal{I}_{k,S_0,\delta}}[f^2] \leq 1$ . Also, let  $g : \prod_{S \in \mathbf{S} \setminus \{S_0\}} \mathcal{X}_S^i \rightarrow \mathbb{R}$  be any functions with  $\mathbf{E}_{\mathcal{I}_{k,S_0,\delta}}[g] = 0, \mathbf{E}_{\mathcal{I}_{k,S_0,\delta}}[g^2] \leq 1$ . From the latter property we get,

$$1 \geq \mathbf{E}_{\mathcal{I}_{k,S_0,\delta}}[g^2] = (1-\delta) \mathbf{E}_{\mathcal{I}_{k,S_0}}[g^2] + \delta \mathbf{E}_{\mathcal{N}_k}[g^2] \geq \delta \mathbf{E}_{\mathcal{N}_k}[g^2]. \quad (10)$$

We now observe that

$$\begin{aligned} \mathbf{E}_{\mathcal{I}_{k,S_0,\delta}}[fg] &= (1-\delta) \mathbf{E}_{\mathcal{I}_{k,S_0}}[fg] + \delta \mathbf{E}_{\mathcal{N}_k}[fg] \\ &\leq (1-\delta) \mathbf{E}_{\mathcal{I}_{k,S_0}}[f] \mathbf{E}_{\mathcal{I}_{k,S_0}}[g] + \delta \sqrt{\mathbf{E}_{\mathcal{N}_k}[f^2]} \sqrt{\mathbf{E}_{\mathcal{N}_k}[g^2]}. \end{aligned}$$

We have equality for the first term because the marginals of  $\mathcal{I}_{k,S_0,\delta}$  on  $\mathcal{X}_{S_0}^i$  and  $\prod_{S \in \mathbf{S} \setminus \{S_0\}} \mathcal{X}_S^i$  are independent. We used Cauchy-Schwarz on the second term. Since  $\mathbf{E}_{\mathcal{I}_{k,S_0}}[f] = 0$  and From (10),

$$\delta \sqrt{\mathbf{E}_{\mathcal{N}_k}[f^2]} \sqrt{\mathbf{E}_{\mathcal{N}_k}[g^2]} \leq \delta \sqrt{1} \sqrt{1/\delta}.$$

Hence,

$$\mathbf{E}_{\mathcal{I}_{k,S_0,\delta}} [fg] \leq \sqrt{\delta}.$$

This completes the proof.  $\square$

## E Correlation lemmas

**Lemma E.1.** For any  $S' \in 2_{\geq 1}^{[k]}$ ,

$$\rho\left(\prod_{S \in 2_{\geq 1}^{[k]} \setminus \{S'\}} \left(\prod_{i=1}^n \mathcal{X}_S^i\right), \prod_{i=1}^n \mathcal{X}_{S'}^i; \mathcal{H}_{k,\delta}^n\right) \leq 1 - \alpha^2/2,$$

where  $\alpha = 2\delta/(2^{k+1} - 2)$

*Proof.* We use Lemma 2.9 of [31] to show that

$$\rho\left(\prod_{S \in 2_{\geq 1}^{[k]} \setminus \{S'\}} \mathcal{X}_S^i, \mathcal{X}_{S'}^i; \mathcal{H}_{k,\delta}\right) \leq 1 - \alpha^2/2 \quad (11)$$

for each  $i$ . First observe that for any  $(a, b) \in \prod_{S \in 2_{\geq 1}^{[k]} \setminus \{S'\}} \mathcal{X}_S^i \times \mathcal{X}_{S'}^i$ , either  $b = 1$  or  $b = -1$  holds.

Next observe that vertices  $(1), (-1) \in \mathcal{X}_{S'}^i$  are connected because there are edges  $(-1, 1), (-1, -1)$  due to the definition of the distribution  $\mathcal{N}_k$ . Thus we can guarantee that the bipartite graph is connected. Easy calculation of  $p, q, r$  shows that the probability of the smallest atom is  $\alpha = 2\delta/(2^{k+1} - 2)$ .

The correlation bound (11), combined with Proposition 2.13 of [31], concludes the proof of Lemma E.1.  $\square$