Exponentially Small Soundness for the Direct Product Z-Test*

Irit Dinur[1] and Inbal Livni Navon[2]

1 **Weizmann Institute of Science, Rehovot, Israel**
irit.dinur@weizmann.ac.il

2 **Weizmann Institute of Science, Rehovot, Israel**
inbal.livni@weizmann.ac.il

Abstract

Given a function $f : [N]^k \to [M]^k$, the Z-test is a three query test for checking if a function f is a direct product, namely if there are functions $g_1, \dots g_k : [N] \to [M]$ such that $f(x_1, \dots, x_k) = (g_1(x_1), \dots g_k(x_k))$ for every input $x \in [N]^k$.

This test was introduced by Impagliazzo et. al. (SICOMP 2012), who showed that if the test passes with probability $\epsilon > \exp(-\sqrt{k})$ then f is $\Omega(\epsilon)$ close to a direct product function in some precise sense. It remained an open question whether the soundness of this test can be pushed all the way down to $\exp(-k)$ (which would be optimal). This is our main result: we show that whenever f passes the Z test with probability $\epsilon > \exp(-k)$, there must be a global reason for this: namely, f must be close to a product function on some $\Omega(\epsilon)$ fraction of its domain.

Towards proving our result we analyze the related (two-query) V-test, and prove a "restricted global structure" theorem for it. Such theorems were also proven in previous works on direct product testing in the small soundness regime. The most recent work, by Dinur and Steurer (CCC 2014), analyzed the V test in the exponentially small soundness regime. We strengthen their conclusion of that theorem by moving from an "in expectation" statement to a stronger "concentration of measure" type of statement, which we prove using hyper-contractivity. This stronger statement allows us to proceed to analyze the Z test.

We analyze two variants of direct product tests. One for functions on ordered tuples, as above, and another for functions on sets, $f : \binom{[N]}{k} \to [M]^k$. The work of Impagliazzo et. al was actually focused only on functions of the latter type, i.e. on sets. We prove exponentially small soundness for the Z-test for both variants. Although the two appear very similar, the analysis for tuples is more tricky and requires some additional ideas.

1998 ACM Subject Classification F.2.2 Nonnumerical Algorithms and Problems

Keywords and phrases Direct Product Testing, Property Testing, Agreement

Digital Object Identifier 10.4230/LIPIcs.CCC.2017.29

1 Introduction

A function $f : [N]^k \to [M]^k$ for $N, M, k \in \mathbb{N}$, is a *direct product function* if $f = (g_1, \dots g_k)$, for $g_i : [N] \to [M]$, i.e. the output of f on each coordinate depends on the input to this coordinate alone. Direct products appear in a variety of contexts in complexity, usually for hardness amplification. In PCPs it underlies the parallel repetition theorem [12] and implicitly appears in other forms of gap amplification, e.g. [4]. The specific task of testing

* Research supported in part by an ISF-UGC grant number 1399/14, and by BSF grant number 2014371.

32nd Computational Complexity Conference (CCC 2017).
Editor: Ryan O'Donnell; Article No. 29; pp. 29:1–29:50
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

direct products as an abstraction of a certain element of PCP constructions was introduced by [8].

The combinatorial question that underlies these works is the direct product testing question: given a function $f : [N]^k \to [M]^k$, is it a direct product function? The setting of interest here is where we query f in the *smallest number of inputs possible*, and decide if is it a direct product function or not.

The direct product testing question is a type of property testing question, yet it is not in the standard property testing parameter regime. In property testing we are generally interested in showing that functions that pass the test with high probability, for example 99%, are close to having the property.

In our case, we are interested in understanding the structure of functions that pass the test with small – but non-trivial – probability, e.g. 1%. The 1% regime is often more challenging than the 99% regime. It plays an important role in PCPs where one needs to prove a large gap. In such arguments one needs to be able to deduce non trivial structure even from a proof that passes a verification test with small probability, e.g. 1%.

There are very few families of tests for which 1% theorems are known. These include algebraic low degree tests and direct product tests. For low degree tests there has been a considerable amount of work in various regimes and in particular towards understanding the extent of the 1% theorems, see e.g. [13, 1, 3] and [2]. It is intriguing to understand more broadly for which tests such theorems can hold. Indeed, as far as we know, there are no other tests that exhibit such strong "structure vs. randomness" behavior, and direct product tests are natural candidates in which to study this question.

We remark that finding new settings where 1% theorems hold (including in particular derandomized direct products) can be potentially useful for constructing locally testable codes and stronger PCPs, see e.g. the recent works of [10, 6]. Towards this goal gaining a more comprehensive understanding of direct product tests, as well as developing tools for proving them, is a natural goal.

1.1 Our Main Result

The main question we study is: if $f : [N]^k \to [M]^k$ passes a certain natural test (Test 1 below) with non-negligible probability, how can f look like? We prove

▶ **Theorem 1** (Main Theorem – Global Structure). *For every $N, M > 1$, there exist small constants $c_1, c_2 > 0$ such that for every constant $\lambda > 0$ and large enough k, if $f : [N]^k \to [M]^k$ is a function that passes Test 1 with probability $\alpha_{Z(\frac{k}{10})}(f) = \epsilon \geq e^{-c_1\lambda^2 k}$, then there exist functions $(g_1, \dots g_k)$, $g_i : [N] \to [M]$ such that*

$$\Pr_{x\in[N]^k}\left[f(x) \overset{\lambda k}{\approx} (g_1(x_1)\dots g_k(x_k))\right] \geq c_2 \cdot \epsilon.$$

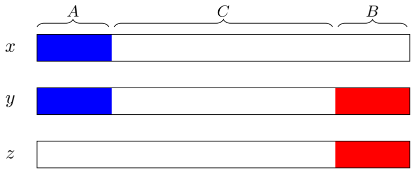
Where $\overset{\lambda k}{\approx}$ means that the strings are equal on all but at most λk coordinates.

The theorem is qualitatively tight with respect to several parameters: (i) Soundness, (i.e. the parameter ϵ), (ii) Approximate equality vs. exact equality (i.e. the parameter λ), (iii) Number of queries in the test. We discuss these next.

(i) Soundness

The soundness of the theorem is the smallest success probability in which the theorem is valid, in our case it is 2^{-ck} for some constant $c > 0$. This is tight up to the constant c, as can be seen by the example bellow.

1. Choose A, B, C to be a random partition of $[k]$, such that $|A| = |B| = t$.
2. Choose uniformly at random $x, y, z \in [N]^k$ such that $x_A = y_A$ and $y_B = z_B$.
3. Reject if $f(x)_A \neq f(y)_A$ or $f(z)_B \neq f(y)_B$, else accept.



Denote by $\alpha_{Z(t)}(f)$ the success probability of f on this test.

Test 1 “Z”-test with parameter t (3-query test).

▶ **Example 2** (Random function). Let $f : [N]^k \to \{0,1\}^k$ be a random function; i.e. for each $x \in [N]^k$ choose $f(x) \in \{0,1\}^k$ uniformly and independently. Two random strings in $\{0,1\}^t$ are equal with probability 2^{-t}, therefore $\alpha_{Z(t)}(f) = 2^{-2t}$, since the test performs two such checks. On the other hand, since f is random, it is not close to any direct product function.

We remark that every function $f : [N]^k \to \{0,1\}^k$ is at least 2^{-k} close to a direct product function [1], so this amount of correlation is meaningless. We conclude that in order to have direct product theorem that is not trivial, the minimal soundness has to be $2^{-c'k}$ for some constant $c' < 1$.

(ii) Approximate equality vs. exact equality

In the theorem, we prove that for $\Omega(\epsilon)$ of the inputs x: $f(x) \overset{\lambda k}{\approx} (g_1(x), \ldots, g_k(x))$. A priori, one could hope for a stronger conclusion in which $f(x) = (g_1(x), \ldots, g_k(x))$ for $\Omega(\epsilon)$ of the x's. However, Example 3 shows that for $t = \frac{k}{10}$, approximate equality is necessary.

▶ **Example 3** (Noisy direct product function). This example is from [5]. Let f be a direct product function, except that on each input x we “corrupt” $f(x)$ on λk random coordinates by changing $f(x)$ on these coordinates into random values. For $\lambda < \frac{1}{10}$, the probability that Test 1 on f missed all the corrupted coordinates is $2^{-\Omega(\lambda k)}$, in which case the test succeeds. Since we have changed $f(x)$ on λk coordinates into random values, no direct product function can approximate f on more than $(1 - \lambda)$ of the coordinates.

From this example we conclude that for f that passes Test 1 for $t = \frac{k}{10}$ with probability $e^{-\delta \lambda k}$, it is not possible to approximate f on more than $(1 - \lambda)$ of the coordinates. Further discussion and examples for different intersection sizes (i.e. t) are in Section 6.

(iii) Number of queries in the test

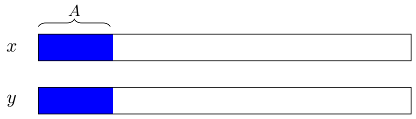
The absolute minimal number of queries for any direct product test is two. Indeed, there is a very natural 2-query test, Test 2.

Dinur and Goldenberg showed that it is not possible to have a direct product theorem with soundness lower than $\frac{1}{poly(k)}$ using the 2-query test [5].

▶ **Example 4** (Localized direct product functions). In this example we assume $N \gg k$. For every $b \in [N]$ we choose a random function $g_b : [N] \to [M]$ independently. For every input $x \in [N]^k$, we choose a random $i_x \in k$, set $b = x_i$ and set $f(x) = (g_b(x_1), \ldots, g_b(x_k))$.

[1] Consider the direct product function constructed incrementally by taking the most common value out of $\{0,1\}$ on each step.

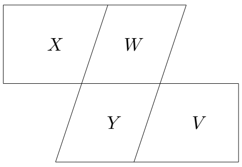
1. Choose $A \subset [k]$ of size t, uniformly at random.
2. Choose uniformly at random $x, y \in [N]^k$ such that $x_A = y_A$.
3. Accept if $f(x)_A = f(y)_A$.



Denote by $\alpha_{V(t)}(f)$ the success probability of f on this test.

Test 2 "V" test with parameter t (2-query test).

1. Choose random $V, W, X, Y \subset [N]$, such that $|W| = |V| = t$, $|X| = |Y| = k - t$ and $X \cap W = Y \cap W = Y \cap V = \emptyset$.
2. Reject if $f(X \cup W)_W \neq f(Y \cup W)_W$ or $f(Y \cup W)_Y \neq f(Y \cup V)_Y$, else accept.



Denote by $\alpha_{Z_{set}(t)}(f)$ the success probability of f on this test.

Test 3 "Z" test for functions over sets, with parameter t (3-queries).

The function f satisfies $\alpha_{V(t)}(f) \geq \frac{1}{k} \cdot \frac{t}{k}$; indeed, for x, y and A chosen in the test, if $i_x = i_y$ and $i_x \in A$, then the test will pass. The probability that $i_x = i_y$ is $\frac{1}{k}$, and the probability that $i_x \in A$ is $\frac{t}{k}$.

For $N \gg k$, the function f is very far from direct product, since it is made up from N different direct product functions. Each piece consisting of roughly $1/N$ fraction of the domain $[N]^k$.

For every t, the function described in the example satisfies $\alpha_{V(t)}(f) \geq \frac{1}{k^2}$, yet there is no direct product function that approximates f when $N \gg k$. In [5] the conclusion from Example 4 was that $1/\operatorname{poly}(k)$ is the limit for small soundness for direct product tests. However, [9] showed that by adding just one more query, this limitation goes away. They introduced a 3-query test, similar to Test 1, and proved a direct product theorem for all $\epsilon > 2^{-k^{\beta}}$ for some constant $\beta \ \leq 1/2$.

Direct product test for functions over sets

Some of the previous direct product works, such as [9] were proven in a slightly different setting, where the function tested is $f : \binom{[N]}{k} \to [M]^k$, i.e. the input to the function f is an unordered set $S \subset [N]$ of k elements. In this work, we also prove a direct product testing theorem for this setting, Test 3 is the analog of Test 1 for functions over sets. In Test 3 (see figure), we pick disjoint sets W, X, Y, Z such that $X \cap W = Y \cap W = Y \cap V = \emptyset$ so that $|X \cup W| = |Y \cup W| = |Y \cup V| = k$ and they can be inputs to the function f.

▶ **Theorem 5** (Global Structure for Sets). *There exist a small constant $c > 0$, such that for every constant $\lambda > 0$, large enough $k \in \mathbb{N}$ and $N > k^2 e^{10c\lambda k}$, if the function $f : \binom{[N]}{k} \to [M]^k$ passes Test 3 with probability $\alpha_{Z_{set}(\frac{k}{10})}(f) = \epsilon > e^{-c\lambda k}$, then there exist a function $g : [N] \to [M]$ such that*

$$\Pr_S\left[f(S) \overset{\lambda k}{\approx} g(S)\right] \geq \epsilon - 4\epsilon^2 .$$

Notice that the probability bound of $\epsilon - 4\epsilon^2$ is better than $\Omega(\epsilon)$, and it is tight as demonstrated by the function f which is a hybrid of $\frac{1}{\epsilon}$ different direct product functions on equals parts of the inputs. f passes Test 3 with probability ϵ, and every direct product function is close to f only on ϵ fraction of the inputs.

We remark that the two theorems are not the same. In Theorem 1, there are k different functions $g_1, \dots, g_k : [N] \to [M]$ whereas in Theorem 5 there is a single one. Furthermore, Theorem 1 holds for any $N, M \in \mathbb{N}$ and large enough k, and Theorem 5 (and other such direct product theorems) only holds for $N \gg k$. The proofs of the theorems are also different, which is discussed later in the introduction.

1.2 Restricted Global Structure

Our proof has two main parts, similar to the structure of the proof of [5, 9]. In the first part, we analyze only Test 2 (which is on tuples) and prove a restricted global structure theorem for it, Theorem 6 below (this was called local structure in [9, 7]). The term "restricted global structure" refers to when we restrict the domain to small (but not trivial) pieces, and show that f is close to a product function on each piece separately. This is the structure of the function in Example 4.

More explicitly, for every $A \in [k]$ of size $\frac{k}{10}$, $r \in [N]^A$ and $\gamma \in [M]^A$, a *restriction* is a triple $\tau = (A, r, \gamma)$. The choice of $t = \frac{k}{10}$ in Theorem 1 is somewhat arbitrary, the theorem can be proven with $t = ck$ for $c < \frac{1}{2}$. The restriction corresponds to the set of inputs

$$\mathcal{V}_\tau = \{w \in [N]^{[k]\setminus A} | f(r, w)_A = \gamma\}.$$

Our next theorem shows that for many restrictions τ there exist a direct product function that is close to f on $\mathcal{V}_\tau$.

▶ **Theorem 6** (Restricted Global Structure – informal). *Let $f : [N]^k \to [M]^k$ be a function that passes Test 2 with probability $\alpha_{V(\frac{k}{10})}(f) = \epsilon > e^{-\delta\lambda k}$, then there exist a natural distribution over restrictions $\tau = (A, r, \gamma)$ such that with probability $\Omega(\epsilon)$, there exist functions $(g_1^\tau, \dots g_{\frac{9k}{10}}^\tau), g_i^\tau : [N] \to [M]$ such that,*

$$\Pr_{w \in [N]^{[k]\setminus A}} \left[f(r, w)_{[k]\setminus A} \overset{\lambda k}{\approx} (g_1^\tau(w_1), \dots g_{\frac{9k}{10}}^\tau(w_{\frac{9k}{10}})) \,\middle|\, w \in \mathcal{V}_\tau \right] \geq 1 - \epsilon^2. \tag{1}$$

Where the distribution over τ is the test distribution, namely choose $A \subset [k], x \in [N]^k$ uniformly, and set $\tau = (A, x_A, f(x)_A)$.

A similar theorem was proven in [9] but only for soundness (i.e. ϵ) at least $\exp(-k^\beta)$ for a constant $\beta \leq 1/2$. This was strengthened to soundness $\exp(-\Omega(k))$ in [7]. Our Theorem 6 improves on the conclusion of [7] . In [7] the probability in (1) was shown to be at least $1 - O(\lambda)$ (recall that λ is a constant), whereas we show it is exponentially close to 1 (when ϵ is that small). This difference may seem minor but in fact it is what prevented [7] from deriving global structure via a three query test (i.e. moving from the V test to the Z test). When we try to move from restricted global structure to global structure, the consistency inside each restriction needs to be very high for the probabilistic arguments to work, as we try to explain below.

The restricted global structure gives us a direct product function that approximates f only on a restricted subset of the inputs. In the proof of the global structure, we use the third query to show that there exists a global function. A key step in the proof of the global structure is to show that for many restrictions τ, the function g^τ is close to f on a much larger subsets of inputs. This is done, intuitively, by claiming that if $f(x)_A = f(y)_A$, then with high probability $f(y) \approx g^\tau(y)$ for $\tau = (A, x_A, f(x)_A)$. Since B is a random set and $f(z)_B = f(y)_B$, then $f(z), g^{\tau}(z)$ are also close. This claim only holds if the success probability on (1) is more than $1-\epsilon$, else it is possible that all the success probability of the test comes from f such that $f(x)_A = f(y)_A$, but $f(y), g^{\tau}(y)$ are far.

1.3 Technical Contribution

In terms of technical contribution our proof consists of two new components.

Domain extension

Our first contribution a new *domain extension* step that facilitates the proof of the restricted global structure. The restricted global structure shows that with probability $\Omega(\epsilon)$, the function f is close to a direct product on the restricted domain $\mathcal{V}_\tau$. A natural way to show that a function is close to a direct product function is to define a direct product function by majority value. However, this method fails when the agreement guaranteed for f is small, as in our case.

This is usually resolved by moving to a restricted domain in which the agreement is much higher, and to define majority there. The first part of our proof is to show that with probability $\Omega(\epsilon)$, over restrictions $\tau = (A, r, \gamma) \sim \mathcal{D}$, the set $\mathcal{V}_\tau$ satisfies the following two properties:

1. Its density is at least $\frac{\epsilon}{2}$.
2. f has very high agreement in $\mathcal{V}_\tau$, informally it means that taking a random pair $w, v \in \mathcal{V}_\tau$ such that $w_J = v_J$, results in agreeing answers, i.e. $f(r,w)_J \approx f(r,v))_J$, with probability greater than $1-\epsilon^{120}$.

We call such restrictions excellent, following [9].

We show that for every excellent restriction $\mathcal{V}_\tau$, the restriction h_τ of f to $\mathcal{V}_\tau$, defined by $h_\tau(w) = f(r,w)_{[k]\setminus A}$, is close to a direct product function. The function h_τ has high agreement, which is good for defining majority, but unfortunately the low density of $\mathcal{V}_\tau$, which can be as low as $\frac{\epsilon}{2}$, which is exponentially small, is where the techniques used in [9] break down. In order to prove that h_τ is close to a direct product function, we use a local averaging operator to *extend* the domain from $\mathcal{V}_\tau$ to $[N]^{[k]\setminus A}$.

The local averaging operator $\mathcal{P}_{\frac{3}{4}}$ is the majority of a $\frac{3}{4}$-correlated neighborhood,

$$\forall w \in [N]^{[k]\setminus A}, i \notin A \qquad \mathcal{P}_{\frac{3}{4}} h_\tau(w)_i = \underset{v \in \mathcal{N}_{\frac{3}{4}}(w), v \in \mathcal{V}_\tau, v_i = w_i}{\text{Plurality}} \{h_\tau(v)_i\},$$

where $v \in \mathcal{N}_{\frac{3}{4}}(w)$ means that v is $\frac{3}{4}$-correlated with w, i.e. we change each coordinate of w with probability $\frac{1}{4}$ independently. The new function, $\mathcal{P}_{\frac{3}{4}} h_\tau$ is defined over all $[N]^{[k]\setminus A}$, unlike h_τ which is defined only on $\mathcal{V}_\tau$.

In order to use $\mathcal{P}_{\frac{3}{4}} h_\tau$ for showing that h_τ is close to a direct product function, we show two things:

1. $\mathcal{P}_{\frac{3}{4}} h_\tau$ and h_τ are similar on $\mathcal{V}_\tau$.
2. $\mathcal{P}_{\frac{3}{4}} h_\tau$ has high agreement, taking a random pair $w, v \in [N]^{[k]\setminus A}$ such that $w_J = v_J$, results in agreeing answers, $\mathcal{P}_{\frac{3}{4}} h_\tau(w)_J \approx \mathcal{P}_{\frac{3}{4}} h_\tau(v)_J$ with probability $1-\epsilon^6$.

To prove that $\mathcal{P}_{\frac{3}{4}} h_\tau$ has high agreement we use reverse hypercontractivity to show that only a few $w \in [N]^{[k]\setminus A}$ have sparse neighborhood (with density less than ϵ^{50}), and use the very high agreement of h_τ.

Lastly, we define a direct product function g_τ by taking the plurality over $\mathcal{P}_{\frac{3}{4}} h_\tau$, and show that it is close to h_τ.

Direct product testing in a dense regime

A second new element comes when stitching the many localized functions into one global direct product function, by using the third query.

We prove two global structure theorems, Theorem 1 for functions on tuples $f : [N]^k \to [M]^k$ and Theorem 5 for functions on sets $f : \binom{[N]}{k} \to [M]^k$.

When we work with f that is defined over sets, we can directly follow the approach of [9] to complete the proof. However, when working with f defined on tuples we reach a combinatorial question that itself resembles a direct product testing question, but in a different (dense) regime. Luckily, the fact that this question is in a dense regime makes it easier to solve, and this leads to our global structure theorem for tuples. An outline of the global structure proofs appears in Section 5.1.

1.4 Agreement Tests and Direct Product Tests

The question of direct product testing fits into a more general family of tests called agreement tests. We next describe this setting formally and explain how direct product tests fit into this framework.

Agreement tests

In all efficient PCPs we break a proof into small overlapping pieces, use a relatively inefficient PCPs (i.e. PCPs that incur a large blowup) to encode each small piece, and then through an *agreement test* put the pieces back together. The agreement test is needed because given the collection of pieces, there is no guarantee that the different pieces come from the same underlying global proof, i.e. that the proofs of each piece can be "put back together again". The PCP system needs to ensure this through *agreement testing*: we take two pieces that have some overlap, and check that they agree.

This situation can be formulated as an agreement testing question as follows. Let V be a ground set, $|V| = N$, and let H be a collection of subsets of V, namely, a set of hyperedges. Let $[M]$ be a finite set of colors, where it is sufficient to think of $M = 2$.

A *local assignment* is a collection $a = \{a_s\}$ of local colorings $a_s : s \to [M]$, one per subset $s \in H$. A local assignment is called *global* if there is a global coloring $g : V \to [M]$ such that

$$\forall s \in H, \qquad a_s \equiv g|_s.$$

An *agreement check* for a pair of subsets s_1, s_2 checks whether their local functions agree, denoted $a_{s_1} \sim a_{s_2}$. Formally,

$$a_{s_1} \sim a_{s_2} \qquad \Leftrightarrow \qquad \forall x \in s_1 \cap s_2, \quad a_{s_1}(x) = a_{s_2}(x).$$

A local assignment that is global passes all agreement checks. The converse is also true: a local assignment that passes *all* agreement checks must be global.

An *agreement test* is specified by giving a distribution $\mathcal{D}$ over pairs (or triples) of subsets s_1, s_2. We define the agreement of a local assignment to be the probability of agreement,

$$agree_{\mathcal{D}}(a) = \Pr_{s_1,s_2 \sim D}\left[a_{s_1} \sim a_{s_2}\right].$$

An agreement theorem shows that if a is a local assignment with $agree_{\mathcal{D}}(a) > \epsilon$ then a is somewhat close to a global assignment. Agreement theorems can be studied for any hypergraph and in this work we prove such theorems for two specific hypergraphs: the k-uniform complete hypergraph, and the k-uniform k-partite complete hypergraph.

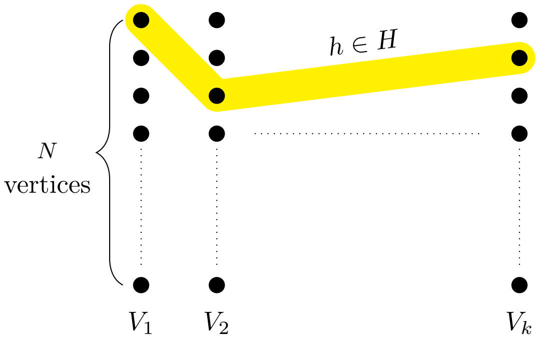


Figure 1 Complete k-uniform k-partite graph.

Relation to direct product testing

Theorem 1 is equivalent to an agreement theorem on the *complete k-uniform k-partite hypergraph* (see Figure 1). Let $G = (V = V_1, \dots V_k, H)$ be the complete k-partite hypergraph with $|V_i| = N$ for $i \in [k]$, and

$$H = \{(v_1, \dots v_k) \mid \ \forall i \in [k], v_i \in V_i\} .$$

There is a bijection between H and $[N]^k$. We shall interpret $f(x_1, \dots, x_k)$ as a local coloring of the vertices $x_1, \dots, x_k$. In this way, we have the following equivalence

$$f : [N]^k \to [M]^k \qquad \Longleftrightarrow \qquad a = \{a_x\}_{x \in H}.$$

Moreover, local assignments which are global, i.e. a such that $a_x = g|_x$ for some global coloring $g : V_1 \cup \dots \cup V_k \to [M]$, correspond exactly to functions f which are direct products, $f = (g_1, \dots, g_k)$ where $g_i = g|_{V_i}$,

$$f = (g_1, \dots, g_k) \qquad \Longleftrightarrow \qquad a \text{ is global.}$$

Finally, Test 2 can be described as taking 2 hyperedges that intersect on t vertices, and check if their local functions agree on the intersection. Similarly, Test 1 can be described as picking three hyperedges, $h_1, h_2, h_3 \in H$ such that h_1, h_2 intersect on t vertices, and h_2, h_3 intersect on a disjoint set of t vertices, and checking agreement.

Our main theorem, Theorem 1, is equivalent to an agreement theorem showing that if a local assignment a passes a certain 3-query agreement test with non-negligible probability, then there exists a global assignment $g : V \to [M]$ with which it agrees non-negligibly.

The k-uniform complete hypergraph (it is non-partite, in contrast to the above), is related to Theorem 5. In this hypergraph the vertex set is $[N]$ and there is a hyperedge for every possible k-element subset of $[N]$. Now we have a similar equivalence between local assignments and functions over sets, i.e. functions where the input is a set $S \subset [N]$ of size k,

$$f : \binom{[N]}{k} \to [M]^k \qquad \Longleftrightarrow \qquad a = \{a_s\}_{s \in \binom{[N]}{k}}.$$

An agreement theorem for this hypergraph is equivalent to Theorem 5, in which f is defined not on “tuples” $[N]^k$ but on “sets” $\binom{[N]}{k}$. A global assignment a or this graph is equivalent to a direct product function over sets, i.e. $f = g : [N] \to [M]$.

1.5 Organization of the Paper

Section 2 contains preliminary notations and definitions. In Section 3 we prove the restricted global structure, Theorem 6. Section 4 is dedicated to the global structure for functions on sets. We show how to deduce a variant of Theorem 6 for sets rather than tuples and then prove the global structure theorem for sets, Theorem 5. In Section 5 we prove the global structure theorem for tuples, Theorem 1. Lastly, in Section 6 we discuss lower bounds for various 3-query direct product tests that were not presented in the introduction.

2 Preliminaries

▶ **Definition 7.** For each two strings $x, y \in [N]^k$ we say that:

1. $x \overset{t}{\approx} y$ if x, y differ in at most t coordinates.
2. $x \overset{t}{\not\approx} y$ if x, y differ in more than t coordinates.

▶ **Definition 8** (Plurality). The plurality of a function f on a distribution $\mathcal{D}$ is its most frequent value

$$\underset{x \sim \mathcal{D}}{\text{Plurality}}(f(x)) = \arg\max_{\beta} \left\{ \Pr_{x \in \mathcal{D}} [f(x) = \beta] \right\}$$

For a set $A \subset [k]$ we denote by $\bar{A}$ the set $[k] \setminus A$.

▶ **Fact 9** (Chernoff bound). *Let $X_1, \dots X_k$ be independent random variables in $\{0,1\}$, let $X = \sum_{i=1}^{k}$, and denote $\mu = \mathbb{E}[X]$, then for every $\delta \in (0,1)$,*

$$\Pr_{X_1,\dots X_k} [X \le (1-\delta)\nu] \le e^{-\frac{\delta^2 \mu^2}{2}},$$

and for every $\delta \in (0,1]$

$$\Pr_{X_1,\dots X_k} [X \ge (1+\delta)\nu] \le e^{-\frac{\delta^2 \mu^2}{3}}.$$

▶ **Corollary 10.** *Let k be a large integer, and let $A \subseteq [k]$ be the set generated by inserting each $i \in [k]$ into A with probability ρ. For every constant $c \in (0,1)$*

$$\Pr_A [|A| \le c\rho k] \le e^{-\frac{(1-c)^2}{2}\rho k},$$

and for every $c' \in [1,2]$,

$$\Pr_A [|A| \ge c'\rho k] \le e^{-\frac{(c'-1)^2}{3}\rho k}.$$

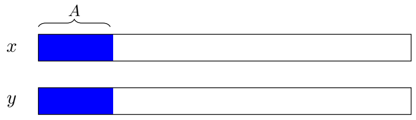
▶ **Claim 11** (Chernoff bound for fixed size subsets). *Let $k \in \mathbb{N}$ be a large integer, $D \subset [k]$ be a fixed subset of size at most $\frac{k}{3}$. Let A be a random subset of size exactly $\frac{k}{10}$, then*

$$\Pr_A \left[|A \cap D| \ge \frac{1}{5}|D| \right] \le e^{-\frac{1}{320}|D|} \tag{2}$$

If $|D| \le \frac{1}{30}k$ then

$$\Pr_A \left[|A \cap D| \le \frac{1}{20}|D| \right] \le e^{-\frac{1}{60}|D|} \tag{3}$$

1. Choose $A \subset [k]$ of size t, uniformly at random.
2. Choose uniformly at random $x, y \in [N]^k$ such that $x_A = y_A$.
3. Accept if $f(x)_A = f(y)_A$.



Denote by $\alpha_{V(t)}(f)$ the success probability of f on this test.

Test 2 "V" test with parameter t (2-query test).

The proof appears in Appendix A.

In our proof we also need Chernoff bound for non-binary random variables.

▶ **Fact 12** (Non-binary Chernoff bound). *Let $X_1, \dots X_k$ be independent random variables in $[0,1]$, let $X = \sum_{i=1}^{k} X_i$, and denote $\mu = \mathbb{E}[X]$ then,*

$$\Pr_{X_1,\dots X_k}[|X - \mu| > t] \leq 2e^{-t^2 k},$$

2.1 Reverse Hypercontractivity

▶ **Definition 13** (ρ-correlated distribution). For each string $y \in [N]^k$ and constant $\rho \in (0,1)$, the ρ correlated distribution from y will be denoted by $(x, J) \in \mathcal{N}_\rho(y)$. For each $i \in [k]$ independently, $i \in J$ with probability ρ, and x is chosen such that $x_J = y_J$, and the rest is uniform.

We quote Proposition 9.2 from [11]:

▶ **Claim 14.** *Let $A, B \subseteq [N]^k$ of sizes $\Pr_{w \in [N]^k}[w \in A] = e^{-\frac{a^2}{2}}$ and $\Pr_{w \in [N]^k}[w \in B] = e^{-\frac{b^2}{2}}$, then*

$$\Pr_{x \in [N]^k, y \in \mathcal{N}_\rho(x)}[x \in A, y \in B] \geq e^{-\frac{(2-\rho)(a^2+b^2)}{4(1-\rho)} - \frac{\rho ab}{2(1-\rho)}}.$$

By changing notations and simplifying, we get the following corollary.

▶ **Corollary 15.** *For $|A| \geq |B|$,*

$$\Pr_{x \in [N]^k, y \in \mathcal{N}_\rho(x)}[x \in A, y \in B] \geq \Pr_{x \in [N]^k}[x \in A]^{1+\frac{\rho}{2(1-\rho)}} \Pr_{x \in [N]^k}[x \in B]^{1+\frac{3\rho}{2(1-\rho)}}.$$

▶ **Claim 16.** *Let $G \subset [N]^k$ be a set of measure ν, then for any $\eta \in (0,1)$ the set $L = \left\{ w \in [N]^k \,\middle|\, \Pr_{(v,J) \in \mathcal{N}_{\frac{3}{4}}(w)}[v \in G] \leq \eta \right\}$ has a measure less than $\nu^{-\frac{11}{9}} \eta^{\frac{2}{9}}$.*

Both proofs appears in Appendix A.

3 Restricted Global Structure

Let $f : [N]^k \to [M]^k$ be such that $\alpha_{V(\frac{k}{10})}(f) = \epsilon \geq e^{-c\lambda k}$, i.e. the success probability of f on Test 2 equals ϵ. To make the reading easy, we write again Test 2 from the introduction.

We show in this section that $\alpha_{V(\frac{k}{10})}(f) = \epsilon$ already implies that f is somewhat structured, namely there are restrictions of the domain $\mathcal{V}_\tau \subset [N]^k$ such that on these restrictions f is roughly a product function.

Recalling the definition from the introduction, we define a *restriction* to be a triple $\tau = (A, r, \gamma)$, for $A \subset [k]$, $r \in [N]^A$ and $\gamma \in [M]^A$. In this section denote by $k' = \frac{9k}{10}$, and recall that $\bar{A} = [k] \setminus A$.

▶ **Definition 17** (Consistent strings). For each restriction $\tau = (A, r, \gamma)$, a string $w \in [N]^{\bar{A}}$ is consistent with τ if $f(r, w)_A = \gamma$. For every τ, let $\mathcal{V}_\tau$ be the set of consistent strings,

$$\mathcal{V}_\tau = \left\{ w \in [N]^{\bar{A}} \,\middle|\, f(r, w)_A = \gamma \right\}.$$

▶ **Definition 18** (Distribution of Restrictions). Let $\mathcal{D}$ be the following distribution over restrictions τ. Pick a uniform set $A \subset [k]$ of size $\frac{k}{10}$, pick a uniform $x \in [N]^k$ and set $r = x_A$ and $\gamma = f(x)_A$.

Note that the distribution $\mathcal{D}$ depends on the function f.

We define good restriction in an analogous way to the definitions of [9].

▶ **Definition 19** (Good restriction). A restriction $\tau = (A, r, \gamma)$ is *good*, if $\Pr_{w \in [N]^{\bar{A}}}[w \in \mathcal{V}_\tau] \geq \frac{\epsilon}{2}$.

▶ **Definition 20** (DP restriction). A restriction $\tau = (A, r, \gamma)$ is a *DP restriction* if it is good, and if there exist functions $(g_1^\tau, \ldots g_{k'}^\tau), g_i^\tau : [N] \to [M]$ such that

$$\Pr_{w \in [N]^{\bar{A}}} \left[f(r, w)_{\bar{A}} \overset{\lambda k}{\not\approx} (g_1^\tau(w_1), \ldots g_{k'}^\tau(w_{k'})) \,\middle|\, w \in \mathcal{V}_\tau \right] \leq \epsilon^2.$$

The main theorem of this section shows that (a) a non-negligible fraction of restrictions are good, and that (b) almost all good restrictions are DP restrictions.

▶ **Theorem 21** (Restricted Global Structure, restated). *There exist a small constant $\delta > 0$, such that for every constant $\lambda > 0$ and large enough $k \in \mathbb{N}$ the following holds. For every function $f : [N]^k \to [M]^k$, if $\alpha_{V(\frac{k}{10})}(f) = \epsilon > e^{-\delta \lambda k}$, then with probability at least $\frac{\epsilon}{2}$, $\tau \sim \mathcal{D}$ is good, and with probability at least $1 - \epsilon^2$ over the good restrictions, τ is a DP restriction. Namely, τ is such that there exist functions $(g_1^\tau, \ldots g_{k'}^\tau), g_i^\tau : [N] \to [M]$ such that*

$$\Pr_{w \in [N]^{\bar{A}}} \left[f(r, w)_{\bar{A}} \overset{\lambda k}{\not\approx} (g_1^\tau(w_1), \ldots g_{k'}^\tau(w_{k'})) \,\middle|\, w \in \mathcal{V}_\tau \right] \leq \epsilon^2.$$

A similar theorem was proven in [7] under the name "local structure". Under the same assumptions [7] showed that f must be close to a product function for many restrictions $\mathcal{V}_\tau$ of the domain. However the closeness was considerably weaker: unlike in our definition of a *DP restriction*, in [7] even in the restricted part of the domain, $\mathcal{V}_\tau \subset [N]^k$, there could be a (small) constant fraction of the inputs on which f differs from the global product function g^τ. In contrast, we only allow an ϵ^2 fraction of disagreeing inputs. As explained in the introduction, in order to extend the restricted global structure into a global one, the set of disagreeing inputs in $\mathcal{V}_\tau$ has to be smaller than ϵ.

3.1 Proof of Theorem 21

In this section we prove Theorem 21, we start by writing a few definitions and lemmas that are used in the proof, and give an intuition for the proof of each lemma. We defer the proofs of these lemmas to the next sections.

The distribution $\mathcal{D}$ over τ is related to the distribution of Test 2. The test can also be written as choose $\tau = (A, r, \gamma) \sim \mathcal{D}$, $w \in [N]^{\bar{A}}$ and accept iff $f(r, w)_A = \gamma$. Therefore, if the function f passes Test 2 with probability ϵ, by a simple averaging argument

$$\Pr_{\tau \sim \mathcal{D}}[\tau \text{ is good}] \geq \frac{\epsilon}{2}. \tag{4}$$

For each τ we define the function h_τ, which is a restriction of f to $\mathcal{V}_\tau$.

▶ **Definition 22.** For each restriction $\tau = (A, r, \gamma)$, let $h_\tau : \mathcal{V}_\tau \to [M]^{\frac{9k}{10}}$ be the function,

$$h_\tau(w) = f(r, w)_{\bar{A}}.$$

We define excellent restriction, in an analogous way to [9],

▶ **Definition 23** (Excellent restriction). Fix a constant $\alpha = \frac{1}{1600}\lambda$, a restriction $\tau = (A, r, \gamma)$ is excellent, if:

1. τ is good.
2. For every $\rho \in \left\{ \frac{a}{b} \;\middle|\; a, b \in \mathbb{N}, a < b \leq k \right\}$, if we pick $w \in [N]^{\bar{A}}$ and $(v, J) \in \mathcal{N}_\rho(w)$ then,

$$\Pr_{w,(v,J)}\left[w, v \in \mathcal{V}_\tau,\ h_\tau(w)_J \overset{\alpha k}{\not\approx} h_\tau(v)_J\right] \leq \left(\frac{9}{10}\right)^{\frac{1}{2}\alpha k}. \tag{5}$$

Note that (5) holds trivially when $\rho < \alpha$, because with high probability $|J| \approx \rho k < \alpha k$, in which it is not possible that $h(w)_J, h(v)_J$ differs in more than αk coordinates. For an excellent τ, the set $\mathcal{V}_\tau$ is of measure at least $\frac{\epsilon}{2}$, and the function f is consistent on $\mathcal{V}_\tau$.

We assume that the constant δ is small enough to satisfy $\left(\frac{9}{10}\right)^{\frac{1}{2}\alpha k} < \epsilon^{120} = e^{-120\delta\lambda k}$, and $\epsilon^{120} > e^{-\frac{\alpha k}{43000}}$.

▶ **Lemma 26.** *For every $\rho \in (0, 1)$, let $\tau = (A, r, \gamma) \sim \mathcal{D}$, let $w \in [N]^{\bar{A}}$ be uniform, and let $(v, J) \in \mathcal{N}_\rho(w)$, then*

$$\Pr_{\tau,w,(v,J)}\left[w, v \in \mathcal{V}_\tau,\ h_\tau(w)_J \overset{\alpha k}{\not\approx} h_\tau(v)_J\right] \leq \left(\frac{9}{10}\right)^{\alpha k}.$$

The proof appears on Section 3.2, the main idea in the proof is that the probability of $w, v \in \mathcal{V}_\tau, h(w)_J \overset{\alpha k}{\not\approx} h(v)_J$ is low when averaging over τ as well. From the definition of h_τ, this is equivalent to $f(r, w)_A = f(r, v)_A = \gamma$ and $f(r, w)_J \overset{\alpha k}{\not\approx} f(r, v)_J$. When r, w, v, A, J are all random, the probability for a uniform A, J to be such that $f(r, w), f(r, v)$ are equal on A but far on J is very small.

▶ **Corollary 24.** *A good $\tau \sim \mathcal{D}$ is excellent with probability larger than $1 - \epsilon^2$.*

Proof. Let $\mu = \left(\frac{9}{10}\right)^{\frac{1}{2}\alpha k}$, and denote by $E(\tau, w, v, J)$ the event of $w, v \in \mathcal{V}_\tau$, $h_\tau(w)_J \overset{\alpha k}{\not\approx} h_\tau(v)_J$. Lemma 26 in these notations is: for every $\rho \in (0, 1)$, $\Pr_{\tau\sim\mathcal{D},w,(v,J)\in\mathcal{N}_\rho(w)}[E] \leq \mu^2$.

For every τ that is good but not excellent, exist $\rho \in \left\{ \frac{a}{b} \;\middle|\; a, b \in \mathbb{N}, a < b \leq k \right\}$ such that,

$$\Pr_{w,(v,J)\in\mathcal{N}_\rho(w)}[E] > \mu.$$

In this case we say that τ is bad for ρ.

Assume towards contradiction that $\Pr_{\tau\sim\mathcal{D}}[\tau \text{ is good but not excellent}] > \epsilon^4$. The set $\left\{ \frac{a}{b} \;\middle|\; a, b \in \mathbb{N}, a < b \leq k \right\}$ contains less than k^2 elements, so there exists ρ in this set such that

$$\Pr_{\tau\sim\mathcal{D}}[\tau \text{ is bad for } \rho] \geq \frac{\epsilon^4}{k^2}.$$

For this ρ,

$$\Pr_{\tau\sim\mathcal{D},w,(v,J)\in\mathcal{N}_\rho(w)}[E] \geq \Pr_{\tau\sim\mathcal{D}}[\tau \text{ is bad for } \rho] \Pr_{w,(v,J)\in\mathcal{N}_\rho(w)}[E \mid \tau \text{ is bad for } \rho] \geq \frac{\epsilon^3}{k^2}\mu.$$

This contradicts Lemma 26, because $\frac{\epsilon^4}{k^2}\mu \gg \mu^2$ (we assume that $\mu < \epsilon^{120}$). Therefore, we conclude that $\Pr_{\tau\sim\mathcal{D}}[\tau$ is good but not excellent$] \le \epsilon^4$

Since $\tau \sim \mathcal{D}$ is good with probability at least $\frac{\epsilon}{2}$, by averaging a good $\tau \sim \mathcal{D}$ is excellent with probability at least $1-\epsilon^2$. ◀

In order to prove Theorem 21, it is enough to show that every excellent restriction is a *DP restriction*. A natural idea is to define a direct product function by taking the plurality of h_τ on $\mathcal{V}_\tau$, because the agreement of h_τ inside $\mathcal{V}_\tau$ is almost 1. However, it is difficult to prove that this function is close to h_τ because the set $\mathcal{V}_\tau$ is very sparse. We define a local averaging operator, which allows us to go from h_τ that is defined on $\mathcal{V}_\tau$, to a function that is defined on $[N]^{\bar{A}}$.

▶ **Definition 25** (Local averaging operator). For every $\rho \in [0,1]$, let $\mathcal{P}_\rho$ be the following function operator. For every subset $\mathcal{V}_\tau \subset [N]^{\bar{A}}$, and every function $h : \mathcal{V}_\tau \to [M]^t$, the function $\mathcal{P}_\rho h : [N]^t \to [M]^t$ satisfies $\forall i \in [k], w \in [N]^t$,

$$\mathcal{P}_\rho h(w)_i = \underset{(v,J)\in\mathcal{N}_\rho(w), v_i = w_i}{\text{Plurality}} (h(v)_i).$$

If there is no v such that $v_i = w_i$ in $\mathcal{V}_\tau$, we define $\mathcal{P}_\rho h(w)_i$ to an arbitrary value.

The local averaging operator of h takes for every w and i the most frequent value $h(v)_i$ over a ρ-correlated neighborhood of w. We note that the function operator is not linear.

In order to prove that h_τ is close to a direct product function, we first show that that $\mathcal{P}_{\frac{3}{4}} h_\tau$ is close to h_τ, and then that $\mathcal{P}_{\frac{3}{4}} h_\tau$ is close to a direct product function. Clearly $\frac{3}{4}$ is an arbitrary constant, our proof works for any constant $\rho > \frac{1}{2}$, and we fix $\rho = \frac{3}{4}$.

▶ **Lemma 27.** *For every excellent τ,*

$$\Pr_{w\in[N]^{\frac{9k}{10}}}\left[h_\tau(w) \overset{12\alpha k}{\not\approx} \mathcal{P}_{\frac{3}{4}} h_\tau(w) \,\middle|\, w \in \mathcal{V}_\tau\right] \le \epsilon^3.$$

The proof is in Section 3.3, and uses the very high consistency of h_τ inside $\mathcal{V}_\tau$ to show that the plurality vote is almost always consistent with $h_\tau(w)$. In the proof we use reverse hypercontractivity [11] to show that the set $\mathcal{V}_\tau$ is not too sparse, such that for almost all $w \in \mathcal{V}_\tau$, the neighborhood $\mathcal{N}_{\frac{3}{4}}(w)$ is not empty.

In a similar way to the proof of Lemma 27, we show that for an excellent τ the function $\mathcal{P}_{\frac{3}{4}} h_\tau$ has high agreement.

▶ **Lemma 28.** *For every excellent τ,*

$$\Pr_{w,(v,J)}\left[\mathcal{P}_{\frac{3}{4}} h(w)_J \overset{20\alpha k}{\approx} \mathcal{P}_{\frac{3}{4}} h(v)_J\right] \ge 1-\epsilon^{10},$$

where $w \in [N]^{\frac{9k}{10}}$ and $(v,J) \in \mathcal{N}_{\frac{1}{2}}(w)$.

The proof of this lemma also appears in Section 3.3, the main idea is that if $\mathcal{P}_{\frac{3}{4}} h(w_1), \mathcal{P}_{\frac{3}{4}} h(w_2)$ disagree on a lot of coordinates, then a large fraction of their $\frac{3}{4}$-correlated neighborhood also disagree on a lot of coordinates. This can only happen for very few inputs w, else we contradict the fact that τ is excellent.

After showing that $\mathcal{P}_{\frac{3}{4}} h_\tau$ has high agreement, we define g^τ to be the plurality vote of $\mathcal{P}_{\frac{3}{4}} h_\tau$, and then use the high agreement, Lemma 28, to show that they g^τ is close to $\mathcal{P}_{\frac{3}{4}} h_\tau$.

▶ **Lemma 29.** *For every excellent restriction τ there exist a direct product function $g^\tau = g^\tau_1 \dots g^\tau_{\frac{9k}{10}} : [N]^{\frac{9k}{10}} \to [M]^{\frac{9k}{10}}$ such that*

$$\Pr_{w \in [N]^{\frac{9k}{10}}}\left[\mathcal{P}_{\frac{3}{4}} h_\tau(w) \overset{1500\alpha k}{\not\approx} g^\tau(w)\right] \leq 3\epsilon^4.$$

The proof is in Section 3.4.

Using the above lemmas we can prove the local structure.

Proof of Theorem 21. Let $f : [N]^k \to [M]^k$ be a function that passes Test 2 with probability ϵ.

From averaging, $\Pr_{\tau \sim \mathcal{D}}[\tau \text{ is good}] \geq \frac{\epsilon}{2}$, Lemma 26 implies that with probability $(1-\epsilon^2)$, a good τ is also excellent.

Fix an excellent τ, by definition the function h_τ has high consistency inside $\mathcal{V}_\tau$, and by Lemma 27, $\mathcal{P}_{\frac{3}{4}} h$ is close to h on $\mathcal{V}_\tau$. Let $E_1(w)$ be the event that $h_\tau(w) \overset{12\alpha k}{\not\approx} \mathcal{P}_{\frac{3}{4}} h_\tau(w)$, in this notation Lemma 27 implies that

$$\Pr_w[E_1 \mid w \in \mathcal{V}_\tau] \leq \epsilon^3. \tag{6}$$

From Lemma 29, there exists a product function g^τ that is similar to $\mathcal{P}_{\frac{3}{4}} h_\tau$. Denote by $E_2(w)$ the event that $\mathcal{P}_{\frac{3}{4}} h_\tau(w) \overset{1500\alpha k}{\not\approx} g^\tau(w)$. In this notation,

$$\Pr_w[E_2] \leq 3\epsilon^4. \tag{7}$$

We want to use (6) and (7) to prove that h_τ is similar to g^τ on $\mathcal{V}_\tau$. In order to do that, we need to bound the probability of E_2 conditioned on $w \in \mathcal{V}_\tau$.

$$\begin{aligned} 3\epsilon^4 &\geq \Pr_w[E_2] \\ &\geq \Pr_w[w \in \mathcal{V}_\tau] \Pr_w[E_2 \mid w \in \mathcal{V}_\tau] && (\tau \text{ is excellent}) \\ &\geq \frac{\epsilon}{2} \Pr_w[E_2 \mid w \in \mathcal{V}_\tau]. \end{aligned}$$

Therefore $\Pr_w[E_2 \mid w \in \mathcal{V}_\tau] \leq 6\epsilon^3$.

If w is such that none of E_1, E_2 happened, then $h_\tau(w), \mathcal{P}_{\frac{3}{4}} h_\tau(w)$ are equal in all but $12\alpha k$ of the coordinates, and $\mathcal{P}_{\frac{3}{4}} h_\tau(w), g^\tau(w)$ are equal in all but $1500\alpha k$ of the coordinates, which means that $h_\tau(w) \overset{1512\alpha k}{\approx} g^\tau(w)$.

$$\begin{aligned} \Pr_w\left[h_\tau(w) \overset{1512\alpha k}{\not\approx} g^\tau(w) \,\middle|\, w \in \mathcal{V}_\tau\right] &\leq \Pr_w[E_1 \vee E_2 \mid w \in \mathcal{V}_\tau] \\ &\leq \Pr_w[E_1 \mid w \in \mathcal{V}_\tau] + \Pr_w[E_2 \mid w \in \mathcal{V}_\tau] \\ &\leq \epsilon^3 + 6\epsilon^3 < \epsilon^2. \end{aligned}$$

By definition, $h_\tau(w) = f(x_A, w)_{\bar{A}}$,

$$\Pr_w\left[f(x_A, w)_{\bar{A}} \overset{1512\alpha k}{\not\approx} g^\tau(w) \,\middle|\, w \in \mathcal{V}_\tau\right] = \Pr_w\left[h_\tau(w) \overset{1512\alpha k}{\not\approx} g^\tau(w) \,\middle|\, w \in \mathcal{V}_\tau\right] < \epsilon^2.$$

Since $\lambda = 1600\alpha$ we are done. ◀

3.2 Good Restrictions are Excellent with High Probability

For convenience, we restate the lemma.

▶ **Lemma 26.** *For every $\rho \in (0,1)$, let $\tau \sim \mathcal{D}$, $w \in [N]^{\frac{9k}{10}}$ and $(v,J) \in \mathcal{N}_\rho(w)$, then*

$$\Pr_{\tau,w,(v,J)}\left[w,v \in \mathcal{V}_\tau,\ h_\tau(w)_J \overset{\alpha k}{\not\approx} h_\tau(v)_J\right] \leq \left(\frac{9}{10}\right)^{\alpha k}.$$

Proof. Fix $\rho \in (0,1)$, let $E_1(\tau,w,v,J)$ be the event in equation (5) of the definition of excellence, Definition 23. More explicitly, $E_1 = 1$ if $w,v \in \mathcal{V}_\tau$ and $h_\tau(w)_J \overset{\alpha k}{\not\approx} h_\tau(v)_J$.

Recall the definition of h_τ for $\tau = (A,r,\gamma)$, for $w \in \mathcal{V}_\tau$, $h_\tau(w) = f(r,w)_{\bar{A}}$. Therefore, the event E_1 can also be written as $f(r,w)_A = f(r,v)_A = \gamma$ and $f(r,w)_{\bar{A}} \overset{\alpha k}{\not\approx} f(r,v)_{\bar{A}}$.

Let E_2 be the event that $f(r,w)_A = f(r,v)_A$ and $f(r,w)_{\bar{A}} \overset{\alpha k}{\not\approx} f(r,v)_{\bar{A}}$. We can easily see that $E_1 \subseteq E_2$, therefore over every distribution $\Pr[E_1] \leq \Pr[E_2]$.

We start by bounding the probability of event E_2, over the distribution $\tau \sim \mathcal{D}$, $w \in [N]^{\frac{9k}{10}}$ uniformly and $(v,J) \in \mathcal{N}_\rho(w)$. Writing the distribution explicitly:

1. Pick $A \subset [k]$ of size $\frac{k}{10}$.
2. Pick $x \in [N]^k$, set $r = x_A$ and $\gamma = f(x)_A$.
3. Pick $J \subset \left[\frac{9k}{10}\right]$ of size $B(\frac{9k}{10},\rho)$ (binomial random variable).
4. Pick uniform $w,v \in [N]^{\frac{9k}{10}}$ such that $w_J = v_J$.

Notice that E_2 is independent of γ, so it does not matter how γ is chosen. We can define an equivalent process for producing the same distribution (without γ):

1. Pick a set $A' \subset [k]$ of size $\frac{k}{10} + B(\frac{9k}{10},\rho)$.
2. Pick $y,z \in [N]^k$ such that $y_{A'} = z_{A'}$.
3. Pick $A \subseteq A'$ of size $\frac{k}{10}$.
4. Set $r = y_A$, $w = y_{\bar{A}}$ and $v = z_{\bar{A}}$.

In order of E_2 to happen, y,z,A' must be such that $f(y)_{A'} \overset{\alpha k}{\not\approx} f(z)_{A'}$. Furthermore, the set A must be chosen such that $f(y)_A = f(z)_A$. As the second random process allows us to see, A is a random subset of A', and each of the αk coordinates i on which $f(y)_i \neq f(z)_i$ has probability of at least $\frac{1}{10}$ to be chosen to A (as $|A| = \frac{k}{10}$ and $|A'| \leq k$). The probability that none of the αk coordinates are in A is at most $\left(\frac{9}{10}\right)^{\alpha k}$, so

$$\Pr_{\tau,w,(v,J)}[E_1] \leq \Pr_{\tau,w,(v,J)}[E_2] \leq \left(\frac{9}{10}\right)^{\alpha k}. \tag{8}$$

◀

3.3 Local Averaging Operator

In this section we prove the two lemmas concerning local averaging operator. We repeat the two lemmas and prove them.

▶ **Lemma 27.** *For every excellent τ,*

$$\Pr_{w\in[N]^{\frac{9k}{10}}}\left[h_\tau(w) \overset{12\alpha k}{\not\approx} \mathcal{P}_{\frac{3}{4}} h_\tau(w) \,\middle|\, w \in \mathcal{V}_\tau\right] \leq \epsilon^3.$$

Proof. Fix an excellent restriction τ, denote by $\mathcal{V} = \mathcal{V}_\tau$, $h = h_\tau$, $\mathcal{P}_{\frac{3}{4}} h = \mathcal{P}_{\frac{3}{4}} h_\tau$ and $k' = \frac{9k}{10}$. In order to simplify the notations, denote by $\mu = \left(\frac{9}{10}\right)^{\frac{1}{2}\alpha k}$ the constant from the definition of excellence (Definition 23).

From the fact that τ is excellent, we know that $\Pr_{w \in [N]^{k'}}[w \in \mathcal{V}] \geq \frac{\epsilon}{2}$ and

$$\Pr_{w,(v,J) \in \mathcal{N}_{\frac{3}{4}}(w)} \left[w, v \in \mathcal{V}, h_J(w) \overset{\alpha k}{\not\approx} h_J(v) \right] \leq \mu.$$

Our goal is to prove that for almost all $w \in \mathcal{V}$, $\mathcal{P}_{\frac{3}{4}} h(w) \approx h(w)$. First, we characterize the "bad" inputs $w \in \mathcal{V}$ for which we can't prove this claim . Then, we prove it on the rest. Fix $\eta = \epsilon^{20}$, the first set of "bad" inputs is the set of inconsistent ones,

$$B = \left\{ w \in \mathcal{V} \;\middle|\; \Pr_{(v,J) \in \mathcal{N}_{\frac{3}{4}}(w)} \left[v \in \mathcal{V}, h(v)_J \overset{\alpha k}{\not\approx} h(w)_J \right] \geq \frac{\eta}{100} \right\}.$$

By averaging, $\Pr_w[w \in B] \leq \frac{100\mu}{\eta}$.

The second set is the set of "lonely" inputs, inputs that have very sparse neighborhood,

$$L = \left\{ w \in \mathcal{V} \;\middle|\; \Pr_{(v,J) \in \mathcal{N}_{\frac{3}{4}}(w)} [v \in \mathcal{V}] \leq \eta \right\}.$$

By hypercontractivity, Claim 16 (uses [11]), $\Pr_w[w \in L] \leq \eta^{\frac{2}{9}} \left(\frac{\epsilon}{2}\right)^{-\frac{11}{9}}$.

Fix an input $w \in \mathcal{V} \setminus \{B \cup L\}$, we will show that $h(w) \overset{12\alpha k}{\approx} \mathcal{P}_{\frac{3}{4}} h(w)$, i.e. $h(w)$ and $\mathcal{P}_{\frac{3}{4}} h(w)$ are equal on all but $12\alpha k$ of the coordinates. Since $\Pr_w [w \notin B \cap L] \leq \frac{100\mu}{\eta} \eta^{\frac{2}{9}} \left(\frac{\epsilon}{2}\right)^{-\frac{11}{9}} \leq \epsilon^3$, this finishes the proof (ϵ is such that $\epsilon^{120} > \mu$).

Denote by D the following set

$$D = \left\{ i \in [k'] \;\middle|\; h(w)_i \neq \mathcal{P}_{\frac{3}{4}} h(w)_i \right\}.$$

D is the set of coordinates in which the local averaging of h doesn't equal h. Since $w \notin B \cup L$, the neighborhood of w is very consistent, and we show that the set D is small.

Assume towards a contradiction that $|D| > 12\alpha k$. For $v \in [N]^{k'}, J \subset [K]$ and $i \in [k]$, let $E(v, J, i)$ be the event

$$E(v, J, i) = (i \in J \wedge h(w)_i \neq h(v)_i).$$

We will reach a contradiction by upper bounding and lower bounding the probability of the event E, under the distribution $i \in D$ and $(v, J) \in \mathcal{N}_{\frac{3}{4}}(w)$, given that $v \in \mathcal{V}$

Lower bound

We look on $E = E_1 \wedge E_2$, where $E_1 = i \in J$ and $E_2 = h(w)_i \neq h(v)_i$. By definition, for every $i \in D$, the value $h(w)_i$ is not the most probable $h(v)_i$ when $(v, J) \in \mathcal{N}_{\frac{3}{4}}(w)$. Therefore,

$$\forall i \in D, \Pr_{(v,J) \in \mathcal{N}_{\frac{3}{4}}(w)} [E_2 \mid E_1, v \in \mathcal{V}] = \Pr_{(v,J) \in \mathcal{N}_{\frac{3}{4}}(w)} [h(w)_i \neq h(v)_i \mid i \in J, v \in \mathcal{V}] \geq \frac{1}{2}. \quad (9)$$

We want to remove the conditioning over E_1, in order to get a bound E. If we choose a uniform $(v, J) \in \mathcal{N}_{\frac{3}{4}}(w)$, the probability of $i \in J$ is exactly $\frac{3}{4}$. If we condition on $v \in \mathcal{V}$, this probability can be different. We start by bounding the probability of $D \cap J$ to be small.

Every $i \in D$ has probability of $\frac{3}{4}$ be be in J independently, by Chernoff bound (Corollary 10), $\Pr_{(v,J)\in\mathcal{N}_{\frac{3}{4}}(w)}\left[|D \cap J| \leq \frac{3}{5}|D|\right] \leq e^{-\frac{\alpha k}{10}}$. If we condition on $v \in \mathcal{V}$, this probability can increase by a factor of at most $\frac{1}{\eta}$, where $\eta \leq \Pr_{(v,J)\in\mathcal{N}_{\frac{3}{4}}(w)}[v \in \mathcal{V}]$.

$$\Pr_{(v,J)\in\mathcal{N}_{\frac{3}{4}}(w)}\left[|D \cap J| \leq \frac{3}{5}|D| \;\middle|\; v \in \mathcal{V}\right] \leq \frac{1}{\eta}e^{-\frac{\alpha k}{10}}. \tag{10}$$

Equation (10) implies that for a typical $i \in D$, the probability E_1 is not very far from $\frac{3}{4}$. If (v, J) are such that $|D \cap J| \geq \frac{3}{5}|D|$, a random $i \in D$ has probability at least $\frac{3}{5}$ to be in J.

$$\begin{aligned}
\Pr_{(v,J)\in\mathcal{N}_{\frac{3}{4}}(w), i\in D}[E_1 \mid v \in \mathcal{V}] &= \Pr_{(v,J)\in\mathcal{N}_{\frac{3}{4}}(w), i\in D}[i \in J \mid v \in \mathcal{V}] \\
&\geq \Pr_{(v,J)\in\mathcal{N}_{\frac{3}{4}}(w), i\in D}\left[i \in J \wedge |D \cap J| \geq \frac{3}{5}|D| \;\middle|\; v \in \mathcal{V}\right] \\
&&\text{(by (10))} \\
&\geq \frac{3}{5}\left(1 - \frac{1}{\eta}e^{-\frac{\alpha k}{10}}\right). && (11)
\end{aligned}$$

Now we can lower bound the probability of E:

$$\begin{aligned}
\Pr_{(v,J)\in\mathcal{N}_{\frac{3}{4}}(w), i\in D}[E \mid v \in \mathcal{V}] &= \Pr_{(v,J)\in\mathcal{N}_{\frac{3}{4}}(w), i\in D}[E_1 \wedge E_2 \mid v \in \mathcal{V}] \\
&= \Pr_{(v,J)\in\mathcal{N}_{\frac{3}{4}}(w), i\in D}[E_1 \mid v \in \mathcal{V}] \Pr_{(v,J)\in\mathcal{N}_{\frac{3}{4}}(w)}[E_2 \mid E_1, v \in \mathcal{V}] \\
&&\text{(by (9))} \\
&\geq \Pr_{(v,J)\in\mathcal{N}_{\frac{3}{4}}(w), i\in D}[E_1 \mid v \in \mathcal{V}]\frac{1}{2} && \text{(by (11))} \\
&\geq \frac{3}{5}\left(1 - \frac{1}{\eta}e^{-\frac{\alpha k}{10}}\right)\frac{1}{2} \geq \frac{1}{5}. && (12)
\end{aligned}$$

Where the last inequality holds since $\eta = \epsilon^{20}$ and ϵ satisfies $\epsilon^{120} > e^{-\frac{\alpha k}{10}}$.

Upper Bound

We want to upper bound the same probability, and reach a contradiction. Since $w \notin L$, $\Pr_{(v,J)\in\mathcal{N}_{\frac{3}{4}}(w)}[v \in \mathcal{V}] \geq \eta$, and from the fact that $w \notin B$ we know that its neighborhood is consistent, i.e. $\Pr_{(v,J)\in\mathcal{N}_{\frac{3}{4}}(w)}\left[v \in \mathcal{V}, h(v)_J \overset{\alpha k}{\not\approx} h(w)_J\right] \leq \frac{\eta}{100}$. Combining both together,

$$\Pr_{(v,J)\in\mathcal{N}_{\frac{3}{4}}(w)}\left[h(v)_J \overset{\alpha k}{\not\approx} h(w)_J \;\middle|\; v \in \mathcal{V}\right] \leq \frac{1}{100}. \tag{13}$$

This implies that with probability at most $\frac{1}{100}$ the chosen (v, J) can be such that $h(v)_J \overset{\alpha k}{\not\approx} h(w)_J$.

Else, $h(v)_J \overset{\alpha k}{\approx} h(w)_J$, so there are at most αk coordinates $i \in J$ in which $h(v)_i \neq h(w)_i$. Since $|D| \geq 12\alpha k$, with probability at most $\frac{1}{12}$ a uniform $i \in D$ is in these αk coordinates.

$$\Pr_{(v,J)\in\mathcal{N}_{\frac{3}{4}}(w), i\in D}[E \mid v \in \mathcal{V}] \leq \frac{1}{100} + \frac{1}{12} < \frac{1}{5}. \tag{14}$$

And we reached a contradiction with (12). ◀

In order to show that the function $\mathcal{P}_{\frac{3}{4}}h_\tau$ is close to a product function, we need to show that it is consistent in a similar way to h_τ (as in the definition of excellence, Definition 23). Lemma 27 only gives us that $\mathcal{P}_{\frac{3}{4}}h_\tau$ is consistent among the inputs in $\mathcal{V}_\tau$, and not in all $[N]^{\frac{9k}{10}}$.

▶ **Lemma 28.** *For every excellent τ,*

$$\Pr_{w,(v,J)}\left[\mathcal{P}_{\frac{3}{4}}h(w)_J \overset{20\alpha k}{\approx} \mathcal{P}_{\frac{3}{4}}h(v)_J\right] \geq 1-\epsilon^{10},$$

where $w \in [N]^{\frac{9k}{10}}$ and $(v,J) \in \mathcal{N}_{\frac{1}{2}}(w)$.

Proof. This proof is similar to the proof of Lemma 27. We fix excellent τ and denote $\mathcal{V} = \mathcal{V}_\tau$, $h = h_\tau$ and $\mathcal{P}_{\frac{3}{4}}h = \mathcal{P}_{\frac{3}{4}}h_\tau$, $k' = \frac{9k}{10}$ and $\mu = \left(\frac{9}{10}\right)^{\frac{1}{2}\alpha k}$.

We characterize the inputs $w,(v,J)$ on which we can't prove that $\mathcal{P}_{\frac{3}{4}}h(w)_J \overset{20\alpha k}{\approx} \mathcal{P}_{\frac{3}{4}}h(v)_J$. Instead of the set B in the proof of Lemma 27, we define a set of two correlated inputs $(w,(v,J))$ that are inconsistent. Fixing $\eta = \epsilon^{51}$, let

$$C = \left\{ w,(v,J) \;\middle|\; \Pr_{(w',J'),(v',J'')}\left[w',v' \in \mathcal{V}, h(w')_{\tilde{J}} \overset{\alpha k}{\not\approx} h(v')_{\tilde{J}}\right] \geq \frac{\eta^2}{4000} \right\}.$$

Where $(w',J') \in \mathcal{N}_{\frac{3}{4}}(w), (v',J'') \in \mathcal{N}_{\frac{3}{4}}(v)$ and $\tilde{J} = J \cap J' \cap J''$.

If w is chosen uniformly in $[N]^{k'}$ and $(v,J) \in \mathcal{N}_{\frac{1}{2}}(w)$, then the marginal distribution on w' is uniform, and $(v',\tilde{J}) \in \mathcal{N}_{(\frac{3}{4})^2\frac{1}{2}}(w')$, since for each i independently, the probability of i to be in $\tilde{J} = J \cap J' \cap J''$ is $\left(\frac{3}{4}\right)^2 \frac{1}{2}$.

Since τ is excellent, $\Pr_{w',(v',\tilde{J})}\left[w',v' \in \mathcal{V}, h(w')_{\tilde{J}} \overset{\alpha k}{\not\approx} h(v')_{\tilde{J}}\right] \leq \mu$. By averaging, it means that $\Pr_{w,(v,J)}[w,(v,J) \in C] \leq \frac{4000\mu}{\eta^2}$.

We define the set of inputs with sparse neighborhood,

$$L = \left\{ w \in [N]^{k'} \;\middle|\; \Pr_{(w',J') \in \mathcal{N}_{\frac{3}{4}}(w)}[w' \in \mathcal{V}] \leq \eta \right\}.$$

From hypercontractivity argument, see Claim 16, $\Pr_w[w \in L] \leq \eta^{\frac{2}{9}}\left(\frac{\epsilon}{2}\right)^{-\frac{11}{9}}$.

For every w and (v,J) such that $w,v \notin L$ and $(w,(v,J)) \notin C$, we show that $\mathcal{P}_{\frac{3}{4}}h(w)_J \overset{20\alpha k}{\not\approx} \mathcal{P}_{\frac{3}{4}}h(v)_J$. This finishes the proof since for $w \in [N]^{k'}$ and $(v,J) \in \mathcal{N}_{\frac{1}{2}}(w)$,

$$\Pr_{w,(v,J)}[(w,(v,J)) \in C \vee w \in L \vee v \in L] \leq \frac{4000\mu}{\eta^2} + 2 \cdot \eta^{\frac{2}{9}}\left(\frac{\epsilon}{2}\right)^{-\frac{11}{9}} \leq \epsilon^{10}.$$

Fix $w,(v,J)$ such that $w,v \notin L$ and $(w,(v,J)) \notin C$, and let $D \subseteq J$ be the set

$$D = \left\{ i \in J \;\middle|\; \mathcal{P}_{\frac{3}{4}}h(w)_i \neq \mathcal{P}_{\frac{3}{4}}h(v)_i \right\}.$$

Similarly to the previous proof, we assume towards a contradiction that $|D| \geq 20\alpha k$.

For every $J',J'' \subset [k'], w',v' \in \mathcal{V}$ and $i \in [k']$, we denote by $E(J',J'',w',v',i)$ the following event:

$$E(J',J'',w',v',i) = \left(h(w')_i \neq h(v')_i \wedge i \in J' \cap J''\right).$$

We upper bound and lower bound the probability of this event, under the distribution $i \in D$ and $(w',J') \in \mathcal{N}_{\frac{3}{4}}(w), (v',J'') \in \mathcal{N}_{\frac{3}{4}}(v)$ given that $w',v' \in \mathcal{V}$.

Lower Bound

We look on $E = E_1 \wedge E_2$, where $E_1 = i \in J' \cap J''$ and $E_2 = h(w')_i \neq h(v')_i$.

For every $i \in D$, $\mathcal{P}_{\frac{3}{4}} h(w)_i \neq \mathcal{P}_{\frac{3}{4}} h(v)_i$, so the most frequent value $h(w')_i$ for $(w', J') \in \mathcal{N}_{\frac{3}{4}}(w)$ doesn't equal the most frequent value $h(v')_i$ for $(v', J'') \in \mathcal{N}_{\frac{3}{4}}(v)$. For every $i \in D$, taking $(w', J') \in \mathcal{N}_{\frac{3}{4}}(w), (v', J'') \in \mathcal{N}_{\frac{3}{4}}(v)$:

$$\Pr_{\substack{(w',J')\\(v',J'')}}[E_2 \mid E_1, w', v' \in \mathcal{V}] = \Pr_{\substack{(w',J')\\(v',J'')}}[h(w')_i \neq h(v')_i \mid i \in J' \cap J'', w', v' \in \mathcal{V}] \geq \frac{1}{2}. \tag{15}$$

In order to prove the lower bound, we need to remove the condition over E_1. To do that, we need to lower bound the size of $D \cap J' \cap J''$. Both J' and J'' are taken by picking each coordinated independently with probability $\frac{3}{4}$. If we do not condition on $w', v' \in \mathcal{V}$ expected value of $|D \cap J' \cap J''|$ is $\left(\frac{3}{4}\right)^2 |D|$. Each $i \in D$ is in $J' \cap J''$ with probability $\left(\frac{3}{4}\right)^2$ independently, so using Chernoff bound (Corollary 10),

$$\Pr_{\substack{(w',J') \in \mathcal{N}_{\frac{3}{4}}(w)\\(v',J'') \in \mathcal{N}_{\frac{3}{4}}(v)}}[|D \cap J' \cap J''| \leq 0.56|D|] \leq e^{-\frac{\alpha k}{9000}}.$$

If we condition on $w' \in \mathcal{V}, v' \in \mathcal{V}$, the probability can increase by a factor of at most $\frac{1}{\eta^2}$, where $\Pr_{(w',J') \in \mathcal{N}_{\frac{3}{4}}(w)}[w' \in \mathcal{V}] \geq \eta$ and $\Pr_{(v',J'') \in \mathcal{N}_{\frac{3}{4}}(v)}[v' \in \mathcal{V}] \geq \eta$,

$$\Pr_{\substack{(w',J') \in \mathcal{N}_{\frac{3}{4}}(w)\\(v',J'') \in \mathcal{N}_{\frac{3}{4}}(v)}}[|D \cap J' \cap J''| \leq 0.56|D| \mid w', v' \in \mathcal{V}] \leq \frac{1}{\eta^2} e^{-\frac{\alpha k}{9000}}. \tag{16}$$

If $|D \cap J' \cap J''| \geq 0.56|D|$, then a uniform $i \in D$ has probability of at least 0.56 to be in $J' \cap J''$,

$$\Pr_{\substack{i \in D, (w',J') \in \mathcal{N}_{\frac{3}{4}}(w)\\(v',J'') \in \mathcal{N}_{\frac{3}{4}}(v)}}[E_1 \mid w', v' \in \mathcal{V}] \geq \left(1 - \frac{1}{\eta^2} e^{-\frac{\alpha k}{9000}}\right) 0.56 \geq 0.55. \tag{17}$$

The last inequality is correct because we assume ϵ is large enough to satisfy $\epsilon^{120} > \frac{1}{\eta^2} e^{-\frac{\alpha k}{9000}}$.

Combining (15) and (17), we can lower bound the probability of E, when $i \in D$, $(w', J') \in \mathcal{N}_{\frac{3}{4}}(w)$ and $(v', J'') \in \mathcal{N}_{\frac{3}{4}}(v)$,

$$\begin{aligned}\Pr_{i,(w',J'),(v',J'')}[E \mid w', v' \in \mathcal{V}] &= \Pr_{i,(w',J'),(v',J'')}[E_1 \wedge E_2, \mid w', v' \in \mathcal{V}] \\ &= \Pr_{i,(w',J'),(v',J'')}[E_1 \mid w', v' \in \mathcal{V}] \qquad (18)\\ &\cdot \Pr_{i,(w',J'),(v',J'')}[E_2 \mid w', v' \in \mathcal{V}, E_1] \\ &\geq \frac{1}{2} \cdot 0.55 > \frac{1}{4}. \qquad (19)\end{aligned}$$

Upper Bound

Since $(w, (v, J)) \notin C$, we know that

$$\Pr_{\substack{(w',J') \in \mathcal{N}_{\frac{3}{4}}(w)\\(v',J'') \in \mathcal{N}_{\frac{3}{4}}(v)}}\left[w', v' \in \mathcal{V}, h(w')_{\tilde{J}} \overset{\alpha k}{\not\approx} h(v')_{\tilde{J}}\right] \leq \frac{\eta^2}{4000},$$

where $\tilde{J} = J \cap J' \cap J''$. From the fact that $w \notin L$, $\Pr_{(w',J')\in\mathcal{N}_{\frac{3}{4}}(w)}[w' \in \mathcal{V}] \geq \eta$ and since $v \notin L$, $\Pr_{(v',J'')\in\mathcal{N}_{\frac{3}{4}}(v)}[v' \in \mathcal{V}] \geq \eta$. This implies that

$$\Pr_{\substack{(w',J')\in\mathcal{N}_{\frac{3}{4}}(w)\\(v',J'')\in\mathcal{N}_{\frac{3}{4}}(v)}}\left[h(w')_{\tilde{J}} \overset{\alpha k}{\not\approx} h(v')_{\tilde{J}} \,\middle|\, w',v' \in \mathcal{V}\right] \leq \frac{1}{4000}.$$

If $h(w')_{\tilde{J}} \overset{\alpha k}{\approx} h(v')_{\tilde{J}}$, then even if all these αk coordinates are in D, a uniform $i \in D$ has probability of at most $\frac{\alpha k}{|D|} \leq \frac{\alpha k}{20\alpha k} \leq \frac{1}{20}$ to be one of these coordinates. Therefore,

$$\Pr_{\substack{i\in D,(w',J')\in\mathcal{N}_{\frac{3}{4}}(w)\\(v',J'')\in\mathcal{N}_{\frac{3}{4}}(v)}}[E] \leq \frac{1}{4000} + \frac{1}{20} < \frac{1}{10},$$

which contradicts (19). ◀

3.4 Direct Product Function

Fixing an excellent τ, we first show that the local average function $\mathcal{P}_{\frac{3}{4}}h_\tau$ is close to a product function g^τ. Then, by Lemma 27, we will conclude that h_τ is close to g^τ. This implies that τ is a DP restriction as needed.

In this section we prove Lemma 29,

▶ **Lemma 29.** *For every excellent restriction τ there exist a product function $g^\tau : [N]^{\frac{9k}{10}} \to [M]^{\frac{9k}{10}}$ such that*

$$\Pr_{w\in[N]^{\frac{9k}{10}}}\left[\mathcal{P}_{\frac{3}{4}}h_\tau(w) \overset{1500\alpha k}{\not\approx} g^\tau(w)\right] \leq 3\epsilon^4.$$

We first define g^τ, the candidate direct product function

▶ **Definition 30.** For each excellent $\tau = (A, r, \gamma)$, let $g^\tau : [N]^{\frac{9k}{10}} \to [M]^{\frac{9k}{10}}$ be the following function, for each $i \notin A$ and $b \in [N]$,

$$g_{\tau,i}(b) = \operatorname*{Plurality}_{w\in[N]^{\frac{9k}{10}} \text{ s.t. } w_i=b} \{\mathcal{P}_{\frac{3}{4}}h(w)_i\},$$

ties are broken arbitrarily.

We prove Lemma 29 using the following few claims.

▶ **Claim 31.**

$$\Pr_{i\in\left[\frac{9k}{10}\right],w,v\in[N]^{\frac{9k}{10}}}\left[\mathcal{P}_{\frac{3}{4}}h(w)_i = \mathcal{P}_{\frac{3}{4}}h(v)_i \,\middle|\, w_i = v_i\right] \geq 1 - 200\alpha.$$

In order to prove Claim 31, we need to define an “almost ρ-correlated”’ distribution.

▶ **Definition 32.** (x, J) are almost ρ-correlated to $y \in [N]^k$, denoted by $(x, J) \in A_\rho(y)$, if they are chosen by the following process:

1. Choose $i \in [k]$ uniformly at random, set $J = \{i\}$.
2. For each $j \neq i$, add j to J with probability ρ independently.
3. Set $x_J = y_J$ and the rest of x is uniform.

▶ **Claim 33.** *For any $\rho \in (0,1)$ and any event $E(y,x,J)$ over $x, y \in [N]^k$ and $J \subseteq [k]$,*

$$\Pr_{y\in[N]^k,(x,J)\in A_\rho(y)}[E(y,x,J)] \leq 2 \Pr_{y\in[N]^k,(x,J)\in\mathcal{N}_\rho(y)}[E(y,x,J)] + 5e^{-\frac{\rho k}{4}}.$$

The proof appears at the end of the section.

Proof of Claim 31. Let $k' = \frac{9k}{10}$. We start by showing that for a uniform $w \in [N]^{k'}$, $(u, J') \in A_{\frac{1}{2}}(w)$ and $i \in J'$, $\Pr_{i,w,(v,J')}\left[\mathcal{P}_{\frac{3}{4}}h(w)_i = \mathcal{P}_{\frac{3}{4}}h(v)_i\right] \geq 1 - 100\alpha$.

Let E_1 be the event that $\mathcal{P}_{\frac{3}{4}}h(w)_i \neq \mathcal{P}_{\frac{3}{4}}h(v)_i$, we further define the following two events, let E_2 to be the event $\mathcal{P}_{\frac{3}{4}}h(w)_{J'} \overset{20\alpha k}{\not\approx} \mathcal{P}_{\frac{3}{4}}h(u)_{J'}$, and let E_3 be the event that $|J'| < \frac{k}{4}$.

If both E_2, E_3 don't happen, then $|J'| \geq \frac{k}{4}$, and there are at most $20\alpha k$ coordinates i in which $\mathcal{P}_{\frac{3}{4}}h(w)_i \neq \mathcal{P}_{\frac{3}{4}}h(v)_i$. Therefore, a uniform $i \in J'$ has probability at most $\frac{20\alpha k}{\frac{k}{4}}$ to satisfy $\mathcal{P}_{\frac{3}{4}}h(w)_i \neq \mathcal{P}_{\frac{3}{4}}h(v)_i$,

$$\Pr_{w,(v,J'),i}[E_1 \mid \neg E_2, \neg E_3] \leq \frac{20\alpha k}{\frac{k}{4}} = 80\alpha. \tag{20}$$

In order to remove the condition over $\neg E_2, \neg E_3$, we bound their probability. For a uniform $w \in [N]^{k'}$ and $(u, J') \in A_{\frac{1}{2}}(\rho)$,

$$\begin{aligned}\Pr_{w,(u,J')\in A_{\frac{1}{2}}(w)}[E_2] &\leq 2 \Pr_{w,(u,J')\in\mathcal{N}_{\frac{1}{2}}(w)}[E_2] + 5e^{-\frac{\rho k}{4}} && \text{(by Claim 33)}\\ &\leq 2\epsilon^{10} + 5e^{-\frac{\rho k}{4}} \leq 3\epsilon^{10}. && \text{(by Lemma 28)}\end{aligned}$$

Similarly, for $w \in [N]^{k'}$ and $(u, J') \in A_{\frac{1}{2}}(w)$,

$$\begin{aligned}\Pr_{w,(u,J')\in A_{\frac{1}{2}}(w)}[E_3] &\leq 2 \Pr_{w,(u,J')\in\mathcal{N}_{\frac{1}{2}}(w)}[E_3] + 5e^{-\frac{\rho k}{4}} && \text{(by Claim 33)}\\ &\leq 2e^{-\frac{k}{100}} + 5e^{-\frac{\rho k}{4}} \leq \epsilon^{10}. && \text{(Chernoff Bound)}\end{aligned}$$

For $(u, J') \in \mathcal{N}_{\frac{1}{2}}(w)$, each coordinate i is in J' with probability $\frac{1}{2}$ independently, so we can use Chernoff bound. If we add a condition on $\neg E_2$, it can increase the probability by a factor of $\frac{1}{\Pr[\neg E_2]} < 2$, therefore $\Pr_{w,(u,J')\in A_{\frac{1}{2}}(w)}[E_3 \mid \neg E_2] \leq 2\epsilon^{10}$.

Combining everything together, for a uniform $w \in [N]^{k'}$, $(u, J') \in A_{\frac{1}{2}}(w)$ and $i \in J'$,

$$\begin{aligned}\Pr_{w,(u,J'),i}[E_1] \leq& \Pr_{w,(u,J'),i}[E_2] + \Pr_{w,(u,J'),i}[E_3 \mid \neg E_2] + \Pr_{w,(u,J'),i}[E_1 \mid \neg E_2, \neg E_3]\\ \leq& 3\epsilon^{10} + 2\epsilon^{10} + 80\alpha \leq 100\alpha\end{aligned} \tag{21}$$

Let $\mathcal{D}' : [k'] \times [N]^{k'} \times [N]^{k'} \times [N]^{k'} \to \{0,1\}$ be the following distribution, generating i, w, v, u:

1. Pick a uniform $i \in [k']$.
2. Pick $w, v \in [N]^{k'}$ such that $w_i = v_i$.
3. For every $j \neq i$, insert j into J with probability $\frac{1}{2}$ independently.
4. For every $j \in [k']$, $u_j = \begin{cases} w_j & j \in J \\ v_j & \text{else} \end{cases}$.

The distribution $\mathcal{D}'$ is built such that the marginal distribution over w, v, i is that $i \in [k']$ uniformly, and w, v are uniform in $[N]^{k'}$ such that $w_i = v_i$. Furthermore, the marginal distribution over $w, (u, J \cup \{i\}), i$ is such that $w \in [N]^{k'}$ uniformly, $(u, J \cap \{i\}) \in A_{\frac{1}{2}}(w)$ and the coordinate i is uniform in $\{i\} \cup J$. Similarly, the marginal distribution over $v, (u, \bar{J})$ is $v \in [N]^{k'}$, $(u, \bar{J}) \in A_{\frac{1}{2}}(v)$ and $i \in \bar{J}$.

Therefore, we can use equation (21) on the pairs $w, (u, J \cup \{i\})$ and $v, (u, \bar{J})$, and by union bound,

$$\begin{aligned}\Pr_{\substack{i \in [k'] \\ w,v \in [N]^{k'}}} \left[\mathcal{P}_{\frac{3}{4}} h(w)_i = \mathcal{P}_{\frac{3}{4}} h(v)_i \;\middle|\; w_i = v_i\right] \geq& \Pr_{i,w,v,u \sim \mathcal{D}'} \left[\mathcal{P}_{\frac{3}{4}} h(w)_i = \mathcal{P}_{\frac{3}{4}} h(v)_i = \mathcal{P}_{\frac{3}{4}} h(u)_i\right] \\ \geq& 1 - 100\alpha - 100\alpha.\end{aligned}$$

◀

▶ **Corollary 34.**

$$\Pr_{w \in [N]^{\frac{9k}{10}}, i \in [\frac{9k}{10}]} \left[\mathcal{P}_{\frac{3}{4}} h(w)_i = g(w)_i\right] \geq 1 - 400\alpha.$$

Proof. For each $w \in [N]^{\frac{9k}{10}}$ and $i \in [\frac{9k}{10}]$ such that $\mathcal{P}_{\frac{3}{4}} h(w)_i \neq g(w)_i$, the value $\mathcal{P}_{\frac{3}{4}} h(w)_i$ is not the most frequent, $\Pr_{v \in [N]^{\frac{9k}{10}}} [\mathcal{P}_{\frac{3}{4}} h(w)_i = \mathcal{P}_{\frac{3}{4}} h(v)_i | w_i = v_i] \leq \frac{1}{2}$. Therefore,

$$\Pr_{w,v \in [N]^{\frac{9k}{10}}, i \in [\frac{9k}{10}]} \left[\mathcal{P}_{\frac{3}{4}} h(w)_i \neq \mathcal{P}_{\frac{3}{4}} h(v)_i \;\middle|\; w_i = v_i\right] \geq \frac{1}{2} \Pr_{w \in [N]^{\frac{9k}{10}}, i \in [\frac{9k}{10}]} \left[\mathcal{P}_{\frac{3}{4}} h(w)_i \neq g(w)_i\right].$$

Using Claim 31 we reach the corollary. ◀

Proof of Lemma 29. Fix an excellent τ, denote $k' = \frac{9k}{10}$.

For each $w \in [N]^{k'}$, let $D_w \subset [k']$ be the set of coordinates in which $g^{\tau}(w), \mathcal{P}_{\frac{3}{4}} h_{\tau}(w)$ differs

$$D_w = \left\{ i \in [k'] \;\middle|\; g^{\tau}(w)_i \neq \mathcal{P}_{\frac{3}{4}} h_{\tau}(w)_i \right\}.$$

Let $C \subset [N]^{k'}$ be the set of inputs such that $g^{\tau}, \mathcal{P}_{\frac{3}{4}} h_{\tau}$ are similar on them,

$$C = \left\{ w \in [N]^{k'} \;\middle|\; |D_w| \leq 500\alpha k \right\}.$$

By Corollary 34 and averaging, $\Pr_w[w \in C] \geq \frac{1}{5}$.

Let $B \subset [N]^{k'}$ be the set of inputs on which $g^{\tau}, \mathcal{P}_{\frac{3}{4}} h_{\tau}$ are far,

$$B = \left\{ w \in [N]^{k'} \;\middle|\; |D_w| \geq 1500\alpha k \right\}.$$

B is the set of inputs in which $\mathcal{P}_{\frac{3}{4}} h_{\tau}(w) \overset{1500\alpha k}{\not\approx} g^{\tau}(w)$, so our goal is to prove that B is small.

Let $E_1(w, v, J)$ be the event that $|J \cap D_w| > 600\alpha k$, and let $E_2(w, v, J)$ be the event that $\mathcal{P}_{\frac{3}{4}} h(w)_J \overset{20\alpha}{\not\approx} \mathcal{P}_{\frac{3}{4}} h(v)_J$. By Lemma 28, $\Pr_{w \in [N]^{k'}, (v,J) \in \mathcal{N}_{\frac{1}{2}}(w)} [E_2] \leq \epsilon^{10}$.

For every v, w, J such that $v_J = w_J$, the function g satisfies $g^{\tau}(w)_J = g^{\tau}(v)_J$ (since g is a product function), and therefore $E_1 \wedge (v \in C) \implies E_2$. This is because if E_2 doesn't hold, then $\mathcal{P}_{\frac{3}{4}} h(w)_J \overset{20\alpha}{\approx} \mathcal{P}_{\frac{3}{4}} h(v)_J$, if E_1 does hold then $|J \cap D_w| > 600\alpha k$, which means that $|D_v \cap J| \geq 580\alpha k$, and $v \notin C$.

We show if B isn't small, then $E_1 \wedge (v \in C)$ happens often, when we pick $w \in [N]^k$, $(v, J) \in \mathcal{N}_{\frac{1}{2}}(w)$.

For $w \in B$, the set D_w is large, $|D_w| \geq 1500\alpha k$, if we take $(v, J) \in \mathcal{N}_{\frac{1}{2}}(w)$, each coordinate $i \in D_w$ is in J with probability $\frac{1}{2}$ independently, so for $w \in B$, by Chernoff bound

$$\Pr_{(v,J)\in\mathcal{N}_{\frac{1}{2}}(w)}[E_1(w)] = \Pr_{(v,J)\in\mathcal{N}_{\frac{1}{2}}(w)}[|J \cap D_w| > 600\alpha k] \geq 1 - e^{-\frac{\rho k}{100}}.$$

From reverse hypercontractivity [11], Corollary 15

$$\Pr_{w,(v,J)\in\mathcal{N}_{\frac{1}{2}}(w)}[w \in B, v \in C] \geq \Pr_w[w \in C]^{\frac{3}{2}} \Pr_w[w \in B]^{\frac{5}{2}}.$$

Therefore,

$$\begin{aligned}\Pr_{w,(v,J)\in\mathcal{N}_{\frac{1}{2}}(w)}[w \in B, v \in C \wedge E_1] &\geq \Pr_w[w \in C]^{\frac{3}{2}} \Pr_w[w \in B]^{\frac{5}{2}} - e^{-\frac{\rho k}{100}} \\ &\geq \left(\frac{1}{5}\right)^{\frac{3}{2}} \Pr_w[w \in B]^{\frac{5}{2}} - e^{-\frac{\rho k}{100}}. \qquad (22)\end{aligned}$$

Where (22) is since $\Pr_w[w \in C] \geq \frac{1}{5}$.

Since $E_1 \wedge (v \in C) \implies E_2$ and by Lemma 28, $\Pr_{w\in[N]^{k'},(v,J)\in\mathcal{N}_{\frac{1}{2}}(w)}[E_2] \leq \epsilon^{10}$, it means that (22) should be smaller than ϵ^{10}, which implies $\Pr_w[w \in B] \leq 3\epsilon^4$ and finishes the proof. ◀

We are left with proving the simple distribution claim – that almost ρ correlated is similar to ρ correlated.

Proof of Claim 33. The proof is based on the fact that the distributions $\mathcal{N}_\rho(y)$, $A_\rho(y)$ are very close, and the probability of an event depending on y, x, J is not much different in both distributions.

By Chernoff bound, ρ-correlated sets are almost always of size about ρk, this holds for almost ρ correlated as well,

$$\Pr_{(x,J)\in\mathcal{N}_\rho(y)}[|J| > 2\rho k] \leq e^{-\frac{\rho k}{3}},$$

$$\Pr_{(x,J)\in A_\rho(y)}[|J| > 2\rho k] \leq e^{-\frac{\rho k}{4}}.$$

For each $y \in [N]^k$, let B_y be the (x, J) that satisfy $E(y, x, J)$

$$B_y = \{(x, J) \mid\ E(y, x, J) = 1\}.$$

Using this notation

$$\Pr_{y\in[N]^k,(x,J)\in\mathcal{N}_\rho(y)}[E(y, x, J)] = \Pr_{y\in[N]^k,(x,J)\in\mathcal{N}_\rho(y)}[(x, J) \in B_y].$$

Fix $y \in [N]^k$, for each $(x, J) \in B_y$, by the definition of ρ-correlation,

$$\Pr_{(z,J')\in\mathcal{N}_\rho(y)}[(z, J') = (x, J)] = \rho^{|J|}(1-\rho)^{k-|J|}\left(\frac{1}{N}\right)^{k-|J|}.$$

By the definition of almost ρ correlation,

$$\Pr_{(z,J')\in A_\rho(y)}[(z,J')=(x,J)] = \frac{|J|}{k}\rho^{|J|-1}(1-\rho)^{k-|J|}\left(\frac{1}{N}\right)^{k-|J|}.$$

Note that for each such $(x,J)\in B_y$ such that $|J|\le 2\rho k$,

$$\Pr_{(z,J')\in A_\rho(y)}[(z,J')=(x,J)] \le 2\Pr_{(z,J')\in \mathcal{N}_\rho(y)}[(z,J')=(x,J)].$$

Therefore

$$\begin{aligned}\Pr_{(x,J)\in A_\rho(y)}[(x,J)\in B_y] \le &\Pr_{(x,J)\in A_\rho(y)}[|J|\ge 2\rho k] + \Pr_{(x,J)\in A_\rho(y)}[(x,J)\in B_y \mid |J|\le 2\rho k]\\ \le &e^{-\frac{\rho k}{4}} + 2\Pr_{(x,J)\in \mathcal{N}_\rho(y)}[(x,J)\in B_y \mid |J|\le 2\rho k]\\ \le &e^{-\frac{\rho k}{4}} + 2\Pr_{(x,J)\in \mathcal{N}_\rho(y)}[(x,J)\in B_y] + 4e^{-\frac{\rho k}{3}}.\end{aligned}$$

When we used conditional probability in the last inequality. This is true for all $y\in[N]^{k'}$, therefore,

$$\begin{aligned}\Pr_{y\in[N]^k,(x,J)\in A_\rho(y)}[E(y,x,J)] = &\Pr_{y\in[N]^k,(x,J)\in A_\rho(y)}[(x,J)\in B_y]\\ \le 2&\Pr_{y\in[N]^k,(x,J)\in \mathcal{N}_\rho(y)}[(x,J)\in B_y] + 5e^{-\frac{\rho k'}{4}}.\end{aligned}$$ ◀

4 Global Structure for Sets

Up until now we have considered functions $f:[N]^k\to[M]^k$ whose inputs are ordered tuples $(x_1\ldots,x_k)\in[N]^k$. We now move to consider functions $f:\binom{[N]}{k}\to[M]^k$ whose inputs are unordered $\{x_1,\ldots,x_k\}\in\binom{[N]}{k}$, and we assume that $N\gg k$ (for tuples no such assumption was made).

To each subset $S=\{s_1,\ldots,s_k\}$ the function f assigns $f(S)\in[M]^k$. $f(S)$ should be viewed as a "local function" on S, assigning a value from $[M]$ to every $a\in S$. We denote by $f(S)_a$ the output of f that corresponds to a. For a subset $W\subset S$, let $f(S)_W$ be the outputs of f corresponding to the elements in W.

There are straightforward analogs to Theorem 1 and Theorem 21 which we present and prove in this section. Interestingly, in the case of sets deducing global structure from restricted global structure is quite easier than it is for tuples.

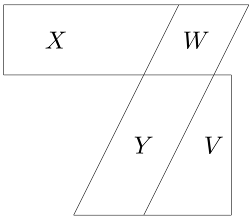
First, let us present the Z test for sets, from [9] when $t=\frac{k}{10}$. Let $\alpha_{Z_{set}(\frac{k}{10})}(f)$ be the success probability of this test. This is the same test as Test 3 from the introduction written differently, it is written this way because it is easier to refer to the test items during the proof.

▶ **Theorem 5.** *There exist a small constant $c>0$, such that for every constant $\lambda>0$, large enough $k\in\mathbb{N}$ and $N>k^2e^{10c\lambda k}$, if the function $f:\binom{[N]}{k}\to[M]^k$ passes Test 3 with probability $\alpha_{Z_{set}(\frac{k}{10})}(f)=\epsilon>e^{-c\lambda k}$, then there exist a function $g:[N]\to[M]$ such that*

$$\Pr_S\left[f(S)\overset{\lambda k}{\approx}g(S)\right]\ge\epsilon-4\epsilon^2.$$

In order to analyze this test, we first need to "translate" the restricted global structure result into this setting, and then prove the global structure in this setting.

1. Choose a random set $W \subset [N]$ of size $\frac{k}{10}$.
2. Choose $X, Y \subset [N] \setminus W$ of size $\frac{9k}{10}$.
3. If $f(X \cup W)_W \neq f(Y \cup W)_W$ reject.
4. Choose $V \subset [N] \setminus Y$ of size $\frac{k}{10}$.
5. If $f(Y \cup W)_Y \neq f(Y \cup V)_Y$ reject, else accept.



Denote by $\alpha_{Z_{set}(\frac{k}{10})}(f)$ the success probability of f on this test.

Test 4 "Z" test for functions over sets, with $t = \frac{k}{10}$ (3-query test).

4.1 Restricted Global Structure for Sets

In this section, we see that for $N \gg k$, the restricted global structure for tuples, Theorem 21, implies restricted global structure for sets. First we define analog definitions for sets, for good restrictions and DP restrictions. To make the reduction proof simpler, we use a constant $\eta \in \left[1 - \frac{k^2}{N}, 1\right]$ (i.e. almost 1) and define good pair using η.

▶ **Definition 35** (Good pair)**.** A pair $X, W \subset [N]$, $|X| = \frac{9k}{10}, |W| = \frac{k}{10}$ is good if

$$\Pr_Y \left[f(X \cup W)_W = f(Y \cup W)_W \mid Y \cap W = \emptyset\right] > \frac{\epsilon}{2}\eta.$$

This definition is analog to Definition 19 of good restriction, the main difference between the definitions is that here we don't have a set of coordinates $A \subset [k]$, because f is defined on sets and not coordinates.

▶ **Definition 36** (DP pair)**.** A pair $X, W \subset [N]$, $|X| = \frac{9k}{10}, |W| = \frac{k}{10}$ is a DP pair if it is good, and if there exist a function $g_{X,W} : [N] \to [M]$ such that

$$\Pr_Y \left[f(Y \cup W)_Y \overset{3\alpha k}{\not\approx} g_{X,W}(Y) \,\middle|\, Y \cap W = \emptyset, f(X \cup W)_W = f(Y \cup W)_W \right] \leq 2\epsilon^2.$$

This definition is analog to Definition 20 of DP restriction, only here there is a single function $g_{X,W}$, instead of $\frac{9k}{10}$ different functions in the case of coordinates.

▶ **Lemma 37** (Restricted global structure for sets)**.** *There exist a small constants $\delta > 0$, such that for every constant $\lambda > 0$ and large enough $k \in \mathbb{N}$ such that $N > k^2 e^{10\delta\lambda k}$, the following holds,*

For every function $f : \binom{[N]}{k} \to [M]^k$, if $\alpha_{Z_{set}(\frac{k}{10})}(f) = \epsilon > e^{-\delta\lambda k}$, then at least $(1 - \epsilon^2 - \frac{k^2}{N})$ of the good pairs $W \in \binom{[N]}{\frac{k}{10}}, X \in \binom{[N]}{\frac{9k}{10}}$ are DP pairs, i.e. there exist $g_{X,W} : [N] \to [M]$ such that

$$\Pr_Y \left[f(Y \cup W)_Y \overset{3\alpha k}{\not\approx} g_{X,W}(Y) \,\middle|\, Y \cap W = \emptyset, f(X \cup W)_W = f(Y \cup W)_W \right] \leq 2\epsilon^2.$$

This lemma for sets is analog to Theorem 21, and we prove it by a reduction from it. For every $f : \binom{[N]}{k} \to [M]^k$ we define a function $f' : [N]^k \to [M]^k \cup \perp$ that equals $\perp$ if the input has two identical coordinates, and identifies with f everywhere else. For $N \gg k$, almost all inputs don't have two identical coordinates, and f', f are equal almost always.

Using Theorem 21, we derive a restricted global structure on f' which gives a direct product function $g^\tau = g_1^\tau, \dots g_{\frac{9k}{10}k}^\tau$ for every excellent τ. Since f equals f' almost always, we find an equivalence between excellent τ and excellent X, W. Then, we build a restricted global function $g_{X,W}$ by taking the most frequent value among the product $g_1^\tau, \ldots g_{\frac{9k}{10}}^\tau$. Note that even though f' is permutation invariant, the functions $g_1^\tau, \ldots g_{\frac{9k}{10}}^\tau$ may not be the same.

Since the proof is technical, and its main points are described in the paragraph above, we defer it to Appendix B.

4.2 Global Structure for Sets

Now we are ready to prove Theorem 5. The proof is very similar to lemma 3.16 in [9].

Proof. Fix a function $f : \binom{[N]}{k} \to [M]^k$ that passes Test 4 with probability $\epsilon > e^{-c\lambda k}$, denote by $\delta = \frac{c}{5}$ and $\alpha = 5\lambda$.

Let $W \in \binom{[N]}{\frac{k}{10}}, X \in \binom{[N]}{\frac{9k}{10}}$ be the subsets chosen on the first two items of the test, if $\Pr_Y[f(X \cup W)_W = f(Y \cup W)_W \mid Y \cap W = \emptyset] < \frac{\epsilon}{2}\eta$, the test rejects in Item 3 with probability at least $1 - \frac{\epsilon}{2}\eta$.

Therefore, in order for f to pass the test with probability ϵ, the test must pass with probability at least ϵ on W, X such that $\Pr_Y[f(X \cup W)_W = f(Y \cup W)_W \mid Y \cap W = \emptyset] > \frac{\epsilon}{2}\eta$, we call these W, X good.

Using Lemma 37, for at least $(1 - 2\epsilon^2 - \frac{k^2}{N})$ of the good W, X there exist a function $g_{W,X} : [N] \to [M]$ such that

$$\Pr_Y\left[f(Y \cup W)_Y \overset{3\alpha k}{\not\approx} g_{X,W}(Y) \;\middle|\; Y \cap W = \emptyset, f(X \cup W)_W = f(Y \cup W)_W\right] \leq 2\epsilon^2.$$

Fix such W, X, let $G = \left\{Y \in \binom{[N]}{\frac{9k}{10}} \;\middle|\; Y \cap W = \emptyset, f(X \cup W)_W = f(Y \cup W)_W\right\}$, and let $g = g_{X,W} : [N] \to [M]$. We want to use the last query to show that this g is in fact a global product function, i.e $f(S) \approx g(S)$ for about an ϵ fraction of $S \in \binom{[N]}{k}$.

For every set S, we say that S is bad if $f(S) \overset{5\alpha k}{\not\approx} g(S)$. Let p be the probability of a uniform S to be bad, i.e. $p = \Pr_{S \in \binom{[N]}{k}}\left[f(S) \overset{5\alpha k}{\not\approx} g(S)\right]$.

Suppose that instead of running Test 4 as is, we choose Y, V by the following process:

1. Choose a uniform $S \in \binom{[N]}{k}$.
2. Choose Y to be a uniform $\frac{9k}{10}$ subset of S.
3. Set $V = S \setminus Y$ and return (Y, V).

We suppose that if the process outputs Y such that $Y \cap W \neq \emptyset$, the test rejects. The probability of this event is less than $\frac{k^2}{N}$, and if it doesn't happen the process generates the test distribution. Therefore, the test on f using this distribution should success with probability at least $\epsilon - \frac{k^2}{N}$.

In order for Test 4 to pass, two checks must hold:

1. $f(X \cup W)_W = f(Y \cup W)_W$, equivalent to $Y \in G$.
2. $f(Y \cup V)_Y = f(Y \cup W)_Y$.

Suppose that S is bad, and we let $Y \cup V = S$ to be the sets used in the test. If $Y \notin G$, the test will fail. If $y \in G$, from the local structure, Lemma 37,

$$\Pr_Y\left[f(Y \cup W)_Y \overset{3\alpha k}{\not\approx} g(Y) \;\middle|\; Y \in G\right] \leq 2\epsilon^2.$$

If we condition on S to be bad, we restrict Y and therefore the probability of this event can increase by a factor of $\frac{1}{p}$.

$$\Pr_Y\left[f(Y \cup W)_Y \overset{3\alpha k}{\not\approx} g(Y) \;\middle|\; Y \in G, S \text{ is bad}\right] \leq \frac{1}{p} 2\epsilon^2. \tag{23}$$

Since S is bad and Y is a uniform $\frac{9k}{10}$ sized set inside S, the probability that less than $3\alpha k$ out of the $5\alpha k$ elements in which $f(S), g(S)$ differ is in Y is exponentially small.

$$\Pr_{Y\subset S}\left[f(S)_Y \overset{3\alpha k}{\approx} g(Y) \;\middle|\; S \text{ is bad}\right] \leq e^{-\frac{1}{320}\alpha k}. \tag{24}$$

The inequality is due to Chernoff bound, using Claim 11 (if D is the set of elements in which $f(S), g(S)$ differ, $f(S)_Y \overset{3\alpha k}{\approx} g(Y) \implies |Y \cap D| \leq \frac{3}{5}|D|$, in the claim we use $A = S \setminus Y$).

From equation (23), we know that with probability $1 - \frac{2\epsilon^2}{p}$, $f(Y \cup W)_Y \overset{3\alpha k}{\approx} g(Y)$. From (24), with probability $1 - e^{-\frac{1}{320}\alpha k}$, $f(S)_Y \overset{3\alpha k}{\not\approx} g(Y)$. If both holds, then $f(S)_Y = f(Y \cup V)_Y \neq f(Y \cup W)_Y$, and the test will fail. Therefore,

$$\Pr_S\left[\text{Test passes} \mid S \text{ is bad}\right] \leq e^{-\frac{1}{320}\alpha k} + \frac{2\epsilon^2}{p} \leq \frac{3\epsilon^2}{p}. \tag{25}$$

The test must pass with probability $\epsilon - \frac{k^2}{N}$,

$$\begin{aligned}\epsilon - \frac{k^2}{N} = \Pr[\text{Test passes}] =& \Pr[S \text{ is bad}] \Pr\left[\text{Test passes} \mid S \text{ is bad}\right] \\ &+ \Pr[S \text{ isn't bad}] \Pr\left[\text{Test passes} \mid S \text{ isn't bad}\right] \\ \leq& p\frac{3\epsilon^2}{p} + (1-p)\end{aligned}$$

Therefore $p = \Pr[S \text{ is bad}] \leq 1 - \epsilon + \frac{k^2}{N} + 3\epsilon^2$, which implies that at least $\epsilon - \frac{k^2}{N} - 3\epsilon^2$ of the test S are not bad, and for such sets $f(S) \overset{5\alpha k}{\approx} g(S)$. We choose $c = \frac{\delta}{5}$ so $\alpha = \frac{1}{5}\lambda$, and notice that $\epsilon - 4\epsilon^2 \leq \epsilon - \frac{k^2}{N} - 3\epsilon^2$ which finishes the proof. ◀

In the introduction, we stressed that in order to extend the restricted global structure into a global structure, the restricted global structure theorem has to be "strong", i.e. the probability of $f(Y \cup W)_Y \overset{3\alpha k}{\not\approx} g_{X,W}(Y)$ should be strictly smaller than ϵ, it is $2\epsilon^2$ in our case. If the local structure was not strong, the bound in (25) would have been larger than ϵ. This means that all the success probability of the test could come from bad sets S. From (25), we see that almost all of the success probability of the test comes from sets that are not bad, this we couldn't have deduced from the restricted structure theorem of [7].

5 Global Structure for Tuples

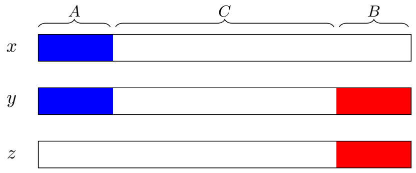
In this section we prove our main theorem – global structure for tuples. The proof uses the restricted global structure, Theorem 21. For convenience we copy the test and theorem from the introduction.

▶ **Theorem 1** (Main theorem – Global Structure for tuples). *For every $N, M > 1$, there exist small constants $c_1, c_2 > 0$ such that for every constant $\lambda > 0$ and large enough k, if $f : [N]^k \to [M]^k$ is a function that passes Test 1 with probability $\alpha_{Z(\frac{k}{10})}(f) = \epsilon \geq e^{-c_1\lambda^2 k}$, then there exist functions $(g_1, \ldots g_k)$, $g_i : [N] \to [M]$ such that*

$$\Pr_{x\in[N]^k}\left[f(x) \overset{\lambda k}{\approx} (g_1(x_1) \ldots g_k(x_k))\right] \geq c_2 \cdot \epsilon.$$

Where $\overset{\lambda k}{\approx}$ means that the strings are equal on all but at most λk coordinates.

1. Choose A, B, C to be a random partition of $[k]$, such that $|A| = |B| = t$.
2. Choose uniformly at random $x, y, z \in [N]^k$ such that $x_A = y_A$ and $y_B = z_B$.
3. Reject if $f(x)_A \neq f(y)_A$ or $f(z)_B \neq f(y)_B$, else accept.



Denote by $\alpha_{Z(t)}(f)$ the success probability of f on this test.

Test 1 "Z"-test with parameter t (3-query test).

5.1 Proof Outline

Our proof of Theorem 1 relies on Theorem 21, which gives us, for many restrictions τ, a product function g^τ that is defined on a set A of $\frac{9k}{10}$ coordinates and approximately equals f on $\mathcal{V}_\tau$. In this section we show how to stitch the restricted functions g^τ together into one global function g. The proof has three parts.

1. In Section 5.2, we show that there exist an $x \in [N]^k$ such that,
   - **a.** On at least $\Omega(\epsilon)$ of the sets A, the tuple $\tau = (A, x_A, f(x)_A)$ is excellent, and Test 1 passes with probability at least $\frac{\epsilon}{3}$.
   - **b.** Taking two such sets A_1, A_2, their functions g^{τ_1}, g^{τ_2} are similar with probability $\Omega(\epsilon^2)$.

   We start from picking x such that the test succeeds on it with probability $\Omega(\epsilon)$, and that for $\Omega(\epsilon)$ of the sets A, $\tau = (A, x_A, f(x)_A)$ satisfies the first item above. We use the third query of the test to show that each g^τ approximates f on $[k] \setminus A$ in $\Omega(\epsilon)$ of the inputs in $[N]^k$. This implies that for many different pairs τ_1, τ_2, both g^{τ_1} and g^{τ_2} are close to f on $[k] \setminus \{A_1 \cup A_2\}$ in $\Omega(\epsilon^2)$ of the inputs, which means that g^{τ_1}, g^{τ_2} are similar to each other.
2. In Section 5.3 we view this situation abstractly as *yet another* agreement question, with a different setting of parameters: given a set of direct product functions, each defined on $\frac{9}{10}k$ coordinates, such that each two are consistent with probability $\Omega(\epsilon^2)$, find a global direct product function $g = (g_1, \ldots g_k)$ that is consistent with $\Omega(\epsilon^2)$ of these functions. We show that such a g can be found, essentially proving an agreement testing theorem for this setting. This may seem circular but in fact the current setting is easier than our original problem because of the density: Since the sets are so large, every two sets intersect.

   In order to solve this agreement question, we build a graph with the functions as nodes, and connect by an edge each two consistent functions. We connect by a "weak edge" each two functions that are somewhat consistent, where we allow a larger difference between the two functions. The weak and strong edges have an "almost transitive" property, if (v_1, v_2) and (v_2, v_3) are connected by a strong edge, then almost surely (v_1, v_3) are connected by a weak edge. We use this property to show that there exist a set of vertices C of size $\Omega(\epsilon^2)$ that is almost a clique, i.e. almost every two functions in C are consistent. We build the global function by taking the plurality over C, and show that it is close to most functions in C.
3. Lastly, in Section 5.4, we connect the two previous items. The functions g^τ for $\tau = (A, x, f(x)_A)$ from the first item are each defined on $\frac{9k}{10}$ coordinates, and each two are similar with probability $\Omega(\epsilon^2)$. This means that they satisfies the conditions of the second item, and there exists a global function g defined on all $[k]$ that is close to $\Omega(\epsilon^2)$ of them. We recall that on Section 5.2 we showed that each g^τ is close to f on $\Omega(\epsilon)$ fraction of the inputs, and conclude that the global function g is also close to f on $\Omega(\epsilon)$ fraction of the input, which finishes the proof.

5.2 Consistency Between Restricted Global Functions

From Theorem 21, we know with probability $1-\epsilon^2$ a good $\tau \sim \mathcal{D}$ is excellent, and for each excellent τ there exist a local direct product function $g^\tau = (g_1^\tau, \dots, g_{\frac{9k}{10}}^\tau)$ that equals f on $\mathcal{V}_\tau$.

▶ **Definition 38.** For every $x \in [N]^k$, let $\mathcal{A}_x$ be the set of subsets $A \subset [k]$ of size $\frac{k}{10}$ such that:

1. Fixing A, x, $\Pr_{y,B,z}$ [Test 1 passed] $\geq \frac{\epsilon}{3}$.
2. $\tau = (A, x_A, f(x)_A)$ is excellent.

▶ **Definition 39.** Let $\tau_1 = (A_1, r_1, \gamma_1), \tau_2 = (A_2, r_2, \gamma_2)$ be two excellent tuples, we say that g^{τ_1}, g^{τ_1} are consistent if for a uniform $i \notin A_1 \cup A_2$ and $u \in [N]$,

$$\Pr_{i,u}\left[g_i^{\tau_1}(u) \neq g_i^{\tau_2}(u)\right] \leq 60\lambda.$$

The main claim we prove in this section is the following,

▶ **Claim 40.** *There exist $x \in [N]^k$, such that:*

1. $\Pr_A\left[A \in \mathcal{A}_x\right] \geq \frac{\epsilon}{4}$.
2. $\Pr_{A_1,A_2 \in \mathcal{A}_x}\left[g^{\tau_1}, g^{\tau_1}\ \textit{are consistent}\right] \geq \frac{\epsilon^2}{32}$, *where the tuples are* $\tau_1 = (A_1, x_{A_1}, f(x)_{A_1})$ *and* $\tau_2 = (A_2, x_{A_2}, f(x)_{A_2})$.

We start from looking for a candidate $x \in [N]^k$.

▶ **Claim 41.** *Let*

$$\mathcal{X}_1 = \left\{x \in [N]^k \;\middle|\; \Pr\left[\textit{Test 1 passed with } x\right] \geq \frac{\epsilon}{4}\right\},$$

$$\mathcal{X}_2 = \left\{x \in [N]^k \;\middle|\; \Pr_A\left[A \in \mathcal{A}_x\right] \geq \frac{\epsilon}{8}\right\}.$$

Then

$$\mathcal{X}_1 \cap \mathcal{X}_2 \neq \emptyset.$$

Proof. Let G be the full weighted bipartite graph, with vertex sets $L = \binom{[k]}{\frac{9k}{10}}$ and $R = [N]^k$. The weight of an edge A, x equals the success probability of Test 1 given that A, x are chosen.

The expected weight of an edge is equal to the test success probability of Test 1, ϵ. For each edges with weight less than $\frac{\epsilon}{2}$, we change its weight to 0. We removed at most half of the total weight, so the expected weight of a uniform edge now is at least $\frac{\epsilon}{2}$.

All the edges that remain with positive weight are of (A, x) such that $\tau = (A, x, f(x)_A)$ is good (there may also be good tuples with weight 0, if Test 2 passed with probability larger than $\frac{\epsilon}{2}$ but Test 1 didn't). We further change to 0 the weight of all the edges A, x such that $\tau = (A, x_A, f(x)_A)$ is not excellent.

From Theorem 21, a random good $\tau \sim \mathcal{D}$ is excellent with probability $1-\epsilon^2$, and the distribution $\tau \sim \mathcal{D}$ corresponds to a uniform choice of $A \in L, x \in R$. Therefore, changing to 0 the wight over these edges means changing to 0 the weight of at most ϵ^2 of the edges in G. The maximal weight of an edge is 1, we have reduced the expected weight by at most ϵ^2. The expected weight now is more than $\frac{\epsilon}{2} - \epsilon^2 \geq \frac{\epsilon}{3}$.

Let x be the vertex with the maximal sum of weights of neighbor edges, then

$$\Pr\left[\text{Test 1 passed given } x\right] \geq \mathop{\mathbb{E}}_A\left[\omega(A, x)\right] \geq \frac{\epsilon}{3}.$$

The inequality is because we have changed to zero the weight some edges.

All edges (A,x) that still have positive weight satisfy $A \in \mathcal{A}_x$,

$$\Pr_A [A \in \mathcal{A}_x] = \Pr_A [\omega(A,x) > 0] \geq \frac{\epsilon}{3},$$

since the maximal weight an edge can have is 1.

Therefore, $x \in \mathcal{X}_1 \cap \mathcal{X}_2$. ◀

In the rest of this section we fix $x \in \mathcal{X}_1 \cap \mathcal{X}_2$, denote $\mathcal{A} = \mathcal{A}_x$ and $g_A = g^\tau = (g_1^\tau, \dots, g_{\frac{9k}{10}}^\tau)$ for $\tau = (A, x, f(x)_A)$, and prove that it fulfills the conditions of Claim 40.

▶ **Definition 42.** An input $z \in [N]^k$ is consistent with a set $A \in \mathcal{A}$ if $f(z)_{\bar{A}} \overset{20\lambda k}{\approx} g_A(z_{\bar{A}})$. Let $\mathcal{Z}_A$ be the set of inputs that are consistent with A.

$$\mathcal{Z}_A = \left\{ z \in [N]^k \;\middle|\; f(z)_{\bar{A}} \overset{20\lambda k}{\approx} g_A(z_{\bar{A}}) \right\}.$$

▶ **Claim 43.** *For every $A \in \mathcal{A}$, $\Pr_z [z \in \mathcal{Z}_A] \geq \frac{\epsilon}{4}$.*

Proof. Assume towards contradiction that the claim does not hold, and fix a set $A \in \mathcal{A}$ such that $\Pr_z [z \in \mathcal{Z}_A] < \frac{\epsilon}{4}$.

We reach a contradiction by showing that conditioning on A, x chosen by the test, $\Pr_{y,B,z} [\text{Test 1 passes}] < \frac{\epsilon}{3}$ contradicting the fact that $A \in \mathcal{A}$.

We define the following events, under the assumption that $y_A = x_A$ and $y_B = z_B$ as in the test.

1. E_1: $f(x)_A = f(y)_A$.
2. E_2: $f(z)_B = f(y)_B$.
3. E_3: $f(y)_B \overset{\lambda k}{\approx} g_A(y_B)$.
4. E_4: $f(z)_B \overset{\lambda k}{\approx} g_A(z_B)$.
5. E_5: $z \notin \mathcal{Z}_A$.

Note that since g_A is a product function and $y_B = z_B$, E_4 can also be written as $f(z)_B \overset{\lambda k}{\approx} g_A(y_B)$. We also notice that $E_2 \wedge E_3 \implies E_4$, since we can switch $f(y)_B$ by $f(z)_B$ in E_3.

By definition, Test 1 succeeds if E_1, E_2 both happened.

$$\begin{aligned}
\Pr_{y,B} [E_1 \wedge E_2] &\leq \Pr_{y,B,z} [E_1 \wedge E_2 \wedge E_3 \wedge E_5] + \Pr_{y,B,z} [E_1 \wedge E_2 \wedge \neg E_3] + \Pr_{y,B,z} [E_1 \wedge E_2 \wedge \neg E_5] \\
&\leq \Pr_{y,B,z} [E_1 \wedge E_4 \wedge E_5] + \Pr_{y,B,z} [E_1 \wedge E_2 \wedge \neg E_3] + \Pr_{y,B,z} [E_1 \wedge E_2 \wedge \neg E_5] \\
&\leq \Pr_{y,B,z} [E_4 \mid E_5] + \Pr_{y,B,z} [\neg E_3 \mid E_1] + \Pr_{y,B,z} [\neg E_5] . \qquad (26)
\end{aligned}$$

We bound each of the three probabilities.

1. If E_5 happened, $z \notin \mathcal{Z}_A$ so $f(z)_{\bar{A}} \overset{20\lambda k}{\not\approx} g_A(z_{\bar{A}})$, let D be the set of coordinates in which $f(z)_{\bar{A}}$ and $g_A(z_{\bar{A}})$ differ

   $$D = \left\{ i \in \bar{A} \;\middle|\; f(z)_i \neq g_{A,i}(z_i) \right\}.$$

   In order to satisfy E_4, the set B should be such that $|B \cap D| \leq \lambda k$, since B is a random set of size $\frac{k}{10}$, using Claim 11

   $$\Pr_{B,y,z} [E_4 \mid E_5] \leq \Pr_B [|B \cap D| \leq \lambda k] \leq e^{-\frac{\lambda k}{60}} < \epsilon^2.$$

2. Since $A \in \mathcal{A}_x$ the tuple $(A, x_A, f(x)_A)$ is excellent, and from Theorem 21

$$\Pr_{y,B,z}[\neg E_3 \mid E_1] = \Pr_{y,B}\left[f(y)_B \overset{\lambda k}{\not\approx} g_A(y_B) \;\middle|\; f(x)_A = f(y)_A \right] \le \epsilon^2,$$

where we use the fact that $B \subseteq \bar{A}$, therefore $f(y)_B \overset{\lambda k}{\not\approx} g_A(y_B)$ implies $f(y)_{\bar{A}} \overset{\lambda k}{\not\approx} g_A(y_{\bar{A}})$.

3. From our assumption,

$$\Pr_{y,B,z}[E_5] = \Pr_z[z \in \mathcal{Z}_A] \le \frac{\epsilon}{4}.$$

Therefore, from (26) we get

$$\Pr_{y,B,z}[\text{Test 1 passes} \mid x, A] = \Pr_{y,B,z}[E_1 \wedge E_2] \le \epsilon^2 + \epsilon^2 + \frac{\epsilon}{4} < \frac{\epsilon}{3},$$

contradicting $A \in \mathcal{A}$. ◀

In the introduction, we explained the difference between our restricted global structure, and the result of [7]. In our result, Theorem 21, $f(y)_{\bar{A}} \approx g^\tau(y)$ for $1-\epsilon^2$ of $y \in \mathcal{V}_\tau$, and it their result it was much less.

▶ **Claim 44.**

$$\Pr_{A_1,A_2 \in \mathcal{A}}\left[|\mathcal{Z}_{A_1} \cap \mathcal{Z}_{A_2}| \ge \frac{\epsilon^2}{32} N^k \right] \ge \frac{\epsilon^2}{32}.$$

Proof. For a uniform pair $A_1, A_2 \in \mathcal{A}$:

$$\mathop{\mathbb{E}}_{A_1,A_2}[|\mathcal{Z}_{A_1} \cap \mathcal{Z}_{A_2}|] = \sum_z \mathop{\mathbb{E}}_{A_1,A_2}[\mathbb{I}(z \in \mathcal{Z}_{A_1} \cap \mathcal{Z}_{A_2})] \ge \sum_z \Pr_{A_1}[z \in \mathcal{Z}_{A_1}]^2 \tag{27}$$

where $\mathbb{I}$ is an indicator. The last inequality holds since A_1, A_2 are independent uniform sets in $\mathcal{A}$, and the square function is convex.

From Claim 43, $\Pr_z[z \in \mathcal{Z}_A] \ge \frac{\epsilon}{4}$ for every $A \in \mathcal{A}$. Therefore, from (27) we get

$$\begin{aligned}\mathop{\mathbb{E}}_{A_1,A_2}[|\mathcal{Z}_{A_1} \cap \mathcal{Z}_{A_2}|] &\ge \sum_z \Pr_{A_1}[z \in \mathcal{Z}_{A_1}]^2 \\ &\ge \left(\sum_z N^{-\frac{k}{2}} \Pr_{A_1}[z \in \mathcal{Z}_{A_1}] \right)^2 && \text{(Cauchy Swartz)} \\ &\ge \left(\frac{\epsilon}{4}\right)^2 N^k. && \text{(Claim 43)}\end{aligned}$$

The maximal value of $|\mathcal{Z}_{A_1} \cap \mathcal{Z}_{A_2}|$ is N^k, therefore by averaging

$$\Pr_{A_1,A_2}\left[|\mathcal{Z}_{A_1} \cap \mathcal{Z}_{A_2}| \ge \frac{\epsilon^2}{32} N^k \right] \ge \frac{\epsilon^2}{32}.$$ ◀

▶ **Claim 45.** *If $A_1, A_2 \in \mathcal{A}$ are such that $|\mathcal{Z}_{A_1} \cap \mathcal{Z}_{A_2}| \ge \frac{\epsilon^2}{32} N^k$, then g_{A_1}, g_{A_2} are consistent, i.e. for a uniform $i \in [k] \setminus \{A_1 \cup A_2\}$ and $u \in [N]$,*

$$\Pr_{i,u}[g_{A_1,i}(u) \ne g_{A_2,i}(u)] \le 60\lambda.$$

Proof. Let $A_1, A_2 \in \mathcal{A}$ be two sets such that $|\mathcal{Z}_{A_1} \cap \mathcal{Z}_{A_2}| \geq \frac{\epsilon^2}{32} N^k$, and let $\mathcal{Z}_{12} = \mathcal{Z}_{A_1} \cap \mathcal{Z}_{A_2}$. In order to simplify the notation, denote $S_1 = [k] \setminus A_1$, $S_2 = [k] \setminus A_2$ and $S_{12} = S_1 \cap S_2 = [k] \setminus \{A_1 \cup A_2\}$. S_{12} is the set of coordinates that both g_{A_1}, g_{A_2} are defined on, $|S_{12}| \geq 0.8k$.

For each $i \in S_{12}$, let

$$p_i = \Pr_{u \in [N]} \left[g_{A_1,i}(u) \neq g_{A_2,i}(u) \right].$$

Let $w \in [N]^{S_{12}}$ uniformly at random, and let I_i be indicator random variable for $g_{A_1,i}(w_i) \neq g_{A_2,i}(w_i)$. Each I_i equals 1 with probability p_i independently. In this notation

$$\mathop{\mathbb{E}}_{w} \left[\mathrm{dist}(g_{A_1}(w), g_{A_2}(w)) \right] = \mathbb{E} \left[\sum_{i \in S_{12}} I_i \right].$$

Assume towards contradiction that $\Pr_{i,b} \left[g_{A_1,i}(u) \neq g_{A_2,i}(u) \right] > 60\lambda$, this will imply that $\mathbb{E}_w \left[\mathrm{dist}(g_{A_1}(w), g_{A_2}(w)) \right] > 60\lambda \cdot 0.8k$.

Using Chernoff bound:

$$\Pr_{w} \left[\mathrm{dist}(g_{A_1}(w), g_{A_2}(w)) \leq 40\lambda k \right] = \Pr \left[\sum_{i \in S_{12}} I_i \leq \frac{5}{6} \mathbb{E} \left[\sum_{i \in S_{12}} I_i \right] \right] \leq e^{-\frac{1}{2}\lambda k}.$$

If instead of taking a completely uniform $w \in [N]^{S_{12}}$, we pick a random $z \in \mathcal{Z}_{12}$, and restrict it to S_{12}, getting $w = z_{S_{12}}$. The probability of any event on w can increase by a factor of at most $\frac{N^k}{|\mathcal{Z}_{12}|} \leq \frac{32}{\epsilon^2}$,

$$\Pr_{z \in \mathcal{Z}_{12}} \left[\mathrm{dist}(g_{A_1}(z_{S_{12}}), g_{A_2}(z_{S_{12}})) \leq 40\lambda k \right] \leq \frac{32}{\epsilon^2} e^{-\frac{1}{2}\lambda k} < \frac{1}{2}. \tag{28}$$

By the definition of $\mathcal{Z}_{A_1}, \mathcal{Z}_{A_2}$, each input $z \in \mathcal{Z}_{12}$ satisfies both $f(z)_{S_1} \overset{20\lambda k}{\approx} g_{A_1}(z_{S_1})$ and $f(z)_{S_2} \overset{20\lambda k}{\approx} g_{A_2}(z_{S_2})$ which implies $g_{A_1}(z_{S_{12}}) \overset{40\lambda k}{\approx} g_{A_2}(z_{S_{12}})$ with probability 1, which contradicts (28). ◀

Combining the last two claims, we prove Claim 40.

5.3 Agreement Theorem in the Dense Case

In this section, we present and prove an abstract problem that will later be used to create the global product function. Given a collection of local functions $\mathcal{F} = \{f_S\}_{S \in \binom{[k]}{\frac{9k}{10}}}$, such that for each $S \in \binom{[k]}{\frac{9k}{10}}$, $f_S : S \to \Sigma$, can we deduce from the agreement of f_S the existence of a single global function $g : [k] \to \Sigma$ that is close to many f_S?

We need to define what exactly agreement means in the case of $\mathcal{F}$, as it is not the setting on which we previously defined agreement on. In order to do so, we assume that we have a bounded distance measure on Σ, i.e. for every $\sigma_1, \sigma_2 \in \Sigma$, $\mathrm{dist}(\sigma_1, \sigma_2) \in [0, 1]$.

▶ **Definition 46.** The difference between $f_{S_1}, f_{S_2} \in \mathcal{F}$, denoted by $\Delta(f_{S_1}, f_{S_2})$ is defined by

$$\Delta(f_{S_1}, f_{S_2}) = \mathop{\mathbb{E}}_{i \in S_1 \cap S_2} \left[\mathrm{dist}(f_{S_1}(i), f_{S_1}(i)) \right].$$

The difference between $f_S \in \mathcal{F}$ to a function $g : [k] \to \Sigma$ is defined by,

$$\Delta(f_S, g) = \mathop{\mathbb{E}}_{i \in S} \left[\mathrm{dist}(f_S(i), g(i)) \right].$$

Note that the difference defined above is not a distance, it may be that $\Delta(f_{S_1}, f_{S_2}) = 0$ for $S_1 \neq S_2$.

Now we are ready to define the agreement, notice that since we are talking on an agreement inside a function set $\mathcal{F}$, the definition is different. The general idea is the same – we check for the agreement of two random elements in $\mathcal{F}$ according to some distribution.

▶ **Definition 47.** The agreement of the collection of local functions $\mathcal{F}$ regarding the uniform distribution with parameter α, denoted by $\text{agree}_\alpha(\mathcal{F})$ is defined by,

$$\text{agree}_\alpha(\mathcal{F}) = \Pr_{f_{S_1}, f_{S_2} \in \mathcal{F}} [\Delta(f_{S_1}, f_{S_2}) < \alpha].$$

▶ **Theorem 48.** *For every small constant $\alpha \in (0,1)$ and $\nu > e^{-\frac{1}{3}\alpha^2 k}$, if a collection of local functions $\mathcal{F}$ has $\text{agree}_\alpha(\mathcal{F}) > \nu$, then there exists a global function $g : [k] \to \Sigma$ such that*

$$\Pr_{S \in \binom{[k]}{\frac{9k}{10}}} [\Delta(f_S, g) \leq 300\alpha] \geq \frac{1}{4}\nu.$$

In order to prove the theorem, it is helpful to look at the elements $S \in \binom{[k]}{\frac{9k}{10}}$ as vertices in a graph. Let $\mathcal{G} = (\mathcal{V}, E_S \cup E_W)$ to be the graph with the vertex set $\mathcal{V} = \binom{[k]}{\frac{9k}{10}}$, and two edge sets, weak edges and strong edges.

▶ **Definition 49.** For every two sets $S_1, S_2 \in \mathcal{V}$:

1. S_1, S_2 are connected by a strong edge, denoted by $S_1 - S_2$, if $\Delta(f_{S_1}, f_{S_2}) < \alpha$.
2. S_1, S_2 are connected by a weak edge, denoted by $S_1 \sim S_2$, if $\Delta(f_{S_1}, f_{S_2}) < 60\alpha$.

We want to find a subset of vertices that is close to a clique in $\mathcal{G}$, such subset will allow us to define a global function g. We start by showing that there exist many vertices of high degree in $\mathcal{G}$.

▶ **Claim 50.** *Exists a set $\mathcal{S} \subset \mathcal{V}$ of measure at least $\frac{\nu}{2}$, such that for every $S \in \mathcal{S}$*

$$\Pr_{S' \in \mathcal{V}} [S - S'] \geq \frac{1}{2}\nu.$$

Proof. Let

$$\mathcal{S} = \left\{ S \subseteq \mathcal{V} \;\middle|\; \Pr_{S'} [S - S'] \geq \frac{1}{2}\nu \right\}.$$

By averaging

$$\begin{aligned}
\nu &\leq \Pr_{S_1, S_2} [S_1 - S_2] \\
&\leq \Pr_{S_1} [S_1 \in \mathcal{S}] \Pr_{S_1, S_2} [S_1 - S_2 \mid S_1 \in \mathcal{S}] + \Pr_{S_1} [S_1 \notin \mathcal{S}] \Pr_{S_1, S_2} [S_1 - S_2 \mid S_1 \notin \mathcal{S}] \\
&\leq \Pr_{S_1} [S_1 \in \mathcal{S}] + \frac{1}{2}\nu \left(1 - \Pr_{S_1} [S_1 \in \mathcal{S}]\right).
\end{aligned}$$

Then $\Pr_{S_1} [S_1 \in \mathcal{S}] \geq \frac{1}{2}\nu$. ◀

Strong connectivity is not transitive, but we can have an “almost transitive” property by considering both strong and weak edges.

▶ **Claim 51.** *For $S, S_1, S_2 \in \mathcal{V}$ uniformly and independently,*

$$\Pr_{S,S_1,S_2}[S - S_1, S - S_2, S_1 \not\sim S_2] \le 2e^{-\alpha^2 k}.$$

Proof. Fix $S_1, S_2 \in \mathcal{V}$ to be two vertices such that $S_1 \not\sim S_2$ (if there are no such vertices, the probability is 0 and we are done).

For every $S \in \mathcal{V}$, we define by d_i, d_i^1, d_i^2 the following distances:

1. For each $i \in S_1 \cap S_2$, $d_i = \mathrm{dist}(f_{S_1}(i), f_{S_2}(i))$.
2. For each $i \in S \cap S_1$, $d_i^1 = \mathrm{dist}(f_S(i), f_{S_1}(i))$.
3. For each $i \in S \cap S_2$, $d_i^2 = \mathrm{dist}(f_S(i), f_{S_2}(i))$.

By the triangle inequality, for every $i \in S \cap S_1 \cap S_2$, $d_i \le d_i^1 + d_i^2$, therefore for every such i, $\max\{d_i^1, d_i^2\} \ge \frac{d_i}{2}$.

Since $S_1 \not\sim S_2$, we know that $\mathbb{E}_{i \in S_1 \cap S_2}[d_i] \ge 60\alpha$, if we look on the sum $\sum_{i \in S_1 \cap S_2} d_i \ge \frac{8k}{10} 60\alpha$ (because $|S_1 \cap S_2| \ge \frac{8}{10}k$). If $S - S_1, S - S_2$, then $\max\{\mathbb{E}_{i \in S \cap S_1}[d_i^1], \mathbb{E}_{i \in S \cap S_2}[d_i^2]\} \le \alpha$, which means that $\max\{\sum_{i \in S \cap S_1} d_i^1, \sum_{i \in S \cap S_2} d_i^2\} \le \frac{9}{10}\alpha k$ (we switched expectation in a sum, $|S \cap S_1| \le \frac{9k}{10}$).

$$\max\left\{\sum_{i \in S \cap S_1} d_i^1, \sum_{i \in S \cap S_2} d_i^2\right\} \ge \frac{1}{2}\sum_{i \in S \cap S_1 \cap S_2} \max\left\{d_i^1, d_i^2\right\} \ge \frac{1}{4}\sum_{i \in S \cap S_1 \cap S_2} d_i \tag{29}$$

The first inequality is since taking the maximum over every i can increase the total sum in a factor of 2 at most from taking maximum of the sum. The second inequality is since $\max\{d_i^1, d_i^2\} \ge \frac{d_i}{2}$.

Notice that the last expression is independent of the function f_S, and depends only on the set S. Let X_S be the random variable $X_S = \frac{1}{4}\sum_{i \in S \cap S_1 \cap S_2} d_i$ for a uniform $S \in \mathcal{V}$. Since the set S is a uniform $\frac{9k}{10}$ sized subset of $[k]$, $\mathbb{E}_S[X_S] = \frac{9}{10}\sum_{i \in S_1 \cap S_2} d_i \ge \frac{9}{10}\frac{8}{10}60\alpha k$. For $S \in \mathcal{V}$ such that $X_S > \alpha k$, by (29) it means that S is not strongly connected to one of S_1, S_2.

To finish the proof, we need to show that $X_S \le \alpha k$ for very few $S \in \mathcal{V}$. Let D contain the $\frac{k}{3}$ indices $i \in S_1 \cap S_2$ with the largest d_i. Obviously $\sum_{i \in D} d_i \ge \frac{1}{3}\sum_{i \in S_1 \cap S_2} d_i \ge 16\alpha k$. In order of $X_S \le \alpha k$, the sum over $i \in D \cap S$ should satisfy, $\sum_{i \in D \cap S} d_i \le 4\alpha k$. By Claim 52, this happens with probability less than $2e^{-\alpha^2 k}$.

Therefore, for every $S_1 \not\sim S_2$, the probability of a uniform $S \in \mathcal{V}$ to be strongly connected to both is at most $2e^{-\alpha^2 k}$. This is true for every $S_1 \not\sim S_2$, it is also true for a random pair. ◀

▶ **Claim 52** (Fixed sized Chernoff bound). *For every constant $\alpha \in (0,1)$, let $k \in \mathbb{N}$ be a large enough integer, $D \subset [k]$ a subset of size at most $\frac{k}{3}$, and for every $i \in D$ let $d_i \in [0,1]$ be constants such that $\sum_{i \in D} d_i > 4\alpha k$.*

Let $S \subset [k]$ be a random subset of size exactly $\frac{9k}{10}$, then

$$\Pr_S\left[\sum_{i \in S \cap D} d_i \le \alpha k\right] \le 2e^{-\alpha^2 k}. \tag{30}$$

Proof. For a set S that is chosen by putting each $i \in [k]$ in S with probability $\frac{9}{10}$ independently, Chernoff bound gives us the required bound easily. Because S has a fixed size, we need to work a little harder.

For each $i \in D$, let S be a uniform set in $\binom{[k]}{\frac{9k}{10}}$, and let I_i be,

$$I_i = \begin{cases} d_i & i \in S \\ 0 & i \notin S \end{cases}.$$

In this notation, $\sum_{i\in S\cap D} d_i = \sum_{i\in D} I_i$. The random variables I_i are not independent, we define the independent random variables J_i,

$$J_i = \begin{cases} d_i & \text{w.p } \frac{1}{2} \\ 0 & \text{w.p } \frac{1}{2} \end{cases}.$$

Since $|D| = \frac{k}{3}$, and S is a uniform $\frac{9}{10}k$ sized subset of $[k]$, even conditioning on all other $j \in D \setminus \{i\}$ to be in S, the probability of i to be in S is at least $\frac{1}{2}$.

$$\Pr\left[I_i = d_i \mid \forall j \in D \setminus \{i\}, I_j > 0\right] \geq \Pr\left[J_i = d_i\right]. \tag{31}$$

So a lower bound for J_i implies a lower bound for I_i.

The random variables J_i satisfies $\mathbb{E}\left[\sum_{i\in D} J_i\right] = \frac{1}{2}\sum_{i\in D} d_i \geq 2\alpha k$.

$$\begin{aligned}\Pr_S\left[\sum_{i\in S\cap D} d_i \leq \alpha k\right] &= \Pr_{I_i}\left[\sum_{i\in D} I_i \leq \alpha k\right] \\ &\leq \Pr_{J_i}\left[\sum_{i\in D} J_i \leq \alpha k\right] \\ &\leq \Pr_{J_i}\left[\left|\sum_{i\in D} J_i - \mathbb{E}\left[\sum_{i\in D} J_i\right]\right| \geq \alpha k\right] && \text{(Chernoff bound)} \\ &\leq 2e^{-\alpha^2 k}. \end{aligned}$$

◀

From the last two claims, Claim 50 and Claim 51, conclude that there is a high degree vertex in $\mathcal{V}$ that its neighbors almost form a clique.

▶ **Claim 53.** *There exists a set $S \in \mathcal{S}$ such that*

$$\Pr_{S_1,S_2\in\mathcal{V}}\left[S_1 \sim S_2 \mid S_1 - S, S_2 - S\right] \geq 1 - \alpha\,.$$

Proof. From Claim 50, we know that if we choose $S, S_1, S_2 \in \mathcal{V}$ independently,

$$\Pr_{S,S_1,S_2}\left[S \in \mathcal{S}, S - S_1, S - S_2\right] \geq \Pr_S\left[S \in \mathcal{S}\right] \Pr_{S,S_1}\left[S - S_1 \mid S \in \mathcal{S}\right]^2 \geq \left(\frac{\nu}{2}\right)^3.$$

From Claim 51, on the same distribution

$$\Pr_{S,S_1,S_2}\left[S - S_1, S - S_2, S_1 \nsim S_2\right] \leq 2e^{-\alpha^2 k}.$$

Therefore

$$\Pr_{S,S_1,S_2}\left[S_1 \nsim S_2 \mid S \in \mathcal{S}, S - S_1, S - S_2\right] \leq \left(\frac{2}{\nu}\right)^3 2e^{-\alpha^2 k} < \alpha\,.$$

The last inequality is since $\log\left(\frac{1}{\nu}\right) \leq \frac{1}{3}\alpha^2 k$.

From averaging, there must be $S \in \mathcal{S}$ that achieves this bound. ◀

Proof of Theorem 48. Let $\tilde{S} \in \mathcal{S}$ be the vertex promised from Claim 53, and denote by C its strong neighbors,

$$C = \left\{ S \in \binom{[k]}{\frac{9k}{10}} \;\middle|\; S - \tilde{S} \right\},$$

since $\tilde{S} \in \mathcal{S}$, the measure of C is at least $\frac{\nu}{2}$. From the claim we also know that $\Pr_{S_1,S_2 \in C}[S_1 \not\sim S_2] \leq \alpha$, so almost every two sets in C have small difference.

The global function $g(i)$ is defined to be $\beta \in \Sigma$ that is closest to $f_S(i)$ over all $S \in C$ that contains i,

$$\forall i \in [k], \quad g(i) = \operatorname*{argmin}_{\beta \in \Sigma} \left\{ \mathop{\mathbb{E}}_{S \in C \text{ s.t. } i \in S} [\mathrm{dist}(f_S(i), \beta)] \right\}.$$

If there is no $S \in C$ such that $i \in S$, we define $g(i)$ to an arbitrary value.

We notice that for every i, by definition

$$\Pr_{S \in C \text{ s.t. } i \in S} [\mathrm{dist}(f_S(i), g(i))] \leq \Pr_{S_1,S_2 \in C \text{ s.t. } i \in S_1,S_2} [\mathrm{dist}(f_{S_1}(i), f_{S_2}(i))]. \tag{32}$$

We know that $S_1, S_2 \in C$ are weakly connected with probability at least $1 - \alpha$, which means that the difference between their functions is small.

$$\begin{aligned} \mathop{\mathbb{E}}_{S_1,S_2 \in C} [\Delta(f_{S_1}, f_{S_2})] \leq & 1 \cdot \Pr_{S_1,S_2 \in C} [S_1 \not\sim S_2] + \mathop{\mathbb{E}}_{S_1,S_2 \in C} [\Delta(f_{S_1}, f_{S_2}) \mid S_1 \sim S_2] \\ \leq & \alpha + 60\alpha \leq 61\alpha. \end{aligned}$$

By the definition of difference, we get that,

$$\begin{aligned} 61\alpha &\geq \mathop{\mathbb{E}}_{S_1,S_2 \in C} [\Delta(f_{S_1}, f_{S_2})] \\ &\geq \mathop{\mathbb{E}}_{S_1,S_2 \in C, i \in S_1 \cap S_2} [\mathrm{dist}(f_{S_1}(i), f_{S_2}(i))]. \end{aligned} \tag{33}$$

Notice that the distribution over i in this expression is not uniform, we define formally the distributions over i that we use.

1. Let $\mathcal{D}_1 : [k] \to [0,1]$ be the distribution that picks $S \in C$ uniformly, then $i \in S$.
2. Let $\mathcal{D}_2 : [k] \to [0,1]$ be the distribution that picks $S_1, S_2 \in C$ uniformly, then $i \in S_1 \cap S_2$ (as $|S_i| = \frac{9k}{10}$ there is always such i).

Using this definition, (33) can also be written as

$$\mathop{\mathbb{E}}_{i \sim \mathcal{D}_2, S_1, S_2 \in C} [\mathrm{dist}(f_{S_1}(i), f_{S_2}(i)) \mid i \in S_1 \cap S_2] \leq 61\alpha. \tag{34}$$

To prove the theorem, we need to prove (34) when $i \sim \mathcal{D}_1$. First, we show that the distributions $\mathcal{D}_1, \mathcal{D}_2$ are close to each other. In order to do so, we define the following set, D,

$$D = \left\{ i \in [k] \;\middle|\; \Pr_{S \in C} [i \in S] < \frac{1}{2} \right\}.$$

By Claim 54, the set D is small $|D| \leq 4\alpha k$. For each $i \notin D$, $\Pr_{S \in C} [i \in S] \in \left[\frac{1}{2}, 1\right]$ which means that for every $i \notin D$,

$$\Pr_{j \sim \mathcal{D}_1} [j = i \mid j \notin D] \leq 2 \Pr_{j \sim \mathcal{D}_2} [j = i \mid j \notin D]. \tag{35}$$

Using (35), (34) and (32), we show that the expected difference between g and f_S for a random $S \in C$ is small,

$$\begin{aligned}
\mathop{\mathbb{E}}_{S\in C}\left[\Delta(f_S, g)\right] &= \mathop{\mathbb{E}}_{S\in C, i\in S}\left[\operatorname{dist}(f(i), g(i))\right] && \text{(by definition of } \mathcal{D}_1\text{)}\\
&= \mathop{\mathbb{E}}_{i\sim\mathcal{D}_1, S\in C}\left[\operatorname{dist}(f_{S_1}(i), g(i)) \mid i \in S\right] && \text{(by (32))}\\
&\le \mathop{\mathbb{E}}_{i\sim\mathcal{D}_1, S_1, S_2\in C}\left[\operatorname{dist}(f_{S_1}(i), f_{S_2}(i)) \mid i \in S_1 \cap S_2\right]\\
&\le \Pr_{i\sim\mathcal{D}_1}\left[i \in D\right] + \mathop{\mathbb{E}}_{i\sim\mathcal{D}_1, S_1, S_2\in C}\left[\operatorname{dist}(f_{S_1}(i), f_{S_2}(i)) \mid i \in S_1 \cap S_2 \setminus D\right]\\
& && \text{(by (35))}\\
&\le 4\alpha + 2 \mathop{\mathbb{E}}_{i\sim\mathcal{D}_2, S_1, S_2\in C}\left[\operatorname{dist}(f_{S_1}(i), f_{S_2}(i)) \mid i \in S_1 \cap S_2 \setminus D\right] && \text{(by (34))}\\
&\le 4\alpha + 2 \cdot \frac{61\alpha}{1-4\alpha} \le 150\alpha. && (36)
\end{aligned}$$

Equation (32) holds for every $i \in [k]$, therefore it holds for expectation over i under any distribution. The last inequality holds because of (34), and because if we condition on $i \notin D$ we can increase the probability by a factor of at most $\Pr_{i\sim\mathcal{D}_2}\left[i \notin D\right]$, which is small.

The only thing left now is a Markov argument, if $\mathbb{E}_{S\in C}\left[\Delta(f_S, g)\right] \le 150\alpha$, then at least half of the sets $S \in C$ satisfies $\Delta(f_S, g) \le 300\alpha$, since the measure of C is $\frac{\nu}{2}$, the measure of half of C is $\frac{\nu}{4}$ and we are done. ◀

▶ **Claim 54.** *Let $C \subset \binom{[k]}{\frac{9k}{10}}$ a subset of fraction size $\frac{\nu}{2}$, then the number of indices $i \in k$ such that $\Pr_{S\in C}\left[i \in S\right] \le \frac{1}{2}$ is at most $4\alpha k$.*

Proof. Let $D \subset [k]$ be this set of indices

$$D = \left\{ i \in [k] \,\middle|\, \Pr_{S\in C}\left[i \in S\right] \le \frac{1}{2} \right\}.$$

If we pick a completely uniform $S' \in \binom{[k]}{\frac{9}{10}k}$,

$$\mathop{\mathbb{E}}_{S'}\left[|S' \cap D|\right] = \frac{9}{10}|D|.$$

From Chernoff, using Claim 11 with $A = [k] \setminus S'$, (if $|D| \ge \frac{k}{3}$, the probability is even smaller)

$$\Pr_{S'}\left[|S' \cap D| \le \frac{2}{3}|D|\right] \le e^{-\frac{|D|}{45}}.$$

If we pick a uniform subset in $S \in C$, instead of a completely uniform set:

$$\Pr_{S\in C}\left[|S \cap D| \le \frac{2}{3}|D|\right] \le \frac{2}{\nu}e^{-\frac{|D|}{45}}.$$

From the definition of D, for each $i \in D$, $\Pr_{S\in C}\left[i \in S\right] \le \frac{1}{2}$, so of course

$$\mathop{\mathbb{E}}_{S\in C}\left[|S \cap D|\right] \le \frac{1}{2}|D|.$$

From averaging

$$\Pr_{S\in C}\left[|S \cap D| \le \frac{2}{3}|D|\right] \ge \frac{1}{4}.$$

This implies that $\frac{2}{\nu}e^{-\frac{|D|}{45}} \ge \frac{1}{4}$, which means that $|D| \le 4\alpha k$ (recall that $\nu > e^{-\frac{1}{150}\alpha k}$) . ◀

5.4 Direct Product Function Inputs

In Section 5.2 we proved Claim 40, let $\mathcal{A} = \mathcal{A}_x$ for this input x. From the claim, we know that for each $A \in \mathcal{A}$ there exists a direct product function g_A such that,

$$\Pr_{A_1,A_2\in\mathcal{A}}[g_{A_1}, g_{A_1} \text{ are consistent}] \geq \frac{\epsilon^2}{32}.$$

We want to use Theorem 48 in order to build a global direct product function. For every $A \in \mathcal{A}$, the direct product function $g_A = (g_{A,1}, \dots g_{A,\frac{9k}{10}})$, $g_{A,i} : [N] \to [M]$ can also be written as $f_S : S \to \Sigma$, where $S = [k] \setminus A$, and $\Sigma = [M]^N$. For every $i \in S$, $f_S(i)$ is the truth table of $g_{A,i}$. The distance measure in Σ is the normalized hamming distance between two strings in $[M]^N$, i.e.

$$\operatorname{dist}(\sigma_1, \sigma_2) = \Pr_{u\in[N]}[\sigma_1(u) \neq \sigma_2(u)].$$

From the definition of consistent, for every consistent A_1, A_2, the functions f_{S_1}, f_{S_2} satisfy $\Delta(f_{S_1}, f_{S_2}) < 60\lambda$.

For every $A \notin \mathcal{A}$, we define a "fake" function f_S for $S = [k] \setminus A$, and assume that its outputs are at distance 1 from any other outputs, i.e. for every $S' \in \binom{[N]}{k}, i \in S \cap S'$, $\operatorname{dist}(f_S(i), f_{S'}(i)) = 1$.

Let $\mathcal{F}$ be the collection of local functions $\{f_S\}_{S\in\binom{[N]}{k}}$ that we have just defined, let $\alpha = 60\lambda$ and $\nu = \left(\frac{\epsilon}{4}\right)^2 \frac{\epsilon^2}{32} = \frac{\epsilon^4}{512}$.

$$\operatorname{agree}_\alpha(\mathcal{F}) = \Pr_{A_1,A_2}[A_1, A_2 \in \mathcal{A}, A_1, A_2 \text{ are consistent}] \geq \left(\frac{\epsilon}{4}\right)^2 \frac{\epsilon^2}{32} = \nu.$$

In order of the theorem to hold, we need $\nu = \frac{\epsilon^4}{32} = \frac{1}{32}e^{-4c_1\lambda^2 k}$ to satisfy $\nu > e^{-\frac{1}{3}\alpha^2 k} = e^{-\frac{1}{3}(60\lambda)^2 k}$, this holds for a small enough c_1.

By Theorem 48, there exists a product function $g' : [k] \to \Sigma$ which is close to $\frac{\nu}{4}$ of the functions f_S. Translating it back to our setting, we can write g' as $g = (g_1, \dots g_k), g_i : [N] \to [M]$, and a set $\mathcal{A}^*$ of size $\frac{\nu}{4} = \frac{1}{2048}\epsilon^4$, such that for each $A \in \mathcal{A}^*$,

$$\Pr_{i\in\bar{A},u\in[N]}[g_i(u_i) \neq g_{A,i}(u_i)] \leq 300\alpha = 18000\lambda.$$

For simplicity of notations, let $\delta = 300\alpha$. Notice that by our definition, for each $A \notin \mathcal{A}$ the function g_A never agrees with any other function, therefore $\mathcal{A}^* \subset \mathcal{A}$.

▶ **Definition 55.** An input z is consistent with a set $A \in \mathcal{A}^*$ with respect to the product function g, denoted by $z \in \mathcal{Z}^g_A$, if $z \in \mathcal{Z}_A$, and $g_A(z_{\bar{A}}) \overset{2\delta k}{\approx} g(z)_{\bar{A}}$.

▶ **Claim 56.** *For each $A \in \mathcal{A}^*$,*

$$\Pr_{z\in[N]^k}[z \in \mathcal{Z}^g_A] \geq \frac{\epsilon}{8}.$$

Proof. Fix $A \in \mathcal{A}^*$, for each $i \in \bar{A}$, denote by p_i the probability of g, g_A to differ on the ith coordinate, $p_i = \Pr_{u\in[N]}[g_{A,i}(u) \neq g_i(u)]$, from Theorem 48 $\mathbb{E}_{i\in\bar{A}}[p_i] \leq \delta$.

Let I_i be the indicator random variable that equals 1 with probability p_i independently for each i. For a uniform $z \in [N]^k$,

$$\mathop{\mathbb{E}}_{z\in[N]^k}[\operatorname{dist}(g_A(z), g(z)_{\bar{A}})] = \sum_{i\in\bar{A}} I_i\,.$$

Using Chernoff

$$\Pr_z\left[g_A(z_{\bar{A}}) \overset{2\delta k}{\not\approx} g(z)_{\bar{A}}\right] \leq \Pr\left[\sum_{i\in[k]\setminus A} I_i \geq 2\,\mathbb{E}\left[\sum_{i\in[k]\setminus A} I_i\right]\right] \leq e^{-\frac{1}{9}\delta k} \leq \frac{\epsilon}{8}.$$

We know that $A \in \mathcal{A}$, therefore $\Pr_z\left[z \in \mathcal{Z}_A\right] \geq \frac{\epsilon}{4}$, therefore

$$\Pr_z\left[z \in \mathcal{Z}^g_A\right] \geq \Pr_z\left[z \in \mathcal{Z}_A, g_A(z_{\bar{A}}) \overset{2\delta k}{\approx} g(z)_{\bar{A}}\right] \geq \frac{\epsilon}{4} - \frac{\epsilon}{8} \geq \frac{\epsilon}{8}.$$

◀

▶ **Claim 57.** *If $z \in [N]^k$ is satisfies $z \in \mathcal{Z}^g_A$ for more than $\frac{\epsilon}{16}$ fraction of the sets $A \in \mathcal{A}^*$, then $f(z) \overset{3\delta k}{\approx} g(z)$.*

Proof. Fix $z \in [N]^k$ such that $z \in \mathcal{Z}^g_A$ for more than $\frac{\epsilon}{16}$ fraction of the sets $A \in \mathcal{A}^*$.

Assume towards contradiction that $f(z) \overset{3\delta k}{\not\approx} g(z)$, and denote by $D \subset [k]$ the set of coordinates in which they differ

$$D = \{i \in [k] \mid \; f(z)_i \neq g(z)_i\}.$$

For each A such that $z \in \mathcal{Z}^g_A$, by definition $g(z)_{\bar{A}} \overset{2\delta k}{\approx} g_A(z_{\bar{A}})$. Since $z \in \mathcal{Z}_A$, we also know that $g_A(z_{\bar{A}}) \overset{20\lambda k}{\approx} f(z)_{\bar{A}}$. Using both,

$$g(z)_{\bar{A}} \overset{2\delta k + 20\lambda k}{\approx} f(z)_{\bar{A}}.$$

By the definition of D, this implies that $|\bar{A} \cap D| \leq 2\delta k + 20\lambda k \leq 2.1\delta k$, the rest of D must be in A, $|A \cap D| \geq |D| - 2.1\delta k$. According to our assumption, $|D| \geq 3\delta k$, which implies that $|A \cap D| \geq \frac{1}{4}|D|$.

From the previous paragraph, all sets A such that $z \in \mathcal{Z}^g_A$ satisfies $|A \cap D| \geq \frac{1}{4}|D|$, and there are $\frac{\epsilon}{16}|\mathcal{A}^*|$ such sets.

From Claim 11, we know that for a random set $A \subset [k]$ of size $\frac{1}{10}k$,

$$\Pr_A\left[|D \cap A| \geq \frac{1}{4}|D|\right] \leq e^{-150\lambda k}.$$

The set $\mathcal{A}^*$ has measure $\frac{\epsilon^4}{2048}$,in order to satisfy the requirements $\frac{\epsilon}{16}\frac{\epsilon^4}{2048} < e^{-150\lambda k}$, and we reach a contradiction. ◀

The previous claims practically finishes the proof

Proof of Theorem 1. From Claim 56, each $A \in \mathcal{A}^*$ satisfies $|\mathcal{Z}^g_A| \geq \frac{\epsilon}{8}N^k$, therefore

$$\mathbb{E}_z\left[|\{A \in \mathcal{A}^* \mid \; z \in \mathcal{Z}^g_A\}|\right] = \sum_{A\in\mathcal{A}^*} \mathbb{E}_z\left[\mathbb{I}(z \in \mathcal{Z}^g_A)\right] = \frac{1}{8}\epsilon|\mathcal{A}^*|.$$

From averaging, a uniform $z \in [N]^k$ satisfies $|\{A \in \mathcal{A}^* \mid \; z \in \mathcal{Z}^g_A\}| \geq \frac{1}{16}\epsilon|\mathcal{A}^*|$ with probability at least $\frac{1}{16}\epsilon$. Using Claim 57, each such input z satisfies $f(z) \overset{3\delta k}{\approx} g(z)$. We chose $\delta = 300\alpha = 18000\lambda$, in order to get that $f(z) \overset{\lambda' k}{\approx} g(z)$ we just need to choose small enough c_1, and substitute $\lambda' = \frac{1}{18000}\lambda$ in the proof. ◀

6 Lower Bounds for Approximate Equality

Our direct product theorem states that if a function $f : [N]^k \to [M]^k$ passes Test 1 with $t = \frac{k}{10}$ with probability $\epsilon > e^{-c_1\lambda^2 k}$, i.e. $\alpha_{Z(\frac{k}{10})}(f) > e^{-c_1\lambda^2 k}$, then there exists a direct product function $g = (g_1, \dots g_k)$ such that

$$\Pr_{x\in[N]^k}\left[f(x) \overset{\lambda k}{\approx} g(x)\right] \geq \Omega(\epsilon).$$

Ideally, we want the stronger conclusion that

$$\Pr_{x\in[N]^k}[f(x) = g(x)] \geq \Omega(\epsilon).$$

i.e., replacing approximate equality with equality.

In the introduction there is an example explaining why approximate equality is necessary for f such that $\alpha_{Z(\frac{k}{10})}(f) \geq e^{-\delta k}$. In this section, we show two extensions.

1. We generalize Test 1 with intersection size t to Test 5 with two intersection parameters $t_1, t_2 \in \mathbb{N}, t_1 + t_2 \leq k$, and show a lower bound for Test 5 with every such t_1, t_2 (Test 5 with $t_1 = t_2$ is equivalent to Test 1).
2. We analyze the triangle test, Test 6, and give a lower bound for this test.

▶ **Definition 58.** We say that functions $f_1, f_2 : [N]^k \to [M]^k$ are (ϵ, δ) close, if

$$\Pr_{x\in[N]^k}\left[f_1(x) \overset{\delta k}{\approx} f_2(x)\right] \geq \epsilon.$$

A function $f : [N]^k \to [M]^k$ is (ϵ, δ) far from direct product, if there is no direct product function $g = (g_1, \dots, g_k) : [N]^k \to [M]^k$ that is (ϵ, δ) close to f.

Recall $w \overset{t}{\approx} w'$ if w, w' are equal in all but t of the coordinates.

In this notation, Theorem 1 states that if $\alpha_{Z(\frac{k}{10})}(f) = \epsilon > e^{-c_1\lambda^2 k}$, then f is $(\Omega(\epsilon), \lambda)$ close to a direct product function. We are interested to know if it is possible to have a direct product theorem such that f is $(\Omega(\epsilon), 0)$ close to a direct product function.

Let h be the function from Example 3 in the introduction, it satisfies $\alpha_{Z(\frac{k}{10})}(h) = \epsilon > e^{-c_1\lambda^2 k}$, but is $(c \cdot \epsilon, \lambda)$-far from a direct product function for any constant c. Therefore, it is not true that $\alpha_{Z(\frac{k}{10})}(h) > e^{-c_1\lambda^2 k}$ implies $(\Omega(\epsilon), 0)$ close to a direct product function.

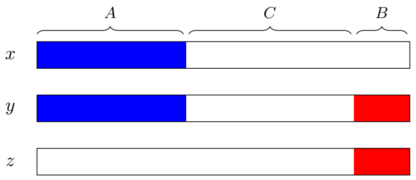
h is a direct product function with noise, on each input $x \in [N]^k$, $h(x)$ is corrupted on λk coordinates. The direct product test with $t = \frac{k}{10}$ does not check all the coordinates of each input, so with probability $e^{-\lambda\delta k}$, non of the corrupted coordinates are checked. However, if we change the parameters of the test from $t = \frac{k}{10}$ to $t = \frac{k}{2}$, in which all coordinates of the input y are checked, the function h no longer passes the test.

Is it possible to prove $(\Omega(\epsilon), 0)$ close for Test 1 with $t = \frac{1}{2}$? the answer is no. For $m = \{0, 1\}$ we don't know the answer, and it remains an open question.

▶ **Claim 59.** *For every constant $\delta > 0$ and $t_1, t_2 \in \mathbb{N}, t_1 + t_2 \leq k$, there exist a constant $\beta > 0$ and a function $f : [N]^k \to [M]^k$ for $N, M \gg k$, such that $\alpha_{Z(t_1,t_2)}(f) = \epsilon \geq e^{-\delta k}$, but f is $(\epsilon^2, \frac{\beta}{\log k})$ far from any direct product function.*

Proof. Test 5 is symmetric with respect to t_1, t_2, so we can assume wlog that $t_1 \geq t_2$. We choose $N, M \geq e^{k^2}$ such that $M \leq \sqrt{N}$. We divide the proof into two cases, depending on t_1. For each of the two cases we construct a function with ℓ corrupted coordinates, such that $\alpha_{Z(t_1,t_2)}(f) = \epsilon \geq e^{-\delta k}$, and show that both these functions are $(\epsilon^2, \frac{\ell}{2k})$ far from direct product function.

1. Choose A, B, C to be a random partition of $[k]$, such that $|A| = t_1, |B| = t_2$.
2. Choose uniformly at random $x, y, z \in [N]^k$ such that $x_A = y_A$ and $y_B = z_B$.
3. Reject if $f(x)_A \neq f(y)_A$ or $f(z)_B \neq f(y)_B$, else accept.



Denote by $\alpha_{Z(t_1,t_2)}(f)$ *the success probability of* f *on this test.*

Test 5 "Z"-test with parameters t_1, t_2 (3-query test).

If $t_1 \leq 0.4k$

This case is similar to Example 3, and we provide here a detailed analysis. Let $f : [N]^k \to [M]^k$ be the constant function 1, i.e. $f(x) = 1, \ldots 1$ for every $x \in [N]^k$, but for every $x \in [N]^k$ we corrupt $f(x)$ on $\ell \leq \frac{1}{10}k$ random coordinates $i_x^{(1)}, \ldots i_x^{(\ell)}$ to random values in $[M] \setminus \{1\}$. The number of corrupted coordinates ℓ is decided later.

Let A, B, C, x, y, z the sets and inputs chosen in Test 5, since $t_2 \leq t_1 \leq 0.4k$, $|C| \geq 0.2k$. If all the corrupted coordinates of x, y, z are not in A and all the corrupted coordinates of y, z not in B, the output of f on all of the corrupted coordinates is not checked and the test passes.

$$\Pr\left[\text{Test passes}\right] \geq \Pr\left[i_x^{(1)}, \ldots i_x^{(\ell)} \notin A, i_y^{(1)}, \ldots i_y^{(\ell)} \notin A \cup B, i_z^{(1)}, \ldots i_z^{(\ell)} \notin B\right] \geq 0.1^{3\ell}.$$

The last inequality is because the corrupted coordinates on x, y, z are independent. For input x and $i_x^{(1)}, \ldots i_x^{(\ell)}$, even conditioning on $i_x^{(1)}, \ldots i_x^{(\ell-1)} \in C$, the probability of $i_x^{(\ell)}$ to be in C is at least 0.1 (since $\ell \leq 0.1k$), same for y, z.

We choose $\ell = \beta k$ for a constant β, such that $0.1^{3\ell} \geq e^{-\delta k}$, this means that f satisfies $\alpha_{Z(t_1,t_2)}(f) = \epsilon \geq e^{-\delta k}$.

We now show that f is $(\epsilon^2, \frac{\ell}{2k})$-far from every direct product function. We do it by describing a property of f, showing that our function satisfies it with high probability and that this property implies $(\epsilon^2, \frac{\ell}{2k})$-far from direct product function.

For every $i \in [k], b \in [N]$ let $G_{i,b}$ be

$$G_{i,b} = \left\{x \in [N]^k \;\middle|\; x_i = b\right\}.$$

The function f is called *balanced* if for every $i \in [k], b \in [N], a \in [M] \setminus \{1\}$,

$$\Pr_{x \in G_{i,b}}\left[f(x)_i = a\right] \leq \frac{2k}{M}.$$

We show that our random function f is balanced with probability almost 1. Fix $i \in [k], b \in [N], a \in [M] \setminus \{1\}$. By the definition of f, $\Pr_{x, \in G_{i,b}}[f(x)_i = a] \leq \frac{1}{M}$, and this is independent for each $x \in G_{i,b}$, therefore using Chernoff bound

$$\Pr\left[\sum_{x \in G_{i,b}} I(f(x)_i = a) \geq \frac{2}{M} N^{k-1}\right] \leq e^{-\frac{1}{3M}N^{k-1}}.$$

Preforming union bound over all $i \in [k], b \in [N], a \in [M]$, the probability that f is balanced is at least $1 - kNMe^{-\frac{1}{3M}N^{k-1}} \geq 1 - e^{-N}$.

Given that f is balanced and has exactly ℓ corrupted coordinates per input, we show it is $(\epsilon^2, \frac{\ell}{2k})$-far from direct product function. Let f be such function, and assume towards contradiction that there exist $g = (g_1, \ldots, g_k)$ that is $(\epsilon^2, \frac{\ell}{2k})$ close to f. Let $F = \left\{ x \in [N]^k \,\middle|\, f(x) \overset{\ell-1}{\approx} g(x) \right\}$, by our assumption $|F| \geq \epsilon^2 N^k$.

Let $F_{i,b} \subseteq G_{i,b}$ be the set

$$F_{i,b} = \{x \in F \mid \ x_i = b, g_i(x_i) = f(x)_i \neq 1\}.$$

Every $x \in F$ has ℓ coordinates $i \in [k]$ in which $f(x)_i \neq 1$. For every $x \in F$, $f(x) \overset{\ell-1}{\approx} g(x)$, so there must be $i \in [k]$ such that $f(x)_i = g_i(x_i) \neq 1$. Therefore, every $x \in F$ must be in at least one $F_{i,b}$, and the sets $\{F_{i,b}\}_{i\in[k],b\in[N]}$ must cover F, i.e. $F \subseteq \bigcup_{i\in[k],b\in[N]} F_{i,b}$.

By definition, all $x \in F_{i,b}$ satisfies $f(x)_i = g_i(b) \in [M] \setminus \{1\}$, since f is balanced, $|F_{i,b}| \leq \frac{2k}{M}|G_{i,b}| \leq \frac{2k}{M}N^{k-1}$.

$$|F| \leq \sum_{i\in[k],b\in[N]} |F_{i,b}| \leq Nk \cdot \frac{2k}{M}N^{k-1} \leq \frac{2k^2}{M}N^k \ll \epsilon^2 N^k$$

and we reached a contradiction to the assumption $|F| \geq \epsilon^2 N^k$.

If $t_1 > 0.4k$

In this case, we can't simply corrupt coordinates to random values, because it is possible that $t_1 + t_2 = t$, and all coordinates of $f(y)$ are checked. Instead, we corrupt coordinates in a more subtle way. We start by constructing a function $f : [N]^k \to [M]^k$ that has a single corrupted coordinate per input, and $\alpha_{Z(t_1,t_2)}(f) = \Omega(\frac{1}{k^2})$.

Let $f : [N]^k \to [M]^k$ be the constant 1 function (i.e. $f(x) = 1, \ldots 1$ for all x), and for every $b \in [N]$, let $p_b : [N] \to [M] \setminus \{1\}$ be a random function. For every input $x \in [N]^k$, we choose two random coordinates $i_x \neq j_x \in [k]$, i_x is the corrupted coordinate, and j_x is the master coordinate. We corrupt $f(x)$ by setting $b = x_{j_x}$ and

$$f(x)_{i_x} = p_b(x_{i_x}).$$

Let A, B, x, y, z be the sets and inputs chosen in the test, if $i_x = i_y$, $j_x = j_y$ and $i_x, j_x \in A$, then $f(x)_A = f(y)_A$ (because the corrupted coordinates are corrupted to the same value). If in addition $i_z \notin B$, then also $f(z)_B = f(y)_B$ (because y, z don't have any corrupted coordinates on B).

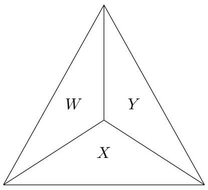
The probability of $i_x = i_y$ and $j_x = j_y$ is $\frac{1}{k^2}$, as they are both random indices in $[k]$. The probability of $i_x, j_x \in A, i_z \notin B$ is at least 0.3^3, therefore $\alpha_{Z(t_1,t_2)}(f) = \Omega(\frac{1}{k^2})$.

Instead of corrupting a single coordinate per input, we can corrupt $\ell \leq 0.1k$ different coordinates, by choosing different $i_x^{(1)}, \ldots, i_x^{(\ell)}$ and $j_x^{(1)}, \ldots, j_x^{(\ell)}$ for every $x \in [N]^k$, and continue as before. A similar probabilistic argument shows that that this function f has $\alpha_{Z(t_1,t_2)}(f) = \Omega(\frac{1}{k^{2\ell}})$ (conditioning on all other $i_x^{(1)}, \ldots i_x^{(\ell)}, j_x^{(1)}, \ldots, j_x^{(\ell-1)} \in A$, the probability of $j_x^{(\ell)} \in A$ is at least 0.2).

Fix a constant $\delta > 0$, in order of the function f to pass the test with probability $e^{-\delta k}$, the number of corrupted coordinates ℓ should satisfy $\frac{c}{k^{2\ell}} > e^{-\delta k}$, which means that we can choose $\ell = \beta \frac{k}{\log k}$ for some constant $\beta > 0$.

The constant function 1 is $(1, \frac{\ell}{k})$ close to f, we show that any direct product function $g = (g_1, \ldots, g_k)$ is $(\epsilon^2, \frac{\ell}{2k})$-far from f. Intuitively, it is true because the corrupted coordinates are corrupted to N different random functions, receiving values in M, for $k \ll N, M$. More

1. Choose disjoint $W, X, Y \subset [N]$ of size $\frac{k}{2}$.
2. Reject if $f(X \cup W)_W \neq f(Y \cup W)_W$, $f(X \cup Y)_Y \neq f(Y \cup W)_Y$ or $f(X \cup W)_X \neq f(X \cup Y)_X$, else accept.



Denote by $\alpha_{T_{set}}(f)$ the success probability of f on this test.

Test 6 Triangle test (3-query test, for even k).

formally, we show that with high probability the function f is also balanced, and use the proof of the previous case.

In the previous case, we showed that f is balanced with high probability by Chernoff bound over the inputs in $G_{i,b}$. This is not possible to do in our case, because for $x, y \in G_{i,b}$, there is a dependence between the values of the corrupted coordinates of x and y. Instead, we look at the random set of functions $\{p_b\}_{b\in[N]}$.

The function set $\{p_b\}_{b\in[N]}$ is called *balanced*, if for every $b' \in [N], a \in [M] \setminus \{1\}$, $\Pr_{b\in[N]}[p_b(b') = a] \leq 2\frac{1}{M}$.

Fix $b' \in [N], a \in [M] \setminus \{1\}$, a random function set $\{p_b\}_{b\in[N]}$ satisfies for every $b \in [N]$, $\Pr_{p_b}[p_b(b') = a] = \frac{1}{M-1}$, independently for each function p_b. Therefore using Chernoff bound,

$$\Pr_{\{p_b\}}\left[\sum_{b\in[N]} I(p_b(b') = a) > \frac{2N}{M}\right] \leq e^{-\frac{N}{4M}} \leq e^{-\frac{\sqrt{N}}{4}}.$$

Preforming union bound over all $b' \in [N], a \in [M] \setminus \{1\}$, a random function set $\{p_b\}_{b\in[N]}$ is balanced with probability at least $1 - NMe^{-\frac{\sqrt{N}}{4}}$.

We now show that a balanced function set $\{p_b\}_{b\in[N]}$ implies a balanced function f. Fix $i \in [k], b' \in [N], a \in [M]$, and let $A = \{b \in [N] \mid\ p_b(b') = a\}$, if $\{p_b\}_{b\in[N]}$ is balanced, then $|A| \leq \frac{2}{M}N$. The set $G_{i,b'}$ is a subcube of dimension $k-1$, so its coordinates are uniform in $[N]$, and by union bound

$$\Pr_{x\in G_{b',i}}[\exists j \in [k] \setminus \{i\} \text{ s.t } x_j \in A] \leq \frac{2k}{M}.$$

If there is no j such that $x_j \in A$, it is impossible that $f(x)_i = a$, because $f(x)_i$ is either 1, or $p_{x_j}(b')$ for some $j \in [k] \setminus \{i\}$. Therefore, at most $\frac{2k}{M}$ of $x \in G_{i,b'}$ can satisfy $f(x)_i = a$, and such f is balanced with probability 1.

The function f is balanced with ℓ corrupted coordinates per input, so by the previous case, f is $(\epsilon^2, \frac{\ell}{2k})$-far from direct product. ◀

Notice that in the proof, the range of $t_1 \leq 0.4k$ has a lower bound of $\ell = \beta k$, whereas in the second case, the lower bound is only $\ell = \frac{\beta k}{\log k}$.

The example in the proof can easily be transformed into a function on sets $f : \binom{[N]}{k} \to [M]^k$, which gives a bound on Test 4. This is done by choosing for each set $S \in \binom{[N]}{k}$ ℓ elements in S to corrupt and ℓ master elements (instead of coordinates) .

In Test 5 with $t_1 + t_2 = k$, we compare $f(y)$ on all coordinates, but only part of the coordinates of $f(x), f(z)$. What if we compare all coordinates of all three inputs? This brings us to the triangle test, ,Test 6, for functions over sets. In this test, every two out of the three inputs share a joint subset of size $\frac{k}{2}$, for this test we must assume that k is even.

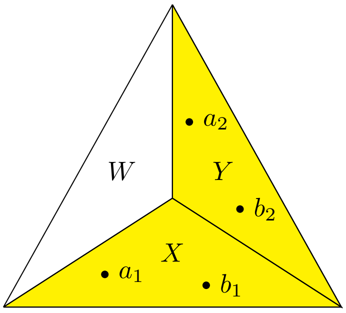


Figure 2 The set $S_2 = X \cup Y$ is marked in yellow.

▶ **Claim 60.** *For every constant $\delta > 0$, there exist a constant $\beta > 0$ and a function $f : \binom{[N]}{k} \to [M]^k$ with $N, M \gg k$, such that $\alpha_{T_{set}}(f) = \epsilon > e^{-\delta k}$, and f is $(\epsilon^2, \frac{\beta}{\log k})$ far from direct product function.*

Proof. The function f that we describe in this proof is similar to the function from the previous proof, we only need to modify it slightly such that there is the same number of corrupted elements in each half of the inputs. We start by describing a function with two corrupted elements per input.

Let $f : \binom{[N]}{k} \to [M]^k$ be the constant function 1, i.e. $f(S) = 1, \dots 1$ for every set S, and for every $b \in [N]$ we choose a random function $p_b : [N] \to [M] \setminus \{1\}$. For every $S \in \binom{[N]}{k}$, we choose two elements to corrupt $a_1, a_2 \in S$ and two master elements $b_1, b_2 \in S$. Then, we set $f(S)_{a_1} = p_{b_1}(a_1)$ and $f(S)_{a_2} = p_{b_2}(a_2)$.

Suppose W, X, Y are the sets chosen in Test 6, fix $a_1, b_1 \in X, a_2, b_2 \in Y$ and $a_3, b_3 \in W$. If the following three events hold, the test passes, see Figure 2.

1. In the set $S_2 = X \cup Y$ the elements chosen to corrupt are a_1, a_2 with the master elements b_1, b_2 respectively.
2. In the set $S_1 = X \cup W$ the elements chosen to corrupt are a_1, a_3 with the master elements b_1 b_3 respectively.
3. In the set $S_3 = Y \cup W$ the elements chosen to corrupt are a_2, a_3 with the master elements b_2, b_3 respectively.

If the three events hold, then on every check of the test, both the corrupted element and its master element are the same in both inputs, so they are corrupted to the same value and the check passes.

The probability of each event is at least $\frac{1}{k^4}$, and the event are independent, since the choice of which elements to corrupt is done independently for each $S \in \binom{[N]}{k}$. Therefore the function f passes Test 6 with probability at least $\frac{1}{k^{12}}$. It is possible to do a more careful analysis and get a higher success probability bound, but it is not important in our case.

If we corrupt 2ℓ elements per set $S \in \binom{[N]}{k}$, similar analysis shows that f satisfies $\alpha_{T_{set}}(f) = \Omega(\frac{1}{k^{12\ell}})$. Setting $\ell = \frac{\beta k}{\log k}$ for some constant β, we get f such that $\alpha_{T_{set}}(f) \geq e^{-\delta k}$.

We show that f with 2ℓ corrupted coordinates is $(\epsilon^2, \frac{\ell}{2k})$-far from direct product function in a very similar way to the previous proof. As we have seen in the proof of Claim 59, the random function set $\{p_b\}_{b \in [N]}$ is balanced with probability at least $1 - MNe^{-\frac{1}{4}\sqrt{N}}$.

For every $b \in [N]$, let $G_b = \{S \subset [N] \mid |S| = k, b \in S\}$, we say that the function $f : \binom{[N]}{k} \to [M]^k$ is *balanced* if for every $b' \in [N], a \in [M] \setminus \{1\}$,

$$\Pr_{S \in G_{b'}} [f(S)_{b'} = a] \leq \frac{2k}{M}.$$

We show that every f with a balanced function set $\{p_b\}_{b\in[N]}$ is balanced. Fix $b'\in[N], a\in[M]$, and let $A=\{b\in[N] \mid \ p_b(b')=a\}$, for a balanced function set, $|A|\leq\frac{2M}{N}$. Like previously, the set $G_{b'}$ is actually equivalent to all subset of size $k-1$ of elements in $[N]\setminus\{b'\}$, therefore a uniform $S\in G_{b'}$ contains $b\in A$ with probability at most $\frac{2k}{M}$, and f is balanced.

Assume towards contradiction that f is $(\epsilon^2,\frac{\ell}{2k})$ close to a direct product function $g:[N]\to[M]$, and let F be set set of inputs in which $f(S)\overset{\ell-1}{\approx}g(S)$. Similar to the previous proof, for every $b\in[N]$ let $F_b=\{S\in F \mid \ b\in S, f(S)_b=g(b)\neq 1\}$.

Since g approximated F up to $\ell-1$ elements, and f has ℓ corrupted elements, every $S\in F$ is in some $F_{b'}$, and $F\subseteq\cup_{b'\in[N]}F_{b'}$. Since f is balanced, for every $b'\in[N]$, $|F_{b'}|\leq\frac{2k}{M}|G_{b'}|$,

$$|F|\leq\sum_{b'\in[N]}|F_{b'}|\leq N\frac{2k}{M}|G_{b'}|\leq\frac{2k^2}{M}\left|\binom{[N]}{k}\right|.$$

The last inequality, is because each $S\in\binom{[N]}{k}$ is in at most k sets $G_{b'}$. This is a contradiction of $|F|\geq\epsilon^2\left|\binom{[N]}{k}\right|$. ◀

References

1 Sanjeev Arora and Madhu Sudan. Improved low-degree testing and its applications. In *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*, pages 485–495. ACM, 1997.

2 Vitaly Bergelson, Terence Tao, and Tamar Ziegler. An inverse theorem for the uniformity seminorms associated with the action of $\mathbb{F}_p^\infty$. *Geometric and Functional Analysis*, 19(6):1539–1596, 2010.

3 Amey Bhangale, Irit Dinur, and Inbal Livni Navon. Cube vs. cube low degree test. In *Proceedings of the 2017 Conference on Innovations in Theoretical Computer Science, ITCS*, 2017.

4 Irit Dinur. The PCP theorem by gap amplification. *Journal of the ACM (JACM)*, 54(3):12, 2007.

5 Irit Dinur and Elazar Goldenberg. Locally testing direct product in the low error range. In *Foundations of Computer Science, 2008. FOCS'08. IEEE 49th Annual IEEE Symposium on*, pages 613–622. IEEE, 2008.

6 Irit Dinur, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. Towards a proof of the 2-to-1 games conjecture? *Electronic Colloquium on Computational Complexity (ECCC)*, 2016. URL: https://eccc.weizmann.ac.il/report/2016/198/.

7 Irit Dinur and David Steurer. Direct product testing. In *Computational Complexity (CCC), 2014 IEEE 29th Conference on*, pages 188–196. IEEE, 2014.

8 Oded Goldreich and Shmuel Safra. A combinatorial consistency lemma with application to proving the PCP theorem. *SIAM Journal on Computing*, 29(4):1132–1154, 2000.

9 Russell Impagliazzo, Valentine Kabanets, and Avi Wigderson. New direct-product testers and 2-query PCPs. *SIAM Journal on Computing*, 41(6):1722–1768, 2012.

10 Subhash Khot, Dor Minzer, and Muli Safra. On Independent Sets, 2-to-2 Games and Grassmann Graphs. *Electronic Colloquium on Computational Complexity (ECCC)*, 2016.

11 Elchanan Mossel, Krzysztof Oleszkiewicz, and Arnab Sen. On reverse hypercontractivity. *Geometric and Functional Analysis*, 23(3):1062–1097, 2013.

12 Ran Raz. A parallel repetition theorem. *SIAM Journal on Computing*, 27(3):763–803, 1998.

13 Ran Raz and Shmuel Safra. A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*, pages 475–484. ACM, 1997.

A Chernoff and Hypercontractivity Proofs

Proof of Claim 11. For each element $i \in D$, we define the indicator random variable I_i to indicate that $i \in A$. In this notation

$$|A \cap D| = \sum_{i \in D} I_i.$$

We want to use Chernoff bound on I_i, but since A is of fixed size, the indicator variables are not independent. Instead, we define for each i the new random variables J_i that are independent.

For (2), let

$$J_i = \begin{cases} 1 & \text{w.p } \frac{3}{20} \\ 0 & \text{w.p } 1 - \frac{3}{20} \end{cases}.$$

For every $i \in D$ and every fixed value $b \in \{0,1\}^{|D|}$ of the indicators $\{I_l, l \neq i\}$, $\Pr\left[I_i = 1 \mid \forall l \neq i, I_l = b_l\right] \leq \Pr\left[J_i = 1\right]$. In the worse case, they are all set to 0 (none is in A), and $\Pr[I_i = 1] = \frac{3}{20}$. Therefore, we can use Chernoff bound on the random variables J_i and get a result for I_i:

$$\Pr_A\left[\sum_{i \in D} I_i \geq \frac{1}{5}|D|\right] \leq \Pr_J\left[\sum_{i \in D} J_i \geq \frac{1}{5}|D|\right] \leq e^{-\frac{1}{320}|D|}$$

For (3), we define

$$J_i = \begin{cases} 1 & \text{w.p } \frac{1}{15} \\ 0 & \text{w.p } 1 - \frac{1}{15} \end{cases}.$$

In this case, for every $i \in D$ and fixed value $b \in \{0,1\}^{|D|}$, $\Pr\left[I_i = 1 \mid \forall l \neq i, I_l = b_l\right] \geq \Pr\left[J_i = 1\right]$, and

$$\Pr_A\left[\sum_{i \in D} I_i \leq \frac{1}{20}|D|\right] \leq \Pr_J\left[\sum_{i \in D} J_i \leq \frac{1}{20}|D|\right] \leq e^{-\frac{1}{60}|D|}. \qquad ◀$$

Proof of Corollary 15. $|A| \geq |B|$ implies $a \leq b$, we know that

$$e^{-\frac{\rho ab}{2(1-\rho)}} \geq e^{-\frac{\rho b^2}{2(1-\rho)}} = \Pr_{x \in [N]^k}[x \in B]^{\frac{\rho}{1-\rho}}.$$

Similarly

$$e^{-\frac{(2-\rho)(a^2+b^2)}{4(1-\rho)}} = e^{\frac{2-\rho}{2(1-\rho)} \cdot \left(-\frac{a^2}{2} - \frac{b^2}{2}\right)} = e^{\left(1+\frac{\rho}{2(1-\rho)}\right) \cdot \left(-\frac{a^2}{2} - \frac{b^2}{2}\right)} =$$
$$\Pr_{x \in [N]^k}[x \in B]^{1+\frac{\rho}{2(1-\rho)}} \Pr_{x \in [N]^k}[x \in A]^{1+\frac{\rho}{2(1-\rho)}}$$

Together we get

$$\Pr_{x,y}[x \in A, y \in B] \geq \Pr_{x \in [N]^k}[x \in A]^{1+\frac{\rho}{2(1-\rho)}} \Pr_{x \in [N]^k}[x \in B]^{1+\frac{3\rho}{2(1-\rho)}}. \qquad ◀$$

Proof of Claim 16. We notice that regardless which of the sets G, L is the largest, by Corollary 15,

$$\Pr_{w\in[N]^k,(v,J)\in\mathcal{N}_{\frac{3}{4}}(w)}[w \in L, v \in G] \geq \left(\Pr_w[w \in L]\right)^{\frac{11}{2}} \nu^{\frac{11}{2}}.$$

By the definition of L,

$$\Pr_{w\in[N]^k,(v,J)\in\mathcal{N}_{\rho}(w)}[w \in L, v \in G] \leq \Pr_w[w \in L]\eta.$$

Therefore

$$\Pr_w[w \in L]^{\frac{9}{2}} \leq \nu^{-\frac{11}{2}}\eta. \qquad ◀$$

B Tuples to Sets Local Structure Proof

In this section we prove Lemma 37, restricted global structure for sets, we restate it bellow.

▶ **Lemma 37.** *There exist a small constants $\delta > 0$, such that for every constant $\lambda > 0$ and large enough $k \in \mathbb{N}$ such that $N > k^2e^{10\delta\lambda k}$, the following holds,*

For every function $f : \binom{[N]}{k} \to [M]^k$, if $\alpha_{Z_{set}(\frac{k}{10})}(f) = \epsilon > e^{-\delta\lambda k}$, then at least $(1-\epsilon^2-\frac{k^2}{N})$ of the good pairs $W \in \binom{[N]}{\frac{k}{10}}, X \in \binom{[N]}{\frac{9k}{10}}$ are DP pairs, i.e. there exist $g_{X,W} : [N] \to [M]$ such that

$$\Pr_Y\left[f(Y \cup W)_Y \overset{3\alpha k}{\not\approx} g_{X,W}(Y) \,\middle|\, Y \cap W = \emptyset, f(X \cup W)_W = f(Y \cup W)_W\right] \leq 2\epsilon^2.$$

In order to prove the lemma, for every function $f : \binom{[N]}{k} \to [M]^k$ we define a function $f' : [N]^k \to [M]^k \cup \bot$. For every $S \subset [N]$, we assume that the output of $f(S)$ is ordered in an ascending order over the elements or S.

In order to simplify the notation, for every string $x \in [N]^k$, we define $U(x) = 1$ if x has unique coordinates, i.e there is no $i \neq j$ such that $x_i = x_j$, else $U(x) = 0$.

▶ **Definition 61.** Given a function $f : \binom{[N]}{k} \to [M]^k$, let $f' : [N]^k \to [M]^k \cup \bot$ be defined as follows. For every $x \in [N]^k$ let X be the set of elements in x,

$$f'(x) = \begin{cases} \pi(f(X)) & U(x) = 1 \\ \bot & U(x) = 0 \end{cases}.$$

Where $\pi \in \mathcal{S}_k$ is the permutation from the ascending order over the elements of X to x.

For a set $S \subset [N]$ of size k and a permutation $\pi \in \mathcal{S}_k$, we denote by $\pi(S) \in [N]^k$ the string generated by applying π on the elements of S ordered in an ascending order. Therefore, for every $X \in \binom{[N]}{k}$, $f'(\pi(X)) = \pi(f(X))$.

▶ **Definition 62.** Let $\mathcal{D} : \binom{[N]}{\frac{k}{10}} \times \binom{[N]}{\frac{9k}{10}} \times \binom{[N]}{\frac{9k}{10}} \to [0,1]$ be the following distribution:

1. Choose $W \subset [N]$ of size $\frac{k}{10}$.
2. Choose $X \subset [N]$ of size $\frac{9k}{10}$ such that $X \cap W = \emptyset$.
3. Choose $Y \subset [N]$ of size $\frac{9k}{10}$ such that $Y \cap W = \emptyset$.

Let $\mathcal{D}' : \binom{[k]}{\frac{k}{10}} \times [N]^k \times [N]^k \to [0,1]$ be the following distribution:

1. Choose a set $A \subset [k]$ of size $\frac{k}{10}$.

2. Choose $x \in [N]^k$ such that $U(x) = 1$.
3. Choose $y \in [N]^k$ such that $x_A = y_A$ and $U(y) = 1$.

Fixing a set $A \subset [k]$ and $x \in [N]^k$ such that $U(x) = 1$, we denote by $\mathcal{D}'|A, x$ the distribution over y, conditioning on A, x being already chosen. Similarly for $W, X \subset [N]$, we define $\mathcal{D}|W, X$ the distribution over Y.

We can easily see that if we pick $(W, X, Y) \sim \mathcal{D}$, then choose a random set A and random permutations $\pi_1 \in \mathcal{S}_{\frac{k}{10}}, \pi_2, \pi_3 \in \mathcal{S}_{\frac{k}{10}}$, and set $x = (\pi_1(W)_A, \pi_2(X)_{\bar{A}})$,$y = (\pi_1(W)_A, \pi_3(Y)_{\bar{A}})$, we get $(A, x, y) \sim \mathcal{D}'$.

For each two sets W, X, let $x = (\pi_1(W)_A, \pi_2(X)_{\bar{A}})$ for an arbitrary $A \subset [k]$ and π_1, π_2, then the distribution $y \sim \mathcal{D}'|A, x$ is the same distribution as $(\pi_1(W)_A, \pi_3(Y)_{\bar{A}})$ for $Y \sim \mathcal{D}|W, X$ and uniform $\pi_3 \in \mathcal{S}_{\frac{9k}{10}}$.

We further notice that the distribution $(W, X, Y) \sim \mathcal{D}$ is the distribution used in Test 4. The distribution $(A, x, y) \sim \mathcal{D}'$ is the distribution of Test 2 with $t = \frac{k}{10}$ conditioning on $U(x) = U(y) = 1$.

Let $p_1 = \Pr_{x \in [N]^k}[U(x) = 0]$. For every $x \in [N]^k$ such that $U(x) = 1$ and a set $A \subset [k]$, let $p_2 = \Pr_y[U(y) = 0 \mid y_A = x_A]$ (p_2 is the same for every A, x such that $U(x) = 1$). We bound the probabilities p_1, p_2.

Choosing a uniform $x \in [N]^k$ can be done coordinate by coordinate. For each coordinate i, the probability that $x_i = x_j$ for $j < i$ is less than $\frac{i-1}{N}$, therefore

$$p_1 = \Pr_{x \in [N]^k}[U(x) = 0] \le \sum_{i=1}^{k} \frac{i-1}{N} \le \frac{k^2}{2N}.$$

Similarly, we can think of picking y given A, x as starting with the fixed y_A (which doesn't contain two identical coordinates as $U(x) = 1$) and choosing coordinates one by one.

$$p_2 = \Pr_y[U(y) = 0 \mid y_A = x_A] \le \sum_{i=\frac{k}{10}}^{k} \frac{i-1}{N} \le \frac{k^2}{2N}.$$

▶ **Claim 63.** *For every function $f : \binom{[N]}{k} \to [M]^k$, the function $f' : [N]^k \to [M]^k$ from Definition 61 satisfies*

$$\alpha_{V(\frac{k}{10})}(f') = (1 - p_1)(1 - p_2) \Pr[f \text{ passes } Item 3 \text{ of } Test\ 4].$$

Proof. Fix a function $f : \binom{[N]}{k} \to [M]^k$, and let $f' : [N]^k \to [M]^k$ be the function from Definition 61.

If either $U(x) = 0$ or $U(y) = 0$, by definition f' outputs $\perp$ and the test fails. If we condition on $U(x) = U(y) = 1$, the test distribution equals $\mathcal{D}'$. Let W be the set of elements of x_A, X of $x_{\bar{A}}$ and Y of $y_{\bar{A}}$, then $(W, X, Y) \sim \mathcal{D}$.

For every A, x, y such that $U(x) = U(y) = 1$ and $x_A = y_A$, the permutation $\pi_1 \in \mathcal{S}_{\frac{k}{10}}$ from the ascending order in W to the order of x_A satisfies $f'(x)_A = \pi_1(f(X, W)_W)$, and $f'(y)_A = \pi_1(f(Y, W)_W)$. Therefore, $f'(x)_A = f'(y)_A \iff f(X, W)_W = f(Y, W)_W$.

This implies that

$$\begin{aligned}\Pr[f' \text{ passes } Test\ 2] &= \Pr_{A,x,y}[f'(x)_A = f'(y)_A \mid x_A = y_A] \\ &= \Pr_{A,x,y}[U(x) = U(y) = 1 \mid x_A = y_A] \Pr_{(A,x,y) \sim \mathcal{D}'}[f'(x)_A = f'(y)_A] \\ &= (1 - p_1)(1 - p_2) \Pr_{(W,X,Y) \sim \mathcal{D}}[f(X, W)_W = f(Y, W)_W] \\ &= (1 - p_1)(1 - p_2) \Pr[f \text{ passes } Item 3 \text{ of } Test\ 4].\end{aligned}$$

Where $\Pr_{A,x,y}\left[U(x)=U(y)=1 \mid x_A=y_A\right]=(1-p_1)(1-p_2)$ by the definition of p_1,p_2. ◀

▶ **Claim 64.** *For every function on sets $f:\binom{[N]}{k}\to[M]^k$, the function $f':[N]^k\to[M]^k$ from Definition 61 satisfies the following. For every disjoint $W\in\binom{[N]}{\frac{k}{10}}, X\in\binom{[N]}{\frac{9k}{10}}$, every set $A\subset[k],|A|=\frac{k}{10}$ and every permutations $\pi_1\in\mathcal{S}_{\frac{k}{10}},\pi_2\in\mathcal{S}_{\frac{9k}{10}}$, the pair $(A,x=(\pi_1(W)_A,\pi_2(X)_{\bar{A}}))$ satisfies*

$$\Pr_y\left[f'(x)_A=f'(y)_A \mid y_A=x_A\right]=(1-p_2)\Pr_{Y\sim\mathcal{D}|W,X}\left[f(X\cup W)_W=f(Y\cup W)_W\right].$$

Proof. Fix a function $f:\binom{[N]}{k}\to[M]^k$, and let $f':[N]^k\to[M]^k$ be the function from Definition 61. Fix two disjoint subsets $W\in\binom{[N]}{\frac{k}{10}}, X\in\binom{[N]}{\frac{9k}{10}}$, a subset $A\subset[k],|A|=\frac{k}{10}$, and permutations $\pi_1\in\mathcal{S}_{\frac{k}{10}},\pi_2\in\mathcal{S}_{\frac{9k}{10}}$. Set $x=(\pi_1(W)_A,\pi_2(X)_{\bar{A}})$, since X,W are disjoint, $U(x)=1$. By the definition of f', $f'(x)_A=\pi_1(f(X,W)_W)$.

Let $y\in[N]^k$ be a random string such that $x_A=y_A$, if $U(y)=0$, then $f'(y)=\perp$ and $f'(x)_A\neq f'(y)_A$. By definition, $p_2=\Pr_y[U(y)=0|x_A=y_A]$. If we condition on $U(y)=1$, the distribution over y is $\mathcal{D}'|A,x$. If we take Y to be the elements of $y_{\bar{A}}$, then the distribution over Y is $\mathcal{D}|W,X$.

For y such that $U(y)=1$, by the definition of f', $f'(y)_A=\pi_1(f(Y,W)_W)$, and therefore $f'(x)_A=f'(y)_A \iff f(X,W)_W=f(Y,W)_W$.

$$\begin{aligned}\Pr_y\left[f'(x)_A=f'(y)_A \mid y_A=x_A\right]=&\Pr_y\left[U(y)=0 \mid x_A=y_A\right]\Pr_{y\sim\mathcal{D}'|A,x}\left[f'(x)_A=f'(y)_A\right]\\ =&(1-p_2)\Pr_{Y\sim\mathcal{D}|W,X}\left[f(X\cup W)_W=f(Y\cup W)_W\right].\end{aligned}$$ ◀

Proof of Lemma 37. Let $f:\binom{[N]}{k}\to[M]^k$ be the function such that $\alpha_{Z_{set}(\frac{k}{10})}(f)=\epsilon>e^{-\delta\lambda k}$, and let $f':[N]^k\to[M]^k$ be the function from Definition 61. By Claim 63, f' passes Test 2 with probability $\epsilon'=(1-p_1)(1-p_2)\epsilon$, therefore, Theorem 21 holds for the function f'.

By Claim 64, for every disjoint $W\in\binom{[N]}{\frac{k}{10}}, X\in\binom{[N]}{\frac{9k}{10}}$,

$$\Pr_y\left[f'(x)_A=f'(y)_A \mid y_A=x_A\right]=(1-p_2)\Pr_{Y\sim\mathcal{D}_{W,X}}\left[f(X\cup W)_W=f(Y\cup W)_W\right].$$

Setting $\eta=1-p_1$, this means that if X,W satisfies $\Pr_Y\left[f(X\cup W)_W=f(Y\cup W)_W\right]\geq\eta\frac{\epsilon}{2}$, then for every set $A\subset[k]$ and permutations π_1,π_2, the pair $(A,x=(\pi_1(W)_A,\pi_2(X)_{\bar{A}}))$ satisfies $\Pr_y\left[f'(x)_A=f'(y)_A \mid y_A=x_A\right]\geq\frac{\epsilon'}{2}$.

Theorem 21 implies that with probability $1-\epsilon'^2$ a good $\tau\sim\mathcal{D}$ (equivalent to A,x that satisfies $\Pr_y\left[f'(x)_A=f'(y)_A \mid y_A=x_A\right]\geq\frac{\epsilon'}{2}$) is a DP-restriction. Since every W,X corresponds for the same number of (A,x), for at least $(1-\epsilon'^2)\geq(1-\epsilon^2-\frac{k^2}{N})$ of the sets W,X, there exist at least one set A and permutations π_1,π_2 such that $\tau=(A,x,f'(x)_A)$ is a DP restriction, for $x=(\pi_1(W)_A,\pi_2(X)_{\bar{A}})$.

Let W,X be such sets, i.e. there exist $A\subset[k]$ and permutations π_1,π_2 such that $\tau=(A,x,f'(x)_A)$ is a DP-restriction, for $x=(\pi_1(W)_A,\pi_2(X)_{\bar{A}})$. We show that (W,X) are a DP-pair. Let $g^\tau=g_1^\tau,\dots g_{\frac{9k}{10}}^\tau,g_i^\tau:[N]\to[M]$ be the direct product function of τ. We define $g_{W,X}:[N]\to[M]$ to be the following function, for every $a\in[N]$, $g_{W,X}(a)$ is the most frequent value $g_i^\tau(a)$, among all $i\in\frac{9k}{10}$.

We recall that $\mathcal{V}_\tau=\left\{w\in[N]^{\bar{A}}\;\middle|\;f'(x_A,w)_A=f'(x)_A\right\}$ and denote by $\mathcal{V}_{W,X}$ the analog in sets,

$$\mathcal{V}_{W,X}=\left\{Y\in\binom{[N]}{\frac{9k}{10}}\;\middle|\;Y\cap W=\emptyset,f(Y,W)_W=f(X,W)_W\right\}.$$

We notice that for every $w \in \mathcal{V}_\tau$, $f'(x_A, w) \neq \perp$, so it has unique coordinates, $U(x_A, w) = 1$.

In these notations, Theorem 21 implies $\Pr_{w \in \mathcal{V}_\tau}\left[f'(x_A, w)_{\bar{A}} \overset{\alpha k}{\not\approx} g^\tau(w)\right] \leq \epsilon'^2$, and we need to prove the analog statement for $Y \in \mathcal{V}_{W,X}$.

We describe the following random process: for every $Y \in \mathcal{V}_{W,X}$, we choose a random permutation π_3 and set $w = \pi_3(Y)$. We notice that for every $Y \in \mathcal{V}_{W,X}$, $f(Y, W)_W = f(X, W)_W$, and by the definition of f' this implies that $f'(x_A, w)_A = f'(x)_A$, so $w \in \mathcal{V}_\tau$. Moreover, for every $w \in \mathcal{V}_\tau$ exists exactly one $Y \in \mathcal{V}_{W,X}$ and permutation π_3 such that $w = \pi_3(Y)$.

Suppose $Y \in \mathcal{V}_{W,X}$ such that $g_{W,X}(Y) \overset{3\alpha k}{\not\approx} f(Y \cup W)_Y$, and let $B \subset Y$ be the set of elements that $g_{W,X}(Y), f(Y \cup W)_Y$ differ on, i.e. for every $b \in B$, $g_{W,X}(b) \neq f(Y \cup W)_b$. Since $g_{W,X}$ is the most frequent value among $g_i^\tau(b)$, for at least half of the locations i, $g_i^\tau(b) \neq f(Y \cup W)_b$.

For a random permutation π_3, each $b \in B$ has probability of at least $\frac{1}{2}$ to fall into a "bad location", i.e i such that $g_i^\tau(b) \neq f(Y \cup W)_b$. Since α is a very small constant, even conditioning on αk of $b \in B$ to be in a bad location, the probability of $b' \in B$ to fall into a bad location is at least $\frac{2}{5}$. By Chernoff bound, with probability larger than $1 - e^{-\frac{1}{100}\alpha k}$, π_3 is such that at least $\frac{1}{3}$ of $b \in B$ are in a "bad location". By the definition of f', this implies that $f'(x_A, w)_{\bar{A}} \overset{\alpha k}{\not\approx} g^\tau(w)$.

Therefore, we get that

$$\Pr_{Y \in \mathcal{V}_{W,X}}\left[g_{W,X}(Y) \overset{3\alpha k}{\not\approx} f(Y \cup W)_Y\right]\left(1 - e^{-\frac{1}{100}\alpha k}\right) \leq \Pr_{w \in \mathcal{V}_\tau}\left[f'(x_A, w)_{\bar{A}} \overset{\alpha k}{\not\approx} g^\tau(w)\right] \leq \epsilon'^2.$$

Which implies that

$$\Pr_{Y \in \mathcal{V}_{W,X}}\left[g_{W,X}(Y) \overset{3\alpha k}{\not\approx} f(Y \cup W)_Y\right] \leq \epsilon'^2 + e^{-\frac{1}{100}\alpha k} \leq 2\epsilon^2. \qquad \blacktriangleleft$$