A Combinatorial Consistency Lemma with application to proving the PCP Theorem

Oded Goldreich[*]
Department of Computer Science
and Applied Mathematics
Weizmann Institute of Science
Rehovot, Israel.

Shmuel Safra
Computer Science Department
Sackler Faculty of Exact Sciences
Tel-Aviv University
Ramat-Aviv, Israel

January 24, 1997

Abstract

The current proof of the PCP Theorem (i.e., $\mathcal{NP} = \mathcal{PCP}(\log, O(1))$) is very complicated. One source of difficulty is the technically involved analysis of low-degree tests. Here, we refer to the difficulty of obtaining *strong* results regarding low-degree tests; namely, results of the type obtained and used by Arora and Safra and Arora et. al.

In this paper, we eliminate the need to obtain such strong results on low-degree tests when proving the PCP Theorem. Although we do not get rid of low-degree tests altogether, using our results it is now possible to prove the PCP Theorem using a simpler analysis of low-degree tests (which yields weaker bounds). In other words, we replace the strong algebraic analysis of low-degree tests presented by Arora and Safra and Arora et. al. by a combinatorial lemma (which does not refer to low-degree tests or polynomials).

Keywords: Parallelization of Probabilistic Proof Systems, Probabilistically Checkable Proofs (PCP), NP, Low-Degree Tests.

[*]Research was supported in part by grant No. 92-00226 from the United States – Israel Binational Science Foundation (BSF), Jerusalem, Israel.

1 Introduction

The characterization of $\mathcal{NP}$ in terms of Probabilistically Checkable Proofs (PCP systems) [AS, ALMSS], hereafter referred to as the PCP Characterization Theorem, is one of the more fundamental achievements of complexity theory. Loosely speaking, this theorem states that membership in any NP-language can be verified probabilistically by a polynomial-time machine which inspects a constant number of bits (in random locations) in a "redundant" NP-witness. Unfortunately, the current proof of the PCP Characterization Theorem is very complicated and, consequently, it has not been assimilated into complexity theory. Clearly, changing this state of affairs is highly desirable.

There are two things which make the current proof (of the PCP Characterization Theorem) difficult. One source of difficulty is the complicated conceptual structure of the proof (most notably the acclaimed 'proof composition' paradigm). Yet, with time, this part seems easier to understand and explain than when it was first introduced. Furthermore, the Proof Composition Paradigm turned out to be very useful and played a central role in subsequent works in this area (cf., [BGLR, BS, BGS, H96]). The other source of difficulty is the technically involved analysis of low-degree tests. Here we refer to the difficulty of obtaining *strong* results regarding low-degree tests; namely, results of the type obtained and used in [AS] and [ALMSS].

In this paper, we eliminate the latter difficulty. Although we do not get rid of low-degree tests altogether, using our results it is now possible to prove the PCP Characterization Theorem using only the weaker and simpler analysis of low-degree tests presented in [GLRSW, RS92, RS96]. In other words, we replace the complicated algebraic analysis of low-degree tests presented in [AS, ALMSS] by a combinatorial lemma (which does not refer to low-degree tests or even to polynomials). We believe that this combinatorial lemma is very intuitive and find its proof much simpler than the algebraic analysis of [AS, ALMSS]. (However, simplicity may be a matter of taste.)

Loosely speaking, our combinatorial lemma provides a way of generating sequences of pairwise independent random points so that any assignment of values to the sequences must induce consistent values on the individual elements. This is obtained by a "consistency test" which samples a constant number of sequences. We stress that the length of the sequences as well as the domain from which the elements are chosen are parameters, which may grow while the number of samples remains fixed.

1.1 Two Combinatorial Consistency Lemmas

The following problem arises frequently when trying to design PCP systems, and in particular when proving the PCP Characterization Theorem. For some sets S and V, one has a procedure, which given (bounded) oracle access to any function $f : S \mapsto V$, tests if f has some desired property. Furthermore, in case f is sufficiently bad (i.e., far from any function having the property), the test detects this with "noticeable" probability. For example, the function f may be the proof-oracle in a basic PCP system which we want to utilize (as an ingredient in the composition of PCP systems). The problem is that we want to increase the detection probability (equivalently, reduce the error probability) without increasing the number of queries, although we are willing to allow more informative queries. For example, we are willing to allow queries in which one supplies a sequence of elements in S and expects to obtain the corresponding sequence of values of f on these elements. The problem is that the sequences of values obtained may not be consistent with any function $f : S \mapsto V$.

We can now phrase a simple problem of testing consistency. One is given access to a function

$F : S^\ell \mapsto V^\ell$ and is asked whether there exists a function $f : S \mapsto V$ so that for most sequences $(x_1, ..., x_\ell) \in S^\ell$,

$$F(x_1, ..., x_\ell) = (f(x_1), ..., f(x_\ell)).$$

Loosely speaking, we prove that querying F on a constant number of related random sequences suffices for testing a relaxion of the above. That is,

Lemma 1.1 (combinatorial consistency – simple case): *For every $\delta > 0$, there exists a constant $c = \mathrm{poly}(1/\delta)$ and a probabilistic oracle machine, T, which on input $(\ell, |S|)$ runs for* $\mathrm{poly}(\ell \cdot \log |S|)$*-time and makes at most c queries to an oracle $F : S^\ell \mapsto V^\ell$, such that*

- *If there exist a function $f : S \mapsto V$ such that $F(x_1, ..., x_\ell) = (f(x_1), ..., f(x_\ell))$, for all $(x_1, ..., x_\ell) \in S^\ell$, then T always accepts when given access to oracle F.*
- *If T accepts with probability at least $\frac{1}{2}$, when given access to oracle F, then there exist a function $f : S \mapsto V$ such that the sequences $F(x_1, ..., x_\ell)$ and $(f(x_1), ..., f(x_\ell))$ agree on at least $\ell - \sqrt{\ell}$ positions, for at least a $1 - \delta$ fraction of all possible $(x_1, ..., x_\ell) \in S^\ell$.*

Specifically, the test examines the value of the function F on random pairs of sequences $((r_1, ..., r_\ell), (s_1, ..., s_\ell))$, where $r_i = s_i$ for $\sqrt{\ell}$ of the i's, and checks that the corresponding values (on these r_i's and s_i's) are indeed equal. For details see Section 4.

Unfortunately, this relatively simple consistency lemma does not suffice for the PCP applications. The reason being that, in that application, error reduction (see above) is done via randomness-efficient procedures such as pairwise-independent sequences (since we cannot afford to utilize $\ell \cdot \log_2 |S|$ random bits as above). Consequently, the function F is not defined on the entire set S^ℓ but rather on a very sparse subset, denoted $\mathbf{S}$. Thus, one is given access to a function $F : \mathbf{S} \mapsto V^\ell$ and is asked whether there exists a function $f : S \mapsto V$ so that for most sequences $(x_1, ..., x_\ell) \in \mathbf{S}$, the sequences $F(x_1, ..., x_\ell)$ and $(f(x_1), ..., f(x_\ell))$ agree on most (continious) subsequences of length $\sqrt{\ell}$. The main result of this paper is

Lemma 1.2 (combinatorial consistency – sparse case): *For every two of integers $s, \ell > 1$, there exists a set $\mathbf{S}_{s,\ell} \subset [s]^\ell$, where $[s] \stackrel{\mathrm{def}}{=} \{1, ..., s\}$, so that the following holds:*

1. *For every $\delta > 0$, there exists a constant $c = \mathrm{poly}(1/\delta)$ and a probabilistic oracle machine, T, which on input (ℓ, s) runs for* $\mathrm{poly}(\ell \cdot \log s)$*-time and makes at most c queries to an oracle $F : \mathbf{S}_{s,\ell} \mapsto V^\ell$, such that*
   - *If there exist a function $f : [s] \mapsto V$ such that $F(x_1, ..., x_\ell) = (f(x_1), ..., f(x_\ell))$, for all $(x_1, ..., x_\ell) \in \mathbf{S}_{s,\ell}$, then T always accepts when given access to oracle F.*
   - *If T accepts with probability at least $\frac{1}{2}$, when given access to oracle F, then there exist a function $f : [s] \mapsto V$ such that for at least a $1 - \delta$ fraction of all possible $(x_1, ..., x_\ell) \in \mathbf{S}_{s,\ell}$ the sequences $F(x_1, ..., x_\ell)$ and $(f(x_1), ..., f(x_\ell))$ agree on at least a $1 - \delta$ fraction of the subsequences of length $\sqrt{\ell}$.*
2. *The individual elements in a uniformly selected sequence in $\mathbf{S}_{s,\ell}$ are uniformly distributed in $[s]$ and are pairwise independent. Furthermore, the set $\mathbf{S}_{s,\ell}$ has cardinality* $\mathrm{poly}(s)$ *and can be constructed in time* $\mathrm{poly}(s, \ell)$.

Specifically, the test examines the value of the function F on related random pairs of sequences $((r_1, ..., r_\ell), (s_1, ..., s_\ell)) \in \mathbf{S}_{s,\ell}$. These sequences are viewed as $\sqrt{\ell} \times \sqrt{\ell}$ matrices, and, loosely speaking, they are chosen to be random extensions of the same random row (or column). For details see Section 2.

In particular, the presentation in Section 2 axiomatizes properties of the set of sequences, $\mathbf{S}_{s,\ell}$, for which the above tester works. Thus, we provide a "parallel repetition theorem" which holds for random but non-independent instances (rather than for independent random instances as in other such results). However, our "parallel repetition theorem" applies only to the case where a single query is asked in the basic system (rather than a pair of related queries as in other results). Due to this limitation, we could not apply our "parallel repetition theorem" directly to the error-reduction of generic proof systems. Instead, as explained below, we applied our "parallel repetition theorem" to derive a relatively strong low-degree test from a weaker low-degree test.

We believe that the combinatorial consistency lemma of Section 2 may play a role in subsequent developments in the area.

1.2 Application to the PCP Characterization Theorem

The currently known proof of the PCP Characterization Theorem [ALMSS] composes proof systems in which the verifier makes a constant number of multi-valued queries. Such verifiers are constructed by "parallelization" of simpler verifiers, and thus the problem of "consistency" arises. This problem is solved by use of low-degree multi-variant polynomials, which in turn requires "high-quality" low-degree testers. Specifically, given a function $f : \mathrm{GF}(p)^n \mapsto \mathrm{GF}(p)$, where p is prime, one needs to test if f is close to some low-degree polynomial (in n variables over the finite field $\mathrm{GF}(p)$). It is required that any function f which disagrees with every d-degree polynomial on at least (say) 1% of the inputs be rejected with (say) probability 99%. The test is allowed to use auxiliary proof oracles (in addition to f) but it may only make a *constant* number of queries and the answers must have length bounded by $\mathrm{poly}(n, d, \log p)$. Using a technical lemma due to Arora and Safra [AS], Arora et. al. [ALMSS] proved such a result.[1] The full proof is quite complex and is algebraic in nature. A weaker result due to Gemmel et. al. [GLRSW] (see [RS96]) asserts the existence of a d-degree test which, using $d+2$ queries, rejects such bad functions with probability at least $\Omega(1/d^2)$. Their proof is much simpler. Combining the result of Gemmel et. al. [GLRSW, RS96] with our combinatorial consistency lemma (i.e., Lemma 1.2), we obtain an alternative proof of the following result

Lemma 1.3 (low-degree tester): *For every $\delta > 0$, there exists a constant c and a probabilistic oracle machine, T, which on input n, p, d runs for* $\mathrm{poly}(n, d, \log p)$*-time and makes at most c queries to both f and to an auxiliary oracle F, such that*

- *If f is a degree-d polynomial, then there exist a function F so that T always accepts.*
- *If T accepts with probability at least $\frac{1}{2}$, when given access to the oracles f and F, then f agrees with some degree-d polynomial on at least a $1 - \Omega(1/d^2)$ fraction of the domain.*[2]

We stress that in contrast to [ALMSS] our proof of the above lemma is mainly combinatorial. Our only reference to algebra is in relying on the result of Gemmel et. al. [GLRSW, RS96] (which is

[1] An improved analysis was later obtained by Friedl and Sudan [FS].
[2] Actually, [ALMSS] only prove agreement on an (arbitrary large) constant fraction of the domain.

weaker and has a simpler proof than that of [ALMSS]). Our tester works by performing many (pairwise independent) instances of the [GLRSW] test in parallel, and by guaranteeing the consistency of the answers obtained in these tests via our combinatorial consistency test (i.e., of Lemma 1.2). In contrast, prior to our work, the only way to guarantee the consistency of these answers resulted in the need to perform a low-degree test of the type asserted in Lemma 1.3 (and using [ALMSS], which was the only alternative known, this meant losing the advantage of utilizing a low-degree tests with a simpler algebraic analysis).

1.3 Related work

We refrain from an attempt to provide an account of the developments which have culminated in the PCP Characterization Theorem. Works which should certainly be mentioned include [GMR, BGKW, FRS, LFKN, S90, BFL, BFLS, FGLSS, AS, ALMSS] as well as [BF, BLR, LS, RS92]. For detailed accounts see surveys by Babai [B94] and Goldreich [G96].

Hastad's recent work [H96] contains a combinatorial consistency lemma which is related to our Lemma 1.1 (i.e., the "simple case" lemma). However, Hastad's lemma refers to the case where the test accepts with very low probability and so its conclusion is weaker (though harder to establish). Raz and Safra [RaSa] claim to have been inspired by our Lemma 1.2 (i.e., the "sparse case" lemma).

1.4 Organization

The (basic) "sparse case" consistency lemma is presented in Section 2. The application to the PCP Characterization Theorem is presented in Section 3. Section 4 contains a proof of Lemma 1.1 (which refers to sequences of totally independent random points).

Remark: This write-up reports work completed in the Spring of 1994, and announced at the Weizmann Workshop on Randomness and Computation (January 1995).

2 The Consistency Lemma (for the sparse case)

In this section we present our main result; that is, a combinatorial consistency lemma which refers to sequences of bounded independence. Specifically, we considered k^2-long sequences viewed as k-by-k matrices. To emphasize the combinatorial nature of our lemma and its proof, we adopt an abstract presentation in which the properties required from the set of matrices are explicitly stated (as axioms). We comment that the set of all k-by-k matrices over S satisfies these axioms. A more important case is given in Construction 2.3: It is based on a standard construction of pairwise-independent sequences (i.e., the matrix is a pairwise-independent sequence of rows, where each row is a pairwise-independent sequence of elements).

General Notation. For a positive integer k, let $[k] \stackrel{\text{def}}{=} \{1, ..., k\}$. For a finite set A, the notation $a \in_{\text{R}} A$ means that a is uniformly selected in A. In case A is a multiset, each element is selected with probability proportional to its multiplicity.

2.1 The Setting

Let S be some finite set, and let k be an integer. Both S and k are parameters, yet they will be implicit in all subsequent notations.

Rows and Columns. Let $\mathbf{R}$ be a multi-set of sequences of length k over S so that every $e \in S$ appears in some sequence of $\mathbf{R}$. For sake of simplicity, think of $\mathbf{R}$ as being a set (i.e., each sequence appears with multiplicity 1). Similarly, let $\mathbf{C}$ be another set of sequences (of length k over S). We neither assume $\mathbf{R} = \mathbf{C}$ nor $\mathbf{R} \neq \mathbf{C}$. We consider matrices having rows in $\mathbf{R}$ and columns in $\mathbf{C}$ (thus, we call the members of $\mathbf{R}$ row-sequences, and those in $\mathbf{C}$ column-sequences). We denote by $\mathbf{M}$ a multi-set of k-by-k matrices with rows in $\mathbf{R}$ and columns in $\mathbf{C}$. Namely,

Axiom 1 *For every $m \in \mathbf{M}$ and $i \in [k]$, the i^{th} row of m is an element of $\mathbf{R}$ and the i^{th} column of m is an element of $\mathbf{C}$.*

For every $i \in [k]$ and $\bar{r} \in \mathbf{R}$, we denote by $\mathbf{M}_i(\bar{r})$ the set of matrices (in $\mathbf{M}$) having $\bar{r}$ as the i^{th} row. Similarly, for $j \in [k]$ and $\bar{c} \in \mathbf{C}$, we denote by $\mathbf{M}^j(\bar{c})$ the set of matrices (in $\mathbf{M}$) having $\bar{c}$ as the j^{th} column. For every $\bar{r} = (r_1, ..., r_k) \in \mathbf{R}$ and every $\bar{c} = (c_1, ..., c_k) \in \mathbf{C}$, so that $r_j = c_i$, we denote by $\mathbf{M}_i^j(\bar{r}, \bar{c})$ the set of matrices having $\bar{r}$ as the i^{th} row and $\bar{c}$ as the j^{th} column (i.e., $\mathbf{M}_i^j(\bar{r}, \bar{c}) = \mathbf{M}_i(\bar{r}) \cap \mathbf{M}^j(\bar{c})$).

Shifts. We assume that $\mathbf{R}$ is "closed" under the shift operator. Namely,

Axiom 2 *For every $\bar{r} = (r_1, ..., r_k) \in \mathbf{R}$ there exists a unique $\bar{s} = (s_1, ..., s_k) \in \mathbf{R}$ satisfying $s_i = r_{i-1}$, for every $2 \leq i \leq k$. We denote this right-shifted sequence by $\sigma(\bar{r})$. Similarly, we assume that there exists a unique $\bar{s} = (s_1, ..., s_k) \in \mathbf{R}$ satisfying $s_i = r_{i+1}$, for every $1 \leq i \leq k-1$. We denote this left-shifted sequence by $\sigma^{-1}(\bar{r})$. Furthermore[3], we assume that shifting each of the rows of a matrix $m \in \mathbf{M}$, to the same direction, yields a matrix m' that is also in $\mathbf{M}$.*

We stress that we do not assume that $\mathbf{C}$ is "closed" under shifts (in an analogous manner). For every (positive) integer i, the notations $\sigma^i(\bar{r})$ and $\sigma^{-i}(\bar{r})$ are defined in the natural way.

Distribution. We now turn to axioms concerning the distribution of rows and columns in a uniformly chosen matrix. We assume that the rows (and columns) of a uniformly chosen matrix are uniformly distributed in $\mathbf{R}$ (and $\mathbf{C}$, respectively).[4] In addition, we assume that the rows (but not necessarily the columns) are also pairwise independent. Specifically,

Axiom 3 *Let m be uniformly selected in $\mathbf{M}$. Then,*

1. *For every $i \in [k]$, the i^{th} column of m is uniformly distributed in $\mathbf{C}$.*
2. *For every $i \in [k]$, the i^{th} row of m is uniformly distributed in $\mathbf{R}$.*
3. *Furthermore, for every $j \neq i$ and $\bar{r} \in \mathbf{R}$, conditioned that the i^{th} row of m equals $\bar{r}$, the j^{th} row of m is uniformly distributed over $\mathbf{R}$.*

Finally, we assume that the columns in a uniformly chosen matrix containing a specific row-sequence are distributed identically to uniformly selected columns with the corresponding entry. A formal statement is indeed in place.

Axiom 4 *For every $i, j \in [k]$ and $\bar{r} = (r_1, ..., r_k) \in \mathbf{R}$, the j^{th} column in a matrix that is uniformly selected among those having $\bar{r}$ as its i^{th} row (i.e., $m \in_{\mathrm{R}} \mathbf{M}_i(\bar{r})$), is uniformly distributed among the column-sequences that have r_j as their i^{th} element.*

[3] The extra axiom is not really necessary; see remark following the definition of the consistency test.
[4] This, in fact, implies Axiom 1.

Clearly, if the j^{th} element of $\bar{r} = (r_1, ..., r_k)$ differs from the i^{th} element of $\bar{c} = (c_1, ..., c_k)$ then $\mathbf{M}_i^j(\bar{r}, \bar{c})$ is empty. Otherwise (i.e., $r_j = c_i$), by the above axiom, $\mathbf{M}_i^j(\bar{r}, \bar{c})$ is not empty. Furthermore, the above axiom implies that (in case $r_j = c_i$) for a uniformly chosen $m \in \mathbf{M}$

$$\begin{aligned}
\text{Prob}(m \in \mathbf{M}_i^j(\bar{r}, \bar{c})) &= \text{Prob}(m \in \mathbf{M}_i(\bar{r})) \cdot \text{Prob}(m \in \mathbf{M}^j(\bar{c}) \mid m \in \mathbf{M}_i(\bar{r})) \\
&= \frac{1}{|\mathbf{R}|} \cdot \frac{1}{|C_i(r_j)|} \\
&> 0
\end{aligned}$$

where $C_i(e)$ denotes the set of column-sequences having e as their i^{th} element. (The second equality is obtained by Axiom 4.)

2.2 The Test

Let Γ be a function assigning matrices in $\mathbf{M}$ (which may be a proper subset of all possible k-by-k matrices over S) values which are k-by-k matrices over some set of values V (i.e., $\Gamma : \mathbf{M} \mapsto V^{k \times k}$). The function Γ is *supposed* to be "consistent" (i.e., assign each element, e, of S the same value, independently of the matrix in which e appears). The purpose of the following test is to check that this property holds in some approximate sense.

Construction 2.1 (Consistency Test):

1. column test: *Select a column-sequence $\bar{c}$ uniformly in $\mathbf{C}$, and $i, j \in_{\mathrm{R}} [k]$. Select two random extensions of this column, namely $m_1 \in_{\mathrm{R}} \mathbf{M}^i(\bar{c})$ and $m_2 \in_{\mathrm{R}} \mathbf{M}^j(\bar{c})$, and test if the i^{th} column of $\Gamma(m_1)$ equals the j^{th} column of $\Gamma(m_2)$.*

2. row test (analogous to the column test): *Select a row-sequence $\bar{r}$ uniformly in $\mathbf{R}$, and $i, j \in_{\mathrm{R}} [k]$. Select two random extensions of this row, namely $m_1 \in_{\mathrm{R}} \mathbf{M}_i(\bar{r})$ and $m_2 \in_{\mathrm{R}} \mathbf{M}_j(\bar{r})$, and test if the i^{th} row of $\Gamma(m_1)$ equals the j^{th} row of $\Gamma(m_2)$.*

3. shift test: *Select a matrix m uniformly in $\mathbf{M}$ and an integer $t \in [k-1]$. Let m' be the matrix obtained from m by shifting each row by t; namely, the i^{th} row of m' is $\sigma^t(\bar{r})$, where $\bar{r}$ denotes the i^{th} row of m. We test if the $k-t$ first columns of $\Gamma(m)$ match the $k-t$ last columns of $\Gamma(m')$.*

The test accepts *if all three* (sub-)tests *succeed.*

Remark: Actually, it suffices to use a seemingly weaker test in which the row-test and shift-test are combined into the following generalized row-test:

> Select a row-sequence $\bar{r}$ uniformly in $\mathbf{R}$, integers $i, j \in_{\mathrm{R}} [k]$ and $t \in_{\mathrm{R}} \{0, 1, ..., k-1\}$. Select a random extension of this row and its shift, namely $m_1 \in_{\mathrm{R}} \mathbf{M}_i(\bar{r})$ and $m_2 \in_{\mathrm{R}} \mathbf{M}_j(\sigma^t(\bar{r}))$, and test if the $(k-t)$-long suffix of the i^{th} row of $\Gamma(m_1)$ equals the $(k-t)$-long prefix of the j^{th} row of $\Gamma(m_2)$.

Our main result asserts that Construction 2.1 is a "good consistency test": Not only that ALMOST ALL ENTRIES *in almost all matrices* are assigned in a consistent manner (which would have been obvious), but ALL ENTRIES IN ALMOST ALL ROWS *of almost all matrices* are assigned in a consistent manner.

Lemma 2.2 *Suppose* $\mathbf{M}$ *satisfies Axioms 1–4. Then, for every constant* $\delta > 0$, *there exist a constant* $\epsilon > 0$ *so that if a function* $\Gamma : \mathbf{M} \mapsto V^{k\times k}$ *passes the consistency test with probability at least* $1-\epsilon$ *then there exists a function* $\tau : S \mapsto V$ *so that, with probability at least* $1-\delta$, *the value assigned by* Γ *to a uniformly chosen matrix matches the values assigned by* τ *to the elements of a uniformly chosen row in this matrix. Namely,*

$$\mathrm{Prob}_{i,m}(\forall j \ : \ \Gamma(m)_{i,j} = \tau(m_{i,j})) \ \ \geq \ \ 1-\delta$$

where $m \in_{\mathrm{R}} \mathbf{M}$ *and* $i \in_{\mathrm{R}} [k]$. *The constant* ϵ *does not depend on* k *and* S. *Furthermore, it is polynomially related to* δ.

As a corollary, we get Part (1) of Lemma 1.2. Part (2) follows from Proposition 2.4 (below).

2.3 Proof of Lemma 2.2

As a motivation towards the proof of Lemma 2.2, consider the following mental experiment. Let $m \in \mathbf{M}$ be an arbitrary matrix and e be its $(i,j)^{\mathrm{th}}$ entry. First, uniformly select a random matrix, denoted m_1, containing the i^{th} row of m. Next, uniformly select a random matrix, denoted m_2, containing the j^{th} column of m_1. The claim is that m_2 is uniformly distributed among the matrices containing the element e. Thus, if Γ passes items (1) and (2) in the consistency test then it must assign consistent values to almost all elements in almost all matrices. Yet, this falls short of even proving that there exists an assignment which matches all values assigned to the elements of some row in some matrix. Indeed, consider a function Γ which assigns 0 to all elements in the first ϵk columns of each matrix and 1's to all other elements. Clearly, Γ passes the row-test with probability 1 and the column-test with probability greater than $1-\epsilon$; yet, there is no $\tau : S \mapsto V$ so that for a random matrix the values assigned by Γ to some row match τ. It is easy to see that the shift-test takes care of this special counter-example. Furthermore, it may be telling to see what is wrong with some naive arguments. A main issue these arguments tend to ignore is that for an "adversarial" choice of Γ and a candidate choice of $\tau : S \mapsto V$, we have no handle on the (column) *location* of the elements in a random matrix on which τ disagrees with Γ. The shift-test plays a central role in getting around this problem; see subsection 2.3.2 and Claim 2.2.14 (below).

Recommendation: The reader may want to skip the proofs of all claims in first reading. We believe that all the claims are quite believable, and that their proofs (though slightly tedious in some cases) are quite straightforward. In contrast, we believe that the ideas underlying the proof of the lemma are to be found in its high level structure; namely, the definitions and the claims made.

Notation: The following notation will be used extensively throughout the proof. For a k-by-k matrix, m, we denote by $\mathrm{row}_i(m)$ the i^{th} row of m and by $\mathrm{col}^j(m)$ the j^{th} column of m. Restating the conditions of the lemma, we have (from the hypothesis that Γ passes the column test)

$$\mathrm{Prob}_{\bar{c},i,j,m_1,m_2}(\mathrm{col}^i(\Gamma(m_1)) = \mathrm{col}^j(\Gamma(m_2))) \geq 1-\epsilon \tag{1}$$

where $\bar{c}, i, j, m_1, m_2$ are uniformly selected in the corresponding sets (i.e., $\bar{c} \in \mathbf{C}$, $i, j \in [k]$, $m_1 \in \mathbf{M}^i(\bar{c})$ and $m_2 \in \mathbf{M}^j(\bar{c})$). Similarly, from the hypothesis that Γ passes the row test, we have

$$\mathrm{Prob}_{\bar{r},i,j,m_1,m_2}(\mathrm{row}_i(\Gamma(m_1)) = \mathrm{row}_j(\Gamma(m_2))) \geq 1-\epsilon \tag{2}$$

where $\bar{r} \in_{\mathrm{R}} \mathbf{R}$, $i, j \in_{\mathrm{R}} [k]$, $m_1 \in_{\mathrm{R}} \mathbf{M}_i(\bar{r})$ and $m_2 \in_{\mathrm{R}} \mathbf{M}_j(\bar{r})$. It will be convenient to extend the shift notation to matrices in the obvious manner; namely, $\sigma^t(m)$ is defined as the matrix m' satisfying

$\text{row}_i(m') = \sigma^t(\text{row}_i(m))$ for every $i \in [k]$. From the hypothesis that Γ passes the shift-test, we obtain

$$\text{Prob}_{m,t}(\forall j \le k-t \quad \text{col}^j(\Gamma(m)) = \text{col}^{j+t}(\Gamma(\sigma^t(m)))) \ge 1-\epsilon \tag{3}$$

where $m \in_{\text{R}} \mathbf{M}$ and $t \in_{\text{R}} [k-1]$. Finally, denoting by $\text{entry}_{i,j}(m)$ the $(i,j)^{\text{th}}$ entry in the matrix m, we restate the conclusion of the lemma as follows

$$\text{Prob}_{i,m}(\exists j \;\text{ so that } \text{entry}_{i,j}(\Gamma(m)) \ne \tau(\text{entry}_{i,j}(m))) \le \delta \tag{4}$$

where $m \in_{\text{R}} \mathbf{M}$ and $i \in_{\text{R}} [k]$.

2.3.1 Stable Rows and Columns – Part 1

For each $\bar{r} \in \mathbf{R}$ and $\bar{\alpha} \in V^k$, we denote by $p_{\bar{r}}(\bar{\alpha})$ the probability that Γ assigns to the row-sequence $\bar{r}$ the value-sequence $\bar{\alpha}$; namely,

$$p_{\bar{r}}(\bar{\alpha}) \stackrel{\text{def}}{=} \text{Prob}_{i,m}(\text{row}_i(\Gamma(m)) = \bar{\alpha})$$

where $i \in_{\text{R}} [k]$ and $m \in_{\text{R}} \mathbf{M}_i(\bar{r})$. Eq. (2) implies that for almost all row-sequences there is a "typical" sequence of values; see Claim 2.2.3 (below).

Definition 2.2.1 (consensus): *The* consensus *of a row-sequence $\bar{r} \in \mathbf{R}$, denoted $\text{con}(\bar{r})$, is defined as the value $\bar{\alpha}$ for which $p_{\bar{r}}(\bar{\alpha})$ is maximum. Namely, $\text{con}(\bar{r}) = \bar{\alpha}$ if $\bar{\alpha}$ is the (lexicographically first) value-sequence for which $p_{\bar{r}}(\bar{\alpha}) = \max_{\bar{\beta}}\{p_{\bar{r}}(\bar{\beta})\}$.*

Definition 2.2.2 (stable sequences): *Let $\epsilon_2 \stackrel{\text{def}}{=} \sqrt{\epsilon}$. We say that the row-sequence $\bar{r}$ is* stable *if $p_{\bar{r}}(\text{con}(\bar{r})) \ge 1-\epsilon_2$. Otherwise, we say that $\bar{r}$ is* unstable.

Clearly, almost all row-sequences are stable. That is,

Claim 2.2.3 *All but at most an ϵ_2 fraction of the row-sequence are stable.*

proof: For each fixed $\bar{r}$ we have

$$\text{Prob}_{i,j,m_1,m_2}(\text{row}_i(\Gamma(m_1)) = \text{row}_j(\Gamma(m_2))) = \sum_{\bar{\alpha}} p_{\bar{r}}(\bar{\alpha})^2$$

where $i,j \in_{\text{R}} [k]$, $m_1 \in_{\text{R}} \mathbf{M}_i(\bar{r})$ and $m_2 \in_{\text{R}} \mathbf{M}_j(\bar{r})$. Taking the expectation over $\bar{r} \in_{\text{R}} \mathbf{R}$, and using Eq. (2), we get

$$\begin{aligned} 1-\epsilon &\le \text{Prob}_{\bar{r},i,j,m_1,m_2}(\text{row}_i(\Gamma(m_1)) = \text{row}_j(\Gamma(m_2))) \\ &= \text{Exp}_{\bar{r}}(\sum_{\bar{\alpha}} p_{\bar{r}}(\bar{\alpha})^2) \\ &\le \text{Exp}_{\bar{r}}(p_{\bar{r}}^{\max}) \end{aligned}$$

where $p_{\bar{r}}^{\max} \stackrel{\text{def}}{=} \max_{\bar{\alpha}}\{p_{\bar{r}}(\bar{\alpha})\}$. Using Markov Inequality, we get

$$\text{Prob}_{\bar{r}}(p_{\bar{r}}^{\max} \le 1-\sqrt{\epsilon}) < \sqrt{\epsilon}$$

and the claim follows. $\Box$

By definition, almost all matrices containing a particular *stable* row-sequence assign this row-sequence the same sequence of values (i.e., its consensus value). We say that such matrices are conforming for this row-sequence.

Definition 2.2.4 (conforming matrix): *Let $i \in [k]$. A matrix $m \in \mathbf{M}$ is called i-*conforming* (or conforming for row-position i) if Γ assigns the i^{th} row of m its consensus value; namely, if $\text{row}_i(\Gamma(m)) = \text{con}(\text{row}_i(m))$. Otherwise, the matrix is called i-*non-conforming* (or non-conforming for row-position i).*

Claim 2.2.5 *The probability that for a uniformly chosen $i \in [k]$ and $m \in \mathbf{M}$, the matrix m is i-non-conforming is at most $\epsilon_3 \stackrel{\text{def}}{=} 2\epsilon_2$. Furthermore, the bound holds also if we require that the i^{th} row of m is stable.*

proof: The stronger bound (on probability) equals the sum of the probabilities of the following two events. The first event is that the i^{th} row of the matrix is unstable; whereas the second event is that the i^{th} row of the matrix is stable and yet the matrix is i-non-conforming. To bound the probability of the first event (by ϵ_2), we fix any $i \in [k]$ and combine Axiom 3 with Claim 2.2.3. To bound the probability of the second event, we fix any stable $\bar{r}$ and use the definition of a stable row. □

Remark: Clearly, an analogous treatment can be applied to column-sequences. In the sequel, we freely refer to the above notions and to the above claims also when discussing column-sequences.

2.3.2 Stable Rows – Part 2 (Shifts)

Now we consider the relation between the consensus of row-sequences and the consensus of their (short) shifts. By a short shift of the row-sequence $\bar{r}$, we mean any row-sequence $\bar{s} = \sigma^d(\bar{r})$ obtained with $d \in \{-(k-1), ..., +(k-1)\}$. Our aim is to show that the consensus (as well as stability) is usually preserved under short shifts.

Definition 2.2.6 (very-stable row): *Let $\epsilon_4 = \sqrt{\epsilon_2}$. We say that a row-sequence $\bar{r}$ is* very stable *if it is stable, and for all but an ϵ_4 fraction of $d \in \{-(k-1), ..., +(k-1)\}$, the row-sequence $\bar{s} \stackrel{\text{def}}{=} \sigma^d(\bar{r})$ is also stable.*

Clearly,

Claim 2.2.7 *All but at most an ϵ_4 fraction of the row-sequence are very-stable.*

proof: By a simple counting argument. □

Definition 2.2.8 (super-stable row): *Let $\epsilon_5 = \sqrt[3]{\epsilon}$ and $\epsilon_6 = 2(\epsilon_4 + \epsilon_5)$. We say that a row-sequence $\bar{r}$ is* super-stable *if it is very-stable, and, for every $j \in [k]$, the following holds*

> *for all but an ϵ_6 fraction of the $t \in [k]$, the row-sequence $\bar{s} \stackrel{\text{def}}{=} \sigma^{t-j}(\bar{r})$ is stable and $\text{con}_j(\bar{r}) = \text{con}_t(\bar{s})$, where $\text{con}_j(\bar{r})$ is the j^{th} element of $\text{con}(\bar{r})$.*

Note that the t^{th} element of $\sigma^{t-j}(\bar{r})$ is $r_{t-(t-j)} = r_j$. Thus, a row-sequence is super-stable if the consensus value of each of its elements is preserved under almost all (short) shifts.

Claim 2.2.9 *All but at most an ϵ_6 fraction of the row-sequence are super-stable.*

proof: We start by proving that almost all row-sequences and almost all their shifts have approximately matching statistics, where the *statistics vector* of $\bar{r} \in \mathbf{R}$ is defined as the k-long sequence (of functions), $p_{\bar{r}}^1(\cdot), ..., p_{\bar{r}}^k(\cdot)$, so that $p_{\bar{r}}^j(v)$ is the probability that Γ assigns the value v to the j^{th} element of the row $\bar{r}$. Namely,

$$p_{\bar{r}}^j(v) \stackrel{\text{def}}{=} \text{Prob}_{i,m}(\text{entry}_{i,j}(\Gamma(m)) = v)$$

where $i \in_{\text{R}} [k]$ and $m \in_{\text{R}} \mathbf{M}_i(\bar{r})$. By the definition of consensus, we know that for every stable row-sequence $\bar{r} \in \mathbf{R}$, we have $p_{\bar{r}}^j(\text{con}_j(\bar{r})) \geq 1 - \epsilon_2$, for every $j \in [k]$. Thus if both $\bar{r}$ and its shift $\bar{s} = \sigma^t(\bar{r})$ are stable and have approximately matching statistics (i.e., the corresponding $(k-t)$-long statistics sub-vectors are close) then their consensus must match (i.e., the corresponding $(k-t)$-long subsequences of the consensus are equal).

subclaim 2.2.9.1: For all but an ϵ_5 fraction of the row-sequences $\bar{r}$, all but an ϵ_5 fraction of the shifts $\bar{s} = \sigma^d(\bar{r})$ (for $d \in [k-1]$), satisfy

$$\sum_v |p_{\bar{r}}^j(v) - p_{\bar{s}}^{j+d}(v)| < 2\epsilon_5 \quad \text{for every } j \leq k-d.$$

proof of subclaim: Let *pref*row$_{i,j}(m)$ denote the j-long prefix of row$_i(m)$ and *suff*row$_{i,j}(m)$ its j-long suffix. By the shift-test (see Eq. (3) and $\epsilon = \epsilon_5^3$)

$$\text{Prob}_{m,i,d}(\mathit{pref}\text{row}_{i,k-d}(\Gamma(m)) = \mathit{suff}\text{row}_{i,k-d}(\Gamma(m'))) \geq 1 - \epsilon_5^3$$

where $i \in_{\text{R}} [k]$, $m \in_{\text{R}} \mathbf{M}$, $d \in_{\text{R}} [k-1]$ and $m' = \sigma^d(m)$. Using Axiom 3 (Part 2) and an averaging argument, we get that all but a ϵ_5 fraction of the $\bar{r} \in \mathbf{R}$, and all but a ϵ_5 fraction of $d \in [k-1]$,

$$\text{Prob}_{i,m}(\mathit{pref}\text{row}_{i,k-d}(\Gamma(m)) = \mathit{suff}\text{row}_{i,k-d}(\Gamma(m'))) \geq 1 - \epsilon_5$$

where $i \in_{\text{R}} [k]$, $m \in_{\text{R}} \mathbf{M}_i(\bar{r})$ and $m' = \sigma^d(m)$. We fix such a pair $\bar{r}$ and d, thus fixing also $\bar{s} = \sigma^d(\bar{r})$. A matrix-pairs (m, m') for which the equality holds contributes equally to the (appropriate $(k-d)$-long portion of the) the statistic vectors of the row-sequences $\bar{r}$ and $\bar{s}$. The contribution of matrix-pairs for which equality does not hold, to the difference $\sum_v |p_{\bar{r}}^j(v) - p_{\bar{s}}^{j+d}(v)|$, is at most $\frac{2}{k \cdot |\mathbf{M}_i(\bar{r})|}$ per each relevant j. The subclaim follows. $\diamond$

As a corollary we get

subclaim 2.2.9.2: Let us call a row-sequence, $\bar{r}$, **infective** if for every $j \in [k]$ all but an $2\epsilon_5$ fraction of the $t \in [k]$ satisfy $\sum_v |p_{\bar{r}}^j(v) - p_{\bar{s}}^t(v)| \leq 2\epsilon_5$, where $\bar{s} = \sigma^{t-j}(\bar{r})$. Then, all but a $2\epsilon_5$ fraction of the row-sequences are infective.

proof of subclaim: The proof is obvious but yet confusing. We say that $\bar{r}$ is *fine1* if for all but an ϵ_5 fraction of the $d \in [k]$ and for every $j \leq k-d$, we have $\sum_v |p_{\bar{r}}^j(v) - p_{\sigma^d(\bar{r})}^{j+d}(v)| \leq 2\epsilon_5$. Now, if $\bar{r}$ is fine1 then for every j there are at most $\epsilon_5 k$ positions $t \in \{j+1, ..., k\}$ so that $\sum_v |p_{\bar{r}}^j(v) - p_{\sigma^{t-j}(\bar{r})}^t(v)| > 2\epsilon_5$. Similarly, $\bar{r}$ is *fine2* if for all but an ϵ_5 fraction of the $d \in [k]$ and for every $j > d$ we have $\sum_v |p_{\bar{r}}^j(v) - p_{\sigma^{-d}(\bar{r})}^{j-d}(v)| \leq 2\epsilon_5$, and whenever $\bar{r}$ is fine2 then for every j there are at most $\epsilon_5 k$ positions $t \in \{1, ..., j-1\}$ so that $\sum_v |p_{\bar{r}}^j(v) - p_{\sigma^{-j+t}(\bar{r})}^t(v)| > 2\epsilon_5$. Thus, if a row-sequence $\bar{r}$ is both fine1 and fine2 then for every $j \in [k]$ all but a $2\epsilon_1$ fraction of the positions $t \in [k]$ satisfy $\sum_v |p_{\bar{r}}^j(v) - p_{\sigma^{t-j}(\bar{r})}^t(v)| \leq 2\epsilon_5$. By subclaim 2.2.9.1, we get that all but an ϵ_5 fraction of the row-sequences are fine1. A similar statement holds for fine2 (since the shift-test can be rewritten as selecting $m' \in_{\text{R}} \mathbf{M}$ and $d \in_{\text{R}} [k-1]$ and setting $m = \sigma^{-d}(m')$). Combining all these trivialities, the subclaim follows. $\diamond$

Clearly, a row-sequence $\bar{r}$ that is both very-stable and infective satisfies, for every $j \in [k]$ and all but at most $\epsilon_4 \cdot (2k-1) + 2\epsilon_5 \cdot k$ of the $t \in [k]$, both

- $\bar{s} \stackrel{\rm def}{=} \sigma^{t-j}(\bar{r})$ is stable; it follows that $p_{\bar{s}}^t(\text{con}_t(\bar{s})) \geq 1-\epsilon_2$ and $p_{\bar{s}}^t(u) \leq \epsilon_2$ for all $u \neq \text{con}_t(\bar{s})$.
- $p_{\bar{s}}^t(v) \geq p_{\bar{r}}^j(v) - 2\epsilon_5$, for every v and in particular for $v = \text{con}_j(\bar{r})$.

It follows that $p_{\bar{s}}^t(\text{con}_j(\bar{r})) \geq 1-\epsilon_2-2\epsilon_5 > \epsilon_2$, and therefore $\text{con}_j(\bar{r}) = \text{con}_t(\bar{s})$ must hold. Thus, such an $\bar{r}$ is super-stable. Combining the lower bounds given by Claim 2.2.7 and subclaim 2.2.9.2, the current claim follows (actually, we get a better bound; i.e., $\epsilon_4 + 2\epsilon_5$). $\Box$

Summary. Before proceeding let us summarize our state of knowledge. The key definitions regarding row-sequences are of stable, very-stable and super-stable row-sequences (i.e., Defs 2.2.2, 2.2.6, and 2.2.8, respectively). Recall that a stable row-sequence is assigned the same value in almost all matrices in which it appear. Furthermore, most prefixes (resp., suffices) of a super-stable row-sequence are assigned the same values in almost all matrices containing these portions (as part of some row). Regarding matrices, we defined a matrix to be i-conforming if it assigns its $i^{\rm th}$ row the corresponding consensus value (i.e., it conforms with the consensus of that row-sequence); cf., Definitions 2.2.4 and 2.2.1. We have seen that almost all row-sequences are super-stable and that almost all matrices are conforming for most of their rows. Actually, we will use the latter fact with respect to columns; that is, almost all matrices are conforming for most columns (cf., Claim 2.2.5 and the remark following it).

2.3.3 Deriving the Conclusion of the Lemma

We are now ready to derive the conclusion of the Lemma. Loosely speaking, we claim that the function τ, defined so that $\tau(e)$ is the value most frequently assigned (by Γ) to e, satisfies Eq. (4). Actually, we use a slightly different definition for the function τ.

Definition 2.2.10 (the function τ): *For a column-sequence $\bar{c}$, we denote by $\text{con}_i(\bar{c})$ the values that $\text{con}(\bar{c})$ assigns to the $i^{\rm th}$ element in $\bar{c}$. We denote by $\mathbf{C}_i(e)$ the set of column-sequences having e as the $i^{\rm th}$ component. Let $q_e(v)$ denote the probability that the consensus of a uniformly chosen column-sequence, containing e, assigns to e the value v. Namely,*

$$q_e(v) \stackrel{\rm def}{=} \text{Prob}_{i,\bar{c}}(\text{con}_i(\bar{c})=v)$$

where $i \in_{\rm R} [k]$ and $\bar{c} \in_{\rm R} \mathbf{C}_i(e)$. We consider $\tau : S \mapsto V$ so that $\tau(e) \stackrel{\rm def}{=} v$ if $q_e(v) = \max_u\{q_e(u)\}$, with ties broken arbitrarily.

Assume, on the contrary to our claim, that Eq. (4) does not hold (for this τ). Namely, for a uniformly chosen $m \in \mathbf{M}$ and $i \in [k]$, the following holds with probability greater that δ

$$\exists j \;\text{ so that } \text{entry}_{i,j}(\Gamma(m)) \neq \tau(\text{entry}_{i,j}(m)) \tag{5}$$

The notion of a annoying row-sequence, defined below, plays a central role in our argument. Using the above (contradiction) hypothesis, we first show that many row-sequences are annoying. Next, we show that lower bounds on the number of annoying row-sequences translate to lower bounds on the probability that a uniformly chosen matrix is non-conforming for a uniformly chosen column position. This yields a contradiction to Claim 2.2.5.

Definition 2.2.11 (row-annoying elements): *An element r_j in $\bar{r} = (r_1, ..., r_k) \in \mathbf{R}$, is said to be* annoying for the row-sequence $\bar{r}$ *if the $j^{\rm th}$ element in $\text{con}(\bar{r})$ differs from $\tau(r_j)$. A row-sequence $\bar{r}$ is said to be* annoying *if $\bar{r}$ contains an element that is annoying for it.*

Using Claim 2.2.9, we get

Claim 2.2.12 *Suppose that Eq.* (4) *does not hold* (for τ). *Then, at least a* $\delta_1 \stackrel{\rm def}{=} \delta - \epsilon_6 - \epsilon_2$ *fraction of the row-sequences are both super-stable and annoying.*

proof: Axiom 3 (part 2) is extensively used throughout this proof (with no explicit reference). Combining Eq. (5) and Claim 2.2.9, with probability at least $\delta - \epsilon_6 - \epsilon_2 = \delta_1$, a uniformly chosen pair $(m, i) \in \mathbf{M} \times [k]$ satisfies the following

1. there exists a j so that $\tau(\mathrm{entry}_{i,j}(m))$ is different from $\mathrm{entry}_{i,j}(\Gamma(m))$;
2. $\mathrm{row}_i(m)$ is super-stable;
3. matrix m is i-conforming; i.e., $\mathrm{entry}_{i,j}(\Gamma(m))$ equals $\mathrm{con}_j(\mathrm{row}_i(m))$, for every $j \in [k]$.

Combining conditions (1) and (3), we get that $e = \mathrm{entry}_{i,j}(m)$ is annoying for the $i^{\rm th}$ row of m. The current claim follows. □

A key observation is that each stable row-sequence which is annoying yields many matrices which are non-conforming for the "annoying column position" (i.e., for the column position containing the element which annoys this row-sequence). Namely,

Claim 2.2.13 *Suppose that a row-sequence* $\bar{r} = (r_1, ..., r_k)$ *is stable and that* r_j *is annoying for* $\bar{r}$. *Then, at least a* $\frac{1}{2} - \epsilon_2$ *fraction of the matrices, containing the row-sequence* $\bar{r}$*, are non-conforming for column-position* j.

We stress that the row-sequence $\bar{r}$ in the above claim is *not* necessarily very-stable (let alone super-stable).

proof: Let us denote by v the value assigned to r_j by the consensus of $\bar{r}$ (i.e., $v \stackrel{\rm def}{=} \mathrm{con}_j(\bar{r})$). Since r_j annoys $\bar{r}$ it follows that v is different from $\tau(r_j)$. Consider the probability space defined by uniformly selecting $i \in [k]$ and $m \in \mathbf{M}_i(\bar{r})$. Since $\bar{r}$ is stable it follows that in almost all of these matrices the value assigned to r_j by the matrix equals v. Namely,

$$\mathrm{Prob}_{i,m}(\mathrm{entry}_{i,j}(\Gamma(m)) = v)) \;\; \geq \;\; 1 - \epsilon_2 \tag{6}$$

where $i \in_{\rm R} [k]$ and $m \in_{\rm R} \mathbf{M}_i(\bar{r})$. By Axiom 4, the $j^{\rm th}$ column of m is uniformly distributed in $\mathbf{C}_i(r_j)$, and thus we may replace $\bar{c} \in_{\rm R} \mathbf{C}_i(r_j)$ by the $j^{\rm th}$ column of $m \in_{\rm R} \mathbf{M}_i(\bar{r})$. Now, using the definition of the function τ and the accompanying notations, we get

$$\mathrm{Prob}_{i,m}(\mathrm{con}_i(\mathrm{col}^j(m)) = v) \;\; = \;\; q_{r_j}(v) \;\; \leq \;\; \frac{1}{2} \tag{7}$$

where, again, $i \in_{\rm R} [k]$ and $m \in_{\rm R} \mathbf{M}_i(\bar{r})$. The inequality holds since $v \neq \tau(r_j)$ and by τ's definition $q_{r_j}(v) \leq q_{r_j}(\tau(r_j))$.
Combining Eq. (6) and (7), we get

$$\mathrm{Prob}_{i,m}(\mathrm{entry}_{i,j}(\Gamma(m)) \neq \mathrm{con}_i(\mathrm{col}^j(m))) \; \geq \; \frac{1}{2} - \epsilon_2$$

and the claim follows. □

Another key observation is that super-stable row-sequences which are annoying have the property of "infecting" almost all their shifts with their annoying positions, and thus spreading the "annoyance" over all column positions. Namely,

Claim 2.2.14 *Suppose that a row-sequence $\bar{r}$ is both super-stable and annoying. In particular, suppose that the j^{th} element of $\bar{r} = (r_1, ..., r_k)$ is annoying for $\bar{r}$. Then, for all but at most an ϵ_6 fraction of the $t \in [k]$, the the row-sequence $\bar{s} = \sigma^{t-j}(\bar{r})$ is stable and its t^{th} element* (which is indeed r_j) *is annoying for $\bar{s}$.*

proof: Since $\bar{r}$ is super-stable, we know that for all but an ϵ_6 fraction of the t's, $\text{con}_j(\bar{r}) = \text{con}_t(\bar{s})$ and $\bar{s}$ is stable (as well), where $\bar{s} = (s_1, ..., s_k) = \sigma^{t-j}(\bar{r})$. Since r_j is annoying for $\bar{r}$, we have $\text{con}_j(\bar{r}) \neq \tau(r_j)$ and $\text{con}_t(\bar{s}) \neq \tau(r_j) = \tau(s_t)$ follows (recall $r_j = s_t$). $\square$

Combining Claims 2.2.12 and 2.2.14, we derive, for almost all positions $t \in [k]$, a lower bound for the number of stable row-sequences that are annoyed by their t^{th} element.

Claim 2.2.15 *Suppose that Eq.* (4) *does not hold* (for τ)*. Then, there exists a set $T \subseteq [k]$ so that $|T| \geq (1 - 2\epsilon_6) \cdot k$ and for every $t \in T$ there is a set of at least $\frac{\delta_1}{2k} \cdot |\mathbf{R}|$ stable row-sequences so that the t^{th} position is annoying for each of these sequences.*

proof: Combining Claims 2.2.12 and 2.2.14, we get that there is a set of super-stable row-sequences $A \subseteq \mathbf{R}$ so that A contains at least a δ_1 fraction of $\mathbf{R}$, and for every $\bar{r} \in A$ there exist a $j_{\bar{r}} \in [k]$ so that for all but a ϵ_6 of the $t \in [k]$, the row-sequence $\bar{s} \stackrel{\text{def}}{=} \sigma^{t-j_{\bar{r}}}(\bar{r})$ is stable and the t^{th} position is annoying for it (i.e., for $\bar{s}$). By a counting argument it follows that there is a set T so that $|T| \geq (1 - 2\epsilon_6) \cdot k$, and for every $t \in T$ at least half of the $\bar{r}$'s in A satisfy the above (i.e., $\bar{s} \stackrel{\text{def}}{=} \sigma^{t-j_{\bar{r}}}(\bar{r})$ is stable and the t^{th} position is annoying for $\bar{s}$). Fixing such a $t \in T$, we consider the set, denoted A_t, containing these $\bar{r}$'s; namely, for every $\bar{r} \in A_t$ the row-sequence $\bar{s} \stackrel{\text{def}}{=} \sigma^{t-j_{\bar{r}}}(\bar{r})$ is stable and the t^{th} position is annoying for it (i.e., for $\bar{s}$). Thus, we have established a mapping from A_t to a set of stable row-sequences which are annoyed by their t^{th} position; specifically, $\bar{r}$ is mapped to $\sigma^{t-j_{\bar{r}}}(\bar{r})$. Each row-sequence in the range of this mapping has at most k preimages (corresponding to the k possible shifts which maintain its t^{th} element). Recalling that A_t contains at least $\frac{\delta_1}{2} \cdot |\mathbf{R}|$ sequences, we conclude that the mapping's range must contain at least $\frac{\delta_1}{2k} \cdot |\mathbf{R}|$ sequences, and the claim follows. $\square$

Combining Claims 2.2.15 and 2.2.13, we get a lower bound on the number of matrices which are non-conforming for the j^{th} column, $\forall j \in T$ (where T is as in Claim 2.2.15). Namely,

Claim 2.2.16 *Let T be as guaranteed by Claim 2.2.15 and suppose that $j \in T$. Then, at least a $\frac{\delta_1}{6}$ fraction of the matrices are non-conforming for column-position j.*

proof: By Claim 2.2.15, there are at least $\frac{\delta_1}{2k} \cdot |\mathbf{R}|$ stable row-sequences that are annoyed by their j^{th} position. Out of these row-sequences, we consider a subset, denoted A, containing exactly $\frac{\delta_1}{2k} \cdot |\mathbf{R}|$ row-sequences. By Claim 2.2.13, for each $\bar{r} \in A$, at least a $\frac{1}{2} - \epsilon_2$ fraction of the matrices containing the row-sequence $\bar{r}$ are non-conforming for column-position j. We claim that almost all of these matrices do not contain another row-sequence in A (here we use the fact that A isn't too large); this will allow us to add-up the matrices guaranteed by each $\bar{r} \in A$ without worrying about multiple counting. Namely,

subclaim 2.2.16.1: For every $\bar{r} \in \mathbf{R}$

$$\text{Prob}_{i,m}(\exists i' \neq i \ \text{ s.t. } \ \text{row}_{i'}(m) \in A) < \frac{\delta_1}{2}$$

where $i \in_{\text{R}} [k]$ and $m \in_{\text{R}} \mathbf{M}_i(\bar{r})$.

proof of subclaim: By Axiom 3 (part 3), we get that for every $i' \neq i$ the i'-th row of $m \in_{\mathrm{R}} \mathbf{M}_i(\bar{r})$ is uniformly distributed in $\mathbf{R}$. Thus, for every $i' \neq i$

$$\mathrm{Prob}_m(\mathrm{row}_{i'}(m) \in A) = \frac{\delta_1}{2k}$$

where $m \in_{\mathrm{R}} \mathbf{M}_i(\bar{r})$. The subclaim follows. $\diamond$

Using the subclaim, we conclude that for each $\bar{r} \in A$, at least a $\frac{1}{2} - \epsilon_2 - \frac{\delta_1}{2}$ ($> \frac{1}{3}$) fraction of the matrices containing the row-sequence $\bar{r}$ are non-conforming for column-position j and do not contain any other row-sequence in A. The desired lower bound now follows. Namely, let B denote the set of matrices which are non-conforming for column-position j, let $B_i(\bar{r}) \stackrel{\mathrm{def}}{=} B \cap \mathbf{M}_i(\bar{r})$ and $B'_i(\bar{r})$ denote the set of matrices in $B_i(\bar{r})$ which do not contain any row in A except for the i^{th} row; then

$$\begin{aligned}
|B| &\geq |\cup_{\bar{r}\in A} \cup_{i=1}^{k} B'_i(\bar{r})| \\
&= \sum_{\bar{r}\in A} \sum_{i=1}^{k} |B'_i(\bar{r})| \\
&> \sum_{\bar{r}\in A} \frac{1}{3} \sum_{i=1}^{k} |\mathbf{M}_i(\bar{r})| \\
&= |A| \cdot \left(\frac{1}{3} \cdot \frac{k \cdot |\mathbf{M}|}{|\mathbf{R}|}\right) \\
&= \frac{\delta_1}{6} \cdot |\mathbf{M}|
\end{aligned}$$

The claim follows. $\Box$

The combination of Claims 2.2.15 and 2.2.16, yields that a uniformly chosen matrix is non-conforming for a uniformly chosen column position with probability at least $(1 - 2\epsilon_6) \cdot \frac{\delta_1}{6}$. For a suitable choice of constants (e.g., $\epsilon = (\delta/30)^4$), this yields contradiction to Claim 2.2.5. Thus, Eq. (4) must hold for τ as defined in Def. 2.2.10, and the lemma follows. $\blacksquare$

2.4 A Construction that Satisfies the Axioms

Clearly, the set of all k-by-k matrices over S satisfies Axioms 1–4. A more interesting and useful set of matrices is defined as follows.

Construction 2.3 (basic construction): *We associate the set S with a finite field and suppose $k \leq |S|$. Furthermore, $[k]$ is associated with k elements of the field so that 1 is the multiplicative unit and $i \in [k]$ is the sum of i such units. Let $\mathbf{M}$ be the set of matrices defined by four field elements as follows. The matrix associated with the quadruple (x, y, x', y') has the $(i, j)^{\mathrm{th}}$ entry equal $(x + jy) + i(x' + jy')$.*

Remark: The column-sequences correspond to the standard pairwise-independent sequences $\{r + is : i \in [k]\}$, where $r, s \in S$. Similarly, the row-sequences are expressed as $\{r + js : j \in [k]\}$, where $r, s \in S$.

Proposition 2.4 *The Basic Construction satisfies Axioms 1–4.*

proof: Axioms 1 is obvious from the above remark. The right-shift of the sequence $\{r+js : j \in [k]\}$ is $\{(r+s)+js : j \in [k]\}$ and Axiom 2 follows. To prove that Axiom 3 holds, we rewrite the i^{th} row as $\{s_i + j \cdot r_i : j \in [k]\}$, where $s_i = x + ix'$ and $r_i = y + iy'$. Now, for every $i \neq i' \in [k]$, when $x, y, x', y' \in_{\mathrm{R}} S$, the pairs (s_i, r_i) and $(s_{i'}, r_{i'})$ are pairwise independent and uniformly distributed in $S \times S$ which corresponds to the set of row-sequences. It remains to prove that Axiom 4 holds. We start by proving the following.

Fact 2.4.1: Consider any $i, j \in [k]$ and two sequences $\bar{r} = (r_1, ..., r_k) \in \mathbf{R}$ and $\bar{c} = (c_1, ..., c_k) \in \mathbf{C}$ so that $r_j = c_i$. Then, $|\mathbf{M}_i^j(\bar{r}, \bar{c})|$ equals $|S|$.

proof of fact: By the construction, there exists a unique pair $(a, b) \in S \times S$ so that $a + j'b = r_{j'}$ for every $j' \in [k]$ (existence is obvious and uniqueness follows by considering any two equations; e.g., $a + b = r_1$ and $a + 2b = r_2$). Similarly, there exist a unique pair (α, β) so that $\alpha + i'\beta = c_{i'}$ for every $i' \in [k]$. We get a system of four linear equations in x, x', y, y' (i.e., $x + ix' = a$, $y + iy' = b$, $x + jy = \alpha$ and $x' + jy' = \beta$). This system has rank 3 and thus $|S|$ solutions, each defining a matrix in $\mathbf{M}_i^j(\bar{r}, \bar{c})$. $\diamond$

Using Fact 2.4.1, Axiom 4 follows since

$$\begin{aligned} \frac{|\mathbf{M}_i^j(\bar{r}, \bar{c})|}{|\mathbf{M}_i(\bar{r})|} &= \frac{|S|}{|S \times S|} \\ &= \frac{1}{|S|} \\ &= \frac{1}{|\mathbf{C}_i(r_j)|} \end{aligned}$$

and so does the proposition. ■

3 A Stronger Consistency Test and the PCP Application

To prove Lemma 1.3, we need a slightly stronger consistency test than the one analyzed in Lemma 2.2. This new test is given access to three related oracles, each supplying assignments to certain classes of sequences over S, and is supposed to establish the consistency of these oracles with one function $\tau : S \mapsto V$. Specifically, one oracle assigns values to k^2-long sequences viewed as two-dimensional arrays (as before). The other two oracles assign values to k^3-long sequences viewed as 3-dimensional arrays, whose slices (along a specific coordinate) correspond to the 2-dimensional arrays of the first oracle. Using Lemma 2.2 (and the auxiliary oracles) we will present a test which verifies that the first oracle is consistent in an even stronger sense than established in Lemma 2.2. Namely, not only that *all entries* IN ALMOST ALL ROWS *of almost all 2-dimensional arrays* are assigned in a consistent manner, but ALL ENTRIES *in almost all 2-dimensional arrays* are assigned in a consistent manner.

3.1 The Setting

Let S, k, $\mathbf{R}$, $\mathbf{C}$ and $\mathbf{M}$ be as in the previous section. We now consider a family, $\mathcal{M}_{\mathsf{c}}$, of k-by-k matrices with entries is $\mathbf{C}$. The family $\mathcal{M}_{\mathsf{c}}$ will satisfy Axioms 1–4 of the previous section. In addition, its induced multi-set of row-sequences, denoted $\mathcal{R}$, will correspond to the multi-set $\mathbf{M}$; namely, each row of a matrix in $\mathcal{M}_{\mathsf{c}}$ will form a matrix in $\mathbf{M}$ (i.e., the sequence of elements of $\mathbf{C}$ corresponding to a row in a $\mathcal{M}_{\mathsf{c}}$-matrix will correspond to a $\mathbf{M}$-matrix). Put formally,

Axiom 5 *For every* $\mathtt{m} \in \mathcal{M}_{\mathtt{c}}$ *and every* $i \in [k]$*, there exists* $m \in \mathbf{M}$ *so that for every* $j \in [k]$*, the* $(i,j)^{\text{th}}$ *entry of* $\mathtt{m}$ *equals the* j^{th} *column of* m (i.e., $\text{entry}_{i,j}(\mathtt{m}) = \text{col}^j(m)$, or, equivalently, $\text{row}_i(\mathtt{m}) \cong m$). *Furthermore, this matrix* m *is unique.*

Analogously, we consider also a family, $\mathcal{M}_{\mathtt{r}}$, of k-by-k matrices the entries of which are elements in $\mathbf{R}$ so that the rows[5] of each $\mathtt{m} \in \mathcal{M}_{\mathtt{r}}$ correspond to matrices in $\mathbf{M}$.

3.2 The Test

As before, Γ is a function assigning (k-by-k) matrices in $\mathbf{M}$ values which are k-by-k matrices over some set of values V (i.e., $\Gamma : \mathbf{M} \mapsto V^{k \times k}$). Let $\Gamma_{\mathtt{c}}$ (resp., $\Gamma_{\mathtt{r}}$) be (the supossedly corresponding) function assigning k-by-k matrices over $\mathbf{C}$ (resp., $\mathbf{R}$) values which are k-by-k matrices over $\overline{V} \stackrel{\text{def}}{=} V^k$ (i.e., $\Gamma_{\mathtt{c}} : \mathcal{M}_{\mathtt{c}} \mapsto \overline{V}^{k \times k}$).

Construction 3.1 (Extended Consistency Test):

1. consistency for sequences: *Apply the consistency test of Construction 2.1 to* $\Gamma_{\mathtt{c}}$*. Same for* $\Gamma_{\mathtt{r}}$*.*
2. correspondence test: *Uniformly select a matrix* $\mathtt{m} \in \mathcal{M}_{\mathtt{c}}$ *and a row* $i \in [k]$*, and compare the* i^{th} *row in* $\Gamma_{\mathtt{c}}(\mathtt{m})$ *to* $\Gamma(m)$*, where* $m \in \mathbf{M}$ *is the matrix formed by the* $\mathbf{C}$*-elements in the* i^{th} *row of* $\mathtt{m}$*. Same for* $\Gamma_{\mathtt{r}}$*.*

The test accepts *if both* (sub-)tests *succeed.*

Lemma 3.2 *Suppose* $\mathbf{M}, \mathcal{M}_{\mathtt{c}}, \mathcal{M}_{\mathtt{r}}$ *satisfy Axioms 1–5. Then, for every constant* $\gamma > 0$*, there exist a constant* ϵ *so that if a function* $\Gamma : \mathbf{M} \mapsto V^{k \times k}$ (together with some functions $\Gamma_{\mathtt{c}} : \mathcal{M}_{\mathtt{c}} \mapsto \overline{V}^{k \times k}$ and $\Gamma_{\mathtt{r}} : \mathcal{M}_{\mathtt{r}} \mapsto \overline{V}^{k \times k}$) *passes the extended consistency test with probability at least* $1 - \epsilon$ *then there exists a function* $\tau : S \mapsto V$ *so that, with probability at least* $1 - \gamma$*, the value assigned by* Γ *to a uniformly chosen matrix* $m \in \mathbf{M}$ *matches the values assigned by* τ *to each of the elements of* m*. Namely,*

$$\text{Prob}_m \left(\forall i,j \;\; \text{entry}_{i,j}(\Gamma(m)) = \tau(\text{entry}_{i,j}(m)) \right) \geq 1 - \gamma$$

where $m \in_{\text{R}} \mathbf{M}$*. The constant* ϵ *does not depend on* k *and* S*. Furthermore, it is polynomially related to* γ*.*

The proof of the lemma starts by applying Lemma 2.2 to derive assignments to $\mathbf{C}$ (resp., $\mathbf{R}$) which are consistent with $\Gamma_{\mathtt{c}}$ (resp., $\Gamma_{\mathtt{r}}$) on almost all rows of almost all k^3-dimensional arrays (ie., $\mathcal{M}_{\mathtt{c}}$ and $\mathcal{M}_{\mathtt{r}}$, respectively). It proceeds by applying a degenerate argument of the kind applied in the proof of Lemma 2.2. Again, the reader may want to skip the proofs of all claims in first reading.

3.3 Proof of Lemma 3.2

We start by considering item (1) in the Extended Consistency Test. By Lemma 2.2, there exists a function $\tau_{\mathtt{c}} : \mathbf{C} \mapsto V^k$ (resp., $\tau_{\mathtt{r}} : \mathbf{R} \mapsto V^k$) so that the value assigned by $\Gamma_{\mathtt{c}}$ (resp., $\Gamma_{\mathtt{r}}$), to a uniformly chosen row in a uniformly chosen matrix $\mathcal{M}_{\mathtt{c}}$ (resp., $\mathcal{M}_{\mathtt{r}}$), matches with high probability the values assigned by $\tau_{\mathtt{c}}$ (resp., $\tau_{\mathtt{r}}$) to each of the $\mathbf{C}$-elements (resp., $\mathbf{R}$-elements) appearing in this

[5] Alternatively, one can consider a family, $\mathcal{M}_{\mathtt{r}}$, of k-by-k matrices the entries of which are elements in $\mathbf{R}$ so that the columns of each $\mathtt{m} \in \mathcal{M}_{\mathtt{r}}$ correspond to matrices in $\mathbf{M}$. However, this would require to modify the basic consistency test (of Construction 2.1), for these matrices, so that it shifts columns instead of rows.

row. Here "with high probability" means with probability at least $1-\delta$, where $\delta > 0$ is a constant, related to ϵ as specified by Lemma 2.2. Namely,

$$\mathrm{Prob}_{i,\mathtt{m}}(\forall j \;\; \mathrm{entry}_{i,j}(\Gamma_{\mathsf{c}}(\mathtt{m})) = \tau_{\mathsf{c}}(\mathrm{entry}_{i,j}(\mathtt{m}))) \;\; \geq \;\; 1-\delta \tag{8}$$

where $i \in_{\mathrm{R}} [k]$ and $\mathtt{m} \in_{\mathrm{R}} \mathcal{M}_{\mathsf{c}}$.

3.3.1 Perfect Matrices and Typical Sequences

Eq. (8) relates τ_{c} to Γ_{c} (resp., τ_{c} to Γ_{c}). Our next step is to relate τ_{c} (resp., τ_{r}) to Γ. This is done easily by referring to item (2) in the Extended Consistency Test. Specifically, it follows that the value assigned by Γ, to a uniformly chosen matrix $m \in \mathbf{M}$, matches, with high probability, the values assigned by τ_{c} (resp., τ_{r}) to each of the columns (resp., rows) of m. That is

Definition 3.2.1 (perfect matrices): *A matrix $m \in \mathbf{M}$ is called* perfect (for columns) *if for every $j \in [k]$, the j^{th} column of $\Gamma(m)$ equals the value assigned by τ_{c} to the j^{th} column of m* (i.e., $\mathrm{col}^j(\Gamma(m)) = \tau_{\mathsf{c}}(\mathrm{col}^j(m))$). *Similarly, $m \in \mathbf{M}$ is called* perfect (for rows) *if* $\mathrm{row}_i(\Gamma(m)) = \tau_{\mathsf{r}}(\mathrm{row}_i(m))$, *for every $i \in [k]$.*

Claim 3.2.2 (perfect matrices): *Let $\delta_1 \stackrel{\mathrm{def}}{=} \delta + \epsilon$.*

(c) *All but a δ_1 fraction of the matrices in* $\mathbf{M}$ *are perfect for columns.*

(r) *All but a δ_1 fraction of the matrices in* $\mathbf{M}$ *are perfect for rows.*

proof: By the Correspondence (sub)Test, with probability at least $1-\epsilon$, a uniformly chosen row in a uniformly chosen $\mathtt{m} \in \mathcal{M}_{\mathsf{c}}$ is "given" the same values by Γ_{c} and by Γ (i.e., $\mathrm{row}_i(\Gamma_{\mathsf{c}}(\mathtt{m})) = \Gamma(m)$, for $i \in_{\mathrm{R}} [k]$, where $m \cong \mathrm{row}_i(\mathtt{m})$). On the other hand, by Eq. (8), with probability at least $1-\delta$, a uniformly chosen row in a uniformly chosen $\mathtt{m} \in \mathcal{M}_{\mathsf{c}}$ is "given" the same values by Γ_{c} and by τ_{c} (i.e., $\mathrm{entry}_{i,j}(\Gamma_{\mathsf{c}}(\mathtt{m})) = \tau_{\mathsf{c}}(\mathrm{entry}_{i,j}(\mathtt{m}))$, for $i \in_{\mathrm{R}} [k]$ and all $j \in [k]$). Thus, with probability at least $1-(\epsilon+\delta)$, a uniformly chosen row in a uniformly chosen $\mathtt{m} \in \mathcal{M}_{\mathsf{c}}$ is "given" the same values by Γ and by τ_{c} (i.e., $\mathrm{col}^j(\Gamma(\mathrm{row}_i(\mathtt{m}))) = \tau_{\mathsf{c}}(\mathrm{entry}_{i,j}(\mathtt{m}))$, for $i \in_{\mathrm{R}} [k]$ and all $j \in [k]$). Using Axiom 3 (part 2 – regarding $\mathcal{M}_{\mathsf{c}}$) and the "furthermore" part of Axiom 5, we get part **(c)** of the claim (i.e., $\mathrm{col}^j(\Gamma(m)) = \tau_{\mathsf{c}}(\mathrm{col}^j(m))$, for all $j \in [k]$). A similar argument holds for part **(r)**. □

A perfect (for columns) matrix "forces" *all* its columns to satisfy some property Π (specifically, the value assigned by τ_{c} to its column-sequences must match the value Γ of the matrix). Recall that we have just shown that almost all matrices are perfect and thus force all their columns to satisfy some property Π. Using a counting argument, one can show that all but at most a $\frac{1}{k}$ fraction of the column-sequences must satisfy Π in *almost all* matrices in which they appear. Namely,

Definition 3.2.3 (typical sequences): *Let $\delta_2 \stackrel{\mathrm{def}}{=} 2\sqrt{\delta_1}$. We say that the column-sequence $\bar{c}$* (resp., row-sequence $\bar{r}$) *is* typical *if*

$$\mathrm{Prob}_{j,m}(\mathrm{col}^j(\Gamma(m)) = \tau_{\mathsf{c}}(\bar{c})) \geq 1-\delta_2$$

where $j \in_{\mathrm{R}} [k]$ and $m \in_{\mathrm{R}} \mathbf{M}^j(\bar{c})$. Otherwise, we say that $\bar{c}$ is non-typical.

Claim 3.2.4 *All but at most an $\frac{\delta_2}{2k}$ fraction of the column-sequence* (resp., row-sequences) *are typical.*

We will only use the bound for the fraction of typical row-sequences.

proof: We mimic part of the counting argument of Claim 2.2.16. Let N be a set of non-typical row-sequences, containing exactly $\frac{\delta_2}{2k}\cdot|\mathbf{R}|$ sequences. Fix any $\bar{r}\in N$ and consider the set of matrices containing $\bar{r}$. By Axiom 3 (part 3 – regarding $\mathbf{M}$), at most a $\frac{\delta_2}{2}$ fraction of these matrices contain some other row in N. On the other hand, by definition (of non-typical row-sequence), at least a δ_2 fraction of the matrices containing $\bar{r}$, have Γ disagree with $\tau_{\mathrm{r}}(\bar{r})$ on $\bar{r}$, and thus are non-perfect (for rows). It follows that at least a $\frac{\delta_2}{2}$ fraction of the matrices containing $\bar{r}$ are non-perfect (for rows) and contain no other row in N. Combining the bounds obtained for all $\bar{r}\in N$, we get that at least a $\frac{\delta_2}{2}\cdot\frac{\delta_2}{2}$ fraction of the matrices are not perfect (for rows). This contradicts Claim 3.2.2(r), and so the current claim follows (for row-sequences and similarly for column-sequences). $\Box$

3.3.2 Deriving the Conclusion of the Lemma

We are now ready to derive the conclusion of the Lemma. Loosely speaking, we claim that the function τ, defined so that $\tau(e)$ is the value most frequently assigned by τ_{c} to e, satisfies the claim of the lemma.

Definition 3.2.5 (the function τ): *Let $\tau_{\mathrm{c}}(\bar{c})_i$ denote the value assigned by τ_{c} to the i^{th} element of $\bar{c}\in\mathbf{C}$. Define*

$$q_e(v)\stackrel{\mathrm{def}}{=}\mathrm{Prob}_{i,\bar{c}}(\tau_{\mathrm{c}}(\bar{c})_i=v)$$

where $i\in_{\mathrm{R}}[k]$ and $\bar{c}\in_{\mathrm{R}}\mathbf{C}_i(e)$ (recall that $\mathbf{C}_i(e)$ denotes the set of column-sequences having e as the i^{th} component). *We consider $\tau:S\mapsto V$ so that $\tau(e)\stackrel{\mathrm{def}}{=}v$ if $q_e(v)=\max_u\{q_e(u)\}$, with ties broken arbitrarily.*

The proof that τ satisfies the claim of Lemma 3.2 is a simplified version of the proof of Lemma 2.2.[6] We assume, on the contrary to our claim, that, for a uniformly chosen $m\in\mathbf{M}$

$$\mathrm{Prob}_m\left(\exists i,j\ \text{ so that }\mathrm{entry}_{i,j}(\Gamma(m))\neq\tau(\mathrm{entry}_{i,j}(m))\right)\ >\ \gamma \tag{9}$$

As in the proof of Lemma 2.2, we define a notion of an *annoying* row-sequence. Using the above (contradiction) hypothesis, we first show that many row-sequences are annoying. Next, we show that lower bounds on the number of annoying row-sequences translate to lower bounds on the probability that a uniformly chosen matrix is non-perfect (for columns). This yields a contradiction to Claim 3.2.2(c).

Definition 3.2.6 (a new definition of annoying rows): *A row-sequence $\bar{r}=(r_1,...,r_k)$ is said to be* annoying *if there exists a $j\in[k]$ so that the j^{th} element in $\tau_{\mathrm{r}}(\bar{r})$ differs from $\tau(r_j)$.*

Using Claim 3.2.2(r), we get

Claim 3.2.7 *Suppose that Eq. (9) hold and let $\gamma_1\stackrel{\mathrm{def}}{=}\gamma-\delta_1$. Then, at least a $\frac{\gamma_1}{k}$ fraction of the row-sequences are annoying.*

[6] The reader may wonder how it is possible that a simpler proof yields a stronger result; as the claim concerning the current τ is stronger. The answer is that the current τ is defined based on a more restricted function over $\mathbf{C}$ and there are also stronger restrictions on Γ. Both restrictions are due to facts that we have inferred using Lemma 2.2 w.r.t Γ_{c} and Γ_{r}.

proof: Combining Eq. (9) and Claim 3.2.2(r), we get that with probability at least $\gamma - \delta_1 = \gamma_1$, a uniformly chosen matrix $m \in \mathbf{M}$ is perfect for rows and contains some entry, denoted (i,j), for which the Γ value is different from the τ value (i.e., $\text{entry}_{i,j}(\Gamma(m)) \neq \tau(\text{entry}_{i,j}(m))$). Since the τ_r-value of all rows of m matches the Γ value, it follows that the i^{th} row of m is annoying. Thus, at least a γ_1 fraction of the matrices contain an annoying row-sequence. Using Axiom 3 (part 2 – regarding $\mathbf{M}$), we conclude that the fraction of annoying row-sequences must be as claimed. $\Box$

A key observation is that each row-sequence that is both typical and annoying yields many matrices which are non-perfect for columns. Namely,

Claim 3.2.8 *Suppose that a row-sequence $\bar{r}$ is both typical and annoying. Then, at least a $\frac{1}{2} - \delta_2$ fraction of the matrices, containing the row-sequence $\bar{r}$, are non-perfect for columns.*

proof: Since $\bar{r} = (r_1, ..., r_k)$ is annoying, there exists a $j \in [k]$ so that the the j^{th} component of $\tau_r(\bar{r})$ (which is the value assigned to r_j) is different from $\tau(r_j)$. Let us denote by v the value $\tau_r(\bar{r})$ assigns to r_j. Note that $v \neq \tau(r_j)$. Consider the probability space defined by uniformly selecting $i \in [k]$ and $m \in \mathbf{M}_i(\bar{r})$. Since $\bar{r}$ is typical it follows that in almost all of these matrices the value assigned to r_j by the Γ equals v; namely,

$$\text{Prob}_{i,m}(\text{entry}_{i,j}(\Gamma(m)) = v) \quad \geq \quad 1 - \delta_2 \tag{10}$$

By Axiom 4 (regarding $\mathbf{M}$), the j^{th} column of m is uniformly distributed in $\mathbf{C}_i(r_j)$. Now, using the definition of the function τ and the accompanying notations, we get

$$\text{Prob}_{i,m}(\tau_c(\text{col}^j(m))_i = v) \quad = \quad q_{r_j}(v) \quad \leq \quad \frac{1}{2} \tag{11}$$

The inequality holds since $v \neq \tau(r_j)$ and by τ's definition $q_{r_j}(v) \leq q_{r_j}(\tau(r_j))$. Combining Eq. (10) and (11), we get

$$\text{Prob}_{i,m}(\text{entry}_{i,j}(\Gamma(m)) \neq \tau_c(\text{col}^j(m))_i) \;\geq\; \frac{1}{2} - \delta_2$$

and the claim follows. $\Box$

Combining Claims 3.2.7, 3.2.4 and 3.2.8, we get a lower bound on the number of matrices which are non-perfect for columns. Namely,

Claim 3.2.9 *Suppose that Eq. (9) hold and let $\gamma_2 \stackrel{\text{def}}{=} \gamma_1 - \frac{\delta_2}{2}$. Then, at least a $\frac{\gamma_2}{3}$ fraction of the matrices are non-perfect for columns.*

proof: By Claims 3.2.7 and 3.2.4, at least a $\frac{\gamma_1}{k} - \frac{\delta_2}{2k}$ $(= \frac{\gamma_2}{k})$ fraction of the row-sequences are both annoying and typical. Let us consider a set of exactly $\frac{\gamma_2}{k} \cdot |\mathbf{R}|$ such row-sequences, denoted A. Mimicking again the counting argument part of Claim 2.2.16, we bound, for each $\bar{r} \in A$, the fraction of non-perfect (for columns) matrices which contain $\bar{r}$ but no other row-sequence in A. Using an adequate setting of δ_2 and γ_2, this fraction is at least $\frac{1}{3}$. Summing the bounds achieved for all $\bar{r} \in A$, the claim follows. $\Box$

Using a suitable choice of γ (as a function of ϵ), Claim 3.2.9 contradicts Claim 3.2.2(c), and so Eq. (9) can not hold. The lemma follows. ■

3.4 Application to Low-Degree Testing

Again, the set of all k-by-k-by-k arrays over S satisfies Axioms 1–5. A more useful set of 3-dimensional arrays is defined as follows.

Construction 3.3 (main construction): *Let* $\mathbf{M}$ *be as in the Basic Construction* (*i.e., Construction 2.3*). *We let* $\mathcal{M}_{\rm c} = \mathcal{M}_{\rm r}$ *be the set of matrices defined by applying the Basic Construction to the element-set* $\mathbf{C} = \mathbf{R}$. *Specifically, a matrix in* $\mathcal{M}_{\rm c}$ *is defined by the quadruple* (x, y, x', y'), *where each of the four elements is a pair over* S, *so that the* $(i,j)^{\rm th}$ *entry in the matrix equals* $(x + jy) + i(x' + jy')$. *Here* x, y, x', y' *are viewed as two-dimensional vectors over the finite field* S *and* i, j *are scalars in* S. *The* $(i,j)^{\rm th}$ *entry is a pair over* S *which represents a pairwise independent sequence* (*which equals an element in* $\mathbf{C} = \mathbf{R}$).

Clearly,

Claim 3.4 *Construction 3.3 satisfies Assumptions 1–5.*

Combining all the above with the low-degree test of [GLRSW, RS96] using the results claimed there[7], we get a low-degree test which is sufficiently efficient to be used in the proof of the PCP-Characterization of NP.

Construction 3.5 (Low Degree Test): *Let* $f : F^n \mapsto F$, *where* F *is a field of prime cardinality, and* d *be an integer so that* $|F| > 4(d+2)^2$. *Let* $\mathbf{M}$, $\mathcal{M}_{\rm c}$ *and* $\mathcal{M}_{\rm r}$ *be as in Construction 3.3, with* $S = F^n$, $V = F$ *and* $k \stackrel{\rm def}{=} 4(d+2)^2$. *Let* $\Gamma : \mathbf{M} \mapsto F^{k\times k}$, $\Gamma_{\rm r} : \mathcal{M}_{\rm r} \mapsto F^{k^3}$ *and* $\Gamma_{\rm c} : \mathcal{M}_{\rm c} \mapsto F^{k^3}$ *be auxiliary tables* (which should contain the corresponding f-values). *The low degree test consists of*

1. *Applying the Extended Consistency Test to* $\Gamma : \mathbf{M} \mapsto F^{k\times k}$, $\Gamma_{\rm r} : \mathcal{M}_{\rm r} \mapsto F^{k^3}$ *and* $\Gamma_{\rm c} : \mathcal{M}_{\rm c} \mapsto F^{k^3}$.

2. *Selecting uniformly a matrix* $m \in \mathbf{M}$ *and testing that the* Polynomial Interpolation Condition (cf., [GLRSW]) *holds for each row; namely, we test that*
$$\sum_{i=1}^{d+2} \alpha_i \cdot \mathrm{entry}_{i,j}(\Gamma(m)) = 0$$
for all $j \in [k]$, *where* $\alpha_i = (-1)^i \cdot \binom{d+1}{i-1}$.

3. *Select uniformly a matrix in* $\mathbf{M}$ *and test matching of random entry to* f. *Namely, select uniformly* $m \in \mathbf{M}$, *and* $i, j \in [k]$, *and check if* $\mathrm{entry}_{i,j}(\Gamma(m)) = f(\mathrm{entry}_{i,j}(m))$.

The test accepts if and only if all the above three sub-tests accept.

Proposition 3.6 *Let* $f : F^n \mapsto F$, *where* F *is a field, and let* $\ell \stackrel{\rm def}{=} n \cdot \log_2 |F|$. *Then, the Low Degree Test of Construction 3.5 requires* $O(\ell)$ *randomness and query length,* $\mathrm{poly}(\ell)$ *answer length and satisfies:*

completeness: *If* f *is a degree-*d *polynomial, then there exist* $\Gamma : \mathbf{M} \mapsto F^{k\times k}$, $\Gamma_{\rm r} : \mathcal{M}_{\rm r} \mapsto F^{k^3}$ *and* $\Gamma_{\rm c} : \mathcal{M}_{\rm c} \mapsto F^{k^3}$ *so that the test always accepts.*

[7] Rather than using much stronger results obtained via a more complicated analysis, as in [ALMSS], which rely on the Lemma of [AS].

soundness: *For every $\delta > 3/(d+2)^2$ there exists an $\epsilon > 0$ so that for every f which is at distance at least δ from any degree-d polynomial and for every $\Gamma : \mathbf{M} \mapsto F^{k\times k}$, $\Gamma_{\mathrm{r}} : \mathcal{M}_{\mathrm{r}} \mapsto F^{k^3}$ and $\Gamma_{\mathrm{c}} : \mathcal{M}_{\mathrm{c}} \mapsto F^{k^3}$, the test rejects with probability at least ϵ. Furthermore, the constant ϵ is a polynomial in δ which does not depend on n, d and F.*

As a corollary, we get Lemma 1.3.

proof: As usual, the completeness clause is easy to establish. We thus turn to the soundness requirement. By Claim 3.4, we may apply Lemma 3.2 to the first sub-test and infer that either the first sub-test fails with some constant probability (say ϵ_1) or there exists a function $\tau : F^n \mapsto F$ so that with very high constant probability (say $1 - \delta_1$)

$$\mathrm{entry}_{i,j}(\Gamma(m)) = \tau(\mathrm{entry}_{i,j}(m)) \tag{12}$$

holds for all $i \in [d+2]$ and $j \in [k]$. On the other hand, by [GLRSW] (see also [S95, Thm 3.3] and [RS96, Thm 5]), either

$$\mathrm{Prob}_{x,y\in F^n}\left(\sum_{i=1}^{d+2} \alpha_i \cdot \tau(x+iy) \neq 0\right) > \frac{1}{2(d+2)^2} \tag{13}$$

or τ is very close (specifically at distance at most $1/(d+2)^2$) to some degree-d polynomial. A key observation is that the Main Construction (i.e., Construction 3.3) has the property that rows in $m \in_{\mathrm{R}} \mathbf{M}$ are distributed identically to the distribution in Eq. (13). Thus, for every $j \in [k]$ either

$$\mathrm{Prob}_{m\in\mathbf{M}}\left(\sum_{i=1}^{d+2} \alpha_i \cdot \tau(\mathrm{entry}_{i,j}(m)) \neq 0\right) > \frac{1}{2(d+2)^2} \tag{14}$$

or τ is at distance at most $\delta_2 \stackrel{\mathrm{def}}{=} 1/(d+2)^2$ from some degree-d polynomial. However, we claim that in case Eq. (14) holds, the second sub-test will reject with constant probability. The claim is proven by first considering $k = 4(d+2)^2$ copies of the GLRSW Test (i.e., the test in Eq. (14)). Using Chebishev's Inequality and the hypothesis by which each copy rejects with probability at least $1/2(d+2)^2$, we conclude that the probability that none of these copies rejects is bounded above by $\frac{2(d+2)^2}{4(d+2)^2} = \frac{1}{2}$. Thus, the second sub-test must reject with probability at least $\epsilon_2 \stackrel{\mathrm{def}}{=} \frac{1}{2} - \delta_1$, where δ_1 accounts for the substitution of the τ values by the entries in $\Gamma(\cdot)$. We conclude that τ must be δ_2-close to a degree-d polynomial or else the test rejects with too high probability (i.e., ϵ_2). Finally, we claim that if f disagrees with τ on $\delta_3 > \delta_1$ of the inputs then the third sub-test rejects with probability at least $\epsilon_3 \stackrel{\mathrm{def}}{=} \delta_3 - \delta_1$ (since the distance from f to τ is bounded by the sum of the distances of f to the matrix and of τ to the matrix). The proposition follows using some arithmetics: Specifically, we set $\delta_1 = \delta/3$, $\delta_3 = 2\delta_1$, $\epsilon_1 = \mathrm{poly}(\delta_1)$ (as in Lemma 3.2), and verify that $\delta_3 + \delta_2 \leq \delta$ (since $\delta/3 \geq (d+2)^{-2} = \delta_2$), and $\epsilon = \min\{\epsilon_1, \epsilon_2, \epsilon_3\} = \mathrm{poly}(\delta)$ (since $\epsilon_2 \geq 1/6$ and $\epsilon_3 = \delta/3$). ■

4 Proof of Lemma 1.1

There should be an easier and direct way of proving Lemma 1.1. However, having proven Lemma 2.2, we can apply it[8] to derive a short proof of Lemma 1.1. To this end we view ℓ-multisets over S

[8] This is indeed an over-kill. For example, we can avoid all complications regarding shifts (in the proof of Lemma 2.2).

as k-by-k matrices, where $k = \sqrt{\ell}$. Recall that the resulting set of matrices satisfies Axioms 1–4. Thus, by Lemma 2.2, in case the test accepts with probability at least $1-\epsilon$, there exists a function $\tau : S \mapsto V$ such that

$$\mathrm{Prob}_{A\in_{\mathrm{R}} S^k, B\in_{\mathrm{R}} E_{k^2}(A)}(\forall e \in A,\, \Gamma(B)_e = \tau(e)) \geq 1-\delta$$

where S^k is the set of all k-multisets over S and $E_l(A)$ is the set of all l-multisets extending A. We can think of this probability space as first selecting $B \in_{\mathrm{R}} S^{k^2}$ and next selecting a k-subset A in B. Thus,

$$\mathrm{Prob}_{B\in_{\mathrm{R}} S^{k^2}, A\in_{\mathrm{R}} C_k(B)}(\exists e \in A \text{ s.t. } \Gamma(B)_e \neq \tau(e)) \leq \delta \tag{15}$$

where $C_k(B)$ denotes the set of all k-multisets contained in B. This implies

$$\mathrm{Prob}_{B\in_{\mathrm{R}} S^{k^2}}(|\{e \in B : \Gamma(B)_e \neq \tau(e)\}| > k) \leq 2\delta$$

as otherwise Eq. (15) is violated. (The probability that a random k-subset hits a subset of density $\frac{1}{k}$ is at least $\frac{1}{2}$.) The lemma follows.

A previous version of this paper [GS96] has stated a stronger version of Lemma 1.1, where the sequences $F(x_1, ..., x_\ell)$ and $(f(x_1), ..., f(x_\ell))$ are claimed to be identical (rather than different on at most k locations), for a $1-\delta$ fraction of all possible $(x_1, ..., x_\ell) \in S^\ell$. Unfortunately, the proof given there was not correct – a mistake in the concluding lines of the proof of Claim 4.2.9 was found by Madhu Sudan. Still we conjecture that the stronger version holds as well, and that it can be established by a test which examines two random $(2k-1)$-extensions of a random k-subset.

Acknowledgment

We are grateful to Madhu Sudan for pointing out an error in an earlier version, and for other helpful comments.

References

[ALMSS] S. Arora, C. Lund, R. Motwani, M. Sudan and M. Szegedy. Proof Verification and Intractability of Approximation Problems. In *33rd IEEE Symposium on Foundations of Computer Science*, pages 14–23, 1992.

[AS] S. Arora and S. Safra. Probabilistic Checkable Proofs: A New Characterization of NP. In *33rd IEEE Symposium on Foundations of Computer Science*, pages 1–13, 1992.

[B94] L. Babai. Transparent Proofs and Limits to Approximation. TR-94-07, Dept. of Computer Science, University of Chicago, 1994.

[BFL] L. Babai, L. Fortnow, and C. Lund. Non-Deterministic Exponential Time has Two-Prover Interactive Protocols. In *31st IEEE Symposium on Foundations of Computer Science*, pages 16–25, 1990.

[BFLS] L. Babai, L. Fortnow, L. Levin, and M. Szegedy. Checking Computations in Polylogarithmic Time. In *23rd ACM Symposium on the Theory of Computing*, pages 21–31, 1991.

[BF] D. Beaver and J. Feigenbaum. Hiding Instances in Multioracle Queries. In *7th Symposium on Theoretical Aspects of Computer Science*, Springer Verlag, LNCS Vol. 415, pages 37–48, 1990.

[BGS] M. Bellare, O. Goldreich and M. Sudan. Free Bits, PCPs and Non-Approximability – Towards Tight Results. Extended abstract in *36th IEEE Symposium on Foundations of Computer Science*, pages 422–431, 1995. Full version available from ECCC at URL `http://www.eccc.uni-trier.de/eccc/`.

[BGLR] M. Bellare, S. Goldwasser, C. Lund and A. Russell. Efficient Probabilistically Checkable Proofs and Applications to Approximation. In *25th ACM Symposium on the Theory of Computing*, pages 294–304, 1993.

[BS] M. Bellare and M. Sudan. Improved Non-Approximability Results. In *26th ACM Symposium on the Theory of Computing*, pages 184–193, 1994.

[BGKW] M. Ben-Or, S. Goldwasser, J. Kilian and A. Wigderson. Multi-Prover Interactive Proofs: How to Remove Intractability. In *20th ACM Symposium on the Theory of Computing*, pages 113–131, 1988.

[BLR] M. Blum, M. Luby and R. Rubinfeld. Self-Testing/Correcting with Applications to Numerical Problems. In *22nd ACM Symposium on the Theory of Computing*, 1990.

[FGLSS] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy. Approximating Clique is almost NP-complete. In *32nd IEEE Symposium on Foundations of Computer Science*, pages 2–12, 1991.

[FRS] L. Fortnow, J. Rompel and M. Sipser. On the Power of Multi-Prover Interactive Protocols. In *Proc. 3rd IEEE Symp. on Structure in Complexity Theory*, pages 156–161, 1988.

[FS] K. Friedl and M. Sudan. Some Improvement to Total Degree Tests. In *Proc. 3rd Israel Symp. on Theory of Computing and Systems*, pages 190–198, 1995.

[GLRSW] P. Gemmell, R. Lipton, R. Rubinfeld, M. Sudan, and A. Wigderson. Self-Testing/Correcting for Polynomials and for Approximate Functions. In *23th ACM Symposium on the Theory of Computing*, pages 32–42, 1991.

[G96] O. Goldreich. A Taxonomy of Proof Systems. To appear in *Complexity Theory Retrospective II*, L.A. Hemaspaandra and A. Selman (eds.), Springer-Verlag, 1996. A copy is available from URL `http://theory.lcs.mit.edu/~oded/`.

[GMW] O. Goldreich, S. Micali and A. Wigderson. Proofs that Yield Nothing but their Validity or All Languages in NP Have Zero-Knowledge Proof Systems. Extended abstract in *27th FOCS*, 1986.

[GS96] O. Goldreich and M. Safra. A Combinatorial Consistency Lemma with application to proving the PCP Theorem. ECCC TR96-047, Version 1, September 1996.

[GMR] S. Goldwasser, S. Micali and C. Rackoff. The Knowledge Complexity of Interactive Proof Systems. *SIAM Journal on Computing*, Vol. 18, pages 186–208, 1989.

[H96] J. Hastad. Clique is Hard to Approximate within $n^{1-\epsilon}$. In *37th IEEE Symposium on Foundations of Computer Science*, 1996.

[LS] D. Lapidot and A. Shamir. Fully Parallelized Multi Prover Protocols for NEXP-time. In *32nd IEEE Symposium on Foundations of Computer Science*, pages 13–18, 1991.

[LFKN] C. Lund, L. Fortnow, H. Karloff, and N. Nisan. Algebraic Methods for Interactive Proof Systems. In *31st IEEE Symposium on Foundations of Computer Science*, pages 2–10, 1990.

[RaSa] R. Raz and S. Safra. A Sub-Constant Error-Probability PCP Characterization of NP. To appear in *29th ACM Symposium on the Theory of Computing*, 1997.

[RS92] R. Rubinfeld and M. Sudan. Testing Polynomial Functions Efficiently and over Rational Domains. In *3rd ACM–SIAM Symposium on Discrete Algorithms*, pages 23–32, 1992.

[RS96] R. Rubinfeld and M. Sudan. Robust Characterization of Polynomials with Application to Program Testing. *SIAM J. of Computing*, Vol. 25, No. 2, pages 252–271, 1996. This paper is the journal version of [GLRSW, RS92].

[S90] A. Shamir. IP=PSPACE. In *31st IEEE Symposium on Foundations of Computer Science*, pages 11–15, 1990.

[S95] M. Sudan. *Efficient Checking of Polynomials and Proofs and the Hardness of Approximation Problems.* ACM Distinguished Theses, Springer-Verlag, LNCS Vol. 1001, 1995.