Building Assignment Testers–DRAFT

Andrew Drucker
University of California, San Diego

Abstract
In this expository paper, we show how to construct Assignment Testers using the Hadamard and quadratic encodings. These algorithms play a key role in Dinur’s recent proof of the PCP theorem.

1. Introduction.
The PCP Theorem of [ALMSS] was a milestone achievement of theoretical computer science. This theorem states that any language $L \in \text{NP}$ (for a definition of $\text{NP}$ and other basic notions in complexity theory, see, e.g. [Pap]) has a 1-round probabilistic proof system consisting of a protocol between a prover $P$ and a polynomially-bounded verifier algorithm $V$. The protocol has the following form and properties:

Given mutual access to an input bitstring $x$, $|x| = n$, $P$ sends a ‘certificate’ bitstring $y$ to $V$ in an attempt to prove that $x \in L$ (which may or may not actually be true). The certificate is ‘short’, i.e., $|y| = O(poly(n))$.

Rather than reading the entire certificate $y$, $V$ generates $O(\log(n))$ random bits, and uses them to determine (in $\text{poly}(n)$ time) a constant ($O(1)$) number of bit-positions at which $V$ (nonadaptively) queries $y$. Based on what it sees, $V$ chooses (again in $\text{poly}(n)$ time) whether to accept or reject the input $x$.

The protocol has a ‘completeness’ property: if $x \in L$, then there exists a certificate $y$ such that if $P$ sends $y$, $V$ accepts $(x, y)$ with probability 1.

The protocol also has a ‘soundness’ property: if $x \in L$, for all certificates $y$, $V$ rejects $(x, y)$ with probability $\Omega(1)$.

The power of this theorem lies in the amazing quantitative strength of its parameters. If the verifier was allowed to query all of $y$, then $y$ might as well be a conventional $\text{NP}$ proof that $x \in L$. The PCP Theorem says something much stronger: we can convince the verifier to ‘within a reasonable doubt’ that $x \in L$ by presenting $V$ with a proof that $V$ barely looks at!

In 2005, Irit Dinur gave a highly original and much simpler proof of the PCP Theorem. The main technical tool in Dinur’s approach is the use and analysis of random walks on a special kind of graphs called expanders, whose powerful properties substitute for the algebraic ‘heavy lifting’ of the original proof. Her proof does, however, retain some of the coding and algebraic-testing ideas used in the original, and these ideas may still be challenging for many readers.
This expository paper aims to present the key algebraic step, the construction of an algorithm called an assignment tester. There are at least two known approaches to building assignment testers, both based on error-correcting codes. The first, used by Dinur, is based on a code called the Long Code, and the proof of its validity (building on work by Håstad [Hås]) uses Fourier analysis. The second approach, which this paper will present, is based on two related codes called the Hadamard and quadratic encodings. This approach is also described in a survey on Dinur’s proof by Radhakrishnan and Sudan ([RS]).

In writing this I have also drawn substantially on the lecture notes from weeks 7-9 of a U. Washington course on PCPs taught by Ryan O’Donnell and Venkatesan Guruswami, available at

http://www.cs.washington.edu/education/courses/533/05au/

My expository goals are twofold: First, to introduce a flexible notion of soundness of tests (‘blockwise soundness’) that I hope gives a clear sense of the relationship of assignment testers to the component tests used in the construction, and more generally to the tests studied in the field called ‘property testing’ ([Fis]).

Second, to streamline the construction of assignment testers by giving a general ‘composition-of-tests’ lemma (Lemma 1) that allows for the hierarchically arranged subtests of the assignment tester to be analyzed in isolation, then brought together according to general principles of test-composition.

The construction still is not particularly simple to describe. However, I hope this presentation will make some of the recurring patterns of testing design and analysis more explicit, and contribute to the general understanding of Dinur’s proof.

2. Assignment Testers.

Definition (adapted from [Din]). A k-query assignment tester is an algorithm $A$ that takes as input a boolean circuit $\Phi(x)$ on some (arbitrarily sized) variable-set $X$ and outputs a list $Y$ of new boolean variable-names and a description of a randomized $k$-query ‘test’ algorithm $t_\Phi(X,Y)$, such that

-(Completeness) If $\Phi(a) = 1$, there exists some assignment $b$ to $Y$ such that $t_\Phi(a,b)$ accepts with probability 1;

-(Soundness) If $d(a,\Phi^{-1}(1)) \geq c \cdot |X|$, $c > 0$, then for all assignments $b$ to $Y$, $t_\Phi(a,b)$ rejects with probability $\geq \Omega(c)$ (over its random choices).

(here $d(\cdot,\cdot)$ denotes Hamming distance, and if $\Phi^{-1}(1)$ is an empty set, this distance is defined as $|X|$).

No restrictions are placed on the runtime of $A$, of $t_\Phi$ or the size of $A$’s output, including $|Y|$.

Main Theorem. There exists an explicit 4-query assignment tester.
The use of assignment testers in Dinur’s proof is described well in the original paper and in the other references mentioned, so we will not discuss it. However, to motivate (but not explain) the importance of assignment testers, we note that they already embody a kind of PCP: given $\Phi(x)$, suppose prover $P$ wants to convince verifier $V$ of the fact that $\Phi$ is satisfiable (circuit satisfiability is a canonical NP-complete problem). Define a protocol where $P$ sends $V$ an assignment $(a, b)$ to $X \cup Y$, where $Y$ are the auxiliary variables returned by $A(\Phi)$. $V$ then runs the $k$-query test $t_\Phi(a, b)$.

If $\Phi$ is in fact satisfiable, say by $a$, then by the completeness property of $A$ there exists a $b$ such that $t_\Phi(a, b)$ accepts with probability 1. This shows the completeness property of the PCP.

On the other hand, suppose $\Phi$ is unsatisfiable; then for any $a$, $d(a, \Phi^{-1}(1)) = |X|$, so that $t_\Phi(a, b)$ rejects with probability $\Omega(1)$. This shows the soundness of the PCP.

The assignment testers we give are efficiently computable, and use only $4 = O(1)$ queries, so all that prevents this construction from yielding the PCP Theorem itself is the fact that $|Y|$ will be exponential, not polynomial, in $|X|$.

**Proof of the Main Theorem:**

We first describe how the construction of a 4-query assignment tester reduces to the construction of a 4-query assignment tester for the special case where we are promised that the circuit $\Phi(x)$ checks that the assignment $a$ to $X$ satisfies some system $P_1(a), P_2(a), \ldots P_m(a)$ of homogeneous quadratic equations over $F_2$ (call such a specialized assignment tester a quadratic assignment tester).

The basic idea (which will need modifying) is, given any circuit $\Phi(x)$ (which need not be of the special type described), to create a auxiliary set of ‘gate-variables’ $X' = \{v_g\}$, one for each non-input gate $g \in \Phi$. Denote also by $v_g$ the variable in $X$ corresponding to an input gate $g$. We can then ‘arithmetize’ $\Phi$ to form a system $P_\Phi$ of quadratic equations over $X \cup X'$, as follows:

- For each $\land$ gate $g = g_1 \land g_2$, add an equation $x_g = x_{g_1} \cdot x_{g_2}$, which we ‘homogenize’, i.e. write as $x_g + x_{g_1} \cdot x_{g_2} = 0$ (we are working mod 2, so $+, -$ are equivalent);

- For each $\lor$ gate $g = g_1 \lor g_2$, add an equation $x_g = 1 - (1 - x_{g_1}) \cdot (1 - x_{g_2})$, which we write as $x_g + x_{g_1} \cdot x_{g_2} = 0$;

- For each ‘NOT’ gate $g = \neg g_1$, add an equation $x_g = 1 + x_{g_1}$, which we write as $x_g + x_{g_1} + 1 = 0$;

- For the output gate $g$, add an equation $x_g = 1$, i.e. $x_g + 1 = 0$.

It is easily verified that if $\Phi(a) = 1$, there exists an assignment $a'$ to $X'$ such that $(a, a')$ satisfy the equations of $P_\Phi$: set each $v_g \in X'$ equal to the output of $g$ on the computation $\Phi(a)$. On the other hand, if $\Phi(a) = 0$, no assignment $a'$ can satisfy all equations in $P_\Phi$.

Here is a ‘first try’ for our reduction to quadratic assignment testing: given
an arbitrary circuit \( \Phi(x) \), construct a circuit \( \Phi'(x, x') : X \cup X' \rightarrow \{0,1\} \) which

decides whether input \((x, x')\) satisfies the equations \( P \).

Then, feed \( \Phi'(x, x') \) to the quadratic assignment tester \( A_q \), yielding a test \( t_0 \)
on \( X \cup X' \cup Y \), where \( Y \) is the new auxiliary variable-set created by \( A_q \). Finally,
return the same test \( t_0 \), except that \( X' \) is included in the auxiliary variable-set.

This construction possesses the completeness property of assignment testers:
if \( \Phi(a) = 1 \), then there exists an assignment \( a' \) to \( X' \) such that \( \Phi'(a, a') \) accepts
(namely, assign to each \( x_g \in X' \) its output in the computation \( \Phi(a) \)). Thus, by
the assumed completeness of \( A_q \), there exists an assignment \( b \) to \( Y \) such that
\( t_0(a, a', b) \) accepts with probability 1.

On the other hand, the soundness property of assignment testers may not be
satisfied. To see why, note that \( |X| \) may be very small in comparison to \( |X'| \), the
number of non-input circuit gates. Then even if \( d(a, \Phi^{-1}(1)) \geq c \cdot |X| \), we could
have \( d(a, a'), \Phi'^{-1}(1) \ll c \cdot (|X| + |X'|) \), if the computation \( \Phi(a) \) looked ‘very
similar’ gate-by-gate to some accepting computation \( \Phi(u) \). Then the soundness
property we assume in \( A_q \) would not guarantee a sufficient rejection probability
in the test we return.

This problem is not too difficult to overcome, and we describe a solution.
Essentially, we duplicate the input variables in an effort to increase their ‘weight’
in the circuit. Given \( \Phi(x) \), we produce a modified circuit \( \Gamma(x) \) which mirrors
\( \Phi(x) \) but which also includes, for each input variable \( x_i \in X \), \(|\Phi| \) many new \( \land 
\) gates \( g_{i,j} = x_i \land x_i \), for
\( j = 1, 2, \ldots |\Phi| \) (we include in \(|\Phi| \) a count of the input gates, i.e. \(|\Phi| = |X| + |X'|\)).

These non-output gates \( g_{i,j} \) are not inputs to any other gates; call them ‘verification-gates’. Denote by \( V = \{v_{i,j}\} \) the arithmetized verification-gates of
\( \Gamma \), with \( v_{i,j} \in V \) arithmetizing the \( j \)th verification gate for \( x_i \).
Let \( X' \) continue to refer to the arithmetized gates other than \( X \cup V \).

Modify our ‘first try’ assignment tester above by passing to \( A_q \), the quadratic
assignment tester, not \( \Phi'(x, x') \), but \( \Gamma'(x, x', v) \), the circuit which checks that
its input satisfies the system \( P \) of quadratic equations. Let \( t_0((x, x', v), y) \) be
the \( k \)-query test returned by \( A_q \), and let \( Y \) be the auxiliary variables introduced
by \( t_0 \).

Let \( A \) return the \((k\)-query\) test \( t \) which, on input \((a, (a', v, y))\) (an assignment
to \( X \cup (X' \cup V \cup Y) \)—note that \( X' \cup V \cup Y \) are the auxiliary variables returned
by \( t \)), generates a random bit \( b \).

If \( b = 0 \), \( t \) runs \( t_0((a, a', v), y) \), accepting iff \( t_0 \) accepts.

If \( b = 1 \), \( t \) picks an \( i \leq |X| \) and a \( j \leq |\Phi| \) at random, and checks that
\( a_i = v_{i,j} \).

Once again, completeness is easily established: Suppose \( \Phi(a) = 1 \). Define an
assignment \( a', v \) to \( X' \cup V \) by letting all arithmetized gate variables equal the output
of their corresponding gates on the computation \( \Gamma(a) \). Then the quadratic
3. Quadratic Assignment Testing.
To construct our quadratic assignment tester, we develop a more general view of tests, with a generalized notion of soundness called ‘blockwise soundness’ based on a different notion of distance between bitvectors, which we introduce now.

**Definition.** Fix integers \( l > 0 \) and \( m_1, m_2, \ldots, m_l > 0 \). Given a bitvector \( x \) of length \( m_1 + m_2 + \ldots + m_l \), consider \( x \) as composed of \( l \) ‘blocks’, with the \( i \)th block \( x_{[i]} \) having length \( m_i \). Given two such bitvectors \( x, y \), define the block-distance
\[
d_{\text{block}}(x, y) = \max_{1 \leq i \leq l} (||x_{[i]} - y_{[i]}|| / m_i),
\]
where \( ||z|| \) denotes the Hamming weight of \( z \).

It is easy to verify that \( d_{\text{block}} \) defines a metric on \( \{0,1\}^{m_1 + \ldots + m_l} \).

By a property \( S \) we mean simply a subset of \( \{0,1\}^\ast \), considered also as a boolean-valued function on bitstrings.

**Definition (Tests).** A \( k \)-query (randomized, nonadaptive) test \( t \) is an algorithm that, given \( x \in \{0,1\}^{m_1 + \ldots + m_l} \), determines at most \( k \) bit positions from \( \{1,2,\ldots,m_1 + \ldots + m_l\} \) (possibly with the aid of randomness), inspects the bits of \( x \) at these positions, and chooses to accept or reject \( x \).

If \( t \) always accepts any \( x \in S^{-1}(1) \), we say that \( t \) has **perfect completeness** for the property \( S \).

Fix a ‘promise’ set \( W \subseteq \{0,1\}^{m_1 + \ldots + m_l} \). If there exists an \( s > 0 \) such that for any \( c \in (0,1], t \) rejects any \( x \in W \) such that \( d_{\text{block}}(x, S^{-1}(1)) \geq c \) with probability at least \( s \cdot c \) (over the random choices made by \( t \)), we say that \( t \) has **blockwise soundness** \( s \) for \( S \) relative to the promise \( x \in W \).

If \( W = \{0,1\}^{m_1 + \ldots + m_l} \), we say that \( t \) has **blockwise soundness** \( s \) for \( S \) without **any promise**. (We could define completeness relative to a promise, but we will not, since all tests in this paper will possess completeness without any promise.)

We can now give the connection between blockwise-sound testing and assignment testers.

**Proposition 1.** Suppose that, given a circuit \( \Phi \) describing a system \( P \) of homogeneous quadratic equations on \( n \) variables, algorithm \( A_q \) produces a description of a property \( S(a, L, Q) \) and a \( k \)-query test \( t \) for \( S(a, L, Q) \), such that

1) \( S(a, L, Q) = 1 \) implies \( a \) satisfies the system \( P \), and for any \( a \) satisfying \( P \), there exist \( L, Q \) such that \( S(a, L, Q) \) holds;

2) \( t \) has perfect completeness and \( \Omega(1) \) blockwise soundness (where \( a, L, Q \) are the three blocks).

Then \( A_q \) is a \( k \)-query quadratic assignment tester.

**Proof:** Consider \( L, Q \) as the auxiliary variables introduced by \( A_q \). The completeness property of \( A_q \) follows immediately from the assumptions. For
soundness of $A_q$, suppose $d(a, \Phi^{-1}(1)) \geq c \cdot n$, $c > 0$.

Then for any $L, Q$,

$$d_{\text{block}}((a, L, Q), S^{-1}(1)) \geq c \text{ (since } S(a', L', Q') = 1 \text{ implies } \Phi(a') = 1),$$

so by the blockwise soundness of $t$, $t(a, L, Q)$ rejects with probability $\Omega(c)$.

Given the quadratic system $P$, we show how to define a property $S_{\text{good}}(a, L, Q)$ and a test $t$ satisfying the hypotheses of Proposition 1. If $a$ is of length $n$ (the number of variables over $F_2$), $L$ will be of length $2^n$ and is interpreted as a function $L(x)$ from $\{0, 1\}^n$ to $\{0, 1\}$; $Q$ will be of length $2^{n^2}$ and is interpreted as a function $Q(X)$ from $n$-by-$n$ 0/1 matrices to $\{0, 1\}$.

Here is the property $S_{\text{good}}(a, L, Q)$:

- $S_{\text{good}}(a, L, Q) = 1$ if and only if $a$ satisfies $P$, $L(x) \equiv a \cdot x$, and $Q(X) \equiv a^T X a$.

The functions $l(x) = a \cdot x$ and $q(X) = a^T X a$ are called the Hadamard encoding and quadratic encoding, respectively, of $a$.

It is clear that $S_{\text{good}}$ satisfies condition 1 in Proposition 1. The challenge lies in building the test $t$ for $S_{\text{good}}$. To do this we define a hierarchy of progressively more restrictive intermediate properties and build tests for each property in turn, each of whose blockwise soundness we analyze relative to the promise that the previous property is satisfied. Then we apply a general composition lemma that will combine the tests into a single test for $S_{\text{good}}$ meeting our requirements.

Specifically, here are the intermediate properties, in increasing order of strictness:

- Let $S_{\text{lin}}(a, L, Q) = 1$ iff $L, Q$ encode linear functions.
- Let $S_{\text{agree}}(a, L, Q) = 1$ iff

$$L(x) \equiv a \cdot x, \quad Q(X) \equiv a^T X a.$$  

Then, $S_{\text{good}}(a, L, Q) = 1$ iff $a$ satisfies $P$ and $S_{\text{agree}}(a, L, Q) = 1$.

We give tests, each with perfect completeness, for:
- $S_{\text{lin}}(a, L, Q)$, with $\Omega(1)$ blockwise soundness relative to no promise;
- $S_{\text{agree}}(a, L, Q)$, with $\Omega(1)$ blockwise soundness relative to the promise $S_{\text{lin}}(a, L, Q)$;
- $S_{\text{good}}(a, L, Q)$, with $\Omega(1)$ blockwise soundness relative to the promise $S_{\text{agree}}(a, L, Q)$.

Furthermore, the latter two tests will have the technical property that the position of each query made is individually uniformly distributed (as a random variable) over a block, after conditioning upon the block chosen. (Uniformity in the joint distribution on query positions is neither required for our purposes, nor achieved.)

These tests will allow the use of the following composition lemma:

**Lemma 1.** Let $R_1(x), R_2(x)$ be properties, such that $R_2(x) \Rightarrow R_1(x)$. Let $k > 0$ be an integer, and suppose that there exist $k$-query tests $t_1(x), t_2(x)$ for properties $R_1(x), R_2(x)$ respectively, such that
1) \( t_1, t_2 \) have perfect completeness for their respective properties;

2) \( t_1 \) has \( \Omega(1) \) blockwise soundness relative to the promise \((x \in W)\);

3) \( t_2 \) has \( \Omega(1) \) blockwise soundness relative to the promise \( R_1(x) \);

4) The query positions of \( t_2(x) \) are (as random variables) each individually uniformly distributed within a block of \( x \), after conditioning on the query block chosen (though the block chosen may or may not be predetermined for each query).

Then there exists a \( k \)-query test \( t'_2(x) \) for the property \( R_2(x) \) with perfect completeness, and with \( \Omega(1) \) blockwise soundness relative to the promise \((x \in W)\).

Note that once we have provided the tests enumerated before the Lemma, two applications of this Lemma will suffice to give a 4-query test for \( S_{\text{good}}(a, L, Q) \) with perfect completeness and \( \Omega(1) \) blockwise soundness relative to no promise:

Namely, if \( t_{\text{lin}}, t_{\text{agree}}, t_{\text{good}} \) are the 4-query tests for \( S_{\text{lin}}, S_{\text{agree}}, S_{\text{good}} \) respectively, with perfect completeness and with \( \Omega(1) \) blockwise soundness relative to the stated promises, we first apply Lemma 1 to \( t_1 = t_{\text{lin}}, t_2 = t_{\text{agree}} \), yielding a test \( t'_2 \) for \( S_{\text{agree}} \) with blockwise soundness \( \Omega(1) \) relative to no promise. \( t'_2 \) then plays the role of \( t_1 \) in a second application of Lemma 1, with \( t_{\text{good}} \) playing the role of \( t_2 \). The resulting derived test gives us the desired test \( t' \) for \( S_{\text{good}} \).

Thus, to apply Proposition 1 and get our quadratic assignment tester, completing the proof of the Main Theorem, it remains only to prove Lemma 1, then give and analyze the tests \( t_{\text{lin}}, t_{\text{agree}}, t_{\text{good}} \).

Proof of Lemma 1:

Let \( t'_2(x) \) act as follows: first, it picks a random bit \( b \in \{0, 1\} \). If \( b = 0 \), it simulates \( t_1(x) \); if \( b = 1 \), it simulates \( t_2(x) \). In either case it accepts if and only if the simulated test accepts.

We first claim \( t'_2(x) \) has perfect completeness for the property \( R_2(x) \). For, if it is capable of rejecting \( x \), it must do so while simulating \( t_1(x) \) or \( t_2(x) \). In the first case, \( x \) must not satisfy \( R_1 \), by the completeness of \( t_1 \) for \( R_1 \); since \( R_2 \Rightarrow R_1 \), this implies that \( x \) does not satisfy \( R_2 \). Similarly, if the simulation of \( t_2(x) \) rejects, \( x \) must not satisfy \( R_2 \) (by the completeness of \( t_2 \) for \( R_2 \)).

Now we analyze the soundness of \( t'_2(x) \) relative to the promise \((x \in W)\). Let test \( t_1(x) \) have soundness \( s_1 > 0 \) relative to the promise \((x \in W)\), and let \( t_2(x) \) have soundness \( s_2 > 0 \) relative to the promise \( R_1(x) \).

Suppose \( x \in W \) and \( d_{\text{block}}(x, R_2^{-1}(1)) \geq c_0 > 0 \).

Case 1: \( d_{\text{block}}(x, R_1^{-1}(1)) \geq \frac{2s_1}{3c_0} \).

Then by the blockwise soundness of \( t_1 \) under the promise \( x \in W \), with probability \( \geq \frac{1}{2}s_1 \cdot \frac{2s_1}{3c_0} \), \( b = 0 \) is chosen and \( t'_2 \), acting as \( t_1 \), proceeds to reject \( x \).
Case II: \( d_{\text{block}}(x, R^{-1}_1(1)) < \frac{s_2 c_0}{3k} \).

Then there exists \( x' \) in \( R^{-1}_1(1) \) such that \( d_{\text{block}}(x, x') < \frac{s_2 c_0}{3k} \).

Let us first consider the behavior of \( t_2(x) \). As the position \( i \) of each individual query \( q \) of \( t_2(x) \) is uniformly distributed over a block of \( x \) after conditioning on the block from which \( q \) is to be taken, and as each block agrees with \( x' \) on a \( > \frac{s_2 c_0}{3k} \) fraction of positions, with probability \( > 1 - \frac{s_2 c_0}{3k} \) it returns a value \( x_{ij}(j) \) that agrees with \( x' \), that is \( x_{ij}(j) = x'_{ij}(j) \).

Then by the intersection bound, with probability greater than \( 1 - (\frac{s_2 c_0}{3k})^j \), \( \text{lin}_ aggregation \text{ blocking} \text{ soundness} \).

\( \bigwedge(\text{block} \text{ agreeing} \text{ with} \text{ both} \text{ blocks}) \).

Thus, by the soundness of \( t_2(y) \) as a test for \( R_2 \) relative to the promise \( y \in R_1 \), \( t_2(x') \) rejects with probability at least \( c_0(1 - \frac{s_2 c_0}{3k})^j \).

By the intersection bound again, it follows that \( t_2(x) \) must reject with probability at least \( c_0(1 - \frac{s_2 c_0}{3k})^j - \frac{s_2 c_0}{3k} \), so that \( t'_2(x) \) chooses \( b = 1 \) and subsequently rejects with probability at least

\[
\frac{1}{2} \left( c_0(1 - \frac{s_2 c_0}{3k})^j - \frac{s_2 c_0}{3k} \right)
= c_0 \cdot \left( \frac{s_2 c_0}{3k} \right)
\]

Combining our two cases, with probability at least \( \frac{2}{3} c_0 \), \( t'_2(x) \) rejects. Thus \( t'_2 \) has blockwise soundness \( \frac{2}{3} c_0 \) relative to the promise \( (x \in W) \), as required. \( \diamond \)

Now we develop and analyze the promised tests \( t_{\text{lin}}, t_{\text{agree}}, t_{\text{good}} \) to complete the proof of the Main Theorem. Our starting point is the Linearity Test of Blum, Luby, and Rubinfeld ([BLR]).

Let \( F \) be a bitstring of length \( 2^m \), interpreted as a boolean function on \( m \) input bits. Say \( F \) is linear if for all \( x, y \in \{0,1\}^m \), \( F(x) + F(y) = F(x + y) \).

Let \( t_{BLR}(F) \) be the (2-query) test that chooses \( x, y \) at random and accepts if \( F(x) + F(y) = F(x + y) \).

**Lemma 2 (Linearity Test) [BLR].** The test \( t_{BLR} \) accepts linear functions with probability 1. Furthermore, if \( F \) is at a Hamming distance at least \( c \cdot 2^m \) from any linear function \( (c > 0) \), \( t_{BLR}(F) \) rejects with probability at least \( c \).

**Proof:** Omitted, see [BLR]. \( \diamond \)

**Corollary 1.** There is a 3-query test \( t_{\text{lin}}(a, L, Q) \) which has perfect completeness for the property \( S_{\text{lin}}(a, L, Q) \), and has blockwise soundness \( \frac{1}{2} \) (with no promise needed). The query positions are uniformly distributed over \( L \) or \( Q \) once their block is determined.
**Proof:** Let \( t_{lin}(a, L, Q) \) choose \( b \in \{0, 1\} \) at random. If \( b = 0 \), \( t_{lin} \) runs the BLR linearity test on \( L \); if \( b = 1 \), \( t_{lin} \) runs the BLR test on \( Q \). In either case, \( t_{lin} \) accepts iff the BLR test accepts.

Completeness follows from completeness of the BLR test; the uniform distribution property is likewise inherited from the BLR test. For blockwise soundness, if \( d_{\text{block}}((a, L, Q), S_{lin}^{-1}(1)) \geq c \), then, as \( a \) is irrelevant to the property \( S_{lin} \), one of \( L, Q \) must be \( c \)-far from any linear function. Then with probability at least \( \frac{1}{2} \), \( t_{lin} \) chooses the \( c \)-far block and rejects (by the soundness property of the BLR test). This shows that \( t_{lin} \) has blockwise soundness \( \frac{1}{2} \). \( \diamond \)

Note that the joint query distribution in this test is not uniform—the third query position is determined by the first two—but that, as mentioned before, joint uniformity is not required in applying Lemma 1.

Note also that, in the proof of Corollary 1, we are applying the same logic as in Lemma 1 to combine tests of \( L \) and of \( Q \); indeed, we could have defined a further intermediate property and applied Lemma 1, but this would not have yielded much simplification since the two tests are essentially independent.

We also used a similar composition idea during the reduction of assignment testing to quadratic assignment testing when we generated the random bit \( b \) in that setting.

**Lemma 3.** There is a 4-query test \( t_{agree}(a, L, Q) \) which has perfect completeness for the property \( S_{agree}(a, L, Q) \), and has blockwise soundness \( \frac{1}{8} \) relative to the promise \( S_{lin}(a, L, Q) \). Furthermore, positions of queries made by \( t_{agree} \) are individually uniformly distributed over a block once their block is determined.

(As should become clear, in building this test we are suppressing another opportunity to define a further intermediate property and apply Lemma 1 an additional time.)

**Proof:**

Let \( t_{agree}(a, L, Q) \) generate a random bit \( b \). If \( b = 0 \), \( t_{agree} \) chooses a random index \( i \leq n \) and a random \( n \)-bit vector \( x \), and checks that \( a_i = L(x) + L(x + e_i) \), where \( e_i \) is the \( i \)th unit vector.

If \( b = 1 \), let \( t_{agree} \) pick random length-\( n \) bitvectors \( s, s' \) and random \( n \)-by-\( n \) matrix \( X \), and accept iff \( L(s)L(s') + Q(X) + Q(X + s(s')^T) = 0 \) (outer addition is mod 2).

(Note that for \( v, v' \in \{0, 1\}^n \), \( v^Tv' \) is a dot product yielding a single bit, whereas \( v(v')^T \) is an \( n \)-by-\( n \) matrix whose \((i, j)\)th entry is \( v_i v'_j \).)

The fact that \( t_{agree} \) makes at most 4 queries, individually distributed either over \( a \), \( L \), or \( Q \) after conditioning on the block chosen, is clear (\( X + Y \) is uniformly distributed whenever \( X \) is).

We show completeness. If \( S_{agree}(a, L, Q) = 1 \), then \( L(x) \equiv a^T \cdot x, Q(X) \equiv \)
so the test accepts. If $b = 1$, then for all $s, s'$, and $X$,

$$L(s)L(s') + Q(X) + Q(X + s(s')^T) = (\Sigma_{i=1}^n a_i s_i) + (\Sigma_{i=1}^n a_i s_i') + a^T X a + a^T (X + s(s')^T) a$$

$$= \Sigma_{i,j \leq n} (a_i a_j) (s_i s_j') + a^T (s(s')^T) a$$

$$= a^T (s(s')^T) a' + a^T (s(s')^T) a = 0.$$

Thus $t_{agree}(a, L, Q)$ must again accept. This shows completeness of $t_{agree}$.

Now we prove the blockwise-soundness claim. Suppose the promise $S_{lia}(a, L, Q)$ is met but $S_{agree}(a, L, Q)$ is false, and $d_{block}((a, L, Q), S^{-1}_{agree}(1)) \geq c > 0$.

We introduce some notation. If $L, Q$ are linear functions, then there exist $w_L \in \{0, 1\}^n$, $M_Q = (q_{i,j}) \in \{0, 1\}^{n^2}$ such that

$$L(x) \equiv w_L^T x, \quad Q(X) \equiv \Sigma_{i,j \leq n} (q_{i,j} x_{i,j}).$$

Case I: $M_Q \neq w_L w_L^T$.

Then $L(s)L(s') + Q(X) + Q(X + s(s')^T) = (s^T \cdot w_L)(w_L^T \cdot s') + Q(s(s')^T)$

$$= s^T (w_L \cdot w_L^T) s + \Sigma_{i,j \leq n} q_{i,j} s_i s_j' = s^T (w_L w_L^T + M_Q) s',$$

and $(w_L w_L^T + M_Q)$ is a nonzero matrix.

Claim. Let $M \in \{0, 1\}^{n^2}$ be a nonzero $n$-by-$n$ matrix. With probability at least $\frac{1}{4}$ over choice of $s, s' \in \{0, 1\}^n$, $s^T Ms' \neq 0$.

Proof of Claim: Let $e_i$ denote the $i$th unit vector. Suppose the $(i, j)$th entry of $M$ is nonzero. Now $(M)_{i,j} = e_i^T M e_j$

$$= s^T M s' + (s + e_i)^T M s' + s^T M (s' + e_j) + (s + e_i)^T M (s' + e_j)$$

for any $s, s' \in \{0, 1\}^n$.

But each of the four pairs of vectors appearing in the terms above, i.e. \{$(s, s'), (s + e_i, s'), (s, s' + e_j), (s + e_i, s' + e_j)$\} are individually uniformly distributed over $\{0, 1\}^n \times \{0, 1\}^n$ since $(s, s')$ are.

With probability 1, at least one of the four terms is nonzero (since the sum above always evaluates to $(M)_{i,j} = 1$), each term, including $s^T M s'$, is individually nonzero with probability at least $\frac{1}{4}$. ∘

So with probability at least $\frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$, the $t_{agree}$ chooses $b = 1$ and rejects $(a, L, Q)$.

Case II: $w_L w_L^T = M_Q$.

Then for $d_{block}((a, L, Q), S^{-1}_{agree}(1)) \geq c$ we must have $||a - w_L|| \geq c \cdot n$. In this case, with probability $\geq \frac{1}{2} \cdot c$, $b = 0$ is chosen and $i$ is chosen such that (for
all $x$) $L(x + e_i) + L(x) = w^T_L \cdot e_i \neq a_i$, and $t_{agree}$ rejects.

Combining the two cases, we conclude that $t_{agree}$ rejects $(a, L, Q)$ with probability $\geq \min(\frac{1}{g}, \frac{c}{2}) \geq \frac{c}{g}$. This shows the $\frac{1}{g}$-blockwise soundness of $t_{agree}$. ◆

Lemma 4. There is a 4-query test $t_{good}(a, L, Q)$ for the property $S_{good}(a, L, Q)$ which has perfect completeness, and has blockwise soundness $\frac{1}{g}$ relative to the promise $S_{agree}(a, L, Q)$. Furthermore, queries made by $t_{good}$ are individually uniformly distributed over a predetermined block.

Proof:
First, we need some definitions and analysis. For each $r \in \{0,1\}^n$, define $P_r(a) = \sum_{i=1}^n r_i \cdot P_i(a)$. As each $P_i(a)$ is quadratic in $a$, so too is $P_r(a)$, and we can express $P_r(a)$ as

$$P_r(a) = b + \sum_{i=1}^n v_i a_i + \sum_{i=1}^n w_{i,j} a_i a_j$$

for $b \in \{0,1\}, v = \{v_i\} \in \{0,1\}^n, W = \{w_{i,j}\} \in \{0,1\}^{n^2}$ all depending on $r$.

Now we define the test $t_{good}(a, L, Q)$: $t_{good}$ picks a random $r \in \{0,1\}^n$. Let $b, v, W$ be the variables defined above, depending on $r$, so that $P_r(a) = b + a^T \cdot v + a^T W a$.

$t_{good}$ picks $x \in \{0,1\}^n, Y \in \{0,1\}^{N^2}$ at random, and accepts iff

$$b + L(v + x) + L(x) + Q(Y + W) + Q(Y) = 0.$$

$t_{good}$ is clearly a 4-query test. Moreover, the individual queries of $t_{good}$ are uniformly distributed either over the entries of $L$ or of $Q$.

For completeness, note that if $S_{good}(a, L, Q)$ holds, then $S_{agree}(a, L, Q)$ holds and

$$L(x) \equiv a^T \cdot x, \quad Q(X) \equiv a^T X a.$$  

For any choice of $r$, $P_r(a) = \sum_{i=1}^m r_i \cdot P_i(a) = \sum_{i=1}^m r_i \cdot 0 = 0$. But as we have shown,

$$P_r(a) = b + a^T \cdot v + a^T W a = b + L(v) + Q(W)$$

$$= u + L(v + x) + L(x) + Q(Y + W) + Q(Y) \quad \text{(by linearity of } L, Q) \text{.}$$

So the last quantity above, evaluated by the test for random $x, Y$, always equals $P_r(a) = 0$. This shows the completeness of $t_{good}$.

For the soundness claim, suppose the promise $S_{agree}(a, L, Q)$ is met, but $a$ does not satisfy every quadratic equation in $P$. Suppose without loss of generality that $P_m(a) \neq 0$.

Writing $P_r(a)$ as

$$P_r(a) = \sum_{i=1}^{m-1} r_i \cdot P_i(a) + P_m(a) \cdot r_m,$$
we find that, regardless of the initial random choices of \(r_1, \ldots, r_{m-1}\), the final term in the sum causes the sum to be nonzero with probability exactly \(\frac{1}{2}\). Thus with probability \(\frac{1}{2}\) over \(r\), \(P(r(a)) \neq 0\).

But just as before (using the promise \(S_{\text{agree}}(a, L, Q)\)),

\[
P(r(a)) = b + a^T \cdot v + a^T W a = u + L(v) + Q(W) = u + L(v + x) + L(x) + Q(Y + W) + Q(Y)
\]

for all choices of \(x, Y\). Thus \(t_{\text{good}}(a, L, Q)\) rejects with probability \(\frac{1}{2}\). So \(t_{\text{good}}\) has blockwise soundness \(\frac{1}{2}\) relative to the promise \(S_{\text{agree}}(a, L, Q)\), as needed. ⋄

As explained before the proof of Lemma 1, the tests \(t_{\text{lin}}, t_{\text{agree}}, t_{\text{good}}\) combine to yield a 4-query test \(t\) for \(S_{\text{good}}(a, L, Q)\) with perfect completeness and \(\Omega(1)\) blockwise soundness relative to no promise. This derived test fulfills the hypotheses of Proposition 1, so the procedure \(A_q\) returning it constitutes a 4-query quadratic assignment tester. We have already shown that a full-fledged 4-query assignment tester can be built, assuming access to a 4-query quadratic assignment tester; thus we have completed our proof of the Main Theorem. ⋄

Bibliography.


[Din] Irit Dinur, The PCP theorem by gap amplification (ECCC 2005, STOC 2006)

[Fis] Eldar Fischer, The art of uninformed decisions: A primer to property testing (BEATCS 75, 2001)

[Häs] Johan Håstad, Some optimal inapproximability results (STOC 1997)
