SIAM J. COMPUT.
Vol. 45, No. 3, pp. 1007–1063

THREE-PLAYER ENTANGLED XOR GAMES ARE NP-HARD TO APPROXIMATE*

THOMAS VIDICK†

Abstract. We show that for any $\varepsilon > 0$ the problem of finding a factor $(2-\varepsilon)$ approximation to the entangled value of a three-player XOR game is NP-hard. Equivalently, the problem of approximating the largest possible quantum violation of a tripartite Bell correlation inequality to within any multiplicative constant is NP-hard. These results are the first constant-factor hardness of approximation results for entangled games or quantum violations of Bell inequalities shown under the sole assumption that P≠NP. They can be thought of as an extension of Håstad's optimal hardness of approximation results for MAX-E3-LIN2 [*J. ACM*, 48 (2001), pp. 798–859] to the entangled-player setting. The key technical component of our work is a soundness analysis of a plane-vs-point low-degree test against entangled players. This extends and simplifies the analysis of the multilinearity test by Ito and Vidick [*Proceedings of the* 53*rd FOCS*, IEEE, Piscataway, NJ, 2012, pp. 243–252]. Our results demonstrate the possibility of efficient reductions between entangled-player games and our techniques may lead to further hardness of approximation results.

Key words. PCP theorem, XOR games, entangled games, Bell inequalities

AMS subject classifications. 81P68, 68Q10

DOI. 10.1137/140956622

1. Introduction.

In quantum mechanics, two or more spatially isolated systems are said to be *entangled* if no complete description of their joint state can be obtained solely from the combination of individual descriptions of each of the subsystems. This intuitive definition is due to Schrödinger,[1] who first coined the term "entangled" in reaction to Einstein, Podolsky, and Rosen's criticism of quantum mechanics as an incomplete theory [EPR35]. It is only through the work of Bell [Bel64], thirty years later, that a mathematically sound and (at least in principle) experimentally verifiable theory for the quantification of the nonlocal effects of entanglement first arose. Bell proposed the use of what are now known as "Bell inequalities." Suppose that each of r subsystems can be locally observed using any one among a set of possible measurements Q, each producing outcomes in a finite set A. For any choice of settings $(q_1, \ldots, q_r) \in Q^r$ the measurements' outcomes can be described by a joint distribution $p(a_1, \ldots, a_r|q_1, \ldots, q_r)$. A *Bell inequality* is a linear inequality in the $|A|^r|Q|^r$ variables $p(a_1, \ldots, a_r|q_1, \ldots, q_r)$ that is satisfied by any product distribution.[2] A quantum state is entangled if and only if there exists a choice of local measurements on its subsystems that give rise to a collection of joint distributions violating a Bell inequality [Gis91].

The study of Bell inequalities has taken an increasingly central role in quantum information theory, from the study of the foundations of quantum mechanics to ap-

*Received by the editors February 11, 2014; accepted for publication (in revised form) November 9, 2015; published electronically June 29, 2016.
http://www.siam.org/journals/sicomp/45-3/95662.html
†Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA (vidick@cms.caltech.edu).

[1] "When two systems, of which we know the states by their respective representatives, enter into temporary physical interaction due to known forces between them, and when after a time of mutual influence the systems separate again, then they can no longer be described in the same way as before, viz., by endowing each of them with a representative of its own. I would not call that *one* but rather *the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought. By the interaction the two representatives have become entangled." [Sch35]

[2] By linearity, the inequality will automatically be satisfied by any *convex combination* of product distributions as well. We refer to distributions in this set as *classical* correlations.

plications in cryptography. Somewhat ignored in the immediate aftermath of Bell's work, interest was revived after the discovery by Clauser et al. [CHSH69] of the first simple inequality that could realistically lead to an experiment (an experiment which was successfully performed by Aspect, Dalibard, and Roger [ADR82] some thirteen years later). Their inequality, the "CHSH inequality," applies to two systems on each of which two binary measurements can be made. It can be stated as follows:

$$\left|\frac{1}{4}\sum_{\substack{(q_1,q_2)\in\{0,1\}^2\\(a_1,a_2)\in\{0,1\}^2}}(-1)^{a_1\oplus a_2=q_1\wedge q_2}p(a_1,a_2|q_1,q_2)\right| \leq \frac{3}{4}. \tag{1.1}$$

Quantum mechanics identifies four measurements (two on the first subsystem and two on the second) which when applied to a system initialized in the joint state $|\Psi\rangle = (1/\sqrt{2})(|00\rangle + |11\rangle)$ are predicted to result in a distribution for which the left-hand side of (1.1) evaluates to $1/2 + \sqrt{2}/4 \approx 0.85$: quantum mechanics violates the inequality. Many Bell inequalities have since been introduced. More than 40 years of investigation, including the extensive use of numerical methods, have led to thousands of papers.[3] These investigations, however, have for the most part been confined to the study of small-scale examples, typically involving at most three subsystems and three or four measurement settings per system. This limitation reflects both the richness of entanglement and the difficulty of obtaining asymptotic results. It raises an obvious question, *What is the computational complexity of Bell inequalities?*

Surprisingly, it is only relatively recently that the question was first precisely formulated by Cleve et al. [CHTW04], who gave a reinterpretation of Bell inequalities in terms of *multiplayer games*. From their use in zero-knowledge proof systems [GMR85] to their role in the proof of the PCP theorem [AS98, ALM+98] multiplayer games have played a central role in computational complexity and cryptography throughout the past quarter century. A multiplayer game is run by the "referee," a trusted classical party, who interacts with $r \geq 2$ "players." The referee chooses questions $(q_1, \ldots, q_r) \in Q^r$ according to a distribution π, and sends question q_i to player i. The players each have to provide an answer a_i to the referee. The referee accepts or rejects the answers he receives according to a criterion $V(a_1, \ldots, a_r|q_1, \ldots, q_r) \in \{0, 1\}$. The rules of the game, including π and V, are public and known to the players, who cooperate in order to win the game. The only restriction on their strategies is that the players are not allowed to exchange any information once the game has started.

In parallel to their use in complexity, multiplayer games have turned out to provide a surprisingly rich framework in which to pursue the study of entanglement initiated by Bell. The "no communication" condition placed upon the players had traditionally been interpreted as the formal requirement that the distribution $p(a_1, \ldots, a_r|q_1, \ldots, q_r)$ on answers that they generate should be a (convex combination of) product distributions. As demonstrated by Bell, however, entanglement *does not* allow for supraluminal communication (quantum mechanics does not violate relativity), but it *does* allow for the generation of distributions that could not be realized locally by classical players, even with the help of shared randomness. The violation of Bell inequalities by quantum mechanics implies the following: there exists games for which entangled-player strategies are strictly more powerful than classical (shared randomness) strategies. Denoting by $\omega^*(G)$ the *entangled value* of a game G (the

[3] Google Scholar finds over 7000; more than 500 papers contain "Bell inequality"' in their title on the quant-ph arXiv alone.

maximum success probability of entangled-player strategies) and by $\omega(G)$ its *classical value* (the maximum success probability of classical players, restricted to using shared randomness as their sole source of correlation), we now know of games for which $\omega^*(G) = 1$ but $\omega(G)$ can be arbitrarily small [Ara02, Raz98].

The question formulated above can thus be restated as follows: *What is the complexity of computing $\omega^*(G)$?* An answer to this question for the case of the classical value $\omega(G)$ is precisely the content of the PCP theorem: $\omega(G)$ is NP-hard to approximate within a multiplicative constant, even for games with two players and binary answers—in fact, it is even hard for so-called *XOR games*, in which the referee's criterion V only depends on the parity of the two answers he receives [Hås01].[4] For the case of the entangled value, however, for a long time little was known. Indeed, nothing can be deduced directly from the classical case, as the sole fact that $\omega(G) \leq \omega^*(G) \leq 1$ does not obviously make the quantum problem any easier or harder.

Interestingly, a series of works have pointed to the quantum problem being *easier* than the classical one, at least for restricted classes of two-player games. Cleve et al., building on work of Tsirelson [Tsi80], gave a polynomial-time algorithm based on the use of semidefinite programming for the exact computation of $\omega^*(G)$ for the case of XOR games [CHTW04]. Kempe, Regev, and Toner [KRT10] also used semidefinite programming to show the existence of an algorithm giving a factor 6 approximation to $1-\omega^*(G)$ for the case of unique games. If one allows so-called *no-signaling* strategies, in which the distribution $p(a_i|q_i)$ is only limited by the condition that the marginal distribution on each subset of players' answers be independent from questions to the other players, then there is again a polynomial-time algorithm, this time based on a linear programming formulation of the problem [Pre07].

Could the computation, or at least approximation, of $\omega^*(G)$ be in BPP? In [KKM$^+$11, IKM09] it was shown that *exact* computation is NP-hard, even for two-player games with two-bit answers from each player. Recently the first strong hardness of approximation result was obtained: the problem of approximating $\omega^*(G)$ to within inverse polylogarithmic accuracy for games with four players is NP-hard under quasi-polynomial reductions [IV12]. This result was obtained as a corollary of the complexity class inclusion NEXP $\subseteq$ MIP*, an entangled-prover analogue of the celebrated NEXP $\subseteq$ MIP [BFL91]. (Here MIP* is the class of languages that have multiprover interactive proof systems with entangled provers.)

The initial discovery of the power of multiple provers, characterized by the equation MIP = NEXP, quickly led to the first hardness of approximation results for problems such as clique and independent set [FGL$^+$96]. Obtaining tight hardness results for constraint satisfaction problems such as 3-SAT [Hås01], however, required much further work and the development of techniques such as low-degree tests [AS98, RS96], composition of verifiers [AS98], and the use of gadgets [BGS98]. Our main contribution is the extension of some of the most important of these techniques to the setting of entangled-player games.[5] We prove the soundness of a variant of the low-degree test against entangled players, provide techniques enabling the composition of verifiers sound against entangled players, and analyze specific gadgets. Motivated by the

[4] Here the input G is always given by an explicit table of values for the distribution π and the predicate V.

[5] Classically multiplayer games and PCPs provide two views on the same object. In the presence of entanglement the equivalence is less clear, and we prefer to work with games. See however [TKRR13] for a possible definition of "no-signaling PCP" which naturally leads to a notion of "entangled PCP" equivalent to the games studied here, and the survey [AAV13] for an extended discussion of the "other quantum PCP" and its relationship to the one studied here.

goal of obtaining strong hardness of approximation results for the simplest possible classes of games, we show the following main result.

THEOREM 1.1. *Let $\varepsilon > 0$ be an arbitrary constant. The following is NP-hard: given a three-player XOR game G, distinguish between $\omega(G) \geq 1 - \varepsilon$ and $\omega^*(G) \leq 1/2 + \varepsilon$.*[6]

As mentioned above, the inclusion NEXP $\subseteq$ MIP* [IV12] can readily be scaled down to a result on the hardness of approximating $\omega^*(G)$. Theorem 1.1 improves on this in the following ways. First, in [IV12] hardness is only obtained for approximation factors $(1 + 1/\operatorname{poly}(\log n))$. Amplifying this to a constant requires sequentially repeating the game a polylogarithmic number of times and induces a superpolynomial blow-up in its size. Second, the scaling down from MIP* results in games which have questions and answers of length $\operatorname{poly}(\log n)$ bits and hence size, as measured by the total number of questions and answers, that is superpolynomial. The reduction behind the NP-hardness result established in Theorem 1.1 produces games with questions of length $O(\log n)$ bits and in which answers consist of a single bit each.

We believe that the soundness analysis of the low-degree test and related tests against entangled players that underlies the proof of Theorem 1.1 is of interest in itself, as these tests may prove useful in other contexts. One may ask, however, if the same conclusion, NP-hardness of three-player entangled games, could not be obtained through a less circuitous route. We view this as an important open question. Examples such as the Mermin-Peres magic square game [Ara02] demonstrate that a black-box use of the classical PCP theorem may be difficult; indeed based on this game it is possible to construct a simple 3-SAT formula, with 24 clauses over 9 variables, that is not satisfiable but such that the standard two-player 3-SAT game (send a clause to one player, and a variable from that clause to the other) has no perfect strategy for classical players but has entangled value 1. Thus some additional structure must be used; in our proof we find such a structure deep in the way approximation-resistant instances of 3-SAT are constructed through the low-degree test, but there may be others. In particular, recent work of Brandao and Harrow [BH13] attempts to leverage the property of entanglement monogamy to obtain a simpler reduction from entangled to classical strategies in the case where there is a sufficient number of players, and Conjecture 4 in [BH13] would imply a very similar (if slightly weaker in terms of parameters) result to our Theorem 1.1.

In terms of Bell inequalities, our main theorem gives the optimal hardness of approximation for inequalities involving three or more systems; indeed no simpler form for such inequalities can be thought of than *correlation inequalities*, which are the equivalent of XOR games. Since such inequalities measure the bias $\beta^*(G) = 2\omega^*(G) - 1$ of a given XOR game (indeed, for XOR games G clearly $\omega^*(G) \geq \omega(G) \geq 1/2$, hence $\beta^*(G)$ is always between 0 and 1), we can state the following immediate corollary of our main theorem.

COROLLARY 1.2. *Given an explicit description of a tripartite Bell correlation inequality, it is NP-hard to give any constant factor multiplicative approximation to the largest possible bias that is allowed by quantum mechanics.*

In addition to the above-mentioned results we also show that for any constant $\delta > 0$ it is NP-hard to distinguish between $\omega(G) = 1$ and $\omega^*(G) < \delta$ for games with three players and constant (depending on δ) answer size.

We note that all our results only apply to games with three or more players. For the case of XOR games the above-mentioned result of Cleve et al. [CHTW04] shows

[6] We note that as an immediate consequence the same hardness statement holds with the condition "$\omega(G) \geq 1 - \varepsilon$" replaced by the stronger condition "$\omega^*(G) \geq 1 - \varepsilon$."

that unless P = NP no hardness result can be expected when there are only two players. Showing hardness of the approximation of $\omega^*(G)$ for two-player non-XOR games (even games with answers of length $O(\log n)$ bits) remains a tantalizing open question (see "soundness of the low-degree test" below for additional discussion).

Techniques and proof overview. Our approach to proving hardness of approximation for entangled-player games is based on two main components. The first is a notion of *equivalence* (or closeness) of entangled-player strategies that is appropriate to *composition*. In analyzing the soundness of a certain game, or test, our goal is to make a statement of the form "any *generic* strategy with success $1-\varepsilon$ in the test must be ε'-*equivalent* to an *ideal* strategy," where the ideal strategy has precisely the type of structure that the test is trying to enforce (for instance, a strategy answering all questions according to a fixed low-degree polynomial). In the case of classical deterministic strategies it is natural to define strategies to be ε'-equivalent when they provide the same answer to all but a fraction at most ε' of questions. In the case of entangled—indeed, even randomized—strategies it is less obvious what the correct notion should be. In particular, it a priori seems impossible to consider single-player strategies by themselves, as, e.g., the marginal distribution on answers that they induce could very well be perfectly uniform, for every possible question. In addition, the notion of equivalence chosen should be appropriate for composition: if one test (for instance, the low-degree test) calls another test as a subprocedure (for instance, instead of checking directly a constraint $\varphi(x_1, \ldots, x_{10})$, the referee transforms φ into a 3-SAT formula over 10 variables and calls a subtest specially designed for the efficient verification of small 3-SAT formulas), then it should be possible to effortlessly combine a soundness analysis of each of the two tests in a soundness analysis of the global test. We give a notion of equivalence that satisfies these requirements, demonstrating "by the example" that it is well-suited to composition.

The second component is a soundness analysis of the plane-vs-point low-degree test from [RS97] with entangled players. Establishing soundness of this test is crucial to obtaining an NP-hardness result for games of polynomial size, rather than quasi-polynomial as in [IV12]. Our analysis follows the same outline as in [RS97],[7] but it requires substantial additional work. In particular, almost all known soundness analyses of the low-degree test rely on a key step of "consolidation" used to bootstrap a form of weak consistency (measured by the number of points of agreement) between a certain function and a more structured one (such as a low-degree polynomial) in order to automatically deduce that a much stronger consistency must hold. The execution of this step requires a deep overhaul, and its extension to entangled-player strategies is one of our main technical contributions.

We briefly expand on each of these two components below, pointing to the aspects of our proof that most differ from previous work done in the classical setting. We note that both components borrow heavily from techniques introduced in [IV12], and our contribution consists in an important extension and simplification of these techniques. We also note that our results make use of a recent parallel repetition theorem for entangled games [KV11], as well as (and independently from its use to obtain parallel repetition) of an "orthonormalization lemma" that played an important role in the proof of the parallel repetition theorem from [KV11].

[7] In contrast to [RS97], which was mostly concerned with obtaining a test with subconstant error, we only analyze the low-error regime. Nevertheless, the analysis given in [RS97] (as detailed and refined in [MR08]) is well-suited to an extension to the case of entangled players.

Equivalence of entangled-player strategies. Suppose we are given a certain game, or test, in which the players are required to answer questions $q \in Q$ with answers $a \in A$. For convenience we focus on a two-player game in which we can assume that both players use the same strategy, defined by a symmetric bipartite state $|\Psi\rangle$ and measurements $\{A_q^a\}_{a\in A}$ for every $q \in Q$. Let $\mathcal{F} \subseteq \{f : Q \to A\}$ be a set of functions having a certain desirable property (for instance, for $\mathbb{F}$ a finite field and m an integer we could have $Q = \mathbb{F}^m$, $A = \mathbb{F}$, and $\mathcal{F}$ the family of all m-variate low-degree polynomials over $\mathbb{F}$). Suppose the test is designed to verify that both players follow the following ideal strategy: there exists a fixed $f \in \mathcal{F}$ such that, upon receiving question q, either player answers it with $f(q)$. Note that if the players are allowed the use of entanglement (or even shared randomness) then it is not realistic to hope for the existence of a *single* f underlying their strategy. Indeed, the players could use shared randomness to select a random $f \in \mathcal{F}$ before computing their answer $f(q)$; no test will distinguish this from an ideal deterministic strategy. We are thus led to the following natural broadening of what is allowed in terms of ideal strategy: there should exist a *self-consistent* measurement $M = \{M^f\}_{f\in\mathcal{F}}$ such that the players are *equivalent* to players who first, measure their respective systems using M, obtaining an outcome f, and second, answer their question q with $f(q)$.

We define *consistent* and *equivalent*. The two notions are related. A pair of measurements $\{M^f\}$ and $\{N^f\}$ are said to be ε-consistent if the following holds:

$$\sum_f \langle\Psi|M^f \otimes N^f|\Psi\rangle \geq 1 - \varepsilon.$$

A measurement $\{M^f\}$ is *ε-self-consistent* if it is ε-consistent with itself. Sometimes we will drop the ε and use "consistent" and "self-consistent" to mean ε-consistent and ε-self-consistent for some small ε that should be clear from context. What consistency means is simply that, whenever two players measure their systems using $\{M^f\}$ and $\{N^f\}$, respectively, they get the same outcome except with probability ε. This is a natural requirement; indeed we are trying to mimic the deterministic case in which both players apply the same fixed function. To define a notion of equivalence we follow the approach from [IV12] and say that a generic strategy $(|\Psi\rangle, A)$ for the players is ε-equivalent to the ideal strategy $\{M^f\}$ if the following holds:

$$\mathrm{E}_{q\in Q} \sum_{a\in A} \sum_{f\in\mathcal{F}:\, f(q)\neq a} \langle\Psi|A_q^a \otimes M^f|\Psi\rangle \leq \varepsilon. \tag{1.2}$$

Note that (1.2) can be interpreted as requiring that the two strategies, generic and ideal, are ε-consistent. The following two arguments point to this notion of equivalence, defined through the property of being ε-consistent for small ε, being the "right" notion.

First, as should already be apparent from the definition, a relation such as (1.2) can be directly linked to quantities that arise naturally in the analysis of a game or test. For instance, success in the plane-vs-point low-degree test immediately implies consistency between the two families of measurements that define a generic entangled strategy for the players: a "points" measurement, designed to answer questions made of a single point, and a "planes" measurement, designed to answer questions about the restriction of the low-degree polynomial to a whole plane (we refer to section 3 for more details). This makes the notion of equivalence defined through (1.2) particularly well-suited to the analysis of multiplayer games.

$(d, m, r, \mathbb{F})$ low-degree test

1. Let $d, m, \mathbb{F}$ be parameters given as input.
2. Choose a random $\boldsymbol{x} \in \mathbb{F}^m$ and two random directions $\vec{y}_1, \vec{y}_2 \in \mathbb{F}^m$. Automatically accept if the two vectors are not linearly independent. Otherwise, let s be the plane $(\boldsymbol{x}; \vec{y}_1, \vec{y}_2)$.
3. Select two players among r at random. Send s to the first, and $\boldsymbol{x}$ to the second.
4. Receive a bivariate degree-d polynomial g defined on s from the first player, and a value $a \in \mathbb{F}$ from the second.
5. Accept if and only if $g(\boldsymbol{x}) = a$.

FIG. 1. *The plane-vs-point low-degree test attempts to verify that the r players answer consistently with a degree-d polynomial defined over* $\mathbb{F}^m$.

Second, equivalence obtained through consistency composes well. Suppose we are given a game obtained by combining two tests, each of which is executed with probability 1/2. In the game each player is asked a pair of questions (q_1, q_2). The first test is meant to verify that whenever a player is asked the pair of questions (q_1, q_2) he or she answers according to a function $f_{q_1} \in \mathcal{F} \subseteq \{f : Q_2 \to A\}$, where $\mathcal{F}$ is a subset of functions that has a certain structure. The second test is meant to check that the function f_{q_1} itself is obtained as $g(q_1)$ for some $g \in \mathcal{G} \subseteq \{g : Q_1 \to \mathcal{F}\}$, where again functions in $\mathcal{G}$ have a certain structure. The overall goal in analyzing such a game is verify that successful players must answer a query (q_1, q_2) with $(g(q_1))(q_2)$. Assuming we already have a proof of soundness for each of the two subtests, this conclusion will clearly hold as well if the players are deterministic. In the quantum, or even randomized, case however it is less immediate. We show that the desired conclusion does hold in the case when the players may apply entangled strategies, provided the soundness analysis of each subtest is based on the notion of equivalence defined by (1.2) (indeed with the right notion it is a simple calculation).

Soundness of the low-degree test with entangled players. Recall that in the plane-vs-point low-degree test the referee chooses a uniformly random affine plane p in $\mathbb{F}^m$, where $\mathbb{F}$ is a large finite field and m an integer, sends p to one player and a uniformly random point $\boldsymbol{x} \in p$ to the second, and expects as answers the description of a polynomial f of total degree at most d defined on p and a value $a \in \mathbb{F}$, respectively, such that $f(\boldsymbol{x}) = a$. (See Figure 1 for a more detailed description of the test.) The goal of the soundness analysis is to show that any generic strategy for the players succeeding with probability at least $1 - \varepsilon$ in the test, for some small fixed $\varepsilon > 0$, is $\mathrm{poly}(\varepsilon)$-equivalent to an ideal strategy in which the set of functions $\mathcal{F}$ is the set $\mathcal{F}_{m,d}$ of m-variate polynomials over $\mathbb{F}$ with total degree at most d. The proof is by induction. First we show that for most lines $\ell \subseteq \mathbb{F}^m$ the players' strategy, when restricted to questions $\boldsymbol{x} \in \ell$, must be $\mathrm{poly}(\varepsilon)$-equivalent to an ideal strategy using polynomials in $\mathcal{F}_{1,d}$. Then we proceed to prove a similar statement for planes, cubes, etc., until the final statement is obtained for $\mathbb{F}^m$. This outline is common to most analyses of the low-degree test.

Here we concentrate on a key difficulty that arises when analyzing entangled-player strategies. In all known proofs by induction of the low-degree test the closeness parameter ε blows up exponentially.[8] (The degree also increases, but we do not

[8] A notable exception is the proof technique from [AS97], which does not use induction but a more direct "bootstrapping" argument.

discuss this issue here.) In the classical, deterministic setting it is possible to argue directly using "robustness" properties of low-degree polynomials that δ-closeness for some sufficiently small δ implies ε'-closeness for some ε' depending only on ε (the failure probability in the test) but independent of δ (the error parameter reached after a number of induction steps). In the entangled-player setting such a statement does not hold. Intuitively, the reason for this is that while a given low-degree polynomial cannot be corrupted at a substantial fraction of points without drastically increasing its degree, for any δ it is possible to "corrupt" a measurement by any arbitrary amount δ, say by performing a small global rotation of the measurement operators. More precisely, (1.2) can fail for a number of reasons. While the measurement $\{M^f\}$ always outputs a low-degree polynomial, it does so probabilistically; hence the final probabilistic outcome $f(q)$ can fully agree with the first player's answer (when she measures using $\{A_q^a\}$) for most questions q, or partially agree for all q, or any combination in-between the two. This difficulty already arises if one attempts to perform the classical analysis directly on randomized strategies, without first "fixing the randomness"; doing so effectively is by itself a nontrivial task. In the case of quantum strategies, an additional difficulty comes from the fact that measurement operators themselves are not discrete, and can be affected by small rotation, or truncation, errors that classical discrete strategies are not subject to. Together with the noncommutativity of the quantum formalism these differences add up to generate technical difficulties that require substantially new arguments.

As a result, the measurements constructed throughout the induction must be modified at each step by performing an *active* correction procedure. Such a procedure was already the most technically challenging step in the proof of [IV12]. Here we build upon their work, but considerably improve and simplify their proof. The main idea is to define the "improved" measurement as the optimum of a particular semidefinite program (SDP)—roughly, one that seeks to minimize (1.2) over all possible measurements $\{M^f\}$. Our analysis makes an important use of duality properties of that SDP. This is one of the main points of departure from [IV12], which relied on a more generic convex optimization procedure for the analogue correction procedure, and did not make use of duality. The improvement in analysis that the new formulation buys us is substantial, and as a result we are able to argue that, provided a reasonably good measurement exists (the one constructed by induction), then there must also exist a much better measurement, in the sense of having much higher consistency properties. However, the resulting measurement may not be defined on the whole Hilbert space (it is not hard to see that this is unavoidable). To overcome this we need to add a layer of recursion by performing the whole analysis again on the parts of the Hilbert space in which the previous step had resulted in unrecoverable failure.

We note that of all our analysis it is only the consolidation procedure that requires the presence of three players (indeed, the low-degree test itself can be defined for two players only). If its correctness was extended to the case of two players one would automatically obtain a hardness result for two-player entangled games. We were unable to achieve this: the fact that the players' entangled state is a tripartite permutation-invariant state seems essential for our proof technique to go through.

Organization of the paper. We start with some useful preliminaries in section 2. In section 3 we introduce the main tests that we analyze: the low-degree test, its self-composition, a simple linearity test, and standard tests geared, respectively, at the verification of 3-SAT formulas and systems of quadratic equations. In section 4 we prove Theorem 1.1 and other hardness of approximation results for entangled-player games, assuming the soundness analysis of the low-degree tests. This analysis is

performed in section 6. The key technical ingredient in the analysis is the consolidation procedure described in section 5. In section 7 we analyze the remaining tests from section 3.

The reader interested in directly understanding the simplest hardness result proved in this paper, Theorem 4.3 on the hardness of general three-player entangled games, may take the following shortest path. The theorem follows almost directly from the analysis of the 3-SAT test stated in Theorem 3.3. The main ingredient in the proof is the soundness of the low-degree test, Theorem 3.1, which is given in section 6 and relies on the consolidation procedure described in section 5. We thus recommend that such reader first familiarize himself or herself with the description of the low-degree test in Figure 1 and the statement of Theorem 3.1, then proceed to its analysis in sections 6, relying on section 5 as a black box, and finally go through the standard reductions that deduce hardness from soundness of the low-degree test; these are standard in the PCP literature and no substantially new ingredient is needed here.

2. Preliminaries.

2.1. Notation. For an integer K, denote $\{1,\dots,K\}$ by $[K]$. Given a finite set X and an integer n, we sometimes use bold font to denote tuples $\boldsymbol{x} = (x_1,\dots,x_n) \in X^n$. We also write $\boldsymbol{x}_{\le i}$ for $(x_1,\dots,x_i) \in X^i$, as well as $\boldsymbol{x}_{<i}$, $\boldsymbol{x}_{\ge i}$, etc. for the obvious tuples. When T is a finite set, we write $\mathrm{E}_{x\in T}$ for the expectation over a uniformly random element x of T. log denotes the logarithm taken in base 2. If B is a boolean variable, 1_B is 1 if B evaluates to true and 0 otherwise. We also let -1_B be 1 if B evaluates to true and -1 otherwise.

It will often be convenient to express "approximate inequalities" as

$$(E) \approx_\delta (F),$$

where here (E), (F) are two expressions that evaluate to complex numbers and $\delta > 0$ a parameter. What this means is that there exists a universal constant C (the constant may be different every time the symbol $\approx$ is used) such that $|(E) - (F)| \le C\,\delta$.

Polynomials and finite fields. $\mathbb{F}$ will always denote a finite field. $\mathbb{F}_2$ is the finite field with two elements. For an integer m we let $\boldsymbol{z} = (z_1,\dots,z_m) \in \mathbb{F}^m$ denote a point, and $\vec{y} = (y_1,\dots,y_m) \in \mathbb{F}^m$ a vector (the distinction is only semantic). Given $\boldsymbol{z}$ and $\vec{y}_1,\dots$, we let $(\boldsymbol{z};\vec{y}_i)$ denote the affine subspace of $\mathbb{F}^m$ containing all points of the form $\boldsymbol{z} + \sum_i \alpha_i \vec{y}_i$ for $\alpha_i \in \mathbb{F}$. Given any such subspace we fix a canonical representation for it, and an associated coordinate system that makes it isomorphic to $\mathbb{F}^d$, where d is the dimension of the space spanned by the $\vec{y}_i$.

For an affine subspace s of $\mathbb{F}^m$ of dimension k and any $0 \le j \le k$ we let $\mathcal{S}_j(s)$ be the set of all j-dimensional affine subspaces of s. When s is clear from context (e.g., $s = \mathbb{F}^m$) we simply write $\mathcal{S}_j$ for $\mathcal{S}_j(s)$. For any affine space s, $\mathcal{P}_d(s)$ is the set of all degree-d polynomials defined on s (in particular, $\mathcal{P}_d(\mathbb{F}^m)$ is the set of all degree-d polynomials in m variables over $\mathbb{F}$). Any such polynomial can be represented by the list of its at most $(d+1)^m$ coefficients over $\mathbb{F}$. We recall the Schwartz–Zippel lemma [Zip79, Sch80], which we will use repeatedly.

LEMMA 2.1 (Schwartz–Zippel). *Let $d, m \ge 1$ be integers and p a nonzero polynomial in m variables of total degree at most d defined over the finite field $\mathbb{F}$. Then p has at most $d|\mathbb{F}|^{m-1}$ zeros.*

States and measurements. We use calligraphic letters, such as $\mathcal{H}$, to denote finite-dimensional Hilbert spaces. For $z \in \mathcal{H}$, $\|z\|$ denotes its Euclidean norm. A *state* is a vector with unit norm. Given an integer $r \ge 1$ and a state $|\Psi\rangle \in \mathcal{H}^{\otimes r}$, we say that $|\Psi\rangle$

is permutation invariant if $\sigma|\Psi\rangle = |\Psi\rangle$, where σ is the linear operator corresponding to any permutation of the r copies of $\mathcal{H}$ (sometimes also called "registers").

Given a permutation-invariant state $|\Psi\rangle$, we will often abuse notation and use the symbol ρ for the reduced density of $|\Psi\rangle$ on any one of the registers (permutation invariance implies that all single-system reduced densities are identical), but also on any two, three, etc., registers. In particular, we also write $\rho = |\Psi\rangle\langle\Psi|$. It will always be clear from context which number of registers is meant. Given a density σ, we write $\mathrm{Tr}_\sigma(A)$ as shorthand for $\mathrm{Tr}(A\sigma)$. Hence, for instance we have the following equivalent ways of writing the same expression:

$$\langle\Psi|A\otimes \mathrm{Id}\otimes\cdots\otimes \mathrm{Id}\,|\Psi\rangle \;=\; \mathrm{Tr}_\rho\big(A\otimes \mathrm{Id}\otimes\cdots\otimes \mathrm{Id}\,\big) \;=\; \mathrm{Tr}_\rho\big(\,\mathrm{Id}\otimes A\big) \;=\; \mathrm{Tr}_\rho(A).$$

Let $\mathrm{L}(\mathcal{H})$ be the set of linear operators on $\mathcal{H}$, and $\|\cdot\|$ the operator norm on $\mathrm{L}(\mathcal{H})$. $\mathrm{Id} = \mathrm{Id}_{\mathrm{L}(\mathcal{H})}$ is the identity operator on $\mathcal{H}$. A *submeasurement* on $\mathcal{H}$ is a finite set $A = \{A_i\}$ of nonnegative definite operators on $\mathcal{H}$ such that $\sum_i A_i \le \mathrm{Id}$. A *measurement* requires that $\sum_i A_i = \mathrm{Id}$.

Let $r \ge 2$ and $|\Psi\rangle$ be a permutation-invariant state on $\mathcal{H}^{\otimes r}$, i.e., $|\Psi\rangle$ is invariant under any permutation of its r subsystems. To $|\Psi\rangle$ we associate a bilinear form on $\mathrm{L}(\mathcal{H})\times\mathrm{L}(\mathcal{H})$ by defining

$$\langle A,B\rangle_\Psi \;:=\; \langle\Psi|A\otimes B\otimes \mathrm{Id}^{\otimes(r-2)}\,|\Psi\rangle \;=\; \mathrm{Tr}_\rho\big(A\otimes B\big) \;\in\mathbb{C} \tag{2.1}$$

for every $A,B\in\mathrm{L}(\mathcal{H})$. The permutation invariance of $|\Psi\rangle$ implies that this expression is independent of the exact registers on which the A and B operators are applied (provided they are distinct). Note that $\langle\cdot,\cdot\rangle_\Psi$ is not an inner product, as it is not positive. We also introduce a seminorm on $\mathrm{L}(\mathcal{H})$ by defining

$$\|A\|_\Psi \;:=\; \big(\langle\Psi|AA^\dagger\otimes \mathrm{Id}^{\otimes(r-1)}\,|\Psi\rangle\big)^{1/2} \;=\; \big(\mathrm{Tr}_\rho\big(AA^\dagger\big)\big)^{1/2}.$$

($\|\cdot\|_\Psi$ is clearly nonnegative, and the triangle inequality can be verified using the Cauchy–Schwarz inequality $\mathrm{Tr}_\rho(AB^\dagger)\le\|A\|_\Psi\,\|B\|_\Psi$.) We note that the order $AA^\dagger$ matters, and one can define an inequivalent norm by ${}_\Psi\|A\|^2 := \langle\Psi|A^\dagger A\otimes \mathrm{Id}^{\otimes(r-1)}\,|\Psi\rangle$. We then have the following inequalities, the first of which follows from Cauchy–Schwarz:

$$\big|\langle A,B\rangle_\Psi\big| \;\le\; \min\big\{\,\|A\|_\Psi\cdot{}_\Psi\|B\|\,,\;{}_\Psi\|A\|\cdot\|B\|_\Psi\,\big\} \;\le\; \min\big\{\,\|A\|_\Psi\cdot\|B\|,\;\|B\|_\Psi\cdot\|A\|\big\}. \tag{2.2}$$

The following inequality will also prove useful: for any $A,X\in\mathrm{L}(\mathcal{H})$,

$$\|AX\|_\Psi \;\le\; \|A\|_\Psi\cdot\|X\|, \tag{2.3}$$

where we used that $A(XX^\dagger)A^\dagger\le\|X\|^2AA^\dagger$ for any square matrices A,X. Finally, we record the following claim for future use.

Claim 2.2. *Let $|\Psi\rangle$ be a permutation-invariant state on $r\ge 3$ registers, and $\{A^a\}$, $\{B^a\}$ two single-register measurements with outcomes in the same set. Then*

$$\sum_a \big\|A^a-B^a\big\|_\Psi^2 \;\le\; O\big(\sqrt{\delta}\big),$$

where

$$\delta := 1-\sum_a \big\langle A^a,B^a\big\rangle_\Psi.$$

Proof. Expand

$$\begin{aligned}\sum_a \|A^a - B^a\|_\Psi^2 &= \sum_a \Big(\mathrm{Tr}_\rho\big((A^a)^2\big) + \mathrm{Tr}_\rho\big((B^a)^2\big) - 2\,\Re\big(\mathrm{Tr}_\rho\big(A^a B^a\big)\big)\Big)\\ &\le 2 - 2\sum_a \Re\big(\mathrm{Tr}_\rho\big(A^a B^a\big)\big),\end{aligned} \tag{2.4}$$

where we used $\sum_a (A^a)^2 \le \sum_a A^a \le \mathrm{Id}$ and similarly for (B^a). Applying the Cauchy–Schwarz inequality to the vectors $B^a \otimes \sqrt{B^b} \otimes \mathrm{Id}\,|\Psi\rangle$ and $A^a \otimes \sqrt{B^b} \otimes \mathrm{Id}\,|\Psi\rangle$ we can bound

$$\begin{aligned}\sum_{b\neq a} \mathrm{Tr}_\rho\big(A^a B^a \otimes B^b\big) &\le \left(\sum_{b\neq a} \mathrm{Tr}_\rho\big((A^a)^2 \otimes B^b\big)\right)^{1/2} \left(\sum_{b\neq a} \mathrm{Tr}_\rho\big((B^a)^2 \otimes B^b\big)\right)^{1/2},\\ &\le \delta^{1/2},\end{aligned} \tag{2.5}$$

where the second inequality uses $(A^a)^2 \le A^a$ and the definition of δ to bound the first term, and $(B^a)^2 \le B^a$ and $\sum_a B^a \le \mathrm{Id}$ to bound the second. Using $\sum_b B^b = \mathrm{Id}$ we may then write

$$\begin{aligned}\sum_a \mathrm{Tr}_\rho\big(A^a B^a\big) &= \sum_{a,b} \mathrm{Tr}_\rho\big(A^a B^a \otimes B^b\big)\\ &\approx_{\sqrt{\delta}} \sum_a \mathrm{Tr}_\rho\big(A^a B^a \otimes B^a\big)\\ &= \sum_{a,c} \mathrm{Tr}_\rho\big(A^a B^a \otimes B^a \otimes A^c\big)\\ &\approx_{\sqrt{\delta}} \sum_a \mathrm{Tr}_\rho\big(A^a B^a \otimes B^a \otimes A^a\big)\\ &\approx_{\sqrt{\delta}} \sum_{a,b,c} \mathrm{Tr}_\rho\big(A^b B^c \otimes B^a \otimes A^a\big)\\ &= \sum_a \mathrm{Tr}_\rho\big(\mathrm{Id} \otimes B^a \otimes A^a\big)\\ &= 1 - \delta,\end{aligned}$$

where the first approximate equality is (2.5), the second approximate inequality is (2.5) with the roles of A and B exchanged, and the third approximate inequality follows from a similar argument, applying (2.5) twice. This lets us upper bound the right-hand side of (2.4) and concludes the proof of the claim. ∎

2.2. Multiplayer games. We study one-round games played by $r \ge 2$ cooperative players against a referee.

DEFINITION 2.3. *A game $G = G(r, \pi, V)$ is given by finite sets Q of* questions *and A of* answers*, together with a distribution $\pi : Q^r \to [0,1]$, and a function $V : A^r \times Q^r \to \{0,1\}$.*[9] *The* size *of the game is defined as $|G| = |Q||A|$.*[10]

[9] We write $V(\cdot,\cdot)$ as $V(\cdot|\cdot)$ to clarify the role of the inputs.

[10] This measure does not explicitly take into account the description size of π, which we always assume to be at most polynomial in $|G|$. It also does not account for the number of players r, which for our purposes will always be a small constant.

The game G is played as follows: the referee samples $(q_1, \ldots, q_r)$ from Q^r according to π, and sends question q_i to player i. The players each reply with an answer $a_i \in A$. We say that the players win the game if $V(a_1, \ldots, a_r | q_1, \ldots, q_r) = 1$; otherwise they lose. The *value* of a game is the maximum winning probability of the players. The players can agree on a strategy before the game starts, but are not permitted to communicate after receiving their questions. We distinguish two different values, depending on the types of strategies allowed for the players: the *classical value* $\omega(G)$, corresponding to the maximum success probability of players using a classical deterministic strategy, and the *entangled value* $\omega^*(G)$, corresponding to the maximum success probability of quantum players allowed to use entanglement.

DEFINITION 2.4. *Let $G = G(r, \pi, V)$ be a multiplayer game. The classical value of G is defined as*

$$\omega(G) := \sup_{f_1,\ldots,f_r : Q \to A} \sum_{(q_1,\ldots,q_r) \in Q^r} \pi(q_1, \ldots, q_r) V\big(f_1(q_1), \ldots, f_r(q_r) | q_1, \ldots, q_r\big).^{11}$$

The entangled value of G is defined as

$$\omega^*(G) := \sup_{|\Psi\rangle, \{A^a_{i,q}\}} \sum_{(q_1,\ldots,q_r) \in Q^r} \pi(q_1, \ldots, q_r) \cdot \sum_{(a_1,\ldots,a_r) \in [A]^r} V(a_1, \ldots, a_r | q_1, \ldots, q_r) \langle \Psi | A^{a_1}_{1,q_1} \otimes \cdots \otimes A^{a_r}_{r,q_r} | \Psi \rangle,$$

where the supremum is taken over all finite-dimensional r-partite states $|\Psi\rangle$ and measurements positive operator-valued measurements (POVM) $\{A^a_{i,q}\}_{a \in A}$ for every $i \in [r]$ and $q \in Q$.

We will most often work with verifiers who treat all the players symmetrically. The next lemma shows that in that case we can always assume that the optimal players' strategy has the same symmetry.

LEMMA 2.5. *Let $G = G(r, \pi, V)$ be a game such that $\pi(q_1, \ldots, q_r)$ is symmetric in $q_1, \ldots, q_r$ and V is symmetric under simultaneous permutation of the questions $(q_1, \ldots, q_r)$ and of the answers $(a_1, \ldots, a_r)$. Then given any strategy $P_1, \ldots, P_r$ with entangled state $|\Psi\rangle$ that succeeds with probability p in G, there exists a strategy $P'_1, \ldots, P'_r$ with entangled state $|\Psi'\rangle$ and success probability p such that $P'_1 = \cdots = P'_r$ and $|\Psi'\rangle$ is invariant with respect to any permutation of its r registers.*

Proof. Let $\mathfrak{S}_r$ be the set of permutations of $\{1, \ldots, r\}$ and assume, by appropriately padding with extra qubits, that all registers of $|\Psi\rangle$ have the same dimension. Define strategies $P'_1, \ldots, P'_r$ as follows: the players share the entangled state $|\Psi'\rangle = \sum_{\sigma \in \mathfrak{S}_r} |\sigma(1)\rangle \ldots |\sigma(r)\rangle \otimes |\Psi^\sigma\rangle$, where the register containing $|\sigma(i)\rangle$ is given to player i and $|\Psi^\sigma\rangle$ is obtained from $|\Psi\rangle$ by swapping the r registers according to σ. For $1 \le i \le r$ player i measures the register containing $|\sigma(i)\rangle$ and applies $P_{\sigma(i)}$. By symmetry of π and V this new strategy achieves the same winning probability p, and $|\Psi'\rangle$ has the required symmetry properties. ∎

3. Protocols. In this section we introduce different games (or "tests") played between the referee and r players. All tests treat the r players symmetrically, and as a consequence of Lemma 2.5 we may assume players use a symmetric strategy; in particular their respective state can be represented using the same Hilbert space $\mathcal{H}$ for each player. In addition, the tests are often (but not always) made of a combination

[11] The supremum can be extended to range over all convex combinations of tuples $(f_1, \ldots, f_r) : Q^r \to A^r$ without changing its value.

of "subtests" in which the marginal distribution on questions to any single player is the same, irrespective of the subtest. It is important to note that, whenever this is the case, the player cannot tell which subtest is being performed and is thus required to apply a measurement that only depends on the question he is asked, but not on the subtest he is being tested on.

In section 3.1 we first introduce a variant of the low-degree test, a test that plays a key role in the construction of efficient PCPs. In sections 3.2 and 3.4 we give standard tests, respectively, for the verification of the satisfiability of a 3-SAT formula and a system of quadratic equations in boolean variables. The latter uses a linearity test for functions $\mathbb{F}_2^n \to \mathbb{F}_2$ given in section 3.3. We note that none of the tests we define are new and all have appeared previously in the PCP literature.

3.1. The low-degree test.

3.1.1. A first protocol. The line-vs-point low-degree test was introduced in [RS96]. Here we analyze a variant from [RS97]. The test is called the "plane-vs-point" low-degree test because it calls for two players to send back the restriction of a low-degree polynomial to a plane and a point chosen randomly in that plane, respectively. The test is described in Figure 1. We summarize its main properties.

Complexity. The longest question is the description of the affine plane s, which requires $3m \log |\mathbb{F}|$ bits. The longest answer is the degree-d bivariate polynomial g, which can be specified using at most $(d+1)^2 \log |\mathbb{F}|$ bits.

Strategies. A strategy for the players in the $(d, m, r, \mathbb{F})$ low-degree test is a triple $(|\Psi\rangle, A, C)$, where

- $|\Psi\rangle$ is a permutation-invariant state on $\mathcal{H}^{\otimes r}$,
- $A = \{A^a_{\boldsymbol{x}}\}$ is a set of "points" measurements $\{A^a_{\boldsymbol{x}}\}_{a \in \mathbb{F}}$ defined for every $\boldsymbol{x} \in \mathbb{F}^m$,
- $C = \{C^g_s\}$ is a set of "planes" measurements $\{C^g_s\}_{g \in \mathcal{P}_d(s)}$ defined for every $s \in \mathcal{S}_2(\mathbb{F}^m)$.

Analysis. We state the soundness of the test as a theorem. The proof is given in section 6. Note that although the test is defined for any $r \geq 2$, the theorem requires $r \geq 3$.

THEOREM 3.1. *Let $0 < \varepsilon \leq 1/2$, $d \geq 1, m \geq 2, r \geq 3$ integers, and $\mathbb{F}$ a finite field of size $|\mathbb{F}| = q$ such that $q \geq (dm/\varepsilon)^{d_1}$, where $d_1 \geq 1$ is a universal constant. Let $(|\Psi\rangle, A, C)$ be a strategy with success $1 - \varepsilon$ in the $(d, m, r, \mathbb{F})$ low-degree test. Then there exists a measurement $\{M^g\}$ with outcomes $g \in \mathcal{P}_d(\mathbb{F}^m)$ such that*

$$\mathrm{E}_{\boldsymbol{x} \in \mathbb{F}^m} \sum_{g \in \mathcal{P}_d(\mathbb{F}^m)} \sum_{a \in \mathbb{F}:\, g(\boldsymbol{x}) \neq a} \langle A^a_{\boldsymbol{x}}, M^g \rangle_\Psi \;\leq\; C_1 \varepsilon^{c_1}, \tag{3.1}$$

where $0 < c_1 \leq 1, C_1 > 0$ are universal constants.

Equation (3.1) serves as a measure of distance between the provers' original strategy, defined by the measurements $A_{\boldsymbol{x}}$, and the new strategy defined by the single measurement M. The equation states that the two measurements are consistent in the sense that, if two players are simultaneously sent the same question $\boldsymbol{x}$, and the first determines his answer by applying the measurement $\{A^a_{\boldsymbol{x}}\}$ while the second first measures using $\{M^g\}$ and then returns $g(\boldsymbol{x})$, then the players will provide identical answers except with probability at most $C_1 \varepsilon^{c_1}$. Hence provers succeeding in the low-degree test are in a sense "equivalent" to provers applying the measurement M to determine a low-degree polynomial g even *before* having looked at their question. This is precisely the sense in which we mean that the low-degree test is

"sound against entangled players". We also note that using Claim 2.2, applied with $(A^a)_a$ in the claim defined as $(|\mathbb{F}^m|^{-1}A^a_{\boldsymbol{x}})_{\boldsymbol{x},a}$ here and $(B^a)_a$ in the claim defined as $(|\mathbb{F}^m|^{-1}\sum_{g\in\mathcal{P}_d(\mathbb{F}^m):\, g(\boldsymbol{x})=a} M^g)_{\boldsymbol{x},a}$ here, (3.1) implies the distance bound

$$\mathrm{E}_{\boldsymbol{x}\in\mathbb{F}^m}\sum_{a\in\mathbb{F}}\left\|A^a_{\boldsymbol{x}}-\sum_{g\in\mathcal{P}_d(\mathbb{F}^m):\, g(\boldsymbol{x})=a} M^g\right\|^2_{\Psi} = O\big(\varepsilon^{c_1/2}\big).$$

3.1.2. A test with reduced answer size. When the low-degree test is used the degree d will typically be polylogarithmic in the input size, so that the answer size (the total number of answers) of the test described in section 3.1.1 is superpolynomial. In this section we show how the previous test can be composed with itself to obtain a test with polynomial answer size. The idea of composition was instrumental in the proof of the PCP theorem [AS98].

Let m, d, q be integers and $\mathbb{F}$ a field of size $|\mathbb{F}| = q$. We first describe how variable substitution (see, e.g., [DFK+11, section 4.4]) can be used to map a degree-d polynomial g over $\mathbb{F}^2$ to a degree-d' polynomial g' over $\mathbb{F}^{m'}$, where $m' = d' := 2\lceil\log(d+1)\rceil$. For $i = 0, \ldots, \lceil\log(d+1)\rceil - 1$ introduce new variables $\tilde{x}_i := x^{2^i}, \tilde{y}_i := y^{2^i}$. Using the base-2 decomposition of k and ℓ, any monomial $x^k y^\ell$ can be written as a product of the $\tilde{x}_i$ and $\tilde{y}_j$, each appearing at most once. Let $g' \in \mathbb{F}[\tilde{x}_i, \tilde{y}_i]$ be such that $g' \to g$ (formally) when $\tilde{x}_i \to x^{b^i}, \tilde{y}_i \to y^{b^i}$. Let

$$\# : \begin{cases} \mathbb{F}^2 & \to \quad \mathbb{F}^{m'}, \\ (x,y) & \mapsto \quad (x, x^2, \ldots, x^d, y, y^2 \ldots, y^d), \end{cases}$$

and note that for any $\boldsymbol{x} \in \mathbb{F}^2$, $g(\boldsymbol{x}) = g'(\#\boldsymbol{x})$.

For any number $r \geq 2$ of players, the $(d, m, r, \mathbb{F})$ two-level low-degree test is described in Figure 2. We summarize its main properties.

Complexity. The longest question is the pair (s, s'), which is $3m\log|\mathbb{F}| + 3m'\log|\mathbb{F}| \leq 6m\log|\mathbb{F}|$ bits. The longest answer is the restriction of the polynomial g' to the plane s', which can be specified using at most $(d')^2\log|\mathbb{F}| = O((\log d)^2\log|\mathbb{F}|)$ bits.

Strategies. The players have the following measurements. For every $\boldsymbol{x} \in \mathbb{F}^m$, a "points" measurement $\{A^a_{\boldsymbol{x}}\}_{a\in\mathbb{F}}$. For every plane $s \in \mathcal{S}_2(\mathbb{F}^m)$ and every $\boldsymbol{x}' \in s$ (where $\boldsymbol{x}'$ is represented as $\#\boldsymbol{x}$ for some $\boldsymbol{x} \in s$), another points measurement $\{B^a_{s,\boldsymbol{x}}\}_{a\in\mathbb{F}}$. For every plane $s \in \mathcal{S}_2(\mathbb{F}^m)$ and every plane $s' \in \mathcal{S}_2(\mathbb{F}^{m'})$, a "planes" measurement $\{C^g_{s,s'}\}$. where g is a degree-d' bivariate polynomial defined on s'.

Analysis. We state the soundness of the test as a theorem. The proof is given in section 6.6.

THEOREM 3.2. *Let $0 < \varepsilon \leq 1/2$, $d \geq 1, m \geq 2, r \geq 3$ integers, and $\mathbb{F}$ a finite field of size $|\mathbb{F}| = q$ such that $q \geq (dm/\varepsilon)^{d_2}$, where $d_2 \geq 1$ is a universal constant. Let $(|\Psi\rangle, A, B, C)$ be an r-player strategy with success $1-\varepsilon$ in the $(d, m, r, \mathbb{F})$ two-level low-degree test. Then there exists a measurement $\{M^g\}$ with outcomes $g \in \mathcal{P}_{dd'}(\mathbb{F}^m)$ such that*

$$\mathrm{E}_{\boldsymbol{x}\in\mathbb{F}^m}\sum_{g\in\mathcal{P}_{dd'}(\mathbb{F}^m)}\sum_{a\in\mathbb{F}:\, g(\boldsymbol{x})\neq a}\langle A^a_{\boldsymbol{x}}, M^g\rangle_\Psi \leq C_2\varepsilon^{c_2},$$

where $c_2 \leq 1, C_2 > 0$ are universal constants.

3.2. The 3-SAT test. Let φ be a 3-SAT formula with n variables and poly(n) clauses. Let $h = \lceil\log n\rceil$ and $m = \lceil\log n/\log\log n\rceil$, so that $(h+1)^m \geq n$. Let $\mathbb{F}$ be a field of size $|\mathbb{F}| = q \geq h+1$, and identify $[n]$ with the subset $\{0, \ldots, h\}^m \subseteq \mathbb{F}^m$.

$(d, m, r, \mathbb{F})$ two-level low-degree test

1. Let $d, m, \mathbb{F}$ be parameters given as input. Set $d' = m' := 2\lceil \log(d+1) \rceil$.
2. The referee chooses a random $\boldsymbol{x} \in \mathbb{F}^m$ and two random directions $\vec{y}_1, \vec{y}_2 \in \mathbb{F}^m$. He automatically accepts if $\vec{y}_1, \vec{y}_2$ are not linearly independent. Let $s := (\boldsymbol{x}; \vec{y}_1, \vec{y}_2)$ be the corresponding affine plane.
3. The referee chooses a random $\boldsymbol{x}' \in \mathbb{F}^{m'}$ and two random directions $\vec{y}'_1, \vec{y}'_2 \in \mathbb{F}^{m'}$. He automatically accepts if $\vec{y}'_1, \vec{y}'_2$ are not linearly independent. Let $s' := (\boldsymbol{x}'; \vec{y}'_1, \vec{y}'_2)$ be the corresponding affine subplane of s.
4. The referee selects two players at random, and performs one of the following two tests, with probability 1/2 each.
   - 4.1 The referee sends $\boldsymbol{x}$ to the first player and $(s, \#\boldsymbol{x})$ to the second. He receives answers $a \in \mathbb{F}$ and $a' \in \mathbb{F}$, respectively, and rejects if $a \neq a'$.
   - 4.2 The referee sends the pair (s, s') to the first player and $(s, \boldsymbol{x}')$ to the second. The first player answers with a degree-d' bivariate polynomial g' over s' and the second with a value $a' \in \mathbb{F}$. The referee rejects if $g'(\boldsymbol{x}') \neq a'$.
5. If the referee has not rejected then he accepts.

FIG. 2. *The $(d, m, r, \mathbb{F})$ two-level low-degree test attempts to verify that the r players answer consistently with a degree-d polynomial defined over $\mathbb{F}^m$. Note that queries to the second player in steps* 4.1 *and* 4.2 *are identically distributed, so that the players cannot distinguish which test is being performed.*

We use boldface $\boldsymbol{x}$ for the element of $\mathbb{F}$ corresponding to the variable x of φ. Let $d := mh$. In the test, the players are supposed to hold a degree-d polynomial g over $\mathbb{F}^m$ obtained as the low-degree extension of a satisfying assignment to the variables of φ: g is the unique m-variate polynomial of degree at most h in each variable such that $g(\boldsymbol{x}) = x$ for every $x \in \{0, \ldots, h\}^m$ associated with a variable $\boldsymbol{x}$ of φ (see, e.g., [BFLS91, Proposition 4.1] for a proof of existence and uniqueness).

A degree-4 curve c in $\mathbb{F}^m$ is specified by m univariate polynomials of degree at most 4 over $\mathbb{F}$, $(c_1, \ldots, c_m)$. The restriction of g to c is a univariate polynomial $g_{|c}(t) = g(c_1(t), \ldots, c_m(t))$ of degree at most $4d$. Using variable substitution as in section 3.1.2, $g_{|c}$ can also be thought of as a polynomial of degree d' in $\mathbb{F}^{m'}$, where $d' = m' = \lceil \log(4d+1) \rceil$. Let $\# : \mathbb{F} \to \mathbb{F}^{m'}$ be the map which performs the variable substitution. The $(\varphi, n, r, \mathbb{F})$ 3-SAT test is described in Figure 3. We note that the use of curves to aggregate the values of a polynomial at different points is standard in the PCP literature; see, e.g., [MR10].

Complexity. Questions can be specified using $O(m \log |\mathbb{F}|)$ bits: questions in the two-level low-degree test require $O(m \log |\mathbb{F}|)$ bits, and in step 2.2 (Figure 3) the longest question is the curve c which takes at most $m \cdot 4 \log |\mathbb{F}|$ bits to specify. The two-level low-degree test has answers of length $O((\log d)^3 \log |\mathbb{F}|)$ bits. The polynomial $g \in \mathcal{P}_{4d'}(c')$ in step 2.2.2 requires $4d' \log |\mathbb{F}|$ bits to specify. Overall, the answers can be specified using $O((\log \log n)^3 \log |\mathbb{F}|)$ bits.

Strategies. The players have a state $|\Psi\rangle$, measurements (A, B, C) corresponding to a strategy in the $(d, m, r, \mathbb{F})$ two-level low-degree test, for every degree-4 curve c in $\mathbb{F}^m$ and $\boldsymbol{w} \in c$ (specified as a point in $\mathbb{F}^{m'}$) a measurement $\{D^a_{c,\boldsymbol{w}}\}_{a \in \mathbb{F}}$, and finally for every degree-4 curve c' in $\mathbb{F}^{m'}$, a "curve" measurement $\{F^g_{c,c'}\}$, where $g \in \mathcal{P}_{4d'}(c')$.

Analysis. It is clear that if φ is satisfiable then the players have a perfect strategy that does not use any entanglement. They can simply define a polynomial g as the

$(\varphi, n, r, \mathbb{F})$ 3-SAT test

1. Let $h = \lceil \log n \rceil$, $m = \lceil \log n / \log\log n \rceil$, $d = mh$, and $d' = m' = \lceil \log(4d+1) \rceil$.
2. Do each of the following with probability 1/2 each:
   - 2.1 Perform the $(d, m, r, \mathbb{F})$ two-level low-degree test.
   - 2.2 Pick a clause $C \in \varphi$ at random. Let $x, y, z \in [n]$ be the three variables in φ, and $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ the associated points in $\mathbb{F}^m$. Let $\boldsymbol{w}$ be a random point in $\mathbb{F}^m$ and c the degree-4 curve through $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{w})$. Do each of the following with probability 1/2 each:
     - 2.2.1 Select two players at random. Send $\boldsymbol{w}$ to the first, receiving $a \in \mathbb{F}$ as answer, and $(c, \#\boldsymbol{w})$ to the second, receiving $a' \in \mathbb{F}$ as answer. Reject if $a \neq a'$.
     - 2.2.2 Pick a random $\boldsymbol{w}' \in \mathbb{F}^{m'}$, and select two players at random. Send $(c, \boldsymbol{w}')$ to the first, receiving $a \in \mathbb{F}$, and (c, c') to the second, where $c' \subseteq \mathbb{F}^{m'}$ is the degree-4 curve going through $(\#\boldsymbol{x}, \#\boldsymbol{y}, \#\boldsymbol{z}, \boldsymbol{w}')$, receiving $g \in \mathcal{P}_{4d'}(c')$ as answer. Reject if $(g(\#\boldsymbol{x}), g(\#\boldsymbol{y}), g(\#\boldsymbol{z}))$ is not a satisfying assignment to the variables in clause C, or if $g(\boldsymbol{w}') \neq a$.

FIG. 3. *The $(\varphi, n, r, \mathbb{F})$ 3-SAT test attempts to verify that the r players answer consistently with a degree-d polynomial over $\mathbb{F}^m$ that is the low-degree extension of a satisfying assignment for φ (encoded in the values of g on $\{0, 1, \ldots, h\}^m$).*

(n, r) linearity test

1. The referee chooses $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}_2^n$ uniformly at random. He selects three players at random and sends them $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y}$, respectively.
2. The players answer with $a, b, c \in \mathbb{F}_2$, respectively. The referee accepts if and only if $c = a + b$.

FIG. 4. *The linearity test attempts to verify that the r players answer consistently with a linear function $f : \mathbb{F}_2^n \to \mathbb{F}_2$.*

degree-d extension of a satisfying assignment to φ, and answer the two-level low-degree test according to g. If a player is asked a query of the form $(c, \#w)$ he answers with $g(w)$. If he is asked for (c, c') he answers with the restriction of g to the curve c', seen as a univariate polynomial of degree at most $4d'$ defined on $c' \subset c \approx \mathbb{F}^{m'}$. We state the soundness of the test as the following theorem. The theorem is proved in section 7.1.

THEOREM 3.3. *Let $0 < \varepsilon \leq K_3$, where $K_3 > 0$ is a universal constant, φ a 3-SAT formula on $n \geq 3$ variables, $r \geq 3$, and $\mathbb{F}$ a field of size $|\mathbb{F}| = q$ such that $q \geq (\log n/\varepsilon)^{d_3}$, where d_3 is a universal constant. Let $(|\Psi\rangle, A, B, C, D, F)$ be an r-player strategy with success $1 - \varepsilon$ in the $(\varphi, n, r, \mathbb{F})$-SAT test. Then there is an assignment to the variables in φ that satisfies all but a fraction at most $C_3\varepsilon^{c_3}$ of the clauses, where $C_3 > 0, 0 < c_3 \leq 1$ are universal constants.*

3.3. The linearity test. Let n be an integer and $\mathbb{F}_2$ the field with two elements. The (n, r) linearity test uses $r \geq 3$ players and is described in Figure 4.

Complexity. Questions have length n bits and answers are a single bit.

Strategies. A strategy for the players in the (n, r) linearity test is given by a state $|\Psi\rangle$ and a family of measurements $\{A_{\boldsymbol{x}}^a\}$ with outcomes $a \in \mathbb{F}_2$.

Analysis. The linearity test was first introduced in [BLR93] in the classical setting. The analysis with entangled players is joint work of the author and Tsuyoshi Ito [Vid11].

THEOREM 3.4. *Let n be an integer, $r \geq 3$, $\varepsilon > 0$, and $(|\Psi\rangle, A)$ a strategy for the players that succeeds with probability $1-\varepsilon$ in the (n,r) linearity test. There exists a measurement $\{M^u\}$ with outcomes $u \in \mathbb{F}_2^n$ such that*

$$\mathrm{E}_{\boldsymbol{x}\in\mathbb{F}_2^n} \sum_{\substack{u\in\mathbb{F}_2^n, a\in\mathbb{F}_2 \\ a\neq u\cdot\boldsymbol{x}}} \langle M^u, A^a_{\boldsymbol{x}}\rangle_\Psi \;=\; O\big(\sqrt{\varepsilon}\big).$$

3.4. The QUADEQ test. Let QUADEQ be the language consisting of all systems of quadratic equations over $\mathbb{F}_2$ that are satisfiable. An instance of QUADEQ over n variables x_i is thus a set of $K = \mathrm{poly}(n)$ quadratic equations of the form

$$\sum_{i,j\in[n]} a_{ij}^{(k)} x_i x_j \;=\; c^{(k)} \qquad (\mathrm{mod}\, 2),$$

for $k = 1, \ldots, K$, that are simultaneously satisfiable. QUADEQ is well known to be NP-complete. Here we recall a standard test for verifying membership in QUADEQ (see, e.g., [AB09, Theorem 11.19]). Let φ be an instance of QUADEQ on n variables and $r \geq 3$. Looking ahead, we assume that the variables of φ are partitioned into three chunks of $n' = n/3$ variables each, labeled ℓ_1, ℓ_2, and ℓ_3 (the labels will be used to identify the chunks among a larger universe of variables). The (φ, n, r) QUADEQ test is described in Figure 5.

Complexity. The maximal question length is n^2 bits plus the length of the labels. Answers are constituted of a single bit each.

Strategies. The players have a state $|\Psi\rangle$, for each label ℓ_i measurements $A_i \equiv A_{\ell_i}$ corresponding to a strategy in the $(n/3, r)$ linearity test, for each pair of labels (ℓ_1, ℓ_2) measurements B_{ℓ_1,ℓ_2} corresponding to strategies in the $(2n/3, r)$ linearity test, and for each triple of labels (ℓ_1, ℓ_2, ℓ_3) measurements $C \equiv C_{\ell_1,\ell_2,\ell_3}$ and $D \equiv S_{\ell_1,\ell_2,\ell_3}$ corresponding to strategies in the (n, r) and (n^2, r) linearity tests, respectively.

Analysis. Suppose φ is satisfiable, and let $x = (x_1, x_2, x_3)$, where $x_i \in \mathbb{F}_2^{n/3}$ contains the assignment to variables from chunk ℓ_i, be a satisfying assignment. Then the players have a perfect strategy that does not use any entanglement. For this, they answer a query of the form (ℓ_i, u) with $u \cdot x_i$; a query of the form (ℓ_1, ℓ_2, v), where $v \in \mathbb{F}_2^{2n/3}$, with $v \cdot (x_1, x_2)$; a query of the form $(\ell_1, \ell_2, \ell_3, w)$, where $w \in \mathbb{F}_2^n$, with $w \cdot (x_1, x_2, x_3)$; a query of the form $(\ell_1, \ell_2, \ell_3, z)$, where $z \in \mathbb{F}_2^{n^2}$, with $z \cdot ((x_1, x_2, x_3) \otimes (x_1, x_2, x_3))$. We state the soundness of the test as the following lemma. The lemma is proved in section 7.2.

LEMMA 3.5. *Let $0 < \varepsilon \leq K_4$, where $K_4 > 0$ is a universal constant, φ a QUADEQ instance on $n \geq 2$ variables, and $r \geq 3$. Let $(|\Psi\rangle, A, B, C, D)$ be an r-player strategy with success $1-\varepsilon$ in the (φ, n, r) QUADEQ test. Then φ is satisfiable. Moreover, suppose a given collection of QUADEQ instances $\varphi_1, \ldots, \varphi_T$, each acting on a triple of chunks of variables chosen from a common universe $\{x_1, \ldots, x_S\}$, where $x_i \in \mathbb{F}_2^{n/3}$. Let $(|\Psi\rangle, (A_i), (B_{i,j}), (C_{i,j,k}), (D_{i,j,k}))$ be such that for every $t \in \{1, \ldots, T\}$ the strategy $(|\Psi\rangle, A_i, B_{i,j}, C_{i,j,k}, D_{i,j,k})$, where φ_t is over chunks x_i, x_j, and x_k, has success $1-\varepsilon$ in the (φ_t, n, r) QUADEQ test. Then there exist measurements $\{M_i^{x_i}\}_{x_i\in\mathbb{F}_2^{n/3}}$ for every $i \in \{1, \ldots, S\}$, such that for every φ_t on (x_i, x_j, x_k) it holds that*

$$\sum_{(x_i,x_j,x_k)\vdash\varphi_t} \langle M_i^{x_i}, M_j^{x_j}\rangle_\Psi \;\geq\; 1 - C_4\varepsilon^{c_4},$$

(φ, n, r) QUADEQ test

1. The referee performs each of the following with probability 1/5 each:
   - 1.1 With probability 1/4 each, do the following:
     - 1.1.1 Choose one of the three labels at random, send it to three players chosen at random and perform the $(n/3, r)$ linearity test.
     - 1.1.2 Send labels (ℓ_1, ℓ_2) to three players chosen at random and perform the $(2n/3, r)$ linearity test.
     - 1.1.3 Send labels (ℓ_1, ℓ_2, ℓ_3) to three players chosen at random and perform the (n, r) linearity test.
     - 1.1.4 Select three players at random and perform the (n^2, r) linearity test.
   - 1.2 Select random $u, v \in \mathbb{F}_2^{n/3}$ and send $(\ell_1, u), (\ell_2, v), (\ell_1, \ell_2, (u, v))$ to three players chosen at random. Receive a, b, c, respectively, and reject if $a + b \neq c$.
   - 1.3 Select random $u \in \mathbb{F}_2^{2n/3}$, $v \in \mathbb{F}_2^{n/3}$, and send (ℓ_1, ℓ_2, u), (ℓ_3, v), $(\ell_1, \ell_2, \ell_3, (u, v))$ to three players chosen at random. Receive a, b, c, respectively, and reject if $a + b \neq c$.
   - 1.4 Select two random vectors $u, v \in \mathbb{F}_2^n$. Send $(\ell_1, \ell_2, \ell_3, u)$, $(\ell_1, \ell_2, \ell_3, v)$, $(\ell_1, \ell_2, \ell_3, u \otimes v)$ to three players chosen at random. Verify that their answers (a, b, c) satisfy $a \cdot b = c$.
   - 1.5 Select a random vector $w \in \mathbb{F}_2^K$ and let $w = \sum_k w_k a^{(k)} \in \mathbb{F}_2^{n^2}$. Send $(\ell_1, \ell_2, \ell_3, w)$ to a randomly chosen player, and check that the answer $a = \sum_k w_k c^{(k)}$.

FIG. 5. *The QUADEQ test attempts to verify that the r players answer consistently with functions $f_{\ell_1}, f_{\ell_2}, f_{\ell_3} : \mathbb{F}_2^{n/3} \to \mathbb{F}_2$ and $f_{\ell_1,\ell_2} : \mathbb{F}_2^{2n/3} \to \mathbb{F}_2$, $f : \mathbb{F}_2^n \to \mathbb{F}_2$, $g : \mathbb{F}_2^{n^2} \to \mathbb{F}_2$ such that $f_{\ell_1,\ell_2}(x_1, x_2) = f_{\ell_1}(x_1) + f_{\ell_2}(x_2)$, $f(x_1, x_2, x_3) = f_{\ell_1,\ell_2}(x_1, x_2) + f_{\ell_3}(x_3)$, and $g = f \otimes f$.*

where $(x_i, x_j, x_k) \vdash \varphi_t$ means that the assignment (x_i, x_j, x_k) satisfies φ_t and $0 < c_4 \leq 1$, $C_4 > 0$ are universal constants.

Note that the "moreover" part of the lemma does not claim that the system $\{\varphi_t\}_{t=1,\ldots,T}$ is simultaneously satisfiable.

4. Hardness results. In this section we prove our main theorem, Theorem 1.1, restated as Corollary 4.10 at the end of the section. In section 4.1 we state and prove a first hardness result based on the 3-SAT test from section 3.2, and whose analysis relies on the (composed) low-degree test from section 3.1.2. In section 4.2 we use the QUADEQ test from section 3.4 to obtain a hardness result for games with constant answer size. In section 4.3 we show that the parallel repetition theorem from [DSV14a] can be adapted to amplify the resulting hardness of the approximation factor. Finally, our main theorem is proven in section 4.4.

Many of the results in this section use *projection games* and associated *two-out-of-three-player games* defined as follows.

DEFINITION 4.1. *A two-player projection game G is specified by question sets U and V, answer sets A and B, a distribution π on $U \times V$, and for every $(u, v) \in U \times V$ a function $\tau_{uv} : B \to A$. The referee accepts answers (a, b) to questions (u, v) if and only if $\tau_{uv}(b) = a$.*

To distinguish the asymmetric role played by both players in a projection game we will call the player with question u and answer a the "left" player, and the player with question v and answer b the "right" player.

Let H be the $|V|\times|V|$ matrix whose (v,v')th entry equals $\pi(v,v') := \sum_v \pi(u)\pi(v|u)\pi(v'|u)$, where $\pi(u) = \sum_v \pi(u,v)$ denotes the marginal distribution on the left player's question and $\pi(v|u) = \pi(u,v)/\pi(u)$ is the conditional distribution on the right player's question. Let D be the diagonal matrix with the degrees $\pi(v)$ on the diagonal, and $L := \mathrm{Id} - D^{-1/2}HD^{-1/2}$ the normalized Laplacian. We say that a family of projection games (G_i) is expanding if the second smallest eigenvalue of L is at least a positive constant independent of the size of G_i.

With any projection game G we associate the following three-player games.

DEFINITION 4.2. *Let G be a projection game with underlying distribution π on $U\times V$ and projection constraints $\tau_{uv} : B \to A$, where U, A and V, B are, respectively, the left and right players' question and answer sets.*

- *The two-out-of-three-player game G' associated with G is the following three-player game: the referee randomly chooses two of the three players and plays the game G with them, ignoring the third player (each player is assigned the role of the left or right player in G at random, and told which is the case).*
- *The cube game G'' associated with G is the following three-player game: the referee chooses a question u for the left player as in G, and three independent questions v, v', v'' for the right player in G (each chosen according to $\pi(\cdot|u)$). He sends v (resp., v', v'') to the first (resp., second, third) player in G''. The players provide answers b, b', and b'', respectively. The referee accepts if and only if $\pi_{uv}(b) = \pi_{uv'}(b') = \pi_{uv''}(b'')$.*

4.1. The basic hardness result. Our first hardness result is the following.

THEOREM 4.3. *There is an $\varepsilon > 0$ such that the following holds. Given a three-player game G in explicit form, it is NP-hard to distinguish between $\omega(G) = 1$ and $\omega^*(G) \le 1-\varepsilon$. Furthermore, the problem is still NP-hard when restricting to games G of size n that are obtained as the two-out-of-three player game associated with a projection game (see Definition 4.2) for which questions and answers can be specified using $O(\log n)$ bits and $\mathrm{poly}(\log\log n)$ bits, respectively.*

Proof. The proof of the theorem follows from the analysis of the 3-SAT test given in Theorem 3.3. First recall that the PCP theorem shows that there exists an $\varepsilon_1 > 0$ such that it is NP-hard to distinguish between a 3-SAT formula being satisfiable or the formula having at most a fraction $1-\varepsilon_1$ of its clauses simultaneously satisfied (see, e.g., [ALM+98, BGLR93, Hås01]). Let n be an integer, $\varepsilon_2 = \min(K_3, (\varepsilon_1/C_3)^{1/c_3})$, and $\mathbb{F}$ a finite field of size $q \in [(\log n/\varepsilon_1)^{d_3}, 2(\log n/\varepsilon_1)^{d_3}]$. Given a 3-SAT formula φ, let $G = G(\varphi)$ be the three-player game corresponding to the $(\varphi, n, 3, \mathbb{F})$ 3-SAT test. With our choice of q questions in G are $O(\log n)$-bit long and answers are $O((\log\log n)^4)$ bits long; in particular the size of G is polynomial in n. Furthermore it is clear that an explicit description of G can be computed in polynomial time from φ.

If φ is satisfiable then $\omega(G) = 1$. Furthermore, Theorem 3.3 implies that if $\omega^*(G) > 1-\varepsilon_2$ then (by definition of ε_2) there is an assignment satisfying more than a fraction $1-\varepsilon_1$ of the clauses of φ. Hence deciding between $\omega(G) = 1$ and $\omega^*(G) \le 1-\varepsilon_2$ is at least as hard as deciding between φ being satisfiable and φ having at most a fraction $1-\varepsilon_1$ of its clauses satisfiable. ∎

4.2. Hardness for games with constant answer size. In this section we combine Theorem 4.3 with the QUADEQ test from section 3.4 to obtain a hardness result for games with binary answers. The result, stated as Corollary 4.5 below, will follow from the following general reduction.

PROPOSITION 4.4. *There is a polynomial-time reduction mapping any two-player game G with n questions per player and in which answers from the players can be specified using m bits to a three-player game G' in which questions to the players have length $O(\log n + m^2)$ bits, answers from the players are restricted to a single bit each, and is such that the following holds. Let G'' be the two-out-of-three player game associated with G. Then $\omega(G) = 1 \implies \omega(G') = 1$ and $\omega^*(G') \leq 1 - \Omega(1 - \omega^*(G''))^c$ for some universal constant c.*

Proof. This transformation is standard in the PCP literature; see, e.g., [AB09, Corollary 22.13]. We proceed with the details.

Let Q be the set of all questions that can be asked in the game G, and π the distribution on $Q \times Q$ with which pairs of questions are chosen. Let m be the maximal length of an answer in G, and write $A = \{0,1\}^m$ for the set of all possible answers. For every $(q_1, q_2) \in Q \times Q$ the referee in G expects a pair of answers $(a_1, a_2) \in A \times A$. He then verifies a certain condition $V(a_1, a_2|q_1, q_2) \in \{0,1\}$. Using NP-completeness of QUADEQ, this condition can be expressed as an instance ψ_{q_1,q_2} of QUADEQ over $2m + m'$ variables. Here the first $2m$ variables correspond to the bits of a_1 and a_2. The additional m' variables are auxiliary variables used in the reduction transforming $V(\cdot,\cdot|q_1, q_2)$ in an instance of QUADEQ. Without loss of generality we can assume $m' = m$. The $3m$ variables can then be split into three chunks of variables, such that the first chunk is associated with a_1, the second with a_2, and the third with the auxiliary variables. Each QUADEQ instance ψ_{q_1,q_2} obtained from $(q_1, q_2) \in Q \times Q$ acts on three chunks of variables taken from a universe of chunks of m binary variables, each labeled using a unique label $\ell(q)$ associated with a single question $q \in Q$ for the "answer" chunks, and a label $\ell(q, q')$ associated with a pair of questions $q, q' \in Q$ for the "auxiliary" chunks. From a classical deterministic strategy in G one can construct an assignment to the variables in all chunks satisfying a fraction of instances ψ_{q_1,q_2} equal to the success probability of the strategy in G.

Consider the following game G':

1. The verifier samples questions (q_1, q_2) as in G.
2. The verifier runs the $(\psi_{q_1,q_2}, m, 3)$ QUADEQ test, where the labels are $\ell_1 = \ell(q_1)$, $\ell_2 = \ell(q_2)$, and $\ell_3 = \ell(q_1, q_2)$.
3. The verifier accepts if and only if the QUADEQ test accepts.

First we note that the length of questions in G' is at most twice that of G (for the labels) plus the square of the answer lengths in G, so it is $O(\log n + m^2)$ bits. The answer length is a single bit.

The discussion above shows that if $\omega(G) = 1$ then $\omega(G') = 1$ as well; in fact it more generally holds that $\omega(G') \geq \omega(G)$. Conversely, suppose that $\omega^*(G') \geq 1 - \varepsilon$, where $\varepsilon > 0$ is to be specified later. Using Markov's inequality, for a fraction at least $1 - \sqrt{\varepsilon}$ of pairs (q_1, q_2) (chosen according to π) the players have success at least $1 - \sqrt{\varepsilon}$ in the $(\psi_{q_1,q_2}, m, 3)$ QUADEQ test. Provided ε is small enough, the "furthermore" part of Lemma 3.5 shows the existence of a family of measurements $\{M_q^a\}_{a \in A}$ such that for each of the "good" pairs (q_1, q_2) it holds that

$$\sum_{(a_1,a_2):\, V(a_1,a_2|q_1,q_2)=1} \langle M_{q_1}^{a_1}, M_{q_2}^{a_2} \rangle_\Psi \;=\; 1 - O\big(\varepsilon_2^{c_4/2}\big).$$

It is then immediate that the strategy $(|\Psi\rangle, \{M_q^a\})$ is a strategy for the players in game G'' with success probability at least $1 - O(\varepsilon^{c_4/2})$. ∎

COROLLARY 4.5. *There is an $\varepsilon > 0$ such that the following holds. Given a three-player game G in explicit form in which answers from the players are restricted to a single bit each, it is NP-hard to distinguish between $\omega(G) = 1$ and $\omega^*(G) \leq 1 - \varepsilon$.*

Proof. Let φ be a 3-SAT formula on n variables, and G the projection game whose existence follows from Theorem 4.3. Letting G'' be the associated two-out-of-three player game, if φ is satisfiable then $\omega(G'') = 1$ whereas if φ is not satisfiable then $\omega^*(G'') \leq 1 - \varepsilon''$ for some $\varepsilon'' > 0$; moreover answers in G'' have length $m = \text{poly}(\log \log n)$. Applying the reduction from Proposition 4.4 results in a three-player game G' with $\text{poly}(n)$ questions and binary answers such that if φ is satisfiable then $\omega(G') = 1$ whereas if φ is not satisfiable then $\omega^*(G') \leq 1 - \Omega((\varepsilon'')^c)$, proving the corollary for a small enough choice of ε. ∎

4.3. Parallel repetition of two-out-of-three player projection games on expanders. The constant ε for which we established NP-hardness in Corollary 4.5 can be very small. In this section we show the following.

COROLLARY 4.6. *Let $\delta > 0$ be an arbitrary constant. Then the following is NP-hard. Given a three-player game G in explicit form, distinguish between $\omega(G) = 1$ and $\omega^*(G) \leq \delta$. Furthermore, the problem is still NP-hard when restricting to games G of size n such that the following hold:*

- *Questions in G have length $O(\log n)$ bits and answers have length $\text{poly}(\delta^{-1})$ bits.*
- *The referee treats all players symmetrically.*

The proof of Corollary 4.6 is based on amplifying the completeness-soundness gap from Corollary 4.5. The standard method for doing so consists in performing parallel repetition: one attempts to argue that, by repeating K instances of a game G in parallel and accepting if and only if the players provide answers valid for each of the K instances, the verifier ensures that the players' success probability decreases roughly as $\omega(G)^K$. Unfortunately there is no known general parallel repetition theorem for two-player games with entangled players (see [CS14, JPY14] for recent partial progress on the question)—let alone for three-player games. Nevertheless in our setting it will be sufficient to amplify the soundness of games that are the cube of an expanding projection game, as defined in Definition 4.2. For this type of game we are able to argue that the results of [DSV14a], which establish parallel repetition for two-player entangled projection games, can be extended in a straightforward manner. This is shown in Lemma 4.7 below. Note that the guarantee provided by the lemma is weaker than standard parallel repetition, as in principle the value of the cube game could be less than 1 while that of the two-out-of-three player game could still be 1 (it is not hard to see that the latter is always at least the former); however it will be sufficient for our purposes.

LEMMA 4.7. *Let G be the cube game associated with an expanding projection game G', and G'' the two-out-of-three-player game associated with G'. The entangled value of G decreases under parallel repetition as follows:*

$$\omega^*\big(G^{\otimes K}\big) \leq \Big(1 - \Omega\big((1 - \omega^*(G''))^2\big)\Big)^K.$$

Combining the lemma with Corollary 4.5 immediately implies Corollary 4.6. The proof of the lemma follows very closely the technique introduced in [DSV14a], and we sketch it below. We note that, although the results of [DSV14a] apply to both expanding and nonexpanding projection games, we do not know whether Lemma 4.7 can be extended to the nonexpanding case. The reason is that the proof given in [DSV14a] for

the nonexpanding case involves the use of a "correlated sampling procedure" through which the players, depending on their respective questions, transform an initial universal bipartite "embezzlement state" into a state that is useful for them to succeed in the game. We do not know how to extend this procedure to the case where two out of the three players are trying to generate the correct shared state from an initial universal tripartite state (even though such states do exist). In contrast, for the expanding case it is possible to show that the same state can be used by the three players, irrespective of their questions.

Proof sketch of Lemma 4.7. We show how the results in [DSV14a] can be adapted to this setting, and for the purposes of this sketch only we assume familiarity with the notation used in [DSV14b, section 4.1]. Proceeding as in [DSV14b], assume that there exists a permutation-invariant strategy $\{A^{b_1\cdots b_K}_{v_1\cdots v_K}, |\psi\rangle\}$ for the three players in $G^{\otimes K}$, where for every K-tuple $(v_1,\ldots,v_K) \in V^K$ the operators $\{A^{b_1\cdots b_K}_{v_1\cdots v_K}\}_{b_1,\ldots,b_K}$ form a POVM and $|\psi\rangle$ is a tripartite permutation-invariant state, such that

$$\mathrm{E}_\omega \mathrm{E}_{v\sim v'\sim v''} \sum_{b\leftrightarrow b'\leftrightarrow b''} \langle\psi| A^b_{\omega v}\otimes A^{b'}_{\omega v'}\otimes A^{b''}_{\omega v''}|\psi\rangle \geq (1-\eta)\mathrm{E}_\omega \mathrm{E}_v \langle\psi| A_{\omega v}\otimes A_{\omega v}\otimes A_{\omega v}|\psi\rangle, \tag{4.1}$$

where the expectation $\mathrm{E}_{v\sim v'\sim v''}$ is over all triples (v,v',v'') obtained by first sampling a question for the left player in G' and then independently sampling three questions v,v',v'' for the right player, and the summation $\sum_{b\leftrightarrow b'\leftrightarrow b''}$ is over all triples of answers (b,b',b'') such that $\pi_{uv}(b) = \pi_{uv'}(b') = \pi_{uv''}(b'')$. (For context, recall that the variable ω in (4.1) serves as a placeholder for all questions and answers in $(K-1)$ repetitions of G, and (v,v',v'') and (b,b',b'') denote the players' respective questions and answers in the Kth repetition.) Equation (4.1) is the analogue of (22) in [DSV14b]; informally it expresses the condition that the value of the game does not decrease as much as would be expected from taking $(K-1)$ to K repetitions (the precise derivation of (4.1) follows exactly that of (22) in [DSV14b]).

To prove the lemma it will suffice to show that from (4.1) we may deduce the existence of a strategy with success $1-O(\sqrt{\eta})$ for the players in the game G''. Due to the presence of the third player, (4.1) takes a slightly different form than (22). To derive an expression that is similar to (22) we show that the third player can be assumed to take a passive role, as follows. Write $A_{\omega v} = \sum_b A^b_{\omega v}$, $A_{\omega u} = \mathrm{E}_{v\sim u} A_{\omega v}$, and observe the identity

$$\begin{aligned}
&\mathrm{E}_\omega \mathrm{E}_{u\sim v}(A_{\omega v}-A_{\omega u})\otimes(A_{\omega v}-A_{\omega u})\otimes A_{\omega v} + A_{\omega u}\otimes A_{\omega v}\otimes A_{\omega v} + A_{\omega v}\otimes A_{\omega u}\otimes A_{\omega v} \\
&\qquad\qquad - 2A_{\omega u}\otimes A_{\omega u}\otimes A_{\omega u} \\
&= \mathrm{E}_\omega\Big(\mathrm{E}_v A_{\omega v}\otimes A_{\omega v}\otimes A_{\omega v} - \mathrm{E}_u A_{\omega u}\otimes A_{\omega u}\otimes A_{\omega u}\Big) \\
&\leq \eta \mathrm{E}_\omega \mathrm{E}_v A_{\omega v}\otimes A_{\omega v}\otimes A_{\omega v}
\end{aligned}$$

by (4.1), from which we deduce using the convexity inequalities

$$\begin{aligned}
\langle\psi|(A_{\omega v}-A_{\omega u})\otimes(A_{\omega v}-A_{\omega u})\otimes A_{\omega v}|\psi\rangle \geq 0, \\
\langle\psi|\big(A_{\omega v}\otimes A_{\omega v} - A_{\omega u}\otimes A_{\omega u}\big)\otimes A_{\omega u}|\psi\rangle \geq 0
\end{aligned}$$

for permutation-invariant $|\psi\rangle$, that both the following bounds must hold:

$$\begin{aligned}
&\mathrm{E}_\omega \mathrm{E}_{u\sim v}\langle\psi|\big(A_{\omega v}\otimes A_{\omega v} - A_{\omega u}\otimes A_{\omega u}\big)\otimes A_{\omega u}|\psi\rangle \\
&\qquad \leq \eta\, \mathrm{E}_\omega \mathrm{E}_v \langle\psi| A_{\omega v}\otimes A_{\omega v}\otimes A_{\omega v}|\psi\rangle, \qquad (4.2) \\
&\mathrm{E}_\omega \mathrm{E}_{u\sim v}\langle\psi|(A_{\omega v}-A_{\omega u})\otimes(A_{\omega v}-A_{\omega u})\otimes A_{\omega v}|\psi\rangle \\
&\qquad \leq \eta\, \mathrm{E}_\omega \mathrm{E}_v \langle\psi| A_{\omega v}\otimes A_{\omega v}\otimes A_{\omega v}|\psi\rangle. \qquad (4.3)
\end{aligned}$$

Using the Cauchy–Schwarz inequality and permutation invariance of $|\psi\rangle$ it follows that, for any $w \in V$,

$$\begin{aligned}
&\mathrm{E}_{u\sim v}\big|\langle\psi|\big(A_{\omega v}\otimes A_{\omega v} - A_{\omega u}\otimes A_{\omega u}\big)\otimes A_{\omega w}|\psi\rangle\big| \\
&\quad = \mathrm{E}_{u\sim v}\big|\langle\psi|\big(A_{\omega v} - A_{\omega u}\big)\otimes A_{\omega u}\otimes A_{\omega w} + A_{\omega u}\otimes\big(A_{\omega v} - A_{\omega u}\big)\otimes A_{\omega w}|\psi\rangle\big| \\
&\quad \le \Big(\mathrm{E}_{u\sim v}\big|\langle\psi|(A_{\omega v} - A_{\omega u})\otimes(A_{\omega v} - A_{\omega u})\otimes A_{\omega v}|\psi\rangle\big|\Big)^{1/2} \\
&\qquad \cdot \Big(\Big(\mathrm{E}_v\langle\psi|A_{\omega w}\otimes A_{\omega w}\otimes A_{\omega v}|\psi\rangle\Big)^{1/2} + \Big(\mathrm{E}_u\langle\psi|A_{\omega w}\otimes A_{\omega w}\otimes A_{\omega u}|\psi\rangle\Big)^{1/2}\Big) \\
&\quad \le \sqrt{2\eta}\Big(\mathrm{E}_v\langle\psi|A_{\omega v}\otimes A_{\omega v}\otimes A_{\omega v}|\psi\rangle\Big)^{1/2} \\
&\qquad \cdot \Big(\mathrm{E}_v\langle\psi|A_{\omega v}\otimes A_{\omega v}\otimes A_{\omega v}|\psi\rangle + \langle\psi|A_{\omega w}\otimes A_{\omega w}\otimes A_{\omega w}|\psi\rangle\Big)^{1/2},
\end{aligned} \tag{4.4}$$

where the last inequality uses (4.3) to bound the first term and the Cauchy–Schwarz inequality to bound the second. Applying (4.4) twice, with $w = v$ and $w = v' \sim v$, and using (4.2) we get

$$\begin{aligned}
&\mathrm{E}_{v\sim v'}\big|\langle\psi|\big(A_{\omega v}\otimes A_{\omega v} - A_{\omega v'}\otimes A_{\omega v'}\big)\otimes A_{\omega v}|\psi\rangle\big| \\
&\qquad\qquad = O\big(\sqrt{\eta}\big)\Big(\mathrm{E}_v\langle\psi|A_{\omega v}\otimes A_{\omega v}\otimes A_{\omega v}|\psi\rangle\Big).
\end{aligned}$$

Iterating this process and using graph expansion as in the proof of [DSV14b, Claim 14] we arrive at

$$\mathrm{E}_{\omega,v,v'}\big|\langle\psi|\big(A_{\omega v}\otimes A_{\omega v} - A_{\omega v'}\otimes A_{\omega v'}\big)\otimes A_{\omega}|\psi\rangle\big| = O\big(\sqrt{\eta}\big)\Big(\mathrm{E}_{\omega,v}\langle\psi|A_{\omega v}\otimes A_{\omega v}\otimes A_{\omega v}|\psi\rangle\Big), \tag{4.5}$$

where here the questions v, v' are sampled independently, we wrote $A_\omega = \mathrm{E}_v A_{\omega v}$, and the constant implicit in the $O(\sqrt{\eta})$ depends on the graph expansion parameter. Using (4.5) in conjunction with (4.1) leads to the inequality

$$\mathrm{E}_\omega\mathrm{E}_u\sum_a \langle\psi|A^a_{\omega u}\otimes A^a_{\omega u}\otimes A_\omega|\psi\rangle \ge \big(1 - O(\sqrt{\eta})\big)\mathrm{E}_\omega\mathrm{E}_v\langle\psi|A_{\omega v}\otimes A_{\omega v}\otimes A_\omega|\psi\rangle. \tag{4.6}$$

Equation (4.6) has precisely the same form as (22) in [DSV14b], where we can think of the bipartite state $|\hat{\Psi}\rangle$ there as $(\mathrm{Id}\otimes\mathrm{Id}\otimes A_\omega^{1/2})|\psi\rangle$ here. Note the latter state is invariant under permutation of the first two registers. Applying [DSV14b, Claim 13] to the present scenario we deduce the existence of a two-player strategy for the game G'', and the state used by this strategy, as defined in (23) in [DSV14b], can be taken as (up to normalization)

$$\sigma = \mathrm{E}_w\Big(U_{\omega w}A^{1/2}_{\omega w}\otimes U_{\omega w}A^{1/2}_{\omega w}\otimes U_{\omega w}A^{1/2}_{\omega w}|\psi\rangle\langle\psi|\big(U_{\omega w}A^{1/2}_{\omega w}\otimes U_{\omega w}A^{1/2}_{\omega w}\otimes U_{\omega w}A^{1/2}_{\omega w}\big)^\dagger\Big),$$

which is invariant under permutation of the three registers (we refer to (18) in [DSV14b] for a definition of the unitaries $U_{\omega w}$). As a consequence an analogous strategy to the one defined in (19) and (20) in [DSV14b] can be used by any two of the three players and leads to a success probability in G'' at least $1-O(\sqrt{\eta})$. Thus $\eta = \Omega((1-\omega^*(G''))^2)$, and recalling that (4.1) follows from the assumption $\omega^*(G^{\otimes K}) \ge (1-\eta)^K$ the lemma is proved. ∎

4.4. Hardness for three-player XOR games. By performing a reduction from (the parallel repetition of) two-player projection games, Håstad [Hås01, Theorem 5.5] showed that for any $\varepsilon > 0$ it is NP-hard to approximate the classical value of a three-player XOR game within a multiplicative factor $2 - \varepsilon$. We prove the following analogue of Håstad's reduction.

THEOREM 4.8. *Let G be the cube game associated with an expanding projection game G' with binary left answer set $A = \{\pm 1\}$, and G'' the two-out-of-three-player game associated with G'. For any $\varepsilon, \delta, \gamma > 0$ there exists a polynomial-time (in $|G|$, ε, and δ, but exponential in γ) transformation from G' to a three-player XOR game $\tilde{G}$ such that, if $\omega(G'') = 1$ then $\omega(\tilde{G}) \geq 1 - \varepsilon$ and if $\omega^*(G'') \leq 1 - \gamma$ then $\omega^*(\tilde{G}) \leq (1 + \delta)/2$.*

Proof. The proof follows the analysis of the test $L_2^{\varepsilon}(u)$ in [Hås01, Lemma 5.2], except we need to perform some slight modifications to account for the fact that we are performing a reduction from the (parallel repetition of) the cube of a projection game, instead of just a two-player projection game as in the case of Håstad's analysis.

We first introduce some notation. Recall that in a projection game the projection constraints maps any answer from the right player to a unique corresponding valid answer from the left player. Let $U = \{u_1, \ldots, u_n\}$ be the set of all possible questions to the left player in G'. Using the projection constraint, any question v for the right player may be formally identified with the set of questions $\{u_i\}$ to the left player such that $\pi(u_i, v) > 0$, where π is the distribution on questions in G', and any possible answer b with the set of $\tau_{u_i v}(b)$, where τ denotes the projection constraint. We will use this formal identification repeatedly in the proof.

For any subset $S \subseteq U$, let $\mathcal{F}_S = \{f : \{\pm 1\}^S \to \{\pm 1\}\}$. For every pair of functions $f, -f \in \mathcal{F}_S$ we select a unique representative and let $R_S \subset \mathcal{F}_S$ be the resulting set.[12] Let K be a parameter to be chosen later. Consider the following three-player XOR game $\tilde{G}$.

1. The referee independently samples K questions (u_k) for the left player as in game G. For $1 \leq k \leq K$ let v_k be a question for the right player chosen according to the conditional distribution $\pi(\cdot|u_k)$. Let $S = \{u_1, \ldots, u_K\}$ and $T = \{v_1, \ldots, v_K\}$. As described above, we identify S with a subset of T by (formally) replacing each $v_k \in T$ by the set of all u such that $\pi(u|v_k) > 0$.
2. The referee chooses a function $\mu \in \mathcal{F}_T$ by setting $\mu(t) = 1$ with probability $1-\varepsilon$ and $\mu(t) = -1$ with probability ε, independently for every $t \in \{\pm 1\}^T$. He chooses $f \in \mathcal{F}_S$, $g_1 \in \mathcal{F}_T$, and $d \in \{\pm 1\}$ uniformly, and sets $g_2 = f g_1 \mu d \in \mathcal{F}_T$ by defining $g_2(t) = f(t_{|S}) g_1(t) \mu(t) d$ for every $t \in \{\pm 1\}^T$.
3. The referee selects a random permutation of the three players and sends $(-1)_{f \in R_S} f$ to the first, $(-1)_{g_1 \in R_T} g_1$ to the second, and $(-1)_{g_2 \in R_T} g_2$ to the third.
4. He receives answers $a, b, c \in \{\pm 1\}$ and accepts if and only if

$$abc = d(-1)_{f \in R_S}(-1)_{g_1 \in R_T}(-1)_{g_2 \in R_T}.$$

First we verify that if $\omega(G'') = 1$ then $\omega(\tilde{G}) \geq 1 - \varepsilon$. Indeed, let $(x, y) \in \{\pm 1\}^U \times \{\pm 1\}^V$ be a perfect strategy for the players in G''. Then in $\tilde{G}$ the players can answer their queries f, g_1, g_2 by $f(x_{|S})$, $g_1(x_{|T})$, $g_2(x_{|T})$, respectively. The players will be accepted if and only if $f g_1 g_2(x) = \mu(x) d = d$, which happens with probability exactly $1 - \varepsilon$ by definition of μ.

To establish soundness of the game $\tilde{G}$ we prove the following.

[12] The role of R_S is to enable an operation known as "folding over true."

CLAIM 4.9. *Suppose that* $\omega^*(\tilde{G}) \geq (1+\delta)/2$. *Then* $\omega^*(G^{\otimes K}) \geq 8\varepsilon^3\delta^4$, *where* $G^{\otimes K}$ *is obtained by repeating the cube game* G *associated with* G' K *times in parallel.*

Proof. We follow the proof of [Hås01, Lemma 5.2]. Using the symmetry in the definition of the game $\tilde{G}$, we can represent any strategy for the three players using a permutation-invariant state $|\Psi\rangle$ with associated density matrix ρ and observables $A_{S,f}$ and $B_{T,g}$ indexed by subsets S, T of questions and functions $f : \{\pm 1\}^S \to \{\pm 1\}$, $g : \{\pm 1\}^T \to \{\pm 1\}$. Suppose we are given such a strategy with success at least $(1+\delta)/2$. Conditioned on the referee choosing sets S, T in the game, the players' success probability is $(1+\delta_{S,T})/2$ with

$$\delta_{S,T} := \mathrm{E}_{f,g_1,g_2} \mathrm{Tr}_\rho\big(d\, A_{S,f} \otimes B_{T,g_1} \otimes B_{T,g_2}\big), \tag{4.7}$$

where the expectation is taken over f, g_1, g_2 distributed as each player's respective question in $\tilde{G}$.

For any set S and $\alpha \subseteq \{\pm 1\}^S$, let $\chi_\alpha : \mathcal{F}_S \to \{\pm 1\}$ be defined by $\chi_\alpha(f) = \prod_{x\in\alpha} f(x)$. We introduce the Fourier transforms

$$\hat{A}_\alpha := \mathrm{E}_{f\in\mathcal{F}_S} \chi_\alpha(f) A_{S,f}, \qquad A_{S,f} = \sum_\alpha \chi_\alpha(f) \hat{A}_\alpha, \tag{4.8}$$

and similarly for $\hat{B}_\beta$ from $B_{T,g}$. Note that we left the dependence of the Fourier coefficients on the sets S, T implicit in the notation. Expanding the observables in (4.7) in the Fourier basis and proceeding as in [Hås01, Proof of Lemma 5.2] we arrive at the following:

$$\delta_{S,T} = \sum_\beta (1-2\varepsilon)^{|\beta|} \mathrm{Tr}_\rho\Big(\hat{A}_{\pi_S(\beta)} \otimes \hat{B}_\beta \otimes \hat{B}_\beta\Big), \tag{4.9}$$

where the summation ranges over all nonempty[13] $\beta \subseteq \{\pm 1\}^T$ and $\pi_S(\beta)$ is defined as the set of $u \in \{\pm 1\}^S$ for which there is an odd number of $v \in \beta$ whose restriction to S is u (recall that the projection constraint of G' lets us identify S with a subset of T).

Recall that in the cube game G, each player receives a $v \in V$; in $G^{\otimes K}$ the ith player for $i \in \{1,2,3\}$, is sent a K-element subset $T_i = \{v_i^1, \ldots, v_i^K\} \subseteq V^K$. The referee also holds a single set $S = \{u^1, \ldots, u^K\} \subseteq U^K$, and he checks that $\pi_{u^k v_1^k}(b_1^k) = \pi_{u^k v_2^k}(b_2^k) = \pi_{u^k v_3^k}(b_3^k)$ for all $k = 1, \ldots, K$, where $(b_i^1, \ldots, b_i^K)$ is player i's answer.

We define a strategy for the players in $G^{\otimes K}$. In this strategy the ith player, for $i \in \{1,2,3\}$, performs the measurement $\{\hat{B}^2_{\beta_i}\}_{\beta_i \subseteq \{\pm 1\}^{T_i}}$ associated with its set T_i, and answers with a random K-tuple $(b_i^1, \ldots, b_i^K) \in \beta_i$. (The fact that this is a well-defined POVM follows from (4.8) and Parseval's formula.)

The probability that the players' answers are accepted is the probability that they are consistent, i.e., their respective answers b_1^k, b_2^k, b_3^k have matching projections onto S for all $k \in \{1, \ldots, K\}$. Further selecting the optimal entangled state, the overall

[13] The fact that only nonempty Fourier coefficients appear is a result of the folding over true operation performed earlier.

success of the above-defined strategy is at least

$$q := \left\| \mathrm{E}_S \sum_{\alpha} \left(\sum_{\beta_1 \colon \pi_S(\beta_1)=\alpha} |\beta_1|^{-1} \hat{B}_{\beta_1}^2 \right) \otimes \left(\sum_{\beta_2 \colon \pi_S(\beta_2)=\alpha} |\beta_2|^{-1} \hat{B}_{\beta_2}^2 \right) \right. \\ \left. \otimes \left(\sum_{\beta_3 \colon \pi_S(\beta_3)=\alpha} |\beta_3|^{-1} \hat{B}_{\beta_3}^2 \right) \right\|_\infty . \tag{4.10}$$

Note that in the right-hand side of the expression above for clarity we suppressed the notation E_{T_i} that should accompany each $\sum_{\beta_i}$. We will keep these expectations implicit for the remainder of the proof, and it should be understood that each time we write $\sum_\beta$ we really mean $\mathrm{E}_T \sum_{\beta \subseteq \{-1,1\}^T}$, where the expectation over T is itself taken with respect to a k-element subset $S \subseteq U$ chosen uniformly by the referee and hidden from the players. We relate (4.10) to the expression appearing in (4.9). For this starting from (4.9) we first apply the inequality

$$\left\| \sum_i X_i \otimes Y_i \right\|_\infty \leq \left\| \sum_i X_i \otimes \overline{X_i} \right\|_\infty^{1/2} \left\| \sum_i Y_i \otimes \overline{Y_i} \right\|_\infty^{1/2} \tag{4.11}$$

(see, e.g., [Pis03, p. 123] for a proof) twice, first with $X_\alpha = \hat{A}_\alpha$ and

$$Y_\alpha = \sum_{\beta \colon \pi_S(\beta)=\alpha} (1-2\varepsilon)^{|\beta|} \hat{B}_\beta \otimes \hat{B}_\beta$$

and second with $X_{\beta_1} = \hat{B}_{\beta_1}$ and

$$Y_{\beta_1} = (1-2\varepsilon)^{|\beta_1|} \hat{B}_{\beta_1} \otimes \sum_{\beta_2 \colon \pi_S(\beta_2)=\pi_S(\beta_1)} (1-2\varepsilon)^{|\beta_2|} \overline{\hat{B}_{\beta_2} \otimes \hat{B}_{\beta_2}},$$

to obtain

$$\begin{aligned} \delta &\leq \left\| \sum_\alpha \hat{A}_\alpha \otimes \overline{\hat{A}_\alpha} \right\|_\infty^{1/2} \left\| \sum_\alpha \left(\sum_{\beta_1 \colon \pi_S(\beta_1)=\alpha} (1-2\varepsilon)^{|\beta|} \hat{B}_{\beta_1} \otimes \hat{B}_{\beta_1} \right) \right. \\ &\qquad \left. \otimes \left(\sum_{\beta_2 \colon \pi_S(\beta_2)=\alpha} (1-2\varepsilon)^{|\beta_2|} \overline{\hat{B}_{\beta_2} \otimes \hat{B}_{\beta_2}} \right) \right\|_\infty^{1/2} \\ &\leq \left\| \sum_\alpha \hat{A}_\alpha \otimes \overline{\hat{A}_\alpha} \right\|_\infty^{1/2} \left\| \sum_{\beta_1} \hat{B}_{\beta_1} \otimes \overline{\hat{B}_{\beta_1}} \right\|_\infty^{1/4} \left\| \sum_{\beta_1} \left(\sum_{\beta_2 \colon \pi_S(\beta_2)=\pi_S(\beta_1)} (1-2\varepsilon)^{|\beta_2|} \hat{B}_{\beta_2} \otimes \hat{B}_{\beta_2} \right) \right. \\ &\qquad \left. \otimes (1-2\varepsilon)^{2|\beta_1|} \hat{B}_{\beta_1} \otimes \overline{\hat{B}_{\beta_1}} \otimes \left(\sum_{\beta_3 \colon \pi_S(\beta_3)=\pi_S(\beta_1)} (1-2\varepsilon)^{|\beta_3|} \overline{\hat{B}_{\beta_3} \otimes \hat{B}_{\beta_3}} \right) \right\|_\infty^{1/4} . \end{aligned}$$

Next using that for any (X_i), $\| \sum_i X_i \otimes X_i \|_\infty \leq \| \sum_i |X_i|^2 \|_\infty$ and $\sum_\alpha |\hat{A}_\alpha|^2 \leq \mathrm{Id}$,

$\sum_\beta |\hat{B}_\beta|^2 \leq \mathrm{Id}$ we obtain

$$\delta \leq \left\| \sum_\alpha \left(\sum_{\beta_1 \colon \pi_S(\beta_1)=\alpha} (1-2\varepsilon)^{|\beta_1|} \left|\hat{B}_{\beta_1}\right|^2 \right) \otimes \left(\sum_{\beta_2 \colon \pi_S(\beta_2)=\alpha} (1-2\varepsilon)^{2|\beta_2|} \left|\hat{B}_{\beta_2}\right|^2 \right) \otimes \left(\sum_{\beta_3 \colon \pi_S(\beta_3)=\alpha} (1-2\varepsilon)^{|\beta_3|} \left|\hat{B}_{\beta_3}\right|^2 \right) \right\|_\infty^{1/4}.$$

Using the inequality $(1-2\varepsilon)^t \leq (2\varepsilon t)^{-1}$, valid for any $t > 0$ and $0 < \varepsilon < 1/2$, with $t = |\beta_i|$ we have shown $\delta^4 \leq q/(8\varepsilon^3)$, where q, as defined in (4.10), is the success probability of the above-defined strategy in $G^{\otimes K}$. Thus $\omega^*(G^{\otimes K}) \geq 8\varepsilon^3\delta^4$, as claimed. □

Let $K = \Theta(\gamma^{-2}\log(1/(\varepsilon\delta)))$ be large enough so that the bound from Lemma 4.7 implies that $\omega^*(G^{\otimes K}) < 8\varepsilon^3\delta^4$ whenever $\omega^*(G'') \leq 1-\gamma$. Combined with Claim 4.9 this implies that if $\omega^*(\tilde{G}) \geq (1+\delta)/2$ then $\omega^*(G'') > 1-\gamma$, establishing soundness of the reduction. □

We end this section with the following corollary, a restatement of Theorem 1.1.

COROLLARY 4.10. *Let $\varepsilon, \delta > 0$ be arbitrary constants. Then the following is NP-hard. Given a three-player XOR game G, distinguish between $\omega(G) \geq 1-\varepsilon$ and $\omega^*(G) \leq (1+\delta)/2$.*

Proof. Let φ be a 3-SAT formula and $G = G(\varphi)$ the game promised by Corollary 4.5. G is a three-player game with binary answers such that, if φ is satisfiable then $\omega(G) = 1$, but if φ is not satisfiable then $\omega^*(G) \leq 1-\varepsilon$ for some $\varepsilon > 0$. In order to apply Theorem 4.8 we need to construct a game G' with the same properties, but that is the two-out-of-three player game associated with an expanding projection game. We construct G' as follows.

- The referee samples a triple of questions (q, q', q'') as in G.
- He chooses two players at random, sends (q, q', q'') to the first and one of q, q', q'', chosen at random among the three possibilities, to the second.
- The referee receives answers (a, a', a'') from the first player and b from the second. He accepts if and only if both the following hold: $V(a, a', a''|q, q', q'') = 1$, where V is the verifier's predicate in G, and b matches the answer that the first player sent in reply to the question that was also sent to the second player.

It is clear that G' is the two-out-of-three player game associated with a projection game H. Furthermore, the game H can be made into an expanding game H' by performing the following trivial modification. With probability 1/2 play G' as defined above. With probability 1/2, select a pair of questions $((q, q', q''), q''')$, where $q''' \in \{q, q', q''\}$ and q, q', q'' are chosen independently and uniformly at random. Accept answers from the players if and only if they are consistent on q'''. Since the bipartite graph associated with this second game is expanding, the overall modified game H' is expanding as well; moreover it obviously holds that $\omega(H) \leq \omega(H') \leq 1-(1-\omega(H))/2$, and the same inequalities hold for both the value and the entangled value of the associated two-out-of-three player game. Therefore we can assume without loss of generality that G' is the two-out-of-three player game associated with an expanding projection game.

Completeness of the reduction is immediate: $\omega(G) = 1 \implies \omega(G') = 1$. To establish soundness, suppose $\omega(G') \geq 1-\varepsilon'$, where $\varepsilon' > 0$ will be specified later. Let $(|\Psi\rangle, A_{q,q',q''}, B_q)$ be a strategy for the players with success $1-\varepsilon'$ in G'. We argue

that, provided ε' is chosen small enough, $(|\Psi\rangle, B_q)$ has success at least $1-\varepsilon$ in G. For every question q and answer a let

$$A_q^a := \mathrm{E}_{(q',q'')} \sum_{a',a'':V(a,a',a''|q,q',q'')=1} A_{q,q',q''}^{a,a',a''},$$

where the expectation is taken according to the marginal distribution on questions (q',q'') when q is fixed (note that the position in which q is placed does not matter as the distribution π on questions in G is symmetric). The tests performed by the referee in game G' enforce that $\{A_q^a\}$ is a submeasurement such that both[14]

$$\mathrm{E}_q \sum_a \langle A_q^a, \mathrm{Id}\rangle_\Psi \geq 1-\varepsilon' \qquad \text{and} \qquad \mathrm{E}_q \sum_a \langle A_q^a, B_q^a\rangle_\Psi \geq 1-\varepsilon'. \tag{4.12}$$

In particular, applying Claim 2.2 to the measurements $\{A_q^a\}$ (completed with the POVM element $\mathrm{Id}-\sum_a A_q^a$) and $\{B_q^a\}$ we have that

$$\mathrm{E}_q \sum_a \left\|A_q^a - B_q^a\right\|_\Psi^2 = O\big(\sqrt{\varepsilon'}\big). \tag{4.13}$$

As a consequence, we can write the following,

$$\begin{aligned}
\omega^*(G) &\geq \mathrm{E}_{(q,q',q'')} \sum_{\substack{(a,a',a'')\\ V(a,a',a''|q,q',q'')=1}} \mathrm{Tr}_\rho\big(B_q^a \otimes B_{q'}^{a'} \otimes B_{q''}^{a''}\big)\\
&\approx_{\varepsilon'^{1/4}} \mathrm{E}_{(q,q',q'')} \sum_{\substack{(a,a',a'')\\ V(a,a',a''|q,q',q'')=1}} \sum_{\substack{(a,b',b'')\\ V(a,b',b''|q,q',q'')=1}} \mathrm{Tr}_\rho\big(A_{q,q',q''}^{a,b',b''} \otimes B_{q'}^{a'} \otimes B_{q''}^{a''}\big)\\
&\approx_{\varepsilon'} \mathrm{E}_{(q,q',q'')} \sum_{\substack{(a,a',a'')\\ V(a,a',a''|q,q',q'')=1}} \mathrm{Tr}_\rho\big(A_{q,q',q''}^{a,a',a''} \otimes B_{q'}^{a'} \otimes B_{q''}^{a''}\big)\\
&\approx_{\varepsilon'} \mathrm{E}_{(q,q',q'')} \sum_{\substack{(a,a',a'')\\ V(a,a',a''|q,q',q'')=1}} \big\langle A_{q,q',q''}^{a,a',a''}, B_{q'}^{a'}\big\rangle_\Psi\\
&\geq 1-\varepsilon',
\end{aligned}$$

where the second line uses (4.13) together with the Cauchy–Schwarz inequality, the third and fourth use (4.12), and the last is by definition of ε'. Hence $\omega^*(G) \geq 1-O((\varepsilon')^{1/4}) \geq 1-\varepsilon$ provided ε' is chosen small enough.

We have derived a reduction from deciding the satisfiability of a 3-SAT formula φ to deciding whether a two-out-of-three player game G' associated with an expanding projection game satisfies $\omega(G')=1$ or $\omega^*(G') \leq 1-\varepsilon'$ for some constant $\varepsilon'>0$. To conclude the proof of the corollary it suffices to apply Theorem 4.8 to the projection game H that underlies our construction of G' for an appropriate choice of the constant γ appearing in the theorem statement. □

5. The consolidation procedure. The proof of Theorem 3.1, which states the soundness of the low-degree plane-vs-point test against entangled players, relies on an induction procedure: the measurement $\{M^g\}$ is constructed, starting from the $\{A_{\boldsymbol{x}}^a\}$, by removing the dependence of $A_{\boldsymbol{x}}$ on the m coordinates of $\boldsymbol{x} \in \mathbb{F}^m$ one at a time. As the induction proceeds the error (as measured by an expression similar to (3.1)) blows up exponentially. To keep it bounded it is necessary to "improve"

[14] We refer to section 2 for the notation used here.

the quality of the measurements constructed at each step of the induction. The main result of this section, stated in Proposition 5.8, shows that this is possible as long as the measurements constructed remain mildly consistent with an underlying "robust" structure (which will eventually be obtained directly from measurements $A_{\boldsymbol{x}}$ passing the low-degree test with high probability). The "robustness" of the structure is used to argue that any measurement mildly consistent with it can be improved to one that is highly consistent.

This consolidation procedure bears some superficial resemblance to similar procedures used in the analysis of low-degree tests in the classical PCP literature. In particular Raz and Safra [RS97], in their analysis of the low-degree test in the small soundness regime, introduce the notion of a "consistency graph" and use it to argue that some form of weak "local consistency" is enough to imply "global consistency." The goal of our consolidation procedure, however, is slightly different: we already have a guarantee of consistency with a global object, the measurement $\{M^g\}$, that has small error, but we want to make the error even smaller, using that some strong form of local consistency holds. Moreover, while the analysis in [RS97] is combinatorial and uses subtle properties of the graph to handle the case of small soundness (large error), ours is algebraic and only uses basic facts about graph expansion and takes advantage of the low-error regime we work in. Thus even though it is possible that a deeper connection exists between the two constructions we do not see it explicitly.

We first define precisely the three properties of a measurement $\{M^g\}$ that we wish to improve. In section 5.2 we introduce the notion of a (δ, μ)-robust triple, the underlying structure that will enable the improvement. In section 5.3 we show how two of the three properties can be improved. Finally, the main result, Proposition 5.8, is proved in section 5.4.

5.1. Consistency parameters. The measurements $\{M^g\}$ constructed throughout the induction will not always be complete measurements, i.e., they will satisfy $0 \leq M^g \leq \mathrm{Id}$ and $\sum^g M^g \leq \mathrm{Id}$, but not necessarily with equality. Whenever these conditions hold we call $\{M^g\}$ a *submeasurement*. The following parameters will be used in our analysis.

DEFINITION 5.1. *Let S be a finite set, $A = \{A^g\}_{g\in S}$ such that $0 \leq A^g \leq \mathrm{Id}$ for every g, and $M = \{M^g\}_{g\in S}$ a submeasurement. For any $\delta, \gamma, \eta > 0$, we say that M is*

- *δ-consistent with A if $\sum_g \langle \mathrm{Id} - A^g, M^g \rangle_\Psi \leq \delta$,*
- *γ-projective if $\langle M, \mathrm{Id} - M \rangle_\Psi \leq \gamma$, where $M := \sum_g M^g$,*
- *η-complete if $\mathrm{Tr}_\rho(M) = \langle M, \mathrm{Id} \rangle_\Psi \geq (1 - \eta)$.*

If M satisfies the first item with $A = M$ we also say that M is δ-self-consistent.

The first property in the definition, consistency, can be understood as a measure of distance: measurements that are consistent are "close" in a precise sense (see the discussion following the statement of Theorem 3.1 for more on this). The second property, projectivity, intuitively measures how far an operator M is from being self-consistent, or "orthogonal" (if $|\Psi\rangle$ was the maximally entangled state on two subsystems, $\langle M, \mathrm{Id} - M \rangle_\Psi$ would be 0 if and only if M is the orthogonal projection on a subspace). The last property, completeness, measures how far a submeasurement is from being complete. Note that complete measurements are automatically 0-projective.

5.2. Robust triples. In this section we define the notion of (δ, μ)-robust triple, and prove some useful properties.

DEFINITION 5.2. *Let $G = (V, E)$ be a graph, S a finite set, $\mathcal{G} \subseteq \{g : V \to S\}$ a set of functions, and for every $v \in V$, $A_v = \{A_v^a\}_{a \in S}$ a measurement with outcomes in S. Given $\delta > 0$ and $0 < \mu \leq 1$, we say that $(G, A, \mathcal{G})$ is a (δ, μ)-robust triple if the following hold:*

1. *(self-consistency) The measurements A are δ-self-consistent, on average over $v \in V$:*
$$\mathrm{E}_{v \in V} \sum_{a \in S} \langle A_v^a, \mathrm{Id} - A_v^a \rangle_\Psi \leq \delta.$$
2. *(small intersection) For any $g \neq g' \in \mathcal{G}$, $\Pr_{v \in V}\big(g(v) = g'(v)\big) \leq \delta$.*
3. *(stability) For any submeasurement $\{R^g\}_{g \in \mathcal{G}}$ it holds that*
$$\mathrm{E}_{v \in V} \mathrm{E}_{v' \in N(v)} \sum_g \langle R^g, (A_v^{g(v)} - A_{v'}^{g(v')})^2 \rangle_\Psi \leq \delta,$$
where $N(v)$ is the set of neighbors of v in G.
4. *(expansion) G has mixing time $O(\mu^{-1})$. Precisely, if for any $v \in V$ we let $p_k(v)$ denote the distribution on V that results from starting a k-step random walk at v, then for any $\delta > 0$ and some $k = O(\log(1/\delta)\log(1/\mu))$ it holds that $\mathrm{E}_{v \in V} \|p_k(v) - u\|_1 \leq \delta$, where u is the uniform distribution on V.*

We will sometimes make the underlying state $|\Psi\rangle$ explicit by writing the triple as $(G, A, \mathcal{G})_\Psi$.

We note a useful property that follows from the definition.

CLAIM 5.3. *Suppose $(G, A, \mathcal{G})_\Psi$ is a (δ, μ)-robust triple. Then the measurements A_v are almost-projective, in the sense that*

$$\mathrm{E}_v \sum_a \langle A_v^a - (A_v^a)^2, \mathrm{Id} \rangle_\Psi = O\big(\sqrt{\delta}\big). \tag{5.1}$$

Furthermore, there exists a $\delta' = O(\delta^{1/2} \log^2(1/\delta) \log^2(1/\mu))$ such that for any submeasurement $\{R^g\}_{g \in \mathcal{G}}$,

$$\sum_g \big\langle R^g, A^g - (A^g)^2 \big\rangle_\Psi \leq \delta',$$

where $A^g := \mathrm{E}_{v \in V} A_v^{g(v)}$.

Proof. The first property in the claim follows from the self-consistency condition. Indeed,

$$\begin{aligned}
\mathrm{E}_v \sum_a \langle A_v^a - (A_v^a)^2, \mathrm{Id} \rangle_\Psi &= \mathrm{E}_v \sum_a \langle A_v^a (\mathrm{Id} - A_v^a), \mathrm{Id} \rangle_\Psi \\
&= \mathrm{E}_v \sum_{a,b} \langle A_v^a (\mathrm{Id} - A_v^a), A_v^b \rangle_\Psi \\
&\approx_{\sqrt{\delta}} \mathrm{E}_v \sum_a \langle A_v^a (\mathrm{Id} - A_v^a), A_v^a \rangle_\Psi \\
&\approx_{\sqrt{\delta}} \mathrm{E}_v \sum_a \langle \mathrm{Id} - A_v^a, A_v^a \rangle_\Psi \\
&\leq \delta,
\end{aligned}$$

where the third and fourth lines follow from the Cauchy–Schwarz inequality and the self-consistency condition.

To show the second property, first recall from Definition 5.2 that the robustness condition implies the stability property

$$\mathrm{E}_{v,v'\in N(v)} \sum_g \big\langle R^g, (A_v^{g(v)} - A_{v'}^{g(v')})^2 \big\rangle_\Psi \le \delta. \tag{5.2}$$

We first observe that this condition implies that for any $k \ge 1$,

$$\mathrm{E}_{v,v'\in N^k(v)} \sum_g \big\langle R^g, (A_v^{g(v)} - A_{v'}^{g(v')})^2 \big\rangle_\Psi \le k^2\,\delta, \tag{5.3}$$

where $N^k(v)$ is the set of all $v' \in V$ that are at distance at most k from v in G, and the distribution on $N^k(v)$ is the one that results from starting a k-step random walk at v. Indeed, (5.3) simply follows from (5.2) and successive applications of the triangle inequality.

The expansion condition in the definition of a (δ,μ)-robust triple implies that for some $k = O(\log(1/\delta)\log(1/\mu))$ the distribution on $N^k(v)$ is δ-close in statistical distance to uniform on V. Applying (5.3) for this k, we get

$$\mathrm{E}_{v,v'} \sum_g \big\langle R^g, (A_v^{g(v)} - A_{v'}^{g(v')})^2 \big\rangle_\Psi \le (k^2+2)\delta,$$

since for any v, v', $\sum_g \langle R^g, (A_v^{g(v)} - A_{v'}^{g(v')})^2\rangle_\Psi \le 2$. Expanding the square and using (5.1), we see that

$$\sum_g \big\langle R^g, (A^g)^2 \big\rangle_\Psi \ge \sum_g \big\langle R^g, A^g \big\rangle_\Psi - \frac{k^2+2}{2}\,\delta - O\big(\sqrt{\delta}\big),$$

which proves the claim. □

5.3. Consistency consolidation. In this section we prove the following lemma, which establishes consolidation for the consistency and projectivity properties (see Definition 5.1). Intuitively, the lemma shows that if a submeasurement Q is mildly consistent with a family of measurements $\{A_v\}$ that underlie a robust triple, then Q can be modified into a submeasurement S that is highly consistent with A and projective.

LEMMA 5.4. *Let $\delta, \eta > 0$ be such that $\delta \le \eta \le 1/2$, $\mu > 0$, and $|\Psi\rangle$ a permutation-invariant state over $r \ge 3$ registers. Let $(G, A, \mathcal{G})_\Psi$ be a (δ,μ)-robust triple and $\{Q^g\}_{g\in\mathcal{G}}$ a submeasurement that is η-consistent with A. Then there exists a submeasurement $\{S^g\}_{g\in\mathcal{G}}$ such that S is η'-consistent with A and projective, for some $\eta' = O(\delta^{1/4}\log^2(1/\delta)\log^2(1/\mu))$ that is independent of η. Moreover, S also satisfies $\langle S, \mathrm{Id}\rangle_\Psi \ge \langle Q, \mathrm{Id}\rangle_\Psi - \eta - \eta'$.*

The remainder of the section is devoted to the proof of the lemma. For any $g \in \mathcal{G}$, let $A^g := \mathrm{E}_{v\in V} A_v^{g(v)}$. We first give the following useful claim.

CLAIM 5.5. *Let $r \ge 2$, $|\Psi\rangle$ an r-register permutation-invariant state, and ρ the reduced density of $|\Psi\rangle\langle\Psi|$ on any one register. Suppose that $\{A_v^a\}$ is a family of*

measurements that is δ-self-consistent, i.e.,

$$\mathrm{E}_v \sum_a \langle A_v^a, \mathrm{Id} - A_v^a \rangle_\Psi \;=\; \mathrm{E}_v \sum_a \langle \Psi | A_v^a \otimes (\mathrm{Id} - A_v^a) \otimes \mathrm{Id}^{\otimes(r-2)} | \Psi \rangle \;\leq\; \delta.$$

Then the following holds:

$$\mathrm{E}_v \sum_a \mathrm{Tr}\big(A_v^a \rho^{1/2} (\mathrm{Id} - \overline{A_v^a}) \rho^{1/2}\big) \;\leq\; 2\delta.$$

Proof. Let $|\Psi\rangle = \sum_i \sqrt{\lambda_i} |u_i\rangle |v_i\rangle$ be the Schmidt decomposition, where the $|u_i\rangle$ are orthonormal vectors on the first register, and $|v_i\rangle$ are on the remaining $(r-1)$ registers. By definition, $\rho = \sum_i \lambda_i |u_i\rangle\langle u_i|$. By self-consistency of A it holds that

$$\mathrm{E}_v \sum_a \langle \Psi | A_v^a \otimes A_v^a \otimes \mathrm{Id} \, | \Psi \rangle \;=\; \mathrm{E}_v \sum_a \sum_{i,j} \sqrt{\lambda_i}\sqrt{\lambda_j} \langle u_i | A_v^a | u_j \rangle \langle v_i | A_v^a \otimes \mathrm{Id} \, | v_j \rangle \;\geq\; 1 - \delta,$$

where here the identity Id acts on all but the first two players' subspaces. Applying the Cauchy–Schwarz inequality, we get

$$\mathrm{E}_v \sum_a \sum_{i,j} \sqrt{\lambda_i}\sqrt{\lambda_j} \big| \langle u_i | A_v^a | u_j \rangle \big|^2 \;\geq\; (1-\delta)^2 \;\geq\; 1 - 2\delta.$$

To conclude, observe that the left-hand side is exactly

$$1 - \mathrm{E}_v \sum_a \mathrm{Tr}(A_v^a \rho^{1/2} (\mathrm{Id} - \overline{A_v^a}) \rho^{1/2}). \qquad \square$$

The proof of Lemma 5.4 relies on the use of the following SDP. Recall that $\rho = \mathrm{Tr}_{\mathcal{H}^{\otimes(r-1)}} |\Psi\rangle\langle\Psi|$ is the reduced density of $|\Psi\rangle$ on any players' subspace, and we may always assume it is invertible (if not, simply restrict all measurement operators to the support of ρ).

Primal SDP

$$\begin{aligned} \omega \;:=\; & \max \sum_g \mathrm{Tr}\big(T^g \rho^{1/2} \overline{A^g} \rho^{1/2}\big) \\ s.t. \quad & \forall g,\, T^g \geq 0, \quad \sum_g T^g \leq \mathrm{Id}\,. \end{aligned} \tag{5.4}$$

Dual SDP

$$\begin{aligned} & \min \mathrm{Tr}(X) \\ s.t. \quad & \forall g,\, X \;\geq\; \rho^{1/2} \overline{A^g} \rho^{1/2}, \\ & X \geq 0. \end{aligned} \tag{5.5}$$

We make a few preliminary observations about the SDP.

Claim 5.6. *Strong duality holds for* (5.4) *and* (5.5). *Let $\{T^g\}$ be an optimal solution to the primal, X a matching dual solution, and $Z := \rho^{-1/2} X \rho^{-1/2}$. Then the following hold:*

1. $\omega = \sum_g \mathrm{Tr}\big(T^g \rho^{1/2} \overline{A^g} \rho^{1/2}\big) \geq \mathrm{Tr}_\rho(Q) - \eta - O(\sqrt{\delta})$;
2. $T := \sum_g T^g = \mathrm{Id}$;
3. $\forall g, \quad T^g \rho^{1/2} Z = T^g \rho^{1/2} \overline{A^g} \quad$ *and* $\quad Z \rho^{1/2} T^g = \overline{A^g} \rho^{1/2} T^g$.

Proof. It is easy to verify that both primal and dual SDPs are strictly feasible, and hence strong duality holds. By choosing $T^g := Q^g$ in (5.4) we get

$$\begin{aligned}\omega &\geq \sum_g \operatorname{Tr}\bigl(Q^g \rho^{1/2} \overline{A^g} \rho^{1/2}\bigr) \\ &\approx_{\sqrt{\delta}} \sum_g \operatorname{Tr}\bigl(\rho^{1/2} Q^g A^g \rho^{1/2}\bigr) \\ &= \sum_g \operatorname{Tr}\bigl(Q^g A^g \otimes \mathrm{Id}\bigr) \\ &\approx_{\sqrt{\delta}} \sum_g \operatorname{Tr}_\rho\bigl(Q^g \otimes A^g\bigr) \\ &\geq \operatorname{Tr}_\rho(Q) - \eta,\end{aligned}$$

where the second line uses the Cauchy–Schwarz inequality and Claim 5.5, the fourth line self-consistency of A, and the fifth follows from η-consistency of Q and A. This proves the first item in the claim.

Let $(\{T^g\}, X)$ be an optimal primal-dual solution pair. Clearly, we may without loss of generality assume that $T = \sum_g T^g = \mathrm{Id}$, as imposing this can only improve the primal objective function. Finally, the last conditions stated in the claim follow from the complementary slackness conditions

$$\forall g, \quad T^g X = T^g \rho^{1/2} \overline{A^g} \rho^{1/2} \quad \text{and} \quad X T^g = \rho^{1/2} \overline{A^g} \rho^{1/2} T^g, \tag{5.6}$$

the definition of Z and the fact that ρ is invertible. ∎

Let $(\{T^g\}, X)$ be an optimal primal-dual solution pair to (5.4)–(5.5), and for every $g \in \mathcal{G}$ define

$$S^g := \mathrm{E}_v A_v^{g(v)} T^g A_v^{g(v)}. \tag{5.7}$$

For every g, we have $0 \leq S^g \leq A^g$ and $S := \sum_g S^g \leq \mathrm{Id}$. We first prove the following about $\{S^g\}$.

Claim 5.7. *Let $(G, A, \mathcal{G})_\Psi$ be a (δ, μ)-robust triple, $\{S^g\}$ as defined in* (5.7)*, and δ' as in Claim* 5.3*. Then $\{S^g\}$ is a submeasurement such that*

1. $\operatorname{Tr}_\rho(S) = \sum_g \operatorname{Tr}_\rho(S^g) \geq \omega - O(\sqrt{\delta})$,
2. *S is $O(\delta')$-consistent with A.*

Proof. We have

$$\begin{aligned}\operatorname{Tr}_\rho(S) &= \mathrm{E}_v \sum_g \operatorname{Tr}\bigl(A_v^{g(v)} T^g A_v^{g(v)} \rho^{1/2} \,\mathrm{Id}\, \rho^{1/2}\bigr) \\ &\approx_{\sqrt{\delta}} \mathrm{E}_v \sum_g \operatorname{Tr}\bigl(T^g \rho^{1/2} \overline{A_v^{g(v)}} \rho^{1/2}\bigr) \\ &\geq \omega,\end{aligned}$$

where the second line follows from Claim 5.5 and the third is by definition of ω. This proves the first item in the claim. To show the second, note that

$$\sum_g \bigl\langle S^g, \mathrm{Id} - A^g \bigr\rangle_\Psi \approx_{\sqrt{\delta}} \sum_g \bigl\langle T^g, A^g (\mathrm{Id} - A^g) A^g \bigr\rangle_\Psi = O(\delta' + \sqrt{\delta}),$$

where the first approximate equality uses self-consistency of A and the last equality follows from Claim 5.3. ∎

The condition on consistency of S and A in Lemma 5.4 now follows from the second item in Claim 5.7 provided $\eta' = \Omega(\delta')$, and the completeness condition follows from the first item in Claim 5.7 together with the first item in Claim 5.6, provided $\eta' = \Omega(\sqrt{\delta})$. To complete the proof of the lemma it only remains to verify the projectivity condition. Recall that $Z = \rho^{-1/2} X \rho^{-1/2}$. We have

$$\begin{aligned}
\langle S, \mathrm{Id} - S \rangle_\Psi &= \mathrm{Tr}_\rho\big(S \otimes (\mathrm{Id} - S) \otimes \mathrm{Id}^{\otimes(r-2)}\big) \\
&\approx_{\sqrt{\delta}} \sum_g \mathrm{Tr}_\rho\big(T^g \otimes (\mathrm{Id} - S) \otimes A^g\big) \\
&\leq \sum_g \mathrm{Tr}_\rho\big(T^g \otimes (\mathrm{Id} - S) \otimes \overline{Z}\big) \\
&= \mathrm{Tr}_\rho\big((\mathrm{Id} - S) \otimes \overline{Z}\big) \\
&\leq \mathrm{Tr}_\rho\big(\overline{Z}\big) - \sum_g \mathrm{Tr}_\rho\big(S^g \otimes A^g\big) \\
&= \sum_g \mathrm{Tr}\big(T^g \rho^{1/2} \overline{A^g} \rho^{1/2}\big) - \mathrm{E}_v \sum_g \mathrm{Tr}_\rho\big(A_v^{g(v)} T^g A_v^{g(v)} \otimes A^g\big) \\
&\approx_{\sqrt{\delta}+\delta'} \sum_g \mathrm{Tr}_\rho\big(A_v^{g(v)} T^g A_v^{g(v)}\big) - \mathrm{E}_v \sum_g \mathrm{Tr}_\rho\big(A_v^{g(v)} T^g A_v^{g(v)}\big) \\
&= O\big(\delta' + \sqrt{\delta}\big),
\end{aligned}$$

where the second line uses the definition of S and self-consistency of A; the third uses the dual constraint and the definition of Z; the fourth item 2 from Claim 5.6; the fifth again uses the dual constraint; the sixth uses strong duality and the fact that $\mathrm{Tr}_\rho(\overline{Z})$ is real (since Z is Hermitian) for the first term, and the definition of S for the second; the seventh uses Claim 5.5 for the first term and Claim 5.3 for the second. This establishes the projectivity condition on S provided $\eta' = \Omega(\delta' + \sqrt{\delta})$.

5.4. Self-consolidation. The following proposition states our main "consolidation" result.

PROPOSITION 5.8 (self-consolidation). *There exists a constant $K > 0$ such that the following holds. Let $r \geq 3$, H a symmetric r-player game, X a finite set, and for every $\boldsymbol{x} \in X$, $G_{\boldsymbol{x}} = (V_{\boldsymbol{x}}, E_{\boldsymbol{x}})$ a graph, $S_{\boldsymbol{x}}$ a set, and $\mathcal{G}_{\boldsymbol{x}} \subseteq \{g : V_{\boldsymbol{x}} \to S_{\boldsymbol{x}}\}$.*

Suppose that for any $0 < \varepsilon < K$ and strategy $(P, |\Psi\rangle)$ for the players that has success $1 - \varepsilon$ in the game H and is ε-self-consistent there exists a collection $A_{\boldsymbol{x}} = \{A^a_{\boldsymbol{x},v}\}_{a \in S}$ of projective measurements defined for every $v \in V_{\boldsymbol{x}}$, possibly depending on P but independent of $|\Psi\rangle$, such that that for all $\boldsymbol{x} \in X$, $(G_{\boldsymbol{x}}, A_{\boldsymbol{x}}, \mathcal{G}_{\boldsymbol{x}})_\Psi$ is a (δ, μ)-robust triple for some $\delta, \mu > 0$ such that $\eta'(\delta, \mu) < 1/4$, where η' is as defined in Lemma 5.4.

Suppose further that for any $\varepsilon' > 0$ there exists $\eta = \eta(\varepsilon')$ such that $\eta \to 0$ as $\varepsilon' \to 0$, and whenever $(P', |\Psi'\rangle)$ is a strategy with success $1 - \varepsilon'$ in H, there exists a family of submeasurements $\{Q^g_{\boldsymbol{x}}\}_{g \in \mathcal{G}}$ that is η-consistent with $A'_{\boldsymbol{x}}$ (obtained from P') and η-complete, on average over $\boldsymbol{x} \in X$.

Then for any small enough $0 < \varepsilon < K$ and strategy $(P, |\Psi\rangle)$ for the players that has success $1 - \varepsilon$ in the game H and is ε-self-consistent there exists a family of (complete) measurements $\{R^g_{\boldsymbol{x}}\}_{g \in \mathcal{G}}$ that is η_c-consistent with $A_{\boldsymbol{x}}$ for some $\eta_c = O(r(\eta')^{1/4})$, on average over $\boldsymbol{x} \in X$.

The significance of the proposition is that the parameter η_c is independent of the parameter η associated with the family $\{Q^g_{\boldsymbol{x}}\}$. It only depends on the robust-

ness parameters δ, μ (which themselves implicitly depend on ε). Before proving the proposition we establish two general claims about symmetric r-player games.

CLAIM 5.9. *Let $r \geq 2$, H a symmetric r-player game, and $\{P_q^a\}$ a set of projective measurements for the players in H (here $q \in Q$ and $a \in A$, respectively, the sets of possible questions and answers in the game). Then there exists an operator $X = X(H,P)$ such that $0 \leq X \leq \mathrm{Id}$ and for any permutation-invariant state $|\Psi\rangle$, the success probability p_s of the strategy $(P, |\Psi\rangle)$ in H satisfies*

$$\left|p_s - \langle\Psi| X \otimes \mathrm{Id}^{\otimes(r-1)} |\Psi\rangle\right| \leq 2(r-1)\sqrt{2\delta},$$

where δ is the self-consistency parameter

$$\delta = \mathrm{E}_q \sum_a \langle\Psi| P_q^a \otimes (\mathrm{Id} - P_q^a) \otimes \mathrm{Id}^{\otimes(r-2)} |\Psi\rangle,$$

and here the expectation is taken according to the marginal distribution of questions on a single player.

Proof. We do the proof for the case $r = 2$; the general case is similar. The players' success probability in H can be expressed as

$$p_s = \mathrm{E}_{(q,q')} \sum_{(a,a')} V(a,a'|q,q') \langle\Psi| P_q^a \otimes P_{q'}^{a'} |\Psi\rangle,$$

where $V(a,a'|q,q') \in \{0,1\}$ are coefficients representing the referee's decision to accept or reject the pair of answers (a,a') to the questions (q,q'). Let

$$X := \mathrm{E}_{(q,q')} \sum_{(a,a')} V(a,a'|q,q') \sqrt{P_q^a} P_{q'}^{a'} \sqrt{P_q^a}.$$

Then $0 \leq X \leq \mathrm{Id}$. Let $Y := \mathrm{E}_{(q,q')} \sum_{(a,a')} V(a,a'|q,q') \sqrt{P_q^a} \otimes P_{q'}^{a'} \sqrt{P_q^a}$. We have

$$\begin{aligned}
&\left|p_s - \langle\Psi|Y|\Psi\rangle\right| \\
&\quad = \left|\mathrm{E}_{(q,q')} \sum_{(a,a')} V(a,a'|q,q') \langle\Psi| \left(\sqrt{P_q^a} \otimes P_{q'}^{a'}\, \mathrm{Id}\right) \left(\sqrt{P_q^a} \otimes \mathrm{Id} - \mathrm{Id} \otimes \sqrt{P_q^a}\right) |\Psi\rangle\right| \\
&\quad \leq \left(\mathrm{E}_{(q,q')} \sum_{a,a'} \langle\Psi| P_q^a \otimes P_{q'}^{a'} |\Psi\rangle\right)^{1/2} \\
&\qquad \cdot \left(\mathrm{E}_{(q,q')} \sum_{a,a'} \langle\Psi| \left(\sqrt{P_q^a} \otimes \mathrm{Id} - \mathrm{Id} \otimes \sqrt{P_q^a}\right) \left(\mathrm{Id} \otimes P_{q'}^{a'}\right) \right. \\
&\qquad \left. \cdot \left(\sqrt{P_q^a} \otimes \mathrm{Id} - \mathrm{Id} \otimes \sqrt{P_q^a}\right) |\Psi\rangle\right)^{1/2} \\
&\quad \leq \left(2 - 2\mathrm{E}_q \sum_a \langle\Psi| \sqrt{P_q^a} \otimes \sqrt{P_q^a} |\Psi\rangle\right)^{1/2} \\
&\quad \leq \sqrt{2\delta},
\end{aligned}$$

where the second line follows from the Cauchy–Schwarz inequality and $V(a,a'|q,q') \leq 1$ and the last uses the definition of δ and $\sqrt{P_q^a} \geq P_q^a$ since $0 \leq P_q^a \leq \mathrm{Id}$ for every q, a. A similar sequence of inequalities results in the bound

$$|\langle\Psi|Y|\Psi\rangle - \langle\Psi| X \otimes \mathrm{Id} |\Psi\rangle| \leq \sqrt{2\delta},$$

and combining the two proves the lemma for $r = 2$. For general r the proof is the same but the operators corresponding to players $2, \dots, r$ need to be brought to the first register one at a time, so that the error is $(r-1)$ times what it is for $r = 2$. ∎

CLAIM 5.10. *Let $\varepsilon > 0$, $r \geq 2$, $0 \leq R \leq \mathrm{Id}$ be such that $\langle R, \mathrm{Id}\rangle_\Psi \geq 1/2$, $\delta := \langle R, \mathrm{Id} - R\rangle_\Psi$, and $|\Phi\rangle := (\mathrm{Id} - R)^{\otimes r}|\Psi\rangle / z$, where $z = \|(\mathrm{Id} - R)^{\otimes r}|\Psi\rangle\|$. Then*

$$\left\|(\mathrm{Id} - R)^{\otimes r}|\Psi\rangle - \left(\mathrm{Id}^{\otimes(r-1)} \otimes (\mathrm{Id} - R)\right)|\Psi\rangle\right\|^2 \leq r^2\delta, \tag{5.8}$$

and it holds that

$$z^2 \geq 1 - \langle R, \mathrm{Id}\rangle_\Psi - 3r\sqrt{\delta}. \tag{5.9}$$

Moreover, there is an $\varepsilon' = O(r(\varepsilon^{1/4} + \delta^{1/4})\langle \mathrm{Id} - R, \mathrm{Id}\rangle_\Psi^{-1/2})$ such that for any symmetric r-player game H and projective symmetric strategy $(P, |\Psi\rangle)$ for the players that has success $1 - \varepsilon$ in H and is ε-self-consistent the strategy $(P, |\Phi\rangle)$ has success at least $1 - \varepsilon'$ in H.

Proof. Let $|\tilde{\Phi}\rangle := (\mathrm{Id} - R)^{\otimes r}|\Psi\rangle$. We first evaluate

$$\begin{aligned}\left\||\tilde{\Phi}\rangle - \left(\mathrm{Id} \otimes (\mathrm{Id} - R)^{\otimes(r-1)}\right)|\Psi\rangle\right\|^2 &= \left\|R \otimes (\mathrm{Id} - R)^{\otimes(r-1)}|\Psi\rangle\right\|^2 \\ &\leq \langle R, \mathrm{Id} - R\rangle_\Psi \\ &= \delta.\end{aligned}$$

Repeating a similar inequality r times and using the triangle inequality shows (5.8). Let X be the operator whose existence is guaranteed by Claim 5.9. Since by assumption the strategy $(P, |\Psi\rangle)$ is ε-self-consistent, the claim shows that

$$\langle \Psi | X \otimes \mathrm{Id}^{\otimes(r-1)} |\Psi\rangle \geq 1 - \varepsilon - 2(r-1)\sqrt{2\varepsilon}. \tag{5.10}$$

Let $|\tilde{\Psi}\rangle = |\Psi\rangle - |\tilde{\Phi}\rangle$. By (5.8) it holds that

$$\left\||\tilde{\Psi}\rangle - \left(\mathrm{Id}^{\otimes(r-1)} \otimes R\right)|\Psi\rangle\right\|^2 \leq r^2\delta, \tag{5.11}$$

so that

$$\||\tilde{\Psi}\rangle\|^2 \leq \left(\langle R^2, \mathrm{Id}\rangle_\Psi^{1/2} + r\sqrt{\delta}\right)^2 \tag{5.12}$$

$$\leq \langle R, \mathrm{Id}\rangle_\Psi + 2r\sqrt{\delta} + r^2\delta, \tag{5.13}$$

where we used $R^2 \leq R$ since $R \leq \mathrm{Id}$. Using that $|\Psi\rangle$ is a unit vector, (5.8) also implies

$$\begin{aligned}z^2 &= \||\tilde{\Phi}\rangle\|^2 \\ &\leq \left\|\left(\mathrm{Id}^{\otimes(r-1)} \otimes (\mathrm{Id} - R)\right)|\Psi\rangle\right\|^2 + 2r\sqrt{\delta} + r^2\delta \\ &\leq \langle \Psi | \mathrm{Id}^{\otimes(r-1)} \otimes (\mathrm{Id} - R)|\Psi\rangle + 2r\sqrt{\delta} + r^2\delta \\ &= 1 - \langle R, \mathrm{Id}\rangle_\Psi + 3r^2\sqrt{\delta}, \qquad (5.14)\end{aligned}$$

where for the third line we used $(\mathrm{Id} - R)^2 \leq (\mathrm{Id} - R)$ since $0 \leq R \leq \mathrm{Id}$. We may also obtain a bound in the other direction as

$$\begin{aligned}z^2 &\geq \left\|\left(\mathrm{Id}^{\otimes(r-1)} \otimes (\mathrm{Id} - R)\right)|\Psi\rangle\right\|^2 - 2r\sqrt{\delta} \\ &\geq \langle \Psi | \mathrm{Id}^{\otimes(r-2)} \otimes (\mathrm{Id} - R) \otimes (\mathrm{Id} - R)|\Psi\rangle - 2r\sqrt{\delta} \\ &\geq 1 - \langle R, \mathrm{Id}\rangle_\Psi - \delta - 2r\sqrt{\delta},\end{aligned}$$

where here the second line follows from the Cauchy–Schwarz inequality and the third uses the definition of δ. This proves (5.9). Combining (5.13) and (5.14) we obtain that

$$z^{-2}(1-\||\tilde{\Psi}\rangle\|^2) \geq 1-6r^2\sqrt{\delta}/\langle R,\mathrm{Id}\rangle_\Psi. \tag{5.15}$$

From (5.10) we get

$$\begin{aligned}1-\varepsilon-2(r-1)\sqrt{2\varepsilon} &\leq \langle\Psi|X|\Psi\rangle\\ &= \langle\tilde{\Psi}|X|\tilde{\Psi}\rangle+\langle\tilde{\Phi}|X|\tilde{\Phi}\rangle+2\Re\langle\tilde{\Psi}|X|\tilde{\Phi}\rangle\\ &\leq \||\tilde{\Psi}\rangle\|^2+z^2\langle\Phi|X|\Phi\rangle\\ &\quad+2|\langle\Psi|X(\mathrm{Id}-R)\otimes(\mathrm{Id}-R)^{\otimes(r-2)}\otimes R(\mathrm{Id}-R)|\Psi\rangle|+2r\sqrt{\delta},\end{aligned}$$

where for the last inequality we used (5.11). Using the Cauchy–Schwarz inequality and consistency of R, the third term above is at most $2\sqrt{\delta}$. Rearranging terms we see that

$$\langle\Phi|X|\Phi\rangle \geq z^{-2}\big(1-\||\tilde{\Psi}\rangle\|^2-(\varepsilon+2(r-1)\sqrt{2\varepsilon}+2(r+1)\sqrt{\delta}\big),$$

which using (5.15) and the assumption on $\langle R,\mathrm{Id}\rangle_\Psi$ shows

$$\langle\Phi|X|\Phi\rangle \geq 1-\Omega(r^2(\sqrt{\varepsilon}+\sqrt{\delta})(\langle\mathrm{Id}-R,\mathrm{Id}\rangle_\Psi^{-1}).$$

The same calculation can be done replacing X by the operator $Y=\mathrm{E}_q\sum_a P_q^a\otimes P_q^a$ that represents the player's consistency, showing that $(P,|\Phi\rangle)$ is $O(r^2(\sqrt{\varepsilon}+\sqrt{\delta})\langle\mathrm{Id}-R,\mathrm{Id}\rangle_\Psi^{-1})$-self-consistent. We may thus apply Claim 5.9 to finish the proof. ∎

We end this section with the proof of Proposition 5.8.

Proof of Proposition 5.8. Let ε, r, H, P, $|\Psi\rangle$, X, and $(G_{\boldsymbol{x}},A_{\boldsymbol{x}},\mathcal{G}_{\boldsymbol{x}})$ be as in the statement of the proposition. Let K be such that for any $\varepsilon'<K$ it holds that $\eta(\varepsilon')\leq 1/4$. This is possible since $\eta(\varepsilon')\to 0$ as $\varepsilon'\to 0$.

By assumption, $(G_{\boldsymbol{x}},A_{\boldsymbol{x}},\mathcal{G}_{\boldsymbol{x}})_\Psi$ is a (δ,μ)-robust triple, and there is a family of submeasurements $\{Q_{\boldsymbol{x}}^g\}$ that is η_1-consistent with $A_{\boldsymbol{x}}$ and η_1-complete, where $\eta_1=\eta(\varepsilon)$. Given our choice of ε, $\eta_1<1/2$, hence we may apply Lemma 5.4 (for every $\boldsymbol{x}$) to deduce the existence of submeasurements $\{\tilde{Q}_{\boldsymbol{x}}^g\}$ that are (on average over $\boldsymbol{x}\in X$) η_2-consistent with $A_{\boldsymbol{x}}$, η_2 projective, and $(\eta_1+\eta_2)$-complete, where $\eta_2=\eta'(\varepsilon)$. Among all submeasurements that are η_2-consistent with $A_{\boldsymbol{x}}$ and η_2-projective, let $\{R_{\boldsymbol{x}}^g\}$ be the one that has the smallest completeness parameter. Note that the existence of $\tilde{Q}_{\boldsymbol{x}}$ implies that necessarily $\langle R,\mathrm{Id}\rangle_\Psi\geq 1-(\eta_1+\eta_2)\geq 1/2$, where as usual $R=\mathrm{E}_{\boldsymbol{x}\in X}\sum_g R_{\boldsymbol{x}}^g$.

If $\langle\mathrm{Id}-R,\mathrm{Id}\rangle_\Psi\leq\eta_2^{1/4}$ then we are done. Otherwise, let $\varepsilon_2:=\varepsilon'(\varepsilon,\eta_2,R)=O(\eta_2^{1/8})$, where ε' is as defined in Claim 5.10. Let $|\tilde{\Phi}\rangle:=(\mathrm{Id}-R)^{\otimes r}|\Psi\rangle$, $z=\||\tilde{\Phi}\rangle\|$, and $|\Phi\rangle:=|\tilde{\Phi}\rangle/z$. By Claim 5.10, the strategy $(P,|\Phi\rangle)$ has success $1-\varepsilon_2$ in H. The assumption made in the proposition thus guarantees the existence of another family of submeasurements $\{S_{\boldsymbol{x}}^g\}$ that is η_3-consistent with $A_{\boldsymbol{x}}$ and η_3-complete for some $\eta_3=\eta(\varepsilon_2)$. Provided K is chosen small enough, using $\eta(\varepsilon')\to 0$ as $\varepsilon'\to 0$ it holds that $\eta_3<1/4$. Consider a new family of submeasurements $T_{\boldsymbol{x}}=\{T_{\boldsymbol{x}}^g\}$, where for every $g\in\mathcal{G}$,

$$T_{\boldsymbol{x}}^g := RR_{\boldsymbol{x}}^gR+(\mathrm{Id}-R)S_{\boldsymbol{x}}^g(\mathrm{Id}-R).$$

The $\{T_{\boldsymbol{x}}^g\}$ are clearly nonnegative, and sum to at most Id since both $\{R_{\boldsymbol{x}}^g\}$ and $\{S_{\boldsymbol{x}}^g\}$ do. Moreover,

$$\begin{aligned}
&\mathrm{E}_{\boldsymbol{x}} \sum_g \langle T_{\boldsymbol{x}}^g, \mathrm{Id} - A_{\boldsymbol{x}}^g \rangle_\Psi \\
&\quad = \mathrm{E}_{\boldsymbol{x}} \sum_g \langle R R_{\boldsymbol{x}}^g R, \mathrm{Id} - A_{\boldsymbol{x}}^g \rangle_\Psi + \mathrm{E}_{\boldsymbol{x}} \sum_g \langle \Psi | (\mathrm{Id} - R) S_{\boldsymbol{x}}^g (\mathrm{Id} - R) \otimes (\mathrm{Id} - A_{\boldsymbol{x}}^g) | \Psi \rangle \\
&\quad \leq \mathrm{E}_{\boldsymbol{x}} \sum_g \langle \Psi | R_{\boldsymbol{x}}^g \otimes \mathrm{Id} - A_{\boldsymbol{x}}^g \otimes R | \Psi \rangle + 3\sqrt{\eta_2} \\
&\qquad + z^2 \mathrm{E}_{\boldsymbol{x}} \sum_g \langle \Phi | S^g \otimes (\mathrm{Id} - A^g) | \Phi \rangle + 2 \big\| |\tilde{\Phi}\rangle - \big((\mathrm{Id} - R) \otimes \mathrm{Id}^{(r-1)} \big) |\Psi\rangle \big\| \\
&\quad \leq \eta_2 + z^2 \eta_3 + 6r\sqrt{\eta_2}, \qquad (5.16)
\end{aligned}$$

where the second line uses projectivity of R for the first term, and the last inequality uses consistency of S with A for the second term and (5.8) for the third. Moreover, we can evaluate

$$\begin{aligned}
\langle T, \mathrm{Id} \rangle_\Psi &= \langle R^3, \mathrm{Id} \rangle_\Psi + \langle (\mathrm{Id} - R) S (\mathrm{Id} - R), \mathrm{Id} \rangle_\Psi \\
&\geq \langle R, \mathrm{Id} \rangle_\Psi - 3\sqrt{\eta_2} + z^2 \langle \Phi | S \otimes \mathrm{Id} \, | \Phi \rangle - 2 \big\| |\tilde{\Phi}\rangle - \big((\mathrm{Id} - R) \otimes \mathrm{Id}^{(r-1)} \big) |\Psi\rangle \big\| \\
&\geq \langle R, \mathrm{Id} \rangle_\Psi + z^2 (1 - \eta_3) - 6r\sqrt{\eta_2}, \qquad (5.17)
\end{aligned}$$

where in the second line we used projectivity of R, and for the last we used (5.8) as well as the completeness condition on S. Applying Lemma 5.4 to the $T_{\boldsymbol{x}}$, we deduce the existence of a family of submeasurements $V_{\boldsymbol{x}}$ that is η_2-consistent with $A_{\boldsymbol{x}}$ and projective, and such that

$$\begin{aligned}
\langle V, \mathrm{Id} \rangle_\Psi &\geq \langle T, \mathrm{Id} \rangle_\Psi - \big(\eta_2 + z^2 \eta_3 + 3r\sqrt{\eta_2} \big) - \eta_2 \\
&\geq \langle R, \mathrm{Id} \rangle_\Psi + z^2 (1 - 2\eta_3) - O\big(r\sqrt{\eta_2} \big),
\end{aligned}$$

where the first line uses (5.16) and the second (5.17). Comparing with our assumption that among all submeasurements that are η_2-consistent with A and projective R had the smallest completeness parameter, and using $\eta_3 \leq 1/4$, from (5.17) we get that $z^2 = O(r\sqrt{\eta_2})$. Using the bound $z^2 \geq 1 - \langle R, \mathrm{Id} \rangle_\Psi - 3r\sqrt{\eta_2}$ which follows from (5.9) we see that necessarily $1 - \langle R, \mathrm{Id} \rangle_\Psi = O(r\sqrt{\eta_2})$. Finally, to conclude we make R into a complete family of measurements by adding $\mathrm{Id} - R_{\boldsymbol{x}}$ to an arbitrary term $R_{\boldsymbol{x}}^g$, for every $\boldsymbol{x}$. ∎

6. Analysis of the low-degree test. In this section we prove Theorems 3.1 and 3.2. We first prove Theorem 3.1 in sections 6.1 to 6.5; the proof of Theorem 3.2 is given in section 6.6.

Fix $0 < \varepsilon \leq 1/2$, $d \geq 1$, $m \geq 2$, and $r \geq 3$. Let $(|\Psi\rangle, A, C)$ be a strategy with success $1 - \varepsilon$ in the $(d, m, r, \mathbb{F})$ low-degree test described in Figure 1. The test accepts in two cases. First, the referee automatically accepts if the two directions $\vec{y}_1, \vec{y}_2$ are not independent, which happens with probability at most $(1 + |\mathbb{F}|)/|\mathbb{F}|^m \leq 2/|\mathbb{F}| \leq \varepsilon$ given the assumption on $q = |\mathbb{F}|$ made in Theorem 3.1. Second, the referee accepts provided the players are consistent. Thus an overall acceptance probability of $1 - \varepsilon$ implies that the following must hold:

$$\mathrm{E}_{p \in \mathcal{S}_2(\mathbb{F}^m)} \mathrm{E}_{\boldsymbol{x} \in p} \sum_{h, a:\, a \neq h(\boldsymbol{x})} \langle A_{\boldsymbol{x}}^a, C_p^h \rangle_\Psi \leq 2\varepsilon, \qquad (6.1)$$

where $\langle \cdot, \cdot \rangle_\Psi$ is defined in section 2 (to which we refer for an introduction to the notation used in this section), the first expectation is taken over the choice of a uniformly random 2-dimensional affine subspace p of $\mathbb{F}^m$, and the second over a uniformly random point $x \in p$.

The proof of Theorem 3.1 is by induction on $m \geq 2$. For $m = 2$ there is a unique plane s in $\mathbb{F}^2$, hence the players have a unique "planes" measurement $\{C^g\}_{g \in \mathcal{P}_d(\mathbb{F}^2)}$. By setting $M^g := C^g$ for every g, (6.1) implies the conclusion of the theorem, provided C_1 is chosen to be at least 2.

We now assume that Theorem 3.1 is true for some dimension $(m-1) \geq 2$ and for every $\varepsilon, d, r, \mathbb{F}$ satisfying the assumptions of the theorem. We will prove the theorem for dimension m (and every $\varepsilon, d, r, \mathbb{F}$ satisfying the assumptions). The proof is divided into three steps. In the first step, carried out in section 6.3, we show that the induction hypothesis implies the existence of a family of measurements $\{Q_s^g\}$, defined for every $s \in \mathcal{S}_{m-1}(\mathbb{F}^m)$ and with outcomes $g \in \mathcal{P}_d(s)$ that are consistent with $\{A_{\boldsymbol{x}}^a\}$ (Lemma 6.8). In section 6.4 we show how the measurements Q_s can be combined together in a sequence of measurements $\{Q_{(s_i)}^{(h_i)}\}$, where $k \geq 1$ and $(s_i)_{1 \leq i \leq k}$ are parallel $(m-1)$-dimensional subspaces, with outcomes $h_i \in \mathcal{P}_d(s_i)$ (Lemma 6.11). Finally, in section 6.5 we show that if k is large enough these measurements can be transformed into a single measurement $\{M^h\}$ with outcomes $h \in \mathcal{P}_d(\mathbb{F}^m)$ that satisfies the conclusion of Theorem 3.1 (Claim 6.15). We begin by stating some useful direct consequences of (6.1).

6.1. The lines measurements. In this section we define a "lines" family of projective measurements $\{B_\ell^g\}_{g \in \mathcal{P}_d(\ell)}$, defined for every $\ell \in \mathcal{S}_1(\mathbb{F}^m)$, and we show that these measurements, together with the points and planes measurements, A and C, that form the players' strategy all enjoy good joint consistency properties.

LEMMA 6.1. *There exists a constant d_c such that the following holds. Let $(d+1)/|\mathbb{F}| < \varepsilon \leq 1/2$ and suppose that $\{A_{\boldsymbol{x}}^a\}$ and $\{C_p^h\}$ are projective measurements such that* (6.1) *holds. Then for every line $\ell \in \mathcal{S}_1(\mathbb{F}^m)$ there exists a projective measurement $\{B_\ell^g\}_{g \in \mathcal{P}_d(\ell)}$ such that, if for every $\boldsymbol{x} \in \mathbb{F}^m$ and $a \in \mathbb{F}$ we define*

$$B_{\boldsymbol{x}}^a := \mathrm{E}_{\ell \in \mathcal{S}_1(\mathbb{F}^m),\, \ell \ni \boldsymbol{x}} \sum_{g \in \mathcal{P}_d(\ell):\, g(\boldsymbol{x}) = a} B_\ell^g \qquad \text{and} \qquad C_{\boldsymbol{x}}^a$$
$$:= \mathrm{E}_{p \in \mathcal{S}_2(\mathbb{F}^m),\, p \ni \boldsymbol{x}} \sum_{h \in \mathcal{P}_d(p):\, h(\boldsymbol{x}) = a} C_p^h,$$

and for $\ell \in \mathcal{S}_1(\mathbb{F}^m)$ and $g \in \mathcal{P}_d(\ell)$,

$$C_\ell^g := \mathrm{E}_{p \in \mathcal{S}_2(\mathbb{F}^m),\, p \supset \ell} \sum_{h:\, h_{|\ell} = g} C_p^h,$$

then the following hold:

$$\max \left\{ \mathrm{E}_{\boldsymbol{x} \in \mathbb{F}^m} \sum_{a \in \mathbb{F}} \left\| B_{\boldsymbol{x}}^a - A_{\boldsymbol{x}}^a \right\|_\Psi^2, \mathrm{E}_{\boldsymbol{x} \in \mathbb{F}^m} \sum_{a \in \mathbb{F}} \left\| C_{\boldsymbol{x}}^a - A_{\boldsymbol{x}}^a \right\|_\Psi^2, \right.$$
$$\left. \mathrm{E}_{\ell \in \mathcal{S}_1(\mathbb{F}^m)} \sum_{g \in \mathcal{P}_d(\ell)} \left\| B_\ell^g - C_\ell^g \right\|_\Psi^2 \right\} = O\big(\varepsilon^{d_c}\big), \tag{6.2}$$

and

$$\max\Bigg\{\mathrm{E}_{\boldsymbol{x}\in\mathbb{F}^m}\sum_{\substack{a,b\in\mathbb{F}\\ a\neq b}}\langle A^a_{\boldsymbol{x}},A^b_{\boldsymbol{x}}\rangle_\Psi,\mathrm{E}_{\ell\in\mathcal{S}_1(\mathbb{F}^m)}\sum_{\substack{g,g'\in\mathcal{P}_d(\ell)\\ g\neq g'}}\langle B^g_\ell,B^{g'}_\ell\rangle_\Psi,$$

$$\mathrm{E}_{p\in\mathcal{S}_2(\mathbb{F}^m)}\sum_{\substack{h,h'\in\mathcal{P}_d(p)\\ h\neq h'}}\langle C^h_p,C^{h'}_p\rangle_\Psi\Bigg\}=O\big(\varepsilon^{d_c}\big). \tag{6.3}$$

We note that although all properties stated in the lemma follow from (6.1), we could have obtained them directly, and with a somewhat better dependence on ε, by assuming the existence of a strategy $(A,B,C,|\Psi\rangle)$ with success $1-\varepsilon$ in the following slight variant of the low-degree test from Figure 1. Let as usual $(d,m,\mathbb{F})$ be input parameters, and perform the following:

1. Choose a random $\boldsymbol{x}\in\mathbb{F}^m$ and two random directions $\vec{y}_1,\vec{y}_2\in\mathbb{F}^m$. Accept if the two vectors are not linearly independent. Otherwise, let p be the plane $(\boldsymbol{x};\vec{y}_1,\vec{y}_2)$ and ℓ the line $(\boldsymbol{x};\vec{y}_1)$.
2. Select two players at random and send them one out of the nine possible pairs of questions $(u,v)\in\{\boldsymbol{x},\ell,p\}\times\{\boldsymbol{x},\ell,p\}$, chosen uniformly at random.
3. Receive $g\in\mathcal{P}_d(u)$ from the first player and $h\in\mathcal{P}_d(v)$ from the second. Accept if and only if g and h are consistent on the intersection of their respective domains.

Nevertheless, we opt to work with the standard low-degree test as described in Figure 1, which is somewhat more concise. We first show the following claim, which establishes the part of the lemma that has to do with the measurements A and C only.

Claim 6.2. *Suppose that $\{A^a_{\boldsymbol{x}}\}$ and $\{C^h_p\}$ are projective measurements satisfying* (6.1)*. Then the following holds:*

$$\mathrm{E}_{\boldsymbol{x}\in\mathbb{F}^m}\sum_a\big\|A^a_{\boldsymbol{x}}-C^a_{\boldsymbol{x}}\big\|^2_\Psi=O\big(\sqrt{\varepsilon}\big). \tag{6.4}$$

Moreover, the families of measurements A and C are both self-consistent:

$$\mathrm{E}_{\boldsymbol{x}\in\mathbb{F}^m}\sum_{a\neq b}\langle A^a_{\boldsymbol{x}},A^b_{\boldsymbol{x}}\rangle_\Psi=O\big(\varepsilon^{1/4}\big), \tag{6.5}$$

$$\mathrm{E}_{p\in\mathcal{S}_2(\mathbb{F}^m)}\sum_{h\neq h'}\langle C^h_p,C^{h'}_p\rangle_\Psi=O\big(\varepsilon^{1/4}\big). \tag{6.6}$$

Proof. We first evaluate

$$\begin{aligned}\mathrm{E}_{\boldsymbol{x}\in\mathbb{F}^m}\sum_a\big\|A^a_{\boldsymbol{x}}\otimes\mathrm{Id}-\mathrm{Id}\otimes C^a_{\boldsymbol{x}}\big\|^2_\Psi&=2-2\,\mathrm{E}_{\boldsymbol{x}\in\mathbb{F}^m}\sum_a\langle A^a_{\boldsymbol{x}},C^a_{\boldsymbol{x}}\rangle_\Psi\\&=2-2\,\mathrm{E}_{\boldsymbol{x}}\mathrm{E}_{p\in\mathcal{S}_2(\mathbb{F}^m):\,p\ni\boldsymbol{x}}\sum_{h,a:\,a=h(\boldsymbol{x})}\langle A^a_{\boldsymbol{x}},C^h_p\rangle_\Psi\\&\leq 4\,\varepsilon,\end{aligned} \tag{6.7}$$

where for the first equality we used the assumption that A and C are projective, for

the second the definition of $C^a_{\boldsymbol{x}}$, and the last inequality follows from (6.1). Thus

$$\begin{aligned}
&\mathrm{E}_{\boldsymbol{x}\in\mathbb{F}^m} \sum_a \left\|A^a_{\boldsymbol{x}} - C^a_{\boldsymbol{x}}\right\|^2_\Psi \\
&\quad = 2 - 2\,\mathrm{E}_{\boldsymbol{x}} \sum_a \Re\big(\mathrm{Tr}_\rho(A^a_{\boldsymbol{x}} C^a_{\boldsymbol{x}})\big) \\
&\quad = 2 - 2\,\mathrm{E}_{\boldsymbol{x}} \sum_a \langle A^a_{\boldsymbol{x}}, A^a_{\boldsymbol{x}}\rangle_\Psi + 2\,\mathrm{E}_{\boldsymbol{x}} \sum_a \Re\big(\mathrm{Tr}_\rho\big((A^a_{\boldsymbol{x}}\otimes \mathrm{Id})(C^a_{\boldsymbol{x}}\otimes \mathrm{Id} - \mathrm{Id}\otimes A^a_{\boldsymbol{x}})\big)\big) \\
&\quad \le 2 - 2\,\mathrm{E}_{\boldsymbol{x}} \sum_a \langle A^a_{\boldsymbol{x}}, A^a_{\boldsymbol{x}}\rangle_\Psi + 2\sqrt{4\varepsilon} \\
&\quad = 2 - 2\,\mathrm{E}_{\boldsymbol{x}} \sum_a \langle A^a_{\boldsymbol{x}}, C^a_{\boldsymbol{x}}\rangle_\Psi + 2\,\mathrm{E}_{\boldsymbol{x}} \sum_a \Re\big(\mathrm{Tr}_\rho\big((A^a_{\boldsymbol{x}}\otimes \mathrm{Id})(A^a_{\boldsymbol{x}}\otimes \mathrm{Id} - \mathrm{Id}\otimes C^a_{\boldsymbol{x}})\big)\big) \\
&\qquad + 2\sqrt{4\varepsilon} \\
&\quad \le 4\varepsilon + 8\sqrt{\varepsilon},
\end{aligned}$$

where for the third and last lines we used the Cauchy–Schwarz inequality and (6.7), and in the fourth line we used that the number of players $r \ge 3$ and the permutation invariance of $|\Psi\rangle$. This proves (6.4). Consistency for A can be verified as

$$\begin{aligned}
\mathrm{E}_{x\in\mathbb{F}^m} \sum_{a\neq b} \langle A^a_{\boldsymbol{x}}, A^b_{\boldsymbol{x}}\rangle_\Psi &= \mathrm{E}_{x\in\mathbb{F}^m} \sum_{a\neq b} \langle A^a_{\boldsymbol{x}}, C^b_{\boldsymbol{x}}\rangle_\Psi + \mathrm{E}_{x\in\mathbb{F}^m} \sum_a \langle A^a_{\boldsymbol{x}}, A^a_{\boldsymbol{x}} - C^a_{\boldsymbol{x}}\rangle_\Psi \\
&\le 2\,\varepsilon + \left(\mathrm{E}_{x\in\mathbb{F}^m} \sum_a \left\|A^a_{\boldsymbol{x}}\right\|^2_\Psi\right)^{1/2} \left(\mathrm{E}_{x\in\mathbb{F}^m} \sum_a \left\|A^a_{\boldsymbol{x}} - C^a_{\boldsymbol{x}}\right\|^2_\Psi\right)^{1/2} \\
&= O\big(\varepsilon^{1/4}\big),
\end{aligned}$$

and a similar chain of inequalities proves consistency for C as well. ∎

We turn to the proof of the lemma.

Proof of Lemma 6.1. For every line $\ell \in \mathcal{S}_1(\mathbb{F}^m)$ and $g \in \mathcal{P}_d(\ell)$, define

$$\tilde{B}^g_\ell := C^g_\ell = \mathrm{E}_{p\in\mathcal{S}_2(\mathbb{F}^m),\, p\supset\ell} \sum_{h:h_{|\ell}=g} C^h_p. \tag{6.8}$$

We verify that the $\tilde{B}$ are self-consistent:

$$\begin{aligned}
\mathrm{E}_\ell \sum_{g\neq g'} \langle \tilde{B}^g_\ell, \tilde{B}^{g'}_\ell\rangle_\Psi &= \mathrm{E}_\ell \mathrm{E}_{p,p'\supset\ell} \sum_{h,h':h_{|\ell}\neq h'_{|\ell}} \langle C^h_p, C^{h'}_{p'}\rangle_\Psi \\
&= \mathrm{E}_\ell \mathrm{E}_{p,p'\supset\ell} \mathrm{E}_{\boldsymbol{x}\in\ell} \sum_{h,h':h_{|\ell}\neq h'_{|\ell}} \sum_{a,b} \langle C^h_p A^a_{\boldsymbol{x}}, C^{h'}_{p'} A^b_{\boldsymbol{x}}\rangle_\Psi \\
&= \mathrm{E}_\ell \mathrm{E}_{p,p'\supset\ell} \mathrm{E}_{\boldsymbol{x}\in\ell} \sum_{h,h':h_{|\ell}\neq h'_{|\ell}} \langle C^h_p A^{h(\boldsymbol{x})}_{\boldsymbol{x}}, C^{h'}_{p'} A^{h'(\boldsymbol{x})}_{\boldsymbol{x}}\rangle_\Psi + O\big(\varepsilon^{1/4}\big) \\
&= \mathrm{E}_\ell \mathrm{E}_{p,p'\supset\ell} \mathrm{E}_{\boldsymbol{x}\in\ell} \sum_{\substack{h,h':h_{|\ell}\neq h'_{|\ell} \\ h(\boldsymbol{x})=h'(\boldsymbol{x})}} \langle C^h_p A^{h(\boldsymbol{x})}_{\boldsymbol{x}}, C^{h'}_{p'} A^{h'(\boldsymbol{x})}_{\boldsymbol{x}}\rangle_\Psi + O\big(\varepsilon^{1/8}\big) \\
&= O\big(\varepsilon^{1/8}\big),
\end{aligned} \tag{6.9}$$

where the third line uses (6.1) and the fact that C is projective, the fourth (6.5) and the Cauchy–Schwarz inequality, and the last that two distinct degree-d polynomials

on ℓ intersect in a fraction at most $(1+d)/|\mathbb{F}| \leq \varepsilon$ of points given our assumption on $q = |\mathbb{F}|$. Applying Markov's inequality, we deduce from (6.9) that for all but a fraction at most $O(\varepsilon^{1/16})$ of lines ℓ, the measurement $\{B_\ell^g\}$ is self-consistent. For these ℓ we can apply the *orthonormalization lemma* from [KV11], which for convenience is restated as Lemma 6.4 below. The lemma guarantees the existence of a universal constant $c_p > 0$ and a family of projective measurements $\{B_\ell^g\}$ such that

$$\sum_g \left\| B_\ell^g - \tilde{B}_\ell^g \right\|_\Psi^2 = O\big(\varepsilon^{c_p/16}\big). \tag{6.10}$$

For the remaining ℓ we define $\{B_\ell^g\}$ arbitrarily (one of them is identity, the others 0). Together with (6.4), the definition of $\tilde{B}$, and the triangle inequality, (6.10) proves (6.2) provided d_c is chosen small enough. Given that self-consistency of A and C have already been proven in Claim 6.2, to conclude the proof of the lemma it remains to establish consistency of B, which follows immediately from (6.9), (6.10), and the Cauchy–Schwarz inequality. ∎

We also state the following useful consequence of Lemma 6.1.

Claim 6.3. *Under the same assumptions as in Lemma* 6.1, *for any family of submeasurements* $\{T_p^g\}_{g\in\mathcal{P}_d(p)}$, *defined for every plane* $p \in \mathcal{S}_2(\mathbb{F}^m)$, *it holds that*

$$\mathrm{E}_{p\in\mathcal{S}_2(\mathbb{F}^m)} \sum_{h\in\mathcal{P}_d(p)} \left\langle T_p^h, \left(\mathrm{E}_{\boldsymbol{x}\in p} A_{\boldsymbol{x}}^{h(\boldsymbol{x})} - C_p^h\right)^2 \right\rangle_\Psi = O\big(\varepsilon^{d_c}\big) \tag{6.11}$$

and

$$\mathrm{E}_{p\in\mathcal{S}_2(\mathbb{F}^m)} \sum_{h\in\mathcal{P}_d(p)} \left\langle T_p^h, \left(\mathrm{E}_{\ell\subset p} B_\ell^{h|_\ell} - C_p^h\right)^2 \right\rangle_\Psi = O\big(\varepsilon^{d_c/2}\big). \tag{6.12}$$

Proof. To show (6.11), expand the square and use

$$\begin{aligned}
&\mathrm{E}_{p\in\mathcal{S}_2(\mathbb{F}^m)} \sum_{h\in\mathcal{P}_d(p)} \mathrm{E}_{\boldsymbol{x}\in p} \big\langle T_p^h, A_{\boldsymbol{x}}^{h(\boldsymbol{x})} C_p^h \big\rangle_\Psi \\
&\quad\approx_{\varepsilon^{d_c}} \mathrm{E}_{p\in\mathcal{S}_2(\mathbb{F}^m)} \sum_{h\in\mathcal{P}_d(p)} \mathrm{E}_{\boldsymbol{x}\in p} \big\langle T_p^h C_p^h, A_{\boldsymbol{x}}^{h(\boldsymbol{x})} C_p^h \big\rangle_\Psi \\
&\quad\approx_{\varepsilon^{d_c}} \mathrm{E}_{p\in\mathcal{S}_2(\mathbb{F}^m)} \sum_{h\in\mathcal{P}_d(p)} \mathrm{E}_{\boldsymbol{x}\in p} \big\langle T_p^h C_p^h, A_{\boldsymbol{x}}^{h(\boldsymbol{x})} \big\rangle_\Psi \\
&\quad\approx_{\varepsilon} \mathrm{E}_{\substack{p\in\mathcal{S}_2(\mathbb{F}^m)\\ x\in p}} \sum_{h\in\mathcal{P}_d(p)} \big\langle T_p^h, A_{\boldsymbol{x}}^{h(\boldsymbol{x})} \big\rangle_\Psi,
\end{aligned}$$

where the first and second lines follow from (6.3) and the third from (6.1). Similar bounds for the other terms appearing in the expansion of the square suffice to prove (6.11). To show (6.12) one proceeds in the same way, using the Cauchy–Schwarz inequality and (6.2) and then (6.3) in lieu of (6.11) to perform the third step above. ∎

We end with a statement of the orthonormalization lemma, a slightly simplified version of [KV10, Lemma 23].

Lemma 6.4 (orthonormalization lemma). *Let* $\{A_i\}$ *be a measurement and* $|\Psi\rangle$ *a permutation-invariant state such that*

$$\sum_i \langle A_i, A_i\rangle_\Psi \geq 1 - \varepsilon.$$

Then there exists a projective measurement $\{B_i\}$ such that

$$\sum_i \big\|A_i - B_i\big\|_\Psi^2 = O(\varepsilon^{c_p}),$$

where $c_p > 0$ is a universal constant.

6.2. Robust triples. The proof of the induction step $(m-1) \to m$ requires successive applications of the consolidation proposition, Proposition 5.8. For this we will use different "robust triples" (see Definition 5.2), which are defined in this section.

Let $1 \le k \le m$, $s \in \mathcal{S}_k(\mathbb{F}^m)$, G_s the complete graph on the vertex set defined by the points in s, and $\mathcal{G}_s = \mathcal{P}_d(s)$. Let $\mathcal{T}_s = (G_s, \{B^a_{\boldsymbol{z}}\}, \mathcal{G}_s)$, where the measurements $\{B^a_{\boldsymbol{z}}\}$ are as defined in Lemma 6.1.

CLAIM 6.5. *The triple $\mathcal{T}_s$ is $(O(\varepsilon^{d_c/4}), 1/2)$-robust for all but a fraction at most $O(\varepsilon^{d_c/4})$ of $s \in \mathcal{S}_k(\mathbb{F}^m)$, where $d_c > 0$ is the constant from Lemma* 6.1.

Proof. Using (6.2) and (6.3) we have

$$\begin{aligned}\mathrm{E}_{s\in\mathcal{S}_k(\mathbb{F}^m)}\mathrm{E}_{\boldsymbol{z}\in s}\sum_{a\in\mathbb{F}}\big\langle B^a_{\boldsymbol{z}}, \mathrm{Id} - B^a_{\boldsymbol{z}}\big\rangle_\Psi &\approx_{\varepsilon^{d_c/2}} \mathrm{E}_{s\in\mathcal{S}_k(\mathbb{F}^m)}\mathrm{E}_{\boldsymbol{z}\in s}\sum_{a\in\mathbb{F}}\big\langle A^a_{\boldsymbol{z}}, \mathrm{Id} - A^a_{\boldsymbol{z}}\big\rangle_\Psi \\ &= O\big(\varepsilon^{d_c}\big).\end{aligned}$$

Applying Markov's inequality, the measurements B are $O(\varepsilon^{d_c/4})$ self-consistent for all but a fraction at most $O(\varepsilon^{d_c/4})$ of subspaces s. For any $s \in \mathcal{S}_k(\mathbb{F}^m)$ let $\{R^g_s\}$ be an arbitrary submeasurement. We have

$$\begin{aligned}&\mathrm{E}_{s\in\mathcal{S}_k(\mathbb{F}^m)}\mathrm{E}_{\boldsymbol{z},\boldsymbol{z}'\in s}\sum_{g\in\mathcal{P}_d(s)}\big\langle R^g_s, \big(B^{g(\boldsymbol{z})}_{\boldsymbol{z}} - B^{g(\boldsymbol{z}')}_{\boldsymbol{z}'}\big)^2\big\rangle_\Psi \\ &\quad\approx_{q^{-1}} 2\,\mathrm{E}_{s\in\mathcal{S}_k(\mathbb{F}^m)}\mathrm{E}_{\boldsymbol{z}\in s}\sum_{g\in\mathcal{P}_d(s)}\big\langle R^g_s, \big(B^{g(\boldsymbol{z})}_{\boldsymbol{z}}\big)^2\big\rangle_\Psi \\ &\quad\quad - 2\Re\Bigg(\mathrm{E}_{\substack{s\in\mathcal{S}_k(\mathbb{F}^m)\\ p\in\mathcal{S}_2(s),\,\boldsymbol{z},\boldsymbol{z}'\in p}}\sum_{g\in\mathcal{P}_d(s)}\big\langle R^g_s, B^{g(\boldsymbol{z})}_{\boldsymbol{z}} B^{g(\boldsymbol{z}')}_{\boldsymbol{z}'}\big\rangle_\Psi\Bigg),\end{aligned} \tag{6.13}$$

where we used that two uniformly distributed $\boldsymbol{z}, \boldsymbol{z}' \in s$ can equivalently (up to an error in statistical distance of $O(1/|\mathbb{F}|)$) be sampled by first choosing a uniformly random plane $p \subset s$ and then two uniform points $\boldsymbol{z}, \boldsymbol{z}' \in p$. To estimate the second term above, write

$$\begin{aligned}&\mathrm{E}_{\substack{s\in\mathcal{S}_k(\mathbb{F}^m)\\ p\in\mathcal{S}_2(s),\,\boldsymbol{z},\boldsymbol{z}'\in p}}\sum_{g\in\mathcal{P}_d(s)}\big\langle R^g_s, B^{g(\boldsymbol{z})}_{\boldsymbol{z}} B^{g(\boldsymbol{z}')}_{\boldsymbol{z}'}\big\rangle_\Psi \\ &\quad\approx_{\varepsilon^{d_c/2}} \mathrm{E}_{\substack{s\in\mathcal{S}_k(\mathbb{F}^m)\\ p\in\mathcal{S}_2(s),\,\boldsymbol{z},\boldsymbol{z}'\in p}}\sum_{g\in\mathcal{P}_d(s)}\big\langle R^g_s, A^{g(\boldsymbol{z})}_{\boldsymbol{z}} A^{g(\boldsymbol{z}')}_{\boldsymbol{z}'}\big\rangle_\Psi \\ &\quad\approx_{\varepsilon^{d_c/2}} \mathrm{E}_{\substack{s\in\mathcal{S}_k(\mathbb{F}^m)\\ p\in\mathcal{S}_2(s)}}\sum_{g\in\mathcal{P}_d(s)}\big\langle R^g_s, C^g_p C^g_p\big\rangle_\Psi \\ &\quad\approx_{\varepsilon^{d_c/2}} \mathrm{E}_{\substack{s\in\mathcal{S}_k(\mathbb{F}^m)\\ p\in\mathcal{S}_2(s),\,\boldsymbol{z}\in p}}\sum_{g\in\mathcal{P}_d(s)}\big\langle R^g_s, B^{g(\boldsymbol{z})}_{\boldsymbol{z}}\big\rangle_\Psi,\end{aligned}$$

where the first line follows from the Cauchy–Schwarz inequality and (6.2), the second from (6.11), and the last uses pojectivity of the $\{C^g_p\}$ and (6.2). Together with (6.13),

we obtain

$$\mathrm{E}_{s\in\mathcal{S}_k(\mathbb{F}^m)}\mathrm{E}_{\boldsymbol{z},\boldsymbol{z}'\in s}\sum_{g\in\mathcal{P}_d(s)}\big\langle R_s^g,\big(B_{\boldsymbol{z}}^{g(\boldsymbol{z})}-B_{\boldsymbol{z}'}^{g(\boldsymbol{z}')}\big)^2\big\rangle_\Psi \;=\; O\big(\varepsilon^{d_c/2}+q^{-1}\big).$$

Applying Markov's inequality and assuming the constant d_1 from Theorem 3.1 is chosen small enough that $dq^{-1}\le\varepsilon^{d_c/2}$, we have thus shown that the measurements B are $O(\varepsilon^{d_c/4})$-stable for all but a fraction at most $O(\varepsilon^{d_c/4})$ of subspaces s. Finally, the small intersection property required in the definition of a robust triple follows from the Schwarz–Zippel lemma, and the expansion property trivially holds for the complete graph G_s. ∎

We generalize the previous construction to tuples of k parallel subspaces $s_i\in\mathcal{S}_{m-1}(\mathbb{F}^m)$. For any $\boldsymbol{z}\in\mathbb{F}^m$ and $a\in\mathbb{F}$ let $X_{\boldsymbol{z}}^a:=B_{\boldsymbol{z}}^a$, and for any $k\ge 2$, k-tuple of aligned points $\boldsymbol{z}_1,\dots,\boldsymbol{z}_k\in\mathbb{F}^m$, and any $a_1,\dots,a_k\in\mathbb{F}$ let

$$X_{(\boldsymbol{z}_i)}^{(a_i)}:=\sum_{\substack{g\in\mathcal{P}_d(\ell(z_i))\\ \forall i,\, g(\boldsymbol{z}_i)=a_i}} B_{\ell(\boldsymbol{z}_i)}^g, \tag{6.14}$$

where $\ell(\boldsymbol{z}_i)$ is the line going through the $\boldsymbol{z}_i$. Let $V_{(s_i)}$ be the set of all k-tuples of aligned points $(\boldsymbol{z}_1,\dots,\boldsymbol{z}_k)$, where $\boldsymbol{z}_i\in s_i$, and $G_{(s_i)}$ the complete graph on V. Let $\mathcal{G}_{(s_i)}\subseteq\{g:V_{(s_i)}\to\mathbb{F}^k\}$ be the set of all k-tuples of degree-d $(m-1)$-variate polynomials (g_i), where $g_i\in\mathcal{P}_d(s_i)$. Finally, let $\mathcal{T}_{(s_i)}=(G_{(s_i)},\{X_{(\boldsymbol{z}_i)}^{(a_i)}\},\mathcal{G}_{(s_i)})$.

LEMMA 6.6. *Let $2\le k\le 2d$. The triple $\mathcal{T}_{(s_i)}$ is $(O(\varepsilon^{d_c/4}),1/2)$-robust for all but a fraction at most $O(\varepsilon^{d_c/4})$ of k-tuples $(s_i)\in(\mathcal{S}_{m-1}(\mathbb{F}^m))^k$.*

Proof. We verify the four properties needed of a robust triple. Expansion is clear, since the graph is complete. The property of small intersection follows from the Schwartz–Zippel lemma. Next we verify self-consistency:

$$\begin{aligned}
&\mathrm{E}_{(s_i),(\boldsymbol{z}_i)\in(s_i)}\sum_{(a_i)\neq(a_i')}\big\langle X_{(\boldsymbol{z}_i)}^{(a_i)},X_{(\boldsymbol{z}_i)}^{(a_i')}\big\rangle_\Psi\\
&\quad=\mathrm{E}_{(s_i),(\boldsymbol{z}_i)\in(s_i)}\sum_{(a_i)\neq(a_i')}\sum_{g,g':\forall i,g(\boldsymbol{z}_i)=a_i,g'(\boldsymbol{z}_i)=a_i'}\big\langle B_{\ell(\boldsymbol{z}_i)}^g,B_{\ell(\boldsymbol{z}_i)}^{g'}\big\rangle_\Psi\\
&\quad\le\mathrm{E}_{(s_i),(\boldsymbol{z}_i)\in(s_i)}\sum_{g\neq g'}\big\langle B_{\ell(\boldsymbol{z}_i)}^g,B_{\ell(\boldsymbol{z}_i)}^{g'}\big\rangle_\Psi\\
&\quad=O(\varepsilon^{d_c}),
\end{aligned}\tag{6.15}$$

where the last equality follows from (6.3). It remains to prove stability. For any k-tuple (s_i), let $\{R_{(s_i)}^g\}$ be an arbitrary submeasurement with outcomes $g=(g_i)\in\mathcal{G}_{(s_i)}$. We abuse notation and also use g to designate the unique polynomial of degree $(d+k-1)$ defined on the whole of $\mathbb{F}^m$ that has degree at most d when restricted to each s_i, and at most $(k-1)$ when restricted to any line $\ell(\boldsymbol{z}_i)$, where $\boldsymbol{z}_i\in s_i$. Such a polynomial can be obtained by interpolation from the g_i. (Uniqueness follows since equality on every line $\ell(\boldsymbol{z}_i)$ implies equality on $\mathbb{F}^m$.) For simplicity of notation, we

also write R^g and omit the expectation over (s_i). We have

$$\begin{aligned}
&\mathrm{E}_{(\boldsymbol{z}_i),(\boldsymbol{z}'_i)\in(s_i)} \sum_g \big\langle R^g, \big(X^{(g(\boldsymbol{z}_i))}_{(\boldsymbol{z}_i)} - X^{(g(\boldsymbol{z}'_i))}_{(\boldsymbol{z}'_i)}\big)^2\big\rangle_\Psi \\
&\quad \le \mathrm{E}_{(\boldsymbol{z}_i),(\boldsymbol{z}'_i)\in(s_i)} \sum_g \big\langle R^g, \big(B^{g_{|\ell(\boldsymbol{z}_i)}}_{\ell(\boldsymbol{z}_i)} - B^{g_{|\ell(\boldsymbol{z}'_i)}}_{\ell(\boldsymbol{z}'_i)}\big)^2\big\rangle_\Psi \\
&\qquad + 2\,\mathrm{E}_{(\boldsymbol{z}_i)} \sum_g \sum_{\substack{h\in\mathcal{P}_d(\ell(\boldsymbol{z}_i)),\, h\neq g_{|\ell} \\ h(\boldsymbol{z}_i)=g(\boldsymbol{z}_i)}} \big\langle R^g, B^h_{\ell(\boldsymbol{z}_i)}\big\rangle_\Psi \\
&\quad \le \mathrm{E}_{\ell,\ell'} \sum_g \big\langle R^g, \big(B^{g_{|\ell}}_\ell - B^{g_{|\ell'}}_{\ell'}\big)^2\big\rangle_\Psi \\
&\qquad + 2\,\mathrm{E}_\ell \sum_{g,h:\, h\neq g_{|\ell}} \Pr_{(\boldsymbol{z}_i)\in\ell}\big(\forall i, h(\boldsymbol{z}_i)=g(\boldsymbol{z}_i)\big)\big\langle R^g_{(s_i)}, B^h_\ell\big\rangle_\Psi \\
&\quad \le \mathrm{E}_{\ell,\ell'} \sum_g \big\langle R^g, \big(B^{g_{|\ell}}_\ell - B^{g_{|\ell'}}_{\ell'}\big)^2\big\rangle_\Psi + O\big(dkq^{-1}\big), \qquad (6.16)
\end{aligned}$$

where for the first inequality we used the definition of X and orthogonality of the B^h_ℓ to separate out those terms for which $h = g_{|\ell}$ and $h \neq g_{|\ell}$ (but still $h(\boldsymbol{z}_i) = g(\boldsymbol{z}_i)$ for every i), and for the last we used the Schwartz–Zippel lemma. This last term can then be bounded exactly as the analogue term was bounded to establish the stability property in the proof of Claim 6.5, using (6.12) instead of (6.11). This establishes the stability property for the X measurements, on average over the choice of the k-tuple (s_i). Applying Markov's inequality proves the lemma (provided the constant d_1 from Theorem 3.1 is chosen small enough that $2d^2q^{-1} \le \varepsilon^{c_d/2}$). ∎

The following claim establishes consistency of X with A.

Claim 6.7. *For any $k \ge 1$ and $(\boldsymbol{z}_1, \dots, \boldsymbol{z}_k) \in (\mathbb{F}^m)^k$ the $\{X^{(a_i)}_{(\boldsymbol{z}_i)}\}_{(a_i)\in\mathbb{F}^k}$ form a projective measurement. Moreover, for any $1 \le j \le k$ we have*

$$\mathrm{E}_{(\boldsymbol{z}_i)\in(\mathbb{F}^m)^k} \sum_{(a_i)\in\mathbb{F}^k} \sum_{a:\, a\neq a_j} \langle X^{(a_i)}_{(\boldsymbol{z}_i)}, A^a_{\boldsymbol{z}_j}\rangle_\Psi \;=\; O\big(\varepsilon^{d_c/2}\big).$$

Proof. If $k = 1$ the claim is immediate by definition of $X^{a_1}_{\boldsymbol{z}_1}$ and (6.2), (6.3). If $k \ge 2$,

$$\begin{aligned}
\mathrm{E}_{(\boldsymbol{z}_i)} \sum_{(a_i)} \sum_{a\neq a_j} \big\langle X^{(a_i)}_{(\boldsymbol{z}_i)}, A^a_{\boldsymbol{z}_j}\big\rangle_\Psi &= \mathrm{E}_{(\boldsymbol{z}_i)} \sum_{(a_i)} \sum_{a\neq a_j} \sum_{g:\forall i,\, g(\boldsymbol{z}_i)=a_i} \big\langle B^g_{\ell(\boldsymbol{z}_i)}, A^a_{\boldsymbol{z}_j}\big\rangle_\Psi \\
&\le \mathrm{E}_\ell \mathrm{E}_{\boldsymbol{z}_j\in\ell} \sum_{g,a:\, a\neq g(\boldsymbol{z}_j)} \big\langle B^g_\ell, A^a_{\boldsymbol{z}_j}\big\rangle_\Psi \\
&= O\big(\varepsilon^{d_c/2}\big).
\end{aligned}$$

Here the inequality follows simply by ignoring the constraint that $g(z_i) = a_i$ for all indices but $i = j$, and the last follows from the case $k = 1$. ∎

6.3. The measurements Q_s. Recall that we set to prove Theorem 3.1 by induction on m. Our first step is to prove that the induction hypothesis can be used to deduce the existence of a family of measurements $\{Q_s\}$ parameterized by $(m-1)$-

dimensional subspaces s of $\mathbb{F}^m$ that are consistent with the measurements $\{A_{\boldsymbol{z}}\}$ coming from the players' strategy.

LEMMA 6.8. *There exists a universal constant $0 < c_\ell \leq d_c/400$ such that the following holds. Under the assumptions of Theorem* 3.1, *for every $s \in \mathcal{S}_{m-1}(\mathbb{F}^m)$ there exists a measurement $\{Q_s^g\}_{g\in\mathcal{P}_d(s)}$ such that*

$$\mathrm{E}_{s\in\mathcal{S}_{m-1}(\mathbb{F}^m)}\mathrm{E}_{\boldsymbol{x}\in s}\sum_{g,a:\, a\neq g(\boldsymbol{x})}\langle Q_s^g, A_{\boldsymbol{x}}^a\rangle_\Psi \;=\; O\big(\varepsilon^{c_\ell}\big).$$

Fix an $s \in \mathcal{S}_{m-1}(\mathbb{F}^m)$. We use the term s-restricted $(d, m-1, \mathbb{F})$ low-degree test to refer to the variant of the low-degree test in which the referee chooses $(z, \vec{y}_1, \vec{y}_2)$ uniformly in s, and then proceeds as in the usual test. We claim the following.

CLAIM 6.9. *For a fraction at least $1-O(\varepsilon^{-2}q^{-1})$ of $(m-1)$-dimensional subspaces s, the strategy $(|\Psi\rangle, A, C)$ has success at least $1-3\varepsilon$ in the s-restricted $(d, m-1, \mathbb{F})$ low-degree test.*

Proof. In the $(d, m, \mathbb{F})$ low-degree test over $\mathbb{F}^m$, the referee picks a triple $(\boldsymbol{z}, \vec{y}_1, \vec{y}_2)$ uniformly at random in $\mathbb{F}^m$, automatically accepts if $(\vec{y}_1, \vec{y}_2)$ are not linearly independent, and proceeds if they are; conditioned on not accepting immediately the resulting plane $p = (\boldsymbol{z}; \vec{y}_1, \vec{y}_2)$ is uniformly distributed in $\mathcal{S}_2(\mathbb{F}^m)$.

Consider now a referee who selects random $(\boldsymbol{z}, \vec{y}_1, \ldots, \vec{y}_{m-1})$ in $\mathbb{F}^m$, automatically accepts if they are not linearly independent, and proceeds with the plane $p = (\boldsymbol{z}; \vec{y}_1, \vec{y}_2)$ if they are. Conditioned on the referee not accepting immediately, the subspace $s = (\boldsymbol{z}; \vec{y}_1, \ldots, \vec{y}_{m-1})$ is uniformly distributed in $\mathcal{S}_{m-1}(\mathbb{F}^m)$, and p is uniformly distributed in $\mathcal{S}_2(s)$. Hence the only difference between the two scenarios is in the probability of accepting immediately. In both cases, it follows from [AS97, Lemma 10] that this probability is upper bounded by $2/q$ (but it is higher in the second scenario). In particular, the players' success probability in the second scenario, conditioned on not having accepted immediately, is at least $1-\varepsilon-2/q$, so that

$$\mathrm{E}_s\left[1-\varepsilon_s\right] \;\geq\; 1-\varepsilon-\frac{2}{q},$$

where ε_s is the players' success in the s-restricted $(d, m-1, \mathbb{F})$ low-degree test. To conclude it suffices to perform a variance analysis exactly as in [AS97, Lemma 12]. One then obtains that, for any $\alpha > 0$,

$$\Pr_s\big(1-\varepsilon_s \;\leq\; (1-\alpha)(1-\varepsilon-2/q)\big) \;\leq\; O\big(\alpha^{-2}q^{-1}\big).$$

Choosing $\alpha = \varepsilon$, provided q is large enough with respect to ε^{-1}, the claim is proved. ∎

We turn to the proof of the lemma.

Proof of Lemma 6.8. Let $s \in \mathcal{S}_{m-1}(\mathbb{F}^m)$. If the strategy $(|\Psi\rangle, A, C)$ has success at least $1-3\varepsilon$ in the s-restricted $(d, m-1, \mathbb{F})$ low-degree test, by applying the induction hypothesis Theorem 3.1 implies the existence of a measurement $\{Q_s^g\}$ which satisfies

$$\mathrm{E}_{\boldsymbol{x}\in s}\sum_{a\in\mathbb{F}}\sum_{g:\, g(\boldsymbol{x})\neq a}\left\langle A_{\boldsymbol{x}}^a, Q_s^g\right\rangle_\Psi \;\leq\; C_1(3\varepsilon)^{c_1}.$$

Using (6.2) and the Cauchy–Schwarz inequality we also obtain that $\{Q_s^g\}$ is δ-consistent with $\{B_{\boldsymbol{z}}^a\}$ for some $\delta = O(\varepsilon^{c_1} + \varepsilon^{d_c/2})$. Define the Q_s arbitrarily for the remaining subspaces s. Let $\mathcal{S}$ be the set of all s for which the triple $\mathcal{T}_s$ defined in Claim 6.5

is $(O(\varepsilon^{d_c/4}), 1/2)$-robust. Claim 6.5 shows that a fraction at least $1 - O(\varepsilon^{d_c/4})$ of s (assuming $d_c \leq 1$) are in $\mathcal{S}$. Using Claim 6.9, we thus get

$$\mathrm{E}_{s\in\mathcal{S}} \sum_{a\in\mathbb{F}} \sum_{g:\, g(\boldsymbol{x})\neq a} \big\langle B^a_{\boldsymbol{x}}, Q^g_s \big\rangle_\Psi \;\leq\; 2\delta,$$

provided the constant d_1 from Theorem 3.1 is chosen large enough.

We are in a position to apply Proposition 5.8, with the set X there being $\mathcal{S}$ here. The proposition shows that, provided ε is small enough, for every $s \in \mathcal{S}$ there exists an "improved" measurement, such that on average over $s \in \mathcal{S}$ the $\{Q^g_s\}$ are $O((\eta'(\varepsilon^{d_c/4}, 1/2))^{1/2})$-consistent with the $\{B^a_{\boldsymbol{x}}\}$, where η' is as in Lemma 5.4. For the remaining subspaces s we define $\{Q^g_s\}$ arbitrarily. Using (6.3) to relate consistency with B to consistency with A, the lemma is proved provided c_ℓ is chosen small enough ($c_\ell = d_c/40$ suffices). ∎

As a corollary of Lemma 6.8, the following claim shows that the measurements $\{Q_s\}$ are self-consistent.

Claim 6.10. *The measurements $\{Q^g_s\}_{g\in\mathcal{P}_d(s)}$ satisfy*

$$\mathrm{E}_{s\in\mathcal{S}_{m-1}(\mathbb{F}^m)} \sum_{g\in\mathcal{P}_d(s)} \langle Q^g_s, (\mathrm{Id} - Q^g_s)\rangle_\Psi \;=\; O\big(\varepsilon^{c_\ell}\big),$$

where c_ℓ is as defined in Lemma 6.8.

Proof. We can write

$$\begin{aligned}
&\mathrm{E}_{s\in\mathcal{S}_{m-1}(\mathbb{F}^m)} \sum_{g,g'\in\mathcal{P}_d(s),\, g\neq g'} \langle Q^g_s, Q^{g'}_s\rangle_\Psi \\
&\qquad \approx_{\varepsilon^{c_\ell}} \mathrm{E}_{s\in\mathcal{S}_{m-1}(\mathbb{F}^m)} \mathrm{E}_{\boldsymbol{x}\in s} \sum_{g,g'\in\mathcal{P}_d(s),\, g\neq g'} \mathrm{Tr}_\rho\big(Q^g_s \otimes Q^{g'}_s \otimes A^{g(\boldsymbol{x})}_{\boldsymbol{x}}\big) \\
&\qquad = O\big(\varepsilon^{c_\ell}\big),
\end{aligned}$$

where both lines follow from consistency of Q with A. ∎

6.4. Pasting the Q_s. In this section we combine the k measurements $\{Q^h_{s_i}\}_{h\in\mathcal{P}_d(s_i)}$, $1 \leq i \leq k$, obtained from Lemma 6.8 for any k-tuples of parallel subspaces $s_i \in \mathcal{S}_{m-1}(\mathbb{F}^m)$ into a single measurement $\{Q^{(h_i)}_{(s_i)}\}_{h_i\in\mathcal{P}_d(s_i)}$. Let $\{X^{(a_i)}_{(z_i)}\}_{a_i\in\mathbb{F}}$ be the measurements defined in (6.14) for every k-tuple of aligned points $\boldsymbol{z}_i \in \mathbb{F}^m$.

Lemma 6.11. *For any $k \geq 1$, $(m-1)$-tuple of linearly independent directions $\vec{y}_1, \ldots, \vec{y}_{m-1} \in \mathbb{F}^m$ and k-tuple of aligned points $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k \in \mathbb{F}^m$, letting $s_i = (z_i; \vec{y}_1, \ldots, \vec{y}_{m-1})$ there exists a family of measurements $\{Q^{(h_i)}_{(s_i)}\}_{(h_i)\in(\mathcal{P}_d(s_i))}$ such that*

$$\mathrm{E}_{\vec{y}_i, \boldsymbol{z}_i} \sum_{\substack{h_i\in\mathcal{P}_d(s_i),\, a_i\in\mathbb{F} \\ \exists i, a_i\neq h_i(\boldsymbol{z}_i)}} \big\langle Q^{(h_i)}_{(s_i)}, X^{(a_i)}_{(\boldsymbol{z}_i)} \big\rangle_\Psi \;=\; O\big(\varepsilon^{c_\ell}\big), \tag{6.17}$$

where c_ℓ is the constant defined in Lemma 6.8.

We first prove the case $k = 1$, which follows almost immediately from Lemma 6.8.

Claim 6.12. *The following holds:*

$$\mathrm{E}_{s\in\mathcal{S}_{m-1}(\mathbb{F}^m)} \mathrm{E}_{\boldsymbol{z}\in s} \sum_{h,a:\, a\neq h(\boldsymbol{z})} \big\langle Q^h_s, X^a_{\boldsymbol{z}} \big\rangle_\Psi \;=\; O\big(\varepsilon^{c_\ell}\big).$$

Proof. By definition,

$$\begin{aligned}
&\mathrm{E}_{s\in\mathcal{S}_{m-1}(\mathbb{F}^m)}\mathrm{E}_{\boldsymbol{z}\in s}\sum_{h,a:\,a\neq h(\boldsymbol{z})}\big\langle Q_s^h, X_{\boldsymbol{z}}^a\big\rangle_\Psi\\
&\quad=\mathrm{E}_{s,\boldsymbol{z}}\sum_{h,a:\,a\neq h(\boldsymbol{z})}\big\langle Q_s^h, A_{\boldsymbol{z}}^a\big\rangle_\Psi+\mathrm{E}_{s,\boldsymbol{z}}\sum_{h,a:\,a\neq h(\boldsymbol{z})}\big\langle Q_s^h, (B_{\boldsymbol{z}}^a-A_{\boldsymbol{z}}^a)\big\rangle_\Psi\\
&\quad\le O(\varepsilon^{c_\ell})+\left(\mathrm{E}_{\boldsymbol{z}}\sum_a\big\|B_{\boldsymbol{z}}^a-A_{\boldsymbol{z}}^a\big\|_\Psi^2\right)^{1/2}\\
&\quad=O(\varepsilon^{c_\ell}),
\end{aligned}$$

where the inequality uses Lemma 6.8 to bound the first term and Cauchy–Schwarz for the second, and the last follows from (6.2) and $c_\ell \le d_c/2$. ∎

We will use the following general "pasting" lemma.

LEMMA 6.13. *Let S_1, S_2 be two disjoint sets and $\{Q^g\}$ and $\{R^h\}$ measurements with outcomes in $\mathcal{G}\subseteq\{g: S_1\to\mathbb{F}\}$ and $\mathcal{H}\subseteq\{h: S_2\to\mathbb{F}\}$, respectively. Suppose that Q and R are each consistent with a family of measurements $\{A_v^a\}_{a\in\mathbb{F}}$ defined for every $v\in S_1\cup S_2$:*

$$\max\left\{\mathrm{E}_{v\in S_1}\sum_{g,a:\,a\neq g(v)}\big\langle Q^g, A_v^a\big\rangle_\Psi,\ \mathrm{E}_{v\in S_2}\sum_{h,a:\,a\neq h(v)}\big\langle R^h, A_v^a\big\rangle_\Psi\right\}\le\delta$$

for some $\delta>0$, and that Q is δ-self-consistent. Then there exists a "pasted" measurement $\{T^f\}$, where $f=(f_1,f_2)$ and $f_1: S_1\to\mathbb{F}$, $f_2: S_2\to\mathbb{F}$, such that T is consistent with A:

$$\mathrm{E}_{v\in S_1\cup S_2}\sum_{f,a:\,a\neq f_1(v)}\big\langle T^f, A_v^a\big\rangle_\Psi=O\big(\sqrt{\delta}\big).$$

Proof. For any $f=(f_1,f_2)$ let $T^f:=\sqrt{Q^{f_1}}R^{f_2}\sqrt{Q^{f_1}}$. Then $\{T^f\}$ is a measurement, and

$$\begin{aligned}
&\mathrm{E}_{v\in S_1\cup S_2}\sum_{f,a:\,a\neq f(v)}\big\langle T^f, A_v^a\big\rangle_\Psi\\
&\quad=\frac{1}{|S_1|+|S_2|}\sum_{v\in S_1\cup S_2}\sum_{f,a:\,a\neq f(v)}\big\langle\sqrt{Q^{f_1}}R^{f_2}\sqrt{Q^{f_1}}, A_v^a\big\rangle_\Psi\\
&\quad\le\frac{1}{|S_1|+|S_2|}\sum_{v\in S_2}\sum_{f,a:\,a\neq f_2(v)}\big\langle\sqrt{Q^{f_1}}R^{f_2}\sqrt{Q^{f_1}}, A_v^a\big\rangle_\Psi+\frac{\delta|S_1|}{|S_1|+|S_2|}\\
&\quad\approx_{\sqrt{\delta}}\frac{1}{|S_1|+|S_2|}\sum_{v\in S_2}\sum_{f,a:\,a\neq f_2(v)}\big\langle R^{f_2}, A_v^aQ^{f_1}\big\rangle_\Psi+\frac{\delta|S_1|}{|S_1|+|S_2|}\\
&\quad=O\big(\sqrt{\delta}\big),
\end{aligned}$$

where the first inequality uses consistency of Q with A, the second uses self-consistency of Q, and the last consistency of R with A. ∎

We turn to the proof of Lemma 6.11.

Proof of Lemma 6.11. The proof of the lemma is by induction on k. The case $k=1$ was proved in Claim 6.12. Assume the lemma true for $k-1$, and for any

$(k-1)$-tuple $(s_i) \in (\mathcal{S}_{m-1}(\mathbb{F}^m))^{k-1}$ of parallel subspaces let $\{Q^g_{s_1\cup\cdots\cup s_{k-1}}\}_g$ be the resulting family of measurements: it holds that

$$\mathrm{E}_{(s_i),\,\boldsymbol{z}_i\in s_i} \sum_{(h_i),(a_i):\,\exists i, a_i\neq h_i(\boldsymbol{z}_i)} \big\langle Q^{(h_i)}_{(s_i)}, X^{(a_i)}_{(\boldsymbol{z}_i)} \big\rangle_\Psi = O\big(\varepsilon^{c_\ell}\big). \tag{6.18}$$

Moreover, from the case $k=1$ the measurements $\{Q^g_s\}$ satisfy

$$\mathrm{E}_{s,\,\boldsymbol{z}\in s} \sum_{h,a:\,a\neq h(\boldsymbol{z})} \big\langle Q^h_s, X^a_{\boldsymbol{z}} \big\rangle_\Psi = O\big(\varepsilon^{c_\ell}\big). \tag{6.19}$$

Let $(s_i) \in (\mathcal{S}_{m-1}(\mathbb{F}^m))^k$ be a k-tuple of parallel subspaces. By (6.18), (6.19), Claim 6.10, Markov's inequality, and a union bound we see that for all but a fraction at most $O(\varepsilon^{c_\ell/2})$ of such tuples, it holds that both measurements $\{Q^{(h_1,\ldots,h_{k-1})}_{(s_1,\ldots,s_{k-1})}\}$ and $\{Q^h_{s_k}\}$ are $O(\varepsilon^{c_\ell/2})$-consistent with the corresponding X measurements, and moreover $\{Q^h_{s_k}\}$ is $O(\varepsilon^{c_\ell/2})$-self-consistent. We are thus in a position to apply Lemma 6.13, with the set S_1 being the set of points in s_k and the set S_2 the set of points in $s_1\cup\cdots\cup s_{i-1}$ (the measurements A being the corresponding X measurements). As a result, the lemma promises the existence of a measurement $\{Q^{(h_1,\ldots,h_k)}_{(s_1,\ldots,s_k)}\}$ for which we can write

$$\begin{aligned}
&\mathrm{E}_{\boldsymbol{z}_i\in s_i} \sum_{\substack{(h_i),\,(a_i)\\ \exists i, a_i\neq h_i(\boldsymbol{z}_i)}} \big\langle Q^{(h_i)}_{(s_i)}, X^{(a_i)}_{(\boldsymbol{z}_i)} \big\rangle_\Psi \\
&= \mathrm{E}_{\boldsymbol{z}_i\in s_i} \sum_{(h_i)} \sum_{\substack{g\in\mathcal{P}_d(\ell(\boldsymbol{z}_i))\\ \exists i, g(\boldsymbol{z}_i)\neq h_i(\boldsymbol{z}_i)}} \big\langle Q^{(h_i)}_{(s_i)}, B^{g_{|\ell(\boldsymbol{z}_i)}}_{\ell(\boldsymbol{z}_i)} \big\rangle_\Psi \\
&\le \mathrm{E}_{\boldsymbol{z}_i\in s_i} \sum_{(h_i)} \sum_{\substack{g\in\mathcal{P}_d(\ell(\boldsymbol{z}_i))\\ \exists i\le k-1, g(\boldsymbol{z}_i)\neq h_i(\boldsymbol{z}_i)}} \big\langle Q^{(h_i)}_{(s_i)}, B^{g_{|\ell(\boldsymbol{z}_i)}}_{\ell(\boldsymbol{z}_i)} \big\rangle_\Psi \\
&\quad + \mathrm{E}_{\boldsymbol{z}_k\in s_k,\,\ell\ni\boldsymbol{z}_k} \sum_{(h_i)} \sum_{\substack{g\in\mathcal{P}_d(\ell)\\ g(\boldsymbol{z}_k)\neq h_k(\boldsymbol{z}_k)}} \big\langle Q^{(h_i)}_{(s_i)}, B^{g_{|\ell}}_{\ell} \big\rangle_\Psi \\
&= \mathrm{E}_{\boldsymbol{z}_i\in s_i} \sum_{(h_i)} \sum_{(a_i):\,\exists i\le k-1, a_i\neq h_i(\boldsymbol{z}_i)} \big\langle Q^{(h_i)}_{(s_i)}, X^{(a_i)}_{(\boldsymbol{z}_i)} \big\rangle_\Psi \\
&\quad + \mathrm{E}_{\boldsymbol{z}_k\in s_k} \sum_{(h_i)} \sum_{a\neq h_k(\boldsymbol{z}_k)} \big\langle Q^{(h_i)}_{(s_i)}, X^a_{\boldsymbol{z}_k} \big\rangle_\Psi \qquad (6.20)\\
&= O\big(\varepsilon^{c_\ell/4}\big), \qquad (6.21)
\end{aligned}$$

where for the first and third lines we used the definition of X, and for the last we used the consistency properties of the Q measurements with the corresponding X promised by Lemma 6.13. For those k-tuples (s_i) for which we could not apply Lemma 6.13 we define $\{Q^{(h_i)}_{(s_i)}\}$ arbitrarily, obtaining as a result that (6.21) holds on average over the choice of a k-tuple (s_i).

To conclude the proof of the lemma we apply Proposition 5.8, with the set X there being the set $\mathcal{S}$ of k-tuples of subspaces for which the triple $\mathcal{T}_{(s_i)}$ introduced in Lemma 6.6 is $(O(\varepsilon^{d_c/4}), 1/2)$-robust. The proposition shows that, provided ε is small enough, for every $(s_i)\in\mathcal{S}$ there exists an "improved" measurement $Q_{(s_i)}$ such that on average over $(s_i)\in\mathcal{S}$ the $\{Q^{(g_i)}_{(s_i)}\}$ are $O((\eta'(\varepsilon^{d_c/4}, 1/2))^{1/2})$-consistent with the

$\{X^{(a_i)}_{(\boldsymbol{z}_i)}\}$, where η' is as in Lemma 5.4. For the remaining subspaces (s_i) we keep the $Q_{(s_i)}$ as previously defined. Using the definition of η' and the fact that $c_\ell \leq d_c/40$, the lemma is proved. ∎

6.5. The measurements M^g. In this section we take the family of measurements $\{Q^{(g_i)}_{(s_i)}\}$ constructed in Lemma 6.11 and show that they can be transformed in a single measurement $\{M^g\}_{g\in\mathcal{P}_d(\mathbb{F}^m)}$ that satisfies the conclusion of Theorem 3.1.

Let $\{Q^{(g_1,\ldots,g_{d+1})}_{(s_1,\ldots,s_{d+1})}\}$ be the family of measurements obtained in Lemma 6.11 for $k = d+1$. By interpolation, from any tuple $g = (g_i)$ we may recover a single polynomial g of degree at most $2d$ defined on $\mathbb{F}^m$ and such that $g_{|s_i} = g_i$ for every $i = 1,\ldots,d+1$. This results in a family of measurements $\{R^g_{(s_1,\ldots,s_{d+1})}\}_{g\in\mathcal{P}_{2d}(\mathbb{F}^m)}$, where here we implicitly select one "representative" outcome $g \in \mathcal{P}_{2d}(\mathbb{F}^m)$ for every tuple (g_i) (as different degree-$2d$ polynomials may interpolate the same tuple). The following claim shows that we can in fact restrict our attention to interpolating polynomials of degree at most d.

Claim 6.14. *For every $(d+1)$-tuple of aligned subspaces (s_i) there exists a measurement $\{S^g_{(s_i)}\}_{g\in\mathcal{P}_d(\mathbb{F}^m)}$ such that*

$$\mathrm{E}_{(s_i)} \sum_{h\neq g_{|\ell}} \langle S^g_{(s_i)}, B^h_\ell\rangle_\Psi \;=\; O\big(\varepsilon^{c_\ell}\big).$$

Proof. First note that Lemma 6.11 and the definition of the measurements X imply the following:

$$\mathrm{E}_{(s_i),\boldsymbol{z}_i\in s_i} \sum_{g\in\mathcal{P}_{2d}(\mathbb{F}^m)} \langle R^g_{(s_i)}, (\mathrm{Id} - B^{g_{|\ell}}_{\ell(\boldsymbol{z}_i)})\rangle_\Psi \;=\; O\big(\varepsilon^{c_\ell}\big). \tag{6.22}$$

Note that B^h_ℓ is only defined for $h \in \mathcal{P}_d(\ell)$, but in general a degree-$(2d)$ polynomial will also have degree $2d$ when restricted to a line. Here though the definition of the X measurements shows that $g_{|\ell}$ should be interpreted as the degree-d polynomial obtained by interpolation from the values $(g(\boldsymbol{z}_i))$.

Next we argue that the contribution of measurement outcomes corresponding to polynomials of degree strictly larger than d must be small:

$$\begin{aligned}
&\mathrm{E}_{(s_i)} \sum_{g,\,\deg(g)>d} \langle R^g_{(s_i)}, \mathrm{Id}\rangle_\Psi \\
&\quad\approx_{\varepsilon^{c_\ell}} \mathrm{E}_{(s_i),\boldsymbol{z},\,\ell,\ell'\ni\boldsymbol{z}} \sum_{g,\,\deg(g)>d} \sum_{\substack{h(\ell\cap s_i)=g(\ell\cap s_i)\\ h'(\ell'\cap s_i)=g(\ell'\cap s_i)}} \mathrm{Tr}_\rho\big(R^g_{(s_i)}\otimes B^h_\ell\otimes B^{h'}_{\ell'}\big) \\
&\quad= O\big(\varepsilon^{d_c/2}\big),
\end{aligned}$$

where for the first line we applied (6.22) twice, and the second follows from the following argument. If g has degree $> d$, its restriction to all but a fraction at most $O(q^{-1})$ of lines ℓ also has degree $> d$ (see, e.g., [MR08, Lemma 6.4]). Hence the degree-d polynomial recovered by interpolation from $(d+1)$ values of g at random aligned points $\boldsymbol{z}_i \in \ell$ is unlikely to agree with $g_{|\ell}$ (since they have different degrees). Thus on the right-hand side of the first line above, the polynomial h (resp., h') will almost certainly disagree with g on a random point in ℓ (resp., ℓ') that is not in $\cup s_i$. The point $\boldsymbol{z}$ of intersection of ℓ and ℓ' is such a point, which gives the conclusion

using (6.3) and (6.2). The measurements S can thus be defined as R when $\deg(g) \le d$, and made into a complete measurement by adding all $R^g_{(s_i)}$ for $\deg(g) > d$ to a single outcome of degree less than d for S (e.g., the $g \equiv 0$ outcome). Using $c_\ell \le d_c/2$, the claim is proved. ∎

Let $g \in \mathcal{P}_d(\mathbb{F}^m)$, and define

$$M^g := \mathrm{E}_{(s_1,\ldots,s_{d+1})} S^g_{(s_i)}, \tag{6.23}$$

where the expectation is taken over all tuples of parallel $(m-1)$-dimensional subspaces s_i. Clearly $M^g \ge 0$ for every g. It also holds that $\sum_g M^g = \mathrm{Id}$. Indeed, for any $g \in \mathcal{P}_d(\mathbb{F}^m)$ and any tuple (s_i), $g_{|s_i}$ has degree at most d; moreover for any tuple $(g_i \in \mathcal{P}_d(s_i))$ there is at most one $g \in \mathcal{P}_d(\mathbb{F}^m)$ that interpolates all the g_i (indeed, any two such g should be 0 on each of $d+1$ parallel $(m-1)$-dimensional subspaces, hence should be 0 on all of $\mathbb{F}^m$). We conclude the proof of Theorem 3.1 by showing the following.

Claim 6.15. *Under the assumptions of Theorem* 3.1 *the measurement* $\{M^g\}_{g\in\mathcal{P}_d(\mathbb{F}^m)}$ *defined in* (6.23) *satisfies*

$$\mathrm{E}_{\boldsymbol{x}} \sum_{g,a:\, a\neq g(\boldsymbol{x})} \big\langle M^g, A^a_{\boldsymbol{x}} \big\rangle_\Psi = O\big(\varepsilon^{c_1}\big),$$

where $c_1 \le 1$ *is a universal constant.*

Proof. We have

$$\begin{aligned}
\mathrm{E}_{\boldsymbol{x}} \sum_{g,a:\, a\neq g(\boldsymbol{x})} \big\langle M^g, A^a_{\boldsymbol{x}} \big\rangle_\Psi &= \mathrm{E}_{(s_i),\,\boldsymbol{x}} \sum_{g,a:\, a\neq g(\boldsymbol{x})} \big\langle S^g_{(s_i)}, A^a_{\boldsymbol{x}} \big\rangle_\Psi \\
&\approx_{\varepsilon^{d_c/2}} \mathrm{E}_{(s_i),\,\boldsymbol{x}} \sum_{g,a:\, a\neq g(\boldsymbol{x})} \big\langle S^g_{(s_i)}, B^a_{\boldsymbol{x}} \big\rangle_\Psi \\
&= \mathrm{E}_{(s_i),\,\boldsymbol{x},\,\ell\ni\boldsymbol{x}} \sum_{g,h:\, h(\boldsymbol{x})\neq g(\boldsymbol{x})} \big\langle S^g_{(s_i)}, B^h_\ell \big\rangle_\Psi \\
&= O\big(\varepsilon^{c_\ell}\big),
\end{aligned}$$

where for the second line we used (6.2) and the last follows from Claim 6.14. ∎

6.6. Analysis of the two-level low-degree test. In this section we prove Theorem 3.2. Let $(|\Psi\rangle, A, B, C)$ be a strategy for the players with success probability at least $1-\varepsilon$ in the $(d, m, r, \mathbb{F})$ two-level low-degree test, as described in Figure 2. The probability that the referee accepts in steps 2 or 3 is at most $(1+|\mathbb{F}|)/|\mathbb{F}|^m \le 2/|\mathbb{F}| \le \varepsilon$ (given the assumption on $q = |\mathbb{F}|$ made in the theorem) for each step. Hence the strategy's success probability in steps 4.1 and 4.2 must be at least $1-6\varepsilon$ each, which immplies the following:

$$\mathrm{E}_{s\in\mathcal{S}_2(\mathbb{F}^m)} \mathrm{E}_{\boldsymbol{x}\in s} \sum_{a,b\in\mathbb{F},\, a\neq b} \big\langle A^a_{\boldsymbol{x}}, B^b_{s,\boldsymbol{x}} \big\rangle_\Psi \le 6\,\varepsilon, \tag{6.24}$$

$$\mathrm{E}_{\substack{s\in\mathcal{S}_2(\mathbb{F}^m)\\ s'\in\mathcal{S}_2(\mathbb{F}^{m'})}} \mathrm{E}_{\boldsymbol{x}'\in s'} \sum_{g'\in\mathcal{P}_{d'}(s')} \sum_{a'\in\mathbb{F},\, a'\neq g'(\boldsymbol{x}')} \big\langle C^{g'}_{s,s'}, B^{a'}_{s,\boldsymbol{x}'} \big\rangle_\Psi \le 6\,\varepsilon. \tag{6.25}$$

The following claim applies the analysis of the low-degree test to measurements B and C, separately for each $s \in \mathcal{S}_2(\mathbb{F}^m)$.

CLAIM 6.16. *For any $s \in \mathcal{S}_2(\mathbb{F}^m)$ there exists a measurement $\{M_s^g\}$ with outcomes $g \in \mathcal{P}_{dd'}(s)$ such that*

$$\mathrm{E}_{s\in\mathcal{S}_2(\mathbb{F}^m)}\mathrm{E}_{\boldsymbol{x}\in s}\sum_{g,a:\, a\neq g(\boldsymbol{x})}\big\langle A_{\boldsymbol{x}}^a, M_s^g\big\rangle_\Psi = O\big(\varepsilon^{c_1}\big), \tag{6.26}$$

where $c_1 > 0$ is the constant from Theorem 3.1.

Proof. Given a plane $s \in \mathcal{S}_2(\mathbb{F}^m)$ let ε_s be the value of the left-hand side of (6.25) (for that s), so that $\mathrm{E}_s[\varepsilon_s] \le 6\varepsilon$. By the definition of ε_s, the strategy $(|\Psi\rangle, B_s, C_s)$ has success at least $1-\varepsilon_s$ in the $(d', m', r, \mathbb{F})$ low-degree test. Choosing the constant d_2 large enough, Markov's inequality implies that a fraction at least $1-\varepsilon$ of subspaces s are such that ε_s satisfies the assumptions of Theorem 3.1. For those s, applying the theorem we obtain a measurement $\{M_s^g\}_{g\in\mathcal{P}_{d'}(\mathbb{F}^{m'})}$ such that

$$\mathrm{E}_{\boldsymbol{x}\in s}\sum_{g,a:\, g(\boldsymbol{x})\neq a}\big\langle M_s^g, B_{s,\boldsymbol{x}}^a\big\rangle_\Psi = O\big(\varepsilon_s^{c_1}\big). \tag{6.27}$$

Translating g back to a function on $s \subseteq \mathbb{F}^m$, we may also think of g as a bivariate polynomial with degree at most dd'. Using concavity of $z \to z^{c_1}$ (since $c_1 \le 1$) and the fact that (6.27) holds for all but an ε fraction of s we obtain

$$\mathrm{E}_{s\in\mathcal{S}_2(\mathbb{F}^m),\,\boldsymbol{x}\in s}\sum_{g\in\mathcal{P}_{dd'}(s)}\sum_{a:\, a\neq g(\boldsymbol{x})}\big\langle M_s^g, B_{s,\boldsymbol{x}}^a\big\rangle_\Psi = O\big(\varepsilon^{c_1}\big),$$

where for those s such that (6.27) we defined $\{M_s^g\}$ arbitrarily. Finally, using (6.24) it is not hard to see that this implies the claim. ∎

We turn to the proof of Theorem 3.2.

Proof of Theorem 3.2. For every $s \in \mathcal{S}_2(\mathbb{F}^m)$ let $\{M_s^g\}_{g\in\mathcal{P}_{dd'}(s)}$ be the measurement promised by Claim 6.16. The consistency relation (6.26) precisely states that the strategy $(|\Psi\rangle, A, M)$ has success probability at least $1-O(\varepsilon^{c_1})$ in the $(dd', m, r, \mathbb{F})$ low-degree test. Provided the constant d_2 is chosen large enough, we may apply Theorem 3.1 to that strategy to obtain a single measurement $\{\tilde{M}^h\}$ with outcomes $h \in \mathcal{P}_{dd'}(\mathbb{F}^m)$ satisfying

$$\mathrm{E}_{\boldsymbol{x}\in\mathbb{F}^m}\sum_{h\in\mathcal{P}_{dd'}(\mathbb{F}^m)}\sum_{a:\, a\neq h(\boldsymbol{x})}\big\langle \tilde{M}^h, A_{\boldsymbol{x}}^a\big\rangle_\Psi = O\big(\varepsilon^{c_1^2}\big),$$

which proves the theorem by choosing $c_2 = c_1^2$ and C_2 large enough. ∎

7. Analysis of additional tests.

7.1. The 3-SAT test. In this section we analyze the protocol for 3-SAT given in section 3.2 and prove Theorem 3.3. We note that the analysis is very standard in the PCP literature, and once the soundness of the low-degree test has been established virtually no additional complications are introduced from the consideration of entangled-player strategies.

Let φ be a 3-SAT formula on n variables, $\varepsilon > 0$, and $(|\Psi\rangle, A, B, C, D, E, F)$ an r-prover strategy with success $1-\varepsilon$ in the $(\varphi, n, r, \mathbb{F})$ 3-SAT test described in Figure 3. The following claim summarizes some initial consequences of the players' success in the test. For any clause $C \in \varphi$ on variables $x, y, z \in [n]$ we let $S(C)$ be the set of all degree-4 curves in $\mathbb{F}^m$ going through the points $\boldsymbol{x}, \boldsymbol{y}$, and $\boldsymbol{z}$. For any $c \in S(C)$, we let

$S(C,c)$ be the set of degree-4 curves in $\mathbb{F}^{m'}$ that go through $\#\boldsymbol{x}, \#\boldsymbol{y}, \#\boldsymbol{z}$, where the coordinate map $\#$ is determined by c.

CLAIM 7.1. *Under the assumptions of Theorem* 3.3, *there exists a measurement* $\{M^g\}_{g\in\mathcal{P}_{dd''}(\mathbb{F}^m)}$, *where* $d''=2\lceil\log(d+1)\rceil$, *such that the following hold:*

$$\mathrm{E}_{\boldsymbol{w}\in\mathbb{F}^m}\sum_{g,a:a\neq g(\boldsymbol{w})}\big\langle M^g, A^a_{\boldsymbol{w}}\big\rangle_\Psi = O\big(\varepsilon^{c_2}\big), \tag{7.1}$$

$$\mathrm{E}_{C\in\varphi}\mathrm{E}_{c\in S(C)}\mathrm{E}_{c'\in S(C,c)}\mathrm{E}_{\boldsymbol{w}\in c'}\sum_{g,a:a\neq g(\#\boldsymbol{w})}\big\langle F^g_{c,c'}, A^a_{\boldsymbol{w}}\big\rangle_\Psi = O\big(\varepsilon\big), \tag{7.2}$$

where $c_2>0$ *is the constant from Theorem* 3.2.

Proof. First we observe that the strategy $(|\Psi\rangle, A, B, C)$ must have success at least $1-2\varepsilon$ in the $(d,m,r,\mathbb{F})$ two-level low-degree test performed in step 2.1. Provided d_3 is chosen small enough compared to d_2, Theorem 3.2 implies the existence of a measurement $\{M^g\}$, with outcomes in $\mathcal{P}_{dd'}(\mathbb{F}^m)$, that is $O(\varepsilon^{c_2})$-consistent with A, proving (7.1). To show (7.2), we first note that in step 2.2.2 the point $\boldsymbol{w}'$ is uniformly distributed in c', hence that test in particular enforces that

$$\mathrm{E}_{C\in\varphi}\mathrm{E}_{c\in S(C)}\mathrm{E}_{c'\in S(C,c)}\mathrm{E}_{\boldsymbol{w}'\in c'}\sum_{g,a:a\neq g(\#\boldsymbol{w}')}\big\langle F^g_{c,c'}, D^a_{c,\#\boldsymbol{w}'}\big\rangle_\Psi = O(\varepsilon). \tag{7.3}$$

Similarly, the first check performed as part of step 2.2.1 in the protocol enforces that

$$\mathrm{E}_{C\in\varphi}\mathrm{E}_{c\in S(C)}\mathrm{E}_{c'\in S(C,c)}\mathrm{E}_{\boldsymbol{w}'\in c'}\sum_{a\neq b}\big\langle A^a_{\boldsymbol{w}'}, D^b_{c,\#\boldsymbol{w}'}\big\rangle_\Psi = O(\varepsilon). \tag{7.4}$$

Equation (7.2) is proved by combining (7.3) and (7.4). ∎

For every polynomial $g\in\mathcal{P}_{dd''}(\mathbb{F}^m)$ and variable $x\in[n]$, let $Z(g,x):=g(\boldsymbol{x})$, where $\boldsymbol{x}\in\mathbb{F}^m$ is the point associated with variable x, be the assignment that g implicitly associates with x. Let $Z(g)\subseteq\{0,1\}^n$ denote the assignment to all variables implied by g. Let $S(\varphi)\subseteq\{0,1\}^n$ be the set assignments satisfying a fraction at least $1-C_3\varepsilon^{c_3}$ of clauses of φ, where c_3, C_3 are constants as in the statement of Theorem 3.3 and defined in the proof of Claim 7.2 below.

CLAIM 7.2. *Under the assumptions of Theorem* 3.3 *it holds that*

$$\sum_{g:\,Z(g)\in S(\varphi)}\big\langle M^g, \mathrm{Id}\big\rangle_\Psi \geq 1-C_3\varepsilon^{c_3}.$$

Note that the claim implies in particular that provided ε is small enough φ *has* an assignment to its variables satisfying a fraction at least $1-C_3\varepsilon^{c_3}$ of clauses, proving Theorem 3.3 provided K_3 is chosen small enough.

Proof. Let $C=(x,y,z)$ be a clause, and $c=c(\boldsymbol{w})$ the degree-4 curve through $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{w})$, where $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in\mathbb{F}^m$ are the points associated with the variables x,y,z, respectively. For $(b,d,e)\in\{0,1\}^3$ we write $(b,d,e)\vdash C$ to indicate that the assignment $(x,y,z):=(b,d,e)$ satisfies the clause C. In step 2.2.2 of the protocol the referee

accepts with probability at least

$$
\begin{aligned}
&1-4\varepsilon\\
&\quad\le \mathrm{E}_{C=(x,y,z)\in\varphi}\mathrm{E}_{\substack{c\in S(C)\\ c'\in S(C,c)}}\sum_{(b,d,e)\vdash C}\sum_{\substack{g:\, g(\#\boldsymbol{x})=b, g(\#\boldsymbol{y})=d,\\ g(\#\boldsymbol{z})=e}}\big\langle F^g_{c,c'},\mathrm{Id}\big\rangle_\Psi\\
&\quad\le \mathrm{E}_{C=(x,y,z)\in\varphi}\mathrm{E}_{\substack{c\in S(C)\\ c'\in S(C,c)\\ \boldsymbol{w}'\in c'}}\sum_{(b,d,e)\vdash C}\sum_{\substack{g:\, g(\#\boldsymbol{x})=b,\\ g(\#\boldsymbol{y})=d,\\ g(\#\boldsymbol{z})=e}}\sum_{h:h(\boldsymbol{w}')=g(\#\boldsymbol{w}')}\mathrm{Tr}_\rho\big(F^g_{c,c'}\otimes M^h\big)\\
&\qquad + O\big(\varepsilon^{c_2}\big), \qquad (7.5)
\end{aligned}
$$

where the second equality follows from (7.1) and (7.2). The restriction of h to the curve c is a univariate polynomial of degree at most $4dd''$. Using variable substitution it is mapped to a polynomial on $\mathbb{F}^{m'}$ of total degree also at most $4dd''$, which when restricted to the degree-4 curve c' has degree at most $16dd''$. This polynomial is either equal to the degree-d' polynomial g, or, by the Schwartz–Zippel lemma, intersects it in a fraction at most $O(dd''/|\mathbb{F}|)$ of points, which is less than ε provided the constant d_3 is chosen large enough. Hence from (7.5) we get

$$
\begin{aligned}
1-O\big(\varepsilon^{c_2}\big) &\le \mathrm{E}_{C=(x,y,z)\in\varphi}\mathrm{E}_{\substack{c\in S(C)\\ c'\in S(C,c)\\ \boldsymbol{w}'\in c'}}\sum_{(b,d,e)\vdash C}\sum_{\substack{g:\, g(\boldsymbol{x})=b, g(\boldsymbol{y})=d,\\ g(\boldsymbol{z})=e}}\big\langle F^{g_{|c,c'}}_{c,c'},M^g\big\rangle_\Psi+\varepsilon\\
&\le \mathrm{E}_{C=(x,y,z)\in\varphi}\mathrm{E}_{\substack{c\in S(C)\\ c'\in S(C,c)\\ \boldsymbol{w}'\in c'}}\sum_{(b,d,e)\vdash C}\sum_{\substack{g:\, g(\boldsymbol{x})=b,\\ g(\boldsymbol{y})=d,\\ g(\boldsymbol{z})=e}}\big\langle M^g,\mathrm{Id}\big\rangle_\Psi+\varepsilon\\
&= \sum_{g\in\mathcal{P}_d(\mathbb{F}^m)}\mathrm{E}_{\substack{C=(x,y,z)\in\varphi\\ (g(\boldsymbol{x}),g(\boldsymbol{y}),g(\boldsymbol{z}))\vdash C}}\big\langle M^g,\mathrm{Id}\big\rangle_\Psi+\varepsilon, \qquad (7.6)
\end{aligned}
$$

where the last line is obtained by simplifying the expression. Given a polynomial g, let $\kappa(g)$ denote the fraction of clauses satisfied by the assignment to the variables of φ implicitly defined by g. Equation (7.6) shows that

$$
\sum_{g\in\mathcal{P}_d(\mathbb{F}^m)}\kappa(g)\langle M^g,\mathrm{Id}\rangle_\Psi\ \ge\ 1-O\big(\varepsilon^{c_2}\big).
$$

Since $(\langle M^g,\mathrm{Id}\rangle_\Psi)$ is a probability distribution over polynomials g, Markov's inequality implies that all but a fraction at most $O(\varepsilon^{c_2/2})$ of g chosen according to this distribution are such that $\kappa(g)\ge 1-O(\varepsilon^{c_2/2})$. This proves the claim for an appropriate choice of constants $c_3=c_2/2$ and C_3 large enough. ∎

7.2. The QUADEQ test. In this section we sketch the proof of Lemma 3.5. The analysis of the QUADEQ test as described in Figure 5 is rather standard (see, e.g., [AB09, Theorem 11.19]). Here the only additional complications introduced by the consideration of entangled players appear in the analysis of the linearity test, which was already stated in Theorem 3.4.

First we note that the players' success probability of $1-\varepsilon$ implies a success probability of at least $1-48\varepsilon$ in each of the six linearity steps performed in step 1.1 of the protocol. Applying Theorem 3.4 six times, for each of the measurements A_1, A_2, A_3, $B_{1,2}$, C, and D there exists a corresponding "linear" measurement $\{M^u_{A,i}\}_{u\in\mathbb{F}_2^{n/3}}$,

$\{M^v_{B,i,j}\}_{v\in\mathbb{F}_2^{2n/3}}$, $\{M^w_C\}_{w\in\mathbb{F}_2^n}$, and $\{M^z_D\}_{z\in\mathbb{F}_2^{n^2}}$, respectively, that is $O(\sqrt{\varepsilon})$-consistent with it. Replacing the players' actions in steps 1.2–1.5 in the protocol by the ones induced by these linear measurements still results in them being accepted with probability at least $1-O(\sqrt{\varepsilon})$.

It is not hard to argue (see the proof of Claim 7.2 for a similar argument) that step 1.5 in the protocol enforces that for a fraction at least $1-O(\varepsilon^{1/4})$ of outcomes z of the measurement M_D (under the distribution given by $(\langle M^z_D, \mathrm{Id}\rangle_\Psi)$) it holds that

$$\Pr_{w\in\mathbb{F}_2^{n^2}}\left(\sum_k w_k\left(\sum_{ij} z_{ij}\, a^{(k)}_{ij}\right) = \sum_k w_k\, c^{(k)}\right) \geq 1-O\big(\varepsilon^{1/4}\big).$$

For any such z, provided ε is small enough it is standard analysis to deduce that z defines an assignment to the n^2 "variables" x_ix_j that must satisfy *all* K equations in φ.

Finally, step 1.4 in the protocol enforces that a fraction at least $1-O(\sqrt{\varepsilon})$ of outcomes z of M_D are of the form $z=x\otimes x$ for some $x\in\mathbb{F}^n$. Indeed, any outcome which does not have this form will fail the test performed in step 1.4 with constant probability (over the choice of the questions and the outcomes of the other two measurements) whenever it is obtained.

Applying a union bound, we deduce that a fraction at least $1-O(\varepsilon^{1/4})$ of outcomes z of M_D are of the form (x,x) for some x defining a satisfying assignment to the variables in φ. Hence

$$\sum_{x\vdash\varphi}\langle M^{(x,x)}_D, \mathrm{Id}\rangle_\Psi = 1-O\big(\varepsilon^{1/4}\big), \tag{7.7}$$

and in particular whenever ε is small enough there must exist at least one such assignment, proving the first part of the lemma provided K_4 is chosen small enough.

To show the "furthermore" part of the lemma, we use steps 1.2 and 1.3 of the protocol. The purpose of the tests performed in those steps is to enforce that the measurement M_C associated with a particular instance φ_t is consistent with the measurements $M_{A,i}$, $M_{A,j}$, and $M_{A,k}$ obtained from the three chunks ℓ_i, ℓ_j, and ℓ_k of variables appearing in φ_t, where here $M_{A,i}$, $M_{A,j}$ depend only on the label ℓ_i, ℓ_j, respectively, but not on the instance φ_t. Hence the players' success $1-4\varepsilon$ in that test together with (7.7) (and the consistency between M_C and M_D enforced in step 1.4) implies that

$$\sum_{(u_i,u_j)\vdash\varphi_t}\langle M^{u_i}_{A,i}, M^{u_j}_{A,j}\rangle_\Psi = 1-O\big(\varepsilon^{1/4}\big).$$

This finishes the proof of the lemma provided the constants c_4, C_4 are chosen appropriately.

Acknowledgments. I am grateful to Dana Moshkovitz for helping me make my way through the classical literature on PCPs, including invaluable advice on the shortest path to obtaining constant-factor hardness of approximation results. I thank Ronald de Wolf for useful conversations that led to the discovery of an inaccuracy in a previous formulation of Corollary 4.6. I am indebted to the anonymous SICOMP referees for many useful comments that greatly helped improve the presentation, including the suggestion to use the parallel repetition theorem from [DSV14a], instead of the one from [KV11], in section 4.3.

REFERENCES

[AAV13] D. Aharonov, I. Arad, and T. Vidick, *The quantum PCP conjecture*, ACM SIGACT News, 44 (2013), pp. 47–79.

[AB09] S. Arora and B. Barak, *Computational Complexity: A Modern Approach*, Cambridge University Press, Cambridge, 2009.

[ADR82] A. Aspect, J. Dalibard, and G. Roger, *Experimental test of Bell's inequalities using time-varying analyzers*, Phys. Rev. Lett., 49 (1982), pp. 1804–1807.

[ALM+98] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy, *Proof verification and the hardness of approximation problems*, J. ACM, 45 (1998), pp. 501–555.

[Ara02] P. K. Aravind, *The Magic Squares and Bell's Theorem*, preprint, arXiv:quant-ph/0206070, 2002.

[AS97] S. Arora and M. Sudan, *Improved low-degree testing and its applications*, in Proceedings of the 30th STOC, ACM, New York, 1997, pp. 485–495.

[AS98] S. Arora and S. Safra, *Probabilistic checking of proofs: A new characterization of NP*, J. ACM, 45 (1998), pp. 70–122.

[Bel64] J. S. Bell, *On the Einstein-Podolsky-Rosen paradox*, Physics, 1 (1964), pp. 195–200.

[BFL91] L. Babai, L. Fortnow, and C. Lund, *Non-deterministic exponential time has two-prover interactive protocols*, Comput. Complexity, 1 (1991), pp. 3–40.

[BFLS91] L. Babai, L. Fortnow, L. A. Levin, and M. Szegedy, *Checking computations in polylogarithmic time*, in Procedings of the 23rd STOC, ACM New York, 1991, pp. 21–32.

[BGLR93] M. Bellare, S. Goldwasser, C. Lund, and A. Russell, *Efficient probabilistically checkable proofs and applications to approximations*, in Proceedings of the 25th STOC, ACM, New York, 1993, pp. 294–304.

[BGS98] M. Bellare, O. Goldreich, and M. Sudan, *Free bits, PCPs, and nonapproximability—towards tight results*, SIAM J. Comput., 27 (1998), pp. 804–915.

[BH13] F. G. S. L. Brandao and A. W. Harrow, *Quantum de Finetti theorems under local measurements with applications*, in Proceedings of the 45th STOC, ACM, New York, 2013, pp. 861–870.

[BLR93] M. Blum, M. Luby, and R. Rubinfeld, *Self-testing/correcting with applications to numerical problems*, J. Comput. System Sci., 47 (1993), pp. 549–595.

[CHSH69] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, *Proposed experiment to test local hidden-variable theories*, Phys. Rev. Lett., 23 (1969), pp. 880–884.

[CHTW04] R. Cleve, P. Høyer, B. Toner, and J. Watrous, *Consequences and limits of nonlocal strategies*, in Proceedings of the 19th IEEE Conference on Computational Complexity (CCC'04), IEEE Computer Society, Los Alamitos, CA, 2004, pp. 236–249.

[CS14] A. Chailloux and G. Scarpa, *Parallel repetition of entangled games with exponential decay via the superposed information cost*, in Automata, Languages, and Programming, J. Esparza, P. Fraigniaud, T. Husfeldt, and E. Koutsoupias, eds., Lecture Notes in Comput. Sci. 8572 , Springer, Berlin, 2014, pp. 296–307.

[DFK+11] I. Dinur, E. Fischer, G. Kindler, R. Raz, and S. Safra, *PCP characterizations of NP: Toward a polynomially-small error-probability*, Comput. Complexity, 20 (2011), pp. 413–504.

[DSV14a] I. Dinur, D. Steurer, and T. Vidick, *A parallel repetition theorem for entangled projection games*, in Proceedings of the 29th IEEE Conference on Computational Complexity (CCC'14), Washington, DC, 2014, IEEE Computer Society, Los Alamitos, CA, pp. 197–208.

[DSV14b] I. Dinur, D. Steurer, and T. Vidick, *A parallel repetition theorem for entangled projection games*, Comput. Complexity, 24 (2015), pp. 201–254.

[EPR35] A. Einstein, B. Podolsky, and N. Rosen, *Can quantum-mechanical description of physical reality be considered complete?*, Phys. Rev., 47 (1935), pp. 777–780.

[FGL+96] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy, *Interactive proofs and the hardness of approximating cliques*, J. ACM, 43 (1996), pp. 268–292.

[Gis91] N. Gisin, *Bell's inequality holds for all non-product states*, Phys. Lett. A, 154 (1991), pp. 201–202.

[GMR85] S. Goldwasser, S. Micali, and C. Rackoff, *The knowledge complexity of interactive proof-systems*, in Proceedings of the 17th STOC, ACM, New York, 1985, pp. 291–304.

[Hås01] J. Håstad, *Some optimal inapproximability results*, J. ACM, 48 (2001), pp. 798–859.

[IKM09] T. Ito, H. Kobayashi, and K. Matsumoto, *Oracularization and two-prover one-round interactive proofs against nonlocal strategies*, in Proceedings of the 24th IEEE Conference on Computational Complexity (CCC'09), IEEE Computer Society, Los Alamitos, CA, 2009, pp. 217–228.

[IV12] T. Ito and T. Vidick, *A multi-prover interactive proof for NEXP sound against entangled provers*, in Proceedings of the 53rd FOCS, IEEE, Piscataway, NJ, 2012, pp. 243–252.

[JPY14] R. Jain, A. Pereszlényi, and P. Yao, *A parallel repetition theorem for entangled two-player one-round games under product distributions*, in Proceedings of the 2014 IEEE 29th Conference on Computational Complexity (CCC), IEEE, Piscataway, NJ, 2014, pp. 209–216.

[KKM+11] J. Kempe, H. Kobayashi, K. Matsumoto, B. Toner, and T. Vidick, *Entangled games are hard to approximate*, SIAM J. Comput., 40 (2011), pp. 848–877.

[KRT10] J. Kempe, O. Regev, and B. Toner, *Unique games with entangled provers are easy*, SIAM J. Comput., 39 (2010), pp. 3207–3229.

[KV10] J. Kempe and T. Vidick, *Parallel Repetition of Entangled Games*, preprint, arXiv:1012.4728v2, 2010.

[KV11] J. Kempe and T. Vidick, *Parallel repetition of entangled games*, in Proceedings of the 43rd STOC, ACM, New York, 2011, pp. 3535–362.

[MR08] D. Moshkovitz and R. Raz, *Sub-constant error low degree test of almost-linear size*, SIAM J. Comput., 38 (2008), pp. 140–180.

[MR10] D. Moshkovitz and R. Raz, *Sub-constant error probabilistically checkable proof of almost-linear size*, Comput. Complexity, 19 (2010), pp. 367–422.

[Pis03] G. Pisier, *Introduction to Operator Space Theory*, Cambridge University Press, Cambridge, 2003.

[Pre07] D. Preda, *private communication*, 2007.

[Raz98] R. Raz, *A parallel repetition theorem*, SIAM J. Comput., 27 (1998), pp. 763–803.

[RS96] R. Rubinfeld and M. Sudan, *Robust characterizations of polynomials with applications to program testing*, SIAM J. Comput., 25 (1996), pp. 252–271.

[RS97] R. Raz and S. Safra, *A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP*, in Proceedings of the 30th STOC, ACM, New York, 1997, pp. 475–484.

[Sch35] E. Schrödinger, *Discussion of probability relations between separated systems*, Math. Proc. Cambridge Philos. Soc., 31 (1935), pp. 555–563.

[Sch80] J. T. Schwartz, *Fast probabilistic algorithms for verification of polynomial identities*, J. ACM, 27 (1980), pp. 707–717.

[TKRR13] Y. Tauman Kalai, R. Raz, and R. D. Rothblum, *Delegation for bounded space*, in Proceedings of the 45th STOC, ACM, New York, 2013, pp. 565–574.

[Tsi80] B. S. Tsirelson, *Quantum generalizations of Bell's inequality*, Lett. Math. Phys., 4 (1980), pp. 93–100.

[Vid11] T. Vidick, *The Complexity of Entangled Games*, Ph.D. thesis, UC Berkeley, Berkeley, CA, 2011.

[Zip79] R. Zippel, *Probabilistic algorithms for sparse polynomials*, in Symbolic and Algebraic Computation: An International Symposium on Symbolic and Algebraic Manipulation (EUROSAM), Lecture Notes in Comput. Sci. 72, Springer, Berlin, 1979, pp. 216–226.