A PARALLEL REPETITION THEOREM FOR ENTANGLED PROJECTION GAMES

IRIT DINUR, DAVID STEURER, AND THOMAS VIDICK

Abstract. We study the behavior of the entangled value of two-player one-round projection games under parallel repetition. We show that for any projection game $G$ of entangled value $1 - \varepsilon < 1$, the value of the $k$-fold repetition of $G$ goes to zero as $O((1 - \varepsilon^c)^k)$, for some universal constant $c \geq 1$. If furthermore the constraint graph of $G$ is expanding, we obtain the optimal $c = 1$. Previously exponential decay of the entangled value under parallel repetition was only known for the case of XOR and unique games. To prove the theorem, we extend an analytical framework introduced by Dinur and Steurer for the study of the classical value of projection games under parallel repetition. Our proof, as theirs, relies on the introduction of a simple relaxation of the entangled value that is perfectly multiplicative. The main technical component of the proof consists in showing that the relaxed value remains tightly connected to the entangled value, thereby establishing the parallel repetition theorem. More generally, we obtain results on the behavior of the entangled value under products of arbitrary (not necessarily identical) projection games.

Relating our relaxed value to the entangled value is done by giving an algorithm for converting a relaxed variant of quantum strategies that we call “vector quantum strategy” to a quantum strategy. The algorithm is considerably simpler in case the bipartite distribution of questions in the game has good expansion properties. When this is not the case, the algorithm relies on a quantum analogue of Holenstein’s correlated sampling lemma which may be of independent interest. Our “quantum correlated sampling lemma” generalizes results of van Dam and Hayden on universal embezzlement to the following approximate scenario: two non-communicating parties, given classical descriptions
of bipartite states $|\psi\rangle$, $|\varphi\rangle$, respectively, such that $|\psi\rangle \approx |\varphi\rangle$, are able to locally generate a joint entangled state $|\Psi\rangle \approx |\psi\rangle \approx |\varphi\rangle$ using an initial entangled state that is independent of their inputs.

**Keywords.** Parallel repetition, entangled games, projection games, correlated sampling, embezzlement.

**Subject classification.** 68Q12.

### 1. Introduction

Two-player one-round games arise naturally in many areas of theoretical computer science. They are prominent in complexity theory, where they are a powerful tool for the study of constraint satisfaction problems, and in cryptography, where they give a polyvalent abstraction used to establish the security of many two-party primitives. They have also recently proven a very convenient framework for the study of some of the deepest issues in quantum mechanics, giving a novel viewpoint on the decades-old study of Bell inequalities (Brunner et al. 2014), which are linear inequalities that must be satisfied by any family of distributions that can be generated locally according to the laws of classical mechanics, but can be violated if the distributions are allowed to be generated using quantum entanglement.

A game $G$ is specified by finite sets $\mathcal{U}, \mathcal{V}$ of questions, $\mathcal{A}, \mathcal{B}$ of answers, a probability distribution $\mu$ on pairs of questions $(u, v) \in \mathcal{U} \times \mathcal{V}$, and an acceptance criterion $V \subseteq \mathcal{A} \times \mathcal{B} \times \mathcal{U} \times \mathcal{V}$ which states, for every possible pair of questions $(u, v)$, which pairs of answers $(a, b) \in \mathcal{A} \times \mathcal{B}$ are valid. The most basic quantity associated with a game is its value. This can be defined operationally as the maximum success probability of two cooperating, but spatially isolated, players in the following game: A trusted party (the “referee”) selects a pair of questions $(u, v)$ according to $\pi$ and sends $u$ to the first player (“Alice”) and $v$ to the second (“Bob”). Each player replies with an answer $a, b$, and the players win the game if and only if $V(a, b, u, v) = 1$.

Remarkably, the precise definition of the value depends on the physical theory used to model the a priori vague assumption that
the players be “spatially isolated.” Under classical theory, isolated players are fully described by the (possibly randomized) functions they each apply to their respective question in order to determine their answer, and this interpretation leads to the classical value $\text{VAL}$ of the game. In contrast, in quantum theory isolated players are allowed any set of strategies that can be implemented by performing local measurements on a shared entangled state. The resulting value is called the entangled value and denoted $\text{VAL}^*(G)$. Clearly for every game, it holds that $\text{VAL} \leq \text{VAL}^*$, and it is the discovery of Einstein et al. (1935) (formalized by Bell (1964), simplified by Clauser et al. (1969) and experimentally verified by Aspect et al. (1981)) that there exist games for which the inequality is strict; indeed, there are families of games $(G_n)$ for which $\text{VAL}(G_n) \to 0$, but $\text{VAL}^*(G_n) = 1$ (Aravind 2002; Raz 1998). One can go even further and consider the non-signaling value $\text{VAL}^{ns}$, which corresponds to players allowed to reproduce any bipartite correlations that do not imply signaling. Here again $\text{VAL}^* \leq \text{VAL}^{ns}$, and there are games, such as the CHSH game (Clauser et al. 1969), for which the inequality is strict.

One of the most fundamental questions one may ask about two-player games is that of the behavior of the value under product. Given games $G$ and $H$, their product $G \otimes H$ is defined as follows: The question and answer sets are the Cartesian product of those from $G$ and $H$; the distribution on questions is the product of the distributions, and the acceptance criterion the AND of those of $G$ and of $H$. How does the value of $G \otimes H$ relate to that of $G$ and $H$? While it is clear that each of the three values defined above satisfies $\text{VAL}(G \otimes H) \geq \text{VAL}(G)\text{VAL}(H)$, the reverse inequality, although intuitive, does not hold in general. In particular, simple constructions of games $G$ are known such that $0 < \text{VAL}(G \otimes G) = \text{VAL}(G) < 1$ (Feige & Lovász 1992); similar constructions exist for $\text{VAL}^*$ (Cleve et al. 2008) and $\text{VAL}^{ns}$ (Kempe & Regev 2010).

In spite of these examples, one may still ask for the behavior of $\text{VAL}(G^{\otimes k})$, for “large” values of $k$. This is known as the parallel repetition question: Given a game $G$ such that $\text{VAL}(G) < 1$, does there exist a $\psi : [0,1] \to [0,1]$ such that $\psi(x) < 1$ whenever $x > 0$.
and $\text{VAL}(G^\otimes k) \leq (\psi(1 - \text{VAL}(G)))^k$? If so, what form does $\psi$ take? Can it be approximately linear in the vicinity of $x = 0$? Answering this question is of importance for many of the applications of two-player games. In cryptography, parallel repetition is a basic primitive using which one may attempt to amplify the security guarantees of a given protocol; in the study of Bell inequalities, it can be used, e.g., to amplify gaps between the quantum and non-signaling values; in complexity theory, it is an important tool for hardness amplification.

For the case of the classical value, a sequence of works (Feige 1991; Feige & Kilian 2000; Verbitsky 1994) over the course of a decade led to the breakthrough by Raz (1998), who was the first to provide a positive answer for general games: Raz showed that one can always take $\psi(x) = (1 - x^c)^{d/\log |A \times B|}$, where $c$ and $d$ are universal constants. Subsequent work focused on obtaining the best possible value for $c$ (the best known for general games is $c = 3$ (Holenstein 2009)) and on removing the dependence on the size of the answer alphabet for specific classes of games (Barak et al. 2009; Rao 2008; Raz & Rosen 2012). For the case of the no-signaling value, Holenstein (2009) showed one can always take $\psi(x) = 1 - Cx^2$ for some constant $C > 0$.

In contrast, for the case of the entangled value in spite of its importance the question is very poorly understood. Strong results are known for some very special classes of games such as XOR games (Cleve et al. 2008), for which repetition is exact (one can take $\psi(x) = 1 - x$) and unique games (Kempe & Regev 2010) (for $\psi(x) = 1 - Cx^2$, where $C > 0$ is a universal constant). However, both these results, as well as related results motivated by cryptographic applications (Hänggi & Renner 2009), rely on the formulation of the entangled value as a semidefinite program, a characterization that is not believed to extend to more general games. Additional results are known, but they only apply to specific games often originating from cryptography (Masanes et al. 2011; Tomamichel et al. 2013). Prior to this work, the most general results were due to Kempe & Vidick (2011), where it is shown that a specific type of repetition inspired by work of Feige & Kilian (2000), in which the original game is mixed with “consistency”
and “free” games, reduces the entangled value at a polynomial rate: Provided $\text{val}^*(G) < 1$, the value $\text{val}^*(G^{FK-\otimes k})$ of $k$ “Feige-Kilian” repetitions of $G$ behaves as $((1 - \text{val}'(G))k)^{-c}$ for some small $c > 0$. (See “related work” below for additional discussion of more recent results that appeared after the initial completion of this work.)

A recent work of Dinur & Steurer (2013) introduces a new approach to the parallel repetition question, focused on the case of projection games. A projection game is one in which the referee’s acceptance criterion has a special form: For any pair of questions $(u, v)$, any answer $b$ from the second player determines at most one valid answer $a = \pi_{uv}(b)$ for the first player. Projection games are among the most interesting and widely studied type of games. In particular, any local constraint satisfaction problem can be made into a projection game as follows: One player is asked for an assignment to all variables appearing in a constraint chosen at random, and the other is asked for an assignment to one of its variables. This simple transformation easily generalizes to convert any two-player game $G$ into a projection game $G'$, while essentially preserving the value: $1 - \text{val}(G') = \Theta(1 - \text{val}(G))$ (see Claim 2.8). In particular, if one is only interested in “amplifying the gap” between $\text{val}(G) = 1$ and $\text{val}(G) < 1$, one can first map $G$ to $G'$ and then consider the parallel repetition of $G'$ itself, and this justifies the predominant role played by projection games in classical complexity theory. This transformation, however, may decrease the entangled value arbitrarily whenever the optimal strategy for the players requires the use of entanglement (though we show that it can never increase the value by too much; see Claim 2.8 for precise bounds). Nevertheless, many of the games studied in quantum information, such as the CHSH game (Clauser et al. 1969) or the magic square game (Aravind 2002), are projection games.

The approach of Dinur & Steurer (2013) is based on the introduction of a relaxation of the game value, denoted $\text{val}_+$. This relaxation can be defined for any game (we give the definition in Section 1.2 below), and it is perfectly multiplicative. Moreover, for the case of projection games $\text{val}_+$ turns out to remain closely related to $\text{val}$, thus leading to a parallel repetition the-
 theorem. Although such a theorem already follows from Raz’s general result (Raz 1998), this arguably simpler approach matches the best parameters currently known (Rao 2008), which are known to be optimal (Raz 2008). In addition, it yields new results for repetitions of games with small value and the case of few repetitions, which has implications for the approximability of the LABEL COVER and SET COVER problems.

1.1. Our results. We extend the analytical framework introduced in Dinur & Steurer (2013) to the case of the entangled value $\text{val}^*$. As a consequence, we obtain the following main theorem on the parallel repetition of the entangled value of projection games.

**Theorem 1.1.** There exists constants $c, C > 0$ such that the following holds. For any projection game $G$,

$$\text{val}^*(G^{\otimes k}) \leq \left(1 - C(1 - \text{val}^*(G))^c\right)^{k/2}.$$ 

Although we do not attempt to fully optimize the constant $c$, the value that comes out of our proof is $c \leq 12$. For the case of expanding games (see definition in Section 2.2), we obtain the optimal $c = 1$.

Parallel repetition results for the classical value were originally motivated by the study of multi-prover interactive proofs (Fortnow et al. 1994), and our result is likewise applicable to the study of classes of multi-prover interactive proofs with entangled provers. Letting $\text{MIP}_{1,s}^\text{er}(k)$ denote the class of languages having $k$-prover 1-round interactive proofs in which completeness $c = 1$ holds with unentangled provers, but soundness $s$ holds even against provers allowed to share entanglement, and Theorem 1.1 implies that $\text{MIP}_{1,s}^\text{er}(2) = \text{MIP}_{1,2-\text{poly}}^\text{er}(2)$ for any $s < 1 - \text{poly}^{-1}$. This is because any protocol in $\text{MIP}_{1,s}^\text{er}(2)$ can be put into a form where the verifier’s test is a projection constraint by following the reduction already discussed above and described in Claim 2.8; this will preserve both perfect completeness (for classical strategies) and soundness bounded away from 1 (for quantum strategies). Prior to our work, it was not known how to amplify soundness to exponentially small without increasing the number of rounds of interaction. It follows from (Ito & Vidick 2012; Vidick 2013) that
MIP_{1,1-poly-1}(3) = \text{NEXP}, but very little is known about the 2-prover class MIP_{1,s}(2).

We believe that our results should find applications to a much wider range of problems. Going beyond the application to the parallel repetition question, our main contribution is the development of a precise framework in which general questions about the behavior of the value under product can be studied. This framework constitutes a comprehensive extension of the one introduced in Dinur & Steurer (2013) for the study of the classical value: As in Dinur & Steurer (2013), we introduce a relaxation $\text{val}^*_+$ of the entangled value, prove that it is perfectly multiplicative, and show that it remains closely related to $\text{val}^*_+$. We find it remarkable that the framework from Dinur & Steurer (2013), introduced in a purely classical context, would find such a direct extension to the case of the entangled value. We hope that the tools developed in this extension will find further applications to the proof of product theorems in areas ranging from cryptography to communication complexity. Even though at a technical level the setting can appear quite different, some of the ideas put forth here could also prove useful to further removed areas such as the additivity conjecture for the minimum output entropy of quantum channels (Amosov et al. 2000; Hastings 2009; Hayden & Winter 2008).

We turn to a more detailed explanation of our framework, hoping to highlight precisely those tools and ideas that may find further application.

1.2. Proof sketch. In order to explain our approach, it is useful to first review the framework introduced in Dinur & Steurer (2013) for the study of the classical value.

Classical strategies. The starting point in Dinur & Steurer (2013) consists in viewing games as operators acting on the space of strategies. In this language, a strategy is simply a vector $|f\rangle$ of nonnegative reals indexed by pairs $(u,a)$ of possible questions and answers: $f(u,a)$ is the probability that the strategy provides answer $a$ to question $u$. To any game, one can associate a matrix $G$ such that, formally, the success probability of strategies $(|f\rangle,|g\rangle)$
for the players equals the vector–matrix–vector product $\langle f|G|g \rangle$. The value of the game is then the norm of $G$ when viewed as an operator on the appropriately normed spaces of strategies.

The first crucial step taken in Dinur & Steurer (2013) consists in relaxing the value of a game $G$ to the value of a symmetrized version of the game, which we call the square $G^\dagger G$ of the game (this notation will be made precise in Section 2.2); we will denote the latter value by $\|G\|\square$. In the square of a game $G$, the referee first samples a question $u$ for the first player as in $G$. He then independently samples two questions $v$ and $v'$ for the second player according to the conditional distribution. The players in $G^\dagger G$ are sent $v$ and $v'$, respectively. They have to provide answers $b$ and $b'$ such that there exists an $a$ such that both $(a, b)$ is a valid answer to $(u, v)$ in $G$ and $(a, b')$ is a valid answer to $(u, v')$. Note that now $G^\dagger G$ is in general not a projection game, even if $G$ was. In particular, $G^\dagger G$ treats both players symmetrically, and it turns out that we may always assume that they both apply the same strategy. For the special case of projection games, it is not hard to show that the value of the game and that of its square are quadratically related:

$$\text{val}(G)^2 \leq \|G\|\square \leq \text{val}(G).$$

Indeed, using the algebraic language introduced above, the first inequality follows from the Cauchy-Schwarz inequality and the second is an easy observation.

The second step consists in observing that the application of the operator corresponding to the product $G \otimes H$, where $G$ and $H$ are arbitrary projection games, can be decomposed as a product $(G \otimes I) \cdot (I \otimes H)$. Starting with a strategy $|f\rangle$ for $G \otimes H$, the result of applying $(I \otimes H)$ to $|f\rangle$ is a new vector which no longer satisfies the strict normalization requirements of strategies. Understanding the new normalization leads to a further relaxation of $\|G\|\square$, denoted $\text{val}_+(G)$, in which the optimization is performed over the appropriate notion of “vector strategies,” which intuitively are vectors that can be obtained by applying game operators to strategies. With the correct definition, it is easy to show that

$$\|G \otimes H\|\square^2 \leq \text{val}_+(G) \cdot \|H\|^2\square.$$
The third and last step, which constitutes most of the technical work in Dinur & Steurer (2013), consists in showing that \( \text{val}_+(G) \) is a good approximation to \( \|G\|\Box \). This is done using a rounding procedure, by which a vector strategy associated with a large \( \text{val}_+ \) is mapped back to an actual strategy for the square game that also has a high value, thus serving as a witness for the value \( \|G\|\Box \) being large as well. Altogether we get a bound on the value of \( G \otimes H \) as a product of a bound on the value of \( G \) and a bound on the value of \( H \). Repeated application of (1.3) then leads to the following chain of inequalities (where the last approximate equality hides a polynomial dependence)

\[
\text{val}(G^{\otimes k})^2 \leq \|G^{\otimes k}\|^2 \leq \text{val}_+(G) \cdot \|G^{\otimes k-1}\|^2 \\
\leq \cdots \\
\leq \text{val}_+(G)^k \approx \text{val}(G)^k,
\]

(1.4)

proving the parallel repetition theorem.

**Quantum strategies.** Our goal now is to extend the above sketch to the case of the entangled value \( \text{val}^* \). There is good reason for optimism. In contrast to most classical proofs that appear in the study of classical two-player games (such as those that go into Dinur’s proof of the PCP theorem (Dinur 2007), or earlier approaches to parallel repetition (Feige & Kilian 2000; Raz 1998; Verbitsky 1994)), which are often information-theoretic or combinatorial in nature, the analytic (one could say linear algebraic) framework introduced in Dinur & Steurer (2013) seems much better suited a priori to an extension to the quantum domain. Indeed, quantum strategies themselves are objects that live in \( d \)-dimensional complex vector space: Instead of a vector \( |f\rangle \) of non-negative reals (giving the probability of answering \( a \) to question \( u \), for every possible \( u \) and \( a \)), a strategy is now a vector \( |A\rangle \) of \( d \)-dimensional positive semidefinite matrices \( A_u^a \) that describe the measurement to be performed upon receiving any question \( u \). The normalization condition is \( \sum_a A_u^a = \text{Id} \) for every \( u \), a constraint dictated by the formalism of measurements in quantum mechanics. Note that taking \( d = 1 \) we recover classical strategies; quantum mechanics allows \( d \) to be arbitrarily large.
At an abstract level, going from the classical to the entangled value thus solely requires us to think of the game $G$ as an operator acting on a bigger space of strategies, “enlarging” the nonnegative reals to the space of $d$-dimensional positive semidefinite matrices. This operation is easily realized by “tensoring with identity,” $G \rightarrow G \otimes \text{Id}_C$ said, extending each of the steps outlined above nevertheless raises a number of challenges unique to the quantum setting, in which far more than in the classical case the strength of strategies usually requires them to be studied in conjunction with the entanglement that enables their unique form of correlation.

It remains to show how to extend each of the steps outlined above. The first step consists in obtaining an analogue of (1.2). As in the classical case, the second inequality is easy and follows by observing that if $|A\rangle$ is a quantum strategy in $G^\dagger G$ then $(G \otimes \text{Id})|A\rangle$ is a valid strategy for the first player in $G$ (this notation will be made precise in Section 2.2.) The first inequality in (1.2) is slightly more subtle. Although it can be shown directly by applying a suitable matrix version of the Cauchy-Schwarz inequality, we note that it can also be proven using known properties of a widely used construction in quantum information theory, the pretty good measurement (PGM) (Hausladen et al. 1996; Hausladen & Wootters 1994). As it turns out, the relaxation $\text{val}^\ast(G) \rightarrow \|G\|_\square$ precisely corresponds to replacing the first player’s optimal choice of strategy in $G$ by a near-optimal choice obtained from the pretty good measurement derived from the post-measurement states, on the first player’s space, that arise from the second player’s measurements. As a consequence, (1.2) extends verbatim:

\[(1^\ast) \quad \text{val}^\ast(G)^2 \leq \|G\|_\square^2 \leq \text{val}^\ast(G).\]

Next we need to find an appropriate notion of vector strategy and corresponding relaxed value $\text{val}^\ast_+$. Here we are helped by the “operational” interpretation of a vector strategy as the result of the application of a game operator to a strategy meant for the product of several games. With the suitable generalization of the definition of classical vector strategies (see Definition 3.11), we also obtain an analogue of (1.3) for $\text{val}^\ast_+$:

\[(2^\ast) \quad \|G \otimes H\|_\square^2 \leq \text{val}^\ast_+(G) \cdot \|H\|_\square^2.\]
Even though this is not directly needed for our purposes, we note that $\text{val}^+$ itself is perfectly multiplicative (see Lemma 3.5 for the easy proof).

Finally, and most arduous, it is to relate the relaxation $\text{val}^+$ back to the value of the square game, $\|G\|_2^\ast$. In the classical case, this involves rounding vector to actual strategies. In the quantum case, rounding has to be performed synchronously by the players and will necessarily involve the use of an entangled state. Intuitively, upon receiving their respective questions in $G$ the players need to initialize themselves in an entangled state that corresponds to the post-measurement state that they would be in, conditioned on having given a particular pair of answers to a given pair of questions in the game $H$ from which the vector strategy is derived (recall that, informally, vector strategies are the result of applying a game operator to a strategy meant for the product of two or more distinct games).

In case the bipartite distribution of questions in the game $G$ has good expansion properties, we can show that this conditioned state is roughly the same regardless of the respective questions received by each player in $G$, so there is a way for players to renormalize their measurements and proceed. For the non-expanding case, the states can differ significantly from question to question. Nevertheless, we can show that based on their respective questions the players are able to agree on classical descriptions of two close states $|\psi\rangle \approx |\varphi\rangle$ that they, respectively, wish to be in.

Since the questions are not known to the players a priori, they need to generate the appropriate entangled states “on the spot,” from an initial shared entangled state that is independent from $|\psi\rangle$ and $|\varphi\rangle$. Our new “quantum correlated sampling” lemma allows the players to do just this: Given classical descriptions of $|\psi\rangle \approx |\varphi\rangle$, respectively, they are able to generate a joint entangled state $|\Psi\rangle \approx |\psi\rangle \approx |\varphi\rangle$ from an initial shared universal “embezzlement state” (van Dam & Hayden 2003) independent of $|\varphi\rangle$ or $|\psi\rangle$, without any communication. The lemma can be seen as a quantum variant of Holenstein’s correlated sampling lemma (Holenstein 2009), as well as a “robust” extension of the results of van Dam & Hayden (2003) on universal embezzlement states. We discuss this lemma and related works in more detail in Section 5.
All steps having been extended, we obtain a direct generalization of the chain of inequalities (1.4) to the case of entangled strategies:

\[
\text{VAL}^*(G^\otimes k)^2 \leq \|G^\otimes k\|_{\text{lin}}^2 \leq \text{VAL}^*(G) \cdot \|G^\otimes (k-1)\|_{\text{lin}}^2
\]

\[
\leq \text{VAL}^+(G)^k \approx (\text{VAL}^+(G))^k.
\]

(3*)

1.3. Additional related work. Although few general results are known, the question of the behavior of the entangled value of a two-player game or protocol under parallel repetition arises frequently. It plays an important role in results on device-independent quantum key distribution (Hänggi & Renner 2009; Masanes et al. 2011) and some cryptographic primitives (Tomamichel et al. 2013). The latter work considers parallel repetition of a game with quantum messages, a setting which is also the focus of Cooney et al. (2011). The approach of Cooney et al. (2011) builds upon Junge et al. (2010), who relate the (classical) value of a two-player one-round game to the norm of the game when viewed as a tensor on the space $\ell_\infty(\ell_1) \otimes \ell_\infty(\ell_1)$. This is similar to our starting point of viewing games as operators acting on strategies, except that it considers the game as a bilinear form rather than an operator; the two points of view are equivalent. This perspective enables the authors to leverage known results on the study of tensor norms in Banach space (resp. operator space) theory to derive results on the classical (resp. entangled) value. To the best of our knowledge, this connection has not led to an alternative approach to proving parallel repetition for general classes of games, although partial results were obtained in Cooney et al. (2011) for the special case of the entangled value of rank-one quantum games.

After the completion of this work, two new results established an exponential parallel repetition theorem for two-player one-round

\footnote{We note, however, that the approximate equality $\text{VAL}^+_{\text{lin}}(G) \approx \text{VAL}^*(G)$ that we obtain in the quantum case, although it suffices for our application to parallel repetition, is weaker than the one from (Dinur & Steurer 2013). In particular, it is probably not tight.}
games with entangled players, in which the distribution on questions is a product distribution. In Chailloux & Scarpa (2014), it is shown that the entangled value of games in which the distribution on questions is uniform decreases as

\[ \text{VAL}^*(G^\otimes k) \leq (1 - (1 - \text{VAL}^*(G))^2)\Omega(k/\log |\mathcal{U}||\mathcal{V}||\mathcal{A}||\mathcal{B}|). \]

Very recently Jain et al. (2013) extended the result to arbitrary product distributions on the questions, while also removing the dependence on the number of questions: They obtained the bound

\[ \text{VAL}^*(G^\otimes k) \leq (1 - (1 - \text{VAL}^*(G))^3)\Omega(k/\log |\mathcal{A}||\mathcal{B}|). \]

Both results are based on the use of information-theoretic techniques. They are incomparable to ours, as they apply to games in which the acceptance predicate is general but the input distribution is required to be product. In addition, both bounds above have a dependence on the number of answers in the game; while for the case of the classical value such a dependence is necessary (Feige & Verbitsky 2002), for the entangled value it is not yet known whether it can be avoided.

1.4. Open questions. We briefly mention several interesting open questions. There still does not exist any parallel repetition result that applies to the entangled value of general, non-projection two-player one-round games, and it would be interesting to investigate whether our techniques could lead to (even relatively weak) results in the general setting. The case of three players is also of interest, and no non-trivial parallel repetition results are known in either the classical or quantum setting. In fact, the closely related question of XOR repetition of three-player games is known to fail dramatically even for the classical value (Briët et al. 2012).

Organization of the paper. We start with some important preliminaries in Section 2. There we introduce the representation of games and strategies that is used throughout the remainder of the paper. In Section 3, we introduce the two relaxations of the entangled value sketched in the introduction and give a more detailed
overview of our proof. In Section 4 we prove the main technical component of our work, the relation between $\text{VAL}_+^*$ and $\|\cdot\|_2$. Finally, in Section 5 we state and prove the quantum correlated sampling lemma.

2. Preliminaries

2.1. Notation. We identify $\mathcal{L}(\mathbb{C}^d, \mathbb{C}^{d'})$, the set of linear operators from $\mathbb{C}^d$ to $\mathbb{C}^{d'}$, with the set of $d \times d'$ matrices with complex entries: If $X \in \mathcal{L}(\mathbb{C}^d, \mathbb{C}^{d'})$ then its matrix has entries $X_{a,b} = \langle a | X | b \rangle$, where $|a\rangle, |b\rangle$ range over the canonical bases for $\mathbb{C}^d, \mathbb{C}^{d'}$, respectively, and we use the bra-ket notation to denote column vectors $|b\rangle$ and row vectors $\langle a | = (|a\rangle)\dagger$, where $\dagger$ denotes the conjugate transpose. We also write $\mathcal{L}(\mathbb{C}^d)$ for $\mathcal{L}(\mathbb{C}^d, \mathbb{C}^d)$.

The space $\mathcal{L}(\mathbb{C}^d, \mathbb{C}^{d'})$ is a Hilbert space for the inner product $\langle A, B \rangle := \text{Tr}(A^\dagger B)$. We let $\|X\|_\infty$ be the operator norm of $X$, its largest singular value. A state $|\Psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^{d'}$ is a vector with norm 1.

The following simple calculation, sometimes known as Ando’s identity, will be useful.

Claim 2.1. Let $X \in \mathcal{L}(\mathbb{C}^{d}), Y \in \mathcal{L}(\mathbb{C}^{d'})$ be two operators and $|\Psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^{d'}$ a bipartite state with Schmidt decomposition $|\Psi\rangle = \sum_i \lambda_i |u_i\rangle |v_i\rangle$, where the $\lambda_i$ are nonnegative reals. Then

\begin{equation}
\langle \Psi | X \otimes Y | \Psi \rangle = \text{Tr}(XKY^T K^\dagger),
\end{equation}

where $K = \sum_i \lambda_i |u_i\rangle \langle v_i|$ and the transpose is taken in the bases specified by the $|u_i\rangle$ and $|v_j\rangle$. In particular, if $|u_i\rangle = |v_i\rangle$ for every $i, K$ is positive semidefinite and (2.2) evaluates to $\text{Tr}(XKY^T K)$.

Proof. The proof follows by direct calculation, expanding the left-hand side of (2.2) using the Schmidt decomposition of $|\Psi\rangle$ and the right-hand side using the definition of $K$. \hfill \Box

We state a matrix analogue of the Cauchy-Schwarz inequality; we include a proof for completeness; see also Pisier (2003, p.123).

Claim 2.3. For any $d$ and operators $A_i \in \mathcal{L}(\mathbb{C}^d), B_i \in \mathcal{L}(\mathbb{C}^{d'})$,

$$\left\| \sum_i A_i \otimes B_i \right\|_\infty^2 \leq \left\| \sum_i A_i \otimes A_i \right\|_\infty \left\| \sum_i B_i \otimes B_i \right\|_\infty.$$
Proof. Let $|\Psi\rangle, |\Phi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ be unit vectors with Schmidt decomposition $|\Psi\rangle = \sum_i \lambda_i |u_i\rangle |v_i\rangle$ and $|\Phi\rangle = \sum_i \mu_i |t_i\rangle |w_i\rangle$. For any $A \in \mathcal{L}(\mathbb{C}^d)$ and $B \in \mathcal{L}(\mathbb{C}^d)$,

$$
\langle \Psi | A \otimes B | \Phi \rangle = \sum_{i,j} \lambda_i \mu_j \langle u_i | A | t_j \rangle \langle v_i | B | w_j \rangle
$$

$$
\leq \left| \sum_{i,j} \lambda_i \mu_j |\langle u_i | A | t_j \rangle|^2 \right|^{1/2} \left| \sum_{i,j} \lambda_i \mu_j |\langle v_i | B | w_j \rangle|^2 \right|^{1/2}
$$

$$
= \left| \langle \Psi_L | A \otimes A | \Phi_L \rangle \right|^{1/2} \left| \langle \Psi_R | B \otimes B | \Phi_R \rangle \right|^{1/2},
$$

where $|\Psi_L\rangle = \sum_i \lambda_i |u_i\rangle |u_i\rangle$, $|\Phi_L\rangle = \sum_j \mu_j |t_j\rangle |t_j\rangle$, $|\Psi_R\rangle = \sum_i \lambda_i |v_i\rangle |v_i\rangle$ and $|\Phi_R\rangle = \sum_j \mu_j |w_j\rangle |w_j\rangle$. Applying the Cauchy-Schwarz inequality once more,

$$
\left| \langle \Psi | \left( \sum_i A_i \otimes B_i \right) | \Phi \rangle \right| \leq \left| \langle \Psi_L | \left( \sum_i A_i \otimes A_i \right) | \Phi_L \rangle \right|^{1/2}
$$

$$
\cdot \left| \langle \Psi_R | \left( \sum_i B_i \otimes B_i \right) | \Phi_R \rangle \right|^{1/2}.
$$

Since (2.4) holds for any $|\Psi\rangle$ and $|\Phi\rangle$, the claim is proved. \Box

2.2. Games and strategies.

Definitions. A two-player game is specified by finite question sets $\mathcal{U}$ and $\mathcal{V}$, finite answer sets $\mathcal{A}$ and $\mathcal{B}$, a distribution $\mu$ on $\mathcal{U} \times \mathcal{V}$, and an acceptance criterion $V \subseteq \mathcal{A} \times \mathcal{B} \times \mathcal{U} \times \mathcal{V}$. We also write $V(a, b, u, v) = 1$ for $(a, b, u, v) \in V$. The game may also be thought of as a bipartite constraint graph, with vertex sets $\mathcal{U}$ and $\mathcal{V}$, edge weights $\mu(u, v)$, and constraints $V(a, b, u, v) = 1$ on each edge $(u, v)$. We will write $\mu_L$ for the marginal distribution of $\mu$ on $\mathcal{U}$, and $\mu_R$ its marginal on $\mathcal{V}$. (We omit the subscripts $L$ and $R$ when they are clear from context.) We also often write $v \sim u$ to mean that $v$ is distributed according to the conditional distribution $\mu(v|u) = \mu(u, v)/\mu_L(u)$. The size of $G$ is defined as $|\mathcal{U}| |\mathcal{V}| |\mathcal{A}| |\mathcal{B}|$.

In this paper, we focus on projection games, which are games for which the acceptance criterion $V$ is such that for every $(u, v, b) \in \mathcal{U} \times \mathcal{V} \times \mathcal{B}$ there is at most one $a \in \mathcal{A}$ such that $V(a, b, u, v) = 1$. 
Equivalently, for every edge \((u, v)\) the associated constraint is a projection constraint \(\pi_{u,v} : \mathcal{B} \to \mathcal{A}\) such that \(\pi_{u,v}(b)\) is the unique \(a\) such that \(V(a, b, u, v) = 1\) if it exists, and a special “fail” symbol \(\bot\) otherwise. When the edge \((u, v)\) is clear from context, we will write \(b \to a\) to mean that \(\pi_{uv}(b) = a\). We also write \(b \leftrightarrow b'\) to mean that there exists an \(a\) such that \(b \to a\) and \(b' \to a\).

Given a projection game \(G\), let \(H\) be the weighted adjacency matrix associated with the square of \(G\): \(H\) is the \(|\mathcal{V}| \times |\mathcal{V}|\) matrix whose \((v, v')\)-th entry equals \(\mu(v, v') := \sum_u \mu(u)\mu(v|u)\mu(v'|u)\). Let \(D\) be the diagonal matrix with the degrees \(\mu_R(v)\) on the diagonal, and \(L := \text{Id} - D^{-1/2}HD^{-1/2}\) the normalized Laplacian associated with the square of \(G\). We say that a family of games \((G_n)\), where \(G_n\) has size \(n\), is expanding if the second smallest eigenvalue of \(L_n = L(G_n)\) is at least a positive constant independent of \(n\).

Projection games as operators. Let \(G\) be a two-player projection game. We will think of \(G\) as a linear operator \(G : \mathbb{C}^{|\mathcal{V}|} \otimes \mathbb{C}^{|\mathcal{B}|} \to \mathbb{C}^{|\mathcal{U}|} \otimes \mathbb{C}^{|\mathcal{A}|}\) defined as follows:

\[
G := \sum_{u,v} \mu(v|u) \sum_{a, b \to a} |u\rangle \langle v| \otimes |a\rangle \langle b| \in \mathcal{L}(\mathbb{C}^{|\mathcal{V}|} \otimes \mathbb{C}^{|\mathcal{B}|}, \mathbb{C}^{|\mathcal{U}|} \otimes \mathbb{C}^{|\mathcal{A}|}).
\]

In other words, for \(|B\rangle \in \mathbb{C}^{|\mathcal{V}|} \otimes \mathbb{C}^{|\mathcal{B}|}\), let \(B^b_v = \langle v, b|B\rangle\) denote the value of \(|B\rangle\) at the coordinates indicated by basis vectors \(|v\rangle \in \mathbb{C}^{\mathcal{V}}\) and \(|b\rangle \in \mathbb{C}^{\mathcal{B}}\). Then

\[
(GB)^a_u := \langle u, a|G|B\rangle = \sum_v \mu(v|u) \sum_{b \to a} B^b_v.
\]

Note that here we adopted the convention that questions \(u \in U\) are summed over, whereas questions \(v \in V\) are weighted by the corresponding conditional probability \(\mu(v|u)\).

Classical strategies. The actions of players in a game \(G\) give rise to a “probabilistic assignment,” a collection of probability distributions \(\{p(a, b|u, v)\}\) such that, for any pair of questions \((u, v)\), \(p(\cdot, \cdot|u, v)\) is a probability distribution on pairs of answers to those questions. We may also represent \(p\) as the rectangular
$|V||B| \times |U||A|$ matrix whose $((v, b), (u, a))$-th entry is $p(a, b|u, v)$. The \textit{value} achieved by $p$ in the game is defined as

$$\text{VAL}(G, p) := \sum_u \mu(u) \sum_v \mu(v|u) \sum_{a,b} p(a, b|u, v) = \text{Tr}_\mu(Gp),$$

where we introduced a trace $\text{Tr}_\mu$ on the set of all $X \in \mathcal{L}(\mathbb{C}^{|U|} \otimes \mathbb{C}^{|A|})$ by defining

$$\text{Tr}_\mu(X) := \sum_u \mu(u) \sum_a X_{(u,a),(u,a)}.$$

In cases of interest, the family of distributions $\{p(a, b|u, v)\}$ is not arbitrary, but has a bipartite structure which reflects the bipartite nature of the game. \textit{Classical} deterministic\footnote{Randomized strategies are convex combinations of deterministic strategies; thus, a randomized strategy can always be replaced by a deterministic one achieving at least as high a value.} strategies correspond to the case when $p(a, b|u, v) = f(a|u)g(b|v)$ for functions $f(\cdot|u) : \mathcal{A} \to \{0, 1\}$ and $g(\cdot|v) : \mathcal{B} \to \{0, 1\}$ taking the value 1 exactly once. The functions $f$ and $g$ may be represented as vectors

$$|f\rangle = \sum_{u,a} f(a|u)|u\rangle|a\rangle \in \mathbb{C}^{|U|} \otimes \mathbb{C}^{|A|}$$

and

$$|g\rangle = \sum_{v,b} g(b|v)|v\rangle|b\rangle \in \mathbb{C}^{|V|} \otimes \mathbb{C}^{|B|},$$

respectively. $p$ is then the rank-one matrix $p = |g\rangle\langle f|$, and we may express the value as

$$\text{VAL}(G, p) = \text{Tr}_\mu(Gp)$$

$$= \langle f, Gg \rangle_{\mu_L}$$

$$= \sum_u \mu_L(u) \sum_v \mu(v|u) \sum_{a,b} f(a|u)g(b|v),$$

where the inner product $\langle \cdot, \cdot \rangle_{\mu_L}$ is defined on $(\mathbb{C}^{|U|} \otimes \mathbb{C}^{|A|}) \times (\mathbb{C}^{|U|} \otimes \mathbb{C}^{|A|})$ by

$$\langle f, g \rangle_{\mu_L} := \sum_u \mu_L(u) \sum_a f(a|u)g(a|u).$$
We may similarly define an inner product $\langle \cdot, \cdot \rangle_{\mu_R}$ on $(\mathbb{C}^V \otimes \mathbb{C}^B) \times (\mathbb{C}^V \otimes \mathbb{C}^B)$, and we will omit the subscripts $L, R$ when they are clear from context. Given a game matrix $G$, we define its adjoint $G^\dagger$ as the unique matrix such that $\langle f, Gg \rangle_{\mu_L} = \langle G^\dagger f, g \rangle_{\mu_R}$ for all $f \in \mathbb{C}^{U \times A}$ and $g \in \mathbb{C}^{V \times B}$. Formally, if $G = \sum_{u,v} \mu(v|u) \sum_{a,b} \langle a| \otimes |b\rangle \langle b| \otimes |a\rangle$ then $G^\dagger = \sum_{u,v} \mu(u|v) \sum_{a,b} \langle b| \otimes |a\rangle \langle a| \otimes |b\rangle$.

Quantum strategies. Next we consider quantum strategies. A quantum strategy is specified by measurements $\{A^a_u\}_a$ for every $u$ and $\{B^b_v\}_b$ for every $v$, where in general a measurement is any collection of positive semidefinite operators, of arbitrary finite dimension $d$, that sum to identity. For any state $|\Psi\rangle$ representing the entanglement between the players, this strategy gives rise to the family of distributions

$$p_{|\Psi\rangle}(a, b|u, v) := \langle \Psi | A^a_u \otimes B^b_v | \Psi \rangle.$$ 

This formula, dictated by the laws of quantum mechanics, corresponds to the probability that the players obtain outcomes $a, b$ when performing the measurements $\{A^a_u\}, \{B^b_v\}$ on their respective share of $|\Psi\rangle$. One can check that positive semidefiniteness of the measurement operators together with the “sum to identity” condition implies that $p_{|\Psi\rangle}(\cdot, \cdot|u, v)$ is a well-defined probability distribution on $A \times B$. To a quantum strategy, we associate vectors

$$|A\rangle = \sum_{u,a} |u\rangle |a\rangle \otimes A^a_u \in \mathbb{C}^{|U|} \otimes \mathbb{C}^{|A|} \otimes \mathcal{L}(\mathbb{C}^d)$$

and

$$|B\rangle = \sum_{v,b} |v\rangle |b\rangle \otimes B^b_v \in \mathbb{C}^{|V|} \otimes \mathbb{C}^{|B|} \otimes \mathcal{L}(\mathbb{C}^d).$$

\[3\] In the literature the state $|\Psi\rangle$ is usually considered to be an integral part of the strategy. However, it will be more convenient for us to not fix it a priori. Given measurement operators for both players in a game, it is always clear what is the optimal choice of entangled state; it is obtained as the largest eigenvector of a given operator depending on the game and the measurements (see below).

\[4\] The complex conjugate on $A$ is not necessary, but for our purposes it is natural to include it in light of the proof of Lemma 3.1.
(Note that these definitions reduce to classical strategies whenever \( d = 1 \).) To express the success probability of this strategy in a game \( G \), we extend the definition of the inner product \( \langle \cdot, \cdot \rangle_\mu \) as follows.

**Definition 2.5 (Extended Inner Product).** The extended inner product

\[
\langle \cdot, \cdot \rangle_\mu_L : \mathbb{C}^{[d]} \otimes \mathbb{C}^{[d]} \otimes \mathcal{L}(\mathbb{C}^d) \mathbb{C}^{[d]} \otimes \mathbb{C}^{[d]} \otimes \mathcal{L}(\mathbb{C}^d) \rightarrow \mathcal{L}(\mathbb{C}^d) \otimes \mathcal{L}(\mathbb{C}^d)
\]

is defined, for \( |A\rangle = \sum_{u,a} |u\rangle |a\rangle \otimes A^a_u \) and \( |B\rangle = \sum_{u,a} |u\rangle |a\rangle \otimes B^a_u \),

by

\[
\langle A, B \rangle_\mu_L := \sum_u \mu_L(u) \sum_a A^a_u \otimes B^a_u.
\]

With this definition, the success probability of the strategy \((|A\rangle, |B\rangle)\) in \( G \) can be expressed as

\[
\text{val}^*(G, |A\rangle, |B\rangle) := \| \langle A, (G \otimes \text{Id}) B \rangle_\mu \|_{\infty}
\]

\[
= \| \sum_{u,a} \mu(u) A^a_u \otimes \left( \sum_v \mu(v) \sum_{b \rightarrow a} B^b_v \right) \|_{\infty}
\]

\[
= \| \sum_{u,v} \mu(u, v) \sum_{a,b \rightarrow a} A^a_u \otimes B^b_v \|_{\infty}
\]

\[
= \max_{|\Psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d, \|\Psi\| = 1} \sum_{u,v} \mu(u, v) \sum_{a,b \rightarrow a} \langle \Psi | A^a_u \otimes B^b_v | \Psi \rangle.
\]

We also define the entangled value of the game, \( \text{val}^*(G) \), to be the highest value achievable by any quantum strategy:

\[
\text{val}^*(G) = \sup_{|A\rangle, |B\rangle} \text{val}^*(G, |A\rangle, |B\rangle)
\]

\[
= \sup_{|A\rangle, |B\rangle} \| \langle A, (G \otimes \text{Id}) B \rangle_\mu \|_{\infty}
\]

\[
= \sup_{\{A^a_u, B^a_v, |\Psi\rangle\}} \sum_{u,v} \mu(u, v) \sum_{a,b \rightarrow a} \langle \Psi | A^a_u \otimes B^b_v | \Psi \rangle
\]

\[
= \sup_{\{A^a_u, B^a_v, |\Psi\rangle\}} \sum_u \mu(u) \sum_{a} \langle \Psi | A^a_u \otimes B^a_u | \Psi \rangle,
\]

(2.6)

\[\text{Note the definition depends on a fixed choice of basis for the spaces \( \mathbb{C}^{[d]} \) and \( \mathbb{C}^{[d]} \).} \]
where here we slightly abuse notation and denote
\[
B^a_u := (\langle u|a|\otimes \Id)(G \otimes \Id)|B\rangle = \sum_v \mu(v|u) \sum_{b \to a} B^b_v.
\]

We note that in the above the supremum may in general not be attained as optimal strategies may require infinite dimensions. In this paper, we always restrict ourselves to finite-dimensional strategies.\(^6\)

It is well-known that any two-player game can be made into a projection game while essentially preserving its classical value. The following claim gives a partial extension of this fact to the case of the entangled value.

**Claim 2.8.** There exists a polynomial-time computable transformation mapping any two-player one-round game \(G\) to a projection game \(G'\) such that the following hold:

\[
1 - \text{val}(G') \leq 1 - \text{val}(G) \leq 2(1 - \text{val}(G')).
\]

In particular, \(\text{val}(G') = 1\) if and only if \(\text{val}(G) = 1\), and \(1 - \text{val}(G') = \Theta(1 - \text{val}(G))\). Moreover, for the entangled value we have the weaker bound

\[
\text{val}^*(G') \leq \sqrt{\frac{1 + \text{val}^*(G)}{2}},
\]

which implies \(1 - \text{val}^*(G') = \Omega(1 - \text{val}^*(G))\).

**Proof.** Let \(G\) be a game with (without loss of generality disjoint) question sets \(U, V\), answer sets \(A, B\), distribution on questions \(\mu\), and acceptance predicate \(V\). Let \(G'\) be the projection game corresponding to the following scenario. The referee selects a pair of questions \((u, v)\) at random from \(\mu\), which it sends to the second player, and then sends either \(u\) or \(v\) to the first player, each with probability \(1/2\). Formally, \(G'\) is defined by question sets \(U' = U \cup V\),

\(^6\)Thus when we say that \((|A\rangle, |B\rangle)\) achieve the value of \(G\), we really mean that \((|A\rangle, |B\rangle)\) are finite-dimensional strategies whose value in \(G\) can be made arbitrarily close to the optimum; for clarity, we ignore this simple technicality in the whole paper.
$\mathcal{Y}' = \mathcal{U} \times \mathcal{V}$, answer sets $\mathcal{A}' = \mathcal{A} \cup \mathcal{B}$, $\mathcal{B}' = \mathcal{A} \times \mathcal{B}$, and a distribution $\mu'$ given by $\mu'(u, (u, v)) = \mu'(u, v)/2$, $\mu'(v, (u, v)) = \mu'(u, v)/2$, and 0 otherwise. For any $(u, v)$ and $(a, b)$, let $\pi_{u,(u,v)}(a, b) = a$ and $\pi_{v,(u,v)}(a, b) = b$ if $V(a, b, u, v) = 1$, and there is no valid answer for the first player if the second player’s answers are such that $V(a, b, u, v) = 0$.

Then clearly $G'$ is a projection game. Let $|f\rangle, |g\rangle$ be classical deterministic strategies for the players such that $\text{VAL}(G, |f\rangle, |g\rangle) = \text{VAL}(G)$. Consider the strategy $(|f'\rangle, |g'\rangle)$ for $G'$ in which $|f'\rangle$ answers as $|f\rangle$ to questions $u \in \mathcal{U}$ and as $|g\rangle$ to questions $v \in \mathcal{V}$, and $|g'\rangle$ answers as $(|f\rangle, |g\rangle)$. Then whenever the strategy $(|f\rangle, |g\rangle)$ provides answers to a pair of questions $(u, v)$ that satisfy the predicate $V$, the strategy $(|f'\rangle, |g'\rangle)$ gives answers to both $(u, (u, v))$ and $(v, (u, v))$ that are accepted in $G'$; hence,

$$\text{VAL}(G') \geq \text{VAL}(G', |f'\rangle, |g'\rangle) \geq \text{VAL}(G, |f\rangle, |g\rangle) = \text{VAL}(G).$$

Conversely, let $(|f'\rangle, |g'\rangle)$ be a strategy for $G'$ such that $\text{VAL}(G') = \text{VAL}(G', |f'\rangle, |g'\rangle)$. Decompose $|f'\rangle$ into a pair of strategies $|f\rangle, |g\rangle$ in $G$, depending on whether the question is $u \in \mathcal{U}$ or $v \in \mathcal{V}$. The pair $(|f\rangle, |g\rangle)$ will give a rejected answer to a pair of questions $(u, v)$ only if $(|f'\rangle, |g'\rangle)$ gave a rejected answer to at least one of the questions $(u, (u, v))$ and $(v, (u, v))$ in $G'$. In the worst case, the $(1 - \text{VAL}(G', |f\rangle, |g\rangle))$ probability that $(|f'\rangle, |g'\rangle)$ provides rejected answers in $G'$ is, say, fully concentrated on questions of the form $(u, (u, v))$. Hence,

$$\text{VAL}(G') \geq \text{VAL}(G, |f\rangle, |g\rangle) \geq 1 - 2(1 - \text{VAL}(G', |f'\rangle, |g'\rangle)) = 1 - 2(1 - \text{VAL}(G')).$$

Finally, let $(|A\rangle, |B\rangle)$ be a pair of quantum strategies such that $\text{VAL}^*(G') = \text{VAL}^*(G', |A\rangle, |B\rangle)$. To $|A\rangle$, we unambiguously associate measurement operators $\{A^a_u\}_a$ for every $u \in \mathcal{U}$, and $\{A^b_v\}_b$ for $v \in \mathcal{V}$. Hence,
\[ \text{VAL}^*(G') = \left\| \mathbf{E} \frac{1}{2} \sum_{(a,b):V(a,b,u,v)=1} \mathbf{A}_u^a \otimes \mathbf{B}_u^{a,b} + \mathbf{A}_v^b \otimes \mathbf{B}_v^{a,b} \right\|_\infty \]
\[ \leq \left\| \mathbf{E} \frac{1}{4} \sum_{(a,b):V(a,b,u,v)=1} \left( (A_u^a + A_v^{b}) \otimes (A_u^a + A_v^{b}) \right)^{1/2} \right\|_\infty \]
\[ \cdot \left\| \mathbf{E} \sum_{(a,b):V(a,b,u,v)=1} \mathbf{B}_u^{a,b} \otimes \mathbf{B}_u^{a,b} \right\|_\infty^{1/2} \]
\[ \leq \left( \frac{1}{2} + \frac{1}{2} \right) \left\| \mathbf{E} \sum_{(a,b):V(a,b,u,v)=1} \mathbf{A}_u^a \otimes \mathbf{A}_v^b \right\|_\infty^{1/2} \]

where the first inequality uses Claim 2.1 and the last uses the triangle inequality for the operator norm and the fact that \( \left\| \sum_i X_i \otimes Y_i \right\|_\infty = \left\| \sum_i Y_i \otimes X_i \right\|_\infty \) for any \( X_i, Y_i \) to bound the first term, and uses \( \sum_{a,b} B_{u,v}^{a,b} \leq \text{Id} \) for every \( u, v \), which implies

\[ \left\| \mathbf{E} \sum_{(a,b):V(a,b,u,v)=1} \mathbf{B}_{u,v}^{a,b} \otimes \mathbf{B}_{u,v}^{a,b} \right\|_\infty \leq \left\| \mathbf{E} \sum_{(a,b):V(a,b,u,v)=1} \text{Id} \otimes \mathbf{B}_{u,v}^{a,b} \right\|_\infty \]
\[ \leq \left\| \text{Id} \otimes \text{Id} \right\|_\infty = 1, \]

to bound the second. Hence, the pair of strategies \((|A_{|\mathcal{U}}\rangle, |A_{|\mathcal{V}}\rangle)\) for \( G \) achieves a value at least

\[ \text{VAL}^*(G) \geq \text{VAL}^*(G, |A_{|\mathcal{U}}\rangle, |A_{|\mathcal{V}}\rangle) \geq 2 \text{VAL}^*(G')^2 - 1, \]

as claimed. \( \Box \)

### 3. Relaxations of the game value

In this section, we introduce two relaxations of the entangled value \( \text{VAL}^*(G) \) of a projection game \( G \). Both are quantum analogues of relaxations in Dinur & Steurer (2013) and are used in the same way. The first relaxation, denoted \( \|G\|_{\text{\#}} \), is related to playing a “squared” version of \( G \) with two players Bob and Bob’ treated symmetrically. It is defined in Section 3.1 and is easily seen to give a good approximation to \( \text{VAL}^* \), as shown in the following lemma (see Section 3.1 for the proof):
**Lemma 3.1.** For any projection game $G$,

$$\text{val}^*(G)^2 \leq \|G\|_b^2 \leq \text{val}^*(G).$$

The second relaxation, denoted $\text{val}_+^*(G)$, is defined in Section 3.2. It will be proven to be a good approximation to $\|G\|_b$ and thus to $\text{val}^*$, although this will require more work.

**Lemma 3.3.** For any projection game $G$,

$$\|G\|_b^2 \leq \text{val}_+^*(G) \leq 1 - C(1 - \|G\|_b^2)^c,$$

for some positive constants $C, c > 0$.

The proof of Lemma 3.3 is given in Section 4. The definition of $\text{val}_+^*$ is motivated by the following multiplicative property.

**Lemma 3.5.** For any two projection games $G$ and $H$,

$$\|G \otimes H\|_b^2 \leq \text{val}_+^*(G) \cdot \|H\|_b^2,$$

and $\text{val}_+^*$ is perfectly multiplicative:

$$\text{val}_+^*(G \otimes H) = \text{val}_+^*(G) \cdot \text{val}_+^*(H).$$

The proof of Lemma 3.5 is given in Section 3.2.

With these three inequalities in hand, we easily derive the parallel repetition theorem, Theorem 1.1, as follows. By repeated applications of (3.6), followed by (3.4), we get

$$\|G^{\otimes k}\|_b^2 = \|G \otimes G^{\otimes k-1}\|_b^2 \leq \text{val}_+^*(G) \cdot \|G^{\otimes k-1}\|_b^2 \leq \cdots \leq (\text{val}_+^*(G))^k.$$  

Combining with (3.2) and (3.4), we get

$$\text{val}^*(G^{\otimes k})^2 \leq \|G^{\otimes k}\|_b^2 \leq (\text{val}_+^*(G))^k \leq (1 - C(1 - \|G\|_b^2)^c)^k \leq (1 - C(1 - \text{val}^*(G))^c)^k,$$

where the last step follows from (3.2) and the monotonicity of $x \mapsto 1 - C(1 - x)^c$ on $[0, 1]$. 
3.1. The square norm.

**Definition 3.8.** For a game $G$ and a quantum strategy $|B\rangle$ write
\[
\|G \otimes \text{Id} |B\rangle\|_\square := \left(\|\langle G \otimes \text{Id} B, G \otimes \text{Id} B \rangle_\mu \|_\infty\right)^{1/2}
\]
and define
\[
\|G\|_\square := \sup_{|B\rangle} \|G \otimes \text{Id} |B\rangle\|_\square,
\]
where the supremum is taken over all $d$ and quantum strategies $|B\rangle \in \mathbb{C}^{|V|} \otimes \mathbb{C}^{|B|} \otimes \mathcal{L} (\mathbb{C}^d)$.

We note that $\|\cdot\|_\square$ is clearly homogeneous and nonnegative. Although we will not use it, one can check that $\|\cdot\|_\square$ is also definite, and hence a norm, by setting $B_v^b = \text{Id}$ for every $v$ and any $b$ such that $(G^\dagger G)_{(v,b),(v,b)} \neq 0$ (when it exists, and for an arbitrary $b$ otherwise).

**Lemma 3.1** claims that $\|G \otimes \text{Id} |B\rangle\|_\square$ gives a good approximation to the maximum success probability in the game, when Bob uses the strategy specified by $|B\rangle$. We give a self-contained proof of the lemma below, but before proceeding readers familiar with quantum information theory may find it interesting to note that a direct proof of the first inequality can be derived using known properties of the pretty good measurement (PGM) (Hausladen et al. 1996; Hausladen & Wootters 1994). We briefly indicate how. Suppose Bob’s strategy in $G$ is fixed to $|B\rangle$. Upon receiving her question $u$, Alice has to decide on an answer $a$. She knows that Bob will receive a question $v$ distributed according to $\mu(\cdot|u)$ and apply his measurement, obtaining an outcome $b$ and resulting in the post-measurement state $\text{Tr}_2(\text{Id} \otimes \sqrt{B_v^b}|\Psi\rangle\langle\Psi| \text{Id} \otimes \sqrt{B_v^b})$ on her system. From her point of view, Alice needs to provide an answer $a$ such that $\pi_{uv}(b) = a$. Only knowing $u$, her task thus amounts to optimally distinguishing between the collection of post-measurement states
\[
\rho_u^a = \mathbb{E}_{v \sim u} \sum_{b \to a} \text{Tr}_2(\text{Id} \otimes \sqrt{B_v^b}|\Psi\rangle\langle\Psi| \text{Id} \otimes \sqrt{B_v^b}).
\]
If, instead of applying the optimal distinguishing measurement, Alice applied the pretty good measurement (PGM) derived from this family of states then it follows from (Barnum & Knill 2002)
that the players’ success probability would be at most quadratically worse than what it would be was Alice to apply the optimal measurement. Using the explicit form of the PGM, one can verify that the resulting value exactly corresponds to $\|G \otimes \text{Id} \, |B\rangle\|_2^2$, which proves the first inequality in (3.2).

**Proof (Proof of Lemma 3.1).** We prove the following inequality, from which (3.2) follows by taking the supremum over all $|B\rangle$:

$$\max_{|A\rangle} \text{val}^*(G, |A\rangle, |B\rangle)^2 \leq \|G \otimes \text{Id} \, |B\rangle\|_2^2 \leq \max_{|A\rangle} \text{val}^*(G, |A\rangle, |B\rangle).$$

(3.9)

For the second inequality, using that $G$ is a projection game we note that for any $d$-dimensional strategy $|B\rangle$ for the second player, $(G \otimes \text{Id})|B\rangle$ is a valid strategy for the first player; hence,

$$\|(G \otimes \text{Id})|B\rangle\|_2^2 = \|\langle G \otimes \text{Id} B, G \otimes \text{Id} B \rangle\|_\infty \leq \max_{|A\rangle} \|\langle A, (G \otimes \text{Id}) B \rangle\|_\infty = \max_{|A\rangle} \text{val}^*(G, |A\rangle, |B\rangle).$$

To show the first, we write the following:

$$\text{val}^*(G, |A\rangle, |B\rangle) = \|\langle A, (G \otimes \text{Id}) B \rangle\|_\infty$$

$$= \left\| \sum_u \mu(u) \sum_a A_u^a \otimes B_u^a \right\|_\infty$$

$$\leq \left\| \sum_u \mu(u) \sum_a A_u^a \otimes A_u^a \right\|_\infty^{1/2}$$

$$\cdot \left\| \sum_u \mu(u) \sum_a B_u^a \otimes B_u^a \right\|_\infty^{1/2}$$

$$\leq \|(G \otimes \text{Id}) B\|_\infty,$$

where for the first inequality we used the matrix Cauchy-Schwarz inequality stated in Claim 2.3, and the last inequality uses $\sum_a A_u^a \leq \text{Id}$ for every $u$. $\square$
3.2. The relaxation $\text{val}^\ast_+(G)$. In order to motivate our definition of $\text{val}^\ast_+$, let us consider two projection games $G, H$ and any quantum strategy $|B\rangle$ for $G \otimes H$ that achieves the optimal value $\|G \otimes H\|_2^2$ in the square game. Letting $\kappa := \|G \otimes H\|_2^2/\|H\|_2^2$, we want to bound $\kappa$ by a quantity that depends on $G$ and not on $H$.

Consider the factorization $G \otimes H = (G \otimes I)(I \otimes H)$ where $I$ is the identity operator on the question and answer spaces associated with the first (resp. second) player in $H$ (resp. $G$); note that $I$ can also be understood as a game in which the two players are asked the same question and win if and only if they return the same answer. The application of $G \otimes H$ thus gives rise to a two-step process

$|A'\rangle \xrightarrow{G \otimes I} |A\rangle \xleftarrow{I \otimes H} |B\rangle$,

mapping $|B\rangle$ to $|A\rangle := ((I \otimes H) \otimes \text{Id})|B\rangle$ and then mapping $|A\rangle$ to $|A'\rangle := ((G \otimes I) \otimes \text{Id})|A\rangle$. Let us view $|B\rangle$ as a table with rows indexed by $\mathcal{V}_G \times \mathcal{B}_G$ and columns indexed by $\mathcal{V}_H \times \mathcal{B}_H$, where $\mathcal{V}_G, \mathcal{V}_H$ and $\mathcal{B}_G, \mathcal{B}_H$ are the question and answer sets associated with the second player in $G$ and $H$, respectively, and whose entries are measurement operators, i.e., elements in $\mathcal{L}(\mathbb{C}^d)$. Then, $|A\rangle$ is the result of applying $H \otimes \text{Id}$ on each row of $|B\rangle$ separately, and we apply $G \otimes \text{Id}$ on each column of $|A\rangle$ separately to get $|A'\rangle = (G \otimes I \otimes \text{Id})|A\rangle$.

It is instructive to view the strategy $|B\rangle$ as an assignment to each $v \in \mathcal{V}_G$ and $b \in \mathcal{B}_G$ of a row vector $(\langle v\langle b | \otimes I \otimes \text{Id} | B\rangle | B\rangle)$ of dimensions $|\mathcal{V}_H||\mathcal{B}_H|$ (whose entries are again in $\mathcal{L}(\mathbb{C}^d)$). Observe that for any $v$, $|B_v\rangle = \sum_b (\langle v\langle b | \otimes I \otimes \text{Id} | B\rangle | B\rangle)$ is a quantum strategy for $H$, since for each question $v'$ for $H$, the sum over answers $b'$ of

$$(\langle v'\langle b' | \otimes \text{Id} | B_{v'}\rangle | B_{v'}\rangle = B^{b',b'}_{v,v'} = \sum_b B^{b,b'}_{v,v'}$$

is $\sum_{b'} B^{b',b'}_{v,v'} = \sum_{b'} \sum_b B^{b,b'}_{v,v'} = \text{Id}$. In particular, $\|H \otimes \text{Id} | B_v\rangle\|_2^2 \leq \|H\|_2^2$. We write

$$|A_v\rangle := \sum_b (\langle v\langle b | \otimes I \otimes \text{Id} | A\rangle | A\rangle)$$

and observe that it is equal to $H \otimes \text{Id} | B_v\rangle$; hence, it satisfies $\| |A_v\rangle\|_2 \leq \|H\|_2$ for every $v$. Thus, the ratio between
As a result of our observations, the ratio $\kappa$ can be upper bounded in a manner that depends only on $G$ and is independent of $H$. Abstracting the set $U_H \times A_H$ associated with pairs of questions and answers for the first player in $H$ as $\Omega$ for some discrete set $\Omega$,\footnote{In order for the extended inner product $\langle \cdot, \cdot \rangle_\mu$ to remain well-defined, we also need to equip $\Omega$ with a measure—here, it would be the Cartesian product of the probability measure $\mu_L$ on $U_H$ and the counting measure on $A_H$.} we are led to the definition of $\text{VAL}_+^*(G)$ as the supremum of $\|G \otimes I_\Omega \otimes \text{Id}_{C^d} |A\rangle\|_2$ ranging over vector quantum strategies $|A\rangle$ with norm $\|A\|_+ \leq 1$ defined as follows.

**Definition 3.11 (Fractional Strategy and Vector Strategy).** Let $G$ be a projection game and $\Omega$ a discrete measured space. An element

$$|A\rangle = \sum_{v,b} |v\rangle |b\rangle \otimes A^b_v \in \mathbb{C}^{[V]} \otimes \mathbb{C}^{[B]} \otimes \mathcal{L}(C^d)$$

is a fractional quantum strategy for $G$ if for every $v, b$ the matrix $A^b_v$ is positive semidefinite and $A_v := \sum_b A^b_v \leq \text{Id}$ for every $v$. A vector quantum strategy is an element

$$|A\rangle = \sum_{\omega \in \Omega} |\omega\rangle |A_\omega\rangle \in \mathbb{C}^{[\Omega]} \otimes \mathbb{C}^{[V]} \otimes \mathbb{C}^{[B]} \otimes \mathcal{L}(C^d)$$

such that each $|A_\omega\rangle$ is a fractional quantum strategy. The norm of a vector quantum strategy is defined as

$$\| |A\rangle \|_+ := (\max_v \| \mathbb{E} A^v_\omega \otimes A^v_\omega \|_\infty )^{1/2}. \tag{3.12}$$

The definition of $\text{VAL}_+^*$ is given by,

**Definition 3.13 (The relaxation $\text{VAL}_+^*$).** Let $G$ be a projection game. Then

$$\text{VAL}_+^*(G) := \sup_{\Omega} \sup_{|A\rangle \in \mathbb{C}^{[V]} \otimes \mathbb{C}^{[B]} \otimes \mathbb{C}^{[\Omega]} \otimes \mathcal{L}(C^d)} \| G \otimes I_\Omega \otimes \text{Id}_{C^d} |A\rangle \|_\square^2$$

where the supremum is taken over all discrete measured spaces $\Omega$.\footnote{In order for the extended inner product $\langle \cdot, \cdot \rangle_\mu$ to remain well-defined, we also need to equip $\Omega$ with a measure—here, it would be the Cartesian product of the probability measure $\mu_L$ on $U_H$ and the counting measure on $A_H$.}
With these definitions in place, we prove Lemma 3.5 relating the square norm of a product of games to $\text{val}_+^*$. 

**Proof (Proof of Lemma 3.5).** Let $|B\rangle$ be an optimal strategy in the square game associated with $G \otimes H$. It follows immediately from our observations above that $|A\rangle = I \otimes H \otimes \text{Id} |B\rangle$ is a vector quantum strategy for $G$ (where the space $\Omega = \mathcal{U}_H \times \mathcal{A}_H$, and the measure is the Cartesian product of the probability measure $\mu_L$ on $\mathcal{U}_H$ and the counting measure on $\mathcal{A}_H$) whose norm is $\| |A\rangle \|_+ \leq \| H \|_\square$. This means that 

$$\| G \otimes H \|_\square^2 = \| G \otimes H \otimes \text{Id} |B\rangle \|_\square^2 = \| G \otimes I \otimes \text{Id} |A\rangle \|_\square^2 \leq \text{val}_+^*(G) \cdot \| H \|_\square^2,$$

where the last inequality comes by observing that $\frac{1}{\| H \|_\square} |A\rangle$ is a vector strategy with norm $\| \cdot \|_+$ at most 1, so its value is at most $\text{val}_+^*(G)$.

Multiplicativity of $\text{val}_+^*$ follows along the same lines. First, we note that $\text{val}_+^*(G \otimes H) \geq \text{val}_+^*(G) \text{val}_+^*(H)$ is clear. To show the converse, proceed as above by first fixing an optimal vector quantum strategy $|B\rangle$ for the square game associated with $G \otimes H$, such that $\| |B\rangle \|_+ = 1$. As in the above, it is easy to see that $|A\rangle = I \otimes H \otimes \text{Id} |B\rangle$ is a vector quantum strategy for $G$ whose norm satisfies $\| |A\rangle \|_+ \leq \text{val}_+^*(H)$. Thus,

$$\text{val}_+^*(G \otimes H) = \| G \otimes H \otimes \text{Id} |B\rangle \|_\square^2 = \| G \otimes I \otimes \text{Id} |A\rangle \|_\square^2 \leq \text{val}_+^*(G) \cdot \text{val}_+^*(H),$$

proving the claim. \qed

### 4. Relating $\text{val}_+^*(G)$ to the square norm

In this section, we prove Lemma 3.3, which states that $\text{val}_+^*(G)$ is a good relaxation of the square norm $\| G \|_\square$ of a projection game and establishes the last step in our proof of the parallel repetition theorem, Theorem 1.1. We will also show that if $G$ is an expanding projection game then one can take $c = 1$ in the bound $\text{val}_+^*(G) \leq 1 - C(1 - \| G \|_\square^2)^c$. 


To prove the lemma, we need to show that the existence of a good vector strategy for the players Bob and Bob’ in the square game $G^\dagger G$ implies that $\|G\|_2^2$ is large; i.e., there also exists a good (standard) quantum strategy for the players Alice and Bob in $G$. We will establish this by describing an explicit rounding procedure mapping the former to the latter. The rounding argument is simpler in case $G$ has the additional property of being expanding (see Section 2.2 for the definition), and we give the proof in that case in Section 4.1. In Section 4.2, we treat the case of general projection games. In that case, the rounding argument is more involved and relies on a “quantum correlated sampling” lemma which is stated and proved in Section 5.

In both cases, the starting point for the rounding procedure is the existence of a vector strategy $|\hat{A}\rangle$ and entangled state $|\hat{\Psi}\rangle$ satisfying inequality (4.2) in the following claim, which is essentially a restatement of the inequality “$\text{VAL}^*_+(G) \geq 1 - \eta$.”

**Claim 4.1.** Let $G$ be a projection game and $\eta > 0$ such that $\text{VAL}^*_+(G) \geq 1 - \eta$. Then, there exists a discrete measured space $\Omega$, an integer $d$, a bipartite state $|\hat{\Psi}\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ and a vector strategy $|\hat{A}\rangle \in \mathcal{C}[\Omega] \otimes \mathcal{C}[V] \otimes \mathcal{C}[B] \otimes \mathcal{L}(\mathbb{C}^d)$ such that for every $\omega$ and $v,b$, $\hat{A}_{\omega v} \geq 0$ and $\hat{A}_{\omega v} = \sum_b \hat{A}_{\omega v}^b \leq \text{Id}$, and

$$
\frac{E}{\omega} \sum_{v \sim v'} \sum_{b \rightarrow b'} \langle \hat{\Psi} | \hat{A}_{\omega v}^b \otimes \hat{A}_{\omega v'}^{b'} | \hat{\Psi} \rangle 
\geq (1 - \eta) \max_v \left\{ \frac{E}{\omega} \langle \hat{\Psi} | \hat{A}_{\omega v} \otimes \hat{A}_{\omega v} | \hat{\Psi} \rangle \right\},
$$

where formally $\sum_{v \sim v'} \sum_{b \rightarrow b'}$ is shorthand for

$$
\sum_u \mu(u) \sum_a \sum_{v,v'} \mu(v|u) \mu(v'|u) \sum_{b \rightarrow a,b' \rightarrow a}.
$$

Furthermore, without loss of generality $|\hat{\Psi}\rangle$ can be chosen so as to have the following symmetry: Its reduced densities on either subsystem are identical, and denoting either by $\hat{\rho}$, for any $X,Y$ it holds that

$$
\langle \hat{\Psi} | X \otimes Y | \hat{\Psi} \rangle = \text{Tr} (X \hat{\rho}^{1/2} Y \hat{\rho}^{1/2}).
$$
Proof. By definition of $\text{val}_+^*$, there exists a discrete measured space $\Omega$ and a vector strategy $|\hat{A}\rangle$ such that $|||\hat{A}|||_+ = 1$ and $\|\text{Id}_\Omega \otimes G \otimes \text{Id} |\hat{A}\rangle\|_2^2 \geq 1 - \eta$. Recalling the definition of $|||\cdot|||_+$ (see Definition 3.11) and of $|||\cdot|||_\infty$ (see Definition 3.8), we may reformulate this statement as the inequality

$$
(4.4) \quad \left\| E_{\omega \sim v} \sum_{b \leftrightarrow b'} \overline{\hat{A}^b_{\omega v}} \otimes \hat{A}^{b'}_{\omega v'} \right\|_{\infty} \geq (1 - \eta) \max_v \left\| E_{\omega \sim v} \overline{\hat{A}_{\omega v}} \otimes \hat{A}_{\omega v} \right\|_{\infty}.
$$

Letting $|\hat{\Psi}\rangle$ be a state which achieves the operator norm on the left-hand side gives (4.2). The fact that $|\hat{\Psi}\rangle$ can be assumed to take the claimed form follows from the symmetry of the left-hand side of (4.4).

Let $|\hat{A}\rangle$ be a vector strategy and $|\hat{\Psi}\rangle$ a state such that (4.2) holds. Our goal is to identify a quantum strategy $|\tilde{A}\rangle$ such that $\|G \otimes \text{Id} |\tilde{A}\rangle\|_2^2 \geq 1 - O(\eta^{1/c})$, which by Claim 4.1 will suffice to prove Lemma 3.3. The “rounding procedure” constructing $|\tilde{A}\rangle$ will differ in the expanding and non-expanding cases. Both cases, however, build on the same measurement operators which we now define.

Fix an arbitrary $\omega \in \Omega$. The only “defect” of $|\hat{A}_\omega\rangle$ that prevents it from directly giving us a quantum strategy is that it is only a fractional strategy, meaning that for any question $v$ the sum $\hat{A}_{\omega v} = \sum_b \hat{A}^b_{\omega v}$ may not equal the identity. It is natural to define a re-normalized strategy as follows. Let $U_{\omega v}$ be a unitary such that

$$
(4.5) \quad U_{\omega v} A_{\omega v}^{1/2} \rho_{1/4}^{\omega v} U_{\omega v}^\dagger = (\rho_{1/4}^{\omega v} A_{\omega v} \rho_{1/4}^{\omega v})^{1/2}
$$

is Hermitian positive semidefinite; such a unitary can be obtained from the singular value decomposition of $A_{\omega v}^{1/2} \rho_{1/4}^{\omega v}$. For every pair of questions $v, v' \in \mathcal{V}$, we introduce the post-measurement state

$$
(4.6) \quad |\Psi_{\omega vv'}\rangle := \overline{U_{\omega v} \hat{A}_{\omega v}}^{1/2} \otimes U_{\omega v'} \hat{A}_{\omega v'}^{1/2} |\hat{\Psi}\rangle.
$$

The state $|\Psi_{\omega vv'}\rangle$ is the post-measurement state that corresponds to applying the binary measurements $\{\overline{A}_{\omega v}, \text{Id} - \overline{A}_{\omega v}\}$ for the first player, $\{\hat{A}_{\omega v'}, \text{Id} - \hat{A}_{\omega v'}\}$ for the second, to $|\hat{\Psi}\rangle$ and conditioning on
both of them obtaining the first outcome. In general, the post-measurement state is only defined up to a local unitary, and this freedom is represented in the unitaries $U_\omega$ and $U_{\omega'}$; our particular choice of unitaries satisfying (4.5) will prove convenient in the analysis. Next for every question $v \in V$ and answer $b \in B$ we define the measurement operator

$$
\tilde{A}^b_{\omega v} := U_\omega \hat{A}^{1/2}_{\omega v} \hat{A}^b_{\omega v} \hat{A}^{1/2}_{\omega v} U_\omega^\dagger,
$$

where here $\hat{A}^{-1/2}_{\omega v}$ denotes the square root of the pseudo-inverse of $\hat{A}^b_{\omega v} = \sum_b \hat{A}^b_{\omega v}$. Again, there is always a unitary degree of freedom in the choice of the square root, and the unitaries $U_\omega$, the same as in (4.6), represent that degree of freedom. With this definition, it is easy to verify that each $\tilde{A}^b_{\omega v}$ is positive semidefinite and that $\sum_b \tilde{A}^b_{\omega v} \leq \text{Id}$; since we may always add a “dummy” outcome in order for the measurement operators to sum to identity, $\{\tilde{A}^b_{\omega v}\}_b$ is easily extended into a well-defined measurement and $|\tilde{A}\rangle := \sum_{v,b} |v,b\rangle \otimes \tilde{A}^b_{\omega v}$ is a valid quantum strategy in $G^\dagger G$.

Now suppose that, upon receiving their respective questions $v$ and $v'$, players Bob and Bob’ in $G^\dagger G$ were to measure their respective share of the (re-normalized) state $|\Psi_{\omega vv'}\rangle$ using the measurements given by the $\{\tilde{A}^b_{\omega v}\}_b$, $\{\tilde{A}^b_{\omega v'}\}_b'$, respectively. The probability that they obtain the pair of outcomes $(b, b')$ is given, up to normalization by $|||\Psi_{\omega vv'}|||^2$, by

$$
\langle \Psi_{\omega vv'} | \tilde{A}^b_{\omega v} \otimes \tilde{A}^{b'}_{\omega v'} | \Psi_{\omega vv'} \rangle \\
= \langle \tilde{\Psi} | (\hat{A}^{1/2}_{\omega v} U_{\omega v'}^\dagger \otimes \hat{A}^{1/2}_{\omega v'} U_{\omega v}^\dagger) (U_{\omega v} \hat{A}^{1/2}_{\omega v} \hat{A}^b_{\omega v} \hat{A}^{1/2}_{\omega v} U_{\omega v}^\dagger \\
\otimes U_{\omega v'} \hat{A}^{1/2}_{\omega v'} \hat{A}^{b'}_{\omega v'} \hat{A}^{1/2}_{\omega v'} U_{\omega v'}^\dagger) (U_{\omega v'} \hat{A}^{1/2}_{\omega v'} \otimes U_{\omega v'} \hat{A}^{1/2}_{\omega v'}) | \tilde{\Psi} \rangle \\
(4.8) = \langle \hat{\Psi} | \hat{A}^b_{\omega v} \otimes \hat{A}^{b'}_{\omega v'} | \hat{\Psi} \rangle,
$$

perfectly reproducing the correlations induced by the fractional strategy $|\hat{A}_{\omega}\rangle$ together with $|\tilde{\Psi}\rangle$. Thus, if it were the case that for all $(v, v')$, $|\Psi_{\omega vv'}\rangle = |\Psi_{\omega}\rangle$, a vector independent of $(v, v')$, then the players could use $|\Psi_{\omega}\rangle$ (for an appropriate, “good” choice of $\omega$) as their initial shared entangled state and perfectly emulate $|\hat{A}_{\omega}\rangle$ using the quantum strategy $|\tilde{A}\rangle$. 
While it may unfortunately not be the case that the $|\Psi_{\omega vv'}\rangle$ are independent of $(v, v')$, the main claim in the proof of Lemma 3.3 will establish that they are close, on average, when $v$ and $v'$ are neighboring vertices in the constraint graph. In the case where the game is expanding, this will be sufficient, as we will be able to conclude that all states $|\Psi_{\omega vv'}\rangle$ are close to a single $|\Psi_{\omega}\rangle$ independent of $v$ and $v'$. In the non-expanding case, we will rely on a more complicated strategy that involves a step of correlated sampling in which the players, after having received their respective $v$ and $v'$, jointly sample an $\omega$ and create the corresponding bipartite state $|\Psi_{\omega vv'}\rangle$ locally.

We first turn to the case of expanding games, for which we can give a simpler (and tighter) analysis.

4.1. The expanding case. Suppose that $G$ is expanding. Our first step consists in fixing a “good” value $\omega \in \Omega$ and restricting our attention to the fractional strategy $|\hat{A}_{\omega}\rangle := (|\omega\rangle \otimes I \otimes \text{Id}) |\hat{A}\rangle$ specified by the operators $\hat{A}^b_{\omega v}$ obtained from that $\omega$. Using that the max is larger than the average, Eq. (4.2) implies

$$\mathbb{E}_{\omega} \left( \mathbb{E}_{v \sim v'} \sum_{b \leftrightarrow b'} \langle \hat{\Psi} | \hat{A}^b_{\omega v} \otimes \hat{A}^{b'}_{\omega v'} | \hat{\Psi} \rangle \right) \geq (1 - \eta) \mathbb{E}_{\omega} \left( \mathbb{E}_{v} \langle \hat{\Psi} | \hat{A}_{\omega v} \otimes \hat{A}_{\omega v} | \hat{\Psi} \rangle \right).$$

For the remainder of this section, fix an $\omega$ such that (4.9) holds for that $\omega$. The only property we will need of the $\{\hat{A}^b_{\omega v}\}$ in order to construct a good strategy in $G^\dagger G$ is that they are positive semidefinite operators which satisfy that inequality. (In contrast, for the non-expanding case, Eq. (4.9) by itself turns out to be too weak an inequality, and we must work with (4.2).)

Having fixed a value for $\omega$, for clarity of notation for every $v, v'$ we let $U_v$ be the unitary defined by (4.5), $|\Psi_{vv'}\rangle$ the state defined in (4.6), and $\hat{A}^b_v$ the measurement operators introduced in (4.7). Let

$$\sigma := \left( \mathbb{E}_{w} \| |\Psi_{ww}\rangle \|^2 \right)^{-1} \mathbb{E}_{w} |\Psi_{ww}\rangle \langle \Psi_{ww}|.$$

The operators $\hat{A}^b_v$, together with the density matrix $\sigma$, form a well-defined strategy for the players in the square game. In order to
prove Lemma 3.3 (for the case of expanding games), it remains to bound the error

\[(4.11) \quad \varepsilon := \mathbb{E}_{v \sim v'} \sum_{b \leftrightarrow b'} \text{Tr} \left( \tilde{A}_v^b \otimes \tilde{A}_{v'}^{b'} \sigma \right), \]

incurred by that strategy, under the assumption that (4.9) holds. We show the following.

**Claim 4.12.** Suppose (4.9) holds, and the constraint graph $G$ is such that the smallest nonzero eigenvalue of the Laplacian $L := \sum_v \langle v | - \sum_{v,v' \sim v} \mu(v,v') \mu(v)^{-1/2} \mu(v')^{-1/2} | v \rangle \langle v |$ is at least $\lambda > 0$, where here $\mu(v,v')$ is the distribution on questions in the square game, as defined in Section 2.2. Let $(\tilde{A}_v^b, \sigma)$ be the strategy defined above. Then,

\[(4.13) \quad \varepsilon = \mathbb{E}_{v \sim v'} \sum_{b \leftrightarrow b'} \text{Tr} \left( \tilde{A}_v^b \otimes \tilde{A}_{v'}^{b'} \sigma \right) = O(\eta/\lambda). \]

Before proceeding with the proof of Claim 4.12, we show that it implies Lemma 3.3.

**Proof** (Proof of Lemma 3.3, expanding case). Let $\eta > 0$ be such that $\text{VAL}_+^*(G) \geq 1 - \eta$. Then, it follows directly from Claim 4.1 that (4.9) holds for this choice of $\eta$. Let $(\tilde{A}_v^b, \sigma)$ be as defined in (4.7) and (4.10). By definition,

\[
\|G \otimes \text{Id} | \tilde{A} \|^2_\infty = \sup_{|\Psi\rangle} \mathbb{E}_{v \sim v'} \sum_{b \leftrightarrow b'} \langle \Psi | \tilde{A}_v^b \otimes \tilde{A}_{v'}^{b'} | \Psi \rangle \\
\geq \mathbb{E}_{v \sim v'} \sum_{b \leftrightarrow b'} \text{Tr} \left( (\tilde{A}_v^b \otimes \tilde{A}_{v'}^{b'}) \sigma \right) \\
= 1 - O(\eta/\lambda),
\]

where the last line follows from (4.13). Using $\|G\|_\infty^2 \geq \|G \otimes \text{Id} | \tilde{A} \|^2_\infty$ concludes the proof of the lemma, with exponent $c = 1$. \qed

It remains to prove Claim 4.12. The proof of the claim will use the expansion properties of $G$ through the following:
Claim 4.14. Suppose (4.9) holds, and $G$ is such that the smallest nonzero eigenvalue of the Laplacian
\[ L := \sum_v |v\rangle\langle v| - \sum_{v,v' \sim v} \mu(v,v')\mu(v)^{-1/2}\mu(v')^{-1/2}|v\rangle\langle v| \]
is at least $\lambda > 0$. Then,
\[ (4.15) \quad E_{v,v'} \langle \hat{\Psi}|\hat{A}_{\omega v} \otimes \hat{A}_{\omega v'}|\hat{\Psi}\rangle \geq (1 - 2\eta / \lambda) E_v \langle \hat{\Psi}|\hat{A}_{\omega v} \otimes \hat{A}_{\omega v}|\hat{\Psi}\rangle. \]

Proof. Using (4.3), we can write
\[ \langle \hat{\Psi}|\hat{A}_{\omega v} \otimes \hat{A}_{\omega v'}|\hat{\Psi}\rangle = \text{Tr}\left( \hat{A}_{\omega v}\rho^{1/2}\hat{A}_{\omega v'}\rho^{1/2} \right), \]
where $\rho$ is the reduced density of $|\hat{\Psi}\rangle$ on either subsystem. Let $\tilde{L} := L \otimes \text{Id}$ and $A := \sum_v \mu(v)^{1/2}|v\rangle \otimes \hat{A}_{\omega v}$. Using that (4.9) holds for our choice of $\omega$,
\[ \text{Tr}(A^\dagger \tilde{L}(\text{Id} \otimes \rho^{1/2})A(\text{Id} \otimes \rho^{1/2})) = E_v \text{Tr}(\hat{A}_{\omega v}\rho^{1/2}\hat{A}_{\omega v'}\rho^{1/2}) - E_{v,v'} \text{Tr}(\hat{A}_{\omega v}\rho^{1/2}\hat{A}_{\omega v'}\rho^{1/2}) \]
\[ \leq \eta E_v \text{Tr}(\hat{A}_{\omega v}\rho^{1/2}\hat{A}_{\omega v'}\rho^{1/2}). \]
\[ (4.16) \]

The normalized Laplacian $L$ has smallest eigenvalue 0, and second smallest $\lambda > 0$. Let the smallest eigenvector of $L$ be $|u_0\rangle = \sum_v \mu(v)^{1/2}|v\rangle$, and write $A = |u_0\rangle \otimes A_0 + \sum_{i>0} |u_i\rangle \otimes A_i$, where the $|u_i\rangle$ are the remaining eigenvectors, with associated eigenvalue $\lambda_i$, of $\tilde{L}$, and $A_0 = \sum_v \mu(v)^{1/2}\hat{A}_{\omega v}$. Then
\[ (4.17) \quad A^\dagger \tilde{L}(\text{Id} \otimes \rho^{1/2})A(\text{Id} \otimes \rho^{1/2}) = \sum_{i>0} \lambda_i A_i^\dagger \rho^{1/2} A_i \rho^{1/2}. \]
Taking the trace, we get
\[ E_v \text{Tr}\left( (\hat{A}_{\omega v} - E_v \hat{A}_{\omega v'})\rho^{1/2}(\hat{A}_{\omega v} - E_v \hat{A}_{\omega v'})\rho^{1/2} \right) \]
\[ = \text{Tr}\left( (A - |v_0\rangle \otimes A_0)^\dagger \rho^{1/2}(A - |v_0\rangle \otimes A_0)\rho^{1/2} \right) \]
\[ = \sum_{i>0} \text{Tr}(A_i^\dagger \rho^{1/2} A_i \rho^{1/2}) \]
\[ \leq \frac{\eta}{\lambda} E_v \text{Tr}(\hat{A}_{\omega v}\rho^{1/2}\hat{A}_{\omega v'}\rho^{1/2}), \]
where the last inequality follows from (4.17) and (4.16). \qed
We conclude this section by giving the proof of Claim 4.12.

PROOF (Proof of Claim 4.12). For any four vertices $v, v', w$ and $w'$ define

$$
\varepsilon_{ww'} := \sum_{b \leftrightarrow b'} \langle \Psi_{ww'} | \bar{A}_v^b \otimes \bar{A}_{v'}^{b'} | \Psi_{ww'} \rangle.
$$

Note also that, given our choice of the unitaries $U_v$ satisfying (4.5) and using (4.3),

$$
\langle \Psi_{ww'} | \tilde{A}_v^b \otimes \tilde{A}_{v'}^{b'} | \Psi_{ww'} \rangle = \operatorname{Tr}(\rho_{1/2}(\hat{A}_{w'}^{1/2}\hat{A}_v^{1/2}\rho_{1/2}(\hat{A}_w^{1/2}\hat{A}_{v'}^{1/2})),
$$

an identity that will prove useful.

By (4.11) and the definition of $\sigma_{wv}$, we have

$$
\varepsilon = (E_{w} |||\Psi_{ww'}|||^2)^{-1} \sum_{v \sim v'} E_{w} \varepsilon_{vv'}.
$$

To prove the claim, it will suffice to show that $\varepsilon = O(\eta/\lambda E_{w} |||\Psi_{ww'}|||^2)$. Eq. (4.9) implies that

$$
E_{v} |||\Psi_{vv'}|||^2 \geq (1 - \eta) E_{v} |||\Psi_{vv'}|||^2,
$$

hence (using (4.9) once more) $E_{v} \varepsilon_{vv'} = O(\eta E_{w} |||\Psi_{ww'}|||^2)$. We relate these quantities by establishing the following three bounds.

$$
E_{v} \varepsilon_{vv'} = O(\eta E_{w} |||\Psi_{ww'}|||^2),
$$

$$
E_{v} \varepsilon_{vv'} = O(\eta E_{w} |||\Psi_{ww'}|||^2),
$$

$$
E_{v} \varepsilon_{vv'} = O(\eta E_{w} |||\Psi_{ww'}|||^2).
$$

It is clear that (4.20) and (4.21) together will conclude the proof. We first show (4.19). Using (4.18),

$$
|\varepsilon_{vv'} - \tilde{A}_v^{1/4} \tilde{A}_{v'}^{1/4} | \leq \left( \operatorname{Tr}(\rho_{1/2}(\hat{A}_w^{1/2} \hat{A}_{v'}^{1/2})) \right)^{1/2}
$$
where the inequality follows from applying the Cauchy-Schwarz inequality to
\[
(\tilde{A}_{v'})^{1/2} \rho^{1/4} (\hat{A}_{v'} - \hat{A}_v) \rho^{1/4} (\tilde{A}_v)^{1/2} \quad \text{and} \quad (\tilde{A}_v)^{1/2} \rho^{1/4} \hat{A}_v \rho^{1/4} (\tilde{A}_{v'})^{1/2}
\]
and using \(\sum_b \tilde{A}_b \leq \text{Id.}\) Taking the expectation over \(v \sim v'\) and using (4.9) together with \(|x - y| \leq \sqrt{ax} \implies x \leq a + 4y\) gives (4.19). To prove (4.20), write
\[
|\varepsilon_{vv'} - \varepsilon_{wv'}| = \left| \sum_{b \leftrightarrow b'} \text{Tr}(\hat{A}_{v'} \rho^{1/4} \tilde{A}_v \rho^{1/4} (\hat{A}_v - \hat{A}_w) \rho^{1/4} \tilde{A}_{v'} \rho^{1/4}) \right|
\leq \left( \text{Tr}((\hat{A}_v - \hat{A}_w) \rho^{1/2}) \rho^{1/2} \right)^{1/2}
\cdot \left( \sum_{b \leftrightarrow b'} \text{Tr}(\hat{A}_{v'} \rho^{1/4} \tilde{A}_v \rho^{1/4} \hat{A}_{v'} \rho^{1/4} \tilde{A}_{v'} \rho^{1/4}) \right)^{1/2},
\]
where the inequality follows from a similar application of the Cauchy-Schwarz inequality as performed above. The second term above is \(\varepsilon_{vv'}\), so using (4.19), and (4.15) to bound the first term, we have proved (4.20). Finally, to prove (4.21) write
\[
|\varepsilon_{vv'} - \varepsilon_{wv'}| = \left| \sum_{b \leftrightarrow b'} \text{Tr}((\hat{A}_{v'} - \hat{A}_{w'}) \rho^{1/4} \tilde{A}_v \rho^{1/4} \hat{A}_v \rho^{1/4} \tilde{A}_{v'} \rho^{1/4}) \right|
\leq \left( \text{Tr}((\hat{A}_{v'} - \hat{A}_{w'}) \rho^{1/2}) \rho^{1/2} \right)^{1/2}
\cdot \left( \sum_{b \leftrightarrow b'} \text{Tr}(\hat{A}_v \rho^{1/4} \tilde{A}_v \rho^{1/4} \hat{A}_v \rho^{1/4} \tilde{A}_{v'} \rho^{1/4}) \right)^{1/2},
\]
which is again bounded using (4.19) and (4.15). Combining (4.20) and (4.21) proves the claim. □

4.2. Non-expanding games. Suppose \(G\) is an arbitrary (not necessarily expanding) projection game. In the game \(G^\dagger G\), the players Bob and Bob’ are always sent neighboring \(v \sim v'\). Using notation from the previous section, we would like to enable the players to take advantage of the possibility of using an arbitrary entangled state in order to initialize themselves in a state that is close to \(|\Psi_{vv'}\rangle\). The difficulty is that this must be done “on the
fly,” as \(|\Psi_{vv'}\rangle\) depends on the questions \(v, v'\); indeed, since \(G\) is not expanding, there may not be a single state close to all \(|\Psi_{vv'}\rangle\) that they could have agreed upon before the start of the game; for instance, \(G\) could be a direct sum of two independent games for which the optimal entangled state and measurements need not bear any relation to each other.

To get around this, we resort to the use of a so-called family of “universal embezzling states” \(|\Gamma_d\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d\). These states, introduced in (van Dam & Hayden 2003), have the property that for any given state \(|\psi\rangle\), there exists a \(d\) and unitaries \(U, V\) such that \(U \otimes V|\Gamma_d\rangle \approx |\psi\rangle|\Gamma_d'\rangle\) for some \(d'\). Hence, if both players have a description of the target state \(|\Psi_{vv'}\rangle\) they can easily generate it locally from the universal state \(|\Gamma_d\rangle\). Our setting presents and additional difficulty: Only the first player, Bob, knows \(v\), and the second, Bob’, knows \(v'\); how to make them agree on which state to embezzle? We will use the following lemma.

**Lemma 4.22.** Let \(|\Phi\rangle\) be a bipartite state invariant under permutation of the two subsystems, \(\rho\) its reduced density on either subsystem, \(0 \leq A_v \leq \text{Id}\), and \(\nu\) a distribution on \(V \times V\) that is symmetric under permutation of the two coordinates (we also denote by \(\nu\) the marginal distribution on either coordinate), such that

\[
E_{(v,v') \sim \nu} \langle \Phi | A_v \otimes A_{v'} | \Phi \rangle \geq (1 - \eta) E_{v \sim \nu} \langle \Phi | A_v \otimes A_v | \Phi \rangle.
\]

Let \(U_v\) be unitaries such that

\[
U_v A_{v}^{1/2} \rho^{1/4} A_{v}^{1/2} U_v^\dagger = (\rho^{1/4} A_v \rho^{1/4})^{1/2},
\]

and let

\[
|\Phi_{vv'}\rangle := U_v A_{v}^{1/2} \otimes U_{v'} A_{v'}^{1/2} |\Phi\rangle.
\]

Then for any \(v'' \in V\),

\[
E_{v \sim v'} \| |\Phi_{vv''}\rangle - |\Phi_{v'v''}\rangle \|^2 = O(\eta^{1/2}) \left( E_v \| |\Phi_{vv'}\rangle \|^2 \right)^{1/2} \| |\Phi_{v''v''}\rangle \|,
\]

and

\[
E_{v \sim v'} \| |\Phi_{vv}\rangle - |\Phi_{v'v}\rangle \|^2 = O(\eta^{1/2}) E_v \| |\Phi_{vv}\rangle \|^2.
\]
The lemma is stated in a stand-alone form, but we may apply it to the present setting by letting $|\Phi\rangle$ be the state $|\tilde{\Psi}\rangle$ and $A_v$ the measurement operators $\hat{A}_{\omega v}$ (for some $\omega$) whose existence is guaranteed by Claim 4.1. Recalling the definition of $|\Psi_{\omega v v'}\rangle$ in (4.6), Lemma 4.22 (together with (4.2) to obtain (4.23)) implies that (for most $\omega$)

$$E_v \| |\Psi_{\omega v v'}\rangle - |\Psi_{\omega vv}\rangle \|^2 = O(\eta^{1/2}) E_v \| |\Psi_{\omega vv}\rangle \|^2$$

; that is, all three states $|\Psi_{\omega v v'}\rangle$, $|\Psi_{\omega vv}\rangle$, and $|\Psi_{\omega v' v'}\rangle$ are close for neighboring $v \sim v'$. Hence, the first player, knowing his question $v$, can compute a classical description of the state $|\Psi_{\omega vv}\rangle$; the second player can compute a classical description of $|\Psi_{\omega v' v'}\rangle$. These two states are close to each other as well as to the target state: Are these conditions sufficient for the two players to successfully embezzle a joint state close to either of the three?

It turns out that, if one naively applies the embezzling procedure described in (van Dam & Hayden 2003), it can fail completely even when the states are arbitrarily close (see Section 5 for an example). Nevertheless, in the next section we state and prove a “quantum correlated sampling lemma,” which extends the results in (van Dam & Hayden 2003) to this “approximate” scenario.

We first prove Lemma 4.22, and then show how the lemma, together with the correlated sampling lemma, Lemma 5.1, imply Lemma 3.3 for the case of general games.

**Proof (Proof of Lemma 4.22).** Let $X_v$ be defined as

$$X_v := U_v A_v^{1/2} \rho^{1/4} = \rho^{1/4} A_v^{1/2} U_v^\dagger.$$  

Using (4.24), $X_v$ is positive semidefinite. With this notation, we have the following useful identities.

**Claim 4.26.** For every $v, v' \in \mathcal{V}$, we have

$$\text{Tr}(X_v^4) = \text{Tr}((X_v X_v^\dagger)^2) = \langle \Phi | \overline{A_v} \otimes A_v | \Phi \rangle = \| |\Phi_{vv}\rangle \|^2$$

and

$$\text{Tr}(X_v^2 X_{v'}^2) = \langle \Phi | \overline{A_v} \otimes A_{v'} | \Phi \rangle.$$
Proof. For (4.27), we use the definition of $X_v$ to write

$$\text{Tr}(X_v^4) = \text{Tr}((X_vX_v^\dagger)^2) = \text{Tr}(A_v\rho^{1/2}A_v\rho^{1/2}) = \langle \Phi | \overline{A_v} \otimes A_v | \Phi \rangle,$$

where the last equality follows from Ando’s identity, Claim 2.1, together with our assumption on $|\Phi\rangle$ being permutation-invariant.

To show (4.28), expand using the definition (4.25)

$$\text{Tr}(X_v^2X_{v'}^2) = \text{Tr}(U_vA_v^{1/2}\rho^{1/2}A_v^{1/2}U_v^\dagger U_{v'}A_{v'}^{1/2}\rho^{1/2}A_{v'}^{1/2}U_{v'}^\dagger) = \text{Tr}(A_v\rho^{1/2}A_{v'}\rho^{1/2}) = \langle \Phi | \overline{A_v} \otimes A_{v'} | \Phi \rangle,$$

where the second equality follows from (4.24) and the last follows from Claim 2.1. □

Now for any three $v, v', v''$,

$$\|\langle \Phi_{vv''} - |\Phi_{v'v''}\rangle\|^2 = \left(\langle \Phi_{vv''} | - \langle \Phi_{v'v''} | \right) \left( |\Phi_{vv''}\rangle - |\Phi_{v'v''}\rangle \right)$$

$$= \langle \Phi | \left(A_v^{1/2}U_v^\dagger - A_{v'}^{1/2}U_{v'}^\dagger \right) \left(U_vA_v^{1/2} - U_{v'}A_{v'}^{1/2} \right) \otimes A_{v''}^{1/2}U_{v''}^\dagger U_{v''}A_{v''}^{1/2} | \Phi \rangle$$

$$= \text{Tr}((X_v - X_{v'})^\dagger(X_v - X_{v'})X_{v''}^\dagger X_{v''})$$

$$\leq \left(\text{Tr}((X_v - X_{v'})^4)\right)^{1/2} \left(\text{Tr}(X_{v''}^4)\right)^{1/2},$$

(4.29)

where the last inequality follows from Cauchy-Schwarz and the fact that the $X_v$ are positive semidefinite. The first term on the right-hand side of (4.29) can be bounded as

$$\text{Tr}((X_v - X_{v'})^4) \leq \text{Tr}((X_v^2 - X_{v'}^2)^2)$$

$$= \langle \Phi | \overline{A_v} \otimes A_v | \Phi \rangle + \langle \Phi | \overline{A_{v'}} \otimes A_{v'} | \Phi \rangle$$

$$- 2\langle \Phi | \overline{A_v} \otimes A_{v'} | \Phi \rangle,$$

where the first inequality can be found as, e.g., Corollary 2 in (Kittaneh 1986) and the equality follows from (4.27) and (4.28). Going back to (4.29), we obtain

$$E_{v \sim v'} \|\langle \Phi_{vv''} - |\Phi_{v'v''}\rangle\|^2 \leq \left(2\eta \ E_{v} \|\langle \Phi_{vv} | \|^2\right)^{1/2} \left(\|\langle \Phi_{v''v''} | \|^2\right)^{1/2},$$
where the first inequality uses the assumption made in the lemma to bound the first term in (4.29) and (4.27) to rewrite the second. This proves the first inequality claimed in the lemma. The second is obtained by taking $v'' = v$ in (4.29), and then the expectation over $v \sim v'$ as in the above. \qed

We conclude this section with the proof of Lemma 3.3.

**Proof (Proof of Lemma 3.3, general case).** Let $|\hat{A}\rangle$ be a vector strategy, and $|\hat{\Psi}\rangle$ a state such that (4.2) holds. Our goal is to identify a quantum strategy $|\tilde{A}\rangle$ such that $\|G \otimes \text{Id} |\tilde{A}\rangle\|_G^2 \geq 1 - O(\eta^{1/\epsilon})$, which by Claim 4.1 will suffice to prove Lemma 3.3.

We define a “re-normalized” vector strategy $|\tilde{A}\rangle$ from which we will later obtain a quantum strategy $|\tilde{A}_\omega\rangle$ by making a good choice of $\omega \in \Omega$. As previously, for every $\omega$ we may define states $|\Psi_{\omega vv}\rangle := U_{\omega v} \hat{A}_{\omega v}^{-1/2} \otimes U_{\omega v'} \hat{A}_{\omega v'}^{-1/2} |\hat{\Psi}\rangle$, (4.30) where the $U_{\omega v}$ are the unitaries given by Lemma 4.22: As a consequence of (4.2) (replacing the max on the right-hand-side by an average), the assumption of the lemma is satisfied, on average over $\omega \in \Omega$, for the states $|\Psi_{\omega vv}\rangle$. The lemma gives the following bound:

$$E_{\omega} E_{v \sim v'} \| |\Psi_{\omega vv}\rangle - |\Psi_{\omega vv'}\rangle\|^2 = O(\eta^{1/2}) E_{\omega} E_{v} \| |\Psi_{\omega vv}\rangle\|^2.$$  

(4.31)

In addition, for every $\omega$ and question $v \in \mathcal{V}$ let $V_{\omega v}$ and $W_{\omega v}$ be the unitaries that are defined in Lemma 5.1, for the (re-normalized) state $|\Psi_{\omega vv}\rangle$ and a choice of $\delta = \eta^2$. By convexity, the lemma gives us that

$$E_{\omega} E_{v \sim v'} \| V_{\omega v} \otimes W_{\omega v'} |\Gamma_{dd'}\rangle - \| |\Psi_{\omega vv}\rangle^{-1} |\Psi_{\omega vv'}\rangle |\Gamma_{d'}\rangle\|^2$$

(4.32) \[ = O \left( E_{v \sim v'} \left( \frac{|\Psi_{\omega vv}\rangle}{\| |\Psi_{\omega vv}\rangle\|} - \frac{|\Psi_{\omega vv'}\rangle}{\| |\Psi_{\omega vv'}\rangle\|} \right)^{2/6} \right). \]

For any question $v \in \mathcal{V}$ and answer $b \in \mathcal{B}$, define measurement operators

$$\tilde{A}_{\omega v}^b := V_{\omega v}^\dagger \left( U_{\omega v} \hat{A}_{\omega v}^{-1/2} \hat{A}_{\omega v}^b \hat{A}_{\omega v}^{-1/2} \otimes \text{Id}_{d'} \right) V_{\omega v},$$

$$\tilde{B}_{\omega v}^b := W_{\omega v}^\dagger \left( U_{\omega v} \hat{A}_{\omega v}^{-1/2} \hat{A}_{\omega v}^b \hat{A}_{\omega v}^{-1/2} \otimes \text{Id}_{d'} \right) W_{\omega v}.$$
It is easy to verify that each \( \tilde{A}_{\omega v} \) and \( \tilde{B}_{\omega w} \) is positive semidefinite and that \( \sum_b \tilde{A}_{\omega v} \leq \text{Id} \). Since we may always add a “dummy” outcome in order for the measurement operators to sum to identity, both \( \{ \tilde{A}_{\omega v} \}_b \) and \( \{ \tilde{B}_{\omega v} \}_b \) are easily made into well-defined measurements, and for every \( \omega \), \( |\tilde{A}_{\omega} \rangle := \sum_{v,b} |v, b \rangle \otimes \tilde{A}_{\omega v} \) and \( |\tilde{B}_{\omega} \rangle := \sum_{v,b} |v, b \rangle \otimes \tilde{B}_{\omega v} \) valid strategies for the players Bob and Bob’ in \( G^G \) (we will soon show that at least one of these strategies must be a good strategy for the square game).

We first bound

\[
\begin{align*}
E_{(\omega, v \sim v')} & E_{\omega} |\tilde{V}_{\omega v} \otimes W_{\omega v'}|_{\Gamma_{dd'}} - |||\Psi_{\omega vv'}\rangle\rangle - 1||\Psi_{\omega vv'}\rangle\rangle |\Gamma_{dd'}\rangle\rangle^2 \\
\leq & E_{(\omega, v \sim v')} E_{\omega} |\tilde{V}_{\omega v} \otimes W_{\omega v'}|_{\Gamma_{dd'}} - |||\Psi_{\omega vv'}\rangle\rangle - 1||\Psi_{\omega vv'}\rangle\rangle |\Gamma_{dd'}\rangle\rangle^2 \\
& + E_{(\omega, v \sim v')} E_{\omega} ||\Psi_{\omega vv'}\rangle - 2||\Psi_{\omega vv'}\rangle - ||\Psi_{\omega vv'}\rangle||^2 \\
= & O\left( E_{(\omega, v \sim v')} E_{\omega} \left|\left|\Psi_{\omega vv'}\rangle\rangle - 1/6 - 1/6\right|\left|\Psi_{\omega vv'}\rangle\rangle\right.\right) + O(\eta^{1/2}) \\
= & O\left( E_{(\omega, v \sim v')} E_{\omega} ||\Psi_{\omega vv'}\rangle - 1/6 - 1/6||\Psi_{\omega vv'}\rangle - 1/6||\Psi_{\omega vv'}\rangle\right) + O(\eta^{1/2}) \\
= & O(\eta^{1/6}),
\end{align*}
\]

where in the second line we used (4.31) and the Cauchy-Schwarz inequality to bound the last term, and (4.32) for the first; in the third line, we used that \( ||\Psi_{\omega vv'}\rangle|| \leq 1 \), and in the last, we again applied (4.31) and the Cauchy-Schwarz inequality. Note that

\[
E_{\omega} \left|\left|G \otimes \text{Id} \right|\tilde{A}_{\omega}\rangle\rangle\right|_{\square} \left|\left|G \otimes \text{Id} \right|\tilde{B}_{\omega}\rangle\rangle\right|_{\square} \]

\[
\begin{align*}
= & \left|\left| E_{(\omega, v \sim v')} E_{\omega} \sum_{b \sim b'} \tilde{A}_{\omega v} \otimes \tilde{A}_{\omega v'} \right|^{1/2}\right| E_{(\omega, v \sim v')} E_{\omega} \sum_{b \sim b'} \tilde{B}_{\omega v} \otimes \tilde{B}_{\omega v'}\right|^{1/2} \\
\geq & \left|\left| E_{(\omega, v \sim v')} E_{\omega} \sum_{b \sim b'} \tilde{A}_{\omega v} \otimes \tilde{B}_{\omega v'} \right|^{1/2}\right|,
\end{align*}
\]

where the last inequality follows from Claim 2.3. Hence,

\[
\left|\left|G\right|\right|_{\square}^2 \geq E_{(\omega, v \sim v')} E_{\omega} \left|\left|G \otimes \text{Id} \right|\tilde{A}_{\omega}\rangle\rangle\right|_{\square} \left|\left|G \otimes \text{Id} \right|\tilde{B}_{\omega}\rangle\rangle\right|_{\square} \]

\[
\begin{align*}
= & E_{(\omega, v \sim v')} E_{\omega} \sum_{b \sim b'} \left(\Gamma_{dd'} |\tilde{A}_{\omega v} \otimes \tilde{B}_{\omega v'}\rangle\rangle |\Gamma_{dd'}\rangle\rangle\right)
\end{align*}
\]
\[
\begin{align*}
\geq & \quad \mathbb{E}_{\omega} \mathbb{E}_{v \sim v'} \sum_{b \leftrightarrow b'} \left\| \Psi_{\omega vv} \right\|^2 - 2 \langle \Psi_{\omega vv'} | U_{\omega v} A_{\omega v}^{-1/2} \hat{A}_{\omega v}^{-1/2} U_{\omega v}^\dagger \rangle \nonumber \\
& \quad \otimes U_{\omega v'} \hat{A}_{\omega v'}^{-1/2} \hat{A}_{\omega v'}^{-1/2} U_{\omega v'}^\dagger |\Psi_{\omega vv'}\rangle - O(\eta^{1/12}) \nonumber \\
& = \mathbb{E}_{\omega} \mathbb{E}_{v \sim v'} \sum_{b \leftrightarrow b'} \left\| \Psi_{\omega vv} \right\|^2 - 2 \langle \hat{\Psi} | \hat{A}_{\omega v}^b \otimes \hat{A}_{\omega v'}^{b'} | \hat{\Psi} \rangle - O(\eta^{1/12}),
\end{align*}
\] (4.34)

where the second line uses the definition of \(\hat{A}_{\omega v}^b\) and (4.33) and the third is by definition of \(|\Psi_{\omega vv'}\rangle\). To conclude, note that applying Markov’s inequality to (4.2) we get that a fraction at least \(1 - \eta^{1/3}\) of \(v \sim v'\) are such that

\[
\mathbb{E}_{\omega} \sum_{b \leftrightarrow b'} \langle \hat{\Psi} | \hat{A}_{\omega v}^b \otimes \hat{A}_{\omega v'}^{b'} | \hat{\Psi} \rangle \geq (1 - \eta^{2/3}) \mathbb{E}_{\omega} \left\| \Psi_{\omega vv} \right\|^2,
\]

where here we crucially used the max on the right-hand side of (4.2) to allow ourselves use the same \(v\) on the right-hand side as on the left-hand side. For any such \(v \sim v'\), a fraction \(1 - \eta^{1/3}\) of \(\omega \in \Omega\) will be such that

\[
\sum_{b \leftrightarrow b'} \langle \hat{\Psi} | \hat{A}_{\omega v}^b \otimes \hat{A}_{\omega v'}^{b'} | \hat{\Psi} \rangle \geq (1 - \eta^{1/3}) \left\| \Psi_{\omega vv} \right\|^2.
\]

For these \(v \sim v'\) and \(\omega\), the right-hand side of (4.34) is at least \(1 - \eta^{1/3} - O(\eta^{1/12})\), and their total weight constitutes at least an \((1 - 2\eta^{1/3})\) fraction of the total. \(\square\)

5. The correlated sampling lemma

In this section, we prove our quantum correlated sampling lemma.

**Lemma 5.1.** Let \(d\) be an integer and \(\delta > 0\). There exists an integer \(d'\), and for every state \(|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d\) unitaries \(V_{\psi}, W_{\psi}\) acting on \(\mathbb{C}^{dd'}\), such that the following holds for any two states \(|\psi\rangle, |\varphi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d\):\(^8\)

\[
\left\| \overline{V_{\psi}} \otimes W_{\varphi} |\Gamma_{dd'}\rangle - |\psi\rangle |\Gamma_{dd'}\rangle \right\| = O\left( \delta^{1/12}, \left\| |\psi\rangle - |\varphi\rangle \right\|^{1/6} \right),
\]

\(^8\)Note that here we implicitly re-ordered the registers, and \(|\psi\rangle |\Gamma_{dd'}\rangle\) should be understood as a bipartite state in \(\mathbb{C}^{dd} \otimes \mathbb{C}^{dd'}\), with the first (resp. second) space \(\mathbb{C}^{dd'}\) being associated with the tensor product of the first (resp. second) spaces, \(\mathbb{C}^d\) and \(\mathbb{C}^{d'}\), respectively associated with \(|\psi\rangle\) and \(|\Gamma_{dd'}\rangle\).
where here $|\Gamma_d\rangle \propto \sum_{1 \leq i \leq d} i^{-1/2} |i\rangle |i\rangle$ is the (properly normalized) $d$-dimensional embezzlement state.

A variant of the lemma holding for the special case of $|\psi\rangle = |\varphi\rangle$ was shown in van Dam & Hayden (2003), where the “embezzlement state” $|\Gamma_d\rangle$ was first introduced. It is not hard to see, however, that the construction of the unitaries $V_\psi$, $W_\varphi$ given in that paper does not satisfy the conclusion of Lemma 5.1. For instance, if $|\psi\rangle = (\sqrt{(1 + \varepsilon)/2} |00\rangle + \sqrt{(1 - \varepsilon)/2} |11\rangle)$ and $|\varphi\rangle = (\sqrt{(1 - \varepsilon)/2} |00\rangle + \sqrt{(1 + \varepsilon)/2} |11\rangle)$ then one can check that for any $\varepsilon > 0$ the unitaries from (van Dam & Hayden 2003) will be such that $\|V_\psi \otimes W_\varphi |\Gamma_{2d'}\rangle - |\psi\rangle|\Gamma_{d'}\rangle\| \geq 1/4$. This is due to our taking advantage of the degenerate spectrum of the reduced density of the EPR pair $(|00\rangle + |11\rangle)/\sqrt{2}$ to split the spectrum of the reduced density matrices of the nearby states $|\psi\rangle, |\varphi\rangle$ in two different ways; our proof of Lemma 5.1 shows that this is essentially the only obstacle that needs to be overcome in order to obtain a robust correlated sampling procedure.

Lemma 5.1 can be seen as a quantum analogue of Holenstein’s correlated sampling lemma (Holenstein 2009), which played an important role in his proof of the classical parallel repetition theorem. There the players, Alice and Bob, receive as inputs a description of a distribution $p, q$, respectively, such that $\|p - q\|_1 = \delta$. Their goal is to sample an element $u \sim p$ for Alice, $v \sim q$ for Bob, such that $u = v$ with probability $1 - O(\delta)$. This task can be reproduced in our setting by giving the states $|\psi\rangle = \sum_u \sqrt{p(u)} |u\rangle |u\rangle$ to Alice and $|\phi\rangle = \sum_v \sqrt{q(v)} |v\rangle |v\rangle$ to Bob. If the players run our procedure and then measure their joint state in the computational basis, they will obtain samples with a distribution close to $p$ and $q$, and moreover, these samples will be identical with high probability (though our proof would require them to use entanglement in order to do so!).

After the completion of this work, Anshu et al. (2014) proposed a different quantum generalization of the classical correlated sampling lemma. In the task they consider, the players are given reduced density matrices $\sigma, \tau$, respectively, such that $\|\sigma - \tau\|_1 = \delta$. Their task is to generate a shared state $|\Psi\rangle_{AA'BB'}$, where Alice holds registers $AA'$ and Bob registers $BB'$, such that the reduced
density of $|\Psi\rangle$ on $A$ (resp. $B$) is $\sigma$ (resp. $\tau$), and furthermore $|\Psi\rangle$ is close to being maximally entangled between $AA'$ and $BB'$. This task does not seem directly related to the one we consider; in particular, Anshu et al. show how their task can be accomplished starting from a sufficiently large number of shared EPR pairs, while our task provably requires a universal embezzlement state to be successfully accomplished.\footnote{Indeed, local operations alone cannot change the Schmidt coefficients, and local operations on a maximally entangled state will only yield maximally entangled states (possibly of varying dimension). See (Leung & Wang 2013) for further discussion of the criteria for universal embezzlement.}

We note that we have not tried to optimize the parameters appearing in the lemma. In particular, from our proof one can verify that taking $d' = 2^{O((d/\delta)^2)}$ in the lemma is sufficient, but this is probably far from optimal. Indeed, the method in van Dam & Hayden (2003) gives $d' = d^{O(1/\delta)}$; it may be possible to achieve such a polynomial dependence on $d$ here as well. (We refer the interested reader to recent work by Leung & Wang (2013) for an investigation of optimal families of embezzlement states, in the sense of van Dam and Hayden.)

**Proof (Proof of Lemma 5.1).** We define the unitaries $V_{\psi}$, $W_{\varphi}$ implicitly through the following procedure, in which two players Alice and Bob receive classical descriptions of two bipartite states $|\psi\rangle$, $|\varphi\rangle$, respectively, each of local dimension $d$, as well as a precision parameter $\delta > 0$. The unitaries $V_{\psi}$ and $W_{\varphi}$ correspond to their respective local quantum operations as described in the procedure. The players’ initial state consists of a classical description of the states $|\psi\rangle$, $|\varphi\rangle$, respectively (where each coefficient is specified with poly log$(\delta, d^{-1})$ bits of precision), a large supply of private qubits initialized in the $|0\rangle$ state, a large supply of shared EPR pairs that they will use as classical shared randomness, and an embezzlement state $|\Gamma_{\delta d'}\rangle$ for some large enough $d'$.

1. Let $d$ be the local dimension of $|\psi\rangle$ and $|\varphi\rangle$, $\delta$ the precision parameter given as part of the input, and $\eta > 0$ a small parameter to be specified later.
2. Using shared randomness, the players jointly compute a sequence \( \tau_0, \ldots, \tau_{K+1} \), where \( K = \lceil \log(d/\delta) \rceil \log(1+\eta) \), as follows. They set \( \tau_0 = 1, \tau_{K+1} = 0 \), and for \( k = 1, \ldots, K \) they jointly sample \( \tau_k \) uniformly at random in the interval \( [(1 + \eta)^{-k}, (1 + \eta)^{-(k+1)}] \).

3. Both players individually compute a classical description of the same (normalized) state

\[
|\xi_0\rangle \propto \sum_{k=0}^{K} \tau_k |k, k\rangle_{AB} |\Phi_d\rangle_{AB},
\]

where \( |\Phi_d\rangle = \sum_{i=1}^{d} |i\rangle|i\rangle \) is the un-normalized maximally entangled state on \( \mathbb{C}^d \otimes \mathbb{C}^d \). Let \( N = \lceil (2\delta d \sum_k \tau_k^2)^{-2} \rceil \). Alice and Bob jointly generate \( N \) copies of \( |\xi_0\rangle \), which they can achieve using the universal embezzling procedure from (van Dam & Hayden 2003) provided \( d' \) is large enough.

4. Alice (resp. Bob) computes the Schmidt decomposition \( |\psi\rangle = \sum_i \lambda_i |u_i\rangle|u_i'\rangle \) (resp. \( |\varphi\rangle = \sum_i \mu_i |v_i\rangle|v_i'\rangle \)). She sets \( S_k \) (resp. \( T_k \)) as the set of those indices \( i \) such that \( \lambda_i \in [\tau_{k+1}, \tau_k) \) (resp. \( \mu_i \in [\tau_{k+1}, \tau_k) \)), \( s_k = |S_k| \) (resp. \( t_k = |T_k| \)), and \( P_k \) (resp. \( Q_k \)) the projector on the span of the \( |u_i\rangle \) for \( i \in S_k \) (resp. \( |v_i\rangle \) for \( i \in T_k \)).

5. Alice measures her share of the first copy of \( |\xi_0\rangle \) using the two-outcome measurement \( \{P_A, \text{Id} - P_A\} \) where \( P_A := \sum_k |k\rangle\langle k| \otimes P_k \). Bob proceeds similarly with \( P_B := \sum_k |k\rangle\langle k| \otimes Q_k \). If either of them obtains the first outcome they proceed to the next step. Otherwise, they repeat this step with the next copy of \( |\xi_0\rangle \). If either player has used up all his or her copies he or she aborts the protocol.

6. Alice (resp. Bob) controls on the second register of \( |\xi_0\rangle \) to erase \( |k\rangle \) in the first register. (This is possible since the \( P_k \) (resp. \( Q_k \)) are orthogonal projections.) The players discard all qubits but the remaining register of \( |\xi_0\rangle \). Bob applies the unitary map \( |v_i\rangle \rightarrow |v_i'\rangle \) to his share.
Throughout the analysis, we assume without loss of generality that \( \delta \geq \| |\psi\rangle - |\varphi\rangle \|^2 \). We will show that with probability at least \( 1 - O(\delta^{1/12}) \) the procedure described above results in a shared state between Alice and Bob that is within trace distance \( O(\delta^{1/12}) \) of both \( |\psi\rangle \) and \( |\varphi\rangle \). Our first claim shows that, based on the \( \tau_k \), the players can each compute a discretized version of their inputs that both have (a slightly re-scaled version of) the \( \tau_k \) as Schmidt coefficients.

**Claim 5.2.** Define

\[
|\Psi\rangle := C \sum_k \tau_k \sum_{i \in S_k} |u_i\rangle |u'_i\rangle \quad \text{and} \quad |\Phi\rangle := C' \sum_k \tau_k \sum_{i \in T_k} |v_i\rangle |v'_i\rangle,
\]

where the \( \tau_k \), \( S_k \), and \( T_k \) are as defined in the protocol, and \( C, C' \) are appropriate normalization constants. Then,

\[
(5.3) \quad (1 + \eta)^{-1} \leq C, C' \leq 1,
\]

and

\[
(5.4) \quad \max \{ \| |\psi\rangle - |\Psi\rangle \|^2, \| |\varphi\rangle - |\Phi\rangle \|^2 \} = O(\eta).
\]

**Proof.** We have

\[
C^{-2} = \sum_k \tau_k^2 s_k \quad \text{which by definition of } S_k \text{ satisfies}
\]

\[
1 = \sum_i \lambda_i^2 \leq \sum_k \tau_k^2 s_k \leq \sum_i (1 + \eta)^2 \lambda_i^2 \leq (1 + \eta)^2.
\]

A similar calculation holds for \( C' \), proving (5.3). Next we bound the first term in (5.4), the second being similar. Using the definition of \( |\Psi\rangle \) and (5.3), we have

\[
\| |\psi\rangle - |\Psi\rangle \|^2 \leq \sum_k \sum_{i \in S_k} (\lambda_i - \tau_k)^2 + O(\eta)
\]

\[
\leq \sum_k \sum_{i \in S_k} \tau_k^2 \left( 1 - \frac{1}{1 + \eta} \right)^2 + O(\eta)
\]

\[
= O(\eta).
\]

Our next claim shows that the subspaces \( P_k, Q_k \) computed by the players are close, in the following sense.
Claim 5.5. The following equality holds with probability at least $1 - O(\delta^{1/6} \eta^{-1/3})$ over the choice of the $\tau_k$:

$$\sum_k \tau_k^2 \text{Tr}(P_k Q_k) = 1 - O(\delta^{1/6} \eta^{-1/3}).$$

(5.6)

Proof. Using Claim 5.2 and $||\psi - \varphi||^2 \leq \delta$, we deduce that $||\langle \Phi | \Psi \rangle|^2 = CC' \sum_{k,k'} \tau_k \tau_k' \text{Tr}(P_k Q_{k'}) = 1 - O(\eta)$. To prove the claim, we bound the contribution of those terms for which $k \neq k'$:

$$\sum_{k \neq k'} \tau_k \tau_k' \text{Tr}(P_k Q_{k'}) = \sum_{k \neq k'} \tau_k \tau_k' \sum_{i \in S_k, j \in T_{k'}} |\langle u_i | v_j \rangle|^2$$

$$\leq (1 + \eta)^2 \left( \sum_{k \neq k', i \in S_k, j \in T_{k'}} \lambda_i \mu_j |\langle u_i | v_j \rangle|^2 \right.$$

$$\left. + \sum_{k \neq k', z, i \in S_k, j \in T_{k'}} \lambda_i \mu_j |\langle u_i | v_j \rangle|^2 \right),$$

(5.7)

where $\theta > 0$ is a parameter to be fixed later. We bound each of the two terms inside the brackets in (5.7) separately. The first term is at most

$$\sum_{i,j} \lambda_i \mu_j |\langle u_i | v_j \rangle|^2 \leq \sum_{i,j} \frac{|\lambda_i - \mu_j|^2}{\theta} |\langle u_i | v_j \rangle|^2$$

$$\leq \theta^{-1} \sum_{i,j} |\lambda_i - \mu_j|^2 |\langle u_i | v_j \rangle|^2$$

$$\leq \theta^{-1} \delta \theta^{-1}.$$

To bound the second term in (5.7), note first that provided $\theta$ is at most a small constant times $\eta$ necessarily $k' = k + 1$ or $k' = k - 1$; our choice of $\theta$ will satisfy this condition. Suppose $k' = k - 1$, the other case being similar. Fix $i, j$ such that $|\sqrt{\lambda_i / \mu_j} - \sqrt{\mu_j / \lambda_i}|^2 < \theta$. This condition implies $|\lambda_i - \mu_j|^2 \leq \theta \mu_j \lambda_i \leq \theta (1 + \eta)^{-3} \tau_k^2$. Since $\tau_k$ is chosen uniformly in an interval of length $\tau_k \eta (1 + \eta)^{-1}$, the expected
fraction of pairs \((i, j)\) such that such that \(|\sqrt{\lambda_i/\mu_j} - \sqrt{\mu_j/\lambda_i}|^2 < \theta\) and \(\lambda_i \leq \tau_k \leq \mu_j\) is at most \(O(\sqrt{\theta}/\eta)\). Hence, on expectation over the choice of the \(\tau_k\) we have

\[
\sum_{k \neq k', i \in S_k, j \in T_{k'}} \lambda_i \mu_j |\langle u_i | v_j \rangle|^2 \leq O(\sqrt{\theta} \eta^{-1}) \sum_{i, j} \lambda_i \mu_j |\langle u_i | v_j \rangle|^2
\]

\[
= O(\sqrt{\theta} \eta^{-1}).
\]

Choosing \(\theta = (\delta \eta)^{2/3}\), we obtain that (5.6) holds, on expectation over the choice of the \(\tau_k\), with a right-hand side of \(1 - O(\delta^{1/3} \eta^{-2/3})\). (The condition that \(\theta \ll \eta\) is equivalent to \(\delta \ll \eta^{1/3}\), which we may assume holds without loss of generality, as otherwise the bound in the claim is trivial.) The left-hand side is at most 1, and applying Markov’s inequality proves the claim.

Our last claim analyzes the outcome of the sampling procedure, proving the lemma.

**Claim 5.8.** Let \(|\psi\rangle, |\varphi\rangle\) be such that \(|||\psi\rangle - |\varphi\rangle||^2 \leq \delta\), and set \(\eta = \delta^{1/4}\). With probability at least \(1 - O(\delta^{1/12})\), the sampling procedure described above terminates with Alice and Bob in a shared state \(|\xi\rangle\) such that \(|||\xi\rangle - |\psi\rangle||^2 = O(\delta^{1/12})\).

**Proof.** Suppose first that (5.6) holds and that Alice and Bob both proceed to the step 6 synchronously. In that case, at the end of the procedure their joint state is

\[
|\xi\rangle := C'' \sum_k \tau_k \sum_{i \in S_k, j \in T_k} \langle u_i | v_j \rangle |u_i\rangle |v_j\rangle,
\]

where the normalization constant \(C''\) satisfies

\[
(C'')^{-2} = \sum_k \tau_k^2 \sum_{i \in S_k, j \in T_k} |\langle u_i | v_j \rangle|^2
\]

\[
= \sum_k \tau_k^2 \text{Tr}(P_k Q_k)
\]

\[
= 1 - O(\eta + \delta^{1/3} \eta^{-2/3})
\]
by Claim 5.5. We can thus evaluate the overlap of $|\xi\rangle$ with $|\Phi\rangle$ as
\[
\langle \xi | \Phi \rangle \geq \sum_k \tau_k^2 \sum_{i \in S_k, j \in T_k} |\langle u_i | v_j \rangle|^2 - O(\delta^{1/3} \eta^{-2/3})
\]
\[
= 1 - O(\delta^{1/6} \eta^{-1/3}),
\]
where for the first equality we used orthogonality of the $|u_i\rangle$, and the last again follows from Claim 5.5.

Next we compute the probability that in step 5 Alice and Bob both obtain the first outcome of their respective POVM in the same iteration. The probability that Alice alone obtains a successful outcome is $\sum_k \tau_k^2 s_k / (d \sum_k \tau_k^2) = (1 + \Theta(\eta))(d \sum_k \tau_k^2)^{-1}$ by (5.3). The same holds for Bob. With probability at least $1 - \delta^2$, both of them obtain a successful outcome before the number $N$ of copies of $|\xi_0\rangle$ runs out. Moreover, the probability that they simultaneously obtain the first outcome is
\[
(d \sum_k \tau_k^2)^{-1} \sum_k \tau_k^2 \text{Tr}(P_k Q_k) \geq (1 - O(\delta^{1/6} \eta^{-1/3}))(d \sum_k \tau_k^2)^{-1}
\]
by Claim 5.5. Hence, the probability that they simultaneously proceed to the third step of the protocol is at least $1 - O(\delta^{1/6} \eta^{-1/3})$. Choosing $\eta = \delta^{1/4}$ proves the lemma.

\textbf{Acknowledgements}

We thank Attila Pereszlényi for comments on an earlier version of this manuscript, and the anonymous CCC and Computational Complexity reviewers for useful comments.

Irit Dinur’s research was supported by ERC grant No. 239985. Part of the work was done while the author was visiting MIT supported by NSF Contract CCF-1018064, and by Simons Investigator Award of Shafi Goldwasser. Part of David Steurer’s work was done at Microsoft Research New England. Thomas Vidick was partially supported by the Ministry of Education, Singapore, under the Tier 3 grant MOE2012-T3-1-009 and the National Science Foundation under Grant No. 0844626.
References

GRIGORI G. AMOSOV, ALEXANDER S. HOLEVO & REINHARDT F. WERNER (2000). On some additivity problems in quantum information theory. Technical report, arXiv:math-ph/0003002.

ANURAG ANSHU, RAHUL JAIN, PRIYANKA MUKHOPADHYAY, ALA SHAYEGHI & PENGHUI YAO (2014). A new operational interpretation of relative entropy and trace distance between quantum states. Technical report, arXiv:1404.1366.

P. K. ARAVIND (2002). The magic squares and Bell’s theorem. Technical report, arXiv:quant-ph/0206070.

ALAIN ASPECT, PHILIPPE GRANGIER & GÉRARD ROGER (1981). Experimental Tests of Realistic Local Theories via Bell’s Theorem. Phys. Rev. Lett. 47(7), 460–463.

BOAZ BARAK, ANUP RAO, RAN RAZ, RICKY ROSEN & RONEN SHALTIEL (2009). Strong Parallel Repetition Theorem for Free Projection Games. In Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, volume 5687, 352–365. Springer Berlin Heidelberg.

HOWARD BARNUM & EMANUEL KNILL (2002). Reversing quantum dynamics with near-optimal quantum and classical fidelity. J. Math. Physics 43(5), 2097–2106.

JOHN S. BELL (1964). On the Einstein-Podolsky-Rosen Paradox. Physics 1, 195–200.

JOP BRIËT, HARRY BURHMAN, TROY LEE & THOMAS VIDICK (2012). Multipartite entanglement in XOR games. Quantum Information and Computation 13(3-4), 334-360.

NICOLAS BRUNNER, DANIEL CAVALCANTI, STEFANO PIRONIO, VALEARIO SCARANI & STEPHANIE WEHNER (2014). Bell nonlocality. Rev. Mod. Phys. 86, 419–478.

ANDRE CHAILLOUX & GIANNICOLA SCARPA (2014). Parallel Repetition of Entangled Games with Exponential Decay via the Superposed Information Cost. In Automata, Languages, and Programming, volume 8572 of Lecture Notes in Computer Science, 296–307. Springer Berlin Heidelberg.
JOHN F. CLAUSER, MICHAEL A. HORNE, ABNER SHIMONY & RICHARD A. HOLT (1969). Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **23**, 880–884.

RICHARD CLEVE, WILLIAM SLOFSTRA, FALK UNGER & SARVAGYA UPADHYAY (2008). Perfect Parallel Repetition Theorem for Quantum XOR Proof Systems. *Comput. Complexity* **17**(2), 282–299.

TOM COONEY, MARIUS JUNGE, CARLOS PALAZUELOS & DAVID PÉREZ-GARCÍA (2011). Rank-one quantum games. Technical report, arXiv:1112.3563.

IRIT DINUR (2007). The PCP theorem by gap amplification. *J. ACM* **54**(3).

IRIT DINUR & DAVID STEURER (2013). Analytical Approach to Parallel Repetition. Technical report, arXiv:1305.1979. To appear in Proceedings STOC’14.

ALBERT EINSTEIN, BORIS PODOLSKY & NATHAN ROSEN (1935). Can quantum-mechanical description of physical reality be considered complete? *Physical Review* **47**, 777–780.

 URIEL FEIGE (1991). On the success probability of two provers in one-round proof systems. In *Proc. 6th IEEE Structure in Complexity Theory*, 116–123.

URIEL FEIGE & JOE KILIAN (2000). Two-Prover Protocols—Low Error at Affordable Rates. *SIAM J. Comput.* **30**(1), 324.

URIEL FEIGE & LÁSZLÓ LOVÁSZ (1992). Two-Prover One-Round Proof Systems: Their Power and Their Problems. In *Proc. 24th STOC*, 733–744.

URIEL FEIGE & OLEG VERBITSKY (2002). Error Reduction by Parallel Repetition – A Negative Result. *Combinatorica* **22**(4), 461–478.

LANCE FORTNOW, JOHN ROMPEL & MICHAEL SIPSER (1988). On the Power of Multi-Prover Interactive Protocols. In *Theoretical Computer Science* **134**(2), 545–557.

ESTHER HÄNGGI & RENATO RENNER (2009). Device-Independent Quantum Key Distribution with Commuting Measurements. Technical report, arXiv:1009.1833.
MATTHEW B HASTINGS (2009). Superadditivity of communication capacity using entangled inputs. *Nature Physics* **5**(4), 255–257.

PAUL HAUSLADEN, RICHARD JOZSA, BENJAMIN SCHUMACHER, MICHAEL WESTMORELAND & WILLIAM K. WOOTTERS (1996). Classical information capacity of a quantum channel. *Phys. Rev. A* **54**, 1869.

PAUL HAUSLADEN & WILLIAM K. WOOTTERS (1994). A ‘Pretty Good’ Measurement for Distinguishing Quantum States. *J. Modern Optics* **41**(12), 2385–2390.

PATRICK HAYDEN & ANDREAS WINTER (2008). Counterexamples to the Maximal $p$-Norm Multiplicativity Conjecture for all $p > 1$. *Comm. Math. Phys.* **284**(1), 263–280.

THOMAS HOLENSTEIN (2009). Parallel Repetition: Simplification and the No-Signaling Case. *Theory of Computing* **5**(1), 141–172.

TSUYOSHI ITO & THOMAS VIDICK (2012). A multi-prover interactive proof for NEXP sound against entangled provers. In *Proc. 53rd FOCS*, 243–252. IEEE Computer Society.

RAHUL JAIN, ATTILA PERESZLÉNYI & PENGHUI YAO (2013). A parallel repetition theorem for entangled two-player one-round games under product distributions. Technical report, arXiv:1311.6309. To appear in CCC’14.

MARIUS JUNGE, CARLOS PALAZUELOS, DAVID PÉREZ-GARCÍA, IGNACIO VILLANUEVA & MICHAEL M. WOLF (2010). Operator Space Theory: A Natural Framework for Bell Inequalities. *Physical Review Letters* **104**, 170 405.

JULIA KEMPE & ODED REGEV (2010). No Strong Parallel Repetition with Entangled and Non-signaling Provers. In *Proc. 25th IEEE Conf. on Computational Complexity (CCC’10)*, 7–15. IEEE Computer Society, Washington, DC, USA.

JULIA KEMPE & THOMAS VIDICK (2011). Parallel Repetition of Entangled Games. In *Proc. 43rd STOC*, 353–362.

FUAD KITTANEH (1986). Inequalities for the Schatten $p$-norm. IV. *Comm. Math. Phys.* **106**(4), 581–585.
Debbie Leung & Bingjie Wang (2013). Characteristics of Universal Embezzling Families. Technical report, arXiv:1311.6842.

Lluís Masanes, Stefano Pironio & Antonio Acín (2011). Secure device-independent quantum key distribution with causally independent measurement devices. Nature Communications 2(238), 7.

Gilles Pisier (2003). Introduction to Operator Space Theory. Cambridge University Press.

Anup Rao (2008). Parallel repetition in projection games and a concentration bound. In Proc. 40th STOC, 1–10. ACM.

Ran Raz (1998). A parallel repetition theorem. SIAM J. Comput. 27, 763–803.

Ran Raz (2008). A Counterexample to Strong Parallel Repetition. In Proc. 49th FOCS, 369–373. IEEE Computer Society.

Ran Raz & Ricky Rosen (2012). A Strong Parallel Repetition Theorem for Projection Games on Expanders. In Proc. 27th IEEE Conf. on Computational Complexity (CCC’12), 247–257.

Marco Tomamichel, Serge Fehr, Jедrzej Kaniewski & Stephanie Wehner (2013). A monogamy-of-entanglement game with applications to device-independent quantum cryptography. New Journal of Physics 15(10), 103 002.

Wim van Dam & Patrick Hayden (2003). Universal entanglement transformations without communication. Phys. Rev. A 67, 060 302(R).

Oleg Verbitsky (1994). Towards the parallel repetition conjecture. Proceedings of IEEE 9th Annual Conference on Structure in Complexity Theory 304–307.

Thomas Vidick (2013). Three-player entangled XOR games are NP-hard to approximate. In Proc. 54th FOCS. IEEE Computer Society.

Manuscript received 17 June 2014
