Parallel Repetition of Multi-party and Quantum Games via Anchoring and Fortification

by

Mohammad Bavarian

B.Sc., University of British Columbia (2011)

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2017

Author . . . .
Department of Mathematics
August 4, 2017

Certified by . . . .
Madhu Sudan
Gordon McKay Professor of Computer Science, Harvard University
Thesis Supervisor

Accepted by . . . .
Jonathan A. Kelner
Chairman, Applied Mathematics Committee

Parallel Repetition of Multi-party and Quantum Games via Anchoring and Fortification

by

Mohammad Bavarian

Submitted to the Department of Mathematics
on August 4, 2017, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Parallel repetition is a fundamental operation for amplifying the hardness inherent in multiplayer games. Through the efforts of many researchers in the past two decades (e.g. Feige, Kilian, Raz, Holentstein, Rao, Braverman, etc.), parallel repetition of two-player classical games has become relatively well-understood. On the other hand, games with entangled players (quantum games), crucial to the study of quantum non-locality and quantum cryptography, and multi-player games were poorly understood until recently.

In this thesis, we resolve some of the major problems regarding the parallel repetition of quantum and multi-player games by establishing the first *exponential-rate hardness amplification results* for these games and hence extend the classes of games where exponential decay rates is known considerably.

We consider two different methods for obtaining these hardness amplification results. For our first method, we draw from the recent work of Moshkovitz on parallel repetition of fortified games. We introduce an *analytic reformulation* of Moshkovitz's fortification framework. This reformulation allows us to expand the scope of the fortification method to new settings. In particular, we prove parallel repetition and fortification theorems for games with players sharing quantum entanglement, and games with more than two players in this new framework. An important component of our work is a variant of the fortification transformation, called ordered fortification, that preserves the entangled value of a game.

For our second method, we introduce a class of games we call *anchored*. Anchoring is a simple transformation on games inspired in part by the transorfmation proposed in the pioneering work of Feige-Kilian. Unlike the Feige-Kilian transformation, our anchoring transformation is completeness preserving. We prove an exponential-decay parallel repetition theorem for anchored games that involve any number of entangled players. We also prove a threshold version of our parallel repetition theorem for anchored games.

Thesis Supervisor: Madhu Sudan
Title: Gordon McKay Professor of Computer Science, Harvard University

Acknowledgments

A PhD is not an easy endeavor. I knew that when I was getting on the plane from Vancouver to Boston (one of the *only* things I knew about grad school). Despite this, the intellectual pull of MIT and research was strong enough to overcome any hesitation I felt at that point. Having reached almost the end of this path, I can confirm that my original belief in the difficulty of the path was more or less accurate. But I can also confirm that in my case getting a PhD was every bit as rewarding as it was difficult.

I am grateful to many people who helped me throughout this process, people who guided me when I was just beginning and had barely any clue, people who graciously shared their knowledge and ideas with me, and people who supported and kept me going through the ups and downs of PhD. Chief among these is my advisor Madhu Sudan. He is a truly remarkable person, and I feel so fortunate to have met him in my first year of graduate school. I still remember the excitement I used to feel before his "three-hour marathon theory reading group" on Fridays where, as a beginning graduate student, I (and many others) first learned the ins and outs of Theoretical Computer Science (TCS). It was in these seminars where I first had the chance to get to know Madhu while learning the newest developments and research directions in TCS from him and other experts in the field.

Madhu's depth of knowledge, intellectual curiosity, and exceptional taste in research were incredibly valuable to me during my studies. However, perhaps the best part of being advised by Madhu was the opportunity to enjoy his contagious lighthearted and optimistic attitude, which in my view was the special ingredient in making him a great adviser and colleague. I truly feel privileged for having had the chance to call myself his student.

Next, I would like to thank Scott Aaronson. My first exposure to quantum computing, the topic of my specialty, came through Scott's class in the fall of my second year. Also, a big portion of my research projects in grad school was directly influenced by him (including two papers I co-authored with him, and two which were inspired by questions he asked me). He is at the same time one of the best scientists and expositors of science that I know, a rare combination. I'm really fortunate to have had him as a mentor in grad school.

The other person that I met in my second year and who also played a pivotal role in my

PhD was Thomas Vidick. Thomas was at the time a freshly-minted PhD out of Berkeley and a postdoc in the quantum computing group at MIT. Being a postdoc, he was more accessible and closer in age to me than all my other mentors, and hence I felt more comfortable exposing my lack of knowledge to him, and he always found a way to educate me about quantum computing or pique my interest in the subject with his ideas and questions. Over time, I became very interested in some of the same research directions and we ended up collaborating quite a bit. In fact, the main results presented in this thesis are the results of a collaboration with him and fellow brilliant student, Henry Yuen. I want to thank of both of them for a wonderful and fruitful research collaboration. I learned so much from you guys.

Beside these, I wanted to thank all my other collaborators over the years, and in particular Andris Ambainis, Arturs Backurs, Dmitry Gavinsky, Tsuyoshi Ito, and Peter W. Shor. Furthermore, I wanted to thank all the other professors and colleagues, especially those in the quantum group, which made MIT such a wonderful environment for research in that area: Shalev Ben-David, Matt Coudron, Aram Harrow, Robin Kothari, Anand Natarajan, and John Wright.

I would also like to thank the amazing theory students at MIT. I interacted with various generations of MIT theory students, and it has always been a real pleasure. From Theory Retreat to Thursday outings to Friday night tea and card games to kayaking and other miscellaneous activities, the adventures we have had together have been incredible. From the people a few years above such as Elette Boyle, Alessandro Chiesa, Michael Forbes, Zeyuan Zhu, to more contemporary and younger theorists such as Aviv Adler, Adam Bouland, Michael Cohen, Alon Cohen, Akshay Degwekar, Andreea Gane, Rati Gelashvili, Daniel Grier, Dhiraj Holden, Justin Holmgren, Gautam Kamath, Pritish Kamath, Jerry Li, Sunoo Park, Madalina Persu, Govind Ramnarayan, Luke Schaeffer, Ludwig Schmidt, Adrian Vladu to the more recent additions such as Heather Berlin, Brynmor Chapman, Saleet Klein, Shibani Santukar, Nicole Wein, Helen Zhou the theory community has been almost like a family.

Special thanks goes to MIT theory Iranian students, Maryam Aliakbarpour, Mohsen Ghaffari, Sepideh Mahabadi, Saeed Mehraban, and Ali Vakilian. They did make G5 and G6 feel much more like home.

My life in grad school was not entirely confined to the theory community. Among my other good friends throughout this time were Zia Ghiasi, Soroush Khaleghi, Yen-Ling Kuo, Corinna Li, Jelena Makovic, Zaeim Mehraban, Mohammad-Hadi Pouransari, Candice Yip, and Yufei Zhao. Special thanks goes to Jelena Markovic. She was truly one of the best friends I've ever had.

Last but not least, I wanted to thank my family: my sisters, Sara, Maryam, Mona, and my parents, Behrouz and Fatemeh. I have been so fortunate to have you guys. Our crazy, big (and now even bigger with the addition of my lovely niece and nephew) family reunions are certainly my favorite times of the year. I wanted to especially thank my parents for always being there for me during the most crucial times. Their kind, accepting, and reassuring voice and their unwavering support are the greatest gifts I've had in my life. I really don't know where I would be without you and I love you very very much. Mom and dad, this thesis is dedicated to you.

Mohammad Bavarian
Cambridge, Massachusetts, USA, August 2017.

Contents

Chapter 1

Introduction

A central concept in theoretical computer science and quantum information is that of a *two-player one-round game*. A two-player game G is specified by question sets $\mathcal{X}$ and $\mathcal{Y}$, answer sets $\mathcal{A}$ and $\mathcal{B}$, a distribution μ over pairs of questions, and a verification predicate $V : \mathcal{A} \times \mathcal{B} \times \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$. The game is played between two cooperating (but non-communicating) players and a referee. The referee samples $(x, y) \in \mathcal{X} \times \mathcal{Y}$ according to μ and sends x and y to each player, who provide answers $a \in A$ and $b \in B$ respectively. The players win the game if their answers satisfy the predicate $V(a, b, x, y)$. More formally, the value of the game, which refers the maximum probability of winning players can achieve, corresponds to the following optimization problem

$$\mathrm{val}(G) = \max_{f,g} \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{(a,b)\in\mathcal{A}\times\mathcal{B}} V(a, b, x, y)\, f(x, a) \cdot g(y, b), \tag{1.1}$$

where $f : \mathcal{X} \times \mathcal{A} \to \mathbb{R}^+$ and $g : \mathcal{Y} \times \mathcal{B} \to \mathbb{R}^+$ with the the normalization condition $\sum_a f(x, a) = \sum_b g(y, b) = 1$.[1]

The main starting point of this thesis is the celebrated Parallel Repetition Theorem of Raz [53], simplified and improved upon by many authors over the years, most notably by Holestein [37].Beside its direct applications to PCPs [35] and the study of interactive proof systems [10, 53], the theorem of Raz/Holenstein is one of the strongest *direct product theorems*

[1] In this thesis, we also heavily study the entangled value game, denoted by $\mathrm{val}^*(G)$, which is an important analogous concept in quantum computing and complexity.

known in the literature, and as such has been very influential as the basic prototype of establishing direct product theorem in variety of models of computations. Parallel repetition theorems (and other direct product theorems) are often used in complexity theory in order to perform some form of *amplification*, such as amplifying the completeness-soundness gap of a proof system. A fundamental question that arises in this context is how the value of a repeated game $G^{\otimes m}$ relates to the value of the original game G and the number of repetitions m.[15, 16, 39]. The Raz/Holentein theorem provides an answer to this question.

Theorem 1.1 (Raz-Holenstein)**.** *Let G be a game and define $G^{\otimes m}$ as the game where the referee selects k tuples $(x_i, y_i)_{i=1}^m \in \mathcal{X} \times \mathcal{Y}$ independently according to G and sends $x = (x_1, x_2, \ldots, x_m)$ and $y = (y_1, y_2, \ldots, y_m)$ to Bob requiring answers in $\mathcal{A}^{\otimes m}$ and $\mathcal{B}^{\otimes m}$ satisfying the predicate $V = \prod_{i=1}^m V(a_i, b_i, x_i, y_i)$. Set $0 \leq \epsilon \leq 1$ by $\mathrm{val}(G) = 1 - \epsilon$. Then,*

$$\mathrm{val}(G^{\otimes m}) \leq (1 - \epsilon^3)^{-\Omega(m/s)}, \tag{1.2}$$

where $s = \log(|\mathcal{A}| \cdot |\mathcal{B}|)$.

Raz's theorem and subsequent improvements [37, 6, 52, 13] provide a satisfactory understanding of the parallel repetition of classical two-player games. However, going beyond the setting of two-player classical games, many of the techniques in these works do not directly apply. The focus of this thesis is to address this problem by providing strong (exponential) hardness amplification results for much more general classes of games. In particular, here we make progress on two of the major open problems regarding the parallel repetition of games, i.e. whether an analogue of Raz's parallel-repetition theorem holds for (a) games with more than two players, and (b) games with quantum players using entanglement. In the next section, we review how we draw from the old idea of *modifying games for parallel repetition*, as first appearing in the work of Feige-Kilian [30], to make progress on these problems.

1.1 Anchoring and fortification

As mentioned before, in this thesis we are interested in obtaining strong hardness amplification and parallel repetition results for general quantum and multiplayer (more than two) games.

The basic idea underlying our approach for achieving this is to modify the game G before applying parallel repetition. For this to be a valid approach, any such transformation should be (i) *efficient*, in the sense of having low algorithmic complexity, and (ii) *non-intrusive*, in the sense of preserving most important properties of the original game G. In particular, since computing the value (and quantum value) of games is a hard algorithmic problem we need the operation to preserve the value in the black-box way (or else change the value in a controllable fashion independent of G).

In this thesis, we introduce two simple transformations on games, namely *fortification* and *anchoring*, that can be used to convert any game to one of these desired formats. Together, our parallel repetition theorem and our transformations [7, 8] provide simple and efficient hardness-amplification in both the classical multiplayer and quantum settings.

1.1.1 Fortification and fortified games

In a recent work, Moshkovitz [48] introduced a simple yet powerful framework called *parallel repetition via fortification.* Like Feige-Kilian and our anchoring technique, the basic idea is to transform the game to make it more amenable to parallel repetition. But in terms of technical ideas and tools, and also final parameters, this framework is very different to the above. One advantage of fortification is that it does not use any of the subtle information theoretic techniques needed in other approaches to parallel repetition, and hence leads to overall to simpler arguments.

Despite its attractive features, the fortification framework [48] has some limitations; for instance it is only applicable to the restricted (though very important) setting of classical two-player projection games. The first contribution of this thesis is to expand the scope of this framework to a wider classes of games. The following is a summary of our main contributions to the fortification framework.

- **Analytic formulation of fortification.** The fortification framework was originally cast in combinatorial terms; Moshkovitz's definition of fortified games, which we describe in Section 3.1, involves a guarantee on the value of every sufficiently large rectangular subgame of a game. In our analytic reformulation, fortified games are defined in terms of *substrategies*, which one can think of as randomized strategies for the game where the probability that the players output an answer may be less than 1. This definition behaves much more "smoothly", allowing us to generalize them to the entangled and multiplayer settings.

- **Fortification of general classical games and games with more than two players.** Next, we show how to fortify a general k-player game G, for any $k \geq 2$. We show that for any two fortified general classical games G' and H', $\mathrm{val}(G' \otimes H') \approx \mathrm{val}(G') \cdot \mathrm{val}(H')$. Together this implies new gap amplification results for general (as opposed to projection) two-player and multiplayer classical games. We note that previously [48, 11] the fortification framework was limited to two-player classical *projection games*, so our extension is significant even for the two-player case.

- **An entangled-value preserving variant of concatenation.** A major obstacle in extending the fortification framework to the quantum setting is that *concatenation*, the main ingredient of the original fortification results, does not in general preserve the entangled value. That is, if G' is the fortification of G, it doesn't generally hold that $\textsc{Val}^*(G') = \textsc{Val}^*(G)$ (even though $\mathrm{val}(G') = \mathrm{val}(G)$). This is problematic for obtaining gap amplification results: if $\textsc{Val}^*(G) = 1$, then $\textsc{Val}^*(G^{\otimes n}) = 1$, but $\textsc{Val}^*(G'^{\otimes n})$ could be exponentially small!

  To resolve this issue, we augment the ordinary concatenation procedure of [48] by giving the players some auxiliary advice input (see Definition 3.4) which helps in keeping the entangled value unchanged. Using this, we define a variant of the fortification transformation which we call *ordered fortification.* As desired, in addition to preserving the classical value, this transformation also preserves the entangled value, which is essential for the completeness of our gap amplification result.

- **Fortification of games with entangled players.** We show that for a general two-player game G, its ordered-fortification G_{OF} is a two-player game such that $\mathrm{val}^*(G_{OF}) = \mathrm{val}^*(G)$, and is also quantumly fortified. We then prove that for any two quantumly fortified games G' and H', $\textsc{Val}^*(G' \otimes H') \approx \textsc{Val}^*(G') \cdot \textsc{Val}^*(H')$. Together this implies a new general gap amplification method for entangled two-player games. This (see Theorem 3.3) is the most technically challenging component of this work.

Let us note that our extensions of the fortification approach, as described in the last three items above, are ultimately enabled by the analytic viewpoint described in point 1.

We describe our main results for fortification in detail in Chapter 3. In particular, see Theorems 3.2, 3.3, 3.5, 3.24 for the formal statements of results of the chapter.

Next, we discuss anchoring approach to parallel repetition. Anchoring in general leads to better parameter performance than fortification (especially in terms of answer alphabet size), but relies on more technical information theoretical arguments. Despite being weaker in terms of parameters, fortification is interesting as it relies on a very different set of techniques than the typically information theoretic parallel repetition results.

1.1.2 Anchoring and anchored games

Next, we study the process of anchoring and anchored games. Using the anchoring transformation, we obtain hardness amplification results that are arguably even stronger than the results for fortified games. In particular, unlike fortification where one needs to know in advance the number of repetition before modifying the game (i.e. the transformation $G \to G'$ depends on number of repetitions m), anchoring is oblivious to the number of rounds of repetitions and has no extra blow-up depending on m. Furthermore, unlike in fortification, there is no alphabet blow-up in anchoring.

Regarding anchoring, we have the following theorem which applies to games with any number of players, with or without entanglement.

Theorem 1.2 (Main Theorem of Chapter 4, informal)**.** *There exists a polynomial-time transformation (called* anchoring*) that takes the description of an arbitrary k-player game G and returns a game $G_\perp$ with the following properties:*

1. *(Classical hardness amplification)*
   If $\mathrm{val}(G) = 1 - \varepsilon$ *then* $\mathrm{val}(G_\perp) = 1 - \frac{3}{4}\varepsilon$ *and* $\mathrm{val}(G_\perp^{\otimes n}) = \exp(-\Omega(\varepsilon^3 \cdot n))$.

2. *(Quantum hardness amplification)*
   If $\mathrm{val}^*(G) = 1 - \delta$ *then* $\mathrm{val}^*(G_\perp) = 1 - \frac{3}{4}\delta$ *and* $\mathrm{val}^*(G_\perp^{\otimes n}) = \exp(-\Omega(\delta^8 \cdot n))$.

The implied constants in the $\Omega(\cdot)$ *only depend on the number of players* k *and the cardinality of the answer sets of* G*.*

We obtain an efficient hardness amplification method from this theorem in the following way: suppose given a k-player game G whose entangled value is either 1 or at most $1-\delta$. By letting $n = \text{poly}(\log \beta^{-1}, \delta^{-1})$, the game $G_\perp^{\otimes n}$ (the n-fold repetition of the anchored game $G_\perp$) has value either 1 or at most β.

We note that even though, unlike fortification, anchoring does not preserve the game value exactly this is not a significant issue since it changes the value in a controllable (linear) fashion (see Definition 1.4). In particular, an important aspect of the anchoring transformation is that it preserves *quantum completeness*, meaning that if $\text{val}^*(G) = 1$, then $\text{val}^*(G_\perp) = 1$. Similar game transformations in previous works (such as the one by Feige and Kilian [30, 43]) do *not* preserve quantum completeness, and thus cannot be used for hardness amplification in the same way.

We also obtain a *threshold* version of the theorem above, which states that the probability that the players win more than an $\text{val}^*(G) + \gamma$ fraction of the n instances of $G_\perp$ in $G_\perp^{\otimes n}$ goes to 0 exponentially fast in n:

Theorem 1.3 (Threshold theorem, informal)**.** *Let* G *be a* k*-player game with* $\text{val}^*(G) = 1-\delta$*, and* $G_\perp$ *the anchored version of* G*. Then for all integer* $n \geq 1$ *the probability that in the game* $G_\perp^{\otimes n}$ *the players can win more than* $(1 - \frac{3}{4}\delta + \gamma)n$ *instances of* $G_\perp$ *is at most* $\exp(-\Omega(\gamma^9 n))$*, where the implied constant only depends on the number of players* k *and the cardinality of the answer sets of* G*.*

The advantage of having a threshold theorem is that it also implies that parallel repetition reduces the *completeness error* in addition to the soundness error. This is useful in situations where we are trying to distinguish between, say, $\text{val}^*(G) \geq 0.99$ and $\text{val}^*(G) \leq 0.5$. The entangled value of $G_\perp^{\otimes n}$ in both cases is exponentially small. However, if the referee instead checks that the number of instances won in $G_\perp^{\otimes n}$ is above a certain threshold, then we can obtain a new game where either the value is exponentially close to 1 or exponentially close to 0. See Theorem 4.16 for a more precise statement.

Finally, we present an application of our threshold theorem to the so-called Quantum

PCP Conjecture. The main application of Raz's parallel repetition theorem is to amplify the completeness/soundness gap of probabilistically checkable proofs, in order to obtain stronger hardness of approximation results (see, e.g., [35]). Similarly, our threshold bound would perform the same function for the multiprover games formulation of the Quantum PCP Conjecture. It is crucial that our threshold bound applies to games with any number of players; so far, it appears that the types of games that arise in approaches to the Quantum PCP Conjecture (games version) involve more than two players [41, 49]. We discuss this in more detail in Section 4.2.

Let us now describe the anchoring transformation. Here, we discuss anchoring for two-player games. We discuss the more general case in Chapter 4 in detail.

Definition 1.4 (Basic anchoring)**.** Let G be a two player game with question distribution μ on $\mathcal{X} \times \mathcal{Y}$, and verification predicate V. Let $0 < \alpha < 1$. In the α-anchored game $G_\perp$ the referee chooses a question pair $(x, y) \in \mathcal{X} \times \mathcal{Y}$ according to μ, and independently and with probability α replaces each of x and y with an auxiliary "anchor" symbol $\perp$ to obtain the pair $(x', y') \in (\mathcal{X} \cup \{\perp\}) \times (\mathcal{Y} \cup \{\perp\})$ which is sent to the players as their respective questions. If any of x', y' is $\perp$ the referee accepts regardless of the players' answers; otherwise, the referee checks the players' answers according to the predicate V.

For a choice of $\alpha = 1 - \frac{\sqrt{3}}{2}$ it holds that both $\mathrm{val}(G_\perp) = \frac{3}{4}\mathrm{val}(G) + \frac{1}{4}$ and $\mathrm{val}^*(G_\perp) = \frac{3}{4}\mathrm{val}^*(G) + \frac{1}{4}$. One can think of $G_\perp$ as playing the original game G with probability 3/4, and a trivial game with probability 1/4. The term "anchored" refers to the fact that question pairs chosen according to μ are all "anchored" by a common question $(\perp, \perp)$. Though the existence of this anchor question makes the game $G_\perp$ *easier* to play than the game G, it facilitates showing that the repeated game $G_\perp^{\otimes n}$ is *hard.* At a high level, the anchor questions provide a convenient way to handle the complicated correlations that may arise when the players use non-product strategies in the repeated game.

Our parallel repetition results more generally apply to a class of games we call *anchored.* The anchoring transformation of Theorem 1.2 produces games of this type; however, anchored games can be more general. We give a full definition of anchored games in Section 4.3. We note that the class of anchored games includes the class of *free games*, a class of games for which quantum parallel repetition theorems were previously shown in [19, 40, 20].

Why anchoring is useful. Given the fact that anchoring games just amounts to adding some simple dummy variable (Definition 1.4), it can be quite surprising that anchoring is useful for proving parallel repetition. This is related to the *quantum dependency-breaking problem* which we briefly describe next.

In most known proofs of classical parallel repetition theorems (except for fortification), the key step consists of bounding the players' success probability in most instances of G in $G^{\otimes n}$, *conditioned on the player winning a significant fraction of the instances.* Let W denote this italicized event. Conditioning on W introduces correlations between the players' questions, making this task non-trivial. Existing proofs rely on a "rounding argument" showing how two players can play the game G as if it were embedded as the i-th instance in the repeated game G, *conditioned on W.* Thus the success probability in the i-th game, conditioned on W, cannot be much higher than the value of G, concluding the proof through a straightforward inductive argument.

In the classical case the rounding argument relies on the ability for Alice and Bob to sample a *dependency-breaking variable* $\Omega_{x,y}$ which a priori depends on both inputs x and y. Once $\Omega_{x,y}$ is sampled by the players they can simulate the i-th instance of G, conditioned on W. The main technical work goes in showing that $\Omega_x \approx \Omega_{x,y} \approx \Omega_y$, where "$\approx$" denotes closeness in statistical distance, and Ω_x (resp. Ω_y) denotes the distribution of $\Omega_{x,y}$ averaged over y (resp. over x). This implies that Ω can be locally sampled by either player without communication through the use of a *correlated sampling* procedure.

In the quantum case the rounding argument seems to require that Alice and Bob jointly sample a *dependency-breaking quantum state* $|\Omega_{x,y}\rangle$, which again depends on both their inputs (although it is technically more complicated, as a first approximation $|\Omega_{x,y}\rangle$ can be thought of as the players' post-measurement state, conditioned on W and their i-th inputs being (x, y)). Designing a state that simultaneously allows Alice and Bob to (a) simulate the execution of the i-th game in $G^{\otimes n}$ conditioned on W, and (b) locally generate $|\Omega_{x,y}\rangle$ without communication is the main obstacle to proving a fully general parallel repetition theorem for entangled games.

The proofs of parallel repetition for free games [40, 19, 20] and projection games [28] resolve this obstacle by arguing for the existence of local unitaries U_x and V_y and a state $|\Omega\rangle$ such that

$$U_x \otimes V_y|\Omega\rangle \approx |\Omega_{x,y}\rangle. \tag{1.3}$$

Thus the players only need to share $|\Omega\rangle$ at the beginning of the simulation, and having received their separate inputs x and y, apply the local unitaries U_x and V_y to obtain (an approximation of) the desired state $|\Omega_{x,y}\rangle$. The argument heavily relies on either the product question distribution assumption for [40, 19, 20] or special symmetries of projection games in the case of [28].

The difficulty in extending the argument for free games to the case of general games is to show that the local unitaries each only depend on the input to a single player. In fact with the definition of $|\Omega_{x,y}\rangle$ used in these works it appears likely that this statement does not hold, thus a different approach must be found. When the game is anchored we are able to use the anchor question in order to show the existence of unitaries U_x and V_y that achieve (1.3) and depend only on a single player's question each. Achieving this requires us to introduce dependency-breaking states $|\Omega_{x,y}\rangle$ that are more complicated than those used in the free games case, and also involve the *classical* dependency-breaking variables of Raz and Holenstein.

We prove (1.3) by proving a sequence of approximate equalities: first we show that for most x there exists U_x such that $(U_x \otimes \mathbb{I})|\Omega_{\perp,\perp}\rangle \approx |\Omega_{x,\perp}\rangle$, where $|\Omega_{\perp,\perp}\rangle$ denotes the dependency-breaking state in the case that both players receive the anchor question "$\perp$", and $|\Omega_{x,\perp}\rangle$ denotes the state when Alice receives x and Bob receives "$\perp$". Then we show that for every y such that $\mu(y|x) > 0$ there exists a unitary V_y such that $(\mathbb{I} \otimes V_y)|\Omega_{x,\perp}\rangle \approx |\Omega_{x,y}\rangle$. Accomplishing this step requires ideas and techniques going beyond those used in the free games case. See Sections 4.3 and 4.4 for more on this topic.

1.2 Related work

Hardness amplification is a central method in complexity theory and cryptography for reducing the soundness error of interactive proofs and argument systems and hence parallel repetition, and other hardness amplification based on it, have been extensively studied [33, 30, 53, 37, 51, 34, 36]. We refer to the surveys by Feige and Raz [29, 54] for an extensive historical account of the classical parallel repetition theorem and its connections to PCPs and multiprover interactive proof systems, and instead focus on more recent results, specifically those pertaining to the quantum or multiplayer parallel repetition.

The study of multi-player games in the quantum setting is due to close connection to Bell inequalities, non-locality in quantum physics [9, 21] which is of great important to quantum cryptography and complexity [56, 47, 24, 58]. We refer to [17, 61, 22] for more on entangled games.

Going back to parallel repetition, the first results regarding the parallel repetition of entangled games was obtained by Cleve et al. [22, 23]. The general idea of modifying the game in order to facilitate the analysis of parallel repetition originates from the work of Feige and Kilian [30] who introduced the confuse/miss-match style repetition of games. The Feige-Kilian type parallel repetition was later extended by Kempe and Vidick [43] to the quantum setting allowing them to obtain the first general parallel repetition theorem for quantum games. However, their transformation did not preserve the entangled value, and hence did not lead to a fully general hardness amplification method for entangled games.

The main result underlying Chapter 3 is Moshkovitz [48], where the framework of parallel repetition via fortification was first introduced. Some simplifications and corrections to the work of Moshkovitz appeared in Bhangale et al. [11]. In particular, an important contribution of [11] was the clarification of the best bounds possible in classical fortification theorems [11, Appendix C].

Another important set of ideas underlying our work is related to the analytic approach to parallel repetition pioneered by Dinur and Steurer [27], further extended by Dinur et al. [28]. Our analytic reformulation of fortification framework is very much inspired by the ideas in these works.

Yet another different stream of work, directly related to Chapter 4, follows the original ideas of Raz and Holenstein [53, 37] by taking a more information theoretic approach to parallel repetition. The first results in this direction were obtained by Chailloux and Scarpa [19] and Jain et al. [40] who prove exponential-decay parallel repetition results for free two-player games. Their analysis, as well as the follow-up work of Chung et al. [20], provided the basis for Chapter 4 of this thesis.

Turning to the multiplayer setting, very little was known prior to the works presented in this thesis. It is folklore that free games with any number of players satisfy an exponential parallel repetition theorem, and this was explicitly proved in both classical and quantum settings in [20]. The only parallel repetition bound that applies to all classical multiplayer games is due to Verbitsky [59], but the rate of decay proved there is very slow – it is essentially an inverse Ackermann-like function. Multiplayer parallel repetition has been studied in the setting of *non-signaling strategies*, a superset of entangled strategies which allows the players to generate any correlations that do not imply communication. Buhrman et al. [18] show that the non-signaling value of a game G with any number of players decays exponentially under parallel repetition, with a rate of decay that depends on the entire description of the game G. Arnon-Friedman et al. [3] and Lancien and Winter [46] achieve similar results using a different technique based on "de Finetti reductions".

Beside the above prior work, we should mention some newer paper which the results of this thesis has inspired. Among these, a notable one is due to Yuen [63] where he uses the ideas in Chapter 4 to show that the entangled value of a general repeated entangled game must decay to 0 polynomially fast (provided the base game has entangled value less than one). Another interesting paper is [26] which establishes exponential-decay bounds for *expander games*, which includes anchored games as a special case.

1.3 Organization

In Chapter 2, we introduce some basic lemmas and definitions. In Chapter 3, we present our fortification framework and results and in Chapter 4, we present our anchoring results.

The organization of Chapter 3 is as follows. In Section 3.2, we introduce the notion of substrategies, induced strategies, and other basic definitions that are used throughout the chapter. In Section 3.3, we present the formal definition of analytically fortified games. The main parallel repetition theorem is proved in Section 3.4. Theorems 3.2 and 3.3 are proved in Sections 3.5 and 3.7, respectively. The main theorem of Chapter 3, Theorem 3.5, is proved by reduction to Theorem 3.3. This reduction is presented in Section 3.6.

The organization of Chapter 4 is as follows. We first introduce the formal definition of anchored games in Section 4.1. In Section 4.2 we give a brief discussion of the Quantum PCP Conjecture, and an application of our threshold theorem (Theorem 4.16) to it. In Section 4.3, we give an overview of the techniques underlying the results of the chapter, mainly focusing on the general ideas and leaving the specifics to each subsequent section. In Section 4.4 we present the proof of the quantum parallel repetition theorem for anchored games, as well as the threshold theorem. In Section 4.5 we present the result on the parallel repetition of multiplayer classical anchored games.

Finally, in Chapter 5 we summarize the main contributions of the thesis and present some open problems.

Chapter 2

Preliminaries

2.1 Games and parallel repetition

In the introduction we introduced the notion of classical value of a two-player game (1.1). Here we recall the notion of entangled value.

For a two-player game $G = (\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mu, V)$, an entangled strategy consists of a state (a vector with ℓ_2 norm 1) $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ and set of positive semi-definite matrices $\{A_x^a\}_{x\in\mathcal{X},a\in\mathcal{A}}$, $\{B_y^b\}_{y\in\mathcal{Y},b\in\mathcal{B}}$ such that

$$\forall x, y \,:\, \sum_a A_x^a = \sum_b B_y^b = \mathbb{I}.$$

The entangled value of G, denoted by $\mathrm{val}^*(G)$, is defined as the supremum value attained by all valid strategies:

$$\textsc{Val}^*(G) = \sup_{\{A_x^a\},\{B_y^b\},|\psi\rangle} \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b:\, V(x,y,a,b)=1} \langle\psi|A_x^a \otimes B_y^b|\psi\rangle.$$

Note that the case of $d = 1$ corresponds to the classical strategies which means $\mathrm{val}(G) \leq \textsc{Val}^*(G)$. It is well-know that for many games [21, 61] this inequality is strict, that is $\mathrm{val}(G) < \textsc{Val}^*(G)$.

2.1.1 Multiplayer games

A k-player game $G = (\mathcal{X}, \mathcal{A}, \mu, V)$ is specified by a question set $\mathcal{X} = \mathcal{X}^1 \times \mathcal{X}^2 \times \cdots \times \mathcal{X}^k$, answer set $\mathcal{A} = \mathcal{A}^1 \times \mathcal{A}^2 \times \cdots \times \mathcal{A}^k$, a probability measure μ on $\mathcal{X}$, and a verification predicate $V : \mathcal{A} \times \mathcal{X} \to \{0, 1\}$. Throughout this thesis, we use superscripts in order to denote which player an input/output symbol is associated with. For example, we write x^1 to denote the input to the first player, and a^t to denote the output of the t-th player. Finally, to denote the tuple of questions/answers to all k players we write $x = (x^1, \ldots, x^k)$ and $a = (a^1, \ldots, a^k)$ respectively.

The *classical value* of a game G is denoted by $\mathrm{val}(G)$ and defined as

$$\mathrm{val}(G) := \sup_{f^1,\ldots,f^k} \mathop{\mathbb{E}}_{(x^1,\ldots,x^k)\sim\mu} \left[V\left((f^1(x^1), \ldots, f^k(x^k), (x^1, \ldots, x^k) \right) \right]$$

where the supremum is over all functions $f_i : \mathcal{X}_i \to \mathcal{A}_i$; these correspond to deterministic strategies used by the players. It is easy to see that the classical value of a game is unchanged if we allow the strategies to take advantage of public or private randomness.

The *entangled value* of G is denoted by $\mathrm{val}^*(G)$ and defined as

$$\mathrm{val}^*(G) := \sup_{\substack{|\psi\rangle \in (\mathbb{C}^d)^{\otimes k} \\ M^1,\ldots,M^k}} \mathop{\mathbb{E}}_{(x^1,\ldots,x^k)\sim\mu} \sum_{\substack{(a^1,\ldots,a^k): \\ V\left((a^1,\ldots,a^k),(x^1,\ldots,x^k)\right)=1}} \langle \psi | M^1(x^1, a^1) \otimes \cdots \otimes M^k(x^k, a^k) | \psi \rangle$$

where the supremum is over all integer $d \geq 2$, k-partite pure states $|\psi\rangle$ in $(\mathbb{C}^d)^{\otimes k}$, and $M^1, \ldots, M^k$ for each player. Each M^t is a set of POVM measurements $\{M(x^t, a^t)\}_{a^t \in \mathcal{A}^t}$ acting on $\mathbb{C}^d$, one for each question $x^t \in \mathcal{X}^t$.

2.1.2 Repeated games

Let $G = (\mathcal{X}, \mathcal{A}, \mu, V)$ be a k-player game, with $\mathcal{X} = \mathcal{X}^1 \times \cdots \times \mathcal{X}^k$ and $\mathcal{A} = \mathcal{A}^1 \times \cdots \times \mathcal{A}^k$. Let $\mu^{\otimes n}$ denote the product probability distribution over $\mathcal{X}^{\otimes n} = \bigotimes_{i=1}^n \mathcal{X}_i$, where each $\mathcal{X}_i$ is a copy of $\mathcal{X}$. Similarly let $\mathcal{A}^{\otimes n} = \bigotimes_{i=1}^n \mathcal{A}_i$ where each $\mathcal{A}_i$ is a copy of $\mathcal{A}$. [1] Let

[1] We will use the tensor product notation ("$\bigotimes$") to denote product across coordinates in a repeated game, and the traditional product notation ("×") to denote product across players.

$V^{\otimes n} : \mathcal{A}^{\otimes n} \times \mathcal{X}^{\otimes n} \to \{0,1\}$ denote the verification predicate that is 1 on question tuple $(x_1, \ldots, x_n) \in \mathcal{X}^{\otimes n}$ and answer tuple $(a_1, \ldots, a_n) \in \mathcal{A}^{\otimes n}$ iff for all i, $V(a_i, x_i) = 1$. We define the n-fold parallel repetition of G to be the k-player game $G^{\otimes n} = (\mathcal{X}^{\otimes n}, \mathcal{A}^{\otimes n}, \mu^{\otimes n}, V^{\otimes n})$. For brevity, we sometimes denote this by G^n.

When working with games with more than 2 players, we use subscripts to denote which game round/coordinate a question/answer symbol is associated with. For example, by x_i^t we mean the question to the t-th player in the i-th round. While this is overloading notation slightly (because superscripts are meant to indicate tuples), we use this convention for the sake of readability. When x^n refers to a tuple $(x_1, \ldots, x_n)$ and when x_i^t refers to the t-th player's question in the i-th coordinate should be clear from context.

2.2 Probability distributions

Given a distribution μ, by $z \sim \mu$ we mean that the random variable z is distributed according to μ. For a set $\mathcal{S}$, by $z \sim \mathcal{S}$ we mean $z \sim U_{\mathcal{S}}$ where $U_{\mathcal{S}}$ is the uniform distribution over $\mathcal{S}$.

We let capital letters denote random variables and lower case letters denote specific samples. We will use subscripted sets to denote tuples, e.g., $X_{[n]} := (X_1, \ldots, X_n)$, $x_{[n]} = (x_1, \ldots, x_n)$, and if $C \subset [n]$ is some subset then X_C will denote the sub-tuple of $X_{[n]}$ indexed by C. We use P_X to denote the probability distribution of random variable X, and $\mathsf{P}_X(x)$ to denote the probability that $X = x$ for some value x. For multiple random variables, e.g., X, Y, Z, $\mathsf{P}_{XYZ}(x, y, z)$ denotes their joint distribution with respect to some probability space understood from context.

We use $\mathsf{P}_{Y|X=x}(y)$ to denote the conditional distribution $\mathsf{P}_{YX}(y, x)/\mathsf{P}_X(x)$, which is defined when $\mathsf{P}_X(x) > 0$. When conditioning on many variables, we usually use the shorthand $\mathsf{P}_{X|y,z}$ to denote the distribution $\mathsf{P}_{X|Y=y,Z=z}$. For example, we write $\mathsf{P}_{V|\omega_{-i},x_i,y_i}$ to denote $\mathsf{P}_{V|\Omega_{-i}=\omega_{-i},X_i=x_i,Y_i=y_i}$. For an event W we let $\mathsf{P}_{XY|W}$ denote the distribution conditioned on W. We use the notation $\mathbb{E}_X f(x)$ and $\mathbb{E}_{\mathsf{P}_X} f(x)$ to denote the expectation $\sum_x \mathsf{P}_X(x) f(x)$.

Let P_{X_0} be a distribution of $\mathcal{X}$, and for every x in the support of P_{X_0}, let $\mathsf{P}_{Y|X_1=x}$ be a conditional distribution defined over $\mathcal{Y}$. We define the distribution $\mathsf{P}_{X_0}\mathsf{P}_{Y|X_1}$ over $\mathcal{X} \times \mathcal{Y}$ as

$$(\mathsf{P}_{X_0}\mathsf{P}_{Y|X_1})(x,y) \;:=\; \mathsf{P}_{X_0}(x) \cdot \mathsf{P}_{Y|X_1=x}(y).$$

Additionally, we write $\mathsf{P}_{X_0 Z}\mathsf{P}_{Y|X_1}$ to denote the distribution $(\mathsf{P}_{X_0 Z}\mathsf{P}_{Y|X_1})(x,z,y) := \mathsf{P}_{X_0 Z}(x,z)\cdot \mathsf{P}_{Y|X_1=x}(y)$.

For two random variables X_0 and X_1 over the same set $\mathcal{X}$, $\mathsf{P}_{X_0} \approx_{\varepsilon} \mathsf{P}_{X_1}$ indicates that the total variation distance between P_{X_0} and P_{X_1},

$$\|\mathsf{P}_{X_0} - \mathsf{P}_{X_1}\| \;:=\; \frac{1}{2}\sum_{x\in\mathcal{X}} |\mathsf{P}_{X_0}(x) - \mathsf{P}_{X_1}(x)|,$$

is at most ε.

The following simple lemma will be used repeatedly.

Lemma 2.1. *Let Q_F and S_F be two probability distributions of some random variable F, and let $\mathsf{R}_{G|F}$ be a conditional probability distribution for some random variable G, conditioned on F. Then*

$$\|\mathsf{Q}_F\mathsf{R}_{G|F} - \mathsf{S}_F\mathsf{R}_{G|F}\| = \|\mathsf{Q}_F - \mathsf{S}_F\|.$$

Proof. Note that $\|\mathsf{Q}_F\mathsf{R}_{G|F} - \mathsf{S}_F\mathsf{R}_{G|F}\|$ is equal to

$$\begin{aligned}\frac{1}{2}\sum_{f,g} |\mathsf{Q}(f)\mathsf{R}(g|f) - \mathsf{S}(f)\mathsf{R}(g|f)| &= \frac{1}{2}\sum_{f} |\mathsf{Q}(f) - \mathsf{S}(f)| \cdot \left(\sum_{g}\mathsf{R}(g|f)\right)\\ &= \frac{1}{2}\sum_{f} |\mathsf{Q}(f) - \mathsf{S}(f)|\\ &= \|\mathsf{Q}_F - \mathsf{S}_F\|.\end{aligned}$$

□

2.3 Some matrix analytic facts

Choi-Jamiolkowski isomorphism. We make use of the correspondence between bipartite states $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ and linear operators $L : \mathcal{H}_2^* \to \mathcal{H}_1$ given by the Choi-Jamiolkowski isomorphism. Explicitly, let $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ be a quantum state and consider a Schmidt basis for $|\psi\rangle$ so we have $|\psi\rangle = \sum_{i=1}^d \sqrt{\lambda_i}|i\rangle|i\rangle$ where $\lambda_i \in \mathbb{R}^{\geq 0}$, up to a local change of basis. Set

$$\rho := \sum_{i=1}^{d} \lambda_i |i\rangle\langle i|. \tag{2.1}$$

Proposition 2.2. *Let Z, W be two linear operators acting on $\mathbb{C}^d$ and let $|\psi\rangle$ and ρ be as above. Then,*

$$\langle\psi|Z \otimes W|\psi\rangle = \mathrm{Tr}(Z\rho^{1/2}W^T\rho^{1/2}).$$

Proof. Both expressions evaluate to $\sum_{i,j=1}^d \sqrt{\lambda_i\lambda_j}\, Z_{ij} \cdot W_{ij}$. □

For a density matrix ρ and a matrix A for convenience we sometimes denote $\mathrm{Tr}(A\rho)$ by $\mathrm{Tr}_\rho(A)$.

Matrix norms and inequalities. The Frobenius norm of a matrix $A \in \mathbb{C}^{n\times m}$ is defined as $\|A\|_F = \sqrt{\mathrm{Tr}(AA^\dagger)}$. The trace norm is defined as $\|A\|_{tr} = \mathrm{Tr}\sqrt{AA^\dagger}$. The following analogue of Proposition 3.20 will be used repeatedly in our argument.

Claim 2.3. *Let M be a bipartite λ-spectral expander on vertex set $X' \cup X$. Let $\{A_{x'}\}_{x'\in X'}$ and ρ be positive semidefinite matrices. For all $x \in X$, define $A_x = \mathbb{E}_{x'\sim N(x)} A_{x'}$ and define $A = \mathbb{E}_{x\sim\mu} A_x$. Then*

$$\mathop{\mathbb{E}}_{x\sim\mu} \mathrm{Tr}_\rho((A_x - A)^2) \leq 2\lambda^2 \cdot \mathop{\mathbb{E}}_{x'\sim\mu'} \mathrm{Tr}_\rho(A_{x'}^2). \tag{2.2}$$

Proof. Define $S_{x'} = \rho^{1/2}A_{x'}$, $S_x = \rho^{1/2}A_x$, and $S = \mathbb{E}_x S_x = \rho^{1/2}A$. Using that M is a bipartite λ-spectral expander, for any fixed entry (i, j)

$$\mathop{\mathbb{E}}_{x} |(S_x)_{ij} - S_{ij}|^2 \leq \lambda^2 \cdot \mathop{\mathbb{E}}_{x'} |(S_{x'})_{ij} - S_{ij}|^2 \leq 2\lambda^2 \cdot \mathop{\mathbb{E}}_{x'} |(S_{x'})_{ij}|^2 \tag{2.3}$$

Summing over all entries,

$$\mathop{\mathbb{E}}_{x} \sum_{i,j} |(S_x)_{ij} - S_{ij}|^2 = \mathop{\mathbb{E}}_{x} \|S_x - S\|_F^2 \leq 2\lambda^2 \mathop{\mathbb{E}}_{x'} \sum_{i,j} |(S_{x'})_{ij}|^2 = 2\lambda^2 \mathop{\mathbb{E}}_{x'} \|S_{x'}\|_F^2. \tag{2.4}$$

Observing that $\mathrm{Tr}_\rho((A_x - A)^2) = \|S_x - S\|_F^2$ and $\|S_{x'}\|_F^2 = \mathrm{Tr}_\rho(A_{x'}^2)$, we obtain the desired result. □

If A has singular value decomposition $A = UJV^\dagger$ its pseudo-inverse is $A^{-1} = VJ^{-1}U^\dagger$, where J^{-1} is obtained from J by taking the reciprocal of non-zero diagonal entries. A simple consequence of the singular value decomposition is the following:

Fact 2.4. *Let A be an $n \times n$ matrix. Then there exists a unitary matrix $\mathcal{U}$ such that $\mathcal{U}A$ is positive semi-definite.*

Proof. Write the SVD as $A = UJV^\dagger$, and choose $\mathcal{U} = VU^\dagger$. □

We make frequent use of the matrix Cauchy-Schwarz inequality.

Proposition 2.5. *For any two matrices S, T we have*

$$\mathrm{Tr}(ST^\dagger) \leq \mathrm{Tr}(SS^\dagger)^{1/2} \cdot \mathrm{Tr}(TT^\dagger)^{1/2} = \|S\|_F \|T\|_F.$$

If S and T are Hermitian,

$$\mathrm{Tr}(STST) \leq \mathrm{Tr}(S^2T^2).$$

Finally, we need a variant of Powers-Størmer inequality from [45].

Lemma 2.6. *Let X, Y be positive semidefinite matrices. Then*

$$\mathrm{Tr}\left((X - Y)^4\right) \leq \mathrm{Tr}\left((X^2 - Y^2)^2\right).$$

This Lemma also played a role in the analysis of Dinur et al. [28] parallel repetition. See Kittaneh [45] for the proof.

2.4 Quantum information theory

For comprehensive references on quantum information we refer the reader to [50, 62].

For a vector $|\psi\rangle$, we use $\||\psi\rangle\|$ to denote its Euclidean length. For a matrix A, we will use $\|A\|_1$ to denote its *trace norm* $\mathrm{Tr}(\sqrt{AA^\dagger})$. A density matrix is a positive semidefinite matrix with trace 1. The *fidelity* between two density matrices ρ and σ is defined as $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$. The Fuchs-van de Graaf inequalities relate fidelity and trace norm as

$$1 - F(\rho, \sigma) \leq \frac{1}{2}\|\rho - \sigma\|_1 \leq \sqrt{1 - F(\rho, \sigma)^2}. \tag{2.5}$$

For Hermitian matrices A, B we write $A \preceq B$ to indicate that $A - B$ is positive semidefinite. We use $\mathbb{I}$ to denote the identity matrix. For an operator X and a density matrix ρ, we write $X[\rho]$ for $X\rho X^\dagger$. A *positive operator valued measurement* (POVM) with outcome set $\mathcal{A}$ is a set of positive semidefinite matrices $\{E^a\}$ labeled by $a \in \mathcal{A}$ that sum to the identity.

We will use the convention that, when $|\psi\rangle$ is a pure state, ψ refers to the rank-1 density matrix $|\psi\rangle\langle\psi|$. We use subscripts to denote system labels; so ρ_{AB} will denote the density matrix on the systems A and B. A *classical-quantum* state ρ_{XE} is classical on X and quantum on E if it can be written as $\rho_{XE} = \sum_x p(x)|x\rangle\langle x|_X \otimes \rho_{E|X=x}$ for some probability measure $p(\cdot)$. The state $\rho_{E|X=x}$ is by definition the E part of the state ρ_{XE}, conditioned on the classical register $X = x$. We write $\rho_{XE|X=x}$ to denote the state $|x\rangle\langle x|_X \otimes \rho_{E|X=x}$. We often write expressions such as $\rho_{E|x}$ as shorthand for $\rho_{E|X=x}$ when it is clear from context which registers are being conditioned on. This will be useful when there are many classical variables to be conditioned on.

For two positive semidefinite operators ρ, σ, the *relative entropy* $S(\rho\|\sigma)$ is defined to be $\mathrm{Tr}(\rho(\log\rho - \log\sigma))$. The *relative min-entropy* $S_\infty(\rho\|\sigma)$ is defined as $\min\{\lambda : \rho \preceq 2^\lambda\sigma\}$.

Let ρ_{AB} be a bipartite state. The mutual information $I(A : B)_\rho$ is defined as $S(\rho^{AB}\|\rho^A \otimes \rho^B)$. For a classical-quantum state ρ_{XAB} that is classical on X and quantum on AB, we write $I(A; B|x)_\rho$ to indicate $I(A; B)_{\rho_x}$.

The following technical lemmas will be used in Section 4.4.

Proposition 2.7 (Pinsker's inequality)**.** *For all density matrices ρ, σ, $\frac{1}{2}\|\rho - \sigma\|_1^2 \leq S(\rho\|\sigma)$.*

Lemma 2.8. *Let $\rho = \sum_z \mathsf{P}_Z(z)|z\rangle\langle z| \otimes \rho_z$, and $\rho' = \sum_z \mathsf{P}_{Z'}(z)|z\rangle\langle z| \otimes \rho'_z$. Then $S(\rho'\|\rho) = S(\mathsf{P}_{Z'}\|\mathsf{P}_Z) + \mathbb{E}_{Z'}\left[S(\rho'_z\|\rho_z)\right]$. In particular, $S(\rho'\|\rho) \geq \mathbb{E}_{Z'}\left[S(\rho'_z\|\rho_z)\right]$.*

We will also use the following Lemma from [20].[2] Here we present an argument that obtains better parameters ([20] proved that $\sum_{i=1}^n I(X_i : A)_\rho \leq 2S(\rho_{XA}\|\sigma_{XA})$.)

[2] Some versions of this lemma, though in a less compact form, also appear in [40, 19].

Lemma 2.9 (Quantum Raz's Lemma). *Let ρ and σ be two CQ states with $\rho_{XA} = \rho_{X_1X_2\dots X_nA}$ and $\sigma = \sigma_{XA} = \sigma_{X_1} \otimes \sigma_{X_2} \otimes \dots \otimes \sigma_{X_n} \otimes \sigma_A$ with $X = X_1X_2\dots X_n$ classical in both states. Then*

$$\sum_{i=1}^{n} I(X_i \,:\, A)_\rho \leq S(\rho_{XA}\,\|\sigma_{XA}). \tag{2.6}$$

The conditions on ρ and σ stated in the lemma are equivalent to them satisfying the following form

$$\rho_{XA} = \sum_x \mathsf{P}_X(x)|x\rangle\langle x| \otimes \rho_{A|X=x}, \qquad \sigma_{XA} = \sum_x \mathsf{P}'_X(x)|x\rangle\langle x| \otimes \sigma_A,$$

where $x = (x_1, x_2, \dots, x_n)$ is an n-tuple, P_X an arbitrary distribution, and $\mathsf{P}'_X(x) = \prod_{i=1}^n \mathsf{P}'_{X_i}(x_i)$ a product distribution.

Proof of Lemma 2.9. By the chain rule (Lemma 2.8) we have

$$S(\rho_{XA}\|\sigma_{XA}) = S(\rho_{X_1}\|\sigma_{X_1}) + \mathop{\mathbb{E}}_{x_1\leftarrow\rho_{X_1}} S(\rho_{X_2|X_1=x_1}\|\sigma_{X_2}) + \dots + \mathop{\mathbb{E}}_{x\leftarrow\rho_{X_1\cdots X_n}} S(\rho_{A|X=x}\|\sigma_A), \tag{2.7}$$

where $x_1 \leftarrow \rho_{X_1}$ means sampling x_1 according to the classical distribution ρ_{X_1}, and similarly for $x \leftarrow \rho_{X_1\cdots X_n}$. Consider any of the first n terms in (2.7). We have

$$\mathop{\mathbb{E}}_{x_{<i}\leftarrow\rho_{X_1X_2\dots X_{i-1}}} S(\rho_{X_i|x_{<i}}\|\sigma_{X_i}) \geq \mathop{\mathbb{E}}_{x_{<i}\leftarrow\rho_{X_1X_2\dots X_{i-1}}} S(\rho_{X_i|x_{<i}}\|\rho_{X_i}) = I(X_1\dots X_{i-1} : X_i)_\rho,$$

where $\rho_{X_i|x_{<i}}$ stands for $\rho_{X_i|X_{<i}=x_{<i}}$. Now consider the last term in (2.7):

$$\begin{aligned}\mathop{\mathbb{E}}_{x\leftarrow\rho_X} S(\rho_{A|X=x}\|\sigma_A) \geq \mathop{\mathbb{E}}_{x\leftarrow\rho_X} S(\rho_{A|X=x}\|\rho_A) &= S(\rho_{XA}\|\rho_X \otimes \rho_A) \\ &= I(X : A)_\rho = \sum_{i=1}^n I(X_i : A|X_1X_2\dots X_{i-1})_\rho.\end{aligned}$$

Summing up the last two equations and using $I(X_i : AX_1\dots X_i) = I(X_i : X_1\dots X_{i-1}) + I(X_i : A|X_1\dots X_{i-1})$ implies

$$S(\rho_{XA}\|\sigma_{XA}) \geq \sum_{i=1}^n I(X_i : AX_1\dots X_{i-1})_\rho \geq \sum_{i=1}^n I(X_i : A)_\rho,$$

where the last inequality follows from strong subadditivity, i.e., $I(X_i : X_1\dots X_{i-1}|A)_\rho \geq 0$. □

Chapter 3

Parallel Repetition via Fortification

This chapter is based on the paper *Parallel repetition via fortification: analytic view and the quantum case*, a joint work with T. Vidick and H. Yuen, published in the Proceedings of the 8th Innovations in Theoretical Computer Science (ITCS 2017) and also presented at the conference on the Theory of Quantum Computation, communication and cryptography (TQC 2017).

3.1 Introduction

As mentioned in the introduction, recently Moshkovitz [48] introduced a simple yet powerful framework for parallel repetition, called *parallel repetition via fortification.* In this framework, a game G is transformed through an operation called "fortification" to a new game G'. This new game G' is equivalent to G in that $\mathrm{val}(G) = \mathrm{val}(G')$, but then Moshkovitz shows that behavior of the value of *fortified* games under parallel repetition is much simpler than the general case, and avoids many of the subtleties encountered in the general case. The main benefits of fortified games are two-fold: first, their behavior under parallel repetition is much simpler than the general case, and second, all games can be easily fortified. Thus for nearly all intents and purposes, it suffices to focus on the parallel repetition of fortified games.

Despite its attractive features, the fortification framework [48] has some limitations; for instance it is only applicable to the restricted (though very important) setting of classical two-player projection games. In this chapter, we continue the study of the fortification approach to parallel repetition and try to expand its scope to wider classes of games. In particular, we give an analytic reformulation of Moshkovitz's framework which is key to expanding the scope of the fortification method to new settings.

Before going into the details of the new analytic framework, it would be useful to review Moshkovitz's original combinatorial fortification framework.

The combinatorial framework. Let G be a two-player game with question sets $\mathcal{X}, \mathcal{Y}$ and acceptance predicate V. For $S \subseteq \mathcal{X}$ and $T \subseteq \mathcal{Y}$, the *subgame* $G_{S\times T}$ is defined as the game where the referee selects $(x, y) \in \mathcal{X} \times \mathcal{Y}$ according to μ conditioned on $x \in S, y \in T$ and checks the players' answers according to the same predicate V (the referee accepts automatically if $\mu(S \times T) = 0$). A game G is said to be (ε, δ)-*combinatorially fortified* if

$$\mathrm{val}(G_{S\times T}) \leq \mathrm{val}(G) + \varepsilon, \qquad \forall\, S \subseteq \mathcal{X},\, T \subseteq \mathcal{Y},\ \text{s.t. } \mu(S \times T) \geq \delta. \tag{3.1}$$

The main insight underlying [48] is that games satisfying (3.1) also satisfy a strong form of parallel repetition (up to some number of rounds depending on ε, δ, and the alphabet size of G). This motivates the following approach to parallel repetition: Given a game G, Moshkovitz transforms the game $G \to G'$ such that $\mathrm{val}(G') \approx \mathrm{val}(G)$ and G' is (ε, δ)-combinatorially fortified for an appropriate choice of (ε, δ). Since fortified games satisfy a strong form of parallel repetition, one expects

$$\mathrm{val}(G'^{\otimes m}) \approx \mathrm{val}(G')^m \approx \mathrm{val}(G)^m. \tag{3.2}$$

Indeed, by appropriately choosing the parameters (ε, δ), [48] can show that the full procedure

$$G \longrightarrow G' \longrightarrow G'^{\otimes m} \tag{3.3}$$

amounts to a size-efficient method of *gap amplification*. That is, we have

$$\begin{array}{lll} \mathrm{val}(G) \geq c & \Rightarrow & \mathrm{val}(G'^{\otimes m}) \gtrsim c^m \\ \mathrm{val}(G) \leq s & \Rightarrow & \mathrm{val}(G'^{\otimes m}) \lesssim s^m \end{array}, \tag{3.4}$$

where we refer to the first condition as completeness and the second as soundness. The gap amplification procedure of Moshkovitz $G \to G' \to G'^{\otimes m}$ from (3.3) has three components: (i) a preprocessing step (biregularization), (ii) fortification, (iii) parallel repetition for fortified games.

The goal of the preprocessing step – the simplest step of the three – is to make the game *biregular* (a game G is called biregular if the marginals of questions on both Alice and Bob sides are uniform), since it is typically easier to analyze the fortification procedure for such games. The second step is *fortification*, which is the main technical ingredient of the whole approach. It is achieved by "concatenating" the game (see Section 3.1.1 below) with appropriate bipartite pseudorandom graphs. The third step $G' \to G'^{\otimes m}$ is the parallel repetition of fortified games, which as observed by [48] is considerably simpler to analyze than the generic (non-fortified) case.

3.1.1 Results and techniques

The main result of our work is the extension of the fortification framework to general classical games (with any number of players) and two-player entangled games. On the way to these results we prove new results on all three components of the fortification framework: (i) biregularization, (ii) fortification, and (iii) parallel repetition. In this subsection, we discuss some of these results in detail.

Parallel repetition. A main contribution of [48] was the realization that the two-player projection fortified games satisfy a strong form of parallel repetition, up to an additive error. This additive error depended on the parameters of fortification as well as the alphabet size of the fortified game. In this work, we prove an improved parallel repetition theorem (Theorem 3.24) which has the same dependance in the parameters of fortification, but instead of the alphabet size of the resulting fortified game, it only depends on the alphabet size of the original game (which has exponentially smaller alphabet size). This new parallel repetition theorem is crucial for extending the fortification framework to the setting of general (as opposed to projection) two-player games.

Let us remark that the reason why alphabet blow-up of fortification does not cause an issue for projection games is because for projection games it suffices to only fortify one side of the game (by working with so-called "square projection" version of the game). As a result there is no alphabet blow-up for the "unfortified" side, which allows the arguments of [48, 11] to go through. This one-sided fortification does not work for general games, which is why we need Theorem 3.24.

Fortification. We start with a definition. Let $G = (\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mu, V)$ be a game, and M and P two bipartite graphs over vertex sets $(\mathcal{X}', \mathcal{X})$ and $(\mathcal{Y}', \mathcal{Y})$ respectively. For each $x \in \mathcal{X}$ or $x' \in \mathcal{X}'$ let $N(x) \subseteq \mathcal{X}'$ and $N(x') \subseteq \mathcal{X}$ denote the set of neighbors of x and x', respectively (similarly for any y, y').

Definition 3.1 (Concatenated game [48])**.** In the concatenated game $G' = (M \circ G \circ P)$, the referee selects questions (x, y) according to μ, and independently selects a random neighbor x' for x using M, and y' for y using P. The players receive questions x' and y' and respond with assignments $a' : N(x') \to \mathcal{A}$ and $b' : N(y') \to \mathcal{B}$ respectively. The players win if $V(a'(x), b'(y), x, y) = 1$.

Our first two main results show how, both in the classical and quantum settings, any game can be fortified by concatenating it with bipartite graphs M and P with sufficiently good spectral expansion. (See Section 3.2.4 for the definition of spectral expanders, and Section 3.3.1 for the notion of weak fortification.)

Theorem 3.2. *Let G be a biregular game and M and P two bipartite λ-spectral expanders. If $\lambda \leq \frac{\varepsilon}{2}\sqrt{\frac{\delta}{2}}$, then the concatenated game $G' = (M \circ G \circ P)$ is (ε, δ)-weakly fortified against classical substrategies.*

Theorem 3.3. *Let G be a biregular game and M and P two bipartite λ-spectral expanders. If $\lambda \leq \frac{\varepsilon^2\delta}{56}$, then the concatenated game $G' = (M \circ G \circ P)$ is (ε, δ)-weakly fortified against entangled strategies.*

We stress that both in the quantum and classical settings the procedure used to fortify a game is precisely the same, i.e. concatenation with spectral expanders, and the only difference is in the resulting parameters. Despite the similarities, the proof of Theorem 3.3 is significantly more involved, requiring several new ideas and substantial matrix analytic arguements.

Next we discuss a distinctively quantum phenomenon which makes the construction of a full quantum gap amplification theorem – quantum analogue of (3.4) – considerably more difficult. As it turns out, even though Theorem 3.3 is sufficient to prove the soundness case of the gap amplification theorem, the concatenation procedure used in the process can undermine the completeness condition (i.e. $\textsc{Val}^*(G'^{\otimes m}) \gtrsim \textsc{Val}^*(G)^m$ in general fails to hold).

The issue is as follows: let G be a game and $G' = (M \circ G \circ P)$ be a concatenated version of G. Classically we have $\mathrm{val}(G') = \mathrm{val}(G)$. Quantumly, even though we still have $\textsc{Val}^*(G') \leq \textsc{Val}^*(G)$ the other direction in general fails: we would have liked to argue that the players in G' are able to utilize the strategy in G to achieve the same success probability in the concatenated game, but this seems impossible: having received $x' \in \mathcal{X}'$ and $y' \in \mathcal{Y}'$, the players have access to lists $N(x') \subseteq \mathcal{X}$ and $N(y') \subseteq \mathcal{Y}$ that they know contain the true questions of the referee, i.e. $x^* \in N(x')$, $y^* \in N(y')$. The players would like to apply their optimal strategy in G to each and every $(x, y) \in N(x') \times N(y')$ simultaneously, but this is in general impossible in the quantum setting.[1]

Note that the same issue does not arises classically because the optimal strategy in G can be taken to be a deterministic one, and the players in G' can use the same labeling suggested by the optimal strategy in G to give labels to all of $N(x')$ and $N(y')$ simultaneously. This strategy however relies on the fact that classically different questions have a simultaneous labeling, a fact which certainly has no quantum analogue.

We resolve the above issue using a novel entangled value-preserving variant of fortification which we call *ordered fortification*. The basic idea for ordered fortification is to give the players some extra advice information which helps in preserving the entangled value.

Let G be a game and $G' = (M \circ G \circ P)$ be a concatenated version of G. There is an extra parameter l in the construction defined as $l = \max\left\{\max_{x' \in \mathcal{X}'} |N(x')|, \max_{y' \in \mathcal{Y}'} |N(y')|\right\}$.

Definition 3.4 (Ordered concatenation)**.** Let G and G' be as above. In G'_{OF}, the referee samples (x, y) according to G and picks random neighbors $x' \sim N(x)$ and $y' \sim N(y)$ independently. She then also picks two random injective maps $r_{x'} : N(x') \to [l]$ and $s_{y'} : N(y') \to [l]$ conditioned on $s_{x'}(x) = r_{y'}(y)$. The referee sends x' and $r_{x'}$ to the first

[1] This is because the measurement operators of different questions do not in general commute which prevents Alice (say) to obtain simultaneous answers for all questions in $N(x')$. As a further illustration of this issue, see Section 3.2.2 for an example of a game where $\textsc{Val}^*(G') < \textsc{Val}^*(G)$.

player, and y' and $s_{y'}$ to the second and accepts if the players' answers $a' : N(x') \to \mathcal{A}$ and $b' : N(y') \to \mathcal{B}$ satisfy $V(a'(x), b'(y), x, y) = 1$.

Here the crucial point is that $r_{x'}$ and $s_{y'}$ are correlated. They give matching labels to true questions x and y. To achieve the same winning probability as in G, the players in G'_{OF} will share l copies of the state $|\psi\rangle$ from the optimal strategy in G. For each $x^* \in N(x')$ with label $i = r_{x'}(x^*)$, the first player will apply the optimal G-strategy for x to the i^{th} copy of $|\psi\rangle$ (similarly for the second player). The fact that $r_{x'}(x) = s_{y'}(y)$ ensures that for the true questions x and y the players apply the optimal G strategies to the same copy of $|\psi\rangle$, and hence are able to win with exactly the same winning probability as in G.

Of course, the crucial part here is that even though the auxiliary information in $r_{x'}$ and $s_{y'}$ is helpful to the players for replicating the winning probability of G, it should not be "too helpful". In particular, we need to still be able to prove that G'_{OF} is fortified with appropriate parameters. This point is established by the following theorem.

Theorem 3.5 (Ordered Fortification)**.** *Let G be a game and M and P be two bipartite graphs as above. Let G'_{OF} be constructed from G and $G' = (M \circ G \circ P)$ as in Definition 3.4. Then, we have*

$$\textsc{Val}^*(G'_{OF}) = \textsc{Val}^*(G).$$

Furthermore if M and P are λ-spectral expanders and $\lambda \leq \frac{\varepsilon^2\delta}{56}$, then G'_{OF} is also (ε, δ) weakly fortified.

We prove Theorem 3.5 in Section 3.6 using a spectral argument that reduces it to Theorem 3.3. Beside the above, we also prove a simple multiplayer fortification in Section 3.5 for classical games. It may be possible to adapt the proofs of Theorem 3.3 and Theorem 3.28 to obtain a *multiplayer fortification theorem* for entangled games. Although plausible, some further technical issues arise in this case which we do not pursue here.

Biregularization. As already mentioned, biregularization is a minor (but necessary) step in the fortification framework. Our biregularization lemmas are presented and proved in Appendix 3.2.3. In terms of final statement, our biregularization lemmas are incomparable with those of [48, 11]. For example, in the case of graphical games, we prove a biregularization lemma which preserves the value exactly but has a cubic blow-up in the number of questions, whereas the biregularization lemmas from [11, 48] had a nearly linear blow-up but only preserved value up to an additive error. Moreover, in this work we prove biregularization for all games whereas [11, 48] only considered graphical games. (See Subsection 3.2.3 for definitions.)

3.2 Preliminaries

3.2.1 Substrategies

The main goal of this section is to introduce the notion of classical and quantum *substrategies* which replace the notion of *subgames* from [48, 11]. As subgames were central in the *combinatorial* framework of [48], substrategies are similarly central to our analytic framework.

Let G be a game with question sets $\mathcal{X}, \mathcal{Y}$, answer sets $\mathcal{A}, \mathcal{B}$, predicate V, and question distribution μ on $\mathcal{X} \times \mathcal{Y}$.

Definition 3.6 (Classical substrategies)**.** Let $G = (\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mu, V)$ be a two-player game. A *classical substrategy* is given by (f, g) where $f : \mathcal{X} \times \mathcal{A} \to [0,1], g : \mathcal{Y} \times \mathcal{B} \to [0,1]$ satisfy

$$\forall x \in \mathcal{X},\ f(x) := \sum_a f(x,a) \le 1, \qquad \forall y \in \mathcal{Y},\ g(y) := \sum_b g(y,b) \le 1.$$

We call (f, g) a *"complete strategy"* (sometimes simply *strategy*) if equality holds in all above inequalities, i.e. $f(x) = g(y) = 1$ for all x, y.

Definition 3.7. Given a substrategy (f, g), the value of G with respect to (f, g) is given by

$$\mathrm{val}(G, f, g) := \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a\in\mathcal{A}, b\in\mathcal{B}} V(a,b,x,y)\, f(x,a) \cdot g(y,b). \tag{3.5}$$

The *classical value* of G is

$$\mathrm{val}(G) := \sup_{f,g} \mathrm{val}(G, f, g), \tag{3.6}$$

where the supremum is taken over all complete strategies f, g.

We note that the definition given by (3.6) can be easily seen to be equivalent to the more traditional definition of the classical value, i.e.

$$\mathrm{val}(G) := \max_{\substack{p:\mathcal{X}\to\mathcal{A}\\ q:\mathcal{Y}\to\mathcal{B}}} \mathop{\mathbb{E}}_{(x,y)\sim\mu} V(p(x), q(y), x, y), \tag{3.7}$$

because any strategy $f : \mathcal{X} \times \mathcal{A} \to [0,1], g : \mathcal{Y} \times \mathcal{B} \to [0,1]$ can be written as convex combination of a collection of strategies of $\{0,1\}$ valued strategies; on the other hand, taking supremum over f, g which are $\{0,1\}$ valued is precisely equivalent to (3.7).

Next, we extend the above notions to the quantum setting.

Definition 3.8 (Quantum substrategies)**.** Let $G = (\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mu, V)$ be a two-player game. A *quantum* (or *entangled*) *substrategy* for G is a tuple $(|\psi\rangle, \{A_x^a\}, \{B_y^b\})$ defined by an integer $d \in \mathbb{N}$, a unit vector $|\psi\rangle \in \mathbb{C}^{d\times d}$ and sets of positive semi-definite matrices $\{A_x^a\}_{x\in\mathcal{X},a\in\mathcal{A}}, \{B_y^b\}_{y\in\mathcal{Y},b\in\mathcal{B}}$ over $\mathbb{C}^d$ satisfying

$$\forall x \in \mathcal{X},\, A_x := \sum_a A_x^a \leq \mathbb{I}, \qquad \forall y \in \mathcal{Y},\, B_y := \sum_b B_y^b \leq \mathbb{I}. \tag{3.8}$$

If $A_x = B_y = \mathbb{I}$ for every x, y the quantum substrategy is called a *"complete strategy"* (sometimes simply *strategy*).

Definition 3.9. Given a quantum substrategy $(|\psi\rangle, \{A_x^a\}, \{B_y^b\})$, the value of G with respect to $(|\psi\rangle, \{A_x^a\}, \{B_y^b\})$ is given by

$$\textsc{Val}^*(G, |\psi\rangle, \{A_x^a\}, \{B_y^b\}) = \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b} V(a,b,x,y)\langle\psi|A_x^a \otimes B_y^b|\psi\rangle.$$

The *entangled value* of G is defined as

$$\textsc{Val}^*(G) = \sup_{|\psi\rangle,\{A_x^a\},\{B_y^b\}} \textsc{Val}^*(G, |\psi\rangle, \{A_x^a\}, \{B_y^b\}), \tag{3.9}$$

where the supremum is taken over all complete strategies $(|\psi\rangle, \{A_x^a\}, \{B_y^b\})$.

3.2.2 Concatenated games

Let $G = (\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mu, V)$ be a game, and M and P two bipartite graphs over vertex sets $(\mathcal{X}', \mathcal{X})$ and $(\mathcal{Y}', \mathcal{Y})$ respectively. For each $x \in \mathcal{X}$ or $x' \in \mathcal{X}'$ let $N(x) \subseteq \mathcal{X}'$ and $N(x') \subseteq \mathcal{X}$ denote the set of neighbors of x and x', respectively (similarly for any y, y'). Recall the definition of Concatenated Games from the introduction.

Definition (Definition 3.1 restated)**.** In the concatenated game $G' = (M \circ G \circ P)$, the referee selects questions (x, y) according to μ, and independently selects a random neighbor x' for x using M, and y' for y using P. The players receive questions x' and y' and respond by assignments $a' : N(x') \to \mathcal{A}$ and $b' : N(y') \to \mathcal{B}$ respectively. The players win if $V(a'(x), b'(y), x, y) = 1$.

For a concatenated game $G' = (M \circ G \circ P)$, we refer to G' as the *outer game* and to G as the *inner game*.

Let $G' = (M \circ G \circ P)$ be a concatenated game. Let $d_{\mathcal{X}'} = \max_{x' \in \mathcal{X}'} |N(x')|$, $d_{\mathcal{Y}'} = \max_{y' \in B'} |N(y')|$. Then, the alphabet of the concatenated game is given by $\mathcal{A}' = \mathcal{A}^{d_{\mathcal{X}'}}$, $\mathcal{B}' = \mathcal{B}^{d_{\mathcal{Y}'}}$. Similarly, it is easy to see that the distribution μ' of questions in G' is given by $\mu'(x', y') = \mathbb{E}_{(x,y)\sim\mu} \frac{1_{x' \in N(x)}}{|N(x)|} \cdot \frac{1_{y' \in N(y)}}{|N(y)|}$.

Definition 3.10. Let $G' = (M \circ G \circ P)$ be a concatenated game. To any pair of substrategies (f, g) for G' we associate the *induced substrategy*[2]

$$f(x, a) := \mathop{\mathbb{E}}_{x' \sim N(x)} \sum_{a' : a'(x) = a} f(x', a'), \qquad g(y, b) := \mathop{\mathbb{E}}_{y' \sim N(y)} \sum_{b' : b'(y) = b} f(y', b'). \tag{3.10}$$

Similarly, given an entangled substrategy $(|\psi\rangle, \{A_{x'}^{a'}\}, \{B_{y'}^{b'}\})$ for G', we define the *induced substrategy* as

$$A_x^a := \mathop{\mathbb{E}}_{x' \sim N(x)} \sum_{a'(x) = a} A_{x'}^{a'}, \qquad B_y^b := \mathop{\mathbb{E}}_{y' \sim N(y)} \sum_{b'(y) = b} B_{y'}^{b'}. \tag{3.11}$$

Intuitively, an induced strategy is a strategy for the inner game in which the players proceed as follows: given question $x \in \mathcal{X}$, $y \in \mathcal{Y}$ and a strategy (f, g) for the outer game, the

[2] Note the slight (but convenient) abuse of notation due to the use of the same letter to represent a substrategy and the corresponding induced substrategy. The more accurate but more cumbersome way of denoting the induced strategies in in [**?**]'s language would have been Mf and Pg.

players select two random neighbors of their questions $x' \in N(x), y' \in N(y)$ independently, and play according to the labeling of x, y suggested by (f, g) at x' and y'.

The following simple proposition will play an important role throughout the paper.

Proposition 3.11. *Let $G' = (M \circ G \circ P)$ be a concatenated game. The value of any classical strategy (f, g) (resp. quantum strategy $(|\psi\rangle, \{A_{x'}^{a'}\}, \{B_{y'}^{b'}\})$) in the outer game G' is equal to the value of the induced strategy in the inner game G:*

$$\mathrm{val}(G', f, g) = \mathrm{val}(G, f, g) \qquad \textit{and} \qquad \mathrm{VAL}^*(G', |\psi\rangle, \{A_{x'}^{a'}\}, \{B_{y'}^{b'}\}) = \mathrm{VAL}^*(G, |\psi\rangle, \{A_x^a\}, \{B_y^b\}). \tag{3.12}$$

As a consequence,

$$\mathrm{val}(G') \leq \mathrm{val}(G), \qquad \textit{and} \qquad \mathrm{VAL}^*(G') \leq \mathrm{VAL}^*(G). \tag{3.13}$$

Furthermore,

$$\mathrm{val}(G') = \mathrm{val}(G). \tag{3.14}$$

Proof. The first equality in (3.12) follows from linearity of expectation and the definition of induced strategies as

$$\begin{aligned}
\mathrm{val}(G', f, g) &= \mathop{\mathbb{E}}_{(x,y)\sim\mu} \mathop{\mathbb{E}}_{x'\sim N(x)} \mathop{\mathbb{E}}_{y'\sim N(y)} \sum_{a'\in\mathcal{A}', b'\in\mathcal{B}'} V(a'(x), b'(y), x, y)\, f(x', a') \cdot g(y', b') \\
&= \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a\in\mathcal{A}, b\in\mathcal{B}} V(a, b, x, y)\, f(x, a) \cdot g(y, b) \\
&= \mathrm{val}(G, f, g).
\end{aligned}$$

The second equality is proved similarly. The two inequalities (3.13) follow directly from (3.12). To show (3.14) it remains to show that $\mathrm{val}(G') \geq \mathrm{val}(G)$. Consider an optimal deterministic strategy for G given by $p : \mathcal{X} \to \mathcal{A}$ and $q : \mathcal{Y} \to \mathcal{B}$. For any $x' \in \mathcal{X}'$, $y' \in \mathcal{Y}'$ define $a' : N(x') \to \mathcal{A}$ according to p and $b' : N(y') \to \mathcal{B}$ according to q. It is easy to see that this achieves the same value in G' as (p, q) did in G. □

As mentioned in the introduction, the quantum analogue of (3.14) does not hold in general. For example, consider the case that M and P are complete bipartite graphs. In this case, the players playing $G' = (M \circ G \circ P)$ need to provide a labeling to all vertices in $\mathcal{X}$ and $\mathcal{Y}$ simultaneously. But this is essentially just a classical strategy as the labelings for $\mathcal{X}, \mathcal{Y}$ are now fixed. Hence, $\textsc{Val}^*(G') = \mathrm{val}(G)$, the classical value, which in many cases could be much smaller than $\textsc{Val}^*(G)$.

3.2.3 Biregularization

As in [48, 11] we prove our fortification theorems for the special class of *biregular games*.

Definition 3.12. A two-prover game $G = (\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mu, V)$ is called biregular if the marginals of μ on $\mathcal{X}$ and $\mathcal{Y}$ are both uniform.

The following lemma justifies that for our purposes we may always assume a game is biregular.

Lemma 3.13 (Biregularization lemma)**.** *Let $G = (\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mu, V)$ be a two-prover game and $\tau \in (0,1)$ a fixed constant. There exists an efficient algorithm that given G produces a biregular game G_{int} with question sets $\mathcal{X}_{int}$ and $\mathcal{Y}_{int}$ of cardinality at most*

$$|\mathcal{X}_{int}| \leq \frac{8|\mathcal{X}|^2|\mathcal{Y}|}{\tau}, \qquad |\mathcal{Y}_{int}| \leq \frac{8|\mathcal{X}||\mathcal{Y}|^2}{\tau}, \tag{3.15}$$

the same answer alphabet size as G, and value satisfying

$$\mathrm{val}(G) \leq \mathrm{val}(G_{int}) \leq \mathrm{val}(G) + \tau, \qquad \textsc{Val}^*(G) \leq \textsc{Val}^*(G_{int}) \leq \textsc{Val}^*(G) + \tau. \tag{3.16}$$

Note that (3.16) implies that applying the Biregularization Lemma to a game never decreases its value, and hence the procedure is completeness preserving.

A widely used class of games in applications are so-called *graphical games*, for which we can get an improved biregularization result that does not require any approximation factor τ.

Definition 3.14. A graphical game G is a game where the questions are given by choosing an edge of a bipartite graph uniformly at random (i.e. $E \subseteq \mathcal{X} \times \mathcal{Y}$ and $\mu(x,y) = \frac{1}{|E|}$ if $(x,y) \in E$ and $\mu(x,y) = 0$ otherwise). The predicate and the answers do not have any restrictions.

Lemma 3.15 (Biregularization lemma, graphical case)**.** *Suppose G is two-prover graphical game with E edges between $(\mathcal{X}, \mathcal{Y})$. There exists an efficient algorithm that given G produces a biregular game G_{int} with question sets $\mathcal{X}_{int}$ and $\mathcal{Y}_{int}$ bounded by*

$$|\mathcal{X}_{int}| \leq |E| \cdot |\mathcal{X}| \leq |\mathcal{X}|^2|\mathcal{Y}|, \qquad |\mathcal{Y}_{int}| \leq |E| \cdot |\mathcal{Y}| \leq |\mathcal{X}||\mathcal{Y}|^2, \tag{3.17}$$

the same answer alphabet size as G, and the value satisfying

$$\mathrm{val}(G) = \mathrm{val}(G_{int}) \qquad \textsc{Val}^*(G) = \textsc{Val}^*(G_{int}). \tag{3.18}$$

Remark 3.16. In the above, we can allow for multiple edges across vertices of G. In this case E must be taken as a multi-set and the bound $|E| \leq |\mathcal{X}||\mathcal{Y}|$ used in (3.17) must be suitably modified.

Interestingly, our technique for proving the biregularization lemmas is concatenation itself! We start by proving the second lemma for graphical games, and derive the general case by reduction.

Graphical games. Suppose G is a graphical game: there is a set of edges $E \subseteq \mathcal{X} \times \mathcal{Y}$ such that $\mu(x, y) = \frac{1}{|E|}$ for all $(x, y) \in E$. In this case we have

$$\mu(x) = \frac{|N(x)|}{|E|}, \qquad \mu(y) = \frac{|N(y)|}{|E|}, \qquad \forall x \in \mathcal{X},\, y \in \mathcal{Y}. \tag{3.19}$$

Let $d_x := |N(x)|$ denote the degree of x, and set S_x be the set $\{x\} \times [d_x]$. Define

$$\mathcal{X}_{int} = \bigcup_{x \sim \mathcal{X}} S_x.$$

Note that $|\mathcal{X}_{int}| = |E| \leq |\mathcal{X}||\mathcal{Y}|$. Define $M_{int}((x, i), x) = \frac{1}{d_x}$ for $i \in \{1, \ldots, d_x\}$, and 0 otherwise. Construct $\mathcal{Y}_{int}$ and P_{int} similarly.

Proposition 3.17. *Let $G_{int} = M_{int} \circ G \circ P_{int}$. Then marginal of μ_{int} induced on $\mathcal{X}_{int}$ and*

$\mathcal{Y}_{int}$ *is uniform. Moreover, we have*

$$\mathrm{val}(G_{int}) = \mathrm{val}(G), \qquad \mathrm{val}^*(G_{int}) = \mathrm{val}^*(G). \tag{3.20}$$

Proof. It is easy to see that for all $\mathcal{X}_{int} = (x, i) \in \mathcal{X}_{int}$ we have $\mu_{int}(x_{int}) = \frac{1}{|E|}$ and similarly for all $y_{int} \in \mathcal{Y}_{int}$. The claims $\mathrm{val}(G_{int}) = \mathrm{val}(G)$ and $\mathrm{val}^*(G_{int}) \leq \mathrm{val}^*(G)$ are true for all concatenated games in general. The final claim $\mathrm{val}^*(G_{int}) \geq \mathrm{val}^*(G)$ follows by considering the strategy $A_{(x,i)} = A_x$, $B_{(y,j)} = B_y$ which achieves the same value as (A_x, B_y) in G. □

General case. Although graphical games include many games considered in applications, it would nevertheless still be nice to extend the above construction to all games. We do not know how to do this exactly, but we can achieve an approximate variant.

The idea is essentially to approximate a general game by a graphical game. More formally, let $\tau \in (0, 1)$ be an error parameter and q an integer such that $\frac{|E|}{\tau} \leq q \leq \frac{2|E|}{\tau}$. We have

$$\frac{\tau}{2|E|} \leq \frac{1}{q} \leq \frac{\tau}{|E|}. \tag{3.21}$$

We would like to define a game $\tilde{G}$ in which all probabilities in the underlying distribution $\tilde{\mu}(x, y)$ are fractions with denominator q. Let $\tilde{\mathcal{X}} = \mathcal{X} \cup \{x_{nul}\}$ and $\tilde{\mathcal{Y}} = \mathcal{Y} \cup \{y_{nul}\}$. For every $(x, y) \in \mathcal{X} \times \mathcal{Y}$ set

$$\tilde{\mu}(x, y) = \frac{\lfloor q \cdot \mu(x, y) \rfloor}{q}. \tag{3.22}$$

Finally let $\tilde{\mu}(x_{nul}, y_{nul})$ such that $\tilde{\mu}$ is a proper probability distribution (i.e. by transferring the excess probabilities to (x_{nul}, y_{nul})) and put an arbitrary winnable predicate on (x_{nul}, y_{nul}).

Proposition 3.18. *The game $\tilde{G}$ is a graphical game with q (possibly parallel) edges. Moreover, we have*

$$\mathrm{val}(G) \leq \mathrm{val}(\tilde{G}) \leq \mathrm{val}(G) + \tau, \qquad \mathrm{val}^*(G) \leq \mathrm{val}^*(\tilde{G}) \leq \mathrm{val}^*(G) + \tau. \tag{3.23}$$

A few remarks are in order: firstly, since the previous construction for graphical games applies equally well in the presence of multiples edges, we can combine it with the above preprocessing to prove Lemma 3.13. Secondly, note that the operation $G \to \tilde{G}$ is value-increasing and hence preserves perfect completeness. Thirdly, note that the right scale for the error parameter τ is $\frac{c-s}{2}$ where $c-s$ is the completeness-soundness gap.

Proof. By construction all $\tilde{\mu}(x,y)$ are integer multiples of $\frac{1}{q}$. This ensures that the same is true for $\tilde{\mu}(x_{nul}, y_{nul})$. Since $\mu(x,y) \geq \tilde{\mu}(x,y)$ for all (x,y), for any strategy (f,g) for G we have

$$1 - \mathrm{val}(G,f,g) = \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{V(a,b,x,y)=0} f(x,a)\cdot g(y,b) \geq 1 - \mathrm{val}(\tilde{G},f,g),$$

which shows that $\mathrm{val}(G) \leq \mathrm{val}(\tilde{G})$. For the other direction, consider an optimal strategy (f,g) for $\tilde{G}$ (which necessarily always wins on (x_{nul}, y_{nul})). We have,

$$\begin{aligned}
1 - \mathrm{val}(G) \leq 1 - \mathrm{val}(G,f,g) &= \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{V(a,b,x,y)=0} f(x,a)\cdot g(y,b) \\
&\leq \sum_{x,y} \tilde{\mu}(x,y) \sum_{V(a,b,x,y)=0} f(x,a)\cdot g(y,b) + \sum_{(x,y)\in E} (\mu(x,y) - \tilde{\mu}(x,y)) \\
&\leq 1 - \mathrm{val}(\tilde{G}) + \tau
\end{aligned}$$

The quantum case is similar. □

3.2.4 Expanders

The method used in [48, 11] for fortifying a game is concatenation with sufficient pseudorandom bipartite graphs. This is done using extractors in [48] whereas expanders are employed in [11].[3] Here we follow the latter approach and use expanders.

Let $M = (\mathcal{X}' \times \mathcal{X}, E)$ be a bipartite graph. For $x \in \mathcal{X}$ let $N(x) \subseteq \mathcal{X}'$ denote the set of neighbors of x and similarly for $x' \in \mathcal{X}'$. We shall work with graphs that are $\mathcal{X}$-regular, i.e. $d = |N(x)|$ for all $x \in \mathcal{X}$. Define distributions μ and μ' on $\mathcal{X}$ and $\mathcal{X}'$ via

$$\mu(x) = \frac{1}{|\mathcal{X}|}, \qquad \mu'(x') = \frac{|N(x')|}{d}.$$

[3] The two approaches however lead to essentially to similar parameters (e.g. $\lambda = O(\varepsilon\sqrt{\delta})$ to get (ε,δ)-fortified graph where λ is the second largest singular value of normalized adjacency matrix of the concatenating graph.); moreover, in the classical setting the approaches are in fact are more or less equivalent. See [11] for more.

for all $x \in \mathcal{X}$ and $x' \in \mathcal{X}'$. Note that $\mu'(x')$ is the probability of obtaining x' by sampling $x \sim \mu$ and taking a random neighbor of (according M) x . Let $\mathcal{M}$ be the following normalized adjacency matrix of M

$$\mathcal{M}(x,x') = \begin{cases} \frac{1}{d} \cdot \sqrt{\frac{\mu(x)}{\mu'(x')}} & \text{if } x' \in N(x) \\ 0 & \text{otherwise} \end{cases}$$

We usually view $\mathcal{M}$ as an operator from $\ell_2(\mathcal{X}')$ to $\ell_2(\mathcal{X})$. Note that when M is a biregular expander we get the simpler definition $\mathcal{M}(x,x') = \frac{1}{d}\sqrt{\frac{|\mathcal{X}'|}{|\mathcal{X}|}}$ for $x' \in N(x)$, and 0 otherwise.

Definition 3.19. A bipartite graph M is called a λ-*spectral expander* if the second-largest singular value of $\mathcal{M}$ is at most λ.

A simple useful proposition for us is the following:

Proposition 3.20. *Let $M = (\mathcal{X}' \times \mathcal{X}, E)$ be a bipartite λ-spectral expander. For $f : \mathcal{X}' \to \mathbb{R}$ and $x \in \mathcal{X}$ let $f(x) = \mathbb{E}_{x' \sim N(x)} f(x')$, and $\bar{f} = \mathbb{E}_{x' \sim \mu'} f(x') = \mathbb{E}_{x \sim \mu} f(x)$. Then*

$$\mathop{\mathbb{E}}_{x \sim \mu}(f(x) - \bar{f})^2 \leq \lambda^2 \mathop{\mathbb{E}}_{x' \sim \mu'}(f(x') - \bar{f})^2. \tag{3.24}$$

Proof. Let $p_\mu \in \mathbb{R}^{\mathcal{X}}, p_{\mu'} \in \mathbb{R}^{\mathcal{X}'}$ to unit vectors defined as $p_\mu(x) := \sqrt{\mu(x)}$ and $p_{\mu'}(x') := \sqrt{\mu(x')}$. Let$q_{\mathcal{X}}(x) := \sqrt{\mu(x)}\, f(x)$, $q_{\mathcal{X}'}(x') := \sqrt{\mu(x')}\, f(x')$.

First, observe that $\mathcal{M}\, p_{\mu'} = p_\mu$ and $\mathcal{M}^t\, p_{\mu'} = p_\mu$. It follows that $(p_\mu, p_{\mu'})$ form a pair of singular vectors of $\mathcal{M}$. Moreover, it is easy to see[4] that these are top singular vectors which shows that $\|\mathcal{M}\|_{op} = 1$. Now notice that

$$\mathop{\mathbb{E}}_{x' \leftarrow \mu'}(f(x') - \bar{f})^2 = \sum_{x'}(\sqrt{\mu(x')}f(x') - \bar{f}\sqrt{\mu(x')})^2 = \|q_{\mathcal{X}'} - \bar{f}p_{\mu'}\|_2^2, \tag{3.25}$$

Second, observe

$$\mathcal{M}\, q_{\mathcal{X}'} = q_{\mathcal{X}}. \tag{3.26}$$

[4] e.g. by appealing to the Perron-Frobenius theorem.

As such, (3.24) precisely corresponds to

$$\|q_{\mathcal{X}} - \bar{f}\, p_{\mu}\|_2^2 = \|\mathcal{M}\,(q_{\mathcal{X}'} - \bar{f}\, p_{\mu'})\|_2^2 \leq \lambda^2 \cdot \|q_{\mathcal{X}'} - \bar{f}\, p_{\mu'}\|_2^2. \tag{3.27}$$

The claim follows by noting the orthogonality property

$$\langle q_{\mathcal{X}'} - \bar{f}\, p_{\mu'}, p_{\mu'}\rangle = \sum_{x'} \mu(x') f(x') - \bar{f} = 0. \tag{3.28}$$

□

3.3 Fortification framework

This section introduces the fortification framework. We define the notion of analytically fortified games and recall our main parallel repetition and fortification theorems. We end by a discussion of the parameters of the resulting gap amplification results.

3.3.1 Analytical fortification

We distinguish between two variants of the notion of fortified games which we call *weakly fortified games* and *strongly fortified games.* Although the difference between the two may seem minor, this difference is in fact quite important in the quantum case.

Definition 3.21 (Fortified games)**.** Let $\varepsilon, \delta \in [0,1]$. A concatenated game $G' = (M \circ G \circ P)$ is called weakly (ε, δ)- fortified against classical substragies if for any substrategy f, g we have

$$\mathrm{val}(G', f, g) \leq (\mathrm{val}(G) + \varepsilon) \cdot \mathop{\mathbb{E}}_{(x,y)\sim\mu} f(x)\, g(y) + \delta. \tag{3.29}$$

Similarly, we define G' to be weakly (ε, δ)-fortified against entangled substrategies if for any substrategy $\{A_{x'}^{a'}\}, \{B_{y'}^{b'}\}$ we have

$$\textsc{Val}^*(G', \{A_{x'}^{a'}\}, \{B_{y'}^{b'}\}) \leq (\textsc{Val}^*(G) + \varepsilon) \cdot \mathop{\mathbb{E}}_{(x,y)\sim\mu} \langle\psi| A_x \otimes B_y |\psi\rangle + \delta. \tag{3.30}$$

If furthermore $\mathrm{val}(G)$ (resp. $\textsc{Val}^*(G)$) can be replaced by $\mathrm{val}(G')$, (resp. $\textsc{Val}^*(G')$) in the above then the game is called “strongly fortified" against classical (resp. quantum) substrategies.

Note that our main results, Theorems 3.2 and 3.3, show how any game can be (weakly) fortified by concatenating it with good-enough spectral expanders.

Two remarks regarding the above definition are in order:

- Using (3.13), we see that strong fortification implies weak fortification, as expected from the terminaology.
- From (3.14) it follows that the two notions in fact coincide in the case of classical fortification, but this is no longer the case for quantum fortification.

Our notion of fortified games and that of [48, 11] are closely related. Essentially, in Definition 3.21 we have replaced the the condition for all δ-large rectangles in (3.1) with a smoother condition. In terms of a precise relation, we can show the following.

Claim 3.22. *Every $(\varepsilon, \varepsilon\delta)$ strongly fortified game is also $(2\varepsilon, \delta)$ combinatorially fortified.*

Proof. Consider a subgame given by $S \subseteq \mathcal{X}, T \subseteq \mathcal{Y}$ in G. To every strategy (p, q) for $G_{S\times T}$, i.e., $p : S \to \mathcal{A}$, $q : T \to \mathcal{B}$, we can associate a natural substrategy (f, g) by

$$f(x,a) = \begin{cases} 1 & \text{if } x \in S \wedge p(x) = a, \\ 0 & \text{otherwise} \end{cases} , \qquad g(y,b) = \begin{cases} 1 & \text{if } y \in T \wedge q(y) = b, \\ 0 & \text{otherwise} \end{cases} . \tag{3.31}$$

Then one can easily see

$$\mathrm{val}(G, f, g) = \mathrm{val}(G_{S\times T}, p, q) \cdot \mu(S \times T).^5 \tag{3.32}$$

Now assuming that rectangle $S \times T$ is δ-large, i.e. $\mu(S \times T) \geq \delta$, and since G is fortified against classical substrategies, we have

$$\mathrm{val}(G_{S\times T}, p, q) = \frac{\mathrm{val}(G, f, g)}{\mu(S \times T)} \tag{3.33}$$

[5] The term $\mu(S \times T) = \mathbb{E}_{(x,y)\sim\mu} f(x)g(y)$ is a natural scaling parameter playing an important role in our discussion as a measure of the “largeness" of a subgame or a substrategy.

$$\leq (\mathrm{val}(G)+\varepsilon)\cdot\frac{\mathbb{E}_{(x,y)\sim\mu} f(x)g(y)}{\mu(S\times T)}+\frac{\delta\varepsilon}{\mu(S\times T)} \tag{3.34}$$

$$\leq \mathrm{val}(G)+2\varepsilon \tag{3.35}$$

where in the second inequality we used $\mu(S\times T)=\mathbb{E}_{(x,y)\sim\mu} f(x)g(y)$. □

We note that in Lemma 3.22, the reverse implication does not hold and the notion of analytically fortified game is strictly stronger. In what follows, in the rare occasion when we call a game fortified (without specifying weak or strong) we mean strongly fortified.

3.3.2 Parallel repetition of fortified games

Using the definition of fortified games, it is straightforward to prove the following parallel repetition theorem.

Theorem 3.23 (Basic parallel repetition). *Let G_2' be a (ε,δ)-fortified game against classical substrategies. Then for any game G_1' we have*

$$\mathrm{val}(G_1'\otimes G_2')\leq (\mathrm{val}(G_2')+\varepsilon)\cdot\mathrm{val}(G_1')+\delta\cdot|\Sigma_{G_1'}|, \tag{3.36}$$

where $\Sigma_{G_1'}$ is the total answer alphabet size (i.e. the product of Alice and Bob's alphabets) of G_1'.

We prove this theorem in Section 3.4 by adapting the proof of the analogous theorems in [48, 11] to the analytic setting. Unfortunately, while this theorem exemplifies the main idea behind our results, it is not directly useful for applications. The reason for this is that the fortification procedure $G\to G'$ via concatenation induces a large blow-up in the alphabet size, $|\Sigma_{G'}|\approx|\Sigma_G|^D$, where $D=\frac{1}{\varepsilon\sqrt{\delta}}$ is the degree of the expander graph chosen. As one iterates the repetition procedure m times, the blow-up due to the additive term in (3.36) will be of order $\delta|\Sigma_{G'}|^{m-1}$. But typically $|\Sigma_{G'}|^{m-1}\gg|\Sigma_G|^{(m-1)/\sqrt{\delta}}$, leading to a term larger than 1 and rendering the theorem useless.

We resolve this problem by proving an improved repetition theorem which exploits the fact that G' takes the form of a concatenated game, whose inner game G has a much smaller alphabet.

Theorem 3.24. *Let G' be a concatenated game, with inner game G, that is (ε, δ)-weakly fortified against classical substrategies. If $\delta \cdot (m-1) \cdot |\Sigma_G|^{m-1} \leq \eta$ then*

$$\mathrm{val}(G'^{\otimes m}) \leq (\mathrm{val}(G) + \varepsilon)^m + \eta. \tag{3.37}$$

Similarly, if G' is (ε, δ) weakly-fortified against entangled substrategies and $\delta \cdot (m-1) \cdot |\Sigma_G|^{m-1} \leq \eta$ then

$$\textsc{Val}^*(G'^{\otimes m}) \leq (\textsc{Val}^*(G) + \varepsilon)^m + \eta. \tag{3.38}$$

The main advantage of Theorem 3.24 compared to Theorem 3.23 is in the additive error, which is now is in terms $|\Sigma_G|$ rather than $|\Sigma_{G'}|$. What is important here is that the size of $|\Sigma_G|$ is independent of the fortification parameters (ε, δ) whereas $|\Sigma_{G'}|$ grows exponentially as δ decreases. Let us also note that Theorem 3.24 is quite general, and in particular applies to the multiplayer case.

3.3.3 Gap amplification

Having stated our main parallel repetition, fortification, and biregularization theorems, all the main components of gap amplification are finally in place. Indeed, using $\mathrm{val}(G) = \mathrm{val}(G')$ Theorem 3.24 implies our final gap amplification for the classical value. This matches the parameters of main results of [48, 11] and extends it to more general settings.

Since quantumly we could have $\textsc{Val}^*(G') < \textsc{Val}^*(G)$, from (3.38) we cannot obtain

$$\textsc{Val}^*(G'^{\otimes m}) \leq (\textsc{Val}^*(G') + \varepsilon)^m + \eta. \tag{3.39}$$

However, Theorem 3.3 and Theorem 3.24 are still sufficient to prove a gap amplification theorem for the case where the completeness holds against classical players and the soundness against the quantum ones.[6] To obtain a fully quantum gap amplification however, we need to appeal to the notion of *ordered fortification* which, as we discussed, is a entangled-value preserving variant of the ordinary fortification.

[6] E.g. as was the case in [38, 60].

Theorem (Theorem 3.5 restated)**.** *Let G be a game and M and P be two bipartite graphs as above. Let G'_{OF} be constructed from G and $G' = (M \circ G \circ P)$ as in Definition 3.4. Then, we have*

$$\textsc{Val}^*(G'_{OF}) = \textsc{Val}^*(G).$$

Furthermore if M and P are λ-spectral expanders and $\lambda \le \frac{\varepsilon^2\delta}{56}$, then G'_{OF} is also (ε, δ) weakly fortified.

We stress that G'_{OF} constructed above is itself a concatenated game with the inner game $G^{\oplus l}$, disjoint union of $l = \text{poly}(\frac{1}{\varepsilon^2\delta})$ copies of G. This means the inner alphabet size of G'_{OF} is precisely the same as G's, and therefore there is fortunately no issue in terms of alphabet blow-up for applying Theorem 3.24 to G'_{OF}. So using G'_{OF} instead of G' in Theorem 3.24, we can finally prove the analogue of (3.39) for G'_{OF}.

3.3.4 Parameters of gap amplification

We can now discuss the parameters of the gap amplification corollaries. As in [48, 11], the parameters are typically very good in terms of question sizes but much worse in terms of alphabet size. Here, we mostly focus our discussion to gap amplification in the classical setting as the calculations in the quantum setting are similar.

To understand the parameters, we need to only consider the soundness case. Suppose we are given a game G with guarantee $\text{val}(G) \le 1 - \tau$ and a target soundness value β. We choose $\varepsilon = \tau/2$ and m such that $(\text{val}(G) + \varepsilon)^m \le \beta/2$. Hence, we have $m = \frac{\log(2/\beta)}{\log(1-\tau/2)} \le \frac{2\log(2/\beta)}{\tau}$. We want

$$\text{val}(G'^{\otimes m}) \le (\text{val}(G) + \varepsilon)^m + \delta \cdot (m-1)|\Sigma_G|^{m-1} \le \beta. \tag{3.40}$$

Hence, we just need to ensure $\delta \cdot |\Sigma_G|^{m-1} \le \beta/2$. So we have $\delta = \frac{\beta}{(m-1)\cdot|\Sigma_G|^{m-1}}$.

So what does the above mean in terms of the size of the final output of gap amplification $G'^{\otimes m}$. The question size is $|\mathcal{X}|^m$ and $|\mathcal{Y}|^m$ (since we have $|\mathcal{X}'| = |\mathcal{X}|$ and $|\mathcal{Y}'| = |\mathcal{Y}|$). Note that m is essentially as small as we can hope for because even given a perfect parallel repetition theorem, we had to take $m \approx \frac{\log(1/\beta)}{\tau}$. Hence, the construction is essentially optimal in terms of question sizes.

For the alphabet size, the situation is much worse. We have $|\Sigma_{G'}| = |\Sigma_G|^D$ where $D = O(\frac{\text{poly}\log(1/\varepsilon^2\delta)}{\varepsilon^2\delta})$. This means (up to dominant factors) that $|\Sigma_{G'^{\otimes m}}| = |\Sigma_G|^{\frac{m^2|\Sigma_G|^{m-1}}{\beta}}$ which means that the alphabet is exponentially worse than basic parallel repetition which results in $|\Sigma_G|^m$. Note that however in typical settings where $|\Sigma_G|$ is constant and β a small constant (or inverse logarithmic in size of G), this exponentially worse behavior of alphabet size does not cause a significant problem.

Next, let us consider the setting where the completeness holds for classical players and soundness against entangled players. In this case, we can just use Theorem 3.3 instead of Theorem 3.2, and hence all the calculations are precisely the same with ε and δ replaced with their squares.

Lastly, in the fully quantum case we need to use Theorem 3.5. In this case, m, ε, δ are chosen in precisely the same way. Alphabet size is also exactly the same as G'_{OF} has the same alphabet size as G'. The only difference is that the question sizes in G'_{OF} are slightly larger than G': we have $|\mathcal{X}'| = |\mathcal{X}| \cdot \text{poly}(\frac{|\Sigma_G|^m}{\beta})$ and $|\mathcal{Y}'| = |\mathcal{Y}| \cdot \text{poly}(\frac{|\Sigma_G|^m}{\beta})$. This is however arguably a minor blow-up since we typically expect that $|\Sigma_G|/\beta$ to be much smaller than *size*$(G) = |\mathcal{X}| \cdot |\mathcal{Y}|$.

3.4 Parallel repetition theorems

In this section we prove our main parallel repetition theorem.

Theorem (Theorem 3.24 restated)**.** *Let G' be a concatenated game (ε, δ)-weakly fortified against classical substrategies with inner game G. If $\delta \cdot (m-1) \cdot |\Sigma_G|^{m-1} \leq \eta$ then*

$$\text{val}(G'^{\otimes m}) \leq (\text{val}(G) + \varepsilon)^m + \eta. \tag{3.41}$$

Similarly, if G' is (ε, δ) weakly-fortified against entangled substrategies and $\delta \cdot (m-1) \cdot |\Sigma_G|^{m-1} \leq \eta$ then

$$\text{V}\small{\text{AL}}^*(G'^{\otimes m}) \leq (\text{V}\small{\text{AL}}^*(G) + \varepsilon)^m + \eta. \tag{3.42}$$

The proof follows directly from the following proposition.

Proposition 3.25. *Let $\{G'_i\}_{i=1}^t$ be a collection of concatenated games with inner games $\{G_i\}_{i=1}^t$. Suppose that G'_t is (ε, δ) weakly fortified against classical substrategies. Then,*

$$\mathrm{val}(G'_1 \otimes G'_2 \otimes \ldots \otimes G'_t) \leq (\mathrm{val}(G_t) + \varepsilon) \cdot \mathrm{val}(G'_1 \otimes G'_2 \otimes \ldots \otimes G'_{t-1}) + \delta \cdot \prod_{i=1}^{t-1} |\Sigma_{G_i}|. \quad (3.43)$$

Similarly, if G'_t is (ε, δ) weakly fortified against quantum substrategies, then

$$\textsc{Val}^*(G'_1 \otimes G'_2 \otimes \ldots \otimes G'_t) \leq (\textsc{Val}^*(G_t) + \varepsilon) \cdot \textsc{Val}^*(G'_1 \otimes G'_2 \otimes \ldots \otimes G'_{t-1}) + \delta \cdot \prod_{i=1}^{t-1} |\Sigma_{G_i}|. \quad (3.44)$$

The key to proving Proposition 3.25 is to work with the induced strategies. This allows us to get an additive error depending just on the alphabet size of the inner game. In the proof, we use the usual notation where a strategy missing an (answer) argument indicates summation over that variable. For example,

$$f(x_1, a_1, \ldots, x_{t-1}, a_{t-1}, x_t) = \sum_{a_t} f(x_1, a_1, \ldots, x_{t-1}, a_{t-1}, x_t, a_t).$$

Proof. We only prove (3.43) as the proof of (3.44) follows the same structure. Also for simplicity, we focus on the case of two-player games as the proof of the multiplayer case is a straightforward extension.

Consider any strategies $f : \mathcal{X}'_1 \times \mathcal{A}'_1 \times \ldots \times \mathcal{X}'_t \times \mathcal{A}'_t \to [0, 1]$, $g : \mathcal{Y}'_1 \times \mathcal{B}'_1 \times \ldots \times \mathcal{Y}'_t \times \mathcal{B}'_t \to [0, 1]$. To clarify notation we will denote tuples $(z_1, \ldots, z_{t-1})$ as $\mathbf{z}_{<t}$. With this notation, $\mathrm{val}(G'_1 \otimes \ldots \otimes G'_t, f, g)$ is precisely

$$\mathop{\mathbb{E}}_{(\mathbf{x}_{\leq t}, \mathbf{y}_{\leq t})} \mathop{\mathbb{E}}_{\mathbf{x'}_{\leq t}} \mathop{\mathbb{E}}_{\mathbf{y'}_{\leq t}} \sum_{\mathbf{a'}_{\leq t}, \mathbf{b'}_{\leq t}} \prod_{i=1}^{t} V(a'_i(x_i), b'_i(y_i), x_i, y_i)\, f(\mathbf{x'}_{\leq t}, \mathbf{a'}_{\leq t}) \cdot g(\mathbf{y'}_{\leq t}, \mathbf{b'}_{\leq t}), \quad (3.45)$$

where the expectations are according to $(x_i, y_i) \sim \mu_i$ and $x'_i \sim N(x_i)$ and $y'_i \sim N(y_i)$ for all $i = 1, \ldots, t$. As usual let

$$f(\mathbf{x}_{<t}, \mathbf{a}_{<t}, x'_t, a'_t) = \mathop{\mathbb{E}}_{\mathbf{x'}_{<t} \sim N(\mathbf{x}_{<t})} \sum_{a'_i(x_i) = a_i,\, i<t} f(\mathbf{x'}_{<t}, \mathbf{a'}_{<t}, x'_t, a'_t). \quad (3.46)$$

Using this notation, we can rewrite (3.45) as

$$\mathop{\mathbb{E}}_{(\mathbf{x}_{<t},\mathbf{y}_{<t})} \sum_{\mathbf{a}_{<t},\mathbf{b}_{<t}} \prod_{i=1}^{t-1} V(a_i,b_i,x_i,y_i)\, S(\mathbf{x}_{<t},\mathbf{y}_{<t},\mathbf{a}_{<t},\mathbf{b}_{<t}), \tag{3.47}$$

where $S(\mathbf{x}_{<t},\mathbf{y}_{<t},\mathbf{a}_{<t},\mathbf{b}_{<t})$ is given by

$$\mathop{\mathbb{E}}_{(x_t,y_t)} \mathop{\mathbb{E}}_{x'_t} \mathop{\mathbb{E}}_{y'_t} \sum_{a'_t,b'_t} V(a'_t(x_t),b'_t(y_t),x_t,y_t)\, f(\mathbf{x}_{<t},\mathbf{a}_{<t},x'_t,a'_t)\cdot g(\mathbf{y}_{<t},\mathbf{b}_{<t},y'_t,b'_t). \tag{3.48}$$

Consider the following substrategy G'_t: fix the first $2(t-1)$ arguments of f to $(\mathbf{x}_{<t},\mathbf{a}_{<t})$ and the first $2(t-1)$ arguments of g to $(\mathbf{y}_{<t},\mathbf{b}_{<t})$. Then (3.48) is precisely the value of this substrategy in G'_t. Since G'_t is (ε,δ) weakly fortified, it follows that

$$(3.48) \le (\mathrm{val}(G_t)+\varepsilon)\cdot \mathop{\mathbb{E}}_{(x_t,y_t)} f(\mathbf{x}_{<t},\mathbf{a}_{<t},x_t)\cdot g(\mathbf{y}_{<t},\mathbf{b}_{<t},y_t)+\delta. \tag{3.49}$$

Plugging this expression back into (3.47), $\mathrm{val}(G'_1\otimes\ldots\otimes G'_t,f,g)$ is bounded by

$$(\mathrm{val}(G_t)+\varepsilon) \mathop{\mathbb{E}}_{(\mathbf{x}_{\le t},\mathbf{y}_{\le t})} \sum_{\mathbf{a}_{<t},\mathbf{b}_{<t}} \prod_{i=1}^{t-1} V(a_i,b_i,x_i,y_i)\, f(\mathbf{x}_{<t},\mathbf{a}_{<t},x_t)\cdot g(\mathbf{y}_{<t},\mathbf{b}_{<t},y_t)+\delta\cdot\prod_{i=1}^{t-1}|\Sigma_{G_i}|.$$

To conclude we observe that

$$\mathop{\mathbb{E}}_{(\mathbf{x}_{\le t},\mathbf{y}_{\le t})} \sum_{\mathbf{a}_{<t},\mathbf{b}_{<t}} \prod_{i=1}^{t-1} V(a_i,b_i,x_i,y_i)\, f(\mathbf{x}_{<t},\mathbf{a}_{<t},x_t)\cdot g(\mathbf{y}_{<t},\mathbf{b}_{<t},y_t) \tag{3.50}$$

is at most $\mathrm{val}(G'_1\otimes\ldots\otimes G'_{t-1})$, as for any fixed (x_t,y_t) the functions $f(\cdot,x_t):\mathcal{X}'_1\times\mathcal{A}'_1\times\ldots\times\mathcal{X}'_{t-1}\times\mathcal{A}'_{t-1}\to[0,1]$ and $g(\cdot,y_t):\mathcal{Y}'_1\times\mathcal{B}'_1\times\ldots\times\mathcal{Y}'_{t-1}\times\mathcal{B}'_{t-1}\to[0,1]$ are valid strategies in $G'_1\otimes\ldots\otimes G'_{t-1}$.

□

Remark 3.26. Theorem 3.23 immediately follows from Proposition 3.25 by taking $t=2$ and considering the trivial concatenation $G'_1=G_1$, $G'_2=G_2$.

Theorem 3.24 follows easily.

Proof of Theorem 3.24. We prove (3.41) as the proof of (3.42) is similar.

The proof is by induction on m. The case $m = 1$ is clear. By the induction hypothesis we have

$$\mathrm{val}(G'^{\otimes(m-1)}) \leq (\mathrm{val}(G) + \varepsilon)^{m-1} + \delta \cdot (m-2)|\Sigma_G|^{m-2}.$$

Note that we can assume $\mathrm{val}(G)+\varepsilon < 1$ otherwise (3.41) holds trivially. Applying Proposition 3.25 we see that

$$\begin{aligned}\mathrm{val}(G'^{\otimes m}) &\leq (\mathrm{val}(G) + \varepsilon) \cdot \mathrm{val}(G'^{\otimes(m-1)}) + \delta \cdot |\Sigma_G|^{m-1} \\ &\leq (\mathrm{val}(G) + \varepsilon)^m + \delta \cdot (\mathrm{val}(G) + \varepsilon) \cdot (m-2)|\Sigma_G|^{m-2} + \delta \cdot |\Sigma_G|^{m-1} \\ &\leq (\mathrm{val}(G) + \varepsilon)^m + \delta \cdot (m-1) \cdot |\Sigma_G|^{m-1}.\end{aligned}$$

□

3.5 Classical fortification

In this section we prove our main theorem regarding the fortification of classical games. Beside providing a short and self-contained treatment of the main result of [48, 11], it serves as preparation for the analysis of Section 3.7.

Theorem (Theorem 3.2 restated). *Let G be a biregular game, M and P two bipartite λ-spectral expanders. If $\lambda \leq \frac{\varepsilon}{2}\sqrt{\frac{\delta}{2}}$, then the concatenated game $G' = (M \circ G \circ P)$ is (ε, δ) weakly fortified against classical substrategies.*

We note that it follows from [11, Appendix C] that the dependence λ and δ in Theorem 3.2 is up to constant factors optimal. On the other hand, the tightness of dependence of ε and δ does not seem to follow from [11] lower bound (however, δ is by far the more significant of the two parameters).

3.5.1 Proof of Theorem 3.2

We start with a simple claim whose proof we will defer to the end of the subsection.

Claim 3.27. *Let $M = (\mathcal{X}' \times \mathcal{X}, E)$ and $N = (\mathcal{Y}' \times \mathcal{Y}, F)$ be two biregular bipartite graphs that are λ-spectral expanders. Let μ be a distribution on $\mathcal{X} \times \mathcal{Y}$ such that both marginals of*

μ are uniform. Let $f : \mathcal{X}' \to \mathbb{R}$ and $g : \mathcal{Y}' \to \mathbb{R}$ be any functions. Define $f : \mathcal{X} \to \mathbb{R}$ and $g : \mathcal{Y} \to \mathbb{R}$ as the induced functions, given by $f : x \mapsto \mathbb{E}_{x' \sim N(x)} f(x')$, $g : y \mapsto \mathbb{E}_{y' \sim N(y)} g(y')$ respectively. Then

$$\mathop{\mathbb{E}}_{(x_1,y_1)\sim\mu} \Big| f(x_1)g(y_1) - \mathop{\mathbb{E}}_{(x_2,y_2)\sim\mu} f(x_2)g(y_2) \Big| \leq 2\sqrt{2}\lambda \Big(\mathop{\mathbb{E}}_{x'\sim\mathcal{X}'} |f(x')|^2 \Big)^{1/2} \Big(\mathop{\mathbb{E}}_{y'\sim\mathcal{Y}'} |g(y')|^2 \Big)^{1/2}$$

and

$$\Big| \mathop{\mathbb{E}}_{x_1\sim\mathcal{X}} f(x) \mathop{\mathbb{E}}_{y_1\sim\mathcal{Y}} g(y_1) - \mathop{\mathbb{E}}_{(x_2,y_2)\sim\mu} f(x_2)g(y_2) \Big| \leq 2\lambda^2 \Big(\mathop{\mathbb{E}}_{x'\sim\mathcal{X}'} |f(x')|^2 \Big)^{1/2} \Big(\mathop{\mathbb{E}}_{y'\sim\mathcal{Y}'} |g(y')|^2 \Big)^{1/2}.$$

We prove a slightly stronger statement which implies Theorem 3.2. Let f, g be any substrategies for G, and let $\gamma = \mathbb{E}_{(x,y)\sim\mu} f(x)g(y)$. We claim that

$$\mathrm{val}(G', f, g) \leq \mathrm{val}(G)\gamma + 2\sqrt{2}\,\lambda\sqrt{\gamma} + 4\,\lambda^2. \tag{3.51}$$

To deduce the bound claimed in Theorem 3.2 from (3.51) we distinguish two cases. Either $\gamma \leq \delta$, in which case using the trivial estimate $\mathrm{val}(G', f, g) \leq \gamma$ the bound immediately follows. Or $\gamma > \delta$, in which case

$$\begin{aligned} \mathrm{val}(G)\gamma + 2\sqrt{2}\lambda\sqrt{\gamma} + 4\lambda^2 &\leq \gamma(\mathrm{val}(G) + 2\sqrt{2}\lambda\delta^{-1/2}) + 4\lambda^2 \\ &\leq \gamma(\mathrm{val}(G) + \varepsilon) + \delta \end{aligned}$$

given the relation between ε, δ and λ expressed in the theorem.

It remains to prove (3.51). Fix substrategies f and g. We have

$$\begin{aligned} \mathrm{val}(G', f, g) &= \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{V(a,b,x,y)=1} f(x,a) \cdot g(y,b) \\ &= \mathop{\mathbb{E}}_{(x,y)\sim\mu} f(x)g(y) \sum_{V(a,b,x,y)=1} \frac{f(x,a)}{f(x)} \cdot \frac{g(y,b)}{g(y)}, \end{aligned}$$

where we adopt the convention that $0/0 = 0$. Using the triangle inequality,

$$\begin{aligned}\mathrm{val}(G', f, g) &\le \gamma \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{V(a,b,x,y)=1} \frac{f(x,a)}{f(x)} \cdot \frac{g(y,b)}{g(y)} + \mathop{\mathbb{E}}_{(x,y)\sim\mu} \Big| f(x)g(y) - \gamma \Big| \\ &\le \gamma \mathrm{val}(G) + \mathop{\mathbb{E}}_{(x,y)\sim\mu} \Big| f(x)g(y) - \gamma \Big|, \end{aligned} \tag{3.52}$$

where the second inequality follows since $(x, a) \mapsto f(x, a)/f(x)$ and $(y, b) \mapsto g(y, b)/g(y)$ form a valid pair of strategies for G. It remains to estimate the second term above. Applying the first inequality in Claim 3.27,

$$\begin{aligned}\mathop{\mathbb{E}}_{(x,y)\sim\mu} \Big| f(x)g(y) - \gamma \Big| &\le 2\sqrt{2}\lambda \Big(\mathop{\mathbb{E}}_{x'\sim\mathcal{X}'} |f(x')|^2 \Big)^{1/2} \Big(\mathop{\mathbb{E}}_{y'\sim\mathcal{Y}'} |g(y')|^2 \Big)^{1/2} \\ &\le 2\sqrt{2}\lambda \Big(\mathop{\mathbb{E}}_{x'\sim\mathcal{X}'} f(x') \mathop{\mathbb{E}}_{y'\sim\mathcal{Y}'} g(y') \Big)^{1/2} \\ &\le 2\sqrt{2}\lambda \sqrt{\gamma + 2\lambda^2} \\ &\le 2\sqrt{2}\lambda (\sqrt{\gamma} + \sqrt{2}\lambda) \\ &= 2\sqrt{2}\lambda\sqrt{\gamma} + 4\lambda^2, \end{aligned}$$

where in the second inequality we used $0 \le f(x'), g(y') \le 1$ for all x', y' and the third uses the second inequality in Claim 3.27. Together with (3.52) this proves (3.51).

Finally, we prove Claim 3.27.

Proof of Claim 3.27. For the first inequality, write

$$\begin{aligned} &\mathop{\mathbb{E}}_{(x_1,y_1)\sim\mu} \Big| f(x_1)g(y_1) - \mathop{\mathbb{E}}_{(x_2,y_2)\sim\mu} f(x_2)g(y_2) \Big| \\ &\quad\le \mathop{\mathbb{E}}_{(x_1,y_1),(x_2,y_2)\sim\mu} \Big(|f(x_1) - f(x_2)||g(y_1)| + |f(x_2)||g(y_1) - g(y_2)| \Big) \\ &\quad\le \Big(\mathop{\mathbb{E}}_{x_1,x_2\sim\mathcal{X}} |f(x_1) - f(x_2)|^2 \Big)^{1/2} \Big(\mathop{\mathbb{E}}_{y_1\sim\mathcal{Y}} |g(y_1)|^2 \Big)^{1/2} \\ &\qquad\qquad + \Big(\mathop{\mathbb{E}}_{x_2\sim\mathcal{X}} |f(x_2)|^2 \Big)^{1/2} \Big(\mathop{\mathbb{E}}_{y_1,y_2\sim\mathcal{Y}} |g(y_1) - g(y_2)|^2 \Big)^{1/2} \\ &\quad\le \lambda \Big(\mathop{\mathbb{E}}_{x_1',x_2'\sim\mathcal{X}'} |f(x_1') - f(x_2')|^2 \Big)^{1/2} \Big(\mathop{\mathbb{E}}_{y_1'\sim\mathcal{Y}'} |g(y_1')|^2 \Big)^{1/2} \\ &\qquad\qquad + \lambda \Big(\mathop{\mathbb{E}}_{x_2'\sim\mathcal{X}'} |f(x_2')|^2 \Big)^{1/2} \Big(\mathop{\mathbb{E}}_{y_1',y_2'\sim\mathcal{Y}'} |g(y_1') - g(y_2')|^2 \Big)^{1/2}, \end{aligned}$$

where the last inequality uses Proposition 3.20. Now note that $\mathbb{E}_{x_1',x_2'\sim\mathcal{X}'}|f(x_1')-f(x_2')|^2 \leq 2\,\mathbb{E}_{x'\sim\mathcal{X}'}|f(x')|^2$. Applying a similar bound for g gives us the first inequality. For the second, write

$$\begin{aligned}
\Big|\mathop{\mathbb{E}}_{(x_2,y_2)\sim\mu} f(x_2)g(y_2) \mathop{\mathbb{E}}_{x_1\sim\mathcal{X}} f(x) \mathop{\mathbb{E}}_{y_1\sim\mathcal{Y}} g(y_1)\Big| \\
&= \Big|\mathop{\mathbb{E}}_{(x_2,y_2)\sim\mu,x_1\sim\mathcal{X},y_1\sim\mathcal{Y}}(f(x_1)-f(x_2))(g(y_1)-g(y_2))\Big| \\
&\leq \Big(\mathop{\mathbb{E}}_{x_1,x_2\sim\mathcal{X}}(f(x_1)-f(x_2))^2\Big)^{1/2}\Big(\mathop{\mathbb{E}}_{y_1,y_2\sim\mathcal{Y}}(g(y_1)-g(y_2))^2\Big)^{1/2} \\
&\leq \lambda^2\Big(\mathop{\mathbb{E}}_{x_1',x_2'\sim\mathcal{X}'}(f(x_1)-f(x_2))^2\Big)^{1/2}\Big(\mathop{\mathbb{E}}_{y_1',y_2'\sim\mathcal{Y}'}(g(y_1')-g(y_2'))^2\Big)^{1/2} \\
&\leq 2\lambda^2\Big(\mathop{\mathbb{E}}_{x'\sim\mathcal{X}'}|f(x')|^2\Big)^{1/2}\Big(\mathop{\mathbb{E}}_{y'\sim\mathcal{Y}'}|g(y')|^2\Big)^{1/2}.
\end{aligned}$$

□

3.5.2 A simple multiplayer fortification

The following is a simple fortification theorem for k-player games. Since Theorem 3.24 applies equally well to the multiplayer setting, we get a hardness amplification result for classical multiplayer games.

Theorem 3.28. *Let G be a k-player game. Suppose G' is given by composing each of the k sides of G by a λ-spectral expander where $\lambda \leq 2\delta/k$. Then G' is a $(0,\delta)$ fortified game.*[7]

Proof. Consider a classical substrategy for G' given by $f_i : \mathcal{X}_i' \times \mathcal{A}_i' \to \mathbb{R}^+$ for $i = 1,2,\ldots,k$. As usual, denote $f_i : \mathcal{X}_i \times \mathcal{A}_i \to \mathbb{R}^+$ the projection of f_i to the inner game G. By definition,

$$\mathrm{val}(G,\{f_i\}_{i=1}^k) = \mathop{\mathbb{E}}_{(x_1,\ldots,x_k)}\sum_{a_1,a_2,\ldots,a_k} V(a_1,\ldots,a_k,x_1,\ldots,x_k)\cdot f_1(x_1,a_1)\cdot f_x(x_2,a_2)\ldots f_k(x_k,a_k).$$

We can rewrite the above as

$$\mathop{\mathbb{E}}_{(x_1,\ldots,x_k)}\prod_{i=1}^k f_i(x_i)\sum_{a_1,\ldots,a_k} V(a_1,\ldots,a_k,x_1,\ldots,x_k)\cdot\frac{f_1(x_1,a_1)\cdot f_x(x_2,a_2)\ldots f_k(x_k,a_k)}{f(x_1)\cdot f(x_2)\ldots\cdot f(x_k)}.$$

[7] Although there is no ε dependence in the above, when applied to 2-player games the theorem is still weaker than Theorem 3.2 because of the worse dependence on δ – which is the more crucial parameter than ε.

Let $\gamma = \mathbb{E}_{(x_1,\dots,x_k)} \prod_{i=1}^k f_i(x_i)$. Applying the triangle inequality,

$$\mathrm{val}(G, \{f_i\}_{i=1}^k) \le \gamma \cdot \mathrm{val}(G) + \mathop{\mathbb{E}}_{(x_1,\dots,x_k)} |\prod_{i=1}^k f_i(x_i) - \gamma|.$$

To conclude it will suffice to show the second term above is at most δ. Let $\overline{f_i} = \mathbb{E}_{x_i} f(x_i)$. Then

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x_1,\dots,x_k)} \left|\prod_{i=1}^k f_i(x_i) - \gamma\right| &\le \mathop{\mathbb{E}}_{x_1,\dots,x_k} \left|\prod_{i=1}^k f_i(x_i) - \prod_{i=1}^k \overline{f_i}\right| + \mathop{\mathbb{E}}_{(x_1,\dots,x_k)} \left|\prod_{i=1}^k \overline{f_i} - \gamma\right| \\
&= \mathop{\mathbb{E}}_{(x_1,\dots,x_k)} \left|\prod_{i=1}^k f_i(x_i) - \prod_{i=1}^k \overline{f_i}\right| + \left|\prod_{i=1}^k \overline{f_i} - \mathop{\mathbb{E}}_{x_1,\dots,x_k} \prod_{i=1}^k f_i(x_i)\right| \\
&\le 2 \cdot \mathop{\mathbb{E}}_{(x_1,\dots,x_k)} \left|\prod_{i=1}^k f_i(x_i) - \prod_{i=1}^k \overline{f_i}\right| \\
&\le 2 \sum_{i=1}^k \mathop{\mathbb{E}}_{x_i} |f_i(x_i) - \overline{f_i}|,
\end{aligned}$$

where the first equality is by definition of γ, the second inequality by convexity of $|\cdot|$, and the last follows from

$$|f_1(x_1)f_2(x_2)\dots f_k(x_k) - \overline{f_1 f_2} \dots \overline{f_k}| \le \sum_{\ell=1}^k \left|\prod_{i=1}^{\ell-1} f_i(x_i) \cdot \prod_{i=\ell}^k \overline{f_i} - \prod_{i=1}^{\ell} f_i(x_i) \cdot \prod_{i=\ell+1}^k \overline{f_i}\right| \le \sum_{i=1}^k \mathop{\mathbb{E}}_{x_i} |f_i(x_i) - \overline{f_i}|.$$

Hence,

$$\mathop{\mathbb{E}}_{x_1,\dots,x_k} \left|\prod_{i=1}^k f_i(x_i) - \gamma\right| \le 2 \sum_{i=1}^k \left(\mathop{\mathbb{E}}_{x_i} (f_i(x_i) - \overline{f_i})^2\right)^{1/2} \le 2\lambda \sum_{i=1}^k \left(\mathop{\mathbb{E}}_{x_i'} (f_i(x_i') - \overline{f_i})^2\right)^{1/2} \le 2\lambda k.$$

The desired result follows. □

3.6 From strong to weak fortification for entangled games

In this section, we start working toward the problem of fortifying games in the entangled case. In particular, we show how Theorem 3.5 follows from Theorem 3.3. Let $G = (\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mu, V)$ be a two-player game.

Definition 3.29. For a game G and integer $l \in \mathbb{N}$ let $G^{\oplus l}$ denote the disjoint union of l copies of G.

Suppose that M and P are regular bipartite graphs over $\mathcal{X}' \times \mathcal{X}$ and $\mathcal{Y}' \times \mathcal{Y}$, respectively. Suppose further that M and P are balanced, i.e. $|\mathcal{X}'| = |\mathcal{X}|$ and $|\mathcal{Y}'| = |\mathcal{Y}|$. Let d_M and d_N denote the degree of vertices M and P, respectively. (Note that since the graphs are balanced and regular, the left and right degrees are the same.)

Following [11], we assume that M and P are explicit bipartite almost-Ramanujan expanders, as provided e.g. by [12], for which the second-largest singular values λ_M and λ_P of A_M and A_P (the normalized adjacency matrices) respectively satisfy

$$\lambda_M = O\left(\frac{\mathrm{poly}(\log d_M)}{\sqrt{d_M}}\right), \qquad \lambda_P = O\left(\frac{\mathrm{poly}(\log d_P)}{\sqrt{d_P}}\right). \tag{3.53}$$

Note that if $d_M, d_P = \widetilde{\Omega}(\frac{1}{\varepsilon^2\delta})$ then Theorem 3.3 implies that $G' = (M \circ G \circ P)$ is (ε, δ) weakly-fortified. Next we recall the definition of G'_{OF} from the introduction.

Ordered fortification. Let G, M, P and $G' = (M \circ G \circ P)$ be as above. let $l = \max\{d_M, d_P\}$. In G'_{OF_l} (or simply G'_{OF} where $l = \max\{d_M, d_P\}$) the referee samples questions (x, y) as in G and selects two random neighbors $x' \in \mathcal{X}'$ and $y' \in \mathcal{Y}'$ of x and y in M and P respectively. Then the referee selects two random injective maps $r_{x'} : N(x') \to [l]$ and $s_{y'} : N(y') \to [l]$ under the condition $r_{x'}(x) = s_{y'}(y)$. Alice's question then is the pair $(x', r_{x'})$ and Bob's is the pair $(y', s_{y'})$. Alice outputs an answer tuple $a' : N(x') \to \mathcal{A}$ and Bob $b' : N(y') \to \mathcal{B}$. The players win if $V(a'(x), b'(y), x, y) = 1$.

Remark 3.30. Note that G'_{OF} has exactly the same answer alphabet size as G', the question sizes $|\mathcal{X}'_{OF}|$ and $|\mathcal{Y}'_{OF}|$ are larger than in G'. This blow-up can be mitigated as follows. It turns out that in Definition 3.31 the use of the complete set $S_{(d,l)}$ is unnecessary. More precisely, from the proof of the main claim of this section, Claim 3.34 below, it will be clear that the only condition required is that the permutations be chosen from a pairwise independent subset of $S_{(d,l)}$. Selecting the smallest possible such subset lets us reduce the blow-up in the size of the question sets from a multiplicative $D!$ down to $\mathrm{poly}(D) = \mathrm{poly}(\frac{1}{\varepsilon^2\delta})$. We omit the details.

Although it may not be immediately apparent, it is possible to view G'_{OF} as a concatenated game. Let $G^{\oplus l}$ be as in Definition 3.29. Note that $G^{\oplus l}$ has exactly the same classical and entangled value as G. Let $S_{(d_M,l)}$ denote the set of all injective maps from $[d_M] \to [l]$. Fix maps $u_{x'} : N(x') \to [d_M]$ and $v_{y'} : N(y') \to [d_M]$ ordering the neighborhoods of each x', y' in an arbitrary way.

Definition 3.31. Let M be a regular bipartite graph over $\mathcal{X}' \times \mathcal{X}$ as above . We define $\widetilde{M}$ as a bipartite graph over $\mathcal{X}'_{OF} := \mathcal{X}' \times S_{(d_M,l)}$ and $\mathcal{X}_{OF} := \mathcal{X} \times [l]$ where

$$(x', \pi) \sim_{\tilde{M}} (x, i) \qquad \iff \qquad \pi(u_{x'}(x)) = i.$$

We define $\widetilde{P}$ from P in a similar way.

Note that here $\pi \circ u_{x'}$ exactly corresponds to $r_{x'} : N(x') \to [l]$ map from the original definition of G'_{OF}. Hence, we obtain the following alternative characterization of G'_{OF}.

Proposition 3.32. *The game G'_{OF} constructed above is a concatenated game given by*

$$G'_{OF} = (\widetilde{M} \circ G^{\oplus l} \circ \widetilde{P}).$$

Next, we show that ordered fortification preserves the entangled value (the classical value is also preserved but that is not important here).

Proposition 3.33. *We have* $\text{VAL}^*(G'_{OF}) = \text{VAL}^*(G)$.

Proof. In one direction we have $\text{VAL}^*(G'_{OF}) \leq \text{VAL}^*(G^{\oplus l}) = \text{VAL}^*(G)$ where we used Propositions 3.11 and 3.32. For the other direction, consider any entangled strategy $(|\psi\rangle, \{A_x^a\}, \{B_y^b\})$ for G. We construct a strategy for $G'_{\oplus}$ that achieves the same value. The provers share l copies of the state $|\psi\rangle$, and each copy is assigned a unique label $i \in [l]$. Alice and Bob receive questions $(x', r_{x'})$ and $(y', s_{y'})$, respectively. For each $x \in N(x')$, Alice applies $\{A_x^a\}$ to the $r_{x'}(x)$-th copy of $|\psi\rangle$. Bob applies a similar strategy.

Since by construction the "true questions" x^* and y^* are given the same label, the distribution of answers obtained for x^* and y^* is identical to the distribution of answers obtained while playing G using $(|\psi\rangle, \{A_x^a\}, \{B_y^b\})$, hence achieving the same winning probability. □

The main technical step in reducing Theorem 3.5 to Theorem 3.3 is an analysis of the singular values of $\widetilde{M}, \widetilde{N}$ in terms of the singular values of M and N. We prove the following.

Claim 3.34. *Let M be a bipartite graph over $\mathcal{X}' \times \mathcal{X}$ as above and let λ_M denote the second largest singular value of M. Let $\tilde{M}$ be as in Definition 3.31. Then,*

$$\lambda_{\tilde{M}} \leq \max\left\{\lambda_M, \frac{1}{\sqrt{d_M - 1}}\right\}.$$

Since in our case $\lambda_M = O\left(\operatorname{poly}(\log d_M)/d_M\right)$, Claim 3.34 implies that $\lambda_{\tilde{M}}$ satisfies the same bound. Also note that a similar statement of course applies to $\tilde{P}$ and $\lambda_{\tilde{P}}$. So we see that Theorem 3.3, Propositions 3.33 and 3.32, and Claim 3.34 together imply Theorem 3.5; it remains to prove the latter.

Proof of Claim 3.34. Recall that by assumption M is a regular balanced bipartite graph. Let $d := d_M$ the degree of vertices in M. The normalized adjacency matrix of $\tilde{M}$ is given by

$$A_{\tilde{M}}((x', \pi), (x, i)) = \begin{cases} \frac{1}{d} \cdot \sqrt{\frac{(l-d)!}{(l-1)!}} & (x', \pi) \sim_{\tilde{M}} (x, i) \\ 0 & (x', \pi) \not\sim_{\tilde{M}} (x, i) \end{cases}. \tag{3.54}$$

We relate the second largest singular value λ_M of A_M and the second largest singular value $\lambda_{\tilde{M}}$ of $\tilde{M}$ by relating the eigenvalues of $B = A_M^\top A_M$ and $C = A_{\tilde{M}}^\top A_{\tilde{M}}$. We can explicitly compute the entries of B and C. For B,

$$B(x_1, x_2) = \frac{|\{x' \in \mathcal{X}' \,:\, \{x_1, x_2\} \subset N(x')\}|}{d^2}, \tag{3.55}$$

and in particular $B(x, x) = \frac{1}{d}$ for all $x \in \mathcal{X}$. To compute entries of C, first note that when $x_1 \neq x_2$ and $i \neq j$ we have

$$C((x_1, i), (x_2, j)) = \frac{|\{x' \in \mathcal{X}' \,:\{x_1, x_2\} \subset N(x')\}| \cdot (l-d)!}{d^2 (l-1)!} \cdot \frac{(l-2)!}{(l-d)!} = \frac{B(x_1, x_2)}{l-1}. \tag{3.56}$$

Finally, observe the following special cases:

- $C((x, i), (x, i)) = \frac{1}{d}$.

- $C((x_1, i), (x_2, i)) = 0$ if $x_1 \neq x_2$.

- $C((x, i), (x, j)) = 0$ when $i \neq j$.

Let $\widetilde{B} = B - \frac{1}{d}\mathbb{I}$ and $\widetilde{C} = C - \frac{1}{d}\mathbb{I}$. Let $\widetilde{J} = \frac{1}{l-1}(J - \mathbb{I})$ be the $l \times l$ matrix that is $(l-1)^{-1}$ in the off-diagonal entries, and 0 along the diagonal. Then it is easy to see that

$$\widetilde{C} = \widetilde{B} \otimes \widetilde{J}. \tag{3.57}$$

The matrix $\widetilde{J}$ has a single eigenvalue equal to 1 and $l-1$ eigenvalues equal to $-\frac{1}{l-1}$, and $\widetilde{B}$ has a single eigenvalue equal to $1 - 1/d$ and the remaining are in the range $[-\frac{1}{d}, \lambda_M^2 - \frac{1}{d}]$. It follows that the top eigenvalue of $C = \widetilde{C} + \frac{1}{d}\mathbb{I}$ is 1 (as expected) and the next one satisfies

$$\lambda_{\tilde{M}} \leq \max\left\{\lambda_M, \sqrt{\frac{1}{d(l-1)} + \frac{1}{d}}\right\},$$

which is bounded by $\max\left\{\lambda_M, \frac{1}{\sqrt{d-1}}\right\}$ since $l \geq d$. □

3.7 Weak fortification of entangled games

In this section is to prove the following.

Theorem (Theorem 3.3 restated)**.** *Let $G' = (M \circ G \circ P)$ be a concatenated game obtained by concatenating two sides of a game G with some λ-spectral expanders M and P. If $\lambda \leq \frac{\varepsilon^2\delta}{56}$, then G' is (ε, δ) weakly-fortified against entangled substrategies.*

At a high level, the proof of Theorem 3.3 follows the same outline as the classical proof of Section 3.5.

Consider a substrategy $\{A_{x'}^{a'}\}_{(x',a')\in\mathcal{X}'\times A'}$, $\{B_{y'}^{b'}\}_{(y',b')\in\mathcal{Y}'\times B'}$ for G'. Define $A_x = \mathbb{E}_{x\sim N(x')} A_{x'}$ and $B_y = \mathbb{E}_{y\sim N(y')} B_{y'}$.[8] Define $A = \mathbb{E}_{x\sim\mu_\mathcal{X}} A_x$ and $B = \mathbb{E}_{y\sim\mu_\mathcal{Y}} B_y$. To prove Theorem 3.3 we

[8] In what follows, we assume without loss of generality that all $A_{x'}$ and $B_{y'}$ are invertible. Note that proving Theorem 3.3 for this subset of substrategies suffices. This follows by a limiting argument because of the continuity of (3.30) in $A_{x'}$ and $B_{y'}$.

must analyze the following expression:

$$\begin{aligned}\textsc{Val}^*(G', \{A^{a'}_{x'}\}, \{B^{b'}_{y'}\}) &= \mathop{\mathbb{E}}_{(x,y)\sim\mu} \mathop{\mathbb{E}}_{x'\sim N(x), y'\sim N(y)} \sum_{a',b'} V(a'_{x'}(x), b'_{y'}(y), x, y) \cdot \mathrm{Tr}(A^{a'}_{x'}\rho^{1/2}B^{b'}_{y'}\rho^{1/2}) \\ &= \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b} V(a,b,x,y) \cdot \mathrm{Tr}(A^a_x\rho^{1/2}B^b_y\rho^{1/2}) \\ &= \mathop{\mathbb{E}}_{(x,y)\sim\mu} \mathrm{Tr}(A_x\rho^{1/2}B_y\rho^{1/2}) \cdot \sum_{V(a,b,x,y)=1} \frac{\mathrm{Tr}(A^a_x\rho^{1/2}B^b_y\rho^{1/2})}{\mathrm{Tr}(A_x\rho^{1/2}B_y\rho^{1/2})},\end{aligned}$$

where A^a_x and B^b_y are defined as in (3.11), and we use the convention that $0/0 = 0$. Our analysis splits into two cases. First let us consider the *small case*. This is handled by the following proposition.

Proposition 3.35. *Suppose* $\mathrm{Tr}(\rho^{1/2}A\rho^{1/2}B) < \delta/2$. *Then* $\textsc{Val}^*(G', \{A^{a'}_{x'}\}, \{B^{b'}_{y'}\}) < \delta$.

Proof. First of all we have

$$\textsc{Val}^*(G', \{A^{a'}_{x'}\}, \{B^{b'}_{y'}\}) = \mathop{\mathbb{E}}_{x,y} \sum_{V(a,b,x,y)=1} \mathrm{Tr}(A^a_x\rho^{1/2}B^b_y\rho^{1/2}) \le \mathop{\mathbb{E}}_{x,y} \mathrm{Tr}(A_x\rho^{1/2}B_y\rho^{1/2}).$$

Subtracting $\mathrm{Tr}(A\rho^{1/2}B\rho^{1/2})$,

$$\mathop{\mathbb{E}}_{x,y} \mathrm{Tr}(A_x\rho^{1/2}B_y\rho^{1/2}) - \mathrm{Tr}(A\rho^{1/2}B\rho^{1/2}) = \mathop{\mathbb{E}}_{x,y} \mathrm{Tr}((A_x - A)\rho^{1/2}(B_y - B)\rho^{1/2}).$$

By applying Cauchy-Schwarz to the latter expression and using Claim 2.3 it follows that

$$\textsc{Val}^*(G', \{A^{a'}_{x'}\}, \{B^{b'}_{y'}\}) \le \delta/2 + \lambda^2;$$

this is smaller than δ by the choice of λ. □

The large case. In this case, the hypothesis of Proposition 3.35 is not satisfied and without loss of generality we assume that

$$\min\Big\{\mathrm{Tr}_\rho(A),\ \mathrm{Tr}_\rho(B)\Big\} \ge \mathrm{Tr}(A\rho^{1/2}B\rho^{1/2}) \ge \delta/2. \tag{3.58}$$

Let

$$\gamma := \mathop{\mathbb{E}}_{(x,y)\sim\mu} \mathrm{Tr}(A_x \rho^{1/2} B_y \rho^{1/2}). \tag{3.59}$$

By the triangle inequality,

$$\textsc{Val}^*(G', A_{x'}, B_{y'}) \leq \mathop{\mathbb{E}}_{(x,y)\sim\mu} |\mathrm{Tr}(A_x \rho^{1/2} B_y \rho^{1/2}) - \gamma| + \gamma \cdot \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{V(a,b,x,y)=1} \frac{\mathrm{Tr}(A_x^a \rho^{1/2} B_y^b \rho^{1/2})}{\mathrm{Tr}(A_x \rho^{1/2} B_y \rho^{1/2})}. \tag{3.60}$$

To bound the first term, we use the triangle inequality to get

$$|\mathrm{Tr}(A_x \rho^{1/2} B_y \rho^{1/2}) - \gamma| \leq |\mathrm{Tr}(A_x \rho^{1/2} B_y \rho^{1/2} - A\rho^{1/2} B \rho^{1/2})| + |\mathrm{Tr}(A\rho^{1/2} B \rho^{1/2}) - \gamma|. \tag{3.61}$$

The first term on the right-hand side of (3.61) can be bounded as

$$\begin{aligned}
&\mathop{\mathbb{E}}_{(x,y)\sim\mu} |\mathrm{Tr}(A_x \rho^{1/2} B_y \rho^{1/2} - A\rho^{1/2} B \rho^{1/2})| \\
&\qquad \leq \mathop{\mathbb{E}}_{(x,y)\sim\mu} |\mathrm{Tr}(A_x \rho^{1/2} (B_y - B) \rho^{1/2})| + \mathop{\mathbb{E}}_{x} |\mathrm{Tr}((A_x - A)\rho^{1/2} B \rho^{1/2})| \\
&\qquad \leq \mathop{\mathbb{E}}_{(x,y)\sim\mu} \Big[\mathrm{Tr}_\rho(A_x^2)^{1/2} \cdot \mathrm{Tr}_\rho((B_y - B)^2)^{1/2}\Big] + \mathop{\mathbb{E}}_{x} \Big[\mathrm{Tr}_\rho(B^2)^{1/2} \cdot \mathrm{Tr}_\rho((A_x - A)^2)^{1/2}\Big] \\
&\qquad \leq \Big(\mathop{\mathbb{E}}_{x} \mathrm{Tr}_\rho(A_x^2)\Big)^{1/2} \cdot \Big(\mathop{\mathbb{E}}_{y} \mathrm{Tr}_\rho((B_y - B)^2)\Big)^{1/2} + \mathrm{Tr}_\rho(B^2)^{1/2} \cdot \Big(\mathop{\mathbb{E}}_{x} \mathrm{Tr}_\rho((A_x - A)^2)\Big)^{1/2} \\
&\qquad \leq 4 \cdot \lambda,
\end{aligned} \tag{3.62}$$

where the first inequality is the triangle inequality, the next two follow from Cauchy-Schwarz, and the last from Claim 2.3 and the trivial bounds $\mathrm{Tr}_\rho(A_x^2), \mathrm{Tr}_\rho(B^2) \leq \mathrm{Tr}(\rho) = 1$. To bound the second term on the right-hand side of (3.61) we note that

$$\begin{aligned}
|\mathrm{Tr}(A\rho^{1/2} B \rho^{1/2}) - \gamma| &= |\mathop{\mathbb{E}}_{(x,y)\sim\mu} \mathrm{Tr}((A_x - A)\rho^{1/2} (B_y - B) \rho^{1/2})| \\
&\leq (\mathop{\mathbb{E}}_{x} \mathrm{Tr}_\rho[(A_x - A)^2])^{1/2} \cdot (\mathop{\mathbb{E}}_{y} \mathrm{Tr}_\rho[(B_y - B)^2])^{1/2},
\end{aligned}$$

and the latter is again bounded by 2λ by Claim 2.3. In total we have

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu} |\mathrm{Tr}(A_x \rho^{1/2} B_y \rho^{1/2}) - \gamma| \leq 4\lambda + 2\lambda^2 \leq \delta, \tag{3.63}$$

which provides an upper bound on the first term in the right-hand side of (3.60).

To bound the second term term in the right-hand side of (3.60) we use a strategy inspired in part by the parallel repetition theorem of [28]. Let U_x, V_y, U, V be a family of unitaries such that the operators

$$\Lambda_x = U_x\sqrt{A_x}\rho^{1/4}, \quad \Lambda = U\sqrt{A}\rho^{1/4}, \quad \Gamma_y = V_y\sqrt{B_y}\rho^{1/4}, \quad \Gamma = V\sqrt{B}\rho^{1/4} \tag{3.64}$$

are all positive semidefinite, which is possible by Fact 2.4. Note that this in particular implies that $\Lambda_x = \Lambda_x^\dagger$ and hence

$$\Lambda_x^2 = \Lambda_x^\dagger\Lambda_x = \rho^{1/4}\sqrt{A_x}U_x^\dagger U_x\sqrt{A_x}\rho^{1/4} = \rho^{1/4}A_x\rho^{1/4}, \tag{3.65}$$

and similarly $\Lambda^2 = \rho^{1/4}A\rho^{1/4}$, $\Gamma^2 = \rho^{1/4}B\rho^{1/4}$ and so on.

Define "rescaled" strategies by

$$\widehat{A_x^a} = U_x A_x^{-1/2}A_x^a A_x^{-1/2}U_x^\dagger, \quad \widehat{B_y^b} = V_y B_y^{-1/2}B_y^b B_y^{-1/2}V_y^\dagger, \tag{3.66}$$

where A_x^{-1}, B_y^{-1}'s are the pseudo-inverses of A_x, B_y. Note that the operators (3.66) satisfy $\widehat{A_x} = \sum_a \widehat{A_x^a}$, $\widehat{B_y} = \sum_b \widehat{B_y^b} \leq \mathbb{I}$ as required. Let

$$K_{xy} = \frac{U_x A_x^{1/2}\rho^{1/2}B_y^{1/2}V_y^\dagger}{\sqrt{\mathrm{Tr}(A_x\rho^{1/2}B_y\rho^{1/2})}}, \qquad K = \frac{UA^{1/2}\rho^{1/2}B^{1/2}V^\dagger}{\sqrt{\mathrm{Tr}(A\rho^{1/2}B\rho^{1/2})}}. \tag{3.67}$$

By definition of $\Lambda_x, \Gamma_y, \mathcal{X}, \mathcal{Y}$ we see that the above is equivalent to

$$K_{xy} = \frac{\Lambda_x\Gamma_y}{\sqrt{\mathrm{Tr}(\Lambda_x^2\Gamma_y^2)}}, \qquad K = \frac{\Lambda\Gamma}{\sqrt{\mathrm{Tr}(\Lambda^2\Gamma^2)}}. \tag{3.68}$$

Now note the following identity

$$\frac{\mathrm{Tr}(A_x^a\rho^{1/2}B_y^b\rho^{1/2})}{\mathrm{Tr}(A_x\rho^{1/2}B_y\rho^{1/2})} = \mathrm{Tr}(\widehat{A_x^a}K_{xy}\widehat{B_y^b}K_{xy}^\dagger). \tag{3.69}$$

So to finish the argument it suffices to estimate

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{V(a,b,x,y)=1} \mathrm{Tr}(\widehat{A_x^a} K_{xy} \widehat{B_y^b} K_{xy}^\dagger). \tag{3.70}$$

To this end note that since $\mathrm{Tr}(KK^\dagger) = 1$ it follows from the definition of $\textsc{Val}^*(G)$ that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{V(a,b,x,y)=1} \mathrm{Tr}(K\widehat{A_x^a} K^\dagger \widehat{B_y^b}) \leq \textsc{Val}^*(G). \tag{3.71}$$

To conclude we use the following proposition.

Proposition 3.36. *Let K_{xy} and K be as above. Then*

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu} \|K_{xy} - K\|_F^2 \leq \frac{12\lambda}{\delta}. \tag{3.72}$$

Before proving the proposition let us see how it implies the desired bound on the second term of (3.60).

$$\begin{aligned} &|\mathrm{Tr}(K_{xy}\widehat{A_x^a} K_{xy}^\dagger \widehat{B_y^b}) - \mathrm{Tr}(K\widehat{A_x^a} K^\dagger \widehat{B_y^b})| \\ &\qquad \leq |\mathrm{Tr}((K_{xy} - K)\widehat{A_x^a} K_{xy}^\dagger \widehat{B_y^b})| + |\mathrm{Tr}(K\widehat{A_x^a}(K_{xy}^\dagger - K^\dagger)\widehat{B_y^b})| \\ &\qquad \leq \mathrm{Tr}((K_{xy} - K)\widehat{A_x^a}(K_{xy} - K)^\dagger \widehat{B_y^b})^{1/2} \cdot \mathrm{Tr}(K_{xy}\widehat{A_x^a} K_{xy}^\dagger \widehat{B_y^b})^{1/2} \\ &\qquad\quad + \mathrm{Tr}((K_{xy} - K)\widehat{A_x^a}(K_{xy} - K)^\dagger \widehat{B_y^b})^{1/2} \cdot \mathrm{Tr}(K\widehat{A_x^a} K^\dagger \widehat{B_y^b})^{1/2}. \end{aligned} \tag{3.73}$$

Averaging with $\mathbb{E}_{(x,y)\sim\mu} \sum_{V(a,b,x,y)=1}$ and applying Cauchy-Schwarz we see that (3.73) is bounded by

$$\begin{aligned} &\Big[\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{V(a,b,x,y)=1} \mathrm{Tr}((K_{xy} - K)\widehat{A_x^a}(K_{xy} - K)^\dagger \widehat{B_y^b})\Big]^{1/2} \cdot \Big[\Big(\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{V(a,b,x,y)=1} \mathrm{Tr}(K_{xy}\widehat{A_x^a} K_{xy}^\dagger \widehat{B_y^b})\Big)^{1/2} \\ &\qquad + \Big(\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{V(a,b,x,y)=1} \mathrm{Tr}(K\widehat{A_x^a} K^\dagger \widehat{B_y^b})\Big)^{1/2}\Big]. \end{aligned} \tag{3.74}$$

We claim that the second term in brackets is at most 2. To see this note that replacing the sum from $\sum_{V(a,b,x,y)=1}$ to a $\sum_{a,b}$ only increase the term, and the claim follows from

$\mathrm{Tr}(K_{xy}K_{xy}^\dagger) = \mathrm{Tr}(KK^\dagger) = 1$. To bound the first term in (3.74), we again relax the summation from $\sum_{V(a,b,x,y)=1}$ to $\sum_{a,b}$. This is valid because all the operators of the form $(K_{xy} - K)\widehat{A_x^a}(K_{xy} - K)^\dagger, B_y \geq 0$ and hence all the additional terms introduced in the sum are nonnegative. The desired result follows because

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b} \mathrm{Tr}((K_{xy}-K)\widehat{A_x^a}(K_{xy}-K)^\dagger \widehat{B_y^b}) \leq \mathop{\mathbb{E}}_{(x,y)\sim\mu} \mathrm{Tr}((K_{x,y}-K)(K_{x,y}-K)^\dagger) = \mathop{\mathbb{E}}_{x,y} \|K_{x,y}-K\|_F^2, \tag{3.75}$$

which is bounded by Proposition 3.36. Combining all bounds, from (3.73) we get

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{V(a,b,x,y)=1} \frac{\mathrm{Tr}(A_x^a\rho^{1/2}B_y^b\rho^{1/2})}{\mathrm{Tr}(A_x\rho^{1/2}B_y\rho^{1/2})} &= \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{V(a,b,x,y)=1} \mathrm{Tr}(\widehat{A_x^a}K_{xy}\widehat{B_y^b}K_{xy}^\dagger) \\
&\leq \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{V(a,b,x,y)=1} \mathrm{Tr}(\widehat{A_x^a}K\widehat{B_y^b}K^\dagger) \\
&\qquad + |\mathrm{Tr}(K_{xy}\widehat{A_x^a}K_{xy}^\dagger\widehat{B_y^b}) - \mathrm{Tr}(K\widehat{A_x^a}K^\dagger\widehat{B_y^b})| \\
&\leq \textsc{Val}^*(G) + 2\cdot\left(\mathop{\mathbb{E}}_{(x,y)\sim\mu} \|K_{xy} - K\|_F^2\right)^{1/2} \\
&\leq \textsc{Val}^*(G) + 2\sqrt{\frac{12\lambda}{\delta}}.
\end{aligned}$$

The latter is bounded by ε by the choice of λ. It only remains to prove Proposition 3.36.

Proof of Proposition 3.36. We have

$$\|K_{xy} - K\|_F \leq \left\| \frac{\Lambda_x\Gamma_y}{\sqrt{\mathrm{Tr}(\Lambda_x^2\Gamma_y^2)}} - \frac{\Lambda\Gamma}{\sqrt{\mathrm{Tr}(\Lambda^2\Gamma^2)}} \right\|_F + \left\| \frac{\Lambda_x\Gamma_y}{\sqrt{\mathrm{Tr}(\Lambda^2\Gamma^2)}} - \frac{\Lambda\Gamma}{\sqrt{\mathrm{Tr}(\Lambda^2\Gamma^2)}} \right\|_F. \tag{3.76}$$

For the first term,

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x,y)\sim\mu} \left\| \frac{\Lambda_x\Gamma_y}{\sqrt{\mathrm{Tr}(\Lambda_x^2\Gamma_y^2)}} - \frac{\Lambda_x\Gamma_y}{\sqrt{\mathrm{Tr}(\Lambda^2\Gamma^2)}} \right\|_F^2 &= \mathop{\mathbb{E}}_{(x,y)\sim\mu} \mathrm{Tr}(\Lambda_x^2\Gamma_y^2)\cdot\left(\frac{1}{\sqrt{\mathrm{Tr}(\Lambda_x^2\Gamma_y^2)}} - \frac{1}{\sqrt{\mathrm{Tr}(\Lambda^2\Gamma^2)}}\right)^2 \\
&= \frac{1}{\mathrm{Tr}(\Lambda^2\Gamma^2)} \mathop{\mathbb{E}}_{(x,y)\sim\mu} \left(\sqrt{\mathrm{Tr}(\Lambda_x^2\Gamma_y^2)} - \sqrt{\mathrm{Tr}(\Lambda^2\Gamma^2)}\right)^2 \\
&\leq \frac{1}{\mathrm{Tr}(\Lambda^2\Gamma^2)} \mathop{\mathbb{E}}_{(x,y)\sim\mu} |\mathrm{Tr}(\Lambda_x^2\Gamma_y^2) - \mathrm{Tr}(\Lambda^2\Gamma^2)| \\
&\leq \frac{1}{\mathrm{Tr}(\Lambda^2\Gamma^2)} \mathop{\mathbb{E}}_{(x,y)\sim\mu} |\mathrm{Tr}((\Lambda_x^2 - \Lambda^2)\Gamma_y^2)| + |\mathrm{Tr}(\Lambda^2(\Gamma_y^2 - \mathcal{Y}^2))|
\end{aligned}$$

$$\leq \frac{1}{\mathrm{Tr}(\Lambda^2\Gamma^2)} \left|\mathop{\mathbb{E}}_{x}[\mathrm{Tr}((\Lambda_x^2-\Lambda^2)^2)]\right|^{1/2} \cdot \left|\mathop{\mathbb{E}}_{y}[\mathrm{Tr}(\Gamma_y^4)]\right|^{1/2} \tag{3.77}$$

$$+ \frac{1}{\mathrm{Tr}(\Lambda^2\Gamma^2)} \left|\mathop{\mathbb{E}}_{x}[\mathrm{Tr}((\Gamma_y^2-\Gamma^2)^2)]\right|^{1/2} \cdot \mathrm{Tr}(\Lambda^4)^{1/2}, \tag{3.78}$$

where the last step follows from two applications of Cauchy-Schwarz. Rewriting the above in terms of A_x, B_y and ρ using (3.65) and its analogues we see that the term in (3.77) equals

$$\frac{1}{\mathrm{Tr}(A\rho^{1/2}B\rho^{1/2})} \left|\mathop{\mathbb{E}}_{x}[\mathrm{Tr}((A_x-A)\rho^{1/2}(A_x-A)\rho^{1/2})]\right|^{1/2} \cdot \left|\mathop{\mathbb{E}}_{y}[\mathrm{Tr}(B_y\rho^{1/2}B_y\rho^{1/2})\right|^{1/2} \tag{3.79}$$

Bounding the last term $\mathrm{Tr}(B_y\rho^{1/2}B_y\rho^{1/2})$ by 1 and the first term by 2λ (which follows by applying Fact 2.5 and Claim 2.3) and doing the same analysis for (3.78) we see that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu} \left\| \frac{\Lambda_x\Gamma_y}{\sqrt{\mathrm{Tr}(\Lambda_x^2\Gamma_y^2)}} - \frac{\Lambda_x\Gamma_y}{\sqrt{\mathrm{Tr}(\Lambda^2\Gamma^2)}} \right\|_F^2 \leq \frac{8\lambda}{\delta}. \tag{3.80}$$

To bound the second term in (3.76) we argue as follows:

$$\begin{aligned} \|\Lambda_x\Gamma_y - \Lambda\Gamma\|_F^2 &\leq 2\cdot\|(\Lambda_x-\Lambda)\Gamma_y\|_F^2 + 2\cdot\|(\Gamma_y-\Gamma)\mathcal{X}\|_F^2 \\ &= 2\cdot\mathrm{Tr}(\Gamma_y^2(\Lambda_x-\Lambda)^2) + 2\cdot\mathrm{Tr}((\Gamma_y-\Gamma)^2\mathcal{X}^2) \qquad (3.81) \\ &\leq 2\cdot\mathrm{Tr}(\Gamma_y^4)^{1/2}\cdot\mathrm{Tr}((\Lambda_x-\Lambda)^4)^{1/2} + 2\cdot\mathrm{Tr}(\Lambda^4)^{1/2}\cdot\mathrm{Tr}((\Gamma_y-\Gamma)^4)^{1/2} \qquad (3.82) \\ &\leq 2\cdot\mathrm{Tr}(\Gamma_y^4)^{1/2}\cdot\mathrm{Tr}[(\Lambda_x^2-\Lambda^2)^2]^{1/2} + 2\cdot\mathrm{Tr}(\Lambda^4)^{1/2}\cdot\mathrm{Tr}[(\Gamma_y^2-\Gamma^2)^2]^{1/2}, \qquad (3.83) \end{aligned}$$

where in the last step we used Lemma 2.6. Using the same bound on the above terms as in the above we see that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu} \|\Lambda_x\Gamma_y - \Lambda\Gamma\|_F^2 \leq 8\lambda. \tag{3.84}$$

Since in the large case $\mathrm{Tr}(A\rho^{1/2}B\rho^{1/2}) \geq \frac{\delta}{2}$ the result follows. □

Chapter 4

Parallel Repetition via Anchoring

This chapter is based on the paper *Hardness amplification for entangled games via anchoring*, a joint work with T. Vidick and H. Yuen, to be published in the Proceedings of 49th Annual ACM Symposium on the Theory of Computing (STOC 2017) and also presented at the Quantum Information Processing Conference (QIP 2016) as a plenary talk.

4.1 Introduction

In this chapter, we study the parallel repetition of a class of games which we call *anchored*. We also prove an exponential-decay parallel repetition theorem for anchored games that involve any number of entangled players. We also prove a threshold version of our parallel repetition theorem for anchored games.

Let us start by giving a definition of anchored games in full generality.

Definition 4.1 (Multiplayer Anchored Games)**.** A game $G = (\mathcal{X}, \mathcal{A}, \mu, V)$ is called α-anchored if there exists $\mathcal{X}^t_\perp \subseteq \mathcal{X}^t$ for all $t \in [k]$ where

1. $\mu(\mathcal{X}^t_\perp) \geq \alpha$ for all $t \in [k]$, and

2. for all $x \in \mathcal{X}$,
$$\mu(x) = \mu(x|_{\overline{F}_x}) \cdot \prod_{t \in F_x} \mu(x^t) \tag{4.1}$$

where for all question tuples $x = (x^1, x^2, \ldots, x^k) \in \mathcal{X}$, $F_x \subseteq [n]$ denotes the set of coordinates of x that lie in the anchor, i.e.

$$F_x = \{t \in [k] \, : x^t \in \mathcal{X}_\perp^t\}$$

and $\overline{F}_x$ denotes the complement, i.e., $[n] - F_x$.

Here for a set $S \subseteq [n]$, $\mu(x|_S)$ denotes the marginal probability of the question tuple x restricted to the coordinates in S, i.e.

$$\mu(x|_S) = \sum_{x'|_S = x|_S} \mu(x').$$

When $k = 2$ this definition coincides with the definition of two-player anchored games in Definition 4.6. Additionally, just like the two-player case, one can easily extend the anchoring transformation given in Definition 1.4 to arbitrary k-player games:

Proposition 4.2. *Let $G = (\mathcal{X}, \mathcal{A}, \mu, V)$ be a k-player game. Let $G_\perp$ be the k-player game where the referee samples $(x^1, x^2, \ldots, x^k)$ according to μ, replaces each x^t with an auxiliary symbol $\perp$ independently with probability α, and checks the players' answers according to V if all $x^t \neq \perp$, and otherwise the referee accepts. Then $G_\perp$ is an α-anchored game satisfying*

$$\mathrm{val}(G_\perp) = 1 - (1-\alpha)^k \cdot (1 - \mathrm{val}(G)), \qquad \mathrm{val}^*(G_\perp) = 1 - (1-\alpha)^k \cdot (1 - \mathrm{val}^*(G)). \quad (4.2)$$

Proof. We give the proof for the classical value; the same argument carries over to the entangled value. First, it is clear that $\mathrm{val}(G_\perp) \geq (1 - (1 - \alpha^k)) + (1-\alpha)^k \cdot \mathrm{val}(G)$. For the other direction, consider an optimal strategy for $G_\perp$. Under this strategy, we can express the entangled value as

$$\mathrm{val}(G_\perp) = (1-\alpha)^k \cdot \Pr(W | \forall t,\, x^t \neq \perp) + (1 - (1-\alpha^k)) \cdot \Pr(W | \exists\, t \text{ s.t. } x^t = \perp)$$

where W is the event that the players win. The optimal strategy for $G_\perp$ yields a strategy for G that wins with probability $\Pr(W | \forall t,\, x^t \neq \perp)$, which can be at most $\mathrm{val}(G)$. Since $\Pr(W | \exists\, t \text{ s.t. } x^t = \perp) = 1$, we obtain the desired equality. □

4.2 Application to the quantum PCP conjecture

Just as the the classical parallel repetition theorem is useful for proving hardness of approximation results, one might expect that a *quantum* parallel repetition theorem would be useful for proving *quantum* hardness of approximation results. However, we do not (yet) have a Quantum PCP theorem; as of writing this is an active field of research. Furthermore, while the classical PCP theorem has three equivalent formulations – one in terms of probabilistically checkable proofs, one in terms of hardness of approximation for constraint satisfaction problems (CSP), and one in terms of games – only two out of the three corresponding formulations of the Quantum PCP Conjecture are known to be equivalent.

The following is the formulation of the Quantum PCP Conjecture that is analogous to the classical CSP formulation. (We refer to the survey [2] for further background on the conjecture, including explanations of the standard technical terms we use below.)

Conjecture 4.3 (Quantum PCP Conjecture, constraint satisfaction formulation)**.** There exists a constant $0 < \gamma < 1$ and integer $k \geq 2$ and $d \geq 2$ for which the following problem is QMA-hard: Given $a, b \in [0,1]$ such that $a - b \geq \gamma$ and a k-local Hamiltonian $H = H_1 + \cdots + H_m$ acting on n qudits of local dimension d such that $0 \leq H \leq \mathbb{I}$, decide whether the smallest eigenvalue of H is at least a or at most b, promised that one is the case.

This problem is known as the k-Local Hamiltonian problem with *constant promise gap*, where by promise gap we mean the gap γ between the thresholds a and b. The problem is only known to be QMA-hard for gaps γ that are inverse polynomial in n [44].

A *games* version of the conjecture is introduced in [32]:

Conjecture 4.4 (Quantum PCP Conjecture, games formulation)**.** There exists a constant $\gamma \in (0,1)$ and integers $s \geq 1, k \geq 2$ for which the following problem is QMA-hard: Given $a, b \in [0,1]$ such that $a - b \geq \gamma$, and a k-player game G where each player answers with s bits, decide whether $\mathrm{val}^*(G) \geq a$ or $\mathrm{val}^*(G) \leq b$, promised that one is the case.

When $\mathrm{val}^*(\cdot)$ is replaced with $\mathrm{val}(\cdot)$, the above conjecture is exactly equivalent to the classical PCP theorem. For constant gap γ it was proved by [60] that the problem of approximating the entangled value of a game is at least NP-hard. For inverse polynomial

γ the problem was shown QMA-hard [41], and very recently it was even shown to be NEXP-hard [42].

Though neither Conjecture 4.3 nor Conjecture 4.4 has been solved, we can nonetheless explore the consequences if they were true. We give a simple application of our parallel repetition for anchored games: assuming the truth of Conjecture 4.4, we can boost its hardness to any desired gap between completeness and soundness.

Proposition 4.5. *If Conjecture 4.4 is true, then for all $\delta > 0$ the following problem is* QMA*-hard: given a description of a k-player game G with answer size that depends only on δ, distinguish between* $\mathrm{val}^*(G) \geq 1-\delta$ *or* $\mathrm{val}^*(G) \leq \delta$*, promised that one is the case.*

Proof. Let $0 \leq b < a \leq 1$ be a promise gap satisfying the conditions of Conjecture 4.4. Define $a' = (1+3a)/4$, and $b' = (1+3b)/4$. Consider the following reduction: given a description of a k-player game G, promised that either $\mathrm{val}^*(G) \leq b$ or $\mathrm{val}^*(G) \geq a$, outputs the description of the following *threshold game* $G_\perp^{t,\geq\tau}$: the referee plays $G_\perp^{\otimes t}$, the t-fold repetition of $G_\perp$, the anchored version of G, but instead accepts iff the players win at least $\tau := (a' - \frac{a'-b'}{4})t$ games. We set parameters $\Delta = (a'-b')/4$ and $t = \frac{s}{c} \cdot \frac{2}{\Delta^9} \cdot \ln \frac{1}{\delta}$, where s is the length of the players' answers in G, and c is the universal constant from Theorem 4.16.

We get that if $\mathrm{val}^*(G) \geq a$, then $\mathrm{val}^*(G_\perp) \geq a'$. One strategy for $G_\perp^{t,\geq\tau}$ is for the players to play each coordinately independently using the optimal strategy for $G_\perp$. By a Chernoff-Hoeffding bound, the probability that they win at least τ games is at least

$$\mathrm{val}^*(G_\perp^{t,\geq\tau}) \geq 1 - \exp(-t\Delta^2/2) \geq 1-\delta.$$

Otherwise, $\mathrm{val}^*(G) \leq b$. Applying Theorem 4.16, we get that

$$\mathrm{val}^*(G_\perp^{t,\geq\tau}) \leq \left(1 - \Delta^9/2\right)^{c_k t/s} \leq \delta.$$

Observe that this reduction is efficient: the size of the description of $G_\perp^{t,\geq\tau}$ is $O(|G|^t)$; assuming the truth of Conjecture 4.4 this means that $a' - b' = \Omega(a-b) = \Omega(1)$, and thus since δ and s are constant, t is constant. The answer size of the new game is still $O(1)$. Thus the reduction runs in time polynomial in the input instance size, so if there were an algorithm that could distinguish between $\text{val}^*(G_\perp^{t,\geq\tau}) \geq 1-\delta$ or $\text{val}^*(G_\perp^{t,\geq\tau}) \leq \delta$, then this would distinguish between whether $\text{val}^*(G) \geq a$ or $\text{val}^*(G) \leq b$, respectively. □

We point out that we used two features of the anchoring transformation: first, that it allows us to analyze the repetition of arbitrary k-player games; second, it yields threshold theorems for parallel repetition.

4.3 Technical overview

We give a technical overview of anchored games and their parallel repetition. For concreteness we focus on the case of two-player games, though all the things discussed in this section can be appropriately generalized to the k-player games.

Definition 4.6 (Two-player anchored games)**.** Let G be a two-player game with question alphabet $\mathcal{X} \times \mathcal{Y}$ and distribution μ. For any $0 < \alpha \leq 1$ we say that G is α-anchored if there exists subsets $\mathcal{X}_\perp \subseteq \mathcal{X}$ and $\mathcal{Y}_\perp \subseteq \mathcal{Y}$ such that, denoting by μ the respective marginals of μ on both coordinates,

1. Both $\mu(\mathcal{X}_\perp), \mu(Y_\perp) \geq \alpha$,

2. Whenever $x \in \mathcal{X}_\perp$ or $y \in \mathcal{Y}_\perp$ it holds that $\mu(x, y) = \mu(x) \cdot \mu(y)$.

Informally, a game is *anchored* if each player independently has a significant probability of receiving a question from the set of "anchor questions" $\mathcal{X}_\perp$ and $\mathcal{Y}_\perp$. An alternative way of thinking about the class of anchored games is to consider the case where μ is uniform over a set of edges in a bipartite graph on vertex set $\mathcal{X} \times \mathcal{Y}$; then the condition is that the induced subgraph on $\mathcal{X}_\perp \times \mathcal{Y}_\perp$ is a complete bipartite graph that is connected to the rest of $\mathcal{X} \times \mathcal{Y}$ and has weight at least α. In other words, a game G is anchored if it contains a free game that is connected to the entire game.

It is easy to see that the games $G_\perp$ output by the anchoring transformation given in Definition 1.4 are α-anchored. Free games are automatically 1-anchored (set $\mathcal{X}_\perp = \mathcal{X}$ and $\mathcal{Y}_\perp = \mathcal{Y}$), but the class of anchored games is much broader; indeed assuming the Exponential Time Hypothesis it is unlikely that there exists a similar (efficient) reduction from general games to free games [1]. Additionally, since free games are anchored games, our parallel repetition theorems automatically reproduce the quantum and multiplayer parallel repetition of free games results of [40, 19, 20], albeit with worse parameters.

Dependency-breaking variables and states. Essentially all known proofs of parallel repetition proceed via reduction, showing how a "too good" strategy for the repeated game $G^{\otimes n}$ can be "rounded" into a strategy for G with success probability strictly greater than $\mathrm{val}(G)$, yielding a contradiction.

Let S^n be a strategy for $G^{\otimes n}$ that has a high success probability. By an inductive argument one can identify a set of coordinates C and an index i such that $\Pr(\text{Players win round } i|W) > \mathrm{val}(G) + \delta$, where W is the event that the players' answers satisfy the predicate V in all instances of G indexed by C. Given a pair of questions (x, y) in G the strategy S embeds them in the i-th coordinate of a n-tuple of questions

$$x_{[n]}y_{[n]} = \begin{pmatrix} x_1, x_2, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n \\ y_1, y_2, \ldots, y_{i-1}, y, y_{i+1}, \ldots, y_n \end{pmatrix}$$

that is distributed according to $\mathsf{P}_{X_{[n]}Y_{[n]}|X_i=x,Y_i=y,W}$. The players then simulate S^n on $x_{[n]}$ and $y_{[n]}$ respectively to obtain answers $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$, and return (a_i, b_i) as their answers in G. The strategy S succeeds with probability precisely $\Pr(\text{Win } i|W)$ in G, yielding the desired contradiction.

As S^n need not be a product strategy, conditioning on W may introduce correlations that make $\mathsf{P}_{X_{[n]}Y_{[n]}|X_i=x,Y_i=y,W}$ impossible to sample exactly. A key insight in Raz' proof of parallel repetition is that it is still possible for the players to *approximately* sample from this distribution. Drawing on the work of Razborov [55], Raz introduced a *dependency-breaking variable* Ω with the following properties:

(a) Given $\omega \sim \mathsf{P}_\Omega$ the players can locally sample $x_{[n]}$ and $y_{[n]}$ according to $\mathsf{P}_{X_{[n]}Y_{[n]}|X_i=x,Y_i=y,W}$,

(b) The players can jointly sample from P_Ω using shared randomness.

In [37] Ω is defined so that a sample ω fixes at least one of $\{x_{i'}, y_{i'}\}$ for each $i' \neq i$. It can then be shown that conditioned on x, Ω is nearly (though not exactly) independent of y, and vice-versa. In other words,

$$\mathsf{P}_{\Omega|X_i=x,W} \approx \mathsf{P}_{\Omega|X_i=x,Y_i=y,W} \approx \mathsf{P}_{\Omega|Y_i=y,W} \tag{4.3}$$

where "$\approx$" denotes closeness in statistical distance. Eq. (4.3) suffices to guarantee that the players can *approximately* sample the same ω from $\mathsf{P}_{\Omega|X_i=x,Y_i=y,W}$ with high probability, achieving point (b) above. This sampling is accomplished through a technique called *correlated sampling.*

This argument relies heavily on the assumption that there are only two players who employ a deterministic strategy. With more than two players, it is not known how to design an appropriate dependency-breaking variable Ω that satisfies requirements (a) and (b) above: in order to be jointly sampleable, Ω needs to fix as few inputs as possible; in order to allow players to locally sample their inputs conditioned on Ω, the variable needs to fix as many inputs as possible. These two requirements are in direct conflict as soon as there are more than two players.

In the quantum case the rounding argument seems to require that Alice and Bob jointly sample a *dependency-breaking state* $|\Omega_{x,y}\rangle$, which again depends on both their inputs. Although it is technically more complicated, as a first approximation $|\Omega_{x,y}\rangle$ can be thought of as the players' post-measurement state, conditioned on W. Designing a state that simultaneously allows Alice and Bob to (a) simulate the execution of the i-th game in $G^{\otimes n}$ conditioned on W, and (b) locally generate $|\Omega_{x,y}\rangle$ without communication is the main obstacle to proving a fully general parallel repetition theorem for entangled games.

It has long been known that in the free games case (i.e. games with product question distributions) these troubles with the dependency-breaking variable disappear, and consequently we have parallel repetition theorems for free games for the multiplayer and quantum settings [20]. With free games involving more than two players, it can be shown that

$$\mathsf{P}_{\Omega|X_i=x,Y_i=y,Z_i=z,\ldots,W} \approx \mathsf{P}_{\Omega|W}, \tag{4.4}$$

on average over question tuples $(x, y, z, \ldots)$. In the quantum case, [40, 19, 20] showed how to construct dependency-breaking states $|\Omega_{X_i=x,Y_i=y,W}\rangle$ and local unitaries U_x and V_y such that

$$(U_x \otimes V_y)|\Omega\rangle \approx |\Omega_{X_i=x,Y_i=y,W}\rangle \tag{4.5}$$

for some fixed quantum state $|\Omega\rangle$. This eliminates the need for the players to use correlated sampling, as they can simply share a sample from $\mathsf{P}_{\Omega|W}$ or the quantum state $|\Omega\rangle$ from the outset.

Breaking correlations in repeated anchored games. Rather than providing a complete extension of the framework of Raz and Holenstein to the multiplayer and quantum settings, we interpolate between the case of free games and the general setting by showing how the same framework of dependency-breaking variables and states can be extended to anchored games – without using correlated sampling. We introduce dependency-breaking variables Ω and states $|\Phi_{x,y}\rangle$ so that we can prove analogous statements to (4.4) and (4.5) in the anchored games setting.

The analysis for anchored games is more intricate than for free games. Proofs of the analogous statements for free games in [40, 19, 20] make crucial use of the fact that all possible question tuples are possible. An anchored game can be far from having this property. Instead, we use the anchors as a "home base" that is connected to all questions. Intuitively, no matter what question tuple $(x, y, z, \ldots)$ we are considering, it is only a few replacements away from the set of anchor questions. Thus the dependency of the variable Ω or state $|\Phi_{x,y}\rangle$ on the questions can be iteratively removed by "switching" each players' question to an anchor as

$$\mathsf{P}_{\Omega|X_i=x,Y_i=y,Z_i=z,W} \approx \mathsf{P}_{\Omega|X_i=x,Y_i=y,Z_i\in\perp,W} \approx \mathsf{P}_{\Omega|X_i=x,Y_i\in\perp,Z_i\in\perp,W} \approx \mathsf{P}_{\Omega|X_i\in\perp,Y_i\in\perp,Z_i\in\perp,W},$$

where "$X_i \in \perp$" is shorthand for the event that $X_i \in \mathcal{X}_\perp$.

Dealing with quantum strategies adds another layer of complexity to the argument. The local unitaries U_x and V_y involved in (4.5) are quite important in the arguments of [40, 19, 20]. The difficulty in extending the argument for free games to the case of general games is to show that these local unitaries each only depend on the input to a single player. In fact with the definition of $|\Omega_{x,y}\rangle$ used in these works it appears likely that this statement does not hold, thus a different approach must be found.

When the game is anchored, however, we are able to use the anchor question in order to show the existence of unitaries U_x and V_y that achieve (4.5) and depend only on a single player's question each. Achieving this requires us to introduce dependency-breaking states $|\Omega_{x,y}\rangle$ that are more complicated than those used in the free games case; in particular they include information about the *classical* dependency-breaking variables of Raz and Holenstein.

We prove (4.5) for anchored games by proving a sequence of approximate equalities: first we show that for most x there exists U_x such that $(U_x \otimes \mathbb{I})|\Omega_{\perp,\perp}\rangle \approx |\Omega_{x,\perp}\rangle$, where $|\Omega_{\perp,\perp}\rangle$ denotes the dependency-breaking state in the case that both Alice and Bob receive the anchor question "$\perp$", and $|\Omega_{x,\perp}\rangle$ denotes the state when Alice receives x and Bob receives "$\perp$". Then we show that for all y such that $\mu(y|x) > 0$ there exists a unitary V_y such that $(\mathbb{I} \otimes V_y)|\Omega_{x,\perp}\rangle \approx |\Omega_{x,y}\rangle$. Accomplishing this step requires ideas and techniques going beyond those in the free games case. Interestingly, a crucial component of our proof is to argue the existence of a local unitary $R_{x,y}$ that depends on *both* inputs x and y. The unitary $R_{x,y}$ is not implemented by Alice or Bob in the simulation, but it is needed to show that V_y maps $|\Omega_{x,\perp}\rangle$ onto $|\Omega_{x,y}\rangle$.

One can view our work as pushing the limits of arguments for parallel repetition that do not require some form of correlated sampling, a procedure that seems inherently necessary to analyze the general case. Our results demonstrate that such procedure is not needed for the purpose of achieving strong gap amplification theorems for multiplayer and quantum games.

4.4 Parallel repetition of anchored games with entangled players

This section is devoted to the analysis of the entangled value of repeated anchored games. The main theorem we prove is the following:

Theorem 4.7. *Let G be a k-player α-anchored game satisfying* $\mathrm{val}^*(G) = 1 - \varepsilon$. *Then*

$$\mathrm{val}^*(G^{\otimes n}) \leq \exp\left(-\Omega\left(\frac{\mathrm{poly}(\alpha^k) \cdot \varepsilon^8 \cdot n}{\mathrm{poly}(k) \cdot s}\right)\right),$$

where s is the total length of the answers output by the players.

For clarity we will focus on the $k = 2$ (two-player) case; we will describe how to extend the proof to arbitrary k at the end. We fix an α-anchored two-player game $G = (\mathcal{X}\times\mathcal{Y}, \mathcal{A}\times\mathcal{B}, \mu, V)$ with entangled value $\mathrm{val}^*(G) = 1 - \varepsilon$ and anchor sets $\mathcal{X}_\perp \subseteq \mathcal{X}$, $\mathcal{Y}_\perp \subseteq \mathcal{Y}$ for Alice and Bob, respectively. We also fix an optimal strategy for $G^{\otimes n}$, consisting of a shared entangled state $|\psi\rangle^{E_A E_B}$ and POVMs $\{A_{x^n}^{a^n}\}$ and $\{B_{y^n}^{b^n}\}$ for Alice and Bob respectively. Without loss of generality we assume that $|\psi\rangle$ is invariant under permutation of the two registers, i.e. there exist basis vectors $\{|v_j\rangle\}_j$ such that $|\psi\rangle = \sum_j \sqrt{\lambda_j}|v_j\rangle|v_j\rangle$.

4.4.1 Setup

We introduce the random variables, entangled states and operators that play an important role in the proof of Theorem 4.7. The section is divided into three parts: first we define the dependency-breaking variable Ω. Then we state useful lemmas about conditioned distributions. Finally we describe the states and operators used in the proof.

Dependency-breaking variables. Let $C \subseteq [n]$ a fixed set of coordinates for the repeated game $G^{\otimes n}$. We will assume that $C = \{m+1, m+2, \ldots, n\}$, where $m = n - |C|$, as this will easily be seen to hold without loss of generality. Let (X^n, Y^n) be distributed according to μ^n and (A^n, B^n) be defined from X^n and Y^n as follows:

$$\mathsf{P}_{A^n B^n | X^n = x^n, Y^n = y^n}(a^n, b^n) = \langle\psi| A_{x^n}^{a^n} \otimes B_{y^n}^{b^n} |\psi\rangle.$$

Let (X_C, Y_C) and $\mathrm{Z} = (A_C, B_C)$ denote the players' questions and answers respectively associated with the coordinates indexed by C. For $i \in [n]$ let W_i denote the event that the players win round i while playing $G^{\otimes n}$. Let $W_C = \bigwedge_{i \in C} W_i$.

We use the same dependency-breaking variable Ω that is used in Holenstein's proof of parallel repetition. In those works, for all $i \in [n]$, Ω_i fixes at least one of X_i or Y_i (and sometimes both, if $i \in C$). Thus, conditioned on Ω, X^n and Y^n are independent of each other.

In more detail, let $D_1, \ldots, D_m$ be independent and uniformly distributed over $\{A, B\}$. Let $M_1, \ldots, M_m$ be independent random variables defined in the following way. If $D_i = A$, then M_i is coupled to X_i (that is, takes the same value as X_i). Otherwise, if $D_i = B$, then M_i is coupled to Y_i. Then $\Omega_i = (D_i, M_i)$, and $\Omega = (\Omega_1, \ldots, \Omega_m, X_C, Y_C)$.

Conditioned distributions. Define $\delta_C := \frac{1}{m}\left(\log 1/\Pr(W_C) + |C| \log |\mathcal{A}||\mathcal{B}|\right)$. For notational convenience we often use the shorthand $X_i \in \perp$ and $Y_i \in \perp$ to stand for $X_i \in \mathcal{X}_\perp$ and $Y_i \in \mathcal{Y}_\perp$, respectively. The following lemma essentially follows from the classical arguments used in [37].

Lemma 4.8. *The following statements hold on, average over i chosen uniformly in* $[m]$*:*

1. $\mathbb{E}_i \left\|\mathsf{P}_{D_i M_i X_i Y_i | W_C} - \mathsf{P}_{D_i M_i X_i Y_i}\right\| \leq O(\sqrt{\delta_C})$
2. $\mathbb{E}_i \left\|\mathsf{P}_{\Omega Z X_i Y_i | W_C} - \mathsf{P}_{\Omega Z | W_C} \mathsf{P}_{X_i Y_i | \Omega}\right\| \leq O(\sqrt{\delta_C})$
3. $\mathbb{E}_i \left\|\mathsf{P}_{X_i Y_i} \mathsf{P}_{\Omega_{-i} Z | X_i \in \perp, Y_i \in \perp, W_C} - \mathsf{P}_{X_i Y_i} \mathsf{P}_{\Omega_{-i} Z | X_i Y_i W_C}\right\| \leq O(\sqrt{\delta_C}/\alpha^2)$
4. $\mathbb{E}_i \left\|\mathsf{P}_{X_i Y_i} \mathsf{P}_{\Omega_{-i} Z | X_i Y_i W_C} - \mathsf{P}_{X_i Y_i \Omega_{-i} Z | W}\right\| \leq O(\sqrt{\delta_C}/\alpha^2)$

Quantum states and operators. Recall that we have fixed an optimal strategy for Alice and Bob in the game $G^{\otimes n}$. This specifies a shared entangled state $|\psi\rangle$, and measurement operators $\{A_{x^n}^{a^n}\}$ for Alice and $\{B_{y^n}^{b^n}\}$ for Bob.

Operators. Define, for all a_C, b_C, x^n, y^n:

$$A_{x^n}^{a_C} := \sum_{a^n | a_C} A_{x^n}^{a^n} \qquad\qquad B_{y^n}^{b_C} := \sum_{b^n | b_C} B_{y^n}^{b^n}$$

where $a^n | a_C$ (resp. $b^n | b_C$) indicates summing over all tuples a^n consistent with the suffix a_C (resp. b^n consistent with suffix b_C). For all i, ω_{-i}, x_i, and y_i define:

$$A_{\omega_{-i}, x_i}^{a_C} = \mathop{\mathbb{E}}_{X^n | \omega_{-i}, x_i} A_{x^n}^{a_C} \qquad\qquad B_{\omega_{-i}, y_i}^{b_C} = \mathop{\mathbb{E}}_{Y^n | \omega_{-i}, y_i} B_{y^n}^{b_C}$$

where recall that $\mathbb{E}_{X^n|\omega_{-i},x_i}$ is shorthand for $\mathbb{E}_{X^n|\Omega_{-i}=\omega_{-i},X_i=x_i}$. Intuitively, these operators represent the "average" measurement that Alice and Bob apply, conditioned on $\Omega_{-i} = \omega_{-i}$, and $X_i = x_i$ and $Y_i = y_i$. Next, define

$$A^{a_C}_{\omega_{-i},\perp} := \mathop{\mathbb{E}}_{X^n|\Omega_{-i}=\omega_{-i}\wedge X_i \in \perp} A^{a_C}_{x^n} \qquad B^{b_C}_{\omega_{-i},\perp} := \mathop{\mathbb{E}}_{Y^n|\Omega_{-i}=\omega_{-i}\wedge Y_i \in \perp} B^{b_C}_{y^n}.$$

These operators represent the "average" measurement performed by Alice and Bob, conditioned on $\Omega_{-i} = \omega_{-i}$ and $M_i = \perp$. Finally, for all $x_i \in \mathcal{X}$ and $y_i \in \mathcal{Y}$, define

$$A^{a_C}_{\omega_{-i},\perp/x_i} := \frac{1}{2}A^{a_C}_{\omega_{-i},\perp} + \frac{1}{2}A^{a_C}_{\omega_{-i},x_i} \qquad B^{b_C}_{\omega_{-i},\perp/y_i} := \frac{1}{2}B^{b_C}_{\omega_{-i},\perp} + \frac{1}{2}B^{b_C}_{\omega_{-i},y_i}.$$

Intuitively, these operators represent the "average" measurements conditioned on $\Omega_{-i} = \omega_{-i}$ and when X_i is x_i with probability 1/2 and $\perp$ with probability 1/2 (or when $Y_i = y_i$ with probability 1/2 and $\perp$ with probability 1/2).

For notational convenience we often suppress the dependence on $(i, \omega_{-i}, z = (a_C, b_C))$ when it is clea from context. Thus, when we refer to an operator such as $A_{\perp/x}$, we really mean the operator $A^{a_C}_{\omega_{-i},\perp/x_i}$.

States. For all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, define the following (unnormalized) states:

$$\begin{aligned} |\Phi_{x,y}\rangle &:= \sqrt{A_x} \otimes \sqrt{B_y}|\psi\rangle & |\Phi_{x,\perp}\rangle &:= \sqrt{A_x} \otimes \sqrt{B_\perp}|\psi\rangle \\ |\Phi_{\perp/x,\perp}\rangle &:= \sqrt{A_{\perp/x}} \otimes \sqrt{B_\perp}|\psi\rangle & |\Phi_{\perp/x,y}\rangle &:= \sqrt{A_{\perp/x}} \otimes \sqrt{B_y}|\psi\rangle \qquad (4.6) \\ |\Phi_{\perp,\perp}\rangle &:= \sqrt{A_\perp} \otimes \sqrt{B_\perp}|\psi\rangle \end{aligned}$$

together with the normalization factors

$$\begin{aligned} \gamma_{x,y} &:= \| |\Phi_{x,y}\rangle \| & \gamma_{x,\perp} &:= \| |\Phi_{x,\perp}\rangle \| \\ \gamma_{\perp/x,\perp} &:= \left\| |\Phi_{\perp/x,\perp}\rangle \right\| & \gamma_{\perp/x,y} &:= \left\| |\Phi_{\perp/x,y}\rangle \right\| \\ \gamma_{\perp,\perp} &:= \| |\Phi_{\perp,\perp}\rangle \| \end{aligned}$$

Note that these normalization factors are the square-roots of the probabilities that a certain pair of answers $z = (a_C, b_C)$ occurred, given the specified inputs and the dependency-breaking variables. For example, revealing the depencies on ω_{-i} and z, we have

$$\gamma^{\omega_{-i},z}_{x_i,y_i} = \sqrt{\mathsf{P}_{\mathsf{Z}|\omega_{-i},x_i,y_i}(z)}.$$

We denote the normalized states by $|\widetilde{\Phi}_{x,y}\rangle = |\Phi_{x,y}\rangle/\gamma_{x,y}$, $|\widetilde{\Phi}_{x,\perp}\rangle = |\Phi_{x,\perp}\rangle/\gamma_{x,\perp}$, $|\widetilde{\Phi}_{\perp/x,\perp}\rangle = |\Phi_{\perp,\perp}\rangle/\gamma_{\perp/x,\perp}$, $|\widetilde{\Phi}_{\perp/x,\perp/y}\rangle = |\Phi_{\perp/x,y}\rangle/\gamma_{\perp/x,y}$, and $|\widetilde{\Phi}_{\perp,\perp}\rangle = |\Phi_{\perp,\perp}\rangle/\gamma_{\perp,\perp}$.

4.4.2 Proof of the parallel repetition theorem

Lemma 4.9. *Let G be an α-anchored two-player game. Let $C \subset [n]$ be a set of coordinates. Then*

$$\mathop{\mathbb{E}}_{i \notin C} \Pr(W_i | W_C) \leq \mathrm{val}^*(G) + O(\delta_C^{1/8}/\alpha^2)$$

where the expectation is over a uniformly chosen $i \in [n] \backslash C$ and $\delta_C = \frac{1}{m}\left(\log 1/\Pr(W_C) + |C| \log |\mathcal{A}||\mathcal{B}|\right)$.

Proof. For every ω_{-i}, $z = (a_C, b_C)$, $x_i \in \mathcal{X}$, $y_i \in \mathcal{Y}$, $a_i \in \mathcal{A}$ and $b_i \in \mathcal{B}$, define

$$\hat{A}^{a_i}_{\omega_{-i},x_i} := \sum_{a^n | a_i, a_C} (A^{a_C}_{\omega_{-i},x_i})^{-1/2} A^{a^n}_{\omega_{-i},x_i} (A^{a_C}_{\omega_{-i},x_i})^{-1/2}$$
$$\hat{B}^{b_i}_{\omega_{-i},y_i} := \sum_{b^n | b_i, b_C} (B^{b_C}_{\omega_{-i},y_i})^{-1/2} B^{b^n}_{\omega_{-i},y_i} (B^{b_C}_{\omega_{-i},y_i})^{-1/2}$$

where $a^n | a_i, a_C$ (resp. $b^n | b_i, b_C$) denotes summing over tuples a^n that are consistent with a_C and a_i (resp. b^n that are consistent with b_C and b_i). Note that the $\{\hat{A}^{a_i}_{\omega_{-i},x_i}\}_{a_i}$ and $\{\hat{B}^{b_i}_{\omega_{-i},y_i}\}_{b_i}$ are positive semidefinite operators that sum to identity, so form valid POVMs.

Consider the following strategy to play game G. Alice and Bob share classical public randomness, and for every setting of i, ω_{-i}, z, the bipartite state $|\widetilde{\Phi}^{\omega_{-i},z}_{\perp,\perp}\rangle$. Upon receiving questions $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ respectively they perform the following:

1. Alice and Bob use public randomness to sample (i, ω_{-i}, z) conditioned on W_C.

2. Alice applies $U_{\omega_{-i},z,x}$ to her register of $|\widetilde{\Phi}^{\omega_{-i},z}_{\perp,\perp}\rangle$.

3. Bob applies $V_{\omega_{-i},z,y}$ to his register of $|\widetilde{\Phi}^{\omega_{-i},z}_{\perp,\perp}\rangle$.

4. Alice measures with POVM operators $\{\hat{A}^{a_i}_{\omega_{-i},x}\}$ and returns the outcome as her answer.

5. Bob measures with POVM operators $\{\hat{B}^{b_i}_{\omega_{-i},y}\}$ and returns the outcome as his answer.

Suppose that, upon receiving questions (x, y) and after jointly picking a uniformly random $i \in [m]$, Alice and Bob could jointly sample ω_{-i}, z from $\mathsf{P}_{\Omega_{-i}\mathrm{Z}|W_C}$ and locally prepare the state $|\widetilde{\Phi}_{\substack{\omega_{-i},z \\ x,y}}\rangle$. For a fixed (x, y), ω_{-i} and z, the distribution of outcomes (a_i, b_i) after measuring $\{\hat{A}^{a_i}_{\omega_{-i},x} \otimes \hat{B}^{b_i}_{\omega_{-i},y}\}_{a_i,b_i}$ will be identical to $\mathsf{P}_{A_iB_i|\omega_{-i},z,x,y}$ (where we mean conditioning on $X_i = x$ and $Y_i = y$). Averaging over $(x, y) \sim \mu$, i, ω_{-i}, and z, the above-defined strategy will win game G with probability at least $\mathbb{E}_i \Pr(W_i|W_C)$.

Next we show that Alice and Bob are able to *approximately* prepare $|\widetilde{\Phi}_{\substack{\omega_{-i},z \\ x,y}}\rangle$ with high probability, and thus produce answers that are approximately distributed according to $\mathsf{P}_{A_iB_i|\omega_{-i},z,x,y}$, allowing them to win game G with probability greater than $1 - \varepsilon$ — a contradiction.

For the remainder of the proof, we will fix C and implicitly carry it around. Let $\delta = \delta_C$. We use the following lemma:

Lemma 4.10. *For every C, i, ω_{-i}, $z = (a_C, b_C)$, x_i and y_i there exists unitaries U_{ω_{-i},z,x_i} acting on E_A and V_{ω_{-i},z,y_i} acting on E_B such that*

$$\frac{1}{m}\sum_i \mathop{\mathbb{E}}_{X_iY_i} \mathop{\mathbb{E}}_{\Omega_{-i}\mathrm{Z}|W} \left\| (U_{\omega_{-i},z,x_i} \otimes V_{\omega_{-i},z,y_i}) \left|\widetilde{\Phi}_{\substack{\omega_{-i},z \\ \perp,\perp}}\right\rangle - \left|\widetilde{\Phi}_{\substack{\omega_{-i},z \\ x_i,y_i}}\right\rangle \right\|^2 = O(\delta^{1/4}/\alpha^4).$$

The proof of Lemma 4.10 is given in Section 4.4.2, and we assume it for now. Using the fact that for two pure states $|\psi\rangle$ and $|\phi\rangle$, $\|\psi - \phi\|_1 \leq \sqrt{2}\| \, |\psi\rangle - |\phi\rangle \|$, as well as Jensen's inequality,

$$\mathop{\mathbb{E}}_{i} \mathop{\mathbb{E}}_{XY} \mathop{\mathbb{E}}_{\Omega_{-i}\mathrm{Z}|W_C} \left\| (U_{\omega_{-i},z,x} \otimes V_{\omega_{-i},z,y}) \left[\widetilde{\Phi}_{\substack{\omega_{-i},z \\ \perp,\perp}}\right] - \widetilde{\Phi}_{\substack{\omega_{-i},z \\ x,y}} \right\|_1 = O\Big(\frac{\delta^{1/8}}{\alpha^2}\Big), \tag{4.7}$$

where the second expectation is over (x, y) drawn from μ, and $(U \otimes V)[\widetilde{\Phi}]$ denotes $(U \otimes V)\widetilde{\Phi}(U \otimes V)^\dagger$. Conditioned on a given pair of questions (x, y) and the players sampling (i, ω_{-i}, z) in Step 1., the state that the players prepare after Step 3. in the protocol is precisely $(U_{\omega_{-i},z,x} \otimes V_{\omega_{-i},z,y})[\widetilde{\Phi}_{\substack{\omega_{-i},z \\ \perp,\perp}}]$. Let $\mathcal{E}_{\substack{\omega_{-i},z \\ x,y}}$ denote the quantum-classical channel on density matrices that performs the measurement $\{\hat{A}^{a_i}_{\omega_{-i},x} \otimes \hat{B}^{b_i}_{\omega_{-i},y}\}_{a_i,b_i}$, and outputs a classical register with the measurement outcome (a_i, b_i). Applying $\mathcal{E}_{\substack{\omega_{-i},z\\x,y}}$ to the expression inside the trace norm in (4.7), using that the trace norm is non-increasing under quantum operations,

$$\mathop{\mathbb{E}}_{i}\ \mathop{\mathbb{E}}_{XY}\ \mathop{\mathbb{E}}_{\Omega_{-i}Z|W_C}\left\|\widetilde{\mathsf{P}}_{A_iB_i|\omega_{-i},v,x,y}-\mathsf{P}_{A_iB_i|\omega_{-i},v,x,y}\right\|\leq O(\delta^{1/8}/\alpha^2).$$

where $\widetilde{\mathsf{P}}_{A_iB_i|\omega_i,z,x,y}(a_i,b_i)$ deontes the probability of outcome (a_i, b_i) in the above strategy, conditioned on questions (x, y) and the players sampling (i, ω_{-i}, z) in Step 1. Thus

$$\begin{aligned}\mathsf{P}_I\cdot\mathsf{P}_{\Omega_{-i}\mathrm{Z}|W_C}\cdot\mathsf{P}_{XY}\cdot\widetilde{\mathsf{P}}_{A_iB_i|\Omega_{-i}\mathrm{Z}X_iY_i}&\approx_{O(\delta^{1/8}/\alpha^2)}\mathsf{P}_I\cdot\mathsf{P}_{\Omega_{-i}\mathrm{Z}|W_C}\cdot\mathsf{P}_{XY}\cdot\mathsf{P}_{A_iB_i|\Omega_{-i}\mathrm{Z}X_iY_i}\\&\approx_{O(\delta^{1/8}/\alpha^2)}\mathsf{P}_I\cdot\mathsf{P}_{\Omega_{-i}\mathrm{Z}X_iY_i|W_C}\cdot\mathsf{P}_{A_iB_i|\Omega_{-i}\mathrm{Z}X_iY_i}\end{aligned}$$

where the X_iY_i in the conditionals is shorthand for $X_i = x, Y_i = y$. The last approximate equality follows from Lemma 4.8. Marginalizing $\Omega_{-i}\mathrm{Z}$, we get

$$\mathsf{P}_I\cdot\mathsf{P}_{XY}\cdot\widetilde{\mathsf{P}}_{A_iB_i|X_iY_i}\approx_{O(\delta^{1/8}/\alpha^2)}\mathsf{P}_I\cdot\mathsf{P}_{X_iY_iA_iB_i|W_C}.\tag{4.8}$$

Under the distribution $\mathsf{P}_{X_iY_iA_iB_i|W_C}$, the probability that $V(x_i, y_i, a_i, b_i) = 1$ is precisely $\Pr(W_i|W_C)$. On the other hand, (4.8) implies that using the protocol described above the players win G with probability at least $\mathbb{E}_i \Pr(W_i|W_C) - O(\delta^{1/8}/\alpha^2)$. This concludes the proof of the lemma. □

Given Lemma 4.9, the proof of Theorem 4.7 (at least the two player case) follows using a standard inductive argument (see, e.g., the argument for Theorem 4.17 given in Section 4.5). Later, in Section 4.4.3, we sketch the changes necessary to adapt the proof to handle an arbitrary number of players.

Proof of the main lemma

This section is devoted to the proof of Lemma 4.10. The proof is based on two lemmas. The first defines the required unitaries.

Lemma 4.11. *For all i, ω_{-i}, z, $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ there exists unitaries $U_{\omega_{-i}zx}$ acting on E_A*

and $V_{\omega_{-i}zy}$, $V_{\substack{\omega_{-i},z\\ x,y}}$ acting on E_B such that

$$\frac{1}{m}\sum_i \mathop{\mathbb{E}}_{\Omega_{-i}\mathsf{Z}|W} \mathop{\mathbb{E}}_{X} \left\| \left|\widetilde{\Phi}_{x,\perp}\right\rangle - U_{\omega_{-i}zx}\left|\widetilde{\Phi}_{\perp,\perp}\right\rangle \right\|^2 = O(\delta^{1/4}/\alpha^2), \tag{4.9}$$

$$\frac{1}{m}\sum_i \mathop{\mathbb{E}}_{\Omega_{-i}\mathsf{Z}|W} \mathop{\mathbb{E}}_{Y} \left\| V_{\omega_{-i}zy}\left|\widetilde{\Phi}_{\perp,\perp}\right\rangle - \left|\widetilde{\Phi}_{\perp,y}\right\rangle \right\|^2 = O(\delta^{1/4}/\alpha^2), \tag{4.10}$$

$$\frac{1}{m}\sum_i \mathop{\mathbb{E}}_{\Omega_{-i}\mathsf{Z}|W} \mathop{\mathbb{E}}_{XY} \left\| V_{\substack{\omega_{-i},z\\ x,y}}\left|\widetilde{\Phi}_{\perp/x,y}\right\rangle - \left|\widetilde{\Phi}_{\perp/x,\perp}\right\rangle \right\|^2 = O(\delta^{1/4}/\alpha^4). \tag{4.11}$$

where $\mathbb{E}_X$, $\mathbb{E}_Y$, and $\mathbb{E}_{XY}$ denote averaging over $\mu(x)$, $\mu(y)$, and $\mu(x,y)$ respectively.

The proof of Lemma 4.11 is given in Section 4.4.2. The second lemma relates the normalization factors $\gamma_{x,y}$, $\gamma_{x,\perp}$, $\gamma_{\perp,y}$, $\gamma_{\perp/x,y}$, $\gamma_{\perp/x,\perp}$, $\gamma_{\perp,\perp}$ that appear in the definition of the corresponding normalized states $|\widetilde{\Phi}\rangle$.

Lemma 4.12. *There exists a set S of triples (i,ω_{-i},z) that has probability $1-\delta^{1/4}$ under $\mathsf{P}_I \cdot \mathsf{P}_{\Omega_{-i}\mathsf{Z}|W}$ such that*

$$\frac{1}{m}\sum_{\substack{x,y\\ (i,\omega_{-i},z)\in S}} \mathsf{P}_{XY}(x,y)\cdot \mathsf{P}_{\Omega_{-i}\mathsf{Z}|W}(\omega_{-i},v)\left|\gamma^2_{\substack{\omega_{-i},z\\ x,y}} - \gamma^2_{\substack{\omega_{-i},z\\ \perp,\perp}}\right| = O\Big(\delta^{1/4}/\alpha^2\Big)\gamma^2, \tag{4.12}$$

where

$$\gamma = \left(\frac{1}{m}\sum_i \sum_{x,y,\omega_{-i},z} \mathsf{P}_{XY}(x,y)\cdot \mathsf{P}_{\Omega_{-i}\mathsf{Z}|W}(\omega_{-i},z)\cdot \gamma^2_{\substack{\omega_{-i},z\\ x,y}}\right)^{1/2}.$$

Furthermore, similar bounds as (4.12) hold where $\gamma_{\substack{\omega_{-i},z\\ x,y}}$ is replaced by any of $\gamma_{\substack{\omega_{-i},z\\ x,\perp}}$, $\gamma_{\substack{\omega_{-i},z\\ \perp,y}}$, $\gamma_{\substack{\omega_{-i},z\\ \perp/x,y}}$, $\gamma_{\substack{\omega_{-i},z\\ \perp/x,\perp}}$.

The proof of Lemma 4.12 uses the following claim.

Claim 4.13.

$$\frac{1}{m}\sum_i \sum_{x,y,(\omega_{-i},z)\in W} \mathsf{P}_{XY}(x,y)\left\| \mathsf{P}_{\Omega_{-i}\mathsf{Z}|X_i=x,Y_i=y}(\omega_{-i},z) - \mathsf{P}_{\Omega_{-i}\mathsf{Z}|X_i\in\perp,Y_i\in\perp}(\omega_{-i},z)\right\| = O\Big(\frac{\sqrt{\delta}}{\alpha^2}\Big)\Pr(W).$$

Proof. First note that

$$\frac{1}{m}\sum_i\sum_{x,y} \mathsf{P}_{XY}(x,y)\left|\Pr(W|X_i=x,Y_i=y)-\Pr(W)\right| = \frac{\Pr(W)}{m}\sum_i \left\|\mathsf{P}_{X_iY_i|W} - \mathsf{P}_{X_iY_i}\right\|$$

$$= O(\sqrt{\delta}) \Pr(W), \tag{4.13}$$

where the second equality follows from Lemma 4.8. Using the triangle inequality and $\Pr(X_i \in \perp, Y_i \in \perp) \geq \alpha^2$ we also get

$$\frac{1}{m}\sum_i \sum_{x,y} \mathsf{P}_{XY}(x,y) \left|\Pr(W|X_i = x, Y_i = y) - \Pr(W|X_i \in \perp, Y_i \in \perp)\right| = O(\sqrt{\delta}/\alpha^2) \Pr(W). \tag{4.14}$$

Using (4.13) and letting $\mathsf{P}_{\Omega_{-i}\mathsf{Z}|x,y,W}$ denote $\mathsf{P}_{\Omega_{-i}\mathsf{Z}|X_i=x,Y_i=y,W}$,

$$\begin{aligned}
&\frac{1}{m}\sum_i \sum_{x,y} \mathsf{P}_{XY}(x,y) \sum_{(\omega_{-i},z)\in W} \left\|\Pr(W) \cdot \mathsf{P}_{\Omega_{-i}\mathsf{Z}|x,y,W}(\omega_{-i},z) - \mathsf{P}_{\Omega_{-i}\mathsf{Z}|x,y}(\omega_{-i},z)\right\| \\
&\approx_{O(\sqrt{\delta})\Pr(W)} \frac{1}{m}\sum_i \sum_{x,y} \mathsf{P}_{XY}(x,y) \sum_{(\omega_{-i},z)\in W} \left\|\mathsf{P}_{\Omega_{-i}\mathsf{Z}\wedge W|x,y}(\omega_{-i},z) - \mathsf{P}_{\Omega_{-i}\mathsf{Z}|x,y}(\omega_{-i},z)\right\| \\
&= 0.
\end{aligned}$$

A similar derivation proves

$$\frac{1}{m}\sum_i \sum_{(\omega_{-i},z)\in W} \left\|\Pr(W) \cdot \mathsf{P}_{\Omega_{-i}\mathsf{Z}|X_i\in\perp,Y_i\in\perp,W}(\omega_{-i},z) - \mathsf{P}_{\Omega_{-i}\mathsf{Z}|X_i\in\perp,Y_i\in\perp}(\omega_{-i},z)\right\| = O(\sqrt{\delta}) \Pr(W).$$

Combining the previous two bounds with the bound

$$\frac{1}{m}\sum_i \Pr(W) \|\mathsf{P}_{X_iY_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|X_i\in\perp,Y_i\in\perp,W} - \mathsf{P}_{X_iY_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|X_iY_iW}\| \leq O(\sqrt{\delta}/\alpha^2) \Pr(W)$$

from Lemma 4.8 with the triangle inequality proves the claim. □

Proof of Lemma 4.12. For any i, x, y and $(\omega_{-i}, z) \in W$ write

$$\begin{aligned}
\mathsf{P}_{XY}(x,y) \cdot \mathsf{P}_{\Omega_{-i}\mathsf{Z}|W}(\omega_{-i},z) \cdot \gamma^2_{\substack{\omega_{-i},z\\x,y}} &= \frac{1}{\Pr(W)} \mathsf{P}_{XY}(x,y) \cdot \mathsf{P}_{\Omega_{-i}\mathsf{Z}}(\omega_{-i},z) \cdot \gamma^2_{\substack{\omega_{-i},z\\x,y}} \\
&= \frac{1}{\Pr(W)} \mathsf{P}_{XY}(x,y) \cdot \mathsf{P}_{\Omega_{-i}|x,y}(\omega_{-i}) \cdot \mathsf{P}_{\mathsf{Z}|\omega_{-i}}(z) \cdot \gamma^2_{\substack{\omega_{-i},z\\x,y}},
\end{aligned}$$

where for the last equality we used $\mathsf{P}_{\Omega_{-i}|X_iY_i} = \mathsf{P}_{\Omega_{-i}}$. From the definition, $\gamma^2_{\substack{\omega_{-i},z\\x,y}} = \mathsf{P}_{\mathsf{Z}|\omega_{-i},x,y}(z)$,

$$= \frac{1}{\Pr(W)}\mathsf{P}_{XY}(x,y)\cdot\mathsf{P}_{\mathsf{Z}|\omega_{-i}}(z)\cdot\mathsf{P}_{\Omega_{-i}\mathsf{Z}|x,y}(\omega_{-i},z), \tag{4.15}$$

where $\mathsf{P}_{\Omega_{-i}\mathsf{Z}|x,y}(\omega_{-i},z)$ denotes $\mathsf{P}_{\Omega_{-i}\mathsf{Z}|X_i=x,Y_i=y}(\omega_{-i},z)$. Similarly, we have

$$\mathsf{P}_{XY}(x,y)\cdot\mathsf{P}_{\Omega_{-i}\mathsf{Z}|W}(\omega_{-i},z)\cdot\gamma^2_{\substack{\omega_{-i},z\\\perp,\perp}} = \frac{1}{\Pr(W)}\mathsf{P}_{XY}(x,y)\cdot\mathsf{P}_{\mathsf{Z}|\omega_{-i}}(z)\cdot\mathsf{P}_{\Omega_{-i}\mathsf{Z}|\perp,\perp}(\omega_{-i},z). \tag{4.16}$$

By definition

$$\gamma^2 = \frac{1}{m}\sum_{i,\omega_{-i},z}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|W}(\omega_{-i},z)\cdot\mathsf{P}_{V|\omega_{-i}}(z),$$

thus for any $\eta > 0$ applying Markov's inequality a fraction at least $1-\eta$ of (i,ω_{-i},z) distributed according to $\mathsf{P}_I\cdot\mathsf{P}_{\Omega_{-i}\mathsf{Z}|W}$ are such that $\mathsf{P}_{\mathsf{Z}|\omega_{-i}}(z) \leq \gamma^2/\eta$. Let S be the set of such triples, and consider summing the difference

$$\mathsf{P}_{XY}(x,y)\cdot\mathsf{P}_{\mathsf{Z}|\omega_{-i}}(z)\cdot\Big|\mathsf{P}_{\Omega_{-i}\mathsf{Z}|x,y}(\omega_{-i},z) - \mathsf{P}_{\Omega_{-i}\mathsf{Z}|\perp,\perp}(\omega_{-i},z)\Big|$$

over all (x,y) and $(i,\omega_{-i},z)\in S$. By lines (4.15) and (4.16), and applying Claim 4.13 we obtain

$$\frac{1}{m}\sum_{\substack{x,y\\(i,\omega_{-i},z)\in S}}\mathsf{P}_{XY}(x,y)\cdot\mathsf{P}_{\Omega_{-i}Z|W}(\omega_{-i},z)\cdot\Big|\gamma^2_{\substack{\omega_{-i},z\\x,y}} - \gamma^2_{\substack{\omega_{-i},z\\\perp,\perp}}\Big| \leq \frac{\gamma^2}{\eta}O\Big(\frac{\sqrt{\delta}}{\alpha^2}\Big).$$

Choosing $\eta = \delta^{1/4}$ proves the lemma. □

Proof of Lemma 4.10. For every (i,ω_{-i},z), $x\in\mathcal{X}$ and $y\in\mathcal{Y}$ let unitaries $U_{\omega_{-i}zx}$, $V_{\omega_{-i}zy}$ and $V_{\substack{\omega_{-i},z\\x,y}}$ be as in Lemma 4.11. For notational convenience we suppress the dependence on (i,ω_{-i},z) when it is clear from context. Call triples (i,ω_{-i},z) that satisfy the conclusion of Lemma 4.12 for $\gamma_{\substack{\omega_{-i},z\\x,y}}$, $\gamma_{\substack{\omega_{-i},z\\x,\perp}}$, $\gamma_{\substack{\omega_{-i},z\\\perp,y}}$, $\gamma_{\substack{\omega_{-i},z\\\perp/x,y}}$, and $\gamma_{\substack{\omega_{-i},z\\\perp/x,\perp}}$ simutaneously *good triples*, and let S denote the set of good triples. Fix $(i,\omega_{-i},z)\in S$. Using $|a-b|^2 \leq |a^2-b^2|$ for $a,b\geq 0$,

$$\sum_{x,y}\mathsf{P}_{XY}(x,y)\cdot\Big\||\widetilde{\Phi}_{x,y}\rangle - \gamma^{-1}|\Phi_{x,y}\rangle\Big\|^2 = \sum_{x,y}\mathsf{P}_{XY}(x,y)\cdot\left|\frac{\gamma-\gamma_{\substack{\omega_{-i},z\\x,y}}}{\gamma}\right|^2$$

$$
\begin{aligned}
&\leq \sum_{x,y} \mathsf{P}_{XY}(x,y) \cdot \left| \frac{\gamma^2 - \gamma^2_{\substack{\omega_{-i},z \\ x,y}}}{\gamma^2} \right| \\
&= O(\delta^{1/4}/\alpha^2), \qquad (4.17)
\end{aligned}
$$

and similar bounds hold for $|\widetilde{\Phi}_{x,\perp}\rangle$, $|\widetilde{\Phi}_{\perp,y}\rangle$ and $|\widetilde{\Phi}_{\perp,\perp}\rangle$. Thus to prove the theorem it will be sufficient to establish that

$$
\frac{1}{m} \sum_{\substack{x,y \\ (i,\omega_{-i},z)\in S}} \mathsf{P}_{XY}(x,y) \cdot \mathsf{P}_{\Omega_{-i},\mathrm{Z}|W}(\omega_{-i},z) \cdot \left\| (U_x \otimes V_y) \left|\Phi_{\perp,\perp}\right\rangle - \left|\Phi_{x,y}\right\rangle \right\|^2 = O\Big(\frac{\delta^{1/4}}{\alpha^2}\Big)\gamma^2. \quad (4.18)
$$

Using the lower bound on the measure of S,

$$
\begin{aligned}
&\frac{1}{m} \sum_{\substack{x,y \\ i,\omega_{-i},v}} \mathsf{P}_{XY}(x,y) \cdot \mathsf{P}_{\Omega_{-i},V|W}(\omega_{-i},v) \cdot \left\| (U_x \otimes V_y) \left|\widetilde{\Phi}_{\perp,\perp}\right\rangle - \left|\widetilde{\Phi}_{x,y}\right\rangle \right\|^2 \\
&\leq \frac{1}{m} \sum_{\substack{x,y \\ (i,\omega_{-i},v)\in S}} \mathsf{P}_{XY}(x,y) \cdot \mathsf{P}_{\Omega_{-i},V|W}(\omega_{-i},v) \cdot \left\| (U_x \otimes V_y) \left|\widetilde{\Phi}_{\perp,\perp}\right\rangle - \left|\widetilde{\Phi}_{x,y}\right\rangle \right\|^2 + O(\delta^{1/4})
\end{aligned}
$$

For each good triple (i,ω_{-i},z), by the triangle inequality

$$
\begin{aligned}
\left\| (U_x \otimes V_y) \left|\widetilde{\Phi}_{\perp,\perp}\right\rangle - \left|\widetilde{\Phi}_{x,y}\right\rangle \right\|^2 &\leq 3 \left\| \left|\widetilde{\Phi}_{\perp,\perp}\right\rangle - \gamma^{-1} \left|\Phi_{\perp,\perp}\right\rangle \right\|^2 + 3 \left\| \left|\widetilde{\Phi}_{x,y}\right\rangle - \gamma^{-1} \left|\Phi_{x,y}\right\rangle \right\|^2 \\
&\qquad + 3\gamma^{-2} \left\| (U_x \otimes V_y) \left|\Phi_{\perp,\perp}\right\rangle - \left|\Phi_{x,y}\right\rangle \right\|^2 \\
&\leq 3\gamma^{-2} \left\| (U_x \otimes V_y) \left|\Phi_{\perp,\perp}\right\rangle - \left|\Phi_{x,y}\right\rangle \right\|^2 + O(\delta^{1/4}/\alpha^2).
\end{aligned}
$$

Using (4.17), the bounds stated in Lemma 4.11 imply the following bounds on the unnormalized vectors:

$$
\begin{aligned}
\frac{1}{m} \sum_{\substack{x \\ (i,\omega_{-i},z)\in S}} \mathsf{P}_{X}(x) \cdot \mathsf{P}_{\Omega_{-i}\mathrm{Z}|W}(\omega_{-i},z) \cdot \left\| \left|\Phi_{x,\perp}\right\rangle - U_{\omega_{-i}zx} \left|\Phi_{\perp,\perp}\right\rangle \right\|^2 &= O\Big(\frac{\delta^{1/4}}{\alpha^2}\Big)\gamma^2, \quad (4.19) \\
\frac{1}{m} \sum_{\substack{y \\ (i,\omega_{-i},z)\in S}} \mathsf{P}_{Y}(y) \cdot \mathsf{P}_{\Omega_{-i}\mathrm{Z}|W}(\omega_{-i},z) \cdot \left\| V_{\omega_{-i}zy} \left|\Phi_{\perp,\perp}\right\rangle - \left|\Phi_{\perp,y}\right\rangle \right\|^2 &= O\Big(\frac{\delta^{1/4}}{\alpha^2}\Big)\gamma^2, \quad (4.20) \\
\frac{1}{m} \sum_{\substack{x,y \\ (i,\omega_{-i},z)\in S}} \mathsf{P}_{XY}(x,y) \cdot \mathsf{P}_{\Omega_{-i}\mathrm{Z}|W}(\omega_{-i},z) \cdot \left\| V_{\substack{\omega_{-i},z \\ x,y}} \left|\Phi_{\perp/x,y}\right\rangle - \left|\Phi_{\perp/x,\perp}\right\rangle \right\|^2 &= O\Big(\frac{\delta^{1/4}}{\alpha^4}\Big)\gamma^2. \quad (4.21)
\end{aligned}
$$

We show how to combine these bounds to establish (4.18). We have

$$\begin{aligned}\left\|U_x|\Phi_{\perp,y}\rangle - |\Phi_{x,y}\rangle\right\|^2 &= \left\|U_x A_{\perp}^{1/2} A_{\perp/x}^{-1/2}|\Phi_{\perp/x,y}\rangle - A_x^{1/2} A_{\perp/x}^{-1/2}|\Phi_{\perp/x,y}\rangle\right\|^2 \\ &= \left\|U_x A_{\perp}^{1/2} A_{\perp/x}^{-1/2} \otimes V_{xy}|\Phi_{\perp/x,y}\rangle - A_x^{1/2} A_{\perp/x}^{-1/2} \otimes V_{xy}|\Phi_{\perp/x,y}\rangle\right\|^2 .\end{aligned}$$

Using the triangle inequality again,

$$\leq 3\left\|\left(U_x A_{\perp}^{1/2} A_{\perp/x}^{-1/2}\right) \otimes V_{xy}|\Phi_{\perp/x,y}\rangle - \left(U_x A_{\perp}^{1/2} A_{\perp/x}^{-1/2}\right)|\Phi_{\perp/x,\perp}\rangle\right\|^2 \tag{4.22}$$

$$+ 3\left\|\left(U_x A_{\perp}^{1/2} A_{\perp/x}^{-1/2}\right)|\Phi_{\perp/x,\perp}\rangle - A_x^{1/2} A_{\perp/x}^{-1/2}|\Phi_{\perp/x,\perp}\rangle\right\|^2 \tag{4.23}$$

$$+ 3\left\|A_x^{1/2} A_{\perp/x}^{-1/2}|\Phi_{\perp/x,\perp}\rangle - A_x^{1/2} A_{\perp/x}^{-1/2} \otimes V_{xy}|\Phi_{\perp/x,y}\rangle\right\|^2 . \tag{4.24}$$

Using $\|U_x A_{\perp}^{1/2} A_{\perp/x}^{-1/2}\| \leq \sqrt{2}$ the term (4.22) can be bounded as

$$\left\|\left(U_x A_{\perp}^{1/2} A_{\perp/x}^{-1/2}\right) \otimes V_{xy}\,|\Phi_{\perp/x,y}\rangle - \left(U_x A_{\perp}^{1/2} A_{\perp/x}^{-1/2}\right)|\Phi_{\perp/x,\perp}\rangle\right\|^2 \leq 2\left\|V_{xy}|\Phi_{\perp/x,y}\rangle - |\Phi_{\perp/x,\perp}\rangle\right\|^2 .$$

The term (4.23) can be re-written as

$$\left\|\left(U_x A_{\perp}^{1/2} A_{\perp/x}^{-1/2}\right)|\Phi_{\perp/x,\perp}\rangle - A_x^{1/2} A_{\perp/x}^{-1/2}|\Phi_{\perp/x,\perp}\rangle\right\|^2 = \left\|U_x|\Phi_{\perp,\perp}\rangle - |\Phi_{x,\perp}\rangle\right\|^2 .$$

Finally, using $\|A_x^{1/2} A_{\perp/x}^{-1/2}\| \leq \sqrt{2}$ the term (4.24) can be bounded as

$$\left\|A_x^{1/2} A_{\perp/x}^{-1/2}|\Phi_{\perp/x,\perp}\rangle - A_x^{1/2} A_{\perp/x}^{-1/2} \otimes V_{xy}\,|\Phi_{\perp/x,y}\rangle\right\|^2 \leq 2\left\||\Phi_{\perp/x,\perp}\rangle - V_{xy}\,|\Phi_{\perp/x,y}\rangle\right\|^2 .$$

Putting the three bounds together, from (4.24) we get

$$\left\|U_x|\Phi_{\perp,y}\rangle - |\Phi_{x,y}\rangle\right\|^2 \leq 3\left\|V_{xy}|\Phi\rangle_{\perp/x,y} - |\Phi_{\perp/x,\perp}\rangle\right\|^2 + 3\left\|U_x|\Phi_{\perp,\perp}\rangle - |\Phi_{x,\perp}\rangle\right\|^2 . \tag{4.25}$$

Using that U_x is unitary,

$$\begin{aligned}\left\|(U_x \otimes V_y)|\Phi_{\perp,\perp}\rangle - |\Phi_{x,y}\rangle\right\|^2 &\leq 2\left\|V_y|\Phi_{\perp,\perp}\rangle - |\Phi_{\perp,y}\rangle\right\|^2 + 2\left\|U_x|\Phi_{\perp,y}\rangle - |\Phi_{x,y}\rangle\right\|^2 \\ &\leq 18\left\|V_{xy}|\Phi_{\perp/x,y}\rangle - |\Phi_{\perp/x,\perp}\rangle\right\|^2 + 6\left\|U_x|\Phi_{\perp,\perp}\rangle - |\Phi_{x,\perp}\rangle\right\|^2\end{aligned}$$

$$+ 2 \left\| V_y |\Phi_{\perp,\perp}\rangle - |\Phi_{\perp,y}\rangle \right\|^2 ,$$

where the last inequality is (4.25). Eqs. (4.19), (4.20) and (4.21) bound the three terms above by $O(\delta^{1/4}/\alpha^4)\gamma^2$ on average over (x, y) weighted by P_{XY}, and $(i, \omega_{-i}, z) \in S$, weighted by $\mathsf{P}_I \cdot \mathsf{P}_{\Omega_{-i}\mathrm{Z}|W}$. This proves (4.18), and the theorem follows. □

Obtaining local unitaries

In this section we give the proof of Lemma 4.11, which states the existence of the local unitary transformations needed for the proof of Theorem 4.7.

Proof of Lemma 4.11. Recall that we let the entangled state $|\psi\rangle$ and POVMs $\{A^{a^n}_{x^n}\}$ and $\{B^{b^n}_{x^n}\}$ constitute an optimal strategy for $G^{\otimes n}$. We refer the reader to Section 4.4.1 for the definitions of operators $A^{a_C}_{\omega}$, etc. We will let ρ denote the reduced density matrix of $|\psi\rangle$ on either system (this is well-defined because we've assumed $|\psi\rangle$ is symmetric).

We first prove (4.10), that is, the existence of the unitary $V_{\omega_{-i}zy_i}$. Recall the notation $\psi = |\psi\rangle\langle\psi|$ and $X[\rho] = X\rho X^\dagger$. Introduce the following states:

$$\Xi_{\Omega Y^n E_A E_B \mathrm{Z}} = \sum_{\omega, y^n, a_C, b_C} \mathsf{P}_{\Omega Y^n}(\omega, y^n)\, |\omega\, y^n\rangle\langle\omega\, y^n| \otimes \left(\sqrt{A^{a_C}_{\omega}} \otimes \sqrt{B^{b_C}_{y^n}}\right)[\psi] \otimes |a_C b_C\rangle\langle a_C b_C|,$$

$$\xi_{\Omega Y^n E_A E_B \mathrm{Z}} = \Xi_{\Omega Y^n E_A E_B \mathrm{Z}|W}, \tag{4.26}$$

$$\xi^{E_A}_{\substack{\omega_{-i},z \\ \perp, y_i}} = \xi_{E_A|\Omega_{-i}=\omega_{-i}, Y_i=y_i, \omega_i=(A,\perp)}. \tag{4.27}$$

The state Ξ is defined so that tracing out the entanglement registers E_A and E_B the resulting state $\Xi_{\Omega Y^n A_C B_C}$ is a classical state that is equivalent to the probability distribution $\mathsf{P}_{\Omega Y^n A_C B_C}$. In (4.26) the conditioning on W is well-defined since the event only involves classical random variables in Ω and Z. In (4.27) only the reduced density on E_A is considered, all other registers being traced out.

The following claim provides the main step of the proof by relating the reduced densities on Alice's registers of states (4.27) associated with different choices for y_i.

Claim 4.14.

$$\frac{1}{m}\sum_i \mathop{\mathbb{E}}_{\Omega_{-i}\mathrm{Z}|W} \; \mathop{\mathbb{E}}_{Y_i} \left\| \xi^{E_A}_{\substack{\omega_{-i},z \\ \perp,y_i}} - \xi^{E_A}_{\substack{\omega_{-i},z \\ \perp,\perp}} \right\|_1^2 = O\Big(\sqrt{\delta}/\alpha^2\Big) \tag{4.28}$$

Proof. First we observe that $\Pr(W)\xi \preceq \Xi$, thus by definition $S(\xi\|\Xi) \leq S_\infty(\xi\|\Xi) \leq \log 1/\Pr(W)$. Using the chain rule for the relative entropy (Lemma 2.8),

$$\mathop{\mathbb{E}}_{\Omega V|W} S(\xi^{Y^n E_A}_{\omega,z}\|\Xi^{Y^n E_A}_{\omega,z}) \leq \log \frac{1}{\Pr(W)}. \tag{4.29}$$

Next we note that for any ω, using Ando's identity

$$\langle\psi|X \otimes Y|\psi\rangle = \mathrm{Tr}(X\sqrt{\rho}Y^\top\sqrt{\rho}),$$

where $|\psi\rangle = \sum\sqrt{\lambda_j}|v_j\rangle|v_j\rangle$, $\rho = \sum \lambda_j|v_j\rangle\langle v_j|$, X, Y are any linear operators and the transpose is taken with respect to the orthonormal basis $\{|v_j\rangle\}$,

$$\begin{aligned}
\Xi^{Y^n E_A A_C B_C}_\omega &= \sum_{y^n,a_C,b_C} \mathsf{P}_{Y^n|\omega}(y^n)\; |y^n\rangle\!\langle y^n| \otimes \sqrt{A^{a_C}_\omega}\sqrt{\rho}\overline{B}^{b_C}_{y^n}\sqrt{\rho}\sqrt{A^{a_C}_\omega} \otimes |a_C b_C\rangle\!\langle a_C b_C| \\
&\preceq \sum_{y^n,a_C,b_C} \mathsf{P}_{Y^n|\omega}(y^n)\; |y^n\rangle\!\langle y^n| \otimes \sqrt{A^{a_C}_\omega}\sqrt{\rho}\overline{B}^{b_C}_{x^n}\sqrt{\rho}\sqrt{A^{a_C}_\omega} \otimes \mathbb{I} \\
&= \sum_{y^n,a_C} \mathsf{P}_{Y^n|\omega}(y^n)\; |y^n\rangle\!\langle y^n| \otimes \sqrt{A^{a_C}_\omega}\rho\sqrt{A^{a_C}_\omega} \otimes \mathbb{I},
\end{aligned} \tag{4.30}$$

where the last equality uses $\sum_{b_C} B^{b_C}_{y^n} = \mathbb{I}$. From (4.30) and the definition of S_∞ it follows that $S_\infty(\Xi^{Y^n E_A}_\omega\|\Xi^{Y^n}_\omega \otimes \Xi^{E_A}_\omega) \leq |C| \cdot \log|\mathcal{A}||\mathcal{B}|$. Applying Lemma 2.9,

$$\begin{aligned}
\frac{1}{m}\sum_i \mathop{\mathbb{E}}_{\Omega\mathrm{Z}|W} I(Y_i;E_A|\omega,z)_\xi &\leq \frac{1}{m}\mathop{\mathbb{E}}_{\Omega\mathrm{Z}|W} S(\xi^{Y^n E_A}_{\omega,z}\|\Xi^{Y^n}_{\omega,z} \otimes \Xi^{E_A}_{\omega,z}) \\
&\leq \frac{1}{m}\Big(\mathop{\mathbb{E}}_{\Omega\mathrm{Z}|W} S(\xi^{Y^n E_A}_{\omega,z}\|\Xi^{Y^n E_A}_{\omega,z}) + \mathop{\mathbb{E}}_{\Omega\mathrm{Z}|W} S_\infty(\Xi^{Y^n E_A}_{\omega,z}\|\Xi^{Y^n}_{\omega,z} \otimes \Xi^{E_A}_{\omega,z})\Big) \\
&\leq \frac{1}{m}\Big(\log\frac{1}{\Pr(W)} + |C| \cdot \log|\mathcal{A}||\mathcal{B}|\Big) = \delta
\end{aligned} \tag{4.31}$$

where in the last line the first term is bounded using (4.29) and the second using (4.30).

Applying Lemma 4.8,

$$\mathop{\mathbb{E}}_{i} \mathsf{P}_{D_iM_i|W}(A,\perp) \approx_{O(\sqrt{\delta})} \mathop{\mathbb{E}}_{i} \mathsf{P}_{D_iM_i}(A,\perp) = \frac{\alpha}{2},$$

thus from (4.31) by conditioning on $\Omega_i = (A,\perp)$ we deduce

$$\frac{1}{m}\sum_i \mathop{\mathbb{E}}_{\Omega\mathsf{Z}|\Omega_i=(A,\perp),W} I\Big(Y_i; E_A|\omega, z\Big)_\xi = O\Big(\delta/\alpha\Big), \tag{4.32}$$

as long as $\alpha = \Omega(\sqrt{\delta})$. Next we apply Pinsker's inequality (Lemma 2.7) and use that Y_i is classical in ξ to write

$$\begin{aligned}\frac{1}{m}\sum_i \mathop{\mathbb{E}}_{\Omega\mathsf{Z}|\Omega_i=(A,\perp),W} \ \mathop{\mathbb{E}}_{Y_i|\omega,z} \left\| \xi^{E_A}_{\substack{\omega_{-i},z\\ \perp,y_i}} - \xi^{E_A}_{\omega,z} \right\|_1^2 &\leq \frac{1}{m}\sum_i \mathop{\mathbb{E}}_{\Omega\mathsf{Z}|\Omega_i=(A,\perp),W} \mathop{\mathbb{E}}_{Y_i|\omega,z} S\Big(\xi^{E_A}_{\substack{\omega_{-i},z\\ \perp,y_i}} \,\Big\|\, \xi^{E_B}_{\omega,z}\Big)\\ &= \frac{1}{m}\sum_i \mathop{\mathbb{E}}_{\Omega\mathsf{Z}|\Omega_i=(A,\perp),W} I(Y_i;E_A|\omega,z)_\xi \\ &= O\Big(\delta/\alpha\Big)\end{aligned}$$

by (4.32). To conclude note that Lemma 4.8 and the classical correlated sampling lemma imply

$$\mathsf{P}_I \cdot \mathsf{P}_{\Omega\mathsf{Z}Y_i|\Omega_i=(A,\perp),W} \approx_{O(\sqrt{\delta}/\alpha^2)} \mathsf{P}_I \cdot \mathsf{P}_{\Omega_{-i}\mathsf{Z}|W} \cdot \mathsf{P}_{Y_i}.$$

□

The proof of (4.9) essentially follows from Claim 4.14 and Uhlmann's theorem. We give the details. First write $\xi^{E_B}_{\substack{\omega_{-i},z\\ \perp,y_i}}$ and $\xi^{E_A}_{\substack{\omega_{-i},z\\ \perp,\perp}}$ explicitly as

$$\begin{aligned}\xi^{E_A}_{\substack{\omega_{-i},z\\ \perp,y_i}} &\propto (A^{a_C}_{\omega_{-i},\perp})^{1/2}\sqrt{\rho}\ \overline{B}^{b_C}_{\omega_{-i},y_i}\ \sqrt{\rho}\ (A^{a_C}_{\omega_{-i},\perp})^{1/2},\\ \xi^{E_A}_{\substack{\omega_{-i},z\\ \perp,\perp}} &\propto (A^{a_C}_{\omega_{-i},\perp})^{1/2}\sqrt{\rho}\ \overline{B}^{b_C}_{\omega_{-i},\perp}\ \sqrt{\rho}\ (A^{a_C}_{\omega_{-i},\perp})^{1/2},\end{aligned}$$

which makes it apparent that the states $\Big|\tilde{\Phi}_{\substack{\omega_{-i},z\\ \perp,y_i}}\Big\rangle$ and $\Big|\tilde{\Phi}_{\substack{\omega_{-i},z\\ \perp,\perp}}\Big\rangle$ introduced in (4.6) purify $\xi^{E_A}_{\substack{\omega_{-i},z\\ \perp,y_i}}$ and $\xi^{E_A}_{\substack{\omega_{-i},z\\ \perp,\perp}}$ respectively. Applying Uhlmann's Theorem, there exists a unitary V_{ω_{-i},z,y_i} acting on E_B such that

$$\begin{aligned}\frac{1}{m}\sum_i \mathop{\mathbb{E}}_{\Omega_{-i}Z|W}\ \mathop{\mathbb{E}}_{Y_i}\left|\left\langle \widetilde{\Phi}_{\substack{\omega_{-i},z\\ \perp,y_i}}\right| V_{\omega_{-i},z,y_i}\left|\widetilde{\Phi}_{\substack{\omega_{-i},z\\ \perp,\perp}}\right\rangle\right| &\geq 1-\frac{1}{m}\sum_i \mathop{\mathbb{E}}_{\Omega_{-i}Z|W}\ \mathop{\mathbb{E}}_{Y_i}\left\|\xi^{E_A}_{\substack{\omega_{-i},z\\ \perp,y_i}} - \xi^{E_A}_{\substack{\omega_{-i},z\\ \perp,\perp}}\right\|_1 \\ &\geq 1 - O(\delta^{1/4}/\alpha),\end{aligned} \tag{4.33}$$

where the first inequality follows from the Fuchs-van de Graaf inequality (2.5) and the second uses Jensen's inequality and (4.28) from Claim 4.14. Expanding out the squared Euclidean norm and making sure that V_{ω_{-i},z,y_i} is chosen so as to ensure that the inner product $\langle\widetilde{\Phi}_{\substack{\omega_{-i},z\\ \perp,y_i}}|V_{\omega_{-i},z,y_i}|\widetilde{\Phi}_{\substack{\omega_{-i},z\\ \perp,\perp}}\rangle$ is positive real, (4.33) proves (4.10).

A nearly identical argument yields (4.9). It remains to show (4.11). Define

$$\xi^{E_A}_{\substack{\omega_{-i},z\\ \perp/x_i,y_i}} = \frac{1}{2}\xi^{E_A}_{\substack{\omega_{-i},z\\ \perp,y_i}} + \frac{1}{2}\xi^{E_A}_{\substack{\omega_{-i},z\\ x_i,y_i}} \qquad \text{and} \qquad \xi^{E_A}_{\substack{\omega_{-i},z\\ \perp/x_i,\perp}} = \frac{1}{2}\xi^{E_A}_{\substack{\omega_{-i},z\\ \perp,\perp}} + \frac{1}{2}\xi^{E_A}_{\substack{\omega_{-i},z\\ x_i,\perp}}$$

For notational clarity, we will suppress mention of ω_{-i} and z; it will be implicitly carried around.

The density matrices $\xi^{E_A}_{\perp/x_i,y_i}$ and $\xi^{E_A}_{\perp/x_i,\perp}$ are purified by $|\widetilde{\Phi}_{\perp/x_i,y_i}\rangle$ and $|\widetilde{\Phi}_{\perp/x_i,\perp}\rangle$ respectively. We will show that these two density matrices are close to together, on average, and hence by Uhlmann's Theorem implies that there exists a unitary V_{x_i,y_i} acting on E_B that moves $|\widetilde{\Phi}_{\perp/x_i,y_i}\rangle$ close to $|\widetilde{\Phi}_{\perp/x_i,\perp}\rangle$. Consider:

$$\begin{aligned}\mathop{\mathbb{E}}_{I}\ \mathop{\mathbb{E}}_{\Omega_{-i}Z|W}\ \mathop{\mathbb{E}}_{X_iY_i}\left\|\xi^{E_A}_{\perp/x_i,y_i} - \xi^{E_A}_{\perp,\perp}\right\|_1 &= \mathop{\mathbb{E}}_{I}\ \mathop{\mathbb{E}}_{\Omega_{-i}Z|W}\ \mathop{\mathbb{E}}_{X_iY_i}\left\|\frac{1}{2}\xi^{E_A}_{\perp,y_i} + \frac{1}{2}\xi^{E_A}_{x_i,y_i} - \xi^{E_A}_{\perp,\perp}\right\|_1 \\ &\leq \mathop{\mathbb{E}}_{I}\ \mathop{\mathbb{E}}_{\Omega_{-i}Z|W}\ \mathop{\mathbb{E}}_{X_iY_i}\left[\frac{1}{2}\left\|\xi^{E_A}_{\perp,y_i} - \xi^{E_A}_{\perp,\perp}\right\|_1 + \frac{1}{2}\left\|\xi^{E_A}_{x_i,y_i} - \xi^{E_A}_{\perp,\perp}\right\|_1\right].\end{aligned}$$

We obtained a bound on the first term in the calculations above. It remains to bound the second term. Again Lemma 4.8 implies

$$\mathsf{P}_I \cdot \mathsf{P}_{\Omega Z Y_i|D_i=A,W} \cong_{O(\sqrt{\delta}/\alpha^2)} \mathsf{P}_I \cdot \mathsf{P}_{\Omega_{-i}\mathrm{Z}|W} \cdot \mathsf{P}_{X_iY_i}$$

where "$\cong$" indicates approximate equality, up to relabeling the random variable M_i with X_i, whose marginals are identical conditioned on $D_i = A$. Thus using the same approach as earlier in the proof, we can obtain the bound

$$\mathop{\mathbb{E}}_{I} \mathop{\mathbb{E}}_{\Omega_{-i}Z|W} \mathop{\mathbb{E}}_{X_iY_i} \left\| \xi^{E_A}_{x_i,y_i} - \xi^{E_A}_{\perp,\perp} \right\|_1 \le O(\sqrt{\delta}/\alpha^4).$$

Thus there exists the desired unitary V_{x_i,y_i} such that

$$\frac{1}{m}\sum_i \mathop{\mathbb{E}}_{\Omega_{-i}Z|W} \mathop{\mathbb{E}}_{X_iY_i} \left\| \left| \widetilde{\Phi}_{\perp/x_i,\perp} \right\rangle - V_{x_i,y_i} \left| \widetilde{\Phi}_{\perp/x_i,y_i} \right\rangle \right\|^2 \le O(\delta^{1/4}/\alpha^4) \tag{4.34}$$

proving (4.11). □

4.4.3 Extending the argument to more than two players

We extend the argument from the previous sections to games with $k > 2$ entangled players. We describe the required modifications to the case of $k = 3$; the only hurdle in handling larger number of players is notational. Furthermore we restrict our attention to the repetition of the game $G_\perp$ obtained by applying the anchor transformation to a game G.

Let G be an arbitrary game involving three players Alice, Bob and Charlie. The players' questions are denoted by X, Y, Z, and their outputs are denoted as A, B, C. We will let $\mu(x, y, z)$ denote the question distribution of the game G. Let $G_\perp$ be the anchoring transformation applied to G (for some α), and let $\mu_\perp(x, y, z)$ denote the question distribution of $G_\perp$. We analyze the behavior of $\mathrm{val}^*(G_\perp^{\otimes n})$. Consider an optimal strategy for $G_\perp^{\otimes n}$, involving a tripartite state $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d \otimes \mathbb{C}^d$ and POVM for each of the players: $\{A^{a^n}_{x^n}\}$ for Alice, $\{B^{b^n}_{y^n}\}$ for Bob, and $\{C^{c^n}_{z^n}\}$ for Charlie. The entangled state $|\psi\rangle$ is supported on three registers E_A, E_B, and E_C.

The subset of coordinates that we condition on winning (formerly called C) will be denoted by S. The answers to rounds in S that we condition on will be denoted together as $Q = (A_S, B_S, C_S)$ (formerly called $Z = (A_C, B_C)$).

The idea behind the proof of the multiplayer extension is to reduce to the two-player case by "combining" two of the three players and treating them as a single player.

Dependency-breaking variable. The dependency-breaking variable Ω is constructed so that for each coordinate $i \notin S$, Ω_i fixes 2 out of 3 questions. That is, D_i is uniformly distributed over $\{\{A, B\}, \{A, C\}, \{B, C\}\}$. The variable D_i indicates which questions M_i is coupled to. For example, if $D_i = \{A, B\}$, then M_i is coupled to the pair (X_i, Y_i). The dependency breaking variable satisfies the property that for all ω, for all i, $\mathsf{P}_{X_i Y_i Z_i | \Omega = \omega}(x, y, z) = \mathsf{P}_{X_i | \Omega = \omega}(x) \cdot \mathsf{P}_{Y_i | \Omega = \omega}(y) \cdot \mathsf{P}_{Z_i | \Omega = \omega}(z)$.

Operators and states. We define the states and operators in a nearly identical way to the two-player case. We also introduce operators corresponding to the third player, $C^{c_S}_{\omega_{-i}, z_i}$, $C^{c_S}_{\omega_{-i}, \perp}$, $C^{c_S}_{\omega_{-i}, \perp/(x_i, y_i)}$, etc., defined in the obvious manner.

The states are also defined in a similar way:

$$|\Phi_{x,y,z}\rangle = \sqrt{A_x} \otimes \sqrt{B_y} \otimes \sqrt{C_z} |\psi\rangle$$

where x, y, and z can be "normal" questions from $\mathcal{X}$, $\mathcal{Y}$, or $\mathcal{Z}$, or they can be $\perp$ or a hybrid such as $\perp/x$.

The analogue of Lemma 4.10 in the three-player setting is the following. We use simplified notation to maximize clarity, and suppress mention of i, ω_{-i}, and $q = (a_S, b_S, c_S)$.

Lemma 4.15. *For all $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, there exist unitaries U_x, V_y, and W_z acting on E_A, E_B, and E_C respectively such that*

$$\mathop{\mathbb{E}}_{XYZ} \|(U_x \otimes V_y \otimes W_z)|\Phi_{\perp,\perp,\perp}\rangle - |\Phi_{x,y,z}\rangle\|^2 = O(\delta^{1/4}/\alpha^{2k}).$$

Proof sketch. Lemma 4.15, as in the two-player case, is proved in two steps. The first step is to establish the existence of unitaries U_x, V_y, and W_z such that $U_x|\Phi_{\perp,\perp,\perp}\rangle \approx |\Phi_{x,\perp,\perp}\rangle$, $V_x|\Phi_{\perp,\perp,\perp}\rangle \approx |\Phi_{\perp,y,\perp}\rangle$, and $W_z|\Phi_{\perp,\perp,\perp}\rangle \approx |\Phi_{\perp,\perp,z}\rangle$, with the unitaries acting on the appropriate spaces.

To prove, say, the existence of U_x, we treat Bob and Charlie as a single player – call him "SuperBob" – and use the analysis from the two-player case where the game G is a two player game involving Alice and SuperBob. Using the same reasoning as in the two-player case, we get that

$$\mathop{\mathbb{E}}_{XY} \|(U_x \otimes V_y \otimes \mathbb{I})|\Phi_{\perp,\perp,\perp}\rangle - |\Phi_{x,y,\perp}\rangle\|^2 = O(\delta^{1/4}/\alpha^{2k}).$$

It then only remains to show that, on average over (x, y, z), $(\mathbb{I} \otimes \mathbb{I} \otimes W_z)|\Phi_{x,y,\perp}\rangle$ is close to $|\Phi_{x,y,z}\rangle$:

$$\begin{aligned}
&\|W_z|\Phi_{x,y,\perp}\rangle - |\Phi_{x,y,z}\rangle\| \\
&= \left\|W_z C_\perp C_{\perp/z}^{-1/2}|\Phi_{x,y,\perp/z}\rangle - C_z C_{\perp/z}^{-1/2}|\Phi_{x,y,\perp/z}\rangle\right\| \\
&= \left\|H_{x,y,z} \otimes W_z C_\perp C_{\perp/z}^{-1/2}|\Phi_{x,y,\perp/z}\rangle - H_{x,y,z} \otimes C_z C_{\perp/z}^{-1/2}|\Phi_{x,y,\perp/z}\rangle\right\| \\
&\approx \left\|W_z C_\perp C_{\perp/z}^{-1/2}|\Phi_{\perp,\perp,\perp/z}\rangle - C_z C_{\perp/z}^{-1/2}|\Phi_{\perp,\perp,\perp/z}\rangle\right\| \\
&= \|W_z|\Phi_{\perp,\perp,\perp}\rangle - |\Phi_{\perp,\perp,z}\rangle\| \\
&\approx 0,
\end{aligned}$$

where $H_{x,y,z}$ is a unitary acting on $E_A E_B$ jointly such that $H_{x,y,z}|\Phi_{x,y,\perp/z}\rangle \approx |\Phi_{\perp,\perp,\perp/z}\rangle$. Such a unitary is analogous to that in (4.11). Taking into account the required normalization factors in to this calculation completes the proof of Lemma 4.15, □

The main theorem for the case of $k > 2$ entangled players follows from Lemma 4.15 using the same steps as in the two-player case.

4.4.4 A threshold theorem

We also observe that our proof nearly immediately yields a *threshold* version of our parallel repetition theorem: we can give an exponentially small bound on the probability that the players are able to win significantly more than a $(1-\varepsilon)n$ coordinates in the repeated game $G_\perp^{\otimes n}$, where $\mathrm{val}^*(G_\perp) = 1 - \varepsilon$. In [52], Rao shows how a Lemma of the form Lemma 4.9 yields not only a parallel repetition theorem, but also gives a concentration bound. Using essentially the same argument, we get the following theorem:

Theorem 4.16. *Let G be an α-anchored k-player game with $\mathrm{val}^*(G) \le 1 - \varepsilon$. Then for all integer $n \ge 1$ the probability that in the game $G^{\otimes n}$ the players can win more than $(1-\varepsilon+\gamma)n$*

games is at most

$$\left(1-\gamma^9/2\right)^{c\,\alpha^{8k}\,n/s}$$

where c is a universal constant and s is the length of the players' answers.

4.5 Classical multiplayer games

Perhaps the most well-known open problem about the classical parallel repetition of games is whether an analogue of Raz's theorem holds for games with more than two players. While the two-player case already presented a number of non-trivial difficulties, proving a parallel repetition theorem for three or more players is believed to require substantially new ideas.

In this section we show that multiplayer anchored games satisfy a classical parallel repetition theorem. Thus, the anchoring transformation along with parallel repetition yields a general hardness amplification technique for classical multiplayer games involving any number of players.[1]

Theorem 4.17. *Let $G = (\mathcal{X}, \mathcal{A}, \mu, V)$ be a k-player α-anchored game such that $\mathrm{val}(G) \leq 1-\varepsilon$. Then*

$$\mathrm{val}(G^{\otimes n}) \leq \exp\left(-\frac{\alpha^{2k}\cdot\varepsilon^3\cdot n}{384\cdot s\cdot k^2}\right), \tag{4.35}$$

where $s = \log|\mathcal{A}|$.

For the remainder of this section we fix a k-player α-anchored game $G = (\mathcal{X}, \mathcal{A}, \mu, V)$, an integer n, and a deterministic strategy for the k players in the repeated game $G^{\otimes n}$ that achieves success probability $\mathrm{val}(G^{\otimes n})$. In Section 4.5.1 we introduce the notation, random variables and basic lemmas for the proof. The proof of Theorem 4.17 itself is given in Section 4.5.2.

[1] There are other ways to perform hardness amplification of classical multiplayer games, including transforming a k-player game G into an equivalent two-player projection game G' (where one player simulates the original k players, and the second player is used to consistently check the answers of the new "super-player"), and then applying Raz's parallel repetition theorem to G'. However, this k-to-2 transformation does not preserve quantum completeness, in general, which may be a useful feature. The anchoring transformation, on the other hand, preserves quantum completeness, and simultaneously supports both classical and quantum hardness parallel repetition.

4.5.1 Breaking classical multipartite correlations

We refer to Section 2.1 for basic notation related to multiplayer games.

Let $C \subseteq [n]$ a fixed set of coordinates for the repeated game $G^{\otimes n}$ of size $|C| = n - m$. It will be convenient to fix $C = \{m+1, m+2, \dots, n\}$; the symmetry of the problem will make it clear that this is without loss of generality. Let $\mathrm{Z} = A_C = (A_C^1, A_C^2, \dots, A_C^k)$ denote the players' answers associated with the coordinates indexed by C.

For $t \in [k]$ let $\mathcal{Y}^t = (\mathcal{X}^t \setminus \mathcal{X}_\perp^t) \cup \{\perp\}$, and define a random variable

$$Y^t = \begin{cases} X^t, & X^t \in \mathcal{X}^t \setminus \mathcal{X}_\perp^t \\ \perp, & X^t \in \mathcal{X}_\perp^t \end{cases}. \tag{4.36}$$

Let $\mathcal{Y} = \mathcal{Y}^1 \times \mathcal{Y}^2 \times \ldots \times \mathcal{Y}^k$ and $Y = (Y^1, Y^2, \dots, Y^k)$. For $G^{\otimes n}$ we write

$$Y^{\otimes n} = (Y_1, Y_2, \dots, Y_n) = \left(\left(Y_1^1, \dots, Y_1^k\right), \left(Y_2^1, \dots, Y_2^k\right), \dots, \left(Y_n^1, \dots, Y_n^k\right)\right).$$

Note that each k-tuple Y_i is a deterministic function of X_i. Furthermore, we will write Y_i^{-t} to denote Y_i with the t-th coordinate Y_i^t omitted.

For $i \in [n]$ let D_i be a subset of $[k]$ of size $k-1$ chosen uniformly at random, and $\overline{D}_i \in [k]$ its complement in $[k]$. Let $M_i = Y_i^{D_i}$ denote the coordinates of Y associated to indices in D_i. Define the *dependence-breaking random variable* Ω_i as

$$\Omega_i = \begin{cases} (D_i, M_i) & i \in \overline{C} \\ X_i & i \in C \end{cases}. \tag{4.37}$$

The importance of Ω is captured in the following lemma.

Lemma 4.18. *(Local Sampling) Let X, Z, Ω be as above. Then $\mathsf{P}_{X_{-i}|X_i\Omega_{-i}\mathrm{Z}}$ is a product distribution across the players:*

$$\mathsf{P}_{X_{-i}|X_i\Omega_{-i}\mathrm{Z}} = \prod_{t=1}^{k} \mathsf{P}_{X_{-i}^t|\Omega_{-i}^t \mathrm{Z}^t X_i^t}.$$

Proof. Conditioned on $M_i = Y_i^{D_i}$ each $X_i = (X_i^1, X_i^2, \dots, X_i^k)$ is a product distribution, hence $\mathsf{P}_{X_{-i}|\Omega_{-i}X_i}$ is product. Since for $t \in [k]$ Z^t is a deterministic function of X^t the same holds of $\mathsf{P}_{X_{-i}|\Omega_{-i}\mathrm{Z}X_i}$. □

Lemma 4.18 crucially relies on the sets D_j being of size $k-1$: if two or more of the players' questions are unconstrained in a coordinate it is no longer necessarily true that $\mathsf{P}_{X_{-i}|\Omega_{-i}\mathrm{Z}X_i}$ is product across all players.

Let $W = W_C = \bigwedge_{i=1}^{C} W_i$ denote the event that the players' answers Z to questions in the coordinates indexed by C satisfy the predicate V. Let

$$\delta = \frac{|C| \log |\mathcal{A}| + \log \frac{1}{\Pr(W_C)}}{m}. \tag{4.38}$$

The following lemma and its corollary are direct consequences of analogous lemmas used in the analysis of repeated two-player games, as stated in e.g. [**?**, Lem. 5] and [**?**, Cor. 6]. They do not depend on the structure of the game, and only rely on W being an event defined only on (X_C, Z).

Lemma 4.19. *We have*

$$(i) \qquad \mathop{\mathbb{E}}_{i\in[m]} \|\mathsf{P}_{X_iY_i\Omega_i|W} - \mathsf{P}_{X_iY_i\Omega_i}\| \le \sqrt{\delta}.$$

$$(ii) \qquad \mathop{\mathbb{E}}_{i\in[m]} \|\mathsf{P}_{X_iY_i\mathrm{Z}\Omega_{-i}|W} - \mathsf{P}_{X_i|Y_i}\mathsf{P}_{Y_i\mathrm{Z}\Omega_{-i}|W}\| \le \sqrt{\delta}$$

$$(iii) \qquad \mathop{\mathbb{E}}_{i\in[m]} \|\mathsf{P}_{Y_i\mathrm{Z}\Omega|W} - \mathsf{P}_{Y_i|\Omega_i}\mathsf{P}_{\mathrm{Z}\Omega|W}\| \le \sqrt{\delta}.$$

Proof. Item (i) follows directly from [37, Lem. 5] by taking $U_i = X_iY_i\Omega_i$. For (ii) apply [37, Cor. 6] with $U_i = X_i$ and $T = (Y_1, Y_2, \dots, Y_m, X_C)$ to get

$$\mathop{\mathbb{E}}_{i\in[m]} \|\mathsf{P}_{X_i\mathrm{Z}Y_{[m]}X_C|W} - \mathsf{P}_{X_i|Y_i}\mathsf{P}_{Y_i\mathrm{Z}Y_{[m]\setminus\{i\}}X_C|W}\| \le \sqrt{\delta}, \tag{4.39}$$

which is stronger than (ii); (ii) follows by marginalizing $Y_i^{\overline{D}_i}$ in each term. Finally, the same corollary applied with $U_i = Y_i$ and $T = \Omega$ shows (iii). □

Corollary 4.20.

$$\mathop{\mathbb{E}}_{i\in[m]} \sum_{t=1}^{k} \|\mathsf{P}_{Y_i}\mathsf{P}_{\mathrm{Z}\Omega_{-i}|WY_i} - \mathsf{P}_{Y_i}\mathsf{P}_{\mathrm{Z}\Omega_{-i}|WY_i^{-t}}\| \le 3k\cdot\sqrt{\delta}.$$

Proof. We have $\mathsf{P}_{Y_i|\Omega_i}\mathsf{P}_{\mathrm{Z}\Omega|W} = \mathsf{P}_{Y_i|\Omega_i}\mathsf{P}_{\Omega_i|W}\mathsf{P}_{\mathrm{Z}\Omega_{-i}|W\Omega_i}$. Applying Lemma 2.1 with $\mathsf{Q}_F = \mathsf{P}_{\Omega_i|W}$, $\mathsf{S}_F = \mathsf{P}_{\Omega_i}$, and $\mathsf{R}_{G|F} = \mathsf{P}_{Y_i|\Omega_i}\mathsf{P}_{\mathrm{Z}\Omega_{-i}|W\Omega_i}$, we see that

$$\mathop{\mathbb{E}}_{i\in[m]} \|\mathsf{P}_{Y_i|\Omega_i}\mathsf{P}_{\mathrm{Z}\Omega|W} - \mathsf{P}_{Y_i\Omega_i}\mathsf{P}_{\mathrm{Z}\Omega_{-i}|W\Omega_i}\| = \mathop{\mathbb{E}}_{i\in[m]} \|\mathsf{P}_{\Omega_i|W} - \mathsf{P}_{\Omega_i}\| \le \sqrt{\delta},$$

where the last inequality follows from Lemma 4.19, item (i). Combining the above with item (iii) of the same Lemma, we have

$$\mathop{\mathbb{E}}_{i\in[m]} \|\mathsf{P}_{Y_i\mathrm{Z}\Omega|W} - \mathsf{P}_{Y_i\Omega_i}\mathsf{P}_{\mathrm{Z}\Omega_{-i}|W\Omega_i}\| \le 2\sqrt{\delta}. \tag{4.40}$$

Noting that Ω_i is determined by Y_i (the D_i are completely independent of everything else), (4.40) implies

$$\begin{aligned}\mathop{\mathbb{E}}_{i\in[m]}\mathop{\mathbb{E}}_{t\in[k]} \|\mathsf{P}_{Y_i\mathrm{Z}\Omega_{-i}|W} - \mathsf{P}_{Y_i}\mathsf{P}_{\mathrm{Z}\Omega_{-i}|WY_i^{-t}}\| &= \mathop{\mathbb{E}}_{i\in[m]} \|\mathsf{P}_{Y_i\mathrm{Z}\Omega_{-i}|W} - \mathsf{P}_{Y_i}\mathsf{P}_{\mathrm{Z}\Omega_{-i}|W\Omega_i}\| \\ &\le 2\sqrt{\delta}.\end{aligned}$$

Finally, notice that Lemmas 2.1 and 4.19 imply $\mathbb{E}_{i\in[m]} \|\mathsf{P}_{Y_i\mathrm{Z}\Omega_{-i}|W} - \mathsf{P}_{Y_i}\mathsf{P}_{\mathrm{Z}\Omega_{-i}|WY_i}\| = \mathbb{E}_{i\in[m]} \|\mathsf{P}_{Y_i} - \mathsf{P}_{Y_i|W}\| \le \sqrt{\delta}$; the desired result follows. □

4.5.2 Proof of the parallel repetition theorem

This section is devoted to the proof of Theorem 4.17. The main ingredient of the proof is given in the next proposition.

Proposition 4.21. *Let $C \subseteq [n]$ and $X, \mathrm{Z}, \Omega_{-i}$ be defined as in Section 4.5.1. Then*

$$\mathop{\mathbb{E}}_{i\in[m]} \left\|\mathsf{P}_{X_i\Omega_{-i}\mathrm{Z}|W} - \mathsf{P}_{X_i}\mathsf{P}_{\Omega_{-i}\mathrm{Z}|W,Y_i=\perp^k}\right\| \le (6k\alpha^{-k}+1)\sqrt{\delta}, \tag{4.41}$$

where δ is defined in (4.38).

Theorem 4.17 follows from this proposition in a relatively standard fashion; this is done at the end of this section. Let us now prove Proposition 4.21 assuming a certain technical statement, Lemma 4.22. This lemma is proved immediately after.

Proof of Proposition 4.21. First observe that

$$\left\|\mathsf{P}_{X_i\Omega_{-i}\mathsf{Z}|W} - \mathsf{P}_{X_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|W,Y_i=\perp^k}\right\| = \left\|\mathsf{P}_{X_iY_i\Omega_{-i}\mathsf{Z}|W} - \mathsf{P}_{X_iY_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|W,Y_i=\perp^k}\right\|$$

as Y_i is a deterministic function of X_i. Applying Lemma 4.19, item (ii) we get

$$\mathop{\mathbb{E}}_{i\in[m]}\left\|\mathsf{P}_{X_iY_i\Omega_{-i}\mathsf{Z}|W} - \mathsf{P}_{X_i|Y_i}\mathsf{P}_{Y_i\Omega_{-i}\mathsf{Z}|W}\right\| \leq \sqrt{\delta}.$$

The latter distribution can be written as $\mathsf{P}_{Y_i|W}\mathsf{P}_{X_i|Y_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|WY_i}$. Applying Lemma 2.1 with $\mathsf{Q}_F = \mathsf{P}_{Y_i|W}$ and $\mathsf{S}_F = \mathsf{P}_{Y_i}$ we see that

$$\left\|\mathsf{P}_{X_i|Y_i}\mathsf{P}_{Y_i\Omega_{-i}\mathsf{Z}|W} - \mathsf{P}_{X_iY_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|WY_i}\right\| = \left\|\mathsf{P}_{Y_i|W} - \mathsf{P}_{Y_i}\right\|,$$

which is bounded by $\sqrt{\delta}$ on average over i by Lemma 4.19, item (i). Hence

$$\begin{aligned}\mathop{\mathbb{E}}_{i\in[m]}\left\|\mathsf{P}_{X_i\Omega_{-i}\mathsf{Z}|W} - \mathsf{P}_{X_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|W,Y_i=\perp^k}\right\| &\leq 2\sqrt{\delta} + \mathop{\mathbb{E}}_{i\in[m]}\left\|\mathsf{P}_{X_iY_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|WY_i} - \mathsf{P}_{X_iY_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|W,Y_i=\perp^k}\right\| \\ &= 2\sqrt{\delta} + \mathop{\mathbb{E}}_{i\in[m]}\left\|\mathsf{P}_{Y_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|WY_i} - \mathsf{P}_{Y_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|W,Y_i=\perp^k}\right\|,\end{aligned}$$

where the equality follows from Lemma 2.1 applied with $\mathsf{R}_{G|F} = \mathsf{P}_{X_i|Y_i}$. Applying the triangle inequality,

$$\begin{aligned}&\mathop{\mathbb{E}}_{i\in[m]}\left\|\mathsf{P}_{X_iY_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|WY_i} - \mathsf{P}_{X_iY_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|W,Y_i=\perp^k}\right\| \\ &\qquad = \mathop{\mathbb{E}}_{i\in[m]}\left\|\mathsf{P}_{Y_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|WY_i} - \mathsf{P}_{Y_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|W,Y_i=\perp^k}\right\| \\ &\qquad \leq \mathop{\mathbb{E}}_{i\in[m]}\sum_{t=1}^{k}\left\|\mathsf{P}_{Y_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|WY_i^{<t}=\perp^{t-1},Y_i^{\geq t}} - \mathsf{P}_{Y_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|WY_i^{\leq t}=\perp^{t},Y_i^{>t}}\right\| && (4.42) \\ &\qquad \leq 6k\alpha^{-k}\cdot\sqrt{\delta}, && (4.43)\end{aligned}$$

where (4.42) is proved by Lemma 4.22 below and (4.43) follows from Corollary 4.20. □

Lemma 4.22. *Let $S \subset [k]$ and $t \in \overline{S}$. Then*

$$\left\|\mathsf{P}_{Y_i}\mathsf{P}_{\Omega_{-i}\mathrm{Z}|WY_i^S=\perp^S,Y_i^{\overline{S}}} - \mathsf{P}_{Y_i}\mathsf{P}_{\Omega_{-i}\mathrm{Z}|WY_i^{S\cup\{t\}}=\perp^{S\cup\{t\}},Y_i^{\overline{S}\setminus\{t\}}}\right\| \\ \leq 2\alpha^{-(|S|+1)} \cdot \left\|\mathsf{P}_{Y_i}\mathsf{P}_{\mathrm{Z}\Omega_{-i}|WY_i} - \mathsf{P}_{Y_i}\mathsf{P}_{\mathrm{Z}\Omega_{-i}|WY_i^{-t}}\right\|. \tag{4.44}$$

Proof. In the proof for ease of notation we omit the subscript i and write Y instead of Y_i. After relabeling we may assume $S = \{1, 2, \dots, r-1\}$ and $t = r$ where $1 \leq r < k$. Expanding the expectation over Y explicitly we can rewrite the left-hand side of (4.44) as

$$\left\|\mathsf{P}_Y \cdot \left(\mathsf{P}_{\Omega_{-i}\mathrm{Z}|W,y^{\geq r},y^{<r}=\perp^{r-1}} - \mathsf{P}_{\Omega_{-i}\mathrm{Z}|W,y^{>r},y^{\leq r}=\perp^r}\right)\right\|. \tag{4.45}$$

Next we use a symmetrization argument to bound the above expression. Consider a random variable $\hat{Y}$ that is a copy of Y, and is coupled to Y in the following way: $\hat{Y}^{-r} = Y^{-r}$, and conditioned on any setting of $Y^r = y^r$, $\hat{Y}^r$ and Y^r are independent. Using the fact that $\Pr[\hat{Y}^r = \perp] \geq \alpha$ conditioned on any value of $Y^{-r} = U^{-r} = y^{-r}$, we get that the expression in (4.45) is at most

$$\alpha^{-1}\left\|\mathsf{P}_{Y^{-r}}\mathsf{P}_{Y^r|Y^{-r}}\mathsf{P}_{\hat{Y}^r|Y^{-r}} \cdot \left(\mathsf{P}_{\Omega_{-i}\mathrm{Z}|W,y^{>r},y^r,y^{<r}=\perp^{r-1}} - \mathsf{P}_{\Omega_{-i}\mathrm{Z}|W,y^{>r},\hat{y}^r,y^{<r}=\perp^{r-1}}\right)\right\|.$$

Using the triangle inequality and symmetry of Y and $\hat{Y}$, this expression can be bounded by

$$2\alpha^{-1} \cdot \left\|\mathsf{P}_Y \cdot \left(\mathsf{P}_{\Omega_{-i}\mathrm{Z}|W,y^{>r},y^r,y^{<r}=\perp^{r-1}} - \mathsf{P}_{\Omega_{-i}\mathrm{Z}|W,y^{>r},y^{<r}=\perp^{r-1}}\right)\right\|,$$

which after noting that the quantity $\|\mathsf{P}_{\Omega_{-i}\mathrm{Z}|W,y^{>r},y^r,y^{<r}=\perp^{r-1}} - \mathsf{P}_{\Omega_{-i}\mathrm{Z}|W,y^{>r},y^{\leq r}=\perp^r}\|$ is independent of the variable $Y^{<r}$, can be rewritten as

$$2\alpha^{-1} \cdot \left\|\mathsf{P}_{Y^{\geq r}} \cdot \left(\mathsf{P}_{\Omega_{-i}\mathrm{Z}|W,y^{>r},y^r,y^{<r}=\perp^{r-1}} - \mathsf{P}_{\Omega_{-i}\mathrm{Z}|W,y^{>r},y^{<r}=\perp^{r-1}}\right)\right\|.$$

Using that the event that $Y^{<r} = \perp^{r-1}$ occurs with probability at least α^{r-1} and $\mathsf{P}_{Y^{\geq r}|Y^{<r}=\perp^{r-1}} =$

$\mathsf{P}_{Y^{\geq r}}$ by the anchor property, we can finally bound (4.45) by

$$2\alpha^{-r}\cdot\|\mathsf{P}_Y\mathsf{P}_{\mathrm{Z}\Omega_{-i}|WY}-\mathsf{P}_Y\mathsf{P}_{\mathrm{Z}\Omega_{-i}|WY^{-r}}\|,$$

which is the desired result. □

We prove Theorem 4.17 by iteratively applying Proposition 4.21 as follows.

Proof of Theorem 4.17. Let $C_0=\emptyset$ and $\delta_0=0$. While $(6k\alpha^{-k}+1)\sqrt{\delta_s}\leq\varepsilon/2$, by Proposition 4.21, we can choose $i\in\overline{C}_s$ with $\left\|\mathsf{P}_{X_i\Omega_{-i}\mathrm{Z}|W}-\mathsf{P}_{X_i}\mathsf{P}_{\Omega_{-i}\mathrm{Z}|W,Y_i=\perp^k}\right\|\leq\varepsilon/2$. Set $C_{s+1}=C_s\cup\{i\}$ and $\delta_{s+1}=(|C_{s+1}|\log|\mathcal{A}|+\log 1/\Pr(W_{C_{s+1}}))/m$. First we show that throughout this process the bound

$$\Pr[W_{C_s}]\leq(1-\varepsilon/2)^{|C_s|}\tag{4.46}$$

holds. Since by the choice of i one has $\left\|\mathsf{P}_{X_i\Omega_{-i}\mathrm{Z}|W_C}-\mathsf{P}_{X_i}\mathsf{P}_{\Omega_{-i}\mathrm{Z}|W_C,Y_i=\perp^k}\right\|\leq\varepsilon/2$, to establish (4.46) it will suffice to show that

$$\Pr(W_i|W_C)\leq\mathrm{val}(G)+\left\|\mathsf{P}_{X_i\Omega_{-i}\mathrm{Z}|W_C}-\mathsf{P}_{X_i}\mathsf{P}_{\Omega_{-i}\mathrm{Z}|W_C,Y_i=\perp^k}\right\|.\tag{4.47}$$

The proof of (4.47) is based on a rounding argument. Consider the following strategy for G: First, the players use shared randomness to obtain a common sample from $\mathsf{P}_{\Omega_{-i}\mathrm{Z}|W_C,Y_i=\perp^k}$. After receiving her question x_t^*, player $t\in[k]$ samples questions for the remaining coordinates according to $\mathsf{P}_{X_{-i}^t|\Omega_{-i}^t\mathrm{Z}^tX_i^t}$, forming the tuple $X^t=(X_{-i}^t,x_t^*)$. She determines her answer $a_i^t\in\mathcal{A}_i^t$ according to the strategy for $G^{\otimes n}$. The distribution over questions X implemented by players following this strategy is

$$\mathsf{P}_{X_i}\mathsf{P}_{\Omega_{-i}\mathrm{Z}|W_CY_i=\perp^k}\prod_{t=1}^k\mathsf{P}_{X_{-i}^t|\Omega_{-i}^t\mathrm{Z}^tX_i^t},$$

which by Lemma 4.18 is equal to

$$\mathsf{P}_{X_i}\mathsf{P}_{\Omega_{-i}\mathrm{Z}|W_CY_i=\perp^k}\mathsf{P}_{X_{-i}|\Omega_{-i}\mathrm{Z}}.$$

On the other hand from the definition of Ω_{-i} we have

$$\mathsf{P}_{X\Omega_{-i}\mathsf{Z}|W_C} = \mathsf{P}_{X_i\Omega_{-i}\mathsf{Z}|W_C}\mathsf{P}_{X_{-i}|\Omega_{-i}\mathsf{Z}W_C} = \mathsf{P}_{X_i\Omega_{-i}\mathsf{Z}|W_C}\mathsf{P}_{X_{-i}|\Omega_{-i}\mathsf{Z}}.$$

Applying Lemma 2.1 with $R = \mathsf{P}_{X_{-i}|\Omega_{-i}\mathsf{Z}}$ it follows that

$$\left\|\mathsf{P}_{X\mathsf{Z}\Omega_{-i}|W_C} - \mathsf{P}_{X_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|W_C Y_i=\perp^k}\mathsf{P}_{X_{-i}|\Omega_{-i}\mathsf{Z}}\right\| = \left\|\mathsf{P}_{X_i\Omega_{-i}\mathsf{Z}|W_C} - \mathsf{P}_{X_i}\mathsf{P}_{\Omega_{-i}\mathsf{Z}|W_C,Y_i=\perp^k}\right\|.$$

Now by definition the winning probability of the extracted strategy for G is at most $\mathrm{val}(G)$, and (4.47) follows.

Let now C be the final set of coordinates when the above-described process stops; at this point we must have

$$\delta = \frac{|C|\log|\mathcal{A}| + \log\frac{1}{\Pr(W_C)}}{n-|C|} > \frac{\alpha^{2k}\varepsilon^2}{48\cdot k^2}.$$

If $|C| \geq n/2$ we are already done by (4.46). Suppose $\frac{|C|\log|\mathcal{A}|+\log(\frac{1}{\Pr[W_C]})}{n} > \frac{\alpha^{2k}\epsilon^2}{96\cdot k^2}$. If $\log(\frac{1}{\Pr(W_C)}) \geq \frac{n\cdot\alpha^{2k}\epsilon^2}{192\cdot k^2}$ we are again done; hence, we can assume

$$\frac{|C|\log|\mathcal{A}|}{n} > \frac{\alpha^{2k}\varepsilon^2}{192\cdot k^2}.$$

Now plugging the lower bound on the size of C in (4.46) we get

$$\mathrm{val}(G^{\otimes n}) \leq \Pr(W_C) \leq \exp\left(-\frac{\alpha^{2k}\cdot\varepsilon^3\cdot n}{384\cdot k^2\cdot s}\right)$$

where $s = \log|\mathcal{A}|$, which completes the proof. □

4.6 Some remarks on multiplayer parallel repetition

We conclude this chapter with some remarks about Theorem 4.17 and the more general problem of multiplayer parallel repetition. Our analysis of repeated anchored games follows the information-theoretic approach of Raz and Holenstein. It is a natural (old) question whether one can extend this framework to prove parallel repetition for general multiplayer games?.

At first sight the Raz/Holenstein framework may seem quite suitable for multiplayer parallel repetition. For instance, it is folklore that classically the approach extends to the case of free games with any number of players, and furthermore, many of the other technical components of the proof readily carry over in much generality. Despite these positive signs, attempts to extend Raz's original argument to the general multiplayer setting have so far failed for different and rather interesting *technical reasons.* Embarrassingly, to our knowledge, it is not even known how to extend the information-theoretic approach to prove that the value of a repeated k-player game decays at all![2]

We give an example of one of the difficulties in proving a multiplayer parallel repetition theorem for general games. Consider the problem of defining an appropriate dependency-breaking variable Ω in the multiplayer setting. There are two competing demands on Ω: on one hand the breaking of dependencies between the players' respective questions seems to require it to contain as many of the players' questions as possible for each coordinate $i \in \overline{C}$. In fact, if the correlations between the players inputs' are generic, it seems hard to avoid the need to keep at least $k-1$ inputs in each Ω_i, as we do in Lemma 4.18. On the other hand, for correlated sampling to be possible, it seems necessary for Ω to specify very few of the questions per coordinate, or in fact in the generic case, at most 1; as soon as $k \geq 3$ both requirements are in direct contradiction.

An insight behind our result is that it is sometimes possible to decouple the above two competing demands on Ω (i.e. the *dependency-breaking* and the *correlated sampling* components). More precisely, when the base game is anchored, we show how to define a useful dependency-breaking variable (or quantum state, in the entangled players setting) that can be sampled *without* correlated sampling. With correlated sampling out of the way, the aforementioned conflict between correlated sampling and dependency-breaking disappears, allowing us to proceed with the argument.

[2] One can modify an argument of Verbitsky based on Hales-Jewett theorem to show that if $\text{val}(G) < 1$, then $\text{val}(G^{\otimes n})$ must go to 0 as n grows [59], but the bound on the rate of decay is extremely poor.

Chapter 5

Conclusion and Open Problems

In this thesis, we showed two methods for proving strong hardness amplification results for entangled games and multiplayer games. Many interesting problems about the parallel repetition of multiplayer and entangled games however remain open. Perhaps the most pressing of these is the problem of obtaining a complete extension of Raz's theorem for general entangled two-player games (see Yuen [63] for a new work in this direction obtaining a polynomial decay for all two-player entangled games). For example, obtaining a fully quantum analogue of Raz's theorem, as was the case for Raz's theorem itself, is likely to have important implications in the setting of communication complexity. One promising candidate approach could be to leverage the recent ideas related to quantum information complexity [57, 14].

Similarly, proving a parallel repetition theorem with exponential decay for *general multiplayer* games remain a fascinating challenge. In our view, however, this problem (even classically) seems more challenging than the two-player entangled case, as its difficulties are related to communication complexity and circuit complexity lower bounds.

An important message of Chpater 3 is that there is a modified form of game concatenation with no adverse effect on the entangled value. This is notable since ordinary concatenation may appear to be not very well-behaved with respect to quantum strategies: we typically do not expect that entangled players would be able to answer a number of questions from a game G simultaneously, while preserving the same question/answer statistics as in G, as players' measurement operators associated with different questions generally do not commute.

The concatenation and composition of games play an important role in the classical setting in the context of PCPs [5, 4, 25]. It remains to be seen whether ideas related to our ordered fortification can be useful in lifting some of these techniques to the quantum world. Even though we believe concatenation will prove a useful tool in quantum complexity and cryptography as has been the case case in classical complexity, in the context of game variant of quantum PCP [31], one should bear in mind the recent work of Ji [42] which rules out some amplification approaches toward this conjecture.

Bibliography

[1] Scott Aaronson, Russell Impagliazzo, and Dana Moshkovitz. AM with multiple merlins. In *Conference on Computational Complexity (CCC)*, pages 44–55, 2014.

[2] Dorit Aharonov, Itai Arad, and Thomas Vidick. Guest column: the quantum PCP conjecture. *Acm sigact news*, 44(2):47–79, 2013.

[3] Rotem Arnon-Friedman, Renato Renner, and Thomas Vidick. Non-signalling parallel repetition using de finetti reductions. *arXiv preprint arXiv:1411.1582*, 2014.

[4] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM (JACM)*, 45(3):501–555, 1998.

[5] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of np. *Journal of the ACM (JACM)*, 45(1):70–122, 1998.

[6] Boaz Barak, Anup Rao, Ran Raz, Ricky Rosen, and Ronen Shaltiel. Strong parallel repetition theorem for free projection games. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 352–365. Springer, 2009.

[7] Mohammad Bavarian, Thomas Vidick, and Henry Yuen. Hardness amplification for entangled games via anchoring. In *Symposium on the Theory of Computing (STOC)*, 2017.

[8] Mohammad Bavarian, Thomas Vidick, and Henry Yuen. Parallel repetition via fortification: analytic view and the quantum case. In *Innovations in Theoretical ComputerScience (ITCS)*, 2017.

[9] John S Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1(3), 1964.

[10] Michael Ben-Or, Shafi Goldwasser, Joe Kilian, and Avi Wigderson. Multi-prover interactive proofs: How to remove intractability assumptions. In *Proceedings of Symposium on Theory of computing (STOC)*, 1988.

[11] Amey Bhangale, Ramprasad Saptharishi, Girish Varma, and Rakesh Venkat. On fortification of projection games. In *RANDOM.* (arXiv:1504.05556), 2015.

[12] Yonatan Bilu and Nathan Linial. Lifts, discrepancy and nearly optimal spectral gap. *Combinatorica*, 26(5):495–519, 2006.

[13] Mark Braverman and Ankit Garg. Small value parallel repetition for general games. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, STOC, 2015.

[14] Mark Braverman, Ankit Garg, Young Kun Ko, Jieming Mao, and Dave Touchette. Near-optimal bounds on bounded-round quantum communication complexity of disjointness. In *Proceedings of Foundations of Computer Science (FOCS)—to appear*, 2015.

[15] Mark Braverman, Anup Rao, Omri Weinstein, and Amir Yehudayoff. Direct products in communication complexity. In *Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on*, pages 746–755. IEEE, 2013.

[16] Mark Braverman and Omri Weinstein. An interactive information odometer with applications. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 21, page 47, 2014.

[17] Harry Buhrman, Richard Cleve, Serge Massar, and Ronald de Wolf. Nonlocality and communication complexity. *Reviews of modern physics*, 82(1):665, 2010.

[18] Harry Buhrman, Serge Fehr, and Christian Schaffner. On the parallel repetition of multi-player games: The no-signaling case. *arXiv preprint arXiv:1312.7455*, 2013.

[19] André Chailloux and Giannicola Scarpa. Parallel repetition of entangled games with exponential decay via the superposed information cost. In *Proceeding of International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 296–307, 2014.

[20] Kai-Min Chung, Xiaodi Wu, and Henry Yuen. Parallel repetition for entangled k-player games via fast quantum search. In *30th Conference on Computational Complexity*, page 512, 2015.

[21] John F Clauser, Michael A Horne, Abner Shimony, and Richard A Holt. Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, 23(15):880–884, 1969.

[22] Richard Cleve, Peter Høyer, Benjamin Toner, and John Watrous. Consequences and limits of nonlocal strategies. In *Computational Complexity, 2004. Proceedings. 19th IEEE Annual Conference on*, pages 236–249. IEEE, 2004.

[23] Richard Cleve, William Slofstra, Falk Unger, and Sarvagya Upadhyay. Perfect parallel repetition theorem for quantum XOR proof systems. *Computational Complexity*, 17(2):282–299, 2008.

[24] Matthew Coudron and Henry Yuen. Infinite randomness expansion with a constant number of devices. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 427–436. ACM, 2014.

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

[26] Irit Dinur, Prahladh Harsha, Rakesh Venkat, and Henry Yuen. Multiplayer parallel repetition for expander games. *arXiv preprint arXiv:1610.08349*, 2016.

[27] Irit Dinur and David Steurer. Analytical approach to parallel repetition. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 624–633. ACM, 2014.

[28] Irit Dinur, David Steurer, and Thomas Vidick. A parallel repetition theorem for entangled projection games. In *the 29th Conference on Computational Complexity, CCC*, pages 197–208, 2014.

[29] Uriel Feige. Error reduction by parallel repetition: The state of the art. *Technical Report CS95-32 of the Weizmann Institute*, 1995.

[30] Uriel Feige and Joe Kilian. Two-prover protocols—low error at affordable rates. *SIAM Journal on Computing*, 30(1):324–346, 2000.

[31] Joseph Fitzsimons and Thomas Vidick. A multiprover interactive proof system for the local hamiltonian problem. *arXiv preprint arXiv:1409.0260*, 2014.

[32] Joseph Fitzsimons and Thomas Vidick. A multiprover interactive proof system for the local hamiltonian problem. In *Proceedings of the Conference on Innovations in Theoretical Computer Science*, pages 103–112. ACM, 2015.

[33] Lance Fortnow, John Rompel, and Michael Sipser. On the power of multi-power interactive protocols. In *Proceedings of Structure in Complexity Theory Conference (CCC)*, pages 156–161, 1988.

[34] Iftach Haitner. A parallel repetition theorem for any interactive argument. In *Foundations of Computer Science, 2009. FOCS'09. 50th Annual IEEE Symposium on*, pages 241–250. IEEE, 2009.

[35] Johan Håstad. Some optimal inapproximability results. *Journal of the ACM (JACM)*, 48(4), 2001.

[36] Johan Håstad, Rafael Pass, Douglas Wikström, and Krzysztof Pietrzak. An efficient parallel repetition theorem. In *Theory of Cryptography Conference*, pages 1–18. Springer, 2010.

[37] Thomas Holenstein. Parallel repetition: simplifications and the no-signaling case. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 411–419. ACM, 2007.

[38] Tsuyoshi Ito and Thomas Vidick. A multi-prover interactive proof for NEXP sound against entangled provers. In *Proceedings of Foundations of Computer Science (FOCS)*, pages 243–252, 2012.

[39] Rahul Jain. New strong direct product results in communication complexity. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 18, page 2, 2011.

[40] Rahul Jain, Attila Pereszlényi, and Penghui Yao. A parallel repetition theorem for entangled two-player one-round games under product distributions. In *Proceedings of Conference on Computational Complexity (CCC)*, pages 209–216, 2014.

[41] Zhengfeng Ji. Classical verification of quantum proofs. *arXiv preprint arXiv:1505.07432*, 2015.

[42] Zhengfeng Ji. Compression of quantum multi-prover interactive proofs. *arXiv preprint arXiv:1610.03133*, 2016.

[43] Julia Kempe and Thomas Vidick. Parallel repetition of entangled games. In *Proceedings of the forty-third annual ACM symposium on Theory of computing (STOC)*, pages 353–362, 2011.

[44] Alexei Yu Kitaev, Alexander Shen, and Mikhail N Vyalyi. *Classical and quantum computation*, volume 47. American Mathematical Society Providence, 2002.

[45] Fuad Kittaneh. Inequalities for the schatten p-norm IV. *Communications in Mathematical Physics*, 106(4):581–585, 1986.

[46] Cécilia Lancien and Andreas Winter. Parallel repetition and concentration for (sub-) no-signalling games via a flexible constrained de finetti reduction. *arXiv preprint arXiv:1506.07002*, 2015.

[47] Carl A Miller and Yaoyun Shi. Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 417–426. ACM, 2014.

[48] Dana Moshkovitz. Parallel repetition from fortification. In *Annual Symposium on Foundations of Computer Science (FOCS)*, pages 414–423, 2014.

[49] Anand Natarajan and Thomas Vidick. Constant-soundness interactive proofs for local hamiltonians. *arXiv preprint arXiv:1512.02090*, 2015.

[50] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

[51] Rafael Pass and Muthuramakrishnan Venkitasubramaniam. An efficient parallel repetition theorem for arthur-merlin games. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 420–429. ACM, 2007.

[52] Anup Rao. Parallel repetition in projection games and a concentration bound. *SIAM Journal on Computing*, 40(6):1871–1891, 2011.

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

[54] Ran Raz. Parallel repetition of two prover games (invited survey). In *2010 25th Annual IEEE Conference on Computational Complexity*, pages 3–6. IEEE, 2010.

[55] Alexander A. Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992.

[56] Ben W Reichardt, Falk Unger, and Umesh Vazirani. Classical command of quantum systems. *Nature*, 496(7446):456–460, 2013.

[57] Dave Touchette. Quantum information complexity and amortized communication. *arXiv preprint arXiv:1404.3733*, 2014.

[58] Umesh Vazirani and Thomas Vidick. Certifiable quantum dice. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1971):3432–3448, 2012.

[59] Oleg Verbitsky. Towards the parallel repetition conjecture. *Theoretical Computer Science*, 157(2):277–282, 1996.

[60] Thomas Vidick. Three-player entangled XOR games are NP-hard to approximate. In *Proceedings of Annual Symposium on Foundations of Computer Science (FOCS)*, pages 766–775. IEEE, 2013.

[61] Reinhard F Werner and Michael M Wolf. Bell inequalities and entanglement. *Quantum Information & Computation*, 1(3):1–25, 2001.

[62] Mark M Wilde. *Quantum information theory.* Cambridge University Press, 2013.

[63] Henry Yuen. A parallel repetition theorem for all entangled games. In *Proceedings of the 43rd International Colloquium on Automata, Languages, and Programming (ICALP)*, 2016.