Quantum soundness of the classical low individual degree test

Zhengfeng Ji*
University of Technology Sydney

Anand Natarajan†
Massachusetts Institute of Technology

Thomas Vidick‡
California Institute of Technology

John Wright§
University of Texas at Austin

Henry Yuen¶
University of Toronto

Abstract

Low degree tests play an important role in classical complexity theory, serving as basic ingredients in foundational results such as MIP = NEXP [BFL91] and the PCP theorem [AS98, ALM+98]. Over the last ten years, versions of these tests which are sound against quantum provers have found increasing applications to the study of nonlocal games and the complexity class MIP*. The culmination of this line of work is the result MIP* = RE [JNV+20].

One of the key ingredients in the first reported proof of MIP* = RE is a two-prover variant of the low degree test, initially shown to be sound against multiple quantum provers in [Vid16]. Unfortunately a mistake was recently discovered in the latter result, invalidating the main result of [Vid16] as well as its use in subsequent works, including [JNV+20].

We analyze a variant of the low degree test called the low individual degree test. Our main result is that the two-player version of this test is sound against quantum provers. This soundness result is sufficient to re-derive several bounds on MIP* that relied on [Vid16], including MIP* = RE.

---

*zhengfeng.ji@uts.edu.au

†anandn@mit.edu. Most of this work performed while affiliated with the California Institute of Technology.

‡vidick@caltech.edu

§wright@cs.utexas.edu. Most of this work performed while also affiliated with the California Institute of Technology.

¶hyen@cs.toronto.edu

Contents

1 Introduction

An m-variate polynomial over the finite field $\mathbb{F}_q$ is a function $g : \mathbb{F}_q^m \to \mathbb{F}_q$ of the form

$$g(x_1, \ldots, x_m) = \sum_{i_1, \ldots, i_m} c_{i_1, \ldots, i_m} \cdot x_1^{i_1} \cdots x_m^{i_m},$$

where each coefficient $c_{i_1, \ldots, i_m}$ is an element of $\mathbb{F}_q$. We say that g has *total degree d* (or *degree d*, for short) if $i_1 + \cdots + i_m \leq d$ for each nonzero coefficient $c_{i_1, \ldots, i_m}$, and *individual degree d* if $i_1, \ldots, i_m \leq d$ for each nonzero coefficient $c_{i_1, \ldots, i_m}$. Low-degree polynomials have a variety of properties which make them useful in theoretical computer science, chief among which is their *distance*: by the Schwartz-Zippel lemma, two nonequal degree d polynomials g and h agree on at most a d/q fraction of the points in $\mathbb{F}_q^m$.

Low (individual) degree testing refers to the task of verifying that an unknown function $g : \mathbb{F}_q^m \to \mathbb{F}_q$ is representable as a polynomial of (individual) degree d by querying g on a small number of points $u \in \mathbb{F}_q^m$. There is a pair of canonical tests for doing so known as the *surface-versus-point low-degree test* and the *low individual degree test*. It is common to frame these tests as games between a referee and two provers. In this setting, the surface-versus-point low degree test, parameterized by an integer $k \geq 1$, is performed by the verifier as follows.

1. Select $\boldsymbol{u} \sim \mathbb{F}_q^m$ uniformly at random. Give it to Prover A. They respond with a value $\boldsymbol{a} \in \mathbb{F}_q$.

2. Select a uniformly random k-dimensional affine surface $\boldsymbol{s}$ in $\mathbb{F}_q^m$ containing $\boldsymbol{u}$. Give it to Prover B. They respond with a degree-d k-variate polynomial $\boldsymbol{f} : \boldsymbol{s} \to \mathbb{F}_q$.

3. Accept if $\boldsymbol{f}(\boldsymbol{u}) = \boldsymbol{a}$.

This test is motivated by the following "local characterization" of low-degree polynomials: a polynomial $g : \mathbb{F}_q^m \to \mathbb{F}_q$ is degree-d if and only if $g|_s$ is degree-d for all k-dimensional surfaces s. Hence, if g is degree-d, then the provers can win with probability 1 by always replying with $\boldsymbol{a} = g(\boldsymbol{u})$ and $\boldsymbol{f} = g|_{\boldsymbol{s}}$. *Soundness* of the low-degree test refers to the converse statement, namely that players who succeed with high probability must be responding based on a low-degree polynomial. This is formalized as follows.

Theorem 1.1 (Raz-Safra [RS97]). *Suppose Provers* A *and* B *pass the $k = 2$ surface-versus-point low-degree test with probability $1 - \epsilon$. Then there exists a degree-d polynomial $g : \mathbb{F}_q^m \to \mathbb{F}_q$ such that*

$$\Pr_{\boldsymbol{u} \sim \mathbb{F}_q^m} [g(\boldsymbol{u}) = \boldsymbol{a}] \geq 1 - \epsilon - \mathrm{poly}(m) \cdot \mathrm{poly}(d/q).$$

A similar "local characterization" of low individual degree polynomials states that a polynomial g has individual degree d if and only if $g|_\ell$ is a univariate degree-d polynomial for all axis-parallel lines ℓ. An axis-parallel line is a line of the form $\ell = \{u + a \cdot e_i \mid i \in \mathbb{F}_q\}$, for $u \in \mathbb{F}_q^m$ and $i \in \{1, \ldots, m\}$. Motivated by this, the low individual degree test follows the same outline as the surface-versus-point low degree test except with the second step substituted with the following.

2. Select a uniformly random axis-parallel line $\boldsymbol{\ell}$ in $\mathbb{F}_q^m$ containing $\boldsymbol{u}$. Give it to Prover B. They respond with a degree-d univariate polynomial $\boldsymbol{f} : \boldsymbol{\ell} \to \mathbb{F}_q$.

When $d = 1$, the low individual degree test is called the *multilinearity test* because a polynomial with individual degree $d = 1$ is a multilinear polynomial. The multilinearity and low individual degree tests were first introduced and proven sound by Babai, Fortnow, and Lund in [BFL91]. The

analysis of its soundness was then improved by [AS98] and then further sharpened by [FHS94]. The best bound follows from the work of Polishchuk and Spielman [PS94]; their work considers only the bivariate $m = 2$ case, but extending it to the multivariate case yields the following result.

Theorem 1.2 (Polishchuk-Spielman [PS94])**.** *Suppose Provers* A *and* B *pass the low individual degree test with probability* $1 - \epsilon$. *Then there exists a polynomial* $g : \mathbb{F}_q^m \to \mathbb{F}_q$ *with individual degree* d *such that*

$$\Pr_{\boldsymbol{u} \sim \mathbb{F}_q^m} [g(\boldsymbol{u}) = \boldsymbol{a}] \geq 1 - \text{poly}(m) \cdot (\text{poly}(\epsilon) + \text{poly}(d/q)).$$

We note that the soundness error the low individual test gives is actually worse than the low degree test, because the low individual degree function g is only $\text{poly}(m) \cdot (\text{poly}(\epsilon) + \text{poly}(d/q))$ close to Player A's strategy, rather than $\epsilon + \text{poly}(m) \cdot \text{poly}(d/q)$. We will discuss this weakness of the low individual degree test below.

The multilinearity test, low individual degree test, and low-degree test form a sequence in which each test generally enables more applications than the previous one. The multilinearity test can be used to show that $\mathsf{MIP} = \mathsf{NEXP}$ using a polynomial number of rounds [BFL91], the low individual degree test can reduce the number of rounds to 1, and the low degree test can be used to "scale this result down" and prove the PCP theorem, i.e. $\mathsf{NP} = \mathsf{MIP}[O(\log(n)), O(1)]$ [AS98, ALM$^+$98].

1.1 Quantum soundness of the low degree tests

The work of Ito and Vidick [IV12] initiated a program of studying these tests in the case when the players are quantum, as a means of proving bounds on the complexity class MIP^*. Because the provers are quantum, they are allowed to share an entangled state, a resource which could potentially allow them to "cheat" the test and win without using a low-degree polynomial. The goal of this program is to show that this is not possible. In other words, the goal is to show that these tests are *quantum sound*, which means that provers who succeed with high success probability must answer their questions according to a low (individual) degree polynomial, even if they are allowed to share quantum entanglement.[1]

Correctly formalizing the notion of quantum soundness is a subtle task, as quantum provers can in fact ace these tests using a broader class of strategies than their classical counterparts. For example, the two provers can use their quantum state $|\psi\rangle$ to simulate shared randomness, which they can use to sample a random low-degree polynomial $\boldsymbol{g}$ to answer their questions with; what makes this still acceptable is that $\boldsymbol{g}$ depends only on their shared randomness and not their questions. The correct formalization of quantum soundness was identified by Ito and Vidick [IV12], which states the following: suppose Provers A and B pass the low-degree test with probability close to 1. For each point question $u \in \mathbb{F}_q^m$, let $A^u = \{A_a^u\}$ be the measurement that Prover A applies to their share of $|\psi\rangle$ to produce the answer $\boldsymbol{a} \in \mathbb{F}_q$. Then the test being quantum sound means that there should be a measurement $G = \{G_g\}$, independent of $u \in \mathbb{F}_q^m$, which outputs degree-d polynomials g and "acts like A". In other words, rather than measuring A^u to produce the outcome $\boldsymbol{a} \in \mathbb{F}_q$, Prover A could have simply measured G, received the polynomial $\boldsymbol{g}$, and outputted its evaluation at $\boldsymbol{u}$, i.e. the value $\boldsymbol{g}(\boldsymbol{u})$. We will measure the similarity between A and G by considering the experiment where Prover A measures with $A^{\boldsymbol{u}}$ to produce $\boldsymbol{a}$, Prover B measures

---

[1]We note that *quantum soundness of the low-degree tests*, which is the focus of this work, is distinct from *soundness of the quantum low-degree test*. The "quantum low-degree test" is a particular test introduced by Natarajan and Vidick in [NV18a] which gets its name from the prominent role that the low-degree test plays as a subroutine, and its "soundness" is simply the result that they prove about it.

	Test shown quantum-sound	Complexity-theoretic consequence	Number of provers
1. [IV12]:	multilinearity test	$\mathsf{NEXP} \subseteq \mathsf{MIP}^*$	3
2. [Vid16]:	low-degree test	$\mathsf{NP} \subseteq \mathsf{MIP}^*[O(\log(n)), O(1)]$	3
3. [NV18b]:	low-degree test	$\mathsf{NP} \subseteq \mathsf{MIP}^*[O(\log(n)), O(1)]$	2
Consequences of [NV18b]:			
(a) [NV18a]:		$\mathsf{QMA} \subseteq \mathsf{MIP}^*[O(\log(n)), O(1)]$ (under randomized reductions)	7
(b) [NW19]:		$\mathsf{NEEXP} \subseteq \mathsf{MIP}^*$	2
(c) [JNV$^+$20]:		$\mathsf{MIP}^* = \mathsf{RE}$	2

Figure 1: Prior work on quantum-sound low degree tests and their complexity-theoretic consequences. The first three works showed a quantum-sound test and an MIP^* bound, both involving the same number of provers indicated in the final column. The last three works use the low-degree test from [NV18b] to show the indicated MIP^* bound.

with G to produce $\boldsymbol{g}$, and we check if $\boldsymbol{g}(\boldsymbol{u}) = \boldsymbol{a}$. This entails studying the quantity

$$\mathop{\mathbf{E}}_{\boldsymbol{u} \sim \mathbb{F}_q^m} \sum_{a \in \mathbb{F}_q} \sum_{g:g(\boldsymbol{u})=a} \langle \psi | \, A_a^{\boldsymbol{u}} \otimes G_g \, | \psi \rangle \, ,$$

which we aim to show is as close to 1 as possible. In this way, the provers' quantum advantage is limited to their ability to select a low-degree polynomial g.

One additional quirk of the quantum setting is that it has been historically useful to consider variants of these tests which feature more than two provers. This allows one to use monogamy of entanglement to reduce the power that entanglement gives to the provers, making it easier to show that a given test is quantum sound. Low degree test results with fewer provers are more difficult to show and have more applications.

The program of showing that these tests are quantum sound was carried out for the 3-prover multilinearity test by Ito and Vidick [IV12] and for the 3-prover low-degree test by Vidick [Vid16], which was later improved to 2-provers by Natarajan and Vidick [NV18b]. This latter result led to a sequence of works which culminated in the proof that $\mathsf{MIP}^* = \mathsf{RE}$ and the refutation of the Connes embedding conjecture in [JNV$^+$20]. We summarize this line of research in Figure 1.

Subsequent to the initial posting of [JNV$^+$20] on the arXiv, an error was discovered in the analysis of the quantum-sound low degree test contained in [Vid16] which was propagated to [NV18b]. The error affects the proof in a manner that appears difficult to fix. As such, we currently do not know if the low-degree test is quantum-sound for any number of provers. The invalidation of this analysis affects every result in Figure 1 except for [IV12].

The purpose of this work is to provide a different soundness analysis, for a variant of the low-degree test, that can nevertheless be used as a replacement for it in most subsequent works. We do so by revisiting the three-player quantum-sound multilinearity test of [IV12] and improving this result in two ways. First, we generalize it to hold for the degree-d low individual degree test, of which the multilinearity test is the $d = 1$ special case. Second, using techniques introduced in [Vid16, NV18b], we reduce the number of provers from 3 to 2. Our main result is as follows.

Theorem 1.3 (Main theorem, informal). *Suppose Provers* A *and* B *pass the two-prover. degree-d low individual degree test with probability* $1 - \epsilon$. *Let* $A = \{A_a^u\}$ *be the measurement the provers perform when they are given the point* $u \in \mathbb{F}_q^m$ *to produce a value* $a \in \mathbb{F}_q$. *Then there exists a projective measurement* $G = \{G_g\}$ *whose outcomes* g *are polynomials of individual degree* d *such that*

$$\mathop{\mathbf{E}}_{\boldsymbol{u} \sim \mathbb{F}_q^m} \sum_{a \in \mathbb{F}_q} \sum_{g : g(\boldsymbol{u}) = a} \langle \psi | A_a^{\boldsymbol{u}} \otimes G_g | \psi \rangle \geq 1 - \mathrm{poly}(m) \cdot (\mathrm{poly}(\epsilon) + \mathrm{poly}(d/q)).$$

In other words, if Prover A *measures according to* $A^{\boldsymbol{u}}$ *to produce* $\boldsymbol{a}$ *and Prover* B *measures according to* G *to produce* $\boldsymbol{g}$, *then* $\boldsymbol{g}(\boldsymbol{u}) = \boldsymbol{a}$ *except with probability* $\mathrm{poly}(m) \cdot (\mathrm{poly}(\epsilon) + \mathrm{poly}(d/q))$.

Thus, we are able to extend Theorem 1.2 to the case of two quantum provers (with some minor caveats; see Theorem 3.10 below for the formal statement of Theorem 1.3).

Although Theorem 1.3 establishes quantum soundness of the low-individual degree test and not of the low-degree test, it is still sufficient to recover the result $\mathsf{NEEXP} \subseteq \mathsf{MIP}^*$ from [NW19] and the result $\mathsf{MIP}^* = \mathsf{RE}$ from [JNV+20]. In addition, we can use it to recover the self-test for an exponential number of EPR pairs from [NV18a]. Edited drafts of these work to account for this change are forthcoming. It remains open whether the complexity-theoretic consequences of [Vid16, NV18b, NV18a] to the "scaled down" setting still hold.

1.2   Total degree versus individual degree

We now contrast the low degree test with the individual degree test and explain why we are only able to prove the latter quantum sound. We begin by explaining why the low individual degree test, unlike the low degree test, requires a $\mathrm{poly}(m) \cdot \mathrm{poly}(\epsilon)$ dependence in the soundness error.

Example 1.4. Consider the degree-$(d+1)$ polynomial $h(x_1, \ldots, x_m) = x_1^{d+1}$, and suppose that Players A and B play according to the following classical strategy.

- (Player A): given $\boldsymbol{u} \in \mathbb{F}_q^m$, return the value $\boldsymbol{a} = h(\boldsymbol{u})$.

- (Player B): given the axis parallel line $\boldsymbol{\ell} = \{\boldsymbol{u} + x \cdot e_{\boldsymbol{i}} \mid x \in \mathbb{F}_q\}$, act as follows. If $\boldsymbol{i} > 1$, then h is a constant function along $\boldsymbol{\ell}$, and so return $\boldsymbol{f} = h|_{\boldsymbol{\ell}}$. Otherwise, if $\boldsymbol{i} = 1$, then h is a degree-$(d+1)$ polynomial along $\boldsymbol{\ell}$. As the verifier expects a degree-d polynomial, simply give up and return $\boldsymbol{f} \equiv 0$.

Using this strategy, $\boldsymbol{f}(\boldsymbol{u}) = h|_{\boldsymbol{\ell}}(\boldsymbol{u}) = h(\boldsymbol{u}) = \boldsymbol{a}$ whenever $\boldsymbol{i} \neq 1$, which occurs with probability $1 - \frac{1}{m}$. Hence, the players pass the degree-d low individual degree test with probability at least $1 - \epsilon$ for $\epsilon = \frac{1}{m}$. However, Player A is responding to their questions using the polynomial h which is degree-$(d+1)$ but not degree-d, and so by the aforementioned Schwartz-Zippel lemma, any degree-d polynomial g will agree with h on at most a $\frac{d+1}{q}$ fraction of all inputs. This means that the agreement between Player A's strategy and any degree-d polynomial g is at most

$$\mathop{\mathbf{Pr}}_{\boldsymbol{u} \sim \mathbb{F}_q^m} [g(\boldsymbol{u}) = \boldsymbol{a}] \leq 1 - m \cdot \epsilon + \frac{d+1}{q}.$$

Example 1.4 shows that the dependence on m and ϵ in Theorem 1.3 is tight up to polynomial factors. This reveals a weakness with the low individual degree test: one can only conclude that the players are using a low individual degree strategy when their failure probability ϵ is tiny—on the order of $\frac{1}{m}$ or smaller. The low (total) degree test is alluring because it has the potential to avoid this dependence on m.

Soundness of the low total and individual degree tests is typically proven by induction on m. For the low individual degree test, each step of the induction incurs an error of $\mathrm{poly}(m) \cdot (\mathrm{poly}(\epsilon) + \mathrm{poly}(d/q))$. Summing over all m steps, this gives a total error of $\mathrm{poly}(m) \cdot (\mathrm{poly}(\epsilon) + \mathrm{poly}(d/q))$, exactly as in Theorem 1.3. For the low degree test, on the other hand, each step of the induction only incurs an error of $\mathrm{poly}(\epsilon) + \mathrm{poly}(d/q)$. However, if summed over all m steps, this still gives a total error of $m \cdot (\mathrm{poly}(\epsilon) + \mathrm{poly}(d/q))$, which is too large.

To account for this, the soundness proofs in [Vid16, NV18b] introduce a technique at the end of each induction step called *Consolidation* in which the growing error is "reset" down to an error $\mathrm{poly}(\epsilon) + \mathrm{poly}(d/q)$ which remains fixed across all levels of the induction. This allows them to conclude with an error that was independent of the dimension m. Consolidation works as follows: if the error of the projective measurement $G = \{G_g\}$ ever grows past this fixed error, the Consolidation step argues that on some portion of the Hilbert space, the measurement G must be performing much worse than expected; it then corrects this by inductively calling the low-degree test soundness to produce a better measurement on this portion of the Hilbert space. Ultimately, however, this creates a cascade of Consolidation steps calling each other with increasing error, which at some point grows so large that the low-degree soundness can no longer be applied. This is the source of the bug.

Fortunately, this work shows that the techniques of these prior works, aside from Consolidation, are still sound. So we can show soundness bounds which grow as a function of m, even if we cannot yet show bounds independent of m. In the end, this "weakness" of the low individual degree test is precisely what allows us to show that it, and not the low degree test, is quantum sound.

That said, we do believe, although we have not rigorously verified, that our techniques are capable of proving a result like Theorem 1.3 for the low degree test, i.e. a soundness bound of $\mathrm{poly}(m) \cdot (\mathrm{poly}(\epsilon) + \mathrm{poly}(d/q))$ rather than $\epsilon + \mathrm{poly}(m) \cdot \mathrm{poly}(d/q)$. This "weak" quantum soundness does not rule out the possibility that quantum provers can significantly outperform their classical counterparts. It is, however, still sufficient for the same applications as the low individual degree test, and it hints towards the possibility of a full quantum soundness of the low degree test.

1.3 Conclusion and open problems

In recent years, the classical low-degree test has played a critical role in the study of nonlocal games and the complexity class MIP^*. In addition to correcting previous proofs of soundness, we hope that this new exposition will invite new researchers to engage with this beautiful area. We conclude with a short list of open problems for future work.

1. Is the classical low-degree test quantum sound? Can the Consolidation step be fixed?

2. Even if the classical low-degree test is shown quantum sound, there are still interesting questions to be answered about the low individual degree test. For example, can the diagonal lines test be removed? (See Section 3 for a description of this subtest.) Doing so would likely simplify the proof of $\mathsf{MIP}^* = \mathsf{RE}$ [JNV$^+$20], as one could replace the complicated "conditional linear functions" used in the proof with a simpler subclass known as "coordinate deletion functions". However, as discussed at the end of Section 2, we know of an example that requires the diagonal lines test for the low individual degree test with parameters $m = 2$, $d = 2$, and $q = 4$. Can we find similar examples for larger q?

3. Classically, the low degree test generalized in various directions, including tests for affine invariant properties [KS08, Sud11] and tests for tensor product codes (see, for example, [CMS17]). Does quantum soundness hold for these tests as well?

4. This work continues the trend of showing that well-studied classical property testers are also quantum sound. Prior to this, the work of [IV12] (see also [Vid11, Chapter 2]) showed that the linearity tester of Blum, Luby, and Rubinfield [BLR93] is also quantum sound. The proofs of these results have so far been case-by-case adaptations of the classical proofs to the quantum setting; in the case of this paper, the proof is highly nontrivial and involves many ad hoc calculations which fortunately go in our favor. Is there a more conceptual reason why quantum soundness holds for these testers? Perhaps a reduction from the quantum case to the classical case?

2   Technical overview

At a high level, our proof follows the approach of [BFL91], and it is useful to start by summarizing their analysis, which applies to classical, *deterministic* strategies. In this version of the test, the provers' strategy is described by a "points function," assigning a value in $\mathbb{F}_q$ to each point in $\mathbb{F}_q^m$, and a "lines function," assigning a low-degree polynomial to each line in $\mathbb{F}_q^m$ queried in the test. From the assumption that the points and lines functions agree at a randomly chosen point with high probability, we would like to construct a global low-degree polynomial that has high agreement with the points function. This is done by inductively constructing "subspace functions" defined on affine axis-aligned subspaces of increasing dimension k. The base case is $k = 1$, and is supplied by the lines function. At each step, to construct the subspace function for a subspace S of dimension $k + 1$, we pick $d + 1$ parallel subspaces of dimension k that lie within S, and compute the unique degree-d polynomial that interpolates them. The analysis shows that at each stage, the function constructed through interpolation has high agreement on average with the points function and lines function. In the end, when we reach $k = m$ the ambient dimension, the subspace function we construct is the desired global function.

For the sake of simplicity, it suffices to consider the case $d = 1$, i.e. multilinear functions. In this case, whenever we perform interpolation, we need to combine 2 parallel subspaces.

The classical zero-error case   To start building intuition, it is useful to think about how to carry out the above program in a highly simplified setting: the classical zero-error case, for $m = 3$. In this case, we assume that we have access to a points function f and lines function g that *perfectly* pass the BFL test. Moreover, we will focus on the *final* step of the induction: thus, we assume that we have already constructed a set of planes functions defined for every axis-parallel plane, that are perfectly consistent with the line and point functions. In the final step of the induction, our goal is to combine these planes functions to create a single global function h over all of $\mathbb{F}_q^3$ that is consistent with the points and lines functions.

To do this, we will interpolate the planes as follows. Let us label the 3 coordinates in the space x, y,and z, and consider planes parallel to the (x, y)-plane. Each plane S_z is specified by a value of the z coordinate:

$$S_z = \{(x, y, z) : (x, y) \in \mathbb{F}_q^2\}.$$

By the induction hypothesis, for every such plane S_z there exists a bilinear function $g_z : \mathbb{F}_q^2 \to \mathbb{F}_q$ that agrees with the points function. To construct a global multilinear function $h : \mathbb{F}_q^3 \to \mathbb{F}_q$, we pick two distinct values $z_1 \neq z_2$, and "paste" the two plane functions g_{z_1} and g_{z_2} together using *polynomial interpolation*. Specifically, we define h to be the unique multilinear polynomial that interpolates between g_{z_1} on the plane S_{z_1} and g_{z_2} on the plane S_{z_2}.

$$h(x, y, z) = \text{interpolate}_{z_1, z_2}(g_{z_1}, g_{z_2}) = \left(\frac{z - z_2}{z_1 - z_2}\right) g_{z_1}(x, y) + \left(\frac{z - z_1}{z_2 - z_1}\right) g_{z_2}(x, y).$$

This procedure defines a global function h, and by the assumption that g_{z_1} and g_{z_2} are multilinear, it follows that h is also a multilinear function. But why is h consistent with the points function? To show this, we need to consider the lines function on lines parallel to the z axis. Given a point (x, y, z), let ℓ be the line parallel to the z axis through this point, and let g_ℓ be the associated lines function. By construction, h agrees with g_ℓ at the two points z_1 and z_2. But g_ℓ and the restriction $h_{|\ell}$ of h to ℓ are both linear functions, and hence if they agree at two points, they must agree *everywhere*. (This is a special case of the *Schwartz-Zippel lemma*, which says that if two degree-d polynomials agree at $d + 1$ points, they must be equal.) Thus, h agrees with g_ℓ at the original point z as well. By success in the test, g_ℓ in turn agrees with the points function f at (x, y, z), and thus, $h(x, y, z) = f(x, y, z)$. Thus, we have shown that the global function h is both multilinear and agrees with the points function f exactly.

Dealing with errors To extend the sketch above to the general case, with nonzero error, requries some modifications. At the most basic level, we may consider what happens when we allow for deterministic classical strategies that succeed with probability less than 1 in the test. Such strategies may have "mislabeling" error: the points function f may be imagined to be a multilinear function that has been corrupted at a small fraction of the points. This type of error is handled by the analysis in [BFL91]. The main modification to the zero-error sketch above is a careful analysis of the probability that the pasting step produces a "good" interpolated polynomial for a randomly chosen pair of planes S_{z_1}, S_{z_2}. This analysis makes use of the Schwartz-Zippel lemma together with combinatorial properties of the point-line test itself (e.g. the expansion properties of the question graph associated with the test).

At the next level of generality, we could consider classical *randomized* strategies. Suppose we are given a randomized strategy that succeeds in the test with probability $1 - \epsilon$. Any randomized strategy can be modeled by first sampling a random seed, and then playing a deterministic strategy conditioned on the value of the seed. A success probability of $1 - \epsilon$ could have two qualitatively different underlying causes: (1) on $O(\epsilon)$ fraction of the seeds, the strategy uses a function which is totally corrupted, and (2) on a large fraction of the seeds, the strategy uses functions which are only ϵ-corrupted. An analysis of randomized strategies could naturally proceed in a "seed-by-seed" fashion, applying the deterministic analysis of [BFL91] to the large fraction of "good" seeds (which are each only ϵ-corrupted), while giving up entirely on the "bad" seeds.

In this document, we consider quantum strategies, which have much richer possibilities for error. Nevertheless, we are able to preserve some intuition from the randomized case, by working with *sub-measurements*: quantum measurements that do not always yield an outcome. Working with a sub-measurement allows us to distinguish two kinds of error: *consistency error* (the probability that a sub-measurement returns a wrong outcome) and *completeness error* (the probability that the sub-measurement fails to return an outcome at all). Roughly speaking, the completeness error corresponds to the probability of obtaining a "bad" seed in the randomized case, while the consistency error corresponds to how well the strategies do on "good" seeds.

The technique of using sub-measurements and managing the two types of error separately goes back to [IV12]. That work developed a crucial tool to convert between these two types of error called the *self-improvement lemma*, and in our analysis we make extensive use of a refined version of this lemma (Theorem 6.2), building on [Vid16, NV18b]. Essentially, the lemma says the following: suppose that (a) the provers pass the test with probability $1 - \epsilon$, and (b) there is a complete measurement G_g whose outcomes are low-degree polynomials that has consistency error ν: that is, G always returns an outcome, but has probability ν of producing an outcome g that disagrees with the points measurement A_a^u at a random point u. Then there exists an "improved" sub-

measurement H_h with consistency error ζ depending *only* on ϵ, and with completeness error (i.e. probability of not producing an outcome at all) of $\nu + \zeta$. Essentially, this lemma says we can always "reset" the consistency error of any measurement we construct at intermediate points in the analysis to a universal function ζ depending only on the provers' success in the test, at the cost of introducing some amount completeness error. Intuitively, one may think of the action of the lemma as correcting G on the portions of Hilbert space where it is only mildly corrupted, while "cutting out" the portions of Hilbert space where G is too corrupted to be correctable. In some sense, this lemma is the quantum analog of the idea of identifying "good" and "bad" random seeds in the classical randomized case. The proof of the lemma uses a semidefinite program together with the combinatorial facts used in [BFL91] (specifically, expansion of the question graph of the test, and the Schwartz-Zippel lemma).

Armed with the self-improvement lemma, we set up the following quantum version of the BFL induction loop: for k running from 1 to m, we construct a family of subspace measurements G_g^S that returns a low-degree polynomial g for every axis-aligned affine subspace S of dimension k.

1. By the induction hypothesis, we know that there exists a measurement G_g^S for every k-dimensional subspace, which has consistency error $\delta(k)$ with the points measurement. (For the base case $k = 1$, this is the lines measurement from the provers' strategy).

2. We apply the self-improvement lemma to these measurements, yielding sub-measurements $\hat{G}_g^S$ that have consistency error ζ independent of $\delta(k)$, and completeness error $\kappa(k) = \delta(k) + \zeta$.

3. For each subspace S of dimension $k + 1$, we construct a pasted sub-measurement, by performing a quantum version of the classical interpolation argument: we define the pasted sub-measurement by sequentially measuring several subspace measurements corresponding to parallel k-dimensional subspaces, and interpolate the resulting outcomes. This pasted sub-measurement has consistency slightly worse than ζ, and completeness error which is slightly worse than $\kappa(k)$. It is at this step that it is crucial to treat the two types of error separately: in particular, we need the consistency error to be low to ensure that the interpolation produces a good result.

4. We convert the resulting sub-measurement into a full measurement, by assigning a random outcome whenever the sub-measurement fails to yield an outcome. This measurement will have consistency error $\delta(k + 1)$ which is larger than $\delta(k)$ by some additive factor.

At the end of the loop, when $k = m$, we obtain a single measurement that returns a global polynomial as desired.

The diagonal lines test An important element of the test we analyze in this document, which is unnecessary in the classical case, is the diagonal lines test. The purpose of this test is to certify that the points measurements used by the provers approximately commute on average over all pairs of points (x, y) in $\mathbb{F}_q^m$—something which is automatically true in the classical case. This is done by asking one prover for a polynomial defined on the line going through x and y, and the other for the function evaluation at either x or y. The line through x and y will not in general be axis-parallel; we refer to these general lines as "diagonal."

The commutation guarantee plays an important role in our analysis of the test, and it is an interesting question whether it is truly necessary to test it directly with the diagonal lines test; might it not automatically follow from success in the axis-parallel lines test? An interesting contrast can be drawn to the Magic Square game [Mer90, Per90, Ara02], in which questions are either cells or

axis-parallel lines in a 3×3 square grid. This has the same question distribution as the axis-parallel line-point test over $\mathbb{F}_3^2$ (although the answers in the Magic Square game are strings in $\mathbb{F}_2$ rather than $\mathbb{F}_3$). For the Magic Square game, "points" measurements along the same axis-parallel "line" commute, but points that are not axis-aligned do *not* commute: indeed, for the perfect strategy, they anticommute. Despite this example, we know that commutation between all pairs of points *can* be deduced from the axis-parallel lines test alone, rending the diagonal lines test unnecessary, at least in the case of the bivariate ($m = 2$) multilinearity test (the low individual degree test when $d = 1$). On the other hand, we know of a quantum strategy using noncommuting measurements which succeeds with probability 1 in the $m = 2$, $d = 2$, $q = 4$ low individual degree test. Whether this counterexample can be extended to larger q remains an open question.

Organization   The rest of this document is organized as follows. In Section 4 we review some preliminaries concerning finite fields, polynomials, and quantum measurements. In Section 5, we present two tools for making quantum measurements projective, which are used in our analysis. In Section 6, we give the inductive argument proving the main theorem. In Sections 7 and 8 we show some properties of the hypercube graph and families of measurements indexed by points on the hypercube, which are used in Section 9 to prove the self-improvement lemma. In Sections 10 and 11, we show commutativity properties of the measurements constructed in the induction, and finally in Section 12 we analyze the pasting step of the induction.

3   The test

Definition 3.1 (Roles)**.** The low individual degree test will be played between two provers and a verifier. The two provers are named Player A and Player B. A *role r* is an element of the set $\{A, B\}$. Given a role r, we write $\bar{r}$ for the other element of the set $\{A, B\}$.

Definition 3.2 (Low individual degree test)**.** Let m and d be nonnegative integers. Let q be a prime power. Then the (m, q, d)-*low individual degree test* is stated in Figure 2.

An important class of strategies are those which are *symmetric*. In this case, the bipartite state $|\psi\rangle$ Alice and Bob share is symmetric. Furthermore, for any question Alice and Bob receive, they apply the same measurement to their share of the state. This means that rather than, for example, keeping track of a separate points measurement for Alice and Bob, we can use a single measurement to refer to both of their strategies, and similarly for the axis-parallel lines and diagonal lines measurements. This is formalized in the following definition, where we also consider strategies which are projective.

Definition 3.3 (Symmetric, projective strategy)**.** A *symmetric, projective strategy* for the (m, q, d)-low individual degree test is a tuple (ψ, A, B, L) defined as follows.

- $|\psi\rangle \in \mathcal{H} \otimes \mathcal{H}$ is a bipartite, permutation-invariant state.

- $A = \{A_a^u\}$ contains a matrix for each $u \in \mathbb{F}_q^m$ and $a \in \mathbb{F}_q$. For each u, A^u is a projective measurement on $\mathcal{H}$.

With probability $\frac{1}{3}$ each, perform one of the following three tests.

1. **Axis-parallel lines test:** Pick a uniformly random role $\boldsymbol{r} \sim \{A, B\}$. Let $\boldsymbol{u} \sim \mathbb{F}_q^m$ be a uniformly random point. Select $\boldsymbol{i} \sim \{1, \ldots, m\}$ uniformly at random, and let $\boldsymbol{\ell} = \{\boldsymbol{u} + t \cdot e_{\boldsymbol{i}} \mid t \in \mathbb{F}_q\}$ be the axis-parallel line which passes through $\boldsymbol{u}$ in the $\boldsymbol{i}$-th direction.

   ○ Player $\boldsymbol{r}$: Give $\boldsymbol{\ell}$; receive the univariate degree-d polynomial $\boldsymbol{f} : \boldsymbol{\ell} \to \mathbb{F}_q$.
   ○ Player $\overline{\boldsymbol{r}}$: Give $\boldsymbol{u}$; receive $\boldsymbol{a} \in \mathbb{F}_q$.

   Accept if $\boldsymbol{f}(\boldsymbol{u}) = \boldsymbol{a}$.

2. **Self-consistency test:** Let $\boldsymbol{u} \sim \mathbb{F}_q^m$ be a uniformly random point.

   ○ Player A: Give $\boldsymbol{u}$; receive $\boldsymbol{a} \in \mathbb{F}_q$.
   ○ Player B: Give $\boldsymbol{u}$; receive $\boldsymbol{b} \in \mathbb{F}_q$.

   Accept if $\boldsymbol{a} = \boldsymbol{b}$.

3. **Diagonal lines test:** Pick a uniformly random role $\boldsymbol{r} \sim \{A, B\}$. Let $\boldsymbol{u} \sim \mathbb{F}_q^m$ be a uniformly random point. Select $\boldsymbol{i} \sim \{1, \ldots, m\}$ uniformly at random, and let $\boldsymbol{v} \in \mathbb{F}_q^m$ be a uniformly random point whose last $m - \boldsymbol{i}$ coordinates are 0. Finally, let $\boldsymbol{\ell} = \{\boldsymbol{u} + t \cdot \boldsymbol{v} \mid t \in \mathbb{F}_q\}$ be the line which passes through $\boldsymbol{u}$ in direction $\boldsymbol{v}$.

   ○ Player $\boldsymbol{r}$: Give $\boldsymbol{\ell}$; receive the univariate degree-md polynomial $\boldsymbol{f} : \boldsymbol{\ell} \to \mathbb{F}_q$.
   ○ Player $\overline{\boldsymbol{r}}$: Give $\boldsymbol{u}$; receive $\boldsymbol{a} \in \mathbb{F}_q$.

   Accept if $\boldsymbol{f}(\boldsymbol{u}) = \boldsymbol{a}$.

Figure 2: The (m, q, d)-low individual degree test.

○ $B = \{B_f^\ell\}$ contains a matrix for each axis-parallel line ℓ in $\mathbb{F}_q^m$ and univariate degree-d polynomial $f : \ell \to \mathbb{F}_q$. For each ℓ, B^ℓ is a projective measurement on $\mathcal{H}$.

○ $L = \{L_f^\ell\}$ contains a matrix for each line ℓ in $\mathbb{F}_q^m$ and univariate degree-md polynomial $f : \ell \to \mathbb{F}_q$. For each ℓ, L^ℓ is a projective measurement on $\mathcal{H}$.

We can also consider more general strategies which are no longer assumed to be symmetric. In this case, Players A and B each have their own versions of the measurements A, B, and L.

Definition 3.4 (General projective strategy). A *projective strategy* for the (m, q, d)-low individual degree test is a tuple $(\psi, A^{\mathrm{A}}, B^{\mathrm{A}}, L^{\mathrm{A}}, A^{\mathrm{B}}, B^{\mathrm{B}}, L^{\mathrm{B}})$ defined as follows.

○ $|\psi\rangle \in \mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}$ is a bipartite state.

Furthermore, for each $w \in \{\mathrm{A}, \mathrm{B}\}$:

○ $A^w = \{A_a^{w,u}\}$ contains a matrix for each $u \in \mathbb{F}_q^m$ and $a \in \mathbb{F}_q$. For each u, $A^{w,u}$ is a projective measurement on $\mathcal{H}_w$.

○ $B^w = \{B_f^{w,\ell}\}$ contains a matrix for each axis-parallel line ℓ in $\mathbb{F}_q^m$ and univariate degree-d polynomial $f : \ell \to \mathbb{F}_q$. For each ℓ, $B^{w,\ell}$ is a projective measurement on $\mathcal{H}_w$.

○ $L^w = \{L_f^{w,\ell}\}$ contains a matrix for each line ℓ in $\mathbb{F}_q^m$ and univariate degree-md polynomial $f : \ell \to \mathbb{F}_q$. For each ℓ, $L^{w,\ell}$ is a projective measurement on $\mathcal{H}_w$.

Remark 3.5. Throughout this work, we will only consider projective strategies. So we will henceforth use the terms "strategy" and "symmetric strategy" to refer exclusively to projective strategies and symmetric, projective strategies, respectively.

In addition, we will spend the vast majority of this work dealing solely with symmetric strategies, as they are notationally simpler to work with. Much of this work will focus on proving Theorem 6.1, a variant of our main theorem for symmetric strategies. Our main theorem for general strategies, Theorem 3.10 below, will be proven in Section 6 by a standard reduction to the symmetric case. We will only see these more cumbersome-to-write general strategies in this section, where we state Theorem 3.10, and in Section 6, where we carry out the reduction.

Definition 3.6 (Good strategy). A strategy is $(\epsilon, \delta, \gamma)$-good if it passes the axis-parallel lines test with probability at least $1 - \epsilon$, the self consistency test with probability at least $1 - \delta$, and the diagonal lines test with probability at least $1 - \gamma$.

Remark 3.7. Using notation which will be introduced in Section 4.4 below, a symmetric strategy is $(\epsilon, \delta, \gamma)$-good if and only if it satisfies the following three conditions. For $\boldsymbol{\ell}$ and $\boldsymbol{u}$ as in the axis-parallel lines test,

$$A_a^u \otimes I \simeq_\epsilon I \otimes B_{[f(u)=a]}^\ell, \quad \text{and} \quad A_a^u \otimes I \simeq_\delta I \otimes A_a^u.$$

And for $\boldsymbol{\ell}$ and $\boldsymbol{u}$ as in the diagonal lines test,

$$A_a^u \otimes I \simeq_\gamma I \otimes L_{[f(u)=a]}^\ell.$$

Notation 3.8. Our proof will be via induction, i.e. proving soundness of the $(m + 1, q, d)$-low individual degree test using the soundness of the (m, q, d)-low individual degree test. To do this, we will frequently use the axis-parallel line test in the specific case of $i = m + 1$. Thus, it will be

14

convenient to introduce the following notation. Let (ψ, A, B, L) be a symmetric strategy for the $(m + 1, q, d)$-low individual degree test, Then for each $u \in \mathbb{F}_q^m$ we will write B_f^u as shorthand for B_f^ℓ, where $\ell = \{(u, x) \mid x \in \mathbb{F}_q\}$. For a function $f : \ell \to \mathbb{F}_q$, we will also sometimes write $f(x)$ as shorthand for $f(u, x)$.

Definition 3.9. Consider the (m, q, d)-low individual degree test. For $j \in \{1, \ldots, m\}$, we refer to the *j-restricted diagonal lines test* as the diagonal lines test conditioned on $\boldsymbol{i} = j$. For example, in the m-restricted diagonal lines test, the line $\boldsymbol{\ell}$ is simply a uniformly random line in $\mathbb{F}_q^m$.

Now we state our main theorem using notation which will be introduced in Section 4 below. This the formal version of Theorem 1.3.

Theorem 3.10 (Main theorem; quantum soundness of the low individual degree test). *Consider a projective strategy $(\psi, A^{\mathrm{A}}, B^{\mathrm{A}}, L^{\mathrm{A}}, A^{\mathrm{B}}, B^{\mathrm{B}}, L^{\mathrm{B}})$ which passes the (m, q, d)-low individual degree test with probability at least $1 - \epsilon$. Let $k \geq md$ be an integer. Let*

$$\nu = 100000 k^2 m^4 \cdot \left(\epsilon^{1/40000} + (d/q)^{1/40000} + e^{-k/(2560000 m^2)} \right).$$

Then there exists projective measurements $G^{\mathrm{A}}, G^{\mathrm{B}} \in \mathrm{PolyMeas}(m, q, d)$ with the following properties:

1. *(Consistency with A): On average over $\boldsymbol{u} \sim \mathbb{F}_q^m$,*

$$A_a^{\mathrm{A}, u} \otimes I \simeq_\nu I \otimes G_{[g(u)=a]}^{\mathrm{B}},$$
$$I \otimes A_a^{\mathrm{B}, u} \simeq_\nu G_{[g(u)=a]}^{\mathrm{A}} \otimes I.$$

2. *(Self-consistency):*

$$G_g^{\mathrm{A}} \otimes I \simeq_\nu I \otimes G_g^{\mathrm{B}}.$$

We note that there is a tradeoff in Theorem 3.10 specified by the parameter k. As k increases, ν's prefactor k^2 increases. On the other hand, as k increases, the term $e^{-k/(2560000 m^2)}$ which also occurs in ν *decreases*. Thus, when applying this theorem, one must select k to balance these competing demands. Typically, choosing $k = \mathrm{poly}(m)$ should be more than sufficient for applications. We believe that this parameter k is an artifact of our proof, and we hope that it will be removed in the future.

4 Preliminaries

We use **boldface font** to denote random variables. For two complex numbers $\alpha, \beta \in \mathbb{C}$, we write $\alpha \approx_\epsilon \beta$ if

$$|\alpha - \beta| \leq \epsilon.$$

We note the following triangle inequality for numbers, which we will use repeatedly:

$$\text{if } \alpha \approx_\epsilon \beta \text{ and } \beta \approx_\delta \gamma, \text{ then } \alpha \approx_{\epsilon+\delta} \gamma. \tag{1}$$

4.1 Finite fields

A *finite field* is a field with a finite number of elements. There is a unique finite field $\mathbb{F}_q$ of q elements for each prime power $q = p^t$, and there are no other finite fields. We write ω for the p-th root of unity $\omega = e^{2\pi i/p}$.

Definition 4.1 (Finite field trace). The *finite field trace* is the function $\mathrm{tr} : \mathbb{F}_q \to \mathbb{F}_p$ defined as

$$\mathrm{tr}[x] = \sum_{\ell=0}^{t-1} x^{p^\ell}.$$

Proposition 4.2. *Let $a \in \mathbb{F}_q$. Then*

$$\mathop{\mathbf{E}}_{x \sim \mathbb{F}_q} \omega^{\mathrm{tr}[x \cdot a]} = \begin{cases} 1 & \text{if } a = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. If $a = 0$, then $\mathrm{tr}[x \cdot a] = 0$ for all $x \in \mathbb{F}_q$. As a result,

$$\mathop{\mathbf{E}}_{x \sim \mathbb{F}_q} \omega^{\mathrm{tr}[x \cdot a]} = \mathop{\mathbf{E}}_{x \sim \mathbb{F}_q} \omega^0 = \mathop{\mathbf{E}}_{x \sim \mathbb{F}_q} 1 = 1.$$

On the other hand, if $a \neq 0$, then there exists a $y \in \mathbb{F}_q$ such that $\mathrm{tr}[a \cdot y] \neq 0$. As a result,

$$C := \mathop{\mathbf{E}}_{x \sim \mathbb{F}_q} \omega^{\mathrm{tr}[x \cdot a]} = \mathop{\mathbf{E}}_{x \sim \mathbb{F}_q} \omega^{\mathrm{tr}[(x+y) \cdot a]} = \mathop{\mathbf{E}}_{x \sim \mathbb{F}_q} (\omega^{\mathrm{tr}[x \cdot a]} \cdot \omega^{\mathrm{tr}[y \cdot a]}) = C \cdot \omega^{\mathrm{tr}[y \cdot a]}.$$

But because $\mathrm{tr}[y \cdot a] \neq 0$, $\omega^{\mathrm{tr}[y \cdot a]} \neq 1$. This implies that $C = 0$. □

Proposition 4.3. *Let $v \in \mathbb{F}_q^m$. Then*

$$\mathop{\mathbf{E}}_{u \sim \mathbb{F}_q^m} \omega^{\mathrm{tr}[u \cdot v]} = \begin{cases} 1 & \text{if } v = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. By the linearity of the trace,

$$\mathop{\mathbf{E}}_{u \sim \mathbb{F}_q^m} \omega^{\mathrm{tr}[u \cdot v]} = \mathop{\mathbf{E}}_{u \sim \mathbb{F}_q^m} (\omega^{\mathrm{tr}[u_1 \cdot v_1]} \cdots \omega^{\mathrm{tr}[u_m \cdot v_m]}) = \prod_{i=1}^{m} (\omega^{\mathrm{tr}[u_i \cdot v_i]}),$$

which by Proposition 4.2 is 1 if $v_i = 0$ for all i and 0 otherwise. □

4.2 Polynomials over finite fields

Definition 4.4 (Polynomials of low individual degree). Let q be a prime power, and let m and d be nonnegative integers. We define $\mathcal{P}(m, q, d)$ to be the set of polynomials in $\mathbb{F}_q^m$ with individual degree d.

Remark 4.5. Note that as defined in Section 1, in this work a polynomial with individual degree d is one in which the degree d_i of each coordinate i is *at most* d. This allows us to say "individual degree d" rather than the wordier "individual degree at most d". For example, under this definition, $\mathcal{P}(m, q, d)$ is contained in $\mathcal{P}(m, q, d+1)$ for each d.

The most important fact about low-degree polynomials is that they have large distance from each other. This is shown by the following lemma.

Lemma 4.6 (Schwartz-Zippel lemma [Sch80, Zip79]). *Let $g, h : \mathbb{F}_q^m \to \mathbb{F}_q$ be two distinct polynomials of total degree d. Then*

$$\Pr_{\boldsymbol{x} \sim \mathbb{F}_q^m} [g(\boldsymbol{x}) = h(\boldsymbol{x})] \leq \frac{d}{q}.$$

Since any polynomial with individual degree d has total degree md, Lemma 4.6 implies the following corollary.

Corollary 4.7 (Schwartz-Zippel for individual degree). *Let $g, h \in \mathcal{P}(m, q, d)$ be distinct. Then*

$$\Pr_{\boldsymbol{x} \sim \mathbb{F}_q^m} [g(\boldsymbol{x}) = h(\boldsymbol{x})] \leq \frac{md}{q}.$$

4.3 Measurements

Definition 4.8 (Measurements and sub-measurements). Let $\mathcal{H}$ be a Hilbert space and $\mathcal{A}$ be a set of outcomes. A *sub-measurement* is a set of Hermitian, positive-semidefinite matrices $A = \{A_a\}_{a \in \mathcal{A}}$ acting on $\mathcal{H}$ such that $\sum_a A_a \leq I$. The sub-measurement is *projective* if $(A_a)^2 = A_a$ for each a. It is a *measurement* if it satisfies the stronger condition $\sum_a A_a = I$.

An important class of sub-measurements are those that output polynomials g of individual degree d. These are defined as follows.

Definition 4.9 (Low-degree polynomial measurements). We write $\mathrm{PolySub}(m, q, d)$ for the set of sub-measurements $G = \{G_g\}$ with outcomes $g \in \mathcal{P}(m, q, d)$. We write $\mathrm{PolyMeas}(m, q, d)$ for the subset of $\mathrm{PolySub}(m, q, d)$ containing only those $G = \{G_g\}$ which are measurements.

Definition 4.10 (Post-processing measurements). Let $A = \{A_a\}_{a \in \mathcal{A}}$ be a set of matrices, and let $f : \mathcal{A} \to \mathcal{B}$ be a function. Then for each $b \in \mathcal{B}$, we define the matrix

$$A_{[f(a)=b]} = \sum_{a : f(a) = b} A_a.$$

Remark 4.11. We note that Definition 4.10 agrees with the notation for post-processing measurements used in [NW19], but it disagrees with the notation used in [JNV+20]. That work uses the notation "$A_{[f(\cdot)=b]}$" rather than the notation "$A_{[f(a)=b]}$" that we use in this work.

An easy-to-prove fact is that measurements remain measurements after post-processing their outcomes.

Proposition 4.12. *Let $A = \{A_a\}_{a \in \mathcal{A}}$ be a set of matrices, and let $f : \mathcal{A} \to \mathcal{B}$ be a function. Then*

$$\sum_a A_a = \sum_b A_{[f(a)=b]}.$$

Thus, if $\{A_a\}$ is a sub-measurement (respectively, measurement), then $\{A_{[f(a)=b]}\}$ is also a sub-measurement (respectively, measurement).

Remark 4.13. We note that there is some ambiguity in the notation $A_{[f(a)=b]}$ because it requires one to know which of f or a is the measurement outcome and which is the function applied to it. For example, given a sub-measurement $G = \{G_g\} \in \mathrm{PolySub}(m, q, d)$, we will often consider evaluating its outputs at a point $u \in \mathbb{F}_q^m$. This entails looking at the sub-measurement

$$\{G_{[g(u)=a]}\}_{a \in \mathbb{F}_q}, \quad \text{where} \quad G_{[g(u)=a]} = \sum_{g : g(u) = a} G_g.$$

17

Here, g is the measurement outcome of G, and the function being applied maps it to its evaluation on the point u, i.e. $g(u)$. In general, it should always be clear from context what the measurement outcome and the function being applied to it are.

Notation 4.14 (Measurements indexed by questions). We will frequently encounter sets of sub-measurements $A^x = \{A_a^x\}$ indexed by elements x from a set of "questions" $\mathcal{X}$. We will write $A = \{A_a^x\}$ for this set. We will typically refer to this set A as a sub-measurement, and we will refer to it as a measurement if each A^x is a measurement. In addition, we will refer to it as projective if each A^x is projective.

Notation 4.15 (Complete part of sub-measurement). Given a sub-measurement $A = \{A_a\}$, we will write $A = \sum_a A_a$. Similarly, given a sub-measurement $A = \{A_a^x\}$, we will write $A^x = \sum_a A_a^x$ and $A = \mathbf{E}_{\boldsymbol{x}} A^{\boldsymbol{x}}$.

The complete part of a sub-measurement contrasts with its *incomplete part*, which is the matrix $I - A^x$. We will sometimes view A as a measurement by throwing in its incomplete part as an additional POVM element. This is formalized in the following definition.

Definition 4.16 (Completing a sub-measurement). Let $A = \{A_a^x\}_{a \in \mathcal{A}}$ be a sub-measurement. We define the *completion of A*, denoted completion(A), to be the measurement $\widehat{A} = \{\widehat{A}_a^x\}$ with outcome set $\widehat{\mathcal{A}} = \mathcal{A} \cup \{\perp\}$ such that for each $a \in \widehat{\mathcal{A}}$,

$$A_a^x = \left\{ \begin{array}{ll} A_a^x & \text{if } a \in \mathcal{A}, \\ I - A^x & \text{if } a = \perp. \end{array} \right.$$

4.4 Comparing measurements

An central problem in this paper is recognizing when two measurements are close to each other. We will survey two methods of doing so, using the *consistency* and the *state dependent distance*. This section largely mirrors Sections 4.4 and 4.5 of [NW19], which prove numerous properties of these two distances. However, we are unable to cite their results directly because they are mostly stated and proven only for measurements, whereas we will need to apply them to sub-measurements as well.

4.4.1 Consistency between measurements

The most basic notion of similarity between two measurements is given by their *consistency*.

Definition 4.17 (Consistency). Let $|\psi\rangle$ be a state in $\mathcal{H}_A \otimes \mathcal{H}_B$. Let $A = \{A_a^x\}$ be a sub-measurement acting on $\mathcal{H}_A$ and $B = \{B_a^x\}$ be a sub-measurement acting on $\mathcal{H}_B$. Finally, let $\mathcal{D}$ be a distribution on the question set $\mathcal{X}$. Then we say that

$$A_a^x \otimes I \simeq_\delta I \otimes B_a^x$$

on state $|\psi\rangle$ and distribution $\mathcal{D}$ if

$$\mathop{\mathbf{E}}_{\boldsymbol{x} \sim \mathcal{D}} \sum_{a \neq b} \langle \psi | A_a^{\boldsymbol{x}} \otimes B_b^{\boldsymbol{x}} | \psi \rangle \leq \delta. \tag{2}$$

This is simply the probability that the provers receive different outcomes when they measure with A and B, assuming the sub-measurements do return an outcome.

Notation 4.18 (Simplifying notation)**.** Because the state $|\psi\rangle$ and distribution $\mathcal{D}$ are typically clear from context, we will often write "$A_a^x \otimes I \simeq_\delta I \otimes B_a^x$" as shorthand for "$A_a^x \otimes I \simeq_\delta I \otimes B_a^x$ on state $|\psi\rangle$ and distribution $\mathcal{D}$". In the case when $\mathcal{D}$ is not clear by context, we might specify it implicitly in terms of a random variable $\boldsymbol{x}$ distributed according to $\mathcal{D}$. For example, suppose the distribution $\mathcal{D}$ is supposed to be uniform on the question set $\mathbb{F}_q$. Then we might say "on average over $\boldsymbol{x} \sim \mathbb{F}_q$,

$$A_a^x \otimes I \simeq_\delta I \otimes B_a^{x}"$$

as shorthand for "$A_a^x \otimes I \simeq_\delta I \otimes B_a^x$ on distribution $\mathcal{D}$".

We note that there are two differences between the definition of consistency in Definition 4.17 and the original definition of consistency given in [NW19, Definition 4.11]. The first of these is that the [NW19] definition allows the right-hand side of Equation (2) to be $O(\delta)$ rather than strictly δ. The benefit of this is that they do not need to keep track of constant prefactors in their proofs; we have elected to use this more concrete definition to make our proofs more easily verifiable, at the expense of tracking these constant prefactors. The second difference is that is that they define their consistency as the quantity

$$1 - \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes B_a^{\boldsymbol{x}} |\psi\rangle \,.$$

As we show below in Proposition 4.19, this agrees with Definition 4.17 when A and B are both measurements. However, these two definitions disagree when A and B are *sub*-measurements, which is a case we will frequently encounter throughout this paper.

Proposition 4.19 (Consistency for measurements)**.** *Let* $A = \{A_a^x\}$ *and* $B = \{B_a^x\}$ *be two measurements. Then*

$$A_a^x \otimes I \simeq_\delta I \otimes B_a^x$$

if and only if

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes B_a^{\boldsymbol{x}} |\psi\rangle \geq 1 - \delta.$$

Proof. We compute

$$
\begin{aligned}
& \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{a \neq b} \langle\psi| A_a^{\boldsymbol{x}} \otimes B_b^{\boldsymbol{x}} |\psi\rangle \\
={}& \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_b \langle\psi| \Big(\sum_{a \neq b} A_a^{\boldsymbol{x}} \Big) \otimes B_b^{\boldsymbol{x}} |\psi\rangle \\
={}& \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_b \langle\psi| (I - A_b^{\boldsymbol{x}}) \otimes B_b^{\boldsymbol{x}} |\psi\rangle && \text{(because } A \text{ is a measurement)} \\
={}& \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_b \langle\psi| I \otimes B_b^{\boldsymbol{x}} |\psi\rangle - \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_b \langle\psi| A_b^{\boldsymbol{x}} \otimes B_b^{\boldsymbol{x}} |\psi\rangle \\
={}& 1 - \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_b \langle\psi| A_b^{\boldsymbol{x}} \otimes B_b^{\boldsymbol{x}} |\psi\rangle \,. && \text{(because } B \text{ is a measurement)}
\end{aligned}
$$

Hence, the first expression is at most δ if and only if the last one is as well, and we are done. $\square$

4.4.2 The state-dependent distance

Suppose we have three measurements $\{A_a^x\}$, $\{B_a^x\}$, and $\{C_a^x\}$, and we know that

$$A_a^x \otimes I \simeq_\epsilon I \otimes C_a^x.$$

What property of A and B allows us to conclude that

$$B_a^x \otimes I \simeq_\epsilon I \otimes C_a^x? \tag{3}$$

The answer is provided by the *state-dependent distance.*

Definition 4.20 (The state-dependent distance)**.** Let $|\psi\rangle$ be a state in $\mathcal{H}$. Let $A = \{A_a^x\}$ and $B = \{B_a^x\}$ be sets of matrices acting on $\mathcal{H}$. Finally, let $\mathcal{D}$ be a distribution on the question set $\mathcal{X}$. Then we say that

$$A_a^x \approx_\delta B_a^x$$

on state $|\psi\rangle$ and distribution $\mathcal{D}$ if

$$\mathop{\mathbf{E}}_{\boldsymbol{x}\sim\mathcal{D}} \sum_a \|(A_a^{\boldsymbol{x}} - B_a^{\boldsymbol{x}})\,|\psi\rangle\,\|^2 \le \delta.$$

We note one odd feature of Definition 4.20 in comparison to the consistency, which is that $|\psi\rangle$ is not assumed to have a bipartition $\mathcal{H}_A \otimes \mathcal{H}_B$ in which A and B are applied on opposite sides. We will address this in Section 4.4.3 below. Before doing so, we will answer our question above, even in the case when C is allowed to be a sub-measurement.

Proposition 4.21 (Transfering "$\simeq$" using "$\approx$")**.** *Let $\{A_a^x\}$ and $\{B_a^x\}$ be measurements, and let $\{C_a^x\}$ be a sub-measurement. Suppose that $A_a^x \otimes I \simeq_\delta I \otimes C_a^x$ and $A_a^x \otimes I \approx_\epsilon B_a^x \otimes I$. Then $B_a^x \otimes I \simeq_{\delta+\sqrt{\epsilon}} I \otimes C_a^x$.*

Proof. First, we can rewrite the consistency between A and C as

$$\delta \ge \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{a \ne b} \langle\psi|\, A_a^{\boldsymbol{x}} \otimes C_b^{\boldsymbol{x}}\, |\psi\rangle = \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi|\, A_a^{\boldsymbol{x}} \otimes (C^{\boldsymbol{x}} - C_a^{\boldsymbol{x}})\, |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi|\, A_a^{\boldsymbol{x}} \otimes C^{\boldsymbol{x}}\, |\psi\rangle - \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi|\, A_a^{\boldsymbol{x}} \otimes C_a^{\boldsymbol{x}}\, |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}} \langle\psi|\, I \otimes C^{\boldsymbol{x}}\, |\psi\rangle - \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi|\, A_a^{\boldsymbol{x}} \otimes C_a^{\boldsymbol{x}}\, |\psi\rangle$$

$$\text{(because } A \text{ is a measurement)}$$

$$= \langle\psi|\, I \otimes C\, |\psi\rangle - \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi|\, A_a^{\boldsymbol{x}} \otimes C_a^{\boldsymbol{x}}\, |\psi\rangle\,.$$

Likewise, we can rewrite the inconsistency between B and C as

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{a \ne b} \langle\psi|\, B_a^{\boldsymbol{x}} \otimes C_b^{\boldsymbol{x}}\, |\psi\rangle = \langle\psi|\, I \otimes C\, |\psi\rangle - \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi|\, B_a^{\boldsymbol{x}} \otimes C_a^{\boldsymbol{x}}\, |\psi\rangle\,.$$

We want to show that the inconsistency between B and C is close to the inconsistency between A and B. In particular, we claim that

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{a \ne b} \langle\psi|\, A_a^{\boldsymbol{x}} \otimes C_b^{\boldsymbol{x}}\, |\psi\rangle \approx_{\sqrt{\epsilon}} \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{a \ne b} \langle\psi|\, B_a^{\boldsymbol{x}} \otimes C_b^{\boldsymbol{x}}\, |\psi\rangle\,.$$

To show this, we bound the magnitude of the difference using Cauchy-Schwarz.

$$\left| \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle \psi | \left(A_a^{\boldsymbol{x}} - B_a^{\boldsymbol{x}} \right) \otimes C_a^{\boldsymbol{x}} | \psi \rangle \right|$$

$$\leq \sqrt{ \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle \psi | \left(A_a^{\boldsymbol{x}} - B_a^{\boldsymbol{x}} \right)^2 \otimes I | \psi \rangle } \cdot \sqrt{ \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle \psi | I \otimes \left(C_a^{\boldsymbol{x}} \right)^2 | \psi \rangle }$$

$$\leq \sqrt{\epsilon} \cdot 1. \qquad\qquad\qquad \text{(because } C \text{ is a sub-measurement)}$$

This completes the proof. □

Next, we show that in the case of measurements, the state-dependent distance is a weakening of the consistency.

Proposition 4.22 ("$\simeq$" implies "$\approx$" for measurements). *Let $A = \{A_a^x\}$ and $B = \{B_a^x\}$ be two measurements such that*

$$A_a^x \otimes I \simeq_\delta I \otimes B_a^x.$$

Then

$$A_a^x \otimes I \approx_{2\delta} I \otimes B_a^x.$$

This is an "if and only if" if A and B are both projective measurements.

Proof. Our goal is to bound

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \left\| \left(A_a^{\boldsymbol{x}} \otimes I - I \otimes B_a^{\boldsymbol{x}} \right) | \psi \rangle \right\|^2$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle \psi | \left(A_a^{\boldsymbol{x}} \otimes I - I \otimes B_a^{\boldsymbol{x}} \right)^2 | \psi \rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle \psi | \left(A_a^{\boldsymbol{x}} \right)^2 \otimes I | \psi \rangle + \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle \psi | I \otimes \left(B_a^{\boldsymbol{x}} \right)^2 | \psi \rangle - 2 \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle \psi | A_a^{\boldsymbol{x}} \otimes B_a^{\boldsymbol{x}} | \psi \rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle \psi | A_a^{\boldsymbol{x}} \otimes I | \psi \rangle + \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle \psi | I \otimes B_a^{\boldsymbol{x}} | \psi \rangle - 2 \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle \psi | A_a^{\boldsymbol{x}} \otimes B_a^{\boldsymbol{x}} | \psi \rangle$$

$$= 2 - 2 \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle \psi | A_a^{\boldsymbol{x}} \otimes B_a^{\boldsymbol{x}} | \psi \rangle \qquad\qquad \text{(because } A \text{ and } B \text{ are measurements)}$$

$$\leq 2 - 2 \cdot (1 - \delta) \qquad \text{(by Proposition 4.19 and the fact that } A \text{ and } B \text{ are measurements)}$$

$$= 2\delta.$$

This completes the proof. When A and B are projective, "if and only if" follows from the first inequality becoming an equality. □

Remark 4.23. We note that Proposition 4.22 certainly does *not* hold for sub-measurements. For example, if $A_a^x = 0$ for all a, then $A_a^x \otimes I \simeq_0 I \otimes B_a^x$, but

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \left\| \left(A_a^{\boldsymbol{x}} \otimes I - I \otimes B_a^{\boldsymbol{x}} \right) | \psi \rangle \right\|^2 = \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \left\| \left(I \otimes B_a^{\boldsymbol{x}} \right) | \psi \rangle \right\|^2,$$

which is nonzero unless $(I \otimes B_a^x) | \psi \rangle = 0$ for all x and a.

4.4.3 Deriving consistency relations from the state-dependent distance

Proposition 4.21 shows the purpose of the state-dependent distance, which is to derive new consistency relations from old ones. In addition, its proof uses a strategy which will recur frequently throughout this paper. In this strategy, we would like to demonstrate a sequence of expressions in which each expression is close to the previous one:

$$\mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| (A_0)_a^{\boldsymbol{x}} (B_0)_a^{\boldsymbol{x}} |\psi\rangle \approx_{\epsilon_1} \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| (A_1)_a^{\boldsymbol{x}} (B_1)_a^{\boldsymbol{x}} |\psi\rangle \approx_{\epsilon_2} \cdots \approx_{\epsilon_t} \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| (A_t)_a^{\boldsymbol{x}} (B_t)_a^{\boldsymbol{x}} |\psi\rangle .$$

By the triangle inequality, we can therefore conclude that the 0-th expression is close to the t-th expression. To show that the i-th quantity is close to the $(i+1)$-st, we will typically arrange for $A_i = A_{i+1}$, and we will swap out B_i for B_{i+1} using an approximation relation such as $(B_i)_a^x \approx (B_{i+1})_a^x$, with the help of the Cauchy-Schwarz inequality (or the same might occur with the roles of A and B reversed). Even if A_0 and B_0 can be written as local measurements applied to either side of a bipartition, e.g. $(A_0)_a^x = A_a^x \otimes I$ and $(B_0)_a^x = I \otimes B_a^x$, and likewise for A_t and B_t, the intermediate steps may feature matrices which do not decompose nicely across a bipartition. This is why Definition 4.20 is phrased so broadly, with no mention of a bipartition.

In general, the A_i's and B_i's encountered in this sequence of steps may be quite unstructured: for example, not sub-measurements, and possibly not even Hermitian. Thus, we are interested in determining which conditions to place on these measurements are sufficient to carry out this proof strategy. The following proposition gives a broad condition under which this can be accomplished.

Proposition 4.24. *Let $\{A_a^x\}$, $\{B_a^x\}$, and $\{C_{a,b}^x\}$ be matrices. Suppose that $A_a^x \approx_\gamma B_a^x$ and that for all x, $\sum_a (\sum_b C_{a,b}^x)(\sum_b C_{a,b}^x)^\dagger \leq I$. Then*

$$\mathbf{E}_{\boldsymbol{x}} \sum_{a,b} \langle\psi| C_{a,b}^{\boldsymbol{x}} A_a^{\boldsymbol{x}} |\psi\rangle \approx_{\sqrt{\gamma}} \mathbf{E}_{\boldsymbol{x}} \sum_{a,b} \langle\psi| C_{a,b}^{\boldsymbol{x}} B_a^{\boldsymbol{x}} |\psi\rangle . \tag{4}$$

Similarly, suppose that $(A_a^x)^\dagger \approx_\gamma (B_a^x)^\dagger$ and that for all x, $\sum_a (\sum_b C_{a,b}^x)^\dagger (\sum_b C_{a,b}^x) \leq I$. Then

$$\mathbf{E}_{\boldsymbol{x}} \sum_{a,b} \langle\psi| A_a^{\boldsymbol{x}} C_{a,b}^{\boldsymbol{x}} |\psi\rangle \approx_{\sqrt{\gamma}} \mathbf{E}_{\boldsymbol{x}} \sum_{a,b} \langle\psi| B_a^{\boldsymbol{x}} C_{a,b}^{\boldsymbol{x}} |\psi\rangle . \tag{5}$$

Proof. We begin by showing Equation (4).

$$\left| \mathbf{E}_{\boldsymbol{x}} \sum_{a,b} \langle\psi| C_{a,b}^{\boldsymbol{x}} (A_a^{\boldsymbol{x}} - B_a^{\boldsymbol{x}}) |\psi\rangle \right|$$

$$= \left| \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| \left(\sum_b C_{a,b}^{\boldsymbol{x}} \right) \cdot (A_a^{\boldsymbol{x}} - B_a^{\boldsymbol{x}}) |\psi\rangle \right|$$

$$\leq \left(\mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| \left(\sum_b C_{a,b}^{\boldsymbol{x}} \right) \left(\sum_b C_{a,b}^{\boldsymbol{x}} \right)^\dagger |\psi\rangle \right)^{1/2} \cdot \left(\mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| (A_a^{\boldsymbol{x}} - B_a^{\boldsymbol{x}})^\dagger (A_a^{\boldsymbol{x}} - B_a^{\boldsymbol{x}}) |\psi\rangle \right)^{1/2}$$

$$\leq \sqrt{\gamma}.$$

The third line uses Cauchy-Schwarz, and the fourth line uses the assumption $\sum_a (\sum_b C_{a,b}^x)(\sum_b C_{a,b}^x)^\dagger \leq I$ to bound the first factor by 1 and the assumption $A_a^x \approx_\gamma B_a^x$ to bound the second factor by $\sqrt{\gamma}$. As for Equation (5), we want to bound

$$\left| \mathbf{E}_{\boldsymbol{x}} \sum_{a,b} \langle\psi| (A_a^{\boldsymbol{x}} - B_a^{\boldsymbol{x}}) C_{a,b}^{\boldsymbol{x}} |\psi\rangle \right| = \left| \mathbf{E}_{\boldsymbol{x}} \sum_{a,b} \langle\psi| (C_{a,b}^{\boldsymbol{x}})^\dagger ((A_a^{\boldsymbol{x}})^\dagger - (B_a^{\boldsymbol{x}})^\dagger) |\psi\rangle \right|.$$

It then follows from Equation (4) that this is at most $\sqrt{\gamma}$. $\qquad\square$

22

Proposition 4.24 is broad enough to capture almost all of our applications of the state-dependent distance. Unfortunately, defining the $C_{a,b}^x$ matrices and showing that they satisfy the inequality $\sum_a (\sum_b C_{a,b}^x)(\sum_b C_{a,b}^x)^\dagger \leq I$ can be somewhat cumbersome. As a result, we will usually carry out these Cauchy-Schwarz calculations by hand. However, we will occasionally use the following proposition which simplifies Proposition 4.24.

Proposition 4.25. *Let $A = \{A_a^x\}$, $B = \{B_a^x\}$, and $C = \{C_a^x\}$ be sub-measurements such that $A_a^x \approx_\delta B_a^x$. Then*

$$\mathbf{E}_{\boldsymbol{x}} \sum_a \langle \psi | A_a^{\boldsymbol{x}} C_a^{\boldsymbol{x}} | \psi \rangle \approx_{\sqrt{\delta}} \mathbf{E}_{\boldsymbol{x}} \sum_a \langle \psi | B_a^{\boldsymbol{x}} C_a^{\boldsymbol{x}} | \psi \rangle .$$

Proof. To show this, we bound the magnitude of the difference.

$$\left| \mathbf{E}_{\boldsymbol{x}} \sum_a \langle \psi | (A_a^{\boldsymbol{x}} - B_a^{\boldsymbol{x}}) \cdot (C_a^{\boldsymbol{x}}) | \psi \rangle \right|$$

$$\leq \sqrt{\mathbf{E}_{\boldsymbol{x}} \sum_a \langle \psi | (A_a^{\boldsymbol{x}} - B_a^{\boldsymbol{x}})^2 | \psi \rangle} \cdot \sqrt{\mathbf{E}_{\boldsymbol{x}} \sum_a \langle \psi | (C_a^{\boldsymbol{x}})^2 | \psi \rangle}$$

$$\leq \sqrt{\delta} \cdot \sqrt{1}.$$

This completes the proof. $\qquad\square$

A proposition similar to Proposition 4.24, but for "$\approx$", holds as well. This is [NW19, Fact 4.20].

Proposition 4.26. *Let $\{A_a^x\}, \{B_a^x\}$, and $\{C_{a,b}^x\}$ be matrices. Suppose that $A_a^x \approx_\delta B_a^x$ and that for all x and a, $\sum_b (C_{a,b}^x)^\dagger (C_{a,b}^x) \leq I$. Then*

$$C_{a,b}^x A_a^x \approx_\delta C_{a,b}^x B_a^x.$$

Proof. The error we wish to bound is

$$\mathbf{E}_{\boldsymbol{x}} \sum_{a,b} \langle \psi | (A_a^{\boldsymbol{x}} - B_a^{\boldsymbol{x}})^\dagger (C_{a,b}^{\boldsymbol{x}})^\dagger C_{a,b}^{\boldsymbol{x}} (A_a^{\boldsymbol{x}} - B_a^{\boldsymbol{x}}) | \psi \rangle \leq \mathbf{E}_{\boldsymbol{x}} \sum_a \langle \psi | (A_a^{\boldsymbol{x}} - B_a^{\boldsymbol{x}})^\dagger (A_a^{\boldsymbol{x}} - B_a^{\boldsymbol{x}}) | \psi \rangle$$

$$\leq \delta. \qquad\square$$

4.4.4 Miscellaneous distance properties

We now state a few miscellaneous properties of our two distances.

Proposition 4.27 (Triangle inequality for vectors squared)**.** *Let $|\psi_1\rangle, \ldots, |\psi_k\rangle$ be vectors. Then*

$$\| |\psi_1\rangle + \cdots + |\psi_k\rangle \|^2 \leq k \cdot (\| |\psi_1\rangle \|^2 + \cdots + \| |\psi_k\rangle \|^2).$$

Proof. First, if $x_1, \ldots, x_k \in \mathbb{R}$, then

$$(x_1 + \cdots + x_k)^2 = \sum_{i,j=1}^k x_i x_j \leq \sum_{i,j=1}^k \frac{x_i^2 + x_j^2}{2} = \sum_{i,j=1}^k x_i^2 = \sum_{i=1}^k k \cdot x_i^2. \tag{6}$$

Next, by the triangle inequality

$$\| |\psi_1\rangle + \cdots + |\psi_k\rangle \|^2 = (\| |\psi_1\rangle + \cdots + |\psi_k\rangle \|)^2$$

$$\leq (\| |\psi_1\rangle \| + \cdots + \| |\psi_k\rangle \|)^2$$

$$\leq k \cdot (\| |\psi_1\rangle \|^2 + \cdots + \| |\psi_k\rangle \|^2).$$

where the last step uses Equation (6) applied to the case of $x_i = \| |\psi_i\rangle \|$. $\qquad\square$

Proposition 4.28 (Triangle inequality for "$\approx_\delta$")**.** *Suppose $A_1 = \{(A_1)_a^x\}, \ldots, A_{k+1} = \{(A_{k+1})_a^x\}$ is a set of matrices such that*

$$(A_i)_a^x \approx_{\delta_i} (A_{i+1})_a^x$$

for all $i \in [k]$. Then

$$(A_1)_a^x \approx_{k \cdot (\delta_1 + \cdots + \delta_k)} (A_{k+1})_a^x.$$

Proof. We want to bound

$$\mathbf{E}_{\boldsymbol{x}} \sum_a \| ((A_1)_a^{\boldsymbol{x}} - (A_{k+1})_a^{\boldsymbol{x}}) \, |\psi\rangle \|^2$$

$$= \mathbf{E}_{\boldsymbol{x}} \sum_a \| (((A_1)_a^{\boldsymbol{x}} - (A_2)_a^{\boldsymbol{x}}) + \cdots + ((A_k)_a^{\boldsymbol{x}} - (A_{k+1})_a^{\boldsymbol{x}})) \, |\psi\rangle \|^2$$

$$\le \mathbf{E}_{\boldsymbol{x}} \sum_a k \cdot (\| ((A_1)_a^{\boldsymbol{x}} - (A_2)_a^{\boldsymbol{x}}) \, |\psi\rangle \|^2 + \cdots + \| ((A_k)_a^{\boldsymbol{x}} - (A_{k+1})_a^{\boldsymbol{x}}) \, |\psi\rangle \|^2)$$

$$\le k \cdot (\delta_1 + \cdots + \delta_k),$$

where the inequality uses Proposition 4.27 applied to the vectors $((A_i)_a^{\boldsymbol{x}} - (A_{i+1})_a^{\boldsymbol{x}}) \, |\psi\rangle$ for $i \in [k]$. $\quad\square$

We note that Proposition 4.28 contrasts with the triangle inequality for "$\approx_\delta$" when applied to numbers, i.e. Equation (1), for which no multiplicative factor of k appears in the error.

The following is Fact 4.29 from [NW19]; however, they incorrectly claimed a final bound of $A_a^x \otimes I \simeq_{\epsilon + \delta + \gamma} I \otimes D_a^x$. We give a new proof of this statement, albeit with a slightly weaker quantitative bound.

Proposition 4.29 (Triangle inequality for "$\simeq$")**.** *Suppose that A, B, C, and D are measurements such that*

$$A_a^x \otimes I \simeq_\epsilon I \otimes B_a^x, \quad C_a^x \otimes I \simeq_\delta I \otimes B_a^x, \quad C_a^x \otimes I \simeq_\gamma I \otimes D_a^x.$$

Then

$$A_a^x \otimes I \simeq_{\epsilon + 2\sqrt{\delta + \gamma}} I \otimes D_a^x.$$

Proof. Because B, C, and D are measurements, Proposition 4.22 implies that

$$C_a^x \otimes I \approx_{2\delta} I \otimes B_a^x, \quad C_a^x \otimes I \approx_{2\gamma} I \otimes D_a^x.$$

The triangle inequality, Proposition 4.28, then implies that

$$I \otimes B_a^x \approx_{4\delta + 4\gamma} I \otimes D_a^x.$$

Finally, Proposition 4.21 implies that

$$A_a^x \otimes I \simeq_{\epsilon + \sqrt{4\delta + 4\gamma}} I \otimes D_a^x.$$

This completes the proof. $\quad\square$

Proposition 4.30 (Data processing for "$\simeq$")**.** *Let $A = \{A_a^x\}$ and $B = \{B_a^x\}$ be two measurements such that*

$$A_a^x \otimes I \simeq_\delta I \otimes B_a^x.$$

Then for any function f,

$$A_{[f(a)=b]}^x \otimes I \simeq_\delta I \otimes B_{[f(a)=b]}^x.$$

24

Proof. We want to bound

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{b \neq b'} \langle\psi| A^{\boldsymbol{x}}_{[f(a)=b]} \otimes B^{\boldsymbol{x}}_{[f(a)=b']} |\psi\rangle = \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{b \neq b'} \sum_{a:f(a)=b} \sum_{a':f(a')=b'} \langle\psi| A^{\boldsymbol{x}}_a \otimes B^{\boldsymbol{x}}_{a'} |\psi\rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{a \neq a'} \langle\psi| A^{\boldsymbol{x}}_a \otimes B^{\boldsymbol{x}}_{a'} |\psi\rangle$$

$$\leq \delta.$$

This completes the proof. $\qquad\square$

The following fact is useful for translating between statements about consistency and closeness between sub-measurements.

Proposition 4.31. *Let $\{A^x_a\}$ be a sub-measurement and let $\{B^x_a\}$ be a measurement such that on average over x,*

$$A^x_a \otimes I \simeq_\gamma I \otimes B^x_a \,.$$

Then the following hold

$$A^x_a \otimes I \approx_\gamma A^x_a \otimes B^x_a \approx_\gamma A^x \otimes B^x_a, \tag{7}$$

where $A^x = \sum_a A^x_a$. As a result, by [Proposition 4.28](#),

$$A^x_a \otimes I \approx_{4\gamma} A^x \otimes B^x_a$$

Proof. We establish the first approximation in [Equation (7)](#):

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| \left(A^{\boldsymbol{x}}_a \otimes (I - B^{\boldsymbol{x}}_a) \right)^2 |\psi\rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A^{\boldsymbol{x}}_a \otimes (I - B^{\boldsymbol{x}}_a) |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{\substack{a,b: \\ b \neq a}} \langle\psi| A^{\boldsymbol{x}}_a \otimes B^{\boldsymbol{x}}_b |\psi\rangle$$

$$\leq \gamma \,.$$

The first inequality folows from the fact that $A^x_a \otimes (I - B^x_a)$ has operator norm at most 1, and the third line follows from the fact that $\{B^x_b\}$ is a complete measurement, and the last line follows from the assumption of consistency between the A and B (sub-)measurements.

To establish the second approximation in [Equation (7)](#), we compute the difference:

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| \left((A^{\boldsymbol{x}} - A^{\boldsymbol{x}}_a) \otimes B^{\boldsymbol{x}}_a \right)^2 |\psi\rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| (A^{\boldsymbol{x}} - A^{\boldsymbol{x}}_a) \otimes B^{\boldsymbol{x}}_a |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{\substack{a,a': \\ a \neq a'}} \langle\psi| A^{\boldsymbol{x}}_{a'} \otimes B^{\boldsymbol{x}}_a |\psi\rangle$$

$$\leq \gamma \,.$$

The second line follows from the fact that $(A^x - A^x_a) \otimes B^x_a$ has operator norm at most 1, and the last inequality follows from the consistency between the A and B (sub-)measurements. $\qquad\square$

Proposition 4.32. *Suppose $\{A_a^x\}$ is a projective sub-measurement satisfying*

$$A_a^x \otimes I \approx_\delta I \otimes A_a^x. \tag{8}$$

Then for any $0 \leq B \leq I$, the following holds:

$$\mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} B A_a^{\boldsymbol{x}} \otimes I |\psi\rangle \approx_{2\sqrt{\delta}} \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| B \otimes A_a^{\boldsymbol{x}} |\psi\rangle \approx_{\sqrt{\delta}} \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| B A_a^{\boldsymbol{x}} \otimes I |\psi\rangle \tag{9}$$

Proof. We will show the first approximation in Equation (9) in two steps. First, we show that

$$\mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} B A_a^{\boldsymbol{x}} \otimes I |\psi\rangle \approx_{\sqrt{\delta}} \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} B \otimes A_a^{\boldsymbol{x}} |\psi\rangle \,. \tag{10}$$

To do so, we bound the magnitude of the difference.

$$\left| \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| (A_a^{\boldsymbol{x}} B \otimes I) \cdot (A_a^{\boldsymbol{x}} \otimes I - I \otimes A_a^{\boldsymbol{x}}) |\psi\rangle \right|$$

$$\leq \left(\mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} B^2 A_a^{\boldsymbol{x}} \otimes I |\psi\rangle \right)^{1/2} \cdot \left(\mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| (A_a^{\boldsymbol{x}} \otimes I - I \otimes A_a^{\boldsymbol{x}})^2 |\psi\rangle \right)^{1/2}$$

$$\leq \sqrt{\delta}. \qquad\qquad\qquad\qquad\qquad (\text{because } B \leq I \text{ and } (8))$$

Next, we show that

$$(10) = \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} B \otimes A_a^{\boldsymbol{x}} |\psi\rangle \approx_{\sqrt{\delta}} \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| B \otimes A_a^{\boldsymbol{x}} |\psi\rangle \,.$$

To do so, we bound the magnitude of the difference.

$$\left| \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| (A_a^{\boldsymbol{x}} \otimes I - I \otimes A_a^{\boldsymbol{x}}) \cdot (B \otimes A_a^{\boldsymbol{x}}) |\psi\rangle \right| \qquad\qquad (\text{because } A \text{ is projective})$$

$$\leq \left(\mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| (A_a^{\boldsymbol{x}} \otimes I - I \otimes A_a^{\boldsymbol{x}})^2 |\psi\rangle \right)^{1/2} \cdot \left(\mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| B^2 \otimes A_a^{\boldsymbol{x}} |\psi\rangle \right)^{1/2}$$

$$\leq \sqrt{\delta}. \qquad\qquad\qquad\qquad\qquad (\text{because } B \leq I \text{ and } (8))$$

Thus, the first approximation in Equation (9) follows. To show the second approximation, we bound the magnitude of the difference.

$$\left| \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| (B \otimes I) \cdot (A_a^{\boldsymbol{x}} \otimes I - I \otimes A_a^{\boldsymbol{x}}) |\psi\rangle \right|$$

$$= \left| \mathbf{E}_{\boldsymbol{x}} \langle\psi| (B \otimes I) \cdot \sum_a (A_a^{\boldsymbol{x}} \otimes I - I \otimes A_a^{\boldsymbol{x}}) |\psi\rangle \right|$$

$$\leq \left(\mathbf{E}_{\boldsymbol{x}} \langle\psi| B^2 \otimes I |\psi\rangle \right)^{1/2} \cdot \left(\mathbf{E}_{\boldsymbol{x}} \sum_{a,b} \langle\psi| (A_a^{\boldsymbol{x}} \otimes I - I \otimes A_a^{\boldsymbol{x}}) \cdot (A_b^{\boldsymbol{x}} \otimes I - I \otimes A_b^{\boldsymbol{x}}) |\psi\rangle \right)^{1/2}.$$

The first term in the product is at most 1 because $B \leq I$. We bound the expression inside the

second square root as follows.

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{a,b} \langle\psi| \left(A_a^{\boldsymbol{x}} \otimes I - I \otimes A_a^{\boldsymbol{x}}\right) \cdot \left(A_b^{\boldsymbol{x}} \otimes I - I \otimes A_b^{\boldsymbol{x}}\right) |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{a,b} \langle\psi| \left(A_a^{\boldsymbol{x}} A_b^{\boldsymbol{x}} \otimes I + I \otimes A_a^{\boldsymbol{x}} A_b^{\boldsymbol{x}} - A_a^{\boldsymbol{x}} \otimes A_b^{\boldsymbol{x}} - A_b^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}}\right) |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| (A_a^{\boldsymbol{x}})^2 \otimes I |\psi\rangle + \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| I \otimes (A_a^{\boldsymbol{x}})^2 |\psi\rangle - 2 \cdot \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{a,b} \langle\psi| A_a^{\boldsymbol{x}} \otimes A_b^{\boldsymbol{x}} |\psi\rangle$$

$$\text{(because } A \text{ is projective)}$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| (A_a^{\boldsymbol{x}})^2 \otimes I |\psi\rangle + \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| I \otimes (A_a^{\boldsymbol{x}})^2 |\psi\rangle - 2 \cdot \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}} |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| \left(A_a^{\boldsymbol{x}} \otimes I - I \otimes A_a^{\boldsymbol{x}}\right)^2 |\psi\rangle$$

$$\leq \delta,$$

where the last step uses Equation (8). □

Proposition 4.33. *Let $A = \{A_a^x\}$ be a sub-measurement and let P be a projective sub-measurement such that $A_a^x \otimes I \approx_\epsilon P_a^x \otimes I$. Then*

$$\langle\psi| A \otimes I |\psi\rangle \geq \langle\psi| P \otimes I |\psi\rangle - 2\sqrt{\epsilon}.$$

Proof. We calculate:

$$\langle\psi| P \otimes I |\psi\rangle = \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| P_a^{\boldsymbol{x}} \otimes I |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| (P_a^{\boldsymbol{x}})^2 \otimes I |\psi\rangle \qquad \text{(because } P \text{ is projective)}$$

$$\approx_{\sqrt{\epsilon}} \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| (A_a^{\boldsymbol{x}} \cdot P_a^{\boldsymbol{x}}) \otimes I |\psi\rangle \qquad \text{(by Proposition 4.25)}$$

$$\approx_{\sqrt{\epsilon}} \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| (A_a^{\boldsymbol{x}})^2 \otimes I |\psi\rangle \qquad \text{(by Proposition 4.25)}$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes I |\psi\rangle$$

$$= \langle\psi| A \otimes I |\psi\rangle.$$

This completes the proof. □

4.4.5 Strong self-consistency

An important property of a sub-measurement $A = \{A_a^x\}$ is that if both provers measure using A, then they receive the same outcome. It seems natural to study this using *self-consistency* of A, i.e. the number δ such that

$$A_a^x \otimes I \simeq_\delta I \otimes A_a^x.$$

However, when A is a sub-measurement, being δ-self consistent only implies the following weaker condition: if both provers measure using A and one of them receives $\boldsymbol{a}$, then the other will most likely either receive $\boldsymbol{a}$ *or not receive any outcome whatsoever*. This motivates defining the following stronger notion of self-consistency.

Definition 4.34 (Strong self consistency). Let $|\psi\rangle$ be a permutation-invariant state in $\mathcal{H} \otimes \mathcal{H}$ and let $A = \{A_a^x\}$ be a sub-measurement acting on $\mathcal{H}$. Then A is δ-*strongly self consistent* if

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}} |\psi\rangle \geq \langle\psi| A \otimes I |\psi\rangle - \delta.$$

We now relate strong self-consistency to our two notions of similarity. First, we show that strong self-consistency is indeed a stronger condition than $A_a^x \otimes I \simeq_\delta I \otimes A_a^x$, at least for sub-measurements.

Proposition 4.35. *Let $|\psi\rangle$ be a permutation-invariant state, and let $A = \{A_a^x\}$ be a sub-measurement. If*

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}} |\psi\rangle \geq \langle\psi| A \otimes I |\psi\rangle - \delta$$

then $A_a^x \otimes I \simeq_\delta I \otimes A_a^x$. This is an "if and only if" if A is a measurement.

Proof. For a sub-measurement A,

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{a \neq b} \langle\psi| A_a^{\boldsymbol{x}} \otimes A_b^{\boldsymbol{x}} |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes (A^{\boldsymbol{x}} - A_a^{\boldsymbol{x}}) |\psi\rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes (I - A_a^{\boldsymbol{x}}) |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes I |\psi\rangle - \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}} |\psi\rangle$$

$$= \langle\psi| A \otimes I |\psi\rangle - \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}} |\psi\rangle.$$

This is at most δ if A is δ-strongly self-consistent. On the other hand, if A is a measurement, then the inequality becomes an equality. Hence, if $A_a^x \otimes I \simeq_\delta I \otimes A_a^x$, then A is δ-strongly self-consistent. $\square$

Next, we show that strong self-consistency is also a stronger condition than $A_a^x \otimes I \approx_{2\delta} I \otimes A_a^x$, at least for non-projective measurements.

Proposition 4.36. *Let $|\psi\rangle$ be a permutation-invariant state, and let $A = \{A_a^x\}$ be a sub-measurement. If*

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}} |\psi\rangle \geq \langle\psi| A \otimes I |\psi\rangle - \delta$$

then $A_a^x \otimes I \approx_{2\delta} I \otimes A_a^x$. This is an "if and only if" if A is projective.

Proof. For general (i.e. not necessarily projective) A

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \| (A_a^{\boldsymbol{x}} \otimes I - I \otimes A_a^{\boldsymbol{x}}) |\psi\rangle \|^2$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| (A_a^{\boldsymbol{x}} \otimes I - I \otimes A_a^{\boldsymbol{x}})^2 |\psi\rangle$$

$$= 2 \cdot \left(\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| (A_a^{\boldsymbol{x}})^2 \otimes I |\psi\rangle - \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}} |\psi\rangle \right)$$

$$\leq 2 \cdot \left(\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes I |\psi\rangle - \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}} |\psi\rangle \right). \tag{11}$$

This is at most $2 \cdot \delta$ if

$$\mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}} |\psi\rangle \geq \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes I |\psi\rangle - \delta.$$

If A is projective, then Equation (11) becomes an equality, and so $A_a^x \otimes I \approx_{2\delta} I \otimes A_a^x$ implies that

$$\mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes I |\psi\rangle - \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}} |\psi\rangle = \frac{1}{2} \cdot \left(\mathbf{E}_{\boldsymbol{x}} \sum_a \|(A_a^{\boldsymbol{x}} \otimes I - I \otimes A_a^{\boldsymbol{x}}) |\psi\rangle \|^2 \right) \leq \delta. \quad \square$$

Hence, we may also refer to the condition $A_a^x \otimes I \approx_{2\delta} I \otimes A_a^x$ as "strong self-consistency" if A is projective.

For the remainder of the section, we will prove various properties of strongly self-consistent sub-measurements.

Proposition 4.37. *Let $|\psi\rangle$ be a permutation-invariant state, and let $A = \{A_a^x\}$ be a sub-measurement such that*

$$\mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}} |\psi\rangle \geq \langle\psi| A \otimes I |\psi\rangle - \delta.$$

Then for any function f, $A_{[f(a)=b]}^x \otimes I \approx_{2\delta} I \otimes A_{[f(a)=b]}^x$.

Proof. Our goal is to bound

$$\mathbf{E}_{\boldsymbol{x}} \sum_b \|(A_{[f(a)=b]}^{\boldsymbol{x}} \otimes I - I \otimes A_{[f(a)=b]}^{\boldsymbol{x}}) |\psi\rangle \|^2$$

$$= \mathbf{E}_{\boldsymbol{x}} \sum_b \langle\psi| (A_{[f(a)=b]}^{\boldsymbol{x}} \otimes I - I \otimes A_{[f(a)=b]}^{\boldsymbol{x}})^2 |\psi\rangle$$

$$= 2 \cdot \left(\mathbf{E}_{\boldsymbol{x}} \sum_b \langle\psi| (A_{[f(a)=b]}^{\boldsymbol{x}})^2 \otimes I |\psi\rangle - \mathbf{E}_{\boldsymbol{x}} \sum_b \langle\psi| A_{[f(a)=b]}^{\boldsymbol{x}} \otimes A_{[f(a)=b]}^{\boldsymbol{x}} |\psi\rangle \right)$$

$$\leq 2 \cdot \left(\mathbf{E}_{\boldsymbol{x}} \sum_b \langle\psi| A_{[f(a)=b]}^{\boldsymbol{x}} \otimes I |\psi\rangle - \mathbf{E}_{\boldsymbol{x}} \sum_b \langle\psi| A_{[f(a)=b]}^{\boldsymbol{x}} \otimes A_{[f(a)=b]}^{\boldsymbol{x}} |\psi\rangle \right).$$

$$= 2 \cdot \left(\langle\psi| A \otimes I |\psi\rangle - \mathbf{E}_{\boldsymbol{x}} \sum_b \langle\psi| A_{[f(a)=b]}^{\boldsymbol{x}} \otimes A_{[f(a)=b]}^{\boldsymbol{x}} |\psi\rangle \right). \tag{12}$$

The second term in Equation (12) can be bounded by

$$\mathbf{E}_{\boldsymbol{x}} \sum_b \langle\psi| A_{[f(a)=b]}^{\boldsymbol{x}} \otimes A_{[f(a)=b]}^{\boldsymbol{x}} |\psi\rangle = \mathbf{E}_{\boldsymbol{x}} \sum_{a,a':f(a)=f(a')} \langle\psi| A_a^{\boldsymbol{x}} \otimes A_{a'}^{\boldsymbol{x}} |\psi\rangle$$

$$\geq \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}} |\psi\rangle$$

$$\geq \langle\psi| A \otimes I |\psi\rangle - \delta.$$

Hence,

$$(12) \leq 2 \cdot \left(\langle\psi| A \otimes I |\psi\rangle - (\langle\psi| A \otimes I |\psi\rangle - \delta) \right) = 2\delta.$$

This concludes the proof. $\square$

Proposition 4.38. *Let $|\psi\rangle$ be a permutation-invariant state, and let $A = \{A_a^x\}$ be a sub-measurement such that*

$$\mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| A_a^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}} |\psi\rangle \geq \langle\psi| A \otimes I |\psi\rangle - \delta.$$

29

In addition, let B be a sub-measurement such that $A_a^x \otimes I \approx_\epsilon B_a^x \otimes I$. Then

$$\langle\psi|\, B \otimes I \,|\psi\rangle \geq \langle\psi|\, A \otimes I \,|\psi\rangle - \delta - 2\sqrt{\epsilon}.$$

Proof. We calculate:

$$\begin{aligned}
\langle\psi|\, B \otimes I \,|\psi\rangle &= \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi|\, B_a^{\boldsymbol{x}} \otimes I \,|\psi\rangle \\
&\geq \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi|\, B_a^{\boldsymbol{x}} \otimes B_a^{\boldsymbol{x}} \,|\psi\rangle \\
&\geq \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi|\, A_a^{\boldsymbol{x}} \otimes B_a^{\boldsymbol{x}} \,|\psi\rangle - \sqrt{\epsilon} && \text{(by Proposition 4.25)} \\
&\geq \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi|\, A_a^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}} \,|\psi\rangle - 2\sqrt{\epsilon} && \text{(by Proposition 4.25)} \\
&\geq \langle\psi|\, A \otimes I \,|\psi\rangle - \delta - 2\sqrt{\epsilon}.
\end{aligned}$$

This completes the proof. $\qquad\square$

Proposition 4.39. *Let $|\psi\rangle$ be a permutation-invariant state, and let $A = \{A_a^x\}$ be a sub-measurement such that*

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_a \langle\psi|\, A_a^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{x}} \,|\psi\rangle \geq \langle\psi|\, A \otimes I \,|\psi\rangle - \delta.$$

In addition, let P be a projective sub-measurement such that $P_a^x \otimes I \approx_\epsilon A_a^x \otimes I$. Then for any function f,

$$P_{[f(a)=b]}^x \otimes I \approx_{8\delta+8\sqrt{\epsilon}} A_{[f(a)=b]}^x \otimes I.$$

Proof. We will begin by showing that

$$P_{[f(a)=b]}^x \otimes I \approx_{2\delta+4\sqrt{\epsilon}} I \otimes A_{[f(a)=b]}^x. \tag{13}$$

To do so, our goal is to bound

$$\begin{aligned}
& \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_b \big\| (P_{[f(a)=b]}^{\boldsymbol{x}} \otimes I - I \otimes A_{[f(a)=b]}^{\boldsymbol{x}}) \,|\psi\rangle \big\|^2 \\
={}& \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_b \langle\psi|\, (P_{[f(a)=b]}^{\boldsymbol{x}} \otimes I - I \otimes A_{[f(a)=b]}^{\boldsymbol{x}})^2 \,|\psi\rangle \\
={}& \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_b \langle\psi|\, \Big((P_{[f(a)=b]}^{\boldsymbol{x}})^2 \otimes I + I \otimes (A_{[f(a)=b]}^{\boldsymbol{x}})^2 - 2 \cdot P_{[f(a)=b]}^{\boldsymbol{x}} \otimes A_{[f(a)=b]}^{\boldsymbol{x}} \Big) \,|\psi\rangle \\
\leq{}& \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_b \langle\psi|\, \Big(P_{[f(a)=b]}^{\boldsymbol{x}} \otimes I + I \otimes A_{[f(a)=b]}^{\boldsymbol{x}} - 2 \cdot P_{[f(a)=b]}^{\boldsymbol{x}} \otimes A_{[f(a)=b]}^{\boldsymbol{x}} \Big) \,|\psi\rangle \\
={}& \langle\psi|\, P \otimes I \,|\psi\rangle + \langle\psi|\, I \otimes A \,|\psi\rangle - 2 \cdot \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_b \langle\psi|\, P_{[f(a)=b]}^{\boldsymbol{x}} \otimes A_{[f(a)=b]}^{\boldsymbol{x}} \,|\psi\rangle. \tag{14}
\end{aligned}$$

By Proposition 4.33, the first term in Equation (14) is at most

$$\langle\psi|\, P \otimes I \,|\psi\rangle \leq \langle\psi|\, A \otimes I \,|\psi\rangle + 2\sqrt{\epsilon}.$$

As for the third term in Equation (14), we can bound it by

$$\mathbf{E}_{\boldsymbol{x}} \sum_b \langle\psi| P^{\boldsymbol{x}}_{[f(a)=b]} \otimes A^{\boldsymbol{x}}_{[f(a)=b]} |\psi\rangle = \mathbf{E}_{\boldsymbol{x}} \sum_{a,a':f(a)=f(a')} \langle\psi| P^{\boldsymbol{x}}_a \otimes A^{\boldsymbol{x}}_{a'} |\psi\rangle$$

$$\geq \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| P^{\boldsymbol{x}}_a \otimes A^{\boldsymbol{x}}_a |\psi\rangle$$

$$\geq \mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| A^{\boldsymbol{x}}_a \otimes A^{\boldsymbol{x}}_a |\psi\rangle - \sqrt{\epsilon} \qquad \text{(by Proposition 4.25)}$$

$$\geq \langle\psi| A \otimes I |\psi\rangle - \delta - \sqrt{\epsilon}.$$

Putting everything together,

$$(14) \leq (\langle\psi| A \otimes I |\psi\rangle + 2\sqrt{\epsilon}) + \langle\psi| A \otimes I |\psi\rangle - 2 \cdot (\langle\psi| A \otimes I |\psi\rangle - \delta - \sqrt{\epsilon}) = 2\delta + 4\sqrt{\epsilon}.$$

This proves Equation (13).

Proposition 4.37 implies that

$$A^x_{[f(a)=b]} \otimes I \approx_{2\delta} I \otimes A^x_{[f(a)=b]}.$$

Hence, by Equation (13),

$$P^x_{[f(a)=b]} \otimes I \approx_{2\delta+4\sqrt{\epsilon}} I \otimes A^x_{[f(a)=b]} \approx_{2\delta} A^x_{[f(a)=b]} \otimes I.$$

Thus, by Proposition 4.28,

$$P^x_{[f(a)=b]} \otimes I \approx_{8\delta+8\sqrt{\epsilon}} A^x_{[f(a)=b]} \otimes I.$$

This concludes the proof. □

Proposition 4.40. *Let $|\psi\rangle$ be a permutation-invariant state, and let $A = \{A_a\}$ be a measurement such that*

$$\mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| A_a \otimes A_a |\psi\rangle \geq \langle\psi| A \otimes I |\psi\rangle - \zeta.$$

Suppose $B = \{B_a\}$ is a sub-measurement such that $A_a \otimes I \approx_\delta B_a \otimes I$. Let $C = \{C_a\}$ be a measurement in which there is an a^* such that $C_{a^*} = B_{a^*} + (I - B)$ and $C_a = B_a$ for all $a \neq a^*$. Then $A_a \otimes I \approx_{2\delta+4\sqrt{\delta}+2\zeta} C_a \otimes I$.

Before proving this, we need the following proposition.

Proposition 4.41. *Let $|\psi\rangle$ be a permutation-invariant state, and let $A = \{A_a\}$ be a sub-measurement such that*

$$\mathbf{E}_{\boldsymbol{x}} \sum_a \langle\psi| A_a \otimes A_a |\psi\rangle \geq \langle\psi| A \otimes I |\psi\rangle - \zeta.$$

Then

$$\sum_a \langle\psi| (A_a)^2 \otimes I |\psi\rangle \geq \sum_a \langle\psi| A_a \otimes I |\psi\rangle - \zeta.$$

Proof. Applying Cauchy-Schwarz, we have

$$\sum_a \langle\psi| A_a \otimes A_a |\psi\rangle = \sum_a \langle\psi| (A_a \otimes I) \cdot (I \otimes A_a) |\psi\rangle$$

$$\leq \sqrt{\sum_a \langle\psi| (A_a)^2 \otimes I |\psi\rangle} \cdot \sqrt{\sum_a \langle\psi| I \otimes (A_a)^2 |\psi\rangle}$$

$$= \sum_a \langle\psi| (A_a)^2 \otimes I |\psi\rangle. \qquad (15)$$

31

As a result,

$$\sum_a \langle\psi|(A_a)^2 \otimes I |\psi\rangle \geq \sum_a \langle\psi| A_a \otimes A_a |\psi\rangle \qquad \text{(by Equation (15))}$$

$$\geq \sum_a \langle\psi| A_a \otimes I |\psi\rangle - \zeta. \qquad \text{(by self-consistency of } A)$$

This completes the proof. $\qquad\square$

Now we prove Proposition 4.40.

Proof of Proposition 4.40. By Proposition 4.27 (the triangle inequality for vectors squared),

$$\sum_a \|(A_a - C_a) \otimes I |\psi\rangle\|^2$$

$$= \sum_{a \neq a^*} \|(A_a - B_a) \otimes I |\psi\rangle\|^2 + \|(A_{a^*} - (I - B + B_{a^*})) \otimes I |\psi\rangle\|^2$$

$$\leq \sum_{a \neq a^*} \|(A_a - B_a) \otimes I |\psi\rangle\|^2 + 2 \cdot \|(A_{a^*} - B_{a^*}) \otimes I |\psi\rangle\|^2 + 2 \cdot \|(I - B) \otimes I |\psi\rangle\|^2$$

$$\leq 2 \sum_{a \neq a^*} \|(A_a - B_a) \otimes I |\psi\rangle\|^2 + 2 \cdot \|(A_{a^*} - B_{a^*}) \otimes I |\psi\rangle\|^2 + 2 \cdot \|(I - B) \otimes I |\psi\rangle\|^2$$

$$= 2 \sum_a \|(A_a - B_a) \otimes I |\psi\rangle\|^2 + 2 \cdot \|(I - B) \otimes I |\psi\rangle\|^2$$

$$\leq 2\delta + 2 \cdot \|(I - B) \otimes I |\psi\rangle\|^2. \qquad \text{(because } A_a \otimes I \approx_\delta B_a \otimes I)$$

The second term we can bound as follows.

$$\|(I - B) \otimes I |\psi\rangle\|^2 = \langle\psi|(I - B)^2 \otimes I |\psi\rangle$$

$$\leq \langle\psi|(I - B) \otimes I |\psi\rangle \qquad \text{(because } B \text{ is a sub-measurement)}$$

$$= 1 - \langle\psi| B \otimes I |\psi\rangle$$

$$= 1 - \sum_a \langle\psi| B_a \otimes I |\psi\rangle$$

$$\leq 1 - \sum_a \langle\psi| B_a^2 \otimes I |\psi\rangle. \qquad (16)$$

Now, by Proposition 4.25 and the fact that $A_a \otimes I \approx_\delta B_a \otimes I$,

$$\sum_a \langle\psi| B_a^2 \otimes I |\psi\rangle \approx_{\sqrt\delta} \sum_a \langle\psi|(A_a \cdot B_a) \otimes I |\psi\rangle$$

$$\approx_{\sqrt\delta} \sum_a \langle\psi| A_a^2 \otimes I |\psi\rangle$$

$$\geq \langle\psi| A \otimes I |\psi\rangle - \zeta \qquad \text{(by Proposition 4.41)}$$

$$= 1 - \zeta. \qquad \text{(because } A \text{ is a measurement)}$$

Hence, by Equation (16),

$$\|(I - B) \otimes I |\psi\rangle\|^2 \leq 1 - \sum_a \langle\psi| B_a^2 \otimes I |\psi\rangle \leq 2\sqrt\delta + \zeta.$$

This concludes the proof. $\qquad\square$

5 Making measurements projective

A recurring theme in MIP^* research is that projective measurements are significantly easier to manipulate than general POVM measurements. As just one example among many, when $A = \{A_a^x\}$ and $B = \{B_a^x\}$ are projective measurements, the "$\approx_\delta$" and "$\simeq_\delta$" distances are equivalent. In other words:

$$A_a^x \otimes I \simeq_\delta I \otimes B_a^x \qquad \Longleftrightarrow \qquad A_a^x \otimes I \approx_{2\delta} I \otimes B_a^x.$$

This is Fact 4.13 from [NW19]. On the other hand when A and B are not projective, one can find examples of measurements where $A_a^x \otimes I \approx_\delta I \otimes B_a^x$ for $\delta \to 0$ but one can show $A_a^x \otimes I \simeq_\epsilon I \otimes B_a^x$ only for $\epsilon \to 1$; this is Remark 4.15 from [NW19]. As a result, it is important to be able to convert POVM measurements to projective measurements whenever possible.

In this section, we will survey two tools for doing so. The first of these is the textbook Naimark dilation theorem. In our setting, it states the following.

Theorem 5.1 (Naimark dilation). *Let $|\psi\rangle$ be a state in $\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B}$. Let $A = \{A_a^x\}$ be a sub-measurement acting on $\mathcal{H}_\mathrm{A}$ and $B = \{B_b^y\}$ be a sub-measurement acting on $\mathcal{H}_\mathrm{B}$. Then there exists*

1. *Hilbert spaces $\mathcal{H}_{\mathrm{A}_\mathrm{aux}}$ and $\mathcal{H}_{\mathrm{B}_\mathrm{aux}}$,*

2. *a state $|\mathsf{aux}\rangle \in \mathcal{H}_{\mathrm{A}_\mathrm{aux}} \otimes \mathcal{H}_{\mathrm{B}_\mathrm{aux}}$,*

3. *and two measurements $\widehat{A} = \{\widehat{A}_a^x\}$ and $\widehat{B} = \{\widehat{B}_b^y\}$ acting on $\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_{\mathrm{A}_\mathrm{aux}}$ and $\mathcal{H}_\mathrm{B} \otimes \mathcal{H}_{\mathrm{B}_\mathrm{aux}}$, respectively,*

such that the following is true. If we write $|\widehat{\psi}\rangle = |\psi\rangle \otimes |\mathsf{aux}\rangle$, then for all x, y, a, b,

$$\langle \psi | A_a^x \otimes B_b^y | \psi \rangle = \langle \widehat{\psi} | \widehat{A}_a^x \otimes \widehat{B}_b^y | \widehat{\psi} \rangle.$$

In addition, $|\mathsf{aux}\rangle$ is a product state, meaning that we can write it as

$$|\mathsf{aux}\rangle = |\mathsf{aux}_\mathrm{A}\rangle \otimes |\mathsf{aux}_\mathrm{B}\rangle,$$

for $|\mathsf{aux}_\mathrm{A}\rangle$ in $\mathcal{H}_{\mathrm{A}_\mathrm{aux}}$ and $|\mathsf{aux}_\mathrm{B}\rangle$ in $\mathcal{H}_{\mathrm{B}_\mathrm{aux}}$.

The second of these is the "orthogonalization lemma" from [KV11]. In the setting of symmetric strategies, it states the following.

Theorem 5.2 (Orthogonalization lemma). *Let $|\psi\rangle$ be a permutation-invariant state, and let $A = \{A_a\}$ be a sub-measurement with strong self-consistency*

$$\sum_a \langle \psi | A_a \otimes A_a | \psi \rangle \geq \sum_a \langle \psi | A_a \otimes I | \psi \rangle - \zeta.$$

Then there exists a projective sub-measurement $P = \{P_a\}$ such that

$$A_a \otimes I \approx_{100\zeta^{1/4}} P_a \otimes I.$$

These two results serve largely the same purpose, and for our applications it will typically suffice to use either one. That said, there are tradeoffs when choosing to use one or the other. Naimark dilation is convenient because it can be applied to any family of sub-measurements and it preserves measurement outcome probabilities exactly. On the other hand, it requires swapping out the state $|\psi\rangle$ and measurements A and B for $|\widehat{\psi}\rangle$ and $\widehat{A}$ and $\widehat{B}$, respectively, which can be notationally

cumbersome. Most works opt to skip adding the hats, and instead say something to the effect of "using Naimark, we can assume that A and B are projective." However, there is subtlety in doing so, which is that Naimark dilation does not necessarily preserve "$\approx_\delta$" statements; in other words, $A_a^x \otimes I \approx_\delta I \otimes B_a^x$ does not necessarily imply $\widehat{A}_a^x \otimes I \approx_\delta I \otimes \widehat{B}_a^x$. (As an example, Example 5.4 below provides a case in which $A_a^x \otimes I \approx_0 I \otimes B_a^x$, but $\widehat{A}_a^x \otimes I \approx_\delta I \otimes \widehat{B}_a^x$ only holds for $\delta \geq 1$.) This can lead to trouble if the same notation is used for A, B and $\widehat{A}, \widehat{B}$, as "$\approx_\delta$" statements derived before applying Naimark might not necessarily hold after applying Naimark. (That said, as pointed out at the end of Section 4.4 of [NW19], "$\approx_\delta$" statements are often derived as a consequence of "$\simeq_\delta$" statements. And as Naimark preserves "$\simeq_\delta$" statements, one could use them to simply rederive any "$\approx_\delta$" statements post-Naimark.)

The downsides of using the orthogonalization lemma are that it can only be applied to measurements which are strongly self-consistent and it introduces additional error. As we will see below, its proof is significantly more complicated than Naimark dilation. On the plus side, it has none of the notational baggage that Naimark brings, and so it is more concrete to use. In addition, it can lead to stronger results, as it does not require introducing an auxiliary state $|\text{aux}\rangle$.

In this work, we will opt to use the orthogonalization lemma rather than Naimark dilation. However, we will include proofs of both for completeness. We will prove Theorem 5.1 in Section 5.1 and then Theorem 5.2 in Section 5.2.

5.1 Naimark dilation

To prove Theorem 5.1, we will first need the following lemma.

Lemma 5.3. *Let $A = \{A_a\}$ be a sub-measurement with k distinct outcomes $a \in \mathcal{A}$, and let $|\text{aux}\rangle \in \mathbb{C}^{k+1}$ be any state. Then there exists a projective sub-measurement $\widehat{A} = \{\widehat{A}_a\}$ such that for each outcome a,*

$$(I \otimes \langle \text{aux}|) \cdot \widehat{A}_a \cdot (I \otimes |\text{aux}\rangle) = A_a.$$

Proof. We will consider an orthonormal basis for $\mathbb{C}^{k+1}$ consisting of a vector $|a\rangle$ for each $a \in \mathcal{A}$ and the vector $|\perp\rangle$. Let U be any unitary such that for each vector $|\psi\rangle \in \mathbb{C}^d$,

$$U \cdot (|\psi\rangle \otimes |\text{aux}\rangle) = \sum_{a \in \mathcal{A}} ((A_a)^{1/2} |\psi\rangle) \otimes |a\rangle + ((1 - A)^{1/2} |\psi\rangle) \otimes |\perp\rangle$$

$$= \Big(\sum_{a \in \mathcal{A}} (A_a)^{1/2} \otimes |a\rangle + (1 - A)^{1/2} \otimes |\perp\rangle \Big) \cdot |\psi\rangle.$$

This implies that

$$U \cdot (I \otimes |\text{aux}\rangle) = \sum_{a \in \mathcal{A}} (A_a)^{1/2} \otimes |a\rangle + (1 - A)^{1/2} \otimes |\perp\rangle.$$

Hence, for any $a \in \mathcal{A}$,

$$(I \otimes \langle a|) \cdot U \cdot (I \otimes |\text{aux}\rangle) = (I \otimes \langle a|) \cdot \Big(\sum_{a \in \mathcal{A}} (A_a)^{1/2} \otimes |a\rangle + (1 - A)^{1/2} \otimes |\perp\rangle \Big) = (A_a)^{1/2}.$$

Then the desired projective sub-measurement is

$$\widehat{A}_a = U^\dagger \cdot (I \otimes |a\rangle \langle a|) \cdot U.$$

34

This is because

$$(I \otimes \langle \mathsf{aux}|) \cdot \widehat{A}_a \cdot (I \otimes |\mathsf{aux}\rangle) = (I \otimes \langle \mathsf{aux}|) \cdot (U^\dagger \cdot (I \otimes |a\rangle \langle a|) \cdot U) \cdot (I \otimes |\mathsf{aux}\rangle)$$
$$= (I \otimes \langle \mathsf{aux}|) \cdot U^\dagger \cdot (I \otimes |a\rangle) \cdot (I \otimes \langle a|) \cdot U \cdot (I \otimes |\mathsf{aux}\rangle)$$
$$= (A_a)^{1/2} \cdot (A_a)^{1/2} = A_a.$$

This completes the proof. $\qquad\qquad\square$

Now we prove Theorem 5.1.

Proof of Theorem 5.1. For each x, let k be the number of outcomes in A^x, and let $|\mathsf{aux}_{\mathrm{A},x}\rangle$ be a state of dimensionality $k + 1$. Let $\widetilde{A}^x$ be the sub-measurement guaranteed by Lemma 5.3. Define $|\mathsf{aux}_{\mathrm{B},x}\rangle$ and $\widetilde{B}^x$ similarly. We define

$$|\widehat{\psi}\rangle = |\psi\rangle \otimes \big(\otimes_x |\mathsf{aux}_{\mathrm{A},x}\rangle \big) \otimes \big(\otimes_x |\mathsf{aux}_{\mathrm{B},x}\rangle \big)$$

and

$$\widehat{A}_a^x = \widetilde{A}_a^x \otimes \big(\otimes_{z \neq x} I_{\mathsf{aux}_{\mathrm{A},z}} \big)$$
$$\widehat{B}_b^y = \widetilde{B}_b^y \otimes \big(\otimes_{z \neq y} I_{\mathsf{aux}_{\mathrm{B},z}} \big).$$

Then for all x, y, z, b,

$$\langle \widehat{\psi}| \, \widehat{A}_a^x \otimes \widehat{B}_b^y \, |\widehat{\psi}\rangle = \langle \psi| \otimes \langle \mathsf{aux}_{\mathrm{A},x}| \otimes \langle \mathsf{aux}_{\mathrm{B},y}| \cdot \widetilde{A}_a^x \otimes \widetilde{B}_b^y \cdot |\psi\rangle \otimes |\mathsf{aux}_{\mathrm{A},x}\rangle \otimes |\mathsf{aux}_{\mathrm{B},y}\rangle$$
$$= \langle \psi| \, A_a^x \otimes B_b^y \, |\psi\rangle \,.$$

This completes the proof. $\qquad\qquad\square$

Example 5.4 (Naimark does not preserve "$\approx_\delta$"). We now carry out a simple example in which Naimark preserves "$\simeq_\delta$" statements without preserving "$\approx_\delta$" statements. Let $|\psi\rangle$ be an arbitrary state in $\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B}$, where $\mathcal{H}_\mathrm{A} = \mathcal{H}_\mathrm{B} = \mathbb{C}^d$. In addition, let $A = \{A_0, A_1\}$ and $B = \{B_0, B_1\}$ be the two-outcome measurements in which $A_0 = A_1 = B_0 = B_1 = \frac{1}{2} \cdot I_{d \times d}$. Then

$$\sum_{a \neq b} \langle \psi| \, A_a \otimes B_b \, |\psi\rangle = \langle \psi| \, A_0 \otimes B_1 \, |\psi\rangle + \langle \psi| \, A_1 \otimes B_0 \, |\psi\rangle$$

$$= \frac{1}{4} \cdot \langle \psi| \, I \otimes I \, |\psi\rangle + \frac{1}{4} \cdot \langle \psi| \, I \otimes I \, |\psi\rangle = \frac{1}{2},$$

and so $A_a \otimes I \simeq_{1/2} I \otimes A_a$. In addition,

$$\sum_a \| (A_a \otimes I - I \otimes B_a) |\psi\rangle \|^2 = \sum_a \left\| \left(\frac{1}{2} \cdot I \otimes I - \frac{1}{2} \cdot I \otimes I \right) |\psi\rangle \right\|^2 = 0,$$

and so $A_a \otimes I \approx_0 I \otimes B_a$.

Now we carry out the steps of Naimark dilation. Because A and B are two-outcome, we will set $\mathcal{H}_{\mathrm{A}_{\mathsf{aux}}} = \mathcal{H}_{\mathrm{B}_{\mathsf{aux}}} = \mathbb{C}^2$, spanned by the basis vectors $|0\rangle$ and $|1\rangle$. (Lemma 5.3 actually asks that the auxiliary space have dimension $2 + 1 = 3$, where the third dimension is spanned by the basis vector $|\bot\rangle$. However, this is unnecessary for this example because A and B are measurements rather than sub-measurements.) We will choose our two local auxiliary states to be

$$|\mathsf{aux}_\mathrm{A}\rangle = |\mathsf{aux}_\mathrm{B}\rangle = |+\rangle = \frac{1}{\sqrt{2}} |0\rangle + \frac{1}{\sqrt{2}} |1\rangle \,.$$

We first construct $\widehat{A}$ by following the proof of Lemma 5.3. To begin, it asks for U to be a unitary such that for any $|\phi\rangle$ in $\mathcal{H}_A$,

$$
\begin{aligned}
U \cdot (|\phi\rangle \otimes |\mathsf{aux_A}\rangle) &= U \cdot (|\phi\rangle \otimes |+\rangle) \\
&= ((A_0)^{1/2} |\phi\rangle) \otimes |0\rangle + ((A_1)^{1/2} |\phi\rangle) \otimes |1\rangle \\
&= \left(\left(\frac{1}{2} \cdot I \right)^{1/2} |\phi\rangle \right) \otimes |0\rangle + \left(\left(\frac{1}{2} \cdot I \right)^{1/2} |\phi\rangle \right) \otimes |1\rangle \\
&= \frac{1}{\sqrt{2}} |\phi\rangle \otimes |0\rangle + \frac{1}{\sqrt{2}} |\phi\rangle \otimes |1\rangle \\
&= |\phi\rangle \otimes |+\rangle .
\end{aligned}
$$

As a result, we can simply take U to be the identity matrix. Thus,

$$
\widehat{A}_a = U^\dagger \cdot (I \otimes |a\rangle \langle a|) \cdot U = I \otimes |a\rangle \langle a| .
$$

By a similar argument, we can take $\widehat{B}_a = I \otimes |a\rangle \langle a|$ for $a \in \{0, 1\}$ as well.

Now we set $|\widehat{\psi}\rangle = |\psi\rangle \otimes |\mathsf{aux_A}\rangle \otimes |\mathsf{aux_B}\rangle$. By Theorem 5.1, we already know that $\widehat{A}_a \otimes I \approx_0 I \otimes \widehat{B}_a$ on state $|\widehat{\phi}\rangle$ because this holds for the un-hatted state and measurements. On the other hand, we will now show that $\widehat{A}_a \otimes I \approx_\delta I \otimes \widehat{B}_a$ only for $\delta \geq 1$, in contrast to the un-hatted case where it holds for $\delta = 0$. To see this, we calculate as follows.

$$
\begin{aligned}
&\sum_a \| (\widehat{A}_a \otimes I_{\mathsf{B,aux_B}} - I_{\mathsf{A,aux_A}} \otimes \widehat{B}_a) |\widehat{\psi}\rangle \|^2 \\
&= \sum_a \| ((I_A \otimes |a\rangle \langle a|) \otimes I_{\mathsf{B,aux_B}} - I_{\mathsf{A,aux_A}} \otimes (I_B \otimes |a\rangle \langle a|)) |\psi\rangle \otimes |+\rangle \otimes |+\rangle \|^2 \\
&= \sum_a \left\| \left(\frac{1}{\sqrt{2}} |\psi\rangle \otimes |a\rangle \otimes |+\rangle - \frac{1}{\sqrt{2}} |\psi\rangle \otimes |+\rangle \otimes |a\rangle \right) \right\|^2 \\
&= \sum_a \frac{1}{2} \cdot \langle \psi \mid \psi \rangle \cdot (\langle a| \otimes \langle +| - \langle +| \otimes \langle a|) \cdot (|a\rangle \otimes |+\rangle - |+\rangle \otimes |a\rangle). \quad (17)
\end{aligned}
$$

For each a, we have that

$$
\begin{aligned}
&(\langle a| \otimes \langle +| - \langle +| \otimes \langle a|) \cdot (|a\rangle \otimes |+\rangle - |+\rangle \otimes |a\rangle) \\
&= 2 - \langle a \mid + \rangle \cdot \langle + \mid a \rangle - \langle + \mid a \rangle \cdot \langle a \mid + \rangle \\
&= 2 - \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \\
&= 1.
\end{aligned}
$$

Plugging this into Equation (17), we arrive at our bound.

5.2 Orthogonalization lemma

To show Theorem 5.2, we first show it for the case when A is a measurement, rather than a submeasurement. Our proof of this case will even work in the more general setting when the strategy is not assumed to be symmetric, meaning that $|\psi\rangle$ is not necessarily permutation-invariant and Player B's measurement B may not be equal to Player A's measurement. This is the content of the following lemma, whose proof we defer till later.

Lemma 5.5 (Orthogonalization lemma for measurements)**.** *Let $|\psi\rangle$ be a state (which is not neces-sarily permutation-invariant), and let $A = \{A_a\}$ and $B = \{B_a\}$ be measurements such that*

$$A_a \otimes I \simeq_\zeta I \otimes B_a$$

on state $|\psi\rangle$. Then there exists a projective sub-measurement $P = \{P_a\}$ such that

$$A_a \otimes I \approx_{84\zeta^{1/4}} P_a \otimes I.$$

(Note that when $|\psi\rangle$ is permutation-invariant and $B = A$, the condition $A_a \otimes I \simeq_\zeta I \otimes B_a$ is equivalent to A_a being ζ-strongly self-consistent, by Proposition 4.35. This is because A is a measurement.)

We now prove Theorem 5.2 by reducing the general (sub-measurement) case to the case when A is a measurement.

Proof of Theorem 5.2 assuming Lemma 5.5. Let $A = \{A_a\}$ be a sub-measurement with outcomes $a \in \mathcal{A}$ whose strong self-consistency is

$$\sum_a \langle\psi| A_a \otimes A_a |\psi\rangle \geq \sum_a \langle\psi| A_a \otimes I |\psi\rangle - \zeta.$$

Note that

$$\langle\psi| A \otimes A |\psi\rangle = \sum_{a,b} \langle\psi| A_a \otimes A_b |\psi\rangle \geq \sum_a \langle\psi| A_a \otimes A_a |\psi\rangle \geq \langle\psi| A \otimes I |\psi\rangle - \zeta.$$

Rearranging,

$$\langle\psi| A \otimes (I - A) |\psi\rangle = \langle\psi| A \otimes I |\psi\rangle - \langle\psi| A \otimes A |\psi\rangle$$
$$\leq (\langle\psi| A \otimes A |\psi\rangle + \zeta) - \langle\psi| A \otimes A |\psi\rangle = \zeta.$$

As a result,

$$\langle\psi| (I - A) \otimes (I - A) |\psi\rangle = \langle\psi| (I - A) \otimes I |\psi\rangle - \langle\psi| (I - A) \otimes A |\psi\rangle$$
$$\geq \langle\psi| (I - A) \otimes I |\psi\rangle - \zeta.$$

Let $\widehat{A}$ be the POVM measurement with outcomes in $\widehat{\mathcal{A}} = \mathcal{A} \cup \{\bot\}$ defined as

$$\widehat{A}_a = \begin{cases} A_a & \text{if } a \in \mathcal{A}, \\ (I - A) & \text{if } a = \bot. \end{cases}$$

Then the strong self-consistency of $\widehat{A}$ is

$$\sum_{a \in \widehat{\mathcal{A}}} \langle\psi| \widehat{A}_a \otimes \widehat{A}_a |\psi\rangle = \sum_{a \in \mathcal{A}} \langle\psi| A_a \otimes A_a |\psi\rangle + \langle\psi| \widehat{A}_\bot \otimes \widehat{A}_\bot |\psi\rangle$$
$$\geq \left(\sum_{a \in \mathcal{A}} \langle\psi| A_a \otimes I |\psi\rangle - \zeta \right) + \langle\psi| (I - A) \otimes (I - A) |\psi\rangle$$
$$\geq \left(\langle\psi| A \otimes I |\psi\rangle - \zeta \right) + \left(\langle\psi| (I - A) \otimes I |\psi\rangle - \zeta \right)$$
$$= 1 - 2\zeta.$$

Because $\widehat{A}$ is a measurement, Proposition 4.35 states that this is equivalent to

$$\widehat{A}_a \otimes I \simeq_{2\zeta} I \otimes \widehat{A}_a.$$

As a result, Lemma 5.5 implies the existence of a projective sub-measurement $\widehat{P} = \{\widehat{P}_a\}_{a \in \widehat{\mathcal{A}}}$ such that

$$\widehat{A}_a \otimes I \approx_{84 \cdot (2\zeta)^{1/4}} \widehat{P}_a \otimes I.$$

If we define the projective sub-measurement $P = \{P_a\}_{a \in \mathcal{A}}$ by $P_a = \widehat{P}_a$ for all $a \in \mathcal{A}$, then

$$\sum_{a \in \mathcal{A}} \| (A_a - P_a) \otimes I \, |\psi\rangle \|^2 \leq \sum_{a \in \widehat{\mathcal{A}}} \| (\widehat{A}_a - \widehat{P}_a) \otimes I \, |\psi\rangle \|^2 \leq 84 \cdot (2\zeta)^{1/4}.$$

Thus, $A_a \otimes I \approx_{84 \cdot (2\zeta)^{1/4}} P_a \otimes I$. We conclude the proof by noting that $84 \cdot (2\zeta)^{1/4} \leq 100 \cdot \zeta^{1/4}$ because $84 \cdot (2)^{1/4} \approx 99.89 \leq 100$. $\qquad\square$

Now we prove Lemma 5.5.

Proof of Lemma 5.5. We note that the lemma as stated is trivial when $\zeta > 1/4$. As a result, we will assume that

$$\zeta \leq 1/4. \tag{18}$$

Let $A = \{A_a\}$ and $B = \{B_a\}$ be POVM measurements such that

$$A_a \otimes I \simeq_{\zeta} I \otimes B_a.$$

By Proposition 4.19, this implies that

$$\sum_a \langle\psi| A_a \otimes B_a |\psi\rangle \geq 1 - \zeta.$$

Applying Cauchy-Schwarz, we have

$$\sum_a \langle\psi| A_a \otimes B_a |\psi\rangle = \sum_a \langle\psi| (A_a \otimes I) \cdot (I \otimes B_a) |\psi\rangle$$

$$\leq \sqrt{\sum_a \langle\psi| (A_a)^2 \otimes I |\psi\rangle} \cdot \sqrt{\sum_a \langle\psi| I \otimes (B_a)^2 |\psi\rangle}$$

$$\leq \sqrt{\sum_a \langle\psi| (A_a)^2 \otimes I |\psi\rangle \cdot 1}.$$

Taking the square of both sides,

$$\sum_a \langle\psi| (A_a)^2 \otimes I |\psi\rangle \geq \Big(\sum_a \langle\psi| A_a \otimes B_a |\psi\rangle \Big)^2 \geq (1 - \zeta)^2 \geq 1 - 2\zeta = \sum_a \langle\psi| A_a \otimes I |\psi\rangle - 2\zeta.$$

(because A is a measurement)

Rearranging,

$$\sum_a \langle\psi| (A_a - (A_a)^2) \otimes I |\psi\rangle \leq 2\zeta. \tag{19}$$

This is all we need the B measurement for; henceforth, we will derive consequences of Equation (19).

In our next lemma, we convert each A_a to a projective matrix R_a by rounding each of A_a's large eigenvalues to 1 and small eigenvalues to 0.

Lemma 5.6 (Rounding to projectors). *There exists a set of projective matrices $\{R_a\}$ such that*

$$A_a \otimes I \approx_{2\sqrt{\zeta}} R_a \otimes I.$$

and

$$R := \sum_a R_a \leq (1 + 2\sqrt{\zeta}) \cdot I.$$

Proof. For each a, we write the eigendecomposition of A_a as follows:

$$A_a = \sum_i \lambda_{a,i} \cdot |u_{a,i}\rangle \langle u_{a,i}|.$$

Let

$$0 < \delta \leq 1/2 \tag{20}$$

be a number to be decided later. Let $\mathsf{trunc}_\delta : [0,1] \to \{0,1\}$ be the truncation function defined as

$$\mathsf{trunc}_\delta(x) = \begin{cases} 1 & \text{if } x \geq 1 - \delta, \\ 0 & \text{otherwise.} \end{cases}$$

Then for each a we define the matrix R_a

$$R_a = \mathsf{trunc}_\delta(A_a) = \sum_i \mathsf{trunc}_\delta(\lambda_{a,i}) \cdot |u_{a,i}\rangle \langle u_{a,i}|.$$

To analyze this, we will require the following technical lemma.

Lemma 5.7. *For any $x \in [0,1]$,*

$$(x - \mathsf{trunc}_\delta(x))^2 \leq \frac{1}{\delta} \cdot (x - x^2).$$

Proof. This is proved by case analysis. First, suppose that $x \geq 1 - \delta$. This implies that $\mathsf{trunc}_\delta(x) = 1$. In addition, because $\delta \leq 1/2$ by Equation (20), $x \geq 1 - \delta \geq 1/2 \geq \delta$. Thus,

$$\begin{aligned} (x - \mathsf{trunc}_\delta(x))^2 &= (1-x)^2 \\ &\leq (1-x) \\ &\leq (1-x) \cdot \frac{x}{\delta} \qquad \text{(because } x \geq \delta\text{)} \\ &= \frac{1}{\delta} \cdot (x - x^2). \end{aligned}$$

Next, suppose that $x < 1 - \delta$. This implies that $\mathsf{trunc}_\delta(x) = 0$. Thus,

$$\begin{aligned} (x - \mathsf{trunc}_\delta(x))^2 &= x^2 \\ &\leq x \\ &\leq x \cdot \frac{(1-x)}{\delta} \qquad \text{(because } x \leq 1 - \delta\text{)} \\ &= \frac{1}{\delta} \cdot (x - x^2). \end{aligned}$$

This concludes the proof. $\qquad \square$

As a result, for each a, Lemma 5.7 implies that

$$(A_a - R_a)^2 = (A_a - \mathsf{trunc}_\delta(A_a))^2 \leq \frac{1}{\delta} \cdot (A_a - (A_a)^2).$$

Thus,

$$\sum_a \|(A_a - R_a) \otimes I \, |\psi\rangle\|^2 = \sum_a \langle\psi| \, (A_a - R_a)^2 \otimes I \, |\psi\rangle \leq \frac{1}{\delta} \cdot \sum_a \langle\psi| \, (A_a - (A_a)^2) \otimes I \, |\psi\rangle$$

$$\leq \frac{1}{\delta} \cdot 2\zeta. \qquad \text{(by Equation (19))}$$

This implies that $A_a \otimes I \approx_{2\zeta/\delta} R_a \otimes I$.

Next, for each $x \in [0, 1]$, it follows from the definition of trunc_δ that

$$\mathsf{trunc}_\delta(x) \leq \left(\frac{1}{1-\delta}\right) \cdot x.$$

Thus, for each a

$$R_a = \mathsf{trunc}_\delta(A_a) \leq \left(\frac{1}{1-\delta}\right) \cdot A_a.$$

Summing over all a,

$$R = \sum_a R_a \leq \left(\frac{1}{1-\delta}\right) \cdot \sum_a A_a = \left(\frac{1}{1-\delta}\right) \cdot A = \left(\frac{1}{1-\delta}\right) \cdot I,$$

because A is a measurement. We note that

$$\frac{1}{1-\delta} = \frac{1}{1-\delta} \cdot \frac{1+2\delta}{1+2\delta} = \frac{1+2\delta}{1+\delta-2\delta^2} \leq 1 + 2\delta,$$

because

$$1 + \delta - 2\delta^2 = 1 + \delta \cdot (1 - 2\delta) \geq 1$$

when $\delta \leq 1/2$, which we assumed in Equation (20). Hence,

$$R \leq (1 + 2\delta) \cdot I.$$

The lemma now follows by setting $\delta = \sqrt{\zeta}$. Note that we required that δ be at most $1/2$, which follows from our assumption that $\zeta \leq 1/4$ from Equation (18). $\qquad \square$

Write d for the dimension of the A matrices. If the $\{R_a\}$ matrices from Lemma 5.6 formed a projective sub-measurement, then their total rank would be at most d. Even if this is not true, the next lemma shows that we can still post-process them to reduce their total rank to at most d.

Lemma 5.8 (Rank reduction). *There exists a set of projection matrices $\{Q_a\}$ such that*

$$A_a \otimes I \approx_{12\sqrt{\zeta}} Q_a \otimes I.$$

and

$$Q := \sum_a Q_a \leq (1 + 2\sqrt{\zeta}) \cdot I.$$

Furthermore, Q has bounded total rank:

$$\sum_a \mathrm{rank}(Q_a) \leq d.$$

Proof. Let $\{R_a\}$ be the set of projective matrices given by Lemma 5.6. For each a, let r_a be the rank of R_a. Let $r = \sum_a r_a$. If $r \leq d$, then the lemma is trivially satisfied by taking Q to be R and applying Lemma 5.6. Henceforth, we will assume that $r > d$. We want to reduce r so that it is at most d. Fortunately, it is already not too much larger than d:

$$r = \sum_a r_a = \sum_a \operatorname{tr}(R_a) = \operatorname{tr}(R) \leq (1 + 2\sqrt{\zeta}) \cdot \operatorname{tr}(I) = (1 + 2\sqrt{\zeta}) \cdot d. \tag{21}$$

Let $|v_{a,1}\rangle, \ldots, |v_{a,r_a}\rangle$ be an orthonormal basis for the range of R_a, so that

$$R_a = \sum_{i=1}^{r_a} |v_{a,i}\rangle \langle v_{a,i}|.$$

To reduce r, we will throw out those $|v_{a,i}\rangle$'s whose overlap with $|\psi\rangle$ is small. For each a and $1 \leq i \leq r_a$, we denote the overlap of $|v_{a,i}\rangle$ and $|\psi\rangle$ by

$$o_{a,i} = \langle \psi | \cdot (|v_{a,i}\rangle \langle v_{a,i}| \otimes I) \cdot |\psi\rangle.$$

The total overlap is given by

$$\begin{aligned}
\sum_a \sum_{i=1}^{r_a} o_{a,i} &= \sum_a \sum_{i=1}^{r_a} \langle \psi | \cdot (|v_{a,i}\rangle \langle v_{a,i}| \otimes I) \cdot |\psi\rangle \\
&= \sum_a \langle \psi | R_a \otimes I | \psi\rangle \\
&= \langle \psi | R \otimes I | \psi\rangle \\
&\leq (1 + 2\sqrt{\zeta}) \cdot \langle \psi | I \otimes I | \psi\rangle \\
&= 1 + 2\sqrt{\zeta}.
\end{aligned} \tag{22}$$

Now we define Large to be the set of large overlaps:

$$\mathsf{Large} = \{(a,i) \mid o_{a,i} \text{ is among the } d \text{ largest of the } o_{b,j}\text{'s}\},$$

where we break ties arbitrarily, and we define Small to be the set containing the remaining (a,i)'s. We note that Large is well-defined and has size d because $r > d$. Hence, Equation (21) implies that

$$|\mathsf{Small}| = r - |\mathsf{Large}| = r - d \leq (1 + 2\sqrt{\zeta}) \cdot d - d = 2\sqrt{\zeta} \cdot d \leq 2\sqrt{\zeta} \cdot r. \tag{23}$$

Thus, the small (a,i)'s have small total overlap:

$$\begin{aligned}
\sum_{(a,i) \in \mathsf{Small}} o_{a,i} &\leq \frac{|\mathsf{Small}|}{r} \sum_{a,i} o_{a,i} \leq 2\sqrt{\zeta} \cdot \sum_{a,i} o_{a,i} && \text{(by Equation (23))} \\
&\leq 2\sqrt{\zeta} \cdot (1 + 2\sqrt{\zeta}) && \text{(by Equation (22))} \\
&\leq 4\sqrt{\zeta}, && \text{(24)}
\end{aligned}$$

where the final step uses the assumption that $\zeta \leq 1/4$ from Equation (18).

For each a we let Large_a to be set of i's such that (a,i) is contained in Large, and we define Small_a similarly. We define the matrix

$$Q_a = \sum_{i \in \mathsf{Large}_a} |v_{a,i}\rangle \langle v_{a,i}|.$$

Then clearly

$$\sum_a \text{rank}(Q_a) = \sum_a |\mathsf{Large}_a| = |\mathsf{Large}| \le d.$$

We can compute the difference

$$R_a - Q_a = \sum_{i=1}^{r_a} |v_{a,i}\rangle \langle v_{a,i}| - \sum_{i \in \mathsf{Large}_a} |v_{a,i}\rangle \langle v_{a,i}| = \sum_{i \in \mathsf{Small}_a} |v_{a,i}\rangle \langle v_{a,i}| .$$

This is a projective Hermitian matrix, which implies that $Q_a \le R_a$. As a result,

$$Q = \sum_a Q_a \le \sum_a R_a = R \le (1 + 2\sqrt{\zeta}) \cdot I.$$

In addition,

$$
\begin{aligned}
\sum_a \| (R_a - Q_a) \otimes I |\psi\rangle \|^2 &= \sum_a \langle\psi| (R_a - Q_a)^2 \otimes I |\psi\rangle \\
&= \sum_a \langle\psi| (R_a - Q_a) \otimes I |\psi\rangle && \text{(because } R_a - Q_a \text{ is a projector)} \\
&= \sum_a \sum_{i \in \mathsf{Small}_a} \langle\psi| \cdot (|v_{a,i}\rangle \langle v_{a,i}| \otimes I) \cdot |\psi\rangle \\
&= \sum_a \sum_{i \in \mathsf{Small}_a} o_{a,i} \\
&\le 4\sqrt{\zeta}. && \text{(by Equation (24))}
\end{aligned}
$$

This means that

$$R_a \otimes I \approx_{4\sqrt{\zeta}} Q_a \otimes I.$$

Since we know that $A_a \otimes I \approx_{2\sqrt{\zeta}} R_a \otimes I$ by Lemma 5.6, Proposition 4.28 implies that

$$A_a \otimes I \approx_{12\sqrt{\zeta}} Q_a \otimes I.$$

This completes the proof. $\qquad\square$

Henceforth, we let $Q = \{Q_a\}$ be the set of projective matrices given by Lemma 5.8. We now derive a few properties of Q that follow as a consequence of Lemma 5.8. To begin, we show that Q is almost as complete as A.

Lemma 5.9 (Completeness of Q).

$$\langle\psi| Q \otimes I |\psi\rangle \ge 1 - 11\zeta^{1/4}.$$

Proof. To begin, we claim that

$$
\begin{aligned}
\langle\psi| Q \otimes I |\psi\rangle &= \sum_a \langle\psi| Q_a \otimes I |\psi\rangle \\
&= \sum_a \langle\psi| (Q_a)^2 \otimes I |\psi\rangle && \text{(because the } Q_a\text{'s are projective)} \\
&\approx_{5\zeta^{1/4}} \sum_a \langle\psi| (Q_a \cdot A_a) \otimes I |\psi\rangle . && (25)
\end{aligned}
$$

To show this, we bound the magnitude of the difference.

$$\left| \sum_a \langle\psi| \left(Q_a \otimes I\right) \cdot \left(\left(Q_a - A_a\right) \otimes I\right) |\psi\rangle \right|$$

$$\leq \sqrt{\sum_a \langle\psi| \left(Q_a\right)^2 \otimes I |\psi\rangle} \cdot \sqrt{\sum_a \langle\psi| \left(Q_a - A_a\right)^2 \otimes I |\psi\rangle}$$

$$\leq \sqrt{1 + 2\sqrt{\zeta}} \cdot \sqrt{12\sqrt{\zeta}} \qquad\qquad \text{(by Lemma 5.8)}$$

$$\leq \sqrt{2} \cdot \sqrt{12\sqrt{\zeta}},$$

where the last line uses the assumption that $\zeta \leq 1/4$ from Equation (18). Next, we claim that

$$(25) = \sum_a \langle\psi| \left(Q_a \cdot A_a\right) \otimes I |\psi\rangle \approx_{4\zeta^{1/4}} \sum_a \langle\psi| \left(A_a\right)^2 \otimes I |\psi\rangle. \qquad (26)$$

To show this, we bound the magnitude of the difference.

$$\left| \sum_a \langle\psi| \left(\left(Q_a - A_a\right) \otimes I\right) \cdot \left(A_a \otimes I\right) |\psi\rangle \right|$$

$$\leq \sqrt{\sum_a \langle\psi| \left(Q_a - A_a\right)^2 \otimes I |\psi\rangle} \cdot \sqrt{\sum_a \langle\psi| \left(A_a\right)^2 \otimes I |\psi\rangle}$$

$$\leq \sqrt{12\sqrt{\zeta}} \cdot 1. \qquad\qquad \text{(by Lemma 5.8)}$$

In conclusion,

$$\langle\psi| Q \otimes I |\psi\rangle \geq \sum_a \langle\psi| \left(A_a\right)^2 \otimes I |\psi\rangle - 9\zeta^{1/4} \qquad \text{(by Equations (25) and (26))}$$

$$\geq \sum_a \langle\psi| A_a \otimes I |\psi\rangle - 2\zeta - 9\zeta^{1/4} \qquad \text{(by Equation (19))}$$

$$\geq \langle\psi| A \otimes I |\psi\rangle - 11\zeta^{1/4}$$

$$= 1 - 11\zeta^{1/4}.$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

We will also need the following bound on the completeness of the *square root* of Q. Note that such a bound follows trivially from Lemma 5.9 when Q is a sub-measurement because $\sqrt{Q} \geq Q$ when $Q \leq I$.

Lemma 5.10 (Completeness of $\sqrt{Q}$).

$$\langle\psi| \sqrt{Q} \otimes I |\psi\rangle \geq 1 - 12\zeta^{1/4}.$$

Proof. Let $Q = \sum_i \nu_i |u_i\rangle \langle u_i|$ be the eigendecomposition of Q. Then because $Q \leq (1 + 2\sqrt{\zeta}) \cdot I$, each eigenvalue ν_i is at most $1 + 2\sqrt{\zeta}$. Thus,

$$\sqrt{Q} = \sum_i \sqrt{\nu_i} |u_i\rangle \langle u_i| \geq \frac{1}{\sqrt{1 + 2\sqrt{\zeta}}} \cdot \sum_i \nu_i |u_i\rangle \langle u_i| = \frac{1}{\sqrt{1 + 2\sqrt{\zeta}}} \cdot Q.$$

We note that

$$\frac{1}{\sqrt{1+2\sqrt{\zeta}}} \geq \frac{1}{\sqrt{1+2\sqrt{\zeta}+\zeta}} = \frac{1}{1+\sqrt{\zeta}} = \frac{1}{1+\sqrt{\zeta}} \cdot \left(\frac{1-\sqrt{\zeta}}{1-\sqrt{\zeta}}\right) = \frac{1-\sqrt{\zeta}}{1-\zeta} \geq 1-\sqrt{\zeta}.$$

Hence, $\sqrt{Q} \geq (1-\sqrt{\zeta}) \cdot Q$. As a result, Lemma 5.9 implies that

$$\langle\psi|\sqrt{Q}\otimes I|\psi\rangle \geq (1-\sqrt{\zeta}) \cdot \langle\psi|Q\otimes I|\psi\rangle \geq (1-\sqrt{\zeta}) \cdot \left(1-11\zeta^{1/4}\right) \geq 1-\sqrt{\zeta}-11\zeta^{1/4}.$$

This concludes the proof. □

Finally, we show the following lemma, which quantifies a sense in which Q is "almost projective".

Lemma 5.11 (Q is almost projective)**.**

$$\sum_a (Q_a \cdot Q \cdot Q_a - Q_a) \leq 4\sqrt{\zeta} \cdot I.$$

Proof. Lemma 5.8 implies that $Q \leq (1+2\sqrt{\zeta}) \cdot I$. As a result,

$$\sum_a Q_a \cdot Q \cdot Q_a - \sum_a Q_a$$
$$\leq (1+2\sqrt{\zeta}) \cdot \sum_a Q_a \cdot I \cdot Q_a - \sum_a Q_a$$
$$= (1+2\sqrt{\zeta}) \cdot \sum_a Q_a - \sum_a Q_a \qquad \text{(because the } Q_a\text{'s are projectors)}$$
$$= (1+2\sqrt{\zeta}) \cdot Q - Q$$
$$= 2\sqrt{\zeta} \cdot Q$$
$$\leq 2\sqrt{\zeta} \cdot (1+2\sqrt{\zeta}) \cdot I$$
$$\leq 2\sqrt{\zeta} \cdot 2 \cdot I. \qquad \text{(by Equation (18))}$$

This completes the proof. □

We now arrive at the most important definition in this proof, which is a natural matrix decomposition for the Q_a matrices.

Definition 5.12 (Matrix decomposition of Q_a)**.** For each a, let m_a be the rank of Q_a. Let $|v_{a,1}\rangle, \ldots, |v_{a,m_a}\rangle$ be an orthonormal basis for the range of Q_a, so that

$$Q_a = \sum_{i=1}^{m_a} |v_{a,i}\rangle\langle v_{a,i}|.$$

Let $m = \sum_a m_a$, and consider an orthonormal basis of $\mathbb{C}^m$ consisting of vectors $|a,i\rangle$ for each a and $1 \leq i \leq m_a$. For each a, define the matrix

$$X_a = \sum_{i=1}^{m_a} |a,i\rangle\langle v_{a,i}|.$$

In addition, define the matrix

$$X = \sum_a X_a = \sum_a \sum_{i=1}^{m_a} |a,i\rangle\langle v_{a,i}|. \tag{27}$$

Finally, we let $T = \{T_a\}$ be the projective measurement on $\mathbb{C}^m$ defined as

$$T_a = \sum_{i=1}^{m_a} |a,i\rangle \langle a,i| \,.$$

The next pair of lemmas will prove some basic properties of the X_a matrices.

Lemma 5.13. *For each* a, $X_a = T_a \cdot X$.

Proof. This is a simple calculation:

$$T_a \cdot X = \Big(\sum_{i}^{m_a} |a,i\rangle \langle a,i| \Big) \cdot \Big(\sum_a \sum_{i=1}^{m_a} |a,i\rangle \langle v_{a,i}| \Big) = \sum_{i=1}^{m_a} |a,i\rangle \langle v_{a,i}| = X_a. \qquad \square$$

Lemma 5.14 (Q_a restated)**.** *For each* a,

$$Q_a = X_a^\dagger \cdot X_a = X^\dagger \cdot T_a \cdot X = X_a^\dagger \cdot X.$$

Proof. The first equality follows from

$$
\begin{aligned}
X_a^\dagger \cdot X_a &= \Big(\sum_{i=1}^{m_a} |v_{a,i}\rangle \langle a,i| \Big) \cdot \Big(\sum_{j=1}^{m_a} |v_{a,j}\rangle \langle a,j| \Big)^\dagger \\
&= \sum_{i,j=1}^{m_a} |v_{a,i}\rangle \langle a,i| \cdot |a,j\rangle \langle v_{a,j}| = \sum_{i=1}^{m_a} |v_{a,i}\rangle \langle v_{a,i}| = Q_a.
\end{aligned}
$$

The remaining equalities follow from Lemma 5.13 and the fact that T is a projective measurement.

$$X_a^\dagger \cdot X_a = (X^\dagger \cdot T_a) \cdot (T_a \cdot X) = X^\dagger \cdot T_a \cdot X = X_a^\dagger \cdot X. \qquad \square$$

Now we introduce our main tool for studying X, which is via its singular value decomposition.

Definition 5.15 (SVD of X)**.** Let $X = U \cdot \Sigma_{m \times d} \cdot V^\dagger$ be the singular value decomposition (SVD) of X. Because X is an $m \times d$ matrix, the definition of the SVD states that U is an $m \times m$ unitary matrix, V is a $d \times d$ unitary matrix, and $\Sigma_{m \times d}$ is an $m \times d$ diagonal matrix with nonnegative real numbers on its diagonal.

Notation 5.16. For positive integers h and w, we will write $I_{h \times w}$ for the $h \times w$ matrix with 1's on its diagonal and 0's everywhere else.

For integers $h, w \geq m$, we also write $\Sigma_{h \times w}$ for the $h \times w$ diagonal matrix whose diagonal agrees with $\Sigma_{m \times d}$'s; namely, $(\Sigma_{h \times w})_{i,i} = (\Sigma_{m \times d})_{i,i}$ for all $1 \leq i \leq m$ and $(\Sigma_{h \times w})_{i,i} = 0$ for all $i > m$. We note that because $\Sigma_{m \times d}$ is a real-valued diagonal matrix, $(\Sigma_{h \times w})^\dagger = \Sigma_{w \times h}$.

With this definition, we can give a helpful expression for the square of X.

Lemma 5.17 (X squared)**.**

$$X \cdot X^\dagger = U \cdot (\Sigma_{m \times m})^2 \cdot U^\dagger, \quad and \quad X^\dagger \cdot X = Q = V \cdot (\Sigma_{d \times d})^2 \cdot V^\dagger.$$

45

Proof. First,

$$
\begin{aligned}
X \cdot X^\dagger &= (U \cdot \Sigma_{m \times d} \cdot V^\dagger) \cdot (U \cdot \Sigma_{m \times d} \cdot V^\dagger)^\dagger \\
&= U \cdot \Sigma_{m \times d} \cdot V^\dagger \cdot V \cdot \Sigma_{d \times m} \cdot U^\dagger \\
&= U \cdot \Sigma_{m \times d} \cdot I_{d \times d} \cdot \Sigma_{d \times m} \cdot U^\dagger && \text{(because } V \text{ is a } d \times d \text{ unitary)} \\
&= U \cdot (\Sigma_{m \times m})^2 \cdot U^\dagger. && \text{(because } m \le d)
\end{aligned}
$$

Second,

$$
\begin{aligned}
X^\dagger \cdot X &= \sum_a X^\dagger \cdot T_a X && \text{(because } T \text{ is a measurement)} \\
&= \sum_a Q_a && \text{(by Lemma 5.14)} \\
&= Q.
\end{aligned}
$$

In addition, we can rewrite $X^\dagger \cdot X$ as

$$
\begin{aligned}
X^\dagger \cdot X &= (U \cdot \Sigma_{m \times d} \cdot V^\dagger)^\dagger \cdot (U \cdot \Sigma_{m \times d} \cdot V^\dagger) \\
&= V \cdot \Sigma_{d \times m} \cdot U^\dagger \cdot U \cdot \Sigma_{m \times d} \cdot V^\dagger \\
&= V \cdot \Sigma_{d \times m} \cdot I_{m \times m} \cdot \Sigma_{m \times d} \cdot V^\dagger && \text{(because } U \text{ is an } m \times m \text{ unitary)} \\
&= V \cdot (\Sigma_{d \times d})^2 \cdot V^\dagger.
\end{aligned}
$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

The following lemma relates an expression in the X_a's with an expression in the Q_a's that appeared previously in Lemma 5.11.

Lemma 5.18. *For each a,*

$$
X_a^\dagger \cdot (X \cdot X^\dagger - I_{m \times m})^2 \cdot X_a = Q_a \cdot Q \cdot Q_a - Q_a.
$$

Proof. Using Lemma 5.17 and Lemma 5.14,

$$
X_a^\dagger \cdot (X \cdot X^\dagger \cdot X \cdot X^\dagger) \cdot X_a = X_a^\dagger \cdot X \cdot Q \cdot X^\dagger \cdot X_a = Q_a \cdot Q \cdot Q_a.
$$

Similarly, because Q_a is projective,

$$
X_a^\dagger \cdot (X \cdot X^\dagger) \cdot X_a = (X_a^\dagger \cdot X) \cdot (X^\dagger \cdot X_a) = Q_a \cdot Q_a = Q_a.
$$

Finally,

$$
X_a^\dagger \cdot (I_{m \times m}) \cdot X_a = X_a^\dagger \cdot X_a = Q_a.
$$

Putting these together,

$$
\begin{aligned}
& X_a^\dagger \cdot (X \cdot X^\dagger - I_{m \times m})^2 \cdot X_a \\
=\ & X_a^\dagger \cdot (X \cdot X^\dagger \cdot X \cdot X^\dagger - 2 \cdot X \cdot X^\dagger + I_{m \times m}) \cdot X_a \\
=\ & Q_a \cdot Q \cdot Q_a - 2 \cdot Q_a + Q_a \\
=\ & Q_a \cdot Q \cdot Q_a - Q_a.
\end{aligned}
$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

Now we are ready to state the projective sub-measurement P which should approximate A. Before doing so, we give some intuition for the construction.

Remark 5.19. Suppose that the Q_a's actually formed a projective measurement. This would imply that the vectors $|v_{a,i}\rangle$, over all a and $1 \leq i \leq m_a$ form an orthonormal set. Then the SVD would actually have already been provided in Equation (27); for each a and $1 \leq i \leq m_a$, the corresponding singular value would be 1 and the corresponding left- and right-singular vectors would be $|a, i\rangle$ and $|v_{a,i}\rangle$, respectively. In particular, we would have $\Sigma = I_{m \times d}$ and $X = U \cdot I_{m \times d} \cdot V^{\dagger}$.

In reality, we don't know that $\Sigma = I_{m \times d}$. However, we will construct $P = \{P_a\}$ as if it were. This motivates the following definition.

Definition 5.20 (Definition of P)**.** Define the matrix

$$\widehat{X} = U \cdot I_{m \times d} \cdot V^{\dagger}.$$

In addition, for each a, define the matrices

$$\widehat{X}_a = T_a \cdot \widehat{X}, \quad P_a = \widehat{X}_a^{\dagger} \cdot \widehat{X}_a.$$

We now give analogues of Lemmas 5.14 and 5.17 for the P matrices.

Lemma 5.21 (P_a restated)**.** *For each a,*

$$P_a = \widehat{X}^{\dagger} \cdot T_a \cdot \widehat{X} = \widehat{X}_a^{\dagger} \cdot \widehat{X}.$$

Proof. This follows from the definition of $\widehat{X}_a^{\dagger}$ and the fact that T is a projective measurement:

$$P_a = \widehat{X}_a^{\dagger} \cdot \widehat{X}_a = (\widehat{X}^{\dagger} \cdot T_a) \cdot (T_a \cdot \widehat{X}) = \widehat{X}^{\dagger} \cdot T_a \cdot \widehat{X} = \widehat{X}_a^{\dagger} \cdot \widehat{X}. \qquad \square$$

Lemma 5.22 ($\widehat{X}$ squared)**.**
$$\widehat{X} \cdot \widehat{X}^{\dagger} = I_{m \times m}.$$

Proof. First,

$$
\begin{aligned}
\widehat{X} \cdot \widehat{X}^{\dagger} &= (U \cdot I_{m \times d} \cdot V^{\dagger}) \cdot (U \cdot I_{m \times d} \cdot V^{\dagger})^{\dagger} \\
&= U \cdot I_{m \times d} \cdot V^{\dagger} \cdot V \cdot I_{d \times m} \cdot U^{\dagger} \\
&= U \cdot I_{m \times d} \cdot I_{d \times d} \cdot I_{d \times m} \cdot U^{\dagger} && \text{(because V is a $d \times d$ unitary)} \\
&= U \cdot I_{m \times m} \cdot U^{\dagger} && \text{(because $m \leq d$)} \\
&= I_{m \times m},
\end{aligned}
$$

where the last step uses the fact that U is an $m \times m$ unitary. $\qquad \square$

Finally, we show two lemmas on quantities involving both X and $\widehat{X}$.

Lemma 5.23 (X times $\widehat{X}$)**.**

$$X \cdot \widehat{X}^{\dagger} = U \cdot \Sigma_{m \times m} \cdot U^{\dagger}, \quad \text{and} \quad X^{\dagger} \cdot \widehat{X} = \sqrt{Q}.$$

Proof. First,

$$
\begin{aligned}
X \cdot \widehat{X}^\dagger &= (U \cdot \Sigma_{m \times d} \cdot V^\dagger) \cdot (U \cdot I_{m \times d} \cdot V^\dagger)^\dagger \\
&= U \cdot \Sigma_{m \times d} \cdot V^\dagger \cdot V \cdot I_{d \times m} \cdot U^\dagger \\
&= U \cdot \Sigma_{m \times d} \cdot I_{d \times d} \cdot I_{d \times m} \cdot U^\dagger \qquad\qquad \text{(because } V \text{ is a } d \times d \text{ unitary)} \\
&= U \cdot \Sigma_{m \times m} \cdot U^\dagger. \qquad\qquad\qquad\qquad\qquad \text{(because } m \le d)
\end{aligned}
$$

Second,

$$
\begin{aligned}
X^\dagger \cdot \widehat{X} &= (U \cdot \Sigma_{m \times d} \cdot V^\dagger)^\dagger \cdot (U \cdot I_{m \times d} \cdot V^\dagger) \\
&= V \cdot \Sigma_{d \times m} \cdot U^\dagger \cdot U \cdot I_{m \times d} \cdot V^\dagger \\
&= V \cdot \Sigma_{d \times m} \cdot I_{m \times m} \cdot I_{m \times d} \cdot V^\dagger \qquad \text{(because } U \text{ is an } m \times m \text{ unitary)} \\
&= V \cdot \Sigma_{d \times d} \cdot V^\dagger \\
&= \sqrt{V \cdot (\Sigma_{d \times d})^2 \cdot V^\dagger} \\
&= \sqrt{Q}. \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(by Lemma 5.17)}
\end{aligned}
$$

This completes the proof. □

Lemma 5.24 (Squared difference).

$$
(X - \widehat{X}) \cdot (X - \widehat{X})^\dagger \le (X \cdot X^\dagger - I_{m \times m})^2.
$$

Proof. By Lemmas 5.17, 5.22 and 5.23,

$$
\begin{aligned}
(X - \widehat{X}) \cdot (X - \widehat{X})^\dagger &= X \cdot X^\dagger - X \cdot \widehat{X}^\dagger - \widehat{X} \cdot X^\dagger + \widehat{X} \cdot \widehat{X}^\dagger \\
&= U \cdot \Sigma_{m \times m}^2 \cdot U^\dagger - 2 \cdot U \cdot \Sigma_{m \times m} \cdot U^\dagger + I_{m \times m} \\
&= U \cdot \left(\Sigma_{m \times m}^2 - 2 \cdot \Sigma_{m \times m} + I_{m \times m} \right) \cdot U^\dagger \\
&= U \cdot (\Sigma_{m \times m} - I_{m \times m})^2 \cdot U^\dagger.
\end{aligned}
$$

Because $\Sigma_{m \times m}$ and $I_{m \times m}$ are commuting,

$$
(\Sigma_{m \times m} - I_{m \times m})^2 \le (\Sigma_{m \times m} + I_{m \times m})^2 \cdot (\Sigma_{m \times m} - I_{m \times m})^2 = (\Sigma_{m \times m}^2 - I_{m \times m})^2.
$$

As a result,

$$
\begin{aligned}
U \cdot (\Sigma_{m \times m} - I_{m \times m})^2 \cdot U^\dagger &\le U \cdot \left(\Sigma_{m \times m}^2 - I_{m \times m} \right)^2 \cdot U^\dagger \\
&= \left(U \cdot \Sigma_{m \times m}^2 \cdot U^\dagger - I_{m \times m} \right)^2 \\
&= (X \cdot X^\dagger - I_{m \times m})^2. \qquad \text{(by Lemma 5.17)}
\end{aligned}
$$

This completes the proof. □

The first property we need of P is that it is a projective sub-measurement. This is shown in the following lemma.

Lemma 5.25 (Projectivity of P). $P = \{P_a\}$ *forms a projective sub-measurement.*

Proof. Let a, b be (possibly distinct) outcomes. Then

$$
\begin{aligned}
P_a \cdot P_b &= (\widehat{X}^\dagger \cdot T_a \cdot \widehat{X}) \cdot (\widehat{X}^\dagger \cdot T_b \cdot \widehat{X}) && \text{(by Lemma 5.21)} \\
&= (\widehat{X}^\dagger \cdot T_a \cdot I_{m \times m} \cdot T_b \cdot \widehat{X}) && \text{(by Lemma 5.22)} \\
&= (\widehat{X}^\dagger \cdot T_a \cdot \widehat{X}) \cdot \mathbf{1}[a = b] && \text{(because } T \text{ is a projective measurement)} \\
&= P_a \cdot \mathbf{1}[a = b]. && \text{(by Lemma 5.21)}
\end{aligned}
$$

This completes the proof. $\square$

The second property we need of P is that it is close to A. We will first show that it is close to Q.

Lemma 5.26 (*P is close to Q*).
$$
Q_a \otimes I \approx_{30\zeta^{1/4}} P_a \otimes I.
$$

Proof. Our goal is to upper-bound the quantity

$$
\sum_a \langle\psi| (Q_a - P_a)^2 \otimes I |\psi\rangle
$$

$$
= \sum_a \langle\psi| (Q_a)^2 \otimes I |\psi\rangle + \sum_a \langle\psi| (P_a)^2 \otimes I |\psi\rangle - \sum_a \langle\psi| Q_a P_a \otimes I |\psi\rangle - \sum_a \langle\psi| P_a Q_a \otimes I |\psi\rangle
$$

$$
= \langle\psi| Q \otimes I |\psi\rangle + \langle\psi| P \otimes I |\psi\rangle - \sum_a \langle\psi| Q_a P_a \otimes I |\psi\rangle - \sum_a \langle\psi| P_a Q_a \otimes I |\psi\rangle, \tag{28}
$$

where the last step uses the projectivity of Q and P. We bound the four terms in Equation (28) separately. First, by Lemma 5.8,

$$
\langle\psi| Q \otimes I |\psi\rangle \le (1 + 2\sqrt{\zeta}) \cdot \langle\psi| I \otimes I |\psi\rangle \le 1 + 2\sqrt{\zeta}.
$$

Second, because P is a sub-measurement,

$$
\langle\psi| P \otimes I |\psi\rangle \le 1.
$$

The third term and fourth terms are significantly more complicated to bound. As they are complex conjugates of each other, we can write their sum as

$$
\sum_a \langle\psi| Q_a P_a \otimes I |\psi\rangle + \sum_a \langle\psi| P_a Q_a \otimes I |\psi\rangle = 2 \cdot \mathfrak{R}\Big(\sum_a \langle\psi| Q_a P_a \otimes I |\psi\rangle \Big). \tag{29}
$$

We now focus on the expression on the right-hand side of Equation (29). To begin, we use Lemma 5.14 to rewrite it as

$$
\sum_a \langle\psi| Q_a P_a \otimes I |\psi\rangle = \sum_a \langle\psi| ((X_a^\dagger \cdot X) \cdot P_a) \otimes I |\psi\rangle.
$$

The main step will be to show that we can exchange the second X for an $\widehat{X}$, i.e.

$$
\sum_a \langle\psi| (X_a^\dagger \cdot X \cdot P_a) \otimes I |\psi\rangle \approx_{2\zeta^{1/4}} \sum_a \langle\psi| (X_a^\dagger \cdot \widehat{X} \cdot P_a) \otimes I |\psi\rangle. \tag{30}
$$

To show this, we bound the magnitude of the difference.

$$\left| \sum_a \langle \psi | \left((X_a^\dagger \cdot (X - \widehat{X})) \otimes I \right) \cdot (P_a \otimes I) \, | \psi \rangle \right|$$

$$\leq \sqrt{\sum_a \langle \psi | \, (X_a^\dagger \cdot (X - \widehat{X}) \cdot (X - \widehat{X})^\dagger \cdot X_a) \otimes I \, | \psi \rangle} \cdot \sqrt{\sum_a \langle \psi | \, (P_a)^2 \otimes I \, | \psi \rangle}.$$

The expression inside the first square root is

$$\sum_a \langle \psi | \, (X_a^\dagger \cdot (X - \widehat{X}) \cdot (X - \widehat{X})^\dagger \cdot X_a) \otimes I \, | \psi \rangle$$

$$\leq \sum_a \langle \psi | \, (X_a^\dagger \cdot (X \cdot X^\dagger - I_{m \times m})^2 \cdot X_a) \otimes I \, | \psi \rangle \qquad \text{(by Lemma 5.24)}$$

$$= \sum_a \langle \psi | \, (Q_a \cdot Q \cdot Q_a - Q_a) \otimes I \, | \psi \rangle \qquad \text{(by Lemma 5.18)}$$

$$\leq 4\sqrt{\zeta} \cdot \langle \psi | \, I \otimes I \, | \psi \rangle \qquad \text{(by Lemma 5.11)}$$

$$= 4\sqrt{\zeta}.$$

The expression inside the second square root is at most 1 because P is a sub-measurement. Next, we claim that the Equation (30) is in fact a real-valued expression. To see this,

$$(30) = \sum_a \langle \psi | \, (X_a^\dagger \cdot \widehat{X} \cdot P_a) \otimes I \, | \psi \rangle$$

$$= \sum_a \langle \psi | \, (X^\dagger \cdot T_a \cdot \widehat{X} \cdot \widehat{X}^\dagger \cdot T_a \cdot \widehat{X}) \otimes I \, | \psi \rangle \qquad \text{(by Lemmas 5.13 and 5.21)}$$

$$= \sum_a \langle \psi | \, (X^\dagger \cdot T_a \cdot I_{m \times m} \cdot T_a \cdot \widehat{X}) \otimes I \, | \psi \rangle \qquad \text{(by Lemma 5.22)}$$

$$= \sum_a \langle \psi | \, (X^\dagger \cdot T_a \cdot \widehat{X}) \otimes I \, | \psi \rangle$$

$$= \langle \psi | \, (X^\dagger \cdot \widehat{X}) \otimes I \, | \psi \rangle \qquad \text{(because } T \text{ is a measurement)}$$

$$= \langle \psi | \, \sqrt{Q} \otimes I \, | \psi \rangle, \qquad \text{(by Lemma 5.23)}$$

which is real-valued because Q is positive semidefinite. As a result, we have

$$\Re\left(\sum_a \langle \psi | \, Q_a P_a \otimes I \, | \psi \rangle \right) \geq \Re\left(\sum_a \langle \psi | \, (X_a^\dagger \cdot \widehat{X} \cdot P_a) \otimes I \, | \psi \rangle \right) - 2\zeta^{1/4} \qquad \text{(by Equation (30))}$$

$$= \langle \psi | \, \sqrt{Q} \otimes I \, | \psi \rangle - 2\zeta^{1/4}$$

$$\geq (1 - 12\zeta^{1/4}) - 2\zeta^{1/4}. \qquad \text{(by Lemma 5.10)}$$

In total, Equation (29) shows that

$$\sum_a \langle \psi | \, Q_a P_a \otimes I \, | \psi \rangle + \sum_a \langle \psi | \, P_a Q_a \otimes I \, | \psi \rangle \geq 2 \cdot (1 - 14\zeta^{1/4}).$$

Putting everything together, we conclude that

$$(28) \leq (1 + 2\sqrt{\zeta}) + 1 - 2 \cdot (1 - 14\zeta^{1/4}) = 2\sqrt{\zeta} + 28\zeta^{1/4} \leq 30\zeta^{1/4}.$$

This completes the proof. $\qquad \square$

Finally, we have that

$$A_a \otimes I \approx_{12\sqrt{\zeta}} Q_a \otimes I \qquad \text{(by Lemma 5.8)}$$
$$\approx_{30\zeta^{1/4}} P_a \otimes I. \qquad \text{(by Lemma 5.26)}$$

Hence, Proposition 4.28 implies that

$$A_a \otimes I \approx_{84\zeta^{1/4}} P_a \otimes I.$$

This completes the proof. $\qquad\qquad\square$

6 The main induction step

We will now carry out the main inductive argument. The inductive hypothesis is stated as follows.

Theorem 6.1 (Main induction). *Let* (ψ, A, B, L) *be an* $(\epsilon, \delta, \gamma)$-*good symmetric strategy for the* (m, q, d) *low individual degree test. Let* $k \geq md$ *be an integer. Then there exists a measurement* $G \in \mathrm{PolyMeas}(m, q, d)$ *such that on average over* $\boldsymbol{u} \sim \mathbb{F}_q^m$,

$$A_a^u \otimes I \simeq_\sigma I \otimes G_{[g(u)=a]},$$

where $\sigma = m^2 \cdot \left(\nu + e^{-k/(80000m^2)}\right)$ *and* $\nu = 1000k^2m^2 \cdot \left(\epsilon^{1/1024} + \delta^{1/1024} + \gamma^{1/1024} + (d/q)^{1/1024}\right)$.

Comparing with our main theorem, Theorem 3.10, Theorem 6.1 produces a measurement which is consistent with A, but it is not projective or self-consistent. In addition, the strategy is assumed to be symmetric. We correct these deficiencies in the following proof.

Proof of Theorem 3.10 assuming Theorem 6.1. Suppose $(\psi, A^{\mathrm{A}}, B^{\mathrm{A}}, L^{\mathrm{A}}, A^{\mathrm{B}}, B^{\mathrm{B}}, L^{\mathrm{B}})$ is a (not necessarily symmetric) strategy which passes the (m, q, d)-low individual degree test with probability $(1 - \epsilon)$. Throughout this proof we will refer to this as the *original strategy*. Then because each of the three subtests occurs with probability $1/3$, $(\psi, A^{\mathrm{A}}, B^{\mathrm{A}}, L^{\mathrm{A}}, A^{\mathrm{B}}, B^{\mathrm{B}}, L^{\mathrm{B}})$ is a $(3\epsilon, 3\epsilon, 3\epsilon)$-good strategy. We now would like to apply Theorem 6.1, but it only applies to strategies which are symmetric. So we will apply a standard construction to "symmetrize" our strategy, apply Theorem 6.1, and then "unsymmetrize" the resulting $\{G_g\}$ measurement to obtain $\{G_g^{\mathrm{A}}\}$ and $\{G_g^{\mathrm{B}}\}$ measurements.

For simplicity, we will assume that Player A and B's Hilbert spaces $\mathcal{H}_{\mathrm{A}}$ and $\mathcal{H}_{\mathrm{B}}$ are both $\mathbb{C}^d$, for some d. (This argument can be extended to the case of different dimensions in a straightforward manner.) We will introduce two additional registers, $\mathcal{H}_{\mathrm{A}'} = \mathcal{H}_{\mathrm{B}'} = \mathbb{C}^2$, referred to as the *role registers*. Then the symmetrized state is given by

$$|\psi_{\mathrm{sym}}\rangle = |0\rangle_{\mathrm{A}'} |1\rangle_{\mathrm{B}'} |\psi\rangle_{\mathrm{A,B}} + |1\rangle_{\mathrm{A}'} |0\rangle_{\mathrm{B}'} |\psi_{\mathrm{swap}}\rangle_{\mathrm{A,B}} \in (\mathbb{C}_{\mathrm{A}'}^2 \otimes \mathbb{C}_{\mathrm{A}}^d) \otimes (\mathbb{C}_{\mathrm{B}'}^2 \otimes \mathbb{C}_{\mathrm{B}}^d),$$

where $|\psi_{\mathrm{swap}}\rangle$ denotes $|\psi\rangle$ with its two registers swapped. Note that the resulting state $|\psi_{\mathrm{sym}}\rangle$ is symmetric under the exchange of its two registers. Next, we define the symmetrized measurement $A_{\mathrm{sym}} = \{(A_{\mathrm{sym}})_a^u\}$ as follows

$$(A_{\mathrm{sym}})_a^u = |0\rangle \langle 0| \otimes A_a^{\mathrm{A},u} + |1\rangle \langle 1| \otimes A_a^{\mathrm{B},u},$$

and we define B_{sym} and L_{sym} similarly. The strategy $(\psi_{\mathrm{sym}}, A_{\mathrm{sym}}, B_{\mathrm{sym}}, L_{\mathrm{sym}})$ is symmetric; we refer to it as the *symmetrized strategy*. It has the following interpretation: the two players measure their respective role registers in the standard basis; the one that receives a "0" will act as Player A

in the original strategy, and the one that receives a "1" will act as Player B in the original strategy. As a result, the symmetrized strategy is also a $(3\epsilon, 3\epsilon, 3\epsilon)$-good strategy.

Now we apply Theorem 6.1 to the symmetrized strategy. To do so, set

$$1000k^2m^2 \cdot \left((3\epsilon)^{1/1024} + (3\epsilon)^{1/1024} + (3\epsilon)^{1/1024} + (d/q)^{1/1024}\right)$$

$$\leq 10000k^2m^2 \cdot \left(\epsilon^{1/1024} + (d/q)^{1/1024}\right) =: \nu,$$

and set $\sigma = m^2 \cdot \left(\nu + e^{-k/(80000m^2)}\right)$. Then Theorem 6.1 produces a measurement $G = \{G_g\} \in \mathrm{PolyMeas}(m, q, d)$ such that

$$(A_{\mathrm{sym}})_a^u \otimes I \simeq_\sigma I \otimes G_{[g(u)=a]}, \quad \text{and} \quad G_{[g(u)=a]} \otimes I \simeq_\sigma I \otimes (A_{\mathrm{sym}})_a^u. \tag{31}$$

Now we unsymmetrize the symmetrized strategy to derive measurements $\{G_g^w\}$ for $w \in \{A, B\}$. Letting I_d denote the $d \times d$ identity operator, define the operators

$$G_g^A = (\langle 0| \otimes I_d) \cdot G_g \cdot (|0\rangle \otimes I_d)$$
$$G_g^B = (\langle 1| \otimes I_d) \cdot G_g \cdot (|1\rangle \otimes I_d),$$

where $|0\rangle$ and $|1\rangle$ act on the $\mathbb{C}^2$ part of G_g. Thus, G_g^A, G_g^B are positive operators acting on $\mathbb{C}^d$, and furthermore they form POVMs:

$$\sum_g G_g^A = \sum_g (\langle 0| \otimes I_d) \cdot G_g \cdot (|0\rangle \otimes I_d) = (\langle 0| \otimes I_d) \cdot \left(\sum_g G_g\right) \cdot (|0\rangle \otimes I_d) = I_d .$$

The same derivation holds for G_g^B.

Next we verify that the $\{G_g^B\}$ measurements are consistent with the $\{A_a^{A,u}\}$ measurements:

$$\mathbf{E}_{\boldsymbol{u}} \sum_{g,a \neq g(\boldsymbol{u})} \langle\psi| A_a^{A,\boldsymbol{u}} \otimes G_g^B |\psi\rangle$$

$$= \mathbf{E}_{\boldsymbol{u}} \sum_{g,a \neq g(\boldsymbol{u})} \langle\psi| A_a^{A,\boldsymbol{u}} \otimes (\langle 1| \otimes I_d) \cdot G_g \cdot (|1\rangle \otimes I_d) |\psi\rangle$$

$$= \mathbf{E}_{\boldsymbol{u}} \sum_{g,a \neq g(\boldsymbol{u})} (\langle 0, 1|_{A'B'} \otimes \langle\psi|_{AB}) \cdot (A_{\mathrm{sym}})_a^{\boldsymbol{u}} \otimes G_g \cdot (|0, 1\rangle_{A'B'} \otimes |\psi\rangle_{AB})$$

$$\leq 2 \cdot \mathbf{E}_{\boldsymbol{u}} \sum_{g,a \neq g(\boldsymbol{u})} \langle\psi_{\mathrm{sym}}| (A_{\mathrm{sym}})_a^{\boldsymbol{u}} \otimes G_g |\psi_{\mathrm{sym}}\rangle$$

$$\leq 2\sigma. \qquad\qquad\qquad\qquad\qquad\qquad\qquad\text{(by Equation (31))}$$

The first inequality follows from the fact that the cross-terms $\langle 0| (A_{\mathrm{sym}})_a^u |1\rangle$ and $\langle 1| (A_{\mathrm{sym}})_a^u |0\rangle$ vanish by construction of A_{sym}. Combined with a similar derivation for the G_g^A and $A_a^{B,u}$ operators, we deduce

$$G_{[g(u)=a]}^A \otimes I \simeq_{2\sigma} I \otimes A_a^{B,u} \tag{32}$$

$$I \otimes G_{[g(u)=a]}^B \simeq_{2\sigma} A_a^{A,u} \otimes I \tag{33}$$

In addition, because the original strategy is $(3\epsilon, 3\epsilon, 3\epsilon)$-good,

$$A_a^{A,u} \otimes I \simeq_{3\epsilon} I \otimes A_a^{B,u}.$$

Hence, Proposition 4.29 implies that

$$G^{\mathrm{A}}_{[g(u)=a]} \otimes I \simeq_{2\sigma+2\sqrt{3\epsilon+2\sigma}} I \otimes G^{\mathrm{B}}_{[g(u)=a]}.$$

This implies that

$$
\begin{aligned}
2\sigma + 2\sqrt{3\epsilon + 2\sigma} &\geq \mathop{\mathbf{E}}_{\bm{u}} \sum_{a\neq b} \langle\psi|\, G^{\mathrm{A}}_{[g(\bm{u})=a]} \otimes G^{\mathrm{B}}_{[g(\bm{u})=b]} \,|\psi\rangle \\
&= \mathop{\mathbf{E}}_{\bm{u}} \sum_{g\neq h} \mathbf{1}[g(\bm{u}) \neq h(\bm{u})] \cdot \langle\psi|\, G^{\mathrm{A}}_g \otimes G^{\mathrm{B}}_h \,|\psi\rangle \\
&= \mathop{\mathbf{E}}_{\bm{u}} \sum_{g\neq h} \langle\psi|\, G^{\mathrm{A}}_g \otimes G^{\mathrm{B}}_h \,|\psi\rangle - \mathop{\mathbf{E}}_{\bm{u}} \sum_{g\neq h} \mathbf{1}[g(\bm{u}) = h(\bm{u})] \cdot \langle\psi|\, G^{\mathrm{A}}_g \otimes G^{\mathrm{B}}_h \,|\psi\rangle \\
&\geq \mathop{\mathbf{E}}_{\bm{u}} \sum_{g\neq h} \langle\psi|\, G^{\mathrm{A}}_g \otimes G^{\mathrm{B}}_h \,|\psi\rangle - \sum_{g\neq h} \frac{md}{q} \cdot \langle\psi|\, G^{\mathrm{A}}_g \otimes G^{\mathrm{B}}_h \,|\psi\rangle \qquad \text{(by Schwartz-Zippel)} \\
&\geq \mathop{\mathbf{E}}_{\bm{u}} \sum_{g\neq h} \langle\psi|\, G^{\mathrm{A}}_g \otimes G^{\mathrm{B}}_h \,|\psi\rangle - \frac{md}{q}.
\end{aligned}
$$

Rearranging, we get

$$G^{\mathrm{A}}_g \otimes I \simeq_{\zeta_1} I \otimes G^{\mathrm{B}}_g. \tag{34}$$

where $\zeta_1 = 2\sigma + 2\sqrt{3\epsilon + 2\sigma} + md/q$.

We have now derived everything we wanted, except that G is not necessarily projective. To remedy this, we apply the orthogonalization lemma for measurements (Lemma 5.5) to Equation (34). It implies the existence of two projective sub-measurements $P^{\mathrm{A}} = \{P^{\mathrm{A}}\}, P^{\mathrm{B}} = \{P^{\mathrm{B}}\} \in \mathrm{PolySub}(m,q,d)$ such that

$$
\begin{aligned}
G^{\mathrm{A}}_g \otimes I &\approx_{100\zeta_1^{1/4}} P^{\mathrm{A}}_g \otimes I, \\
I \otimes G^{\mathrm{B}}_g &\approx_{100\zeta_1^{1/4}} I \otimes P^{\mathrm{B}}_g.
\end{aligned}
$$

Hence, Proposition 4.40 implies that we can complete P^{A} and P^{B} to projective measurements $Q^{\mathrm{A}} = \{Q^{\mathrm{A}}\}, Q^{\mathrm{B}} = \{Q^{\mathrm{B}}\} \in \mathrm{PolyMeas}(m,q,d)$ such that

$$
\begin{aligned}
G^{\mathrm{A}}_g \otimes I &\approx_{\zeta_2} Q^{\mathrm{A}}_g \otimes I, \\
I \otimes G^{\mathrm{B}}_g &\approx_{\zeta_2} I \otimes Q^{\mathrm{B}}_g.
\end{aligned} \tag{35}
$$

where $\zeta_2 = 200\zeta_1^{1/4} + 40\zeta_1^{1/8}$. Now, Equation (34) and Proposition 4.22 imply that

$$G^{\mathrm{A}}_g \otimes I \approx_{2\zeta_1} I \otimes G^{\mathrm{B}}_g.$$

By the triangle inequality (Proposition 4.28),

$$Q^{\mathrm{A}}_g \otimes I \approx_{\zeta_3} I \otimes Q^{\mathrm{B}}_g,$$

where $\zeta_3 = 6\zeta_1 + 6\zeta_2$. Because both of these measurements are projective, Proposition 4.22 then implies that

$$Q^{\mathrm{A}}_g \otimes I \simeq_{\zeta_3/2} I \otimes Q^{\mathrm{B}}_g, \tag{36}$$

By the data processing inequality (Proposition 4.30),

$$Q^{\mathrm{A}}_{[g(u)=a]} \otimes I \simeq_{\zeta_3/2} I \otimes Q^{\mathrm{B}}_{[g(u)=a]}, \tag{37}$$

Next, Proposition 4.21, applied to Equation (34) and Equation (35) implies that

$$Q_g^{\mathrm{A}} \otimes I \simeq_{\zeta_1} I \otimes G_g^{\mathrm{B}}.$$

By data processing (Proposition 4.30),

$$Q_{[g(u)=a]}^{\mathrm{A}} \otimes I \simeq_{\zeta_1} I \otimes G_{[g(u)=a]}^{\mathrm{B}}. \tag{38}$$

Now we apply the triangle inequality (Proposition 4.29) to Equations (33), (37) and (38), which implies that

$$A_a^{\mathrm{A},u} \otimes I \simeq_{\zeta_4} I \otimes Q_{[g(u)=a]}^{\mathrm{B}}, \tag{39}$$

where $\zeta_4 = 2\sigma + 2\sqrt{\zeta_1 + \zeta_3/2}$. A similar argument shows that

$$I \otimes A_a^{\mathrm{B},u} \simeq_{\zeta_4} Q_{[g(u)=a]}^{\mathrm{A}} \otimes I. \tag{40}$$

Now we calculate the error. First,

$$\sigma = m^2 \cdot \left(10000k^2m^2 \cdot \left(\epsilon^{1/1024} + (d/q)^{1/1024}\right) + e^{-k/(80000m^2)}\right)$$
$$\leq 10000k^2m^4 \cdot \left(\epsilon^{1/1024} + (d/q)^{1/1024} + e^{-k/(80000m^2)}\right).$$

Next, using $10000^{1/2} = 100$,

$$\zeta_1 = 2\sigma + 2\sqrt{3\epsilon + 2\sigma} + md/q$$
$$\leq 2 \cdot \left(10000k^2m^4 \cdot \left(\epsilon^{1/1024} + (d/q)^{1/1024} + e^{-k/(80000m^2)}\right)\right)$$
$$\quad + 2 \cdot \left(3\epsilon + 10000k^2m^4 \cdot \left(\epsilon^{1/1024} + (d/q)^{1/1024} + e^{-k/(80000m^2)}\right)\right)^{1/2} + md/q$$
$$\leq 2 \cdot \left(10000k^2m^4 \cdot \left(\epsilon^{1/1024} + (d/q)^{1/1024} + e^{-k/(80000m^2)}\right)\right)$$
$$\quad + 2 \cdot \left(2\epsilon^{1/2} + 100km^2 \cdot \left(\epsilon^{1/2048} + (d/q)^{1/2048} + e^{-k/(160000m^2)}\right)\right) + md/q$$
$$\leq 20204k^2m^4 \cdot \left(\epsilon^{1/2048} + (d/q)^{1/2048} + e^{-k/(160000m^2)}\right).$$

Next, using the fact that $20204^{1/4} \leq 12$ and $20204^{1/8} \leq 4$,

$$\zeta_2 = 200\zeta_1^{1/4} + 40\zeta_1^{1/8}$$
$$\leq 200\left(20204k^2m^4 \cdot \left(\epsilon^{1/2048} + (d/q)^{1/2048} + e^{-k/(160000m^2)}\right)\right)^{1/4}$$
$$\quad + 40\left(20204k^2m^4 \cdot \left(\epsilon^{1/2048} + (d/q)^{1/2048} + e^{-k/(160000m^2)}\right)\right)^{1/8}$$
$$\leq 200\left(12km \cdot \left(\epsilon^{1/8192} + (d/q)^{1/8192} + e^{-k/(640000m^2)}\right)\right)$$
$$\quad + 40\left(4km \cdot \left(\epsilon^{1/16384} + (d/q)^{1/16384} + e^{-k/(1280000m^2)}\right)\right)$$
$$\leq 2560km \cdot \left(\epsilon^{1/16384} + (d/q)^{1/16384} + e^{-k/(1280000m^2)}\right).$$

Next, using $6 \cdot 20204 + 6 \cdot 2560 \leq 150000$,

$$\zeta_3 = 6\zeta_1 + 6\zeta_2 \leq 150000k^2m^4 \cdot \left(\epsilon^{1/16384} + (d/q)^{1/16384} + e^{-k/(1280000m^2)}\right).$$

Next, using $\sqrt{20204} \le 143$, $\sqrt{150000} \le 388$, and $2 \cdot (10000 + 143 + 388) \le 40000$,

$$
\begin{aligned}
\zeta_4 &= 2\sigma + 2\sqrt{\zeta_1 + \zeta_3/2} \\
&\le 2 \cdot \left(10000k^2m^4 \cdot \left(\epsilon^{1/1024} + (d/q)^{1/1024} + e^{-k/(80000m^2)} \right) \right) \\
&\quad + 2 \cdot \left(20204k^2m^4 \cdot \left(\epsilon^{1/2048} + (d/q)^{1/2048} + e^{-k/(160000m^2)} \right) \right. \\
&\qquad \left. + 150000k^2m^4 \cdot \left(\epsilon^{1/16384} + (d/q)^{1/16384} + e^{-k/(1280000m^2)} \right) \right)^{1/2} \\
&\le 2 \cdot \left(10000k^2m^4 \cdot \left(\epsilon^{1/1024} + (d/q)^{1/1024} + e^{-k/(80000m^2)} \right) \right) \\
&\quad + 2 \cdot \left(143km^2 \cdot \left(\epsilon^{1/4096} + (d/q)^{1/4096} + e^{-k/(320000m^2)} \right) \right. \\
&\qquad \left. + 388km^2 \cdot \left(\epsilon^{1/32768} + (d/q)^{1/32768} + e^{-k/(2560000m^2)} \right) \right) \\
&\le 40000k^2m^4 \cdot \left(\epsilon^{1/32768} + (d/q)^{1/32768} + e^{-k/(2560000m^2)} \right) \Big) .
\end{aligned}
$$

Both this and $\zeta_3/2$ are less than

$$
100000k^2m^4 \cdot \left(\epsilon^{1/40000} + (d/q)^{1/40000} + e^{-k/(2560000m^2)} \right) \Big) .
$$

Hence, Equations (36), (39) and (40) provide the three bounds we want. This concludes the proof. $\qquad\square$

The remainder of this section is organized as follows: first, in Section 6.1, we define the two main steps in the proof of Theorem 6.1 known as self-improvement and pasting. Following that, we prove Theorem 6.1 in Section 6.2.

6.1 Self-improvement and pasting

There are two main steps in the proof of Theorem 6.1. The first is self-improvement, which is stated as follows.

Theorem 6.2 (Self-improvement). *Let (ψ, A, B, L) be an $(\epsilon, \delta, \gamma)$-good symmetric strategy for the (m, q, d) low individual degree test. Let $G \in \mathrm{PolySub}(m, q, d)$ be a sub-measurement with the following properties:*

1. (Consistency with A): On average over $\boldsymbol{u} \sim \mathbb{F}_q^m$,

$$
A_a^u \otimes I \simeq_\nu I \otimes G_{[g(u)=a]}.
$$

Let

$$
\zeta = 3000m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + (d/q)^{1/32} \right).
$$

Then there exists a projective sub-measurement $H \in \mathrm{PolySub}(m, q, d)$ with the following properties:

1. (Completeness): If $H = \sum_h H_h$, then

$$
\langle \psi | \, H \otimes I \, | \psi \rangle \ge (1 - \nu) - \zeta.
$$

2. (Consistency with A): On average over $\boldsymbol{u} \sim \mathbb{F}_q^m$,

$$
A_a^u \otimes I \simeq_\zeta I \otimes H_{[h(u)=a]}.
$$

3. *(Strong self-consistency):*

$$H_h \otimes I \approx_\zeta I \otimes H_h.$$

4. *(Boundedness):* *There exists a positive-semidefinite matrix Z such that*

$$\langle\psi| Z \otimes (I - H) |\psi\rangle \leq \zeta$$

and for each $h \in \mathcal{P}(m, q, d)$,

$$Z \geq \left(\mathbf{E}_{\boldsymbol{u}} A^{\boldsymbol{u}}_{h(\boldsymbol{u})}\right).$$

We note that by Proposition 4.36 the condition in Item 3 is equivalent to H's strong self-consistency because H is projective. Self-improvement states that we can take a measurement G whose consistency error with A is ν and produce another measurement H which has negligible consistency error with A and incompleteness ν. Hence, we have "moved" G's consistency error onto H's incompleteness.

The second main step is pasting, which is stated as follows.

Theorem 6.3 (Pasting). *Let (ψ, A, B, L) be an $(\epsilon, \delta, \gamma)$-good symmetric strategy for the $(m + 1, q, d)$ low individual degree test. Let $\{G^x\}_{x \in \mathbb{F}_q}$ denote a set of projective sub-measurements in $\mathrm{PolySub}(m, q, d)$ with the following properties:*

1. *(Completeness):* *If $G = \mathbf{E}_{\boldsymbol{x}} \sum_g G^{\boldsymbol{x}}_g$, then*

$$\langle\psi| G \otimes I |\psi\rangle \geq 1 - \kappa.$$

2. *(Consistency with A):* *On average over $(\boldsymbol{u}, \boldsymbol{x}) \sim \mathbb{F}^{m+1}_q$,*

$$A^{u,x}_a \otimes I \simeq_\zeta I \otimes G^x_{[g(u)=a]}.$$

3. *(Strong self-consistency):* *On average over $\boldsymbol{x} \sim \mathbb{F}_q$,*

$$G^x_g \otimes I \approx_\zeta I \otimes G^x_g.$$

4. *(Boundedness):* *There exists a positive-semidefinite matrix Z^x for each $x \in \mathbb{F}_q$ such that*

$$\mathbf{E}_{\boldsymbol{x}} \langle\psi| (I - G^{\boldsymbol{x}}) \otimes Z^{\boldsymbol{x}} |\psi\rangle \leq \zeta$$

and for each $x \in \mathbb{F}_q$ and $g \in \mathcal{P}(m, q, d)$,

$$Z^x \geq \left(\mathbf{E}_{\boldsymbol{u}} A^{\boldsymbol{u},x}_{g(\boldsymbol{u})}\right).$$

Let $k \geq 400md$ be an integer. Let

$$\nu = 100k^2m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right),$$

$$\sigma = \kappa \cdot \left(1 + \frac{1}{100m}\right) + 2\nu + e^{-k/(80000m^2)}.$$

Then there exists a "pasted" measurement $H \in \mathrm{PolyMeas}(m + 1, q, d)$ which satisfies the following property.

1. *(Consistency with A):* On average over $\boldsymbol{u} \sim \mathbb{F}_q^{m+1}$,

$$A_a^u \otimes I \simeq_\sigma I \otimes H_{[h(u)=a]}.$$

Intuitively, σ should be thought of as being roughly κ, plus a small amount of error which does not depend on κ. Hence, pasting states that we can take a family of sub-measurements $\{G^x\}$ with incompleteness κ and produce a pasted measurement H whose consistency error with A is roughly κ, plus a small amount of new error. Thus, the overall inductive step looks as follows: given a family of measurements with some inconsistency error, we "move" the error into the incompleteness using self-improvement, and then we paste the measurements together to form a single measurement whose error is roughly the same as the original error.

We note that the main error term ν depends only on the "small" parameters ϵ, δ, ζ, γ, and d/q and not on the "large" parameter κ.

6.2 Proof of Theorem 6.1

Definition 6.4. Let $x \in \mathbb{F}_q$. For each line $\ell \in \mathbb{F}_q^m$, we define $\mathrm{append}_x(\ell)$ to be the line in $\mathbb{F}_q^{m+1}$ containing every point (u, x) such that $u \in \ell$. In addition, for each function $f : \ell \to \mathbb{F}_q$, we define $\mathrm{append}_x(f) : \mathrm{append}_x(\ell) \to \mathbb{F}_q$ such that for each $(u, x) \in \mathrm{append}_x(\ell)$, $\mathrm{append}_x(f)(u, x) = f(u)$.

Definition 6.5 (*x*-restricted low-degree strategy). Let (ψ, A, B, L) be a symmetric strategy the $(m+1, q, d)$-low individual degree test. Given $x \in \mathbb{F}_q$, we define the *x*-restricted strategy (ψ, A^x, B^x, L^x) for the (m, q, d)-low individual degree test as follows.

1. For each $u \in \mathbb{F}_q^m$, $(A^x)_a^u = A_a^{u,x}$.

2. For each axis parallel line $\ell \in \mathbb{F}_q^m$, $(B^x)_f^\ell = B_{\mathrm{append}_x(f)}^{\mathrm{append}_x(\ell)}$.

3. For each line $\ell \in \mathbb{F}_q^m$, $(L^x)_f^\ell = L_{\mathrm{append}_x(f)}^{\mathrm{append}_x(\ell)}$.

Lemma 6.6. *Let (ψ, A, B, L) be an $(\epsilon, \delta, \gamma)$-good symmetric strategy for the $(m+1, d, q)$-low individual degree test. For each $x \in \mathbb{F}_q$, let (ψ, A^x, B^x, L^x) be the corresponding x-restricted strategy. In addition, write ϵ_x for the probability that it fails the axis-parallel lines test, δ_x for the probability it fails the self-consistency test, and γ_x for the probability it fails the diagonal lines test. Then*

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \epsilon_{\boldsymbol{x}} \leq \left(\frac{m+1}{m}\right) \cdot \epsilon, \quad \mathop{\mathbf{E}}_{\boldsymbol{x}} \delta_{\boldsymbol{x}} \leq \delta, \quad \mathop{\mathbf{E}}_{\boldsymbol{x}} \gamma_{\boldsymbol{x}} \leq \left(\frac{m+1}{m}\right) \cdot \gamma.$$

Proof. We begin with the self-consistency test. Here, both provers are given a uniformly random point $(\boldsymbol{u}, \boldsymbol{x})$, they measure using $A^{\boldsymbol{u}, \boldsymbol{x}} = (A^{\boldsymbol{x}})^{\boldsymbol{u}}$, and they succeed if their outcomes are the same. This is equivalent to performing the self-consistency test on the $\boldsymbol{x}$-restricted strategy, averaged over $\boldsymbol{x}$, and so $\mathbf{E}_{\boldsymbol{x}} \delta_{\boldsymbol{x}} \leq \delta$. (It is a "$\leq$" rather than an "$=$" because δ is just an upper-bound on the failure probability.)

Next, we consider the axis-parallel lines test. Suppose the provers are sent the line $\boldsymbol{\ell}$ and the point $(\boldsymbol{u}, \boldsymbol{x}) \in \boldsymbol{\ell}$. With probability $\frac{1}{m+1}$, $\boldsymbol{\ell}$ is parallel to the $(m+1)$-st direction. When it is not, then $\boldsymbol{\ell} = \mathrm{append}_{\boldsymbol{x}}(\boldsymbol{\ell}')$, where $\boldsymbol{\ell}'$ is an axis-parallel line in $\mathbb{F}_q^m$. In this case, the points prover measures with the measurement $A^{\boldsymbol{u}, \boldsymbol{x}} = (A^{\boldsymbol{x}})^{\boldsymbol{u}}$ and receives an outcome $\boldsymbol{a}$, the lines prover measures with the measurement $B^{\boldsymbol{\ell}} = (B^{\boldsymbol{x}})^{\boldsymbol{\ell}'}$ and receives an outcome $\boldsymbol{f} = \mathrm{append}_{\boldsymbol{x}}(\boldsymbol{f}')$, and they succeed if $\boldsymbol{f}(\boldsymbol{u}, \boldsymbol{x}) = \boldsymbol{a}$, or, equivalently, if $\boldsymbol{f}'(\boldsymbol{u}) = \boldsymbol{a}$. Hence, the probability that they succeed is equal to

the probability that the $\boldsymbol{x}$-restricted strategy passes the (m, q, d)-low individual degree test, which is $\epsilon_{\boldsymbol{x}}$. As a result,

$$\epsilon \geq \Pr_{\boldsymbol{\ell}, \boldsymbol{u}, \boldsymbol{x}} [A \text{ and } B \text{ succeed given } \boldsymbol{\ell}, (\boldsymbol{u}, \boldsymbol{x})]$$

$$\geq \left(\frac{m}{m+1} \right) \cdot \Pr_{\boldsymbol{\ell}, \boldsymbol{u}, \boldsymbol{x}} [A \text{ and } B \text{ succeed given } \boldsymbol{\ell}, (\boldsymbol{u}, \boldsymbol{x}) \mid \boldsymbol{\ell} \text{ is not parallel to direction } m+1]$$

$$= \left(\frac{m}{m+1} \right) \cdot \mathbf{E}_{\boldsymbol{x}} \, \epsilon_{\boldsymbol{x}}.$$

Finally, the analysis of the diagonal lines test follows the same proof as the axis-parallel lines test, and we omit it here. $\qquad\square$

Proof of Theorem 6.1. We note that the bound we are proving is trivial when at least one of ϵ, δ, γ, or d/q is ≥ 1, as ν is at least 1 in that case. Hence, we may assume that $\epsilon, \delta, \gamma, d/q \leq 1$. This will aid us when carrying out the error calculations, as it allows us to bound terms like $(d/q)^{1/2}$ by terms like $(d/q)^{1/4}$.

The proof is by induction on m. The base case is when $m = 1$. In this case, there is only one axis-parallel line ℓ in $\mathbb{F}_q^m$, and so $B^\ell \in \mathrm{PolyMeas}(m, q, d)$. Because this strategy fails the axis-parallel line test with probability at most ϵ,

$$A_a^u \otimes I \simeq_\epsilon I \otimes B_{[f(u)=a]}^\ell,$$

on average over $\boldsymbol{u} \sim \mathbb{F}_q$. We note that this bound holds independent of the value of k. This is even better than the theorem demands, and so the theorem is proved.

Now we perform the induction step. Assuming that Theorem 6.1 holds for $m \geq 1$, we will show that it holds for $m + 1$ as well. Let (ψ, A, B, L) be an $(\epsilon, \delta, \gamma)$-good symmetric strategy for the $(m + 1, q, d)$ low individual degree test. Let $k \geq (m + 1)d$ be an integer.

For each $x \in \mathbb{F}_q$, let (ψ, A^x, B^x, L^x) be the corresponding x-restricted strategy. In addition, write ϵ_x for the probability that it fails the axis-parallel lines test, δ_x for the probability it fails the self-consistency test, and γ_x for the probability it fails the diagonal lines test.

For each $x \in \mathbb{F}_q$, we apply the inductive hypothesis to the x-restricted strategy with the same integer k. This is possible because $k \geq (m + 1)d \geq md$. Let

$$\nu_x = 1000k^2 m^2 \cdot \left(\epsilon_x^{1/1024} + \delta_x^{1/1024} + \gamma_x^{1/1024} + (d/q)^{1/1024} \right),$$

and

$$\sigma_x = m^2 \cdot \left(\nu_x + e^{-k/(80000m^2)} \right)$$

Then the inductive hypothesis states that there exists a measurement $G^x \in \mathrm{PolyMeas}(m, q, d)$ such that

$$(A^x)_a^u \otimes I \simeq_{\sigma_x} I \otimes G_{[g(u)=a]}^x.$$

Next, we apply self-improvement to each G^x. Let

$$\zeta_x = 3000m \cdot \left(\epsilon_x^{1/32} + \delta_x^{1/32} + (d/q)^{1/32} \right).$$

Then Theorem 6.2 produces a projective sub-measurement $\widehat{G}^x \in \mathrm{PolySub}(m, q, d)$ such that for each $x \in \mathbb{F}_q$, the following statements hold.

1. (Completeness): If $\widehat{G}^x = \sum_g \widehat{G}^x_g$, then

$$\langle\psi|\,\widehat{G}^x \otimes I\,|\psi\rangle \geq (1 - \sigma_x) - \zeta_x.$$

2. (Consistency with A^x): On average over $\boldsymbol{u} \sim \mathbb{F}_q^m$,

$$(A^x)^u_a \otimes I \simeq_{\zeta_x} I \otimes \widehat{G}^x_{[g(u)=a]}.$$

3. (Strong self-consistency):
$$\widehat{G}^x_g \otimes I \approx_{\zeta_x} I \otimes \widehat{G}^x_g.$$

4. (Boundedness): There exists a positive-semidefinite matrix Z^x such that

$$\langle\psi|\,Z^x \otimes (I - \widehat{G}^x)\,|\psi\rangle \leq \zeta_x$$

and for each $g \in \mathcal{P}(m, q, d)$,

$$Z^x \geq \Big(\,\mathop{\mathbf{E}}_{\boldsymbol{u}}(A^x)^{\boldsymbol{u}}_{g(\boldsymbol{u})}\Big).$$

Having produced the $\widehat{G}^x$'s, we would like to paste them together. To do so, we need bounds for the above four properties which are stated on average over $\boldsymbol{x} \sim \mathbb{F}_q$ rather than for each $x \in \mathbb{F}_q$ individually. This involves computing "averaged" versions of our error parameters ν_x, σ_x, and ζ_x. In these derivations, we will crucially use the fact that $\alpha \mapsto \alpha^c$ is concave when $c \leq 1$, and hence $\mathbf{E}(\boldsymbol{\alpha})^c \leq (\mathbf{E}\,\boldsymbol{\alpha})^c$.

$$\begin{aligned}
\mathop{\mathbf{E}}_{\boldsymbol{x}} \nu_{\boldsymbol{x}} &= \mathop{\mathbf{E}}_{\boldsymbol{x}} \Big(1000k^2 m^2 \cdot \big(\epsilon_{\boldsymbol{x}}^{1/1024} + \delta_{\boldsymbol{x}}^{1/1024} + \gamma_{\boldsymbol{x}}^{1/1024} + (d/q)^{1/1024} \big) \Big) \\
&\leq 1000k^2 m^2 \cdot \Big((\mathop{\mathbf{E}}_{\boldsymbol{x}} \epsilon_{\boldsymbol{x}})^{1/1024} + (\mathop{\mathbf{E}}_{\boldsymbol{x}} \delta_{\boldsymbol{x}})^{1/1024} + (\mathop{\mathbf{E}}_{\boldsymbol{x}} \gamma_{\boldsymbol{x}})^{1/1024} + (d/q)^{1/1024} \Big) \quad \text{(by concavity)} \\
&\leq 1000k^2 m^2 \cdot \Big(\big(\tfrac{(m+1)}{m} \cdot \epsilon\big)^{1/1024} + \delta^{1/1024} + \big(\tfrac{(m+1)}{m} \cdot \gamma\big)^{1/1024} + (d/q)^{1/1024} \Big) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(by Lemma 6.6)} \\
&\leq 1000k^2 (m+1)^2 \cdot \Big(\epsilon^{1/1024} + \delta^{1/1024} + \gamma^{1/1024} + (d/q)^{1/1024} \Big).
\end{aligned}$$

We call this value ν. Next, if we define

$$\sigma = m^2 \cdot \Big(\nu + e^{-k/(80000m^2)} \Big),$$

then

$$\sigma \geq m^2 \cdot \Big(\mathop{\mathbf{E}}_{\boldsymbol{x}} \nu_{\boldsymbol{x}} + e^{-k/(80000m^2)} \Big) = \mathop{\mathbf{E}}_{\boldsymbol{x}} \sigma_{\boldsymbol{x}}.$$

Finally,

$$\begin{aligned}
\mathop{\mathbf{E}}_{\boldsymbol{x}} \zeta_{\boldsymbol{x}} &= \mathop{\mathbf{E}}_{\boldsymbol{x}} \Big(3000m \cdot \big(\epsilon_{\boldsymbol{x}}^{1/32} + \delta_{\boldsymbol{x}}^{1/32} + (d/q)^{1/32} \big) \Big) \\
&\leq 3000m \cdot \Big((\mathop{\mathbf{E}}_{\boldsymbol{x}} \epsilon_{\boldsymbol{x}})^{1/32} + (\mathop{\mathbf{E}}_{\boldsymbol{x}} \delta_{\boldsymbol{x}})^{1/32} + (d/q)^{1/32} \Big) \qquad\qquad \text{(by concavity)} \\
&\leq 3000m \cdot \Big(\big(\tfrac{(m+1)}{m} \cdot \epsilon\big)^{1/32} + \delta^{1/32} + (d/q)^{1/32} \Big) \qquad\qquad \text{(by Lemma 6.6)} \\
&\leq 3000(m+1) \cdot \Big(\epsilon^{1/32} + \delta^{1/32} + (d/q)^{1/32} \Big).
\end{aligned}$$

We call this value ζ. We note for later that

$$
\begin{aligned}
\zeta &= 3000(m+1) \cdot \left(\epsilon^{1/32} + \delta^{1/32} + (d/q)^{1/32} \right) \\
&\leq 1000k^2(m+1)^2 \cdot \left(\epsilon^{1/32} + \delta^{1/32} + (d/q)^{1/32} \right) \qquad \text{(because } m \geq 2) \\
&\leq 1000k^2(m+1)^2 \cdot \left(\epsilon^{1/1024} + \delta^{1/1024} + \gamma^{1/1024} + (d/q)^{1/1024} \right) \\
&= \nu.
\end{aligned}
\tag{41}
$$

Having defined these, the following statements hold.

1. (Completeness): If $\widehat{G} = \mathbf{E}_{\boldsymbol{x}} \sum_g \widehat{G}_g^{\boldsymbol{x}}$, then

$$
\langle \psi | \, \widehat{G} \otimes I \, | \psi \rangle \geq (1 - \sigma) - \zeta.
$$

2. (Consistency with A): On average over $(\boldsymbol{u}, \boldsymbol{x}) \sim \mathbb{F}_q^{m+1}$,

$$
A_a^{u,x} \otimes I \simeq_\zeta I \otimes \widehat{G}_{[g(u)=a]}^x.
$$

3. (Strong self-consistency): On average over $\boldsymbol{x} \sim \mathbb{F}_q$,

$$
\widehat{G}_g^x \otimes I \approx_\zeta I \otimes \widehat{G}_g^x.
$$

4. (Boundedness): There exists a positive-semidefinite matrix Z^x for each $x \in \mathbb{F}_q$ such that

$$
\mathbf{E}_{\boldsymbol{x}} \langle \psi | \, Z^{\boldsymbol{x}} \otimes (I - \widehat{G}^{\boldsymbol{x}}) \, | \psi \rangle \leq \zeta
$$

and for each $x \in \mathbb{F}_q$ and $g \in \mathcal{P}(m, q, d)$,

$$
Z^x \geq \left(\mathbf{E}_{\boldsymbol{u}} A_{g(\boldsymbol{u})}^{\boldsymbol{u},x} \right).
$$

We are now ready to apply Theorem 6.3. To do so, we note that because $(3000)^{1/32} \leq 2$ and $32 \cdot 32 = 1024$,

$$
\begin{aligned}
\zeta^{1/32} &= \left(3000(m+1) \cdot \left(\epsilon^{1/32} + \delta^{1/32} + (d/q)^{1/32} \right) \right)^{1/32} \\
&\leq 2(m+1) \cdot \left(\epsilon^{1/1024} + \delta^{1/1024} + (d/q)^{1/1024} \right).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
& 100k^2 m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32} \right) \\
\leq \;\; & 100k^2 m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + 2(m+1) \cdot \left(\epsilon^{1/1024} + \delta^{1/1024} + (d/q)^{1/1024} \right) + (d/q)^{1/32} \right) \\
\leq \;\; & 200k^2 m(m+1) \cdot \left(\epsilon^{1/1024} + \delta^{1/1024} + \gamma^{1/1024} + \epsilon^{1/1024} + \delta^{1/1024} + (d/q)^{1/1024} + (d/q)^{1/1024} \right) \\
\leq \;\; & 1000k^2(m+1)^2 \cdot \left(\epsilon^{1/1024} + \delta^{1/1024} + \gamma^{1/1024} + (d/q)^{1/1024} \right) \\
= \;\; & \nu.
\end{aligned}
$$

Then Theorem 6.3 implies the existence of a pasted measurement $H \in \text{PolySub}(m+1, q, d)$ which satisfies the following property. On average over $\boldsymbol{u} \sim \mathbb{F}_q^{m+1}$,

$$A_a^u \otimes I \simeq_{\sigma^*} I \otimes H_{[h(u)=a]},$$

where

$$\sigma^* = (\sigma + \zeta) \cdot \left(1 + \frac{1}{100m}\right) + 2\nu + e^{-k/(80000m^2)}.$$

The consistency with A is as guaranteed in the theorem statement. Hence, we need only verify that the completeness bound implies the one in the theorem statement as well.

$$\sigma^* = (\sigma + \zeta) \cdot \left(1 + \frac{1}{100m}\right) + 2\nu + e^{-k/(80000m^2)}$$

$$\leq (\sigma + \nu) \cdot \left(1 + \frac{1}{100m}\right) + 2\nu + e^{-k/(80000m^2)} \qquad \text{(by Equation (41))}$$

$$= \left(1 + \frac{1}{100m}\right) \cdot \left(m^2 \cdot \left(\nu + e^{-k/(80000m^2)}\right) + \nu\right) + 2\nu + e^{-k/(80000m^2)}$$

$$\leq \left(1 + \frac{1}{100m}\right) \cdot (m^2 + 3) \cdot \left(\nu + e^{-k/(80000m^2)}\right). \qquad (42)$$

Now, because $m \geq 2$,

$$\frac{1}{100m} \cdot (m^2 + 3) \leq \frac{1}{100m} \cdot (m^2 + 4m - 5) = \frac{1}{100m} \cdot (m-1)(m+5) \leq m - 1 \leq 2(m-1).$$

Hence,

$$\left(1 + \frac{1}{100m}\right) \cdot (m^2 + 3) = m^2 + 3 + \frac{1}{100m} \cdot (m^2 + 3) \leq m^2 + 3 + 2(m-1) = m^2 + 2m + 1 = (m+1)^2.$$

As a result,

$$(42) \leq (m+1)^2 \cdot \left(\nu + e^{-k/(80000m^2)}\right) \leq (m+1)^2 \cdot \left(\nu + e^{-k/(80000(m+1)^2)}\right).$$

This is the bound guaranteed by the theorem and so it completes the proof. $\qquad \square$

7 Expansion in the hypercube graph

Definition 7.1 (Hypercube graph). The *hypercube graph* $C = (V, E)$ is the graph with vertex set $V = \mathbb{F}_q^m$ and an edge between $u, v \in V$ whenever u and v disagree in at most one coordinate (so that every vertex is connected to itself). A random edge in C, denoted $(\boldsymbol{u}, \boldsymbol{v}) \sim C$, is distributed as follows: draw $\boldsymbol{u} \sim \mathbb{F}_q^m$, $\boldsymbol{i} \sim \{1, \ldots, m\}$, and $\boldsymbol{x} \sim \mathbb{F}_q$, all uniformly at random, and set $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{x} \cdot e_{\boldsymbol{i}}$.

We will use M to denote the number of vertices in C, i.e. $M = q^m$.

7.1 Eigenvalues of the hypercube graph

Definition 7.2 (Adjacency matrix). The *normalized adjacency matrix of C* is the matrix K defined as

$$K = \mathop{\mathbf{E}}_{(\boldsymbol{u}, \boldsymbol{v}) \sim C} |\boldsymbol{u}\rangle \langle \boldsymbol{v}|.$$

The *Laplacian of C* is the matrix

$$L = \frac{1}{M} \cdot I - K.$$

The following proposition gives another convenient way of writing the Laplacian of C.

Proposition 7.3. $L = \frac{1}{2} \cdot \underset{(\boldsymbol{u},\boldsymbol{v}) \sim C}{\mathbf{E}} (|\boldsymbol{u}\rangle - |\boldsymbol{v}\rangle) \cdot (\langle \boldsymbol{u}| - \langle \boldsymbol{v}|)$.

Proof. If we draw $(\boldsymbol{u}, \boldsymbol{v}) \sim C$, then both $\boldsymbol{u}$ and $\boldsymbol{v}$ are distributed as uniformly random elements of $\mathbb{F}_q^m$. As a result,

$$\frac{1}{M} \cdot I = \underset{\boldsymbol{u} \in \mathbb{F}_q^m}{\mathbf{E}} |\boldsymbol{u}\rangle \langle \boldsymbol{u}| = \frac{1}{2} \cdot \underset{(\boldsymbol{u},\boldsymbol{v}) \sim C}{\mathbf{E}} |\boldsymbol{u}\rangle \langle \boldsymbol{u}| + |\boldsymbol{v}\rangle \langle \boldsymbol{v}| .$$

In addition, $(\boldsymbol{u}, \boldsymbol{v})$ is distributed identically to $(\boldsymbol{v}, \boldsymbol{u})$. As a result,

$$K = \underset{(\boldsymbol{u},\boldsymbol{v}) \sim C}{\mathbf{E}} |\boldsymbol{u}\rangle \langle \boldsymbol{v}| = \frac{1}{2} \cdot \underset{(\boldsymbol{u},\boldsymbol{v}) \sim C}{\mathbf{E}} |\boldsymbol{u}\rangle \langle \boldsymbol{v}| + |\boldsymbol{v}\rangle \langle \boldsymbol{u}| .$$

Combining these two,

$$L = \frac{1}{2} \cdot \underset{(\boldsymbol{u},\boldsymbol{v}) \sim C}{\mathbf{E}} [|\boldsymbol{u}\rangle \langle \boldsymbol{u}| + |\boldsymbol{v}\rangle \langle \boldsymbol{v}| - |\boldsymbol{u}\rangle \langle \boldsymbol{v}| - |\boldsymbol{v}\rangle \langle \boldsymbol{u}|] = \frac{1}{2} \cdot \underset{(\boldsymbol{u},\boldsymbol{v}) \sim C}{\mathbf{E}} (|\boldsymbol{u}\rangle - |\boldsymbol{v}\rangle) \cdot (\langle \boldsymbol{u}| - \langle \boldsymbol{v}|). \quad \square$$

The most important properties of the adjacency matrix are its eigenvalues and eigenvectors. These are provided in the next proposition, which is standard in the literature.

Proposition 7.4. *For each $\alpha \in \mathbb{F}_q^m$, define the vector*

$$|\varphi_\alpha\rangle := \frac{1}{M^{1/2}} \cdot \sum_{u \in \mathbb{F}_q^m} \omega^{\mathrm{tr}[u \cdot \alpha]} \cdot |u\rangle .$$

Then the following two statements are true.

1. *The $|\varphi_\alpha\rangle$'s form an orthonormal basis of $\mathbb{C}^V$.*

2. *For each $\alpha \in \mathbb{F}_q^m$, $|\varphi_\alpha\rangle$ is an eigenvector for K with eigenvalue $\frac{1}{M} \cdot \frac{m - |\alpha|}{m}$, where $|\alpha|$ is the number of nonzero coordinates in α.*

Proof. First, we prove Item 1. Given $\alpha, \beta \in \mathbb{F}_q^m$,

$$\langle \varphi_\alpha \mid \varphi_\beta \rangle = \frac{1}{M} \sum_{u \in \mathbb{F}_q^m} \omega^{\mathrm{tr}[u \cdot (\beta - \alpha)]} = \begin{cases} 1 & \text{if } \alpha = \beta, \\ 0 & \text{otherwise.} \end{cases} \qquad \text{(by Proposition 4.3)}$$

As a result, the $|\varphi_\alpha\rangle$ vectors form an orthonormal basis of $\mathbb{C}^V$.

Next, we prove Item 2. Given $\alpha \in \mathbb{F}_q^m$,

$$K \cdot |\varphi_\alpha\rangle = \left(\underset{(\boldsymbol{u},\boldsymbol{v}) \sim C}{\mathbf{E}} |\boldsymbol{u}\rangle \langle \boldsymbol{v}| \right) \cdot \left(\frac{1}{M^{1/2}} \cdot \sum_{u \in \mathbb{F}_q^m} \omega^{\mathrm{tr}[u \cdot \alpha]} \cdot |u\rangle \right) = \frac{1}{M^{1/2}} \cdot \underset{(\boldsymbol{u},\boldsymbol{v}) \sim C}{\mathbf{E}} \omega^{\mathrm{tr}[\boldsymbol{v} \cdot \alpha]} |\boldsymbol{u}\rangle . \qquad (43)$$

By definition of a random edge, we can replace $\boldsymbol{v}$ with $\boldsymbol{u} + \boldsymbol{x} \cdot e_{\boldsymbol{i}}$, where $\boldsymbol{i}$ is a uniformly random index in $\{1, \ldots, m\}$ and $\boldsymbol{x}$ is a uniformly random element of $\mathbb{F}_q$. As a result,

$$(43) = \frac{1}{M^{1/2}} \cdot \underset{\boldsymbol{u},\boldsymbol{i},\boldsymbol{x}}{\mathbf{E}} \omega^{\mathrm{tr}[(\boldsymbol{u} + \boldsymbol{x} \cdot e_{\boldsymbol{i}}) \cdot \alpha]} |\boldsymbol{u}\rangle = \left(\underset{\boldsymbol{i},\boldsymbol{x}}{\mathbf{E}} \omega^{\mathrm{tr}[(\boldsymbol{x} \cdot e_{\boldsymbol{i}}) \cdot \alpha]} \right) \cdot \frac{1}{M^{1/2}} \cdot \underset{\boldsymbol{u}}{\mathbf{E}} \omega^{\mathrm{tr}[\boldsymbol{u} \cdot \alpha]} |\boldsymbol{u}\rangle = \frac{1}{M} \left(\underset{\boldsymbol{i},\boldsymbol{x}}{\mathbf{E}} \omega^{\mathrm{tr}[\boldsymbol{x} \cdot \alpha_{\boldsymbol{i}}]} \right) \cdot |\varphi_\alpha\rangle .$$

Hence, $|\varphi_\alpha\rangle$ is an eigenvector of K with eigenvalue

$$\frac{1}{M} \cdot \underset{\boldsymbol{i}}{\mathbf{E}} \left[\underset{\boldsymbol{x}}{\mathbf{E}} \omega^{\mathrm{tr}[\boldsymbol{x} \cdot \alpha_{\boldsymbol{i}}]} \right] = \frac{1}{M} \cdot \underset{\boldsymbol{i}}{\mathbf{E}} \left[\mathbf{1}[\alpha_{\boldsymbol{i}} = 0] \right] = \frac{1}{M} \cdot \frac{m - |\alpha|}{m}. \qquad \text{(by Proposition 4.2)}$$

This concludes the proof. $\qquad \square$

We will use the following corollary of Proposition 7.4.

Corollary 7.5. *Let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$ be the eigenvalues of L. Then $\lambda_1 = 0$ and $\lambda_2 = \frac{1}{mM}$.*

Proof. Let $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_M$ be the eigenvalues of K. Then $\mu_i = \frac{1}{M} - \lambda_i$. Thus, it suffices to show that $\mu_1 = \frac{1}{M}$ and $\mu_2 = \frac{1}{M} \cdot \frac{m-1}{m}$. By Proposition 7.4, $|\varphi_\alpha\rangle$ has eigenvalue $\frac{1}{M}$ when $|\alpha| = 0$ and eigenvalue $\frac{1}{M} \cdot \frac{m-1}{m}$ when $|\alpha| = 1$. $\square$

7.2 Local and global variance

In this section, $|\psi\rangle$ will denote a vector (not necessarily normalized) in $\mathcal{H}_A \otimes \mathcal{H}_B$, and for each $u \in \mathbb{F}_q^m$, $0 \leq A^u \leq I$ will be a matrix acting on $\mathcal{H}_A$.

Definition 7.6. The *local variance of A on $|\psi\rangle$* is defined as

$$\mathbf{Var}_{\text{local}}(A, \psi) := \frac{1}{2} \cdot \underset{(\boldsymbol{u},\boldsymbol{v}) \sim C}{\mathbf{E}} \langle \psi | (A^{\boldsymbol{u}} - A^{\boldsymbol{v}})^2 \otimes I | \psi \rangle.$$

The *global variance of A on $|\psi\rangle$* is defined as

$$\mathbf{Var}_{\text{global}}(A, \psi) := \frac{1}{2} \cdot \underset{\boldsymbol{u},\boldsymbol{v} \sim \mathbb{F}_q^m}{\mathbf{E}} \langle \psi | (A^{\boldsymbol{u}} - A^{\boldsymbol{v}})^2 \otimes I | \psi \rangle.$$

The global variance differs from the local variance because $\boldsymbol{u}, \boldsymbol{v}$ are chosen independently from $\mathbb{F}_q^m$ rather than from the edges of C. A standard fact from spectral graph theory allows us to use the expansion of C to relate these two quantities.

Lemma 7.7. $\mathbf{Var}_{\text{global}}(A, \psi) \leq m \cdot \mathbf{Var}_{\text{local}}(A, \psi)$.

Before proving Lemma 7.7, we will give nice expressions for the local and global variances. To begin, we show how to rewrite the local variance in terms of the Laplacian of C.

Lemma 7.8. *Define the matrix*

$$A_{\text{combine}} = \sum_{u \in \mathbb{F}_q^m} |u\rangle \otimes A^u \otimes I.$$

Then

$$\text{Tr}(A_{\text{combine}}^\dagger \cdot (L \otimes |\psi\rangle \langle\psi|) \cdot A_{\text{combine}}) = \mathbf{Var}_{\text{local}}(A, \psi).$$

Proof. For any $u, v \in \mathbb{F}_q^m$,

$$((\langle u| - \langle v|) \otimes \langle \psi|) \cdot A_{\text{combine}} = ((\langle u| - \langle v|) \otimes \langle \psi|) \cdot \sum_{w \in \mathbb{F}_q^m} |w\rangle \otimes A^w \otimes I$$

$$= \langle \psi| \cdot ((A^u - A^v) \otimes I). \tag{44}$$

As a result,

$$A_{\text{combine}}^\dagger \cdot L \otimes |\psi\rangle \langle\psi| \cdot A_{\text{combine}}$$

$$= \frac{1}{2} \cdot \underset{(\boldsymbol{u},\boldsymbol{v}) \sim C}{\mathbf{E}} A_{\text{combine}}^\dagger ((|\boldsymbol{u}\rangle - |\boldsymbol{v}\rangle) \cdot (\langle\boldsymbol{u}| - \langle\boldsymbol{v}|) \otimes |\psi\rangle \langle\psi|) \cdot A_{\text{combine}} \qquad \text{(by Proposition 7.3)}$$

$$= \frac{1}{2} \cdot \underset{(\boldsymbol{u},\boldsymbol{v}) \sim C}{\mathbf{E}} ((A^{\boldsymbol{u}} - A^{\boldsymbol{v}}) \otimes I) \cdot |\psi\rangle \langle\psi| \cdot ((A^{\boldsymbol{u}} - A^{\boldsymbol{v}}) \otimes I). \qquad \text{(by (44))}$$

Thus, if we take the trace,

$$\text{Tr}(A_{\text{combine}}^\dagger \cdot (L \otimes |\psi\rangle \langle \psi|) \cdot A_{\text{combine}}) = \frac{1}{2} \cdot \mathop{\mathbf{E}}_{(\boldsymbol{u},\boldsymbol{v})\sim C} \langle \psi | (A^{\boldsymbol{u}} - A^{\boldsymbol{v}})^2 \otimes I \cdot |\psi\rangle = \mathbf{Var}_{\text{local}}(A, \psi).$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Next, we give a simple expression for the global variance.

Lemma 7.9. *Expand A_{combine} as*

$$A_{\text{combine}} = |\varphi_0\rangle \otimes A_0 + |\varphi_\perp\rangle \otimes A_\perp,$$

where $|\varphi_\perp\rangle$ is orthogonal to $|\varphi_0\rangle$. (Here we are writing $|\varphi_0\rangle$ for the vector $|\varphi_\alpha\rangle$ from Proposition 7.4 in the case of $\alpha = (0, \ldots, 0)$.) Then

$$\frac{1}{M} \cdot \text{Tr}(\langle \varphi_\perp| \otimes A_\perp \cdot (I \otimes |\psi\rangle \langle \psi|) \cdot |\varphi_\perp\rangle \otimes A_\perp) = \mathbf{Var}_{\text{global}}(A, \psi).$$

Proof. We begin by computing A_0:

$$A_0 = \langle \varphi_0| \otimes I \cdot A_{\text{combine}} = \left(\frac{1}{M^{1/2}} \cdot \sum_{u\in\mathbb{F}_q^m} \langle u| \right) \otimes I \cdot \left(\sum_{u\in\mathbb{F}_q^m} |u\rangle \otimes A^u \otimes I \right) = \frac{1}{M^{1/2}} \cdot \sum_{u\in\mathbb{F}_q^m} A^u \otimes I.$$

Then

$$|\varphi_0\rangle \otimes A_0 = \left(\frac{1}{M^{1/2}} \cdot \sum_{u\in\mathbb{F}_q^m} |u\rangle \right) \otimes \left(\frac{1}{M^{1/2}} \cdot \sum_{u\in\mathbb{F}_q^m} A^u \otimes I \right) = \frac{1}{M} \sum_{u\in\mathbb{F}_q^m} |u\rangle \otimes \sum_{v\in\mathbb{F}_q^m} A^v \otimes I$$

$$= \sum_{u\in\mathbb{F}_q^m} |u\rangle \otimes A_{\text{avg}} \otimes I,$$

where we have written $A_{\text{avg}} = \mathbf{E}_{\boldsymbol{u}} A^{\boldsymbol{u}}$. As a result,

$$|\varphi_\perp\rangle \otimes A_\perp = A_{\text{combine}} - |\varphi_0\rangle \otimes A_0 = \sum_{u\in\mathbb{F}_q^n} |u\rangle \otimes (A^u - A_{\text{avg}}) \otimes I.$$

Thus,

$$\langle \varphi_\perp| \otimes A_\perp \cdot (I \otimes |\psi\rangle \langle \psi|) \cdot |\varphi_\perp\rangle \otimes A_\perp$$

$$= \sum_{u,v\in\mathbb{F}_q^m} \langle u \mid v\rangle \otimes ((A^u - A_{\text{avg}}) \otimes I) \cdot |\psi\rangle \langle \psi| \cdot ((A^v - A_{\text{avg}}) \otimes I)$$

$$= \sum_{u\in\mathbb{F}_q^m} ((A^u - A_{\text{avg}}) \otimes I) \cdot |\psi\rangle \langle \psi| \cdot ((A^u - A_{\text{avg}}) \otimes I).$$

As a result, if we take the trace,

$$\frac{1}{M} \cdot \text{Tr}(\langle \varphi_\perp| \otimes A_\perp \cdot (I \otimes |\psi\rangle \langle \psi|) \cdot |\varphi_\perp\rangle \otimes A_\perp) = \mathop{\mathbf{E}}_{\boldsymbol{u}\sim\mathbb{F}_q^m} \langle \psi | (A^{\boldsymbol{u}} - A_{\text{avg}})^2 \otimes I |\psi\rangle . \qquad (45)$$

We can rewrite the squared expression as

$$\mathop{\mathbf{E}}_{\boldsymbol{u}\sim\mathbb{F}_q^m}(A^{\boldsymbol{u}} - A_{\mathrm{avg}})^2 = \mathop{\mathbf{E}}_{\boldsymbol{u}\sim\mathbb{F}_q^m}((A^{\boldsymbol{u}})^2 + (A_{\mathrm{avg}})^2 - A^{\boldsymbol{u}}\cdot A_{\mathrm{avg}} - A_{\mathrm{avg}}\cdot A^{\boldsymbol{u}})$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u}\sim\mathbb{F}_q^m}((A^{\boldsymbol{u}})^2 - (A_{\mathrm{avg}})^2)$$

$$= \frac{1}{2}\cdot\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}\sim\mathbb{F}_q^m}((A^{\boldsymbol{u}})^2 + (A^{\boldsymbol{v}})^2 - A^{\boldsymbol{u}}\cdot A^{\boldsymbol{v}} - A^{\boldsymbol{v}}\cdot A^{\boldsymbol{u}})$$

$$= \frac{1}{2}\cdot\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}\sim\mathbb{F}_q^m}(A^{\boldsymbol{u}} - A^{\boldsymbol{v}})^2.$$

Thus,

$$(45) = \frac{1}{2}\cdot\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}\sim\mathbb{F}_q^m}\langle\psi|\,(A^{\boldsymbol{u}} - A^{\boldsymbol{v}})^2\otimes I\,|\psi\rangle = \mathbf{Var}_{\mathrm{global}}(A,\psi).$$

This completes the proof. $\qquad\qquad\square$

Now we prove Lemma 7.7.

Proof of Lemma 7.7. We begin by computing

$$A^{\dagger}_{\mathrm{combine}}\cdot L\otimes|\psi\rangle\langle\psi|\cdot A_{\mathrm{combine}} = (\langle\varphi_0|\otimes A_0 + \langle\varphi_\perp|\otimes A_\perp)\cdot L\otimes|\psi\rangle\langle\psi|\cdot(|\varphi_0\rangle\otimes A_0 + |\varphi_\perp\rangle\otimes A_\perp)$$

$$= \langle\varphi_\perp|\otimes A_\perp\cdot L\otimes|\psi\rangle\langle\psi|\cdot|\varphi_\perp\rangle\otimes A_\perp$$

$$= \langle\varphi_\perp|\,L\,|\varphi_\perp\rangle\cdot A_\perp\,|\psi\rangle\langle\psi|\,A_\perp, \qquad\qquad (46)$$

where the second step follows from the fact that $|\varphi_0\rangle$ is a 0-eigenvector for L. Note that because $|\varphi_\perp\rangle$ is orthogonal to $|\varphi_0\rangle$,

$$\langle\varphi_\perp|\,L\,|\varphi_\perp\rangle \geq \frac{1}{mM}\cdot\langle\varphi_\perp\mid\varphi_\perp\rangle$$

by Corollary 7.5. As a result,

$$\mathbf{Var}_{\mathrm{local}}(A,\psi) = \mathrm{Tr}(A^{\dagger}_{\mathrm{combine}}\cdot L\otimes|\psi\rangle\langle\psi|\cdot A_{\mathrm{combine}}) \qquad\qquad\text{(by Lemma 7.8)}$$

$$= \langle\varphi_\perp|\,L\,|\varphi_\perp\rangle\cdot\mathrm{Tr}(A_\perp\cdot|\psi\rangle\langle\psi|\cdot A_\perp) \qquad\qquad\text{(by Equation (46))}$$

$$\geq \frac{1}{mM}\cdot\langle\varphi_\perp\mid\varphi_\perp\rangle\cdot\mathrm{Tr}(A_\perp\cdot|\psi\rangle\langle\psi|\cdot A_\perp)$$

$$= \frac{1}{mM}\cdot\mathrm{Tr}(\langle\varphi_\perp|\otimes A_\perp\cdot(I\otimes|\psi\rangle\langle\psi|)\cdot|\varphi_\perp\rangle\otimes A_\perp)$$

$$= \frac{1}{m}\cdot\mathbf{Var}_{\mathrm{global}}(A,\psi). \qquad\qquad\text{(by Lemma 7.9)}$$

This concludes the proof. $\qquad\qquad\square$

8 Global variance of the points measurements

Throughout this section, (ψ, A, B, L) will denote a fixed $(\epsilon, \delta, \gamma)$-good symmetric strategy for the (m, q, d)-low individual degree test.

Lemma 8.1. *Let $G \in \mathrm{PolySub}(m, q, d)$. Then*

$$B^{\ell}_{[f(u)=g(u)]}\otimes(G_g)^{1/2} \approx_{md/q} B^{\ell}_{g|_\ell}\otimes(G_g)^{1/2}$$

on the axis-parallel lines test distribution.

Proof. We want to bound the quantity

$$\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{\ell}} \sum_{g \in \mathcal{P}(m,d,q)} \| (B^{\boldsymbol{\ell}}_{[f(\boldsymbol{u})=g(\boldsymbol{u})]} - B^{\boldsymbol{\ell}}_{g|_{\boldsymbol{\ell}}}) \otimes (G_g)^{1/2} | \psi \rangle \|^2$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{\ell}} \sum_{g \in \mathcal{P}(m,d,q)} \langle \psi | \left(\sum_{f : f \neq g|_{\boldsymbol{\ell}}} \mathbf{1}[f(\boldsymbol{u}) = g(\boldsymbol{u})] \cdot B^{\boldsymbol{\ell}}_f \right)^2 \otimes G_g | \psi \rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{\ell}} \sum_{g \in \mathcal{P}(m,d,q)} \langle \psi | \left(\sum_{f : f \neq g|_{\boldsymbol{\ell}}} \mathbf{1}[f(\boldsymbol{u}) = g(\boldsymbol{u})] \cdot B^{\boldsymbol{\ell}}_f \right) \otimes G_g | \psi \rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{\ell}} \sum_{g \in \mathcal{P}(m,d,q)} \sum_{f : f \neq g|_{\boldsymbol{\ell}}} \langle \psi | B^{\boldsymbol{\ell}}_f \otimes G_g | \psi \rangle \cdot \left(\mathop{\mathbf{E}}_{\boldsymbol{u}} \mathbf{1}[f(\boldsymbol{u}) = g(\boldsymbol{u})] \right)$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{\ell}} \sum_{g \in \mathcal{P}(m,d,q)} \sum_{f : f \neq g|_{\boldsymbol{\ell}}} \langle \psi | B^{\boldsymbol{\ell}}_f \otimes G_g | \psi \rangle \cdot \frac{md}{q} \qquad \text{(by Schwartz-Zippel)}$$

$$\leq \frac{md}{q}. \qquad \qquad \square$$

Lemma 8.2. *Let $G \in \mathrm{PolySub}(m,q,d)$. Then*

$$A^u_{g(u)} \otimes (G_g)^{1/2} \approx_{24 \cdot (\epsilon + \delta + \frac{md}{q})} A^v_{g(v)} \otimes (G_g)^{1/2} \tag{47}$$

on the distribution $(\boldsymbol{u}, \boldsymbol{v}) \sim C$.

Proof. Let $\boldsymbol{u}$ and $\boldsymbol{\ell}$ be distributed as in the axis-parallel lines test, and sample $\boldsymbol{v} \sim \boldsymbol{\ell}$. Then $\boldsymbol{v}$ and $\boldsymbol{\ell}$ are also distributed as in the axis-parallel lines test. As a result,

$$A^u_{g(u)} \otimes (G_g)^{1/2} \approx_{2\delta} I \otimes (G_g)^{1/2} A^u_{g(u)} \qquad \text{(by Proposition 4.22)}$$

$$\approx_{2\epsilon} B^{\boldsymbol{\ell}}_{[f(u)=g(u)]} \otimes (G_g)^{1/2} \qquad \text{(by Proposition 4.22)}$$

$$\approx_{\frac{md}{q}} B^{\boldsymbol{\ell}}_{g|_{\boldsymbol{\ell}}} \otimes (G_g)^{1/2} \qquad \text{(by Lemma 8.1)}$$

$$\approx_{\frac{md}{q}} B^{\boldsymbol{\ell}}_{[f(v)=g(v)]} \otimes (G_g)^{1/2} \qquad \text{(by Lemma 8.1)}$$

$$\approx_{2\epsilon} I \otimes (G_g)^{1/2} A^v_{g(v)} \qquad \text{(by Proposition 4.22)}$$

$$\approx_{2\delta} A^v_{g(v)} \otimes (G_g)^{1/2}. \qquad \text{(by Proposition 4.22)}$$

Steps 1, 2, 5, and 6 are also using Remark 3.7 and Proposition 4.26. The lemma now follows from Proposition 4.28. $\qquad \square$

We note that Equation (47) is equivalent to the statement that

$$\sum_{g \in \mathcal{P}(m,q,d)} \mathop{\mathbf{E}}_{(\boldsymbol{u},\boldsymbol{v}) \sim C} \langle \psi | (A^{\boldsymbol{u}}_{g(\boldsymbol{u})} - A^{\boldsymbol{v}}_{g(\boldsymbol{v})})^2 \otimes G_g | \psi \rangle \leq 24 \left(\epsilon + \delta + \frac{md}{q} \right). \tag{48}$$

This can be viewed as a form of local variance for the points measurements. We now derive the corresponding expression for the global variance of the points measurements.

Lemma 8.3. *Let $G \in \mathrm{PolySub}(m,q,d)$. Then*

$$A^u_{g(u)} \otimes (G_g)^{1/2} \approx_{24m \cdot (\epsilon + \delta + \frac{md}{q})} A^v_{g(v)} \otimes (G_g)^{1/2} \tag{49}$$

on the distribution $\boldsymbol{u}, \boldsymbol{v} \sim \mathbb{F}^m_q$.

Proof. We want to bound

$$\underset{\boldsymbol{u},\boldsymbol{v}\sim\mathbb{F}_q^m}{\mathbf{E}} \sum_{g\in\mathcal{P}(m,q,d)} \|(A_{g(\boldsymbol{u})}^{\boldsymbol{u}} - A_{g(\boldsymbol{v})}^{\boldsymbol{v}}) \otimes (G_g)^{1/2} |\psi\rangle \|^2 = \underset{\boldsymbol{u},\boldsymbol{v}\sim\mathbb{F}_q^m}{\mathbf{E}} \sum_{g\in\mathcal{P}(m,q,d)} \langle\psi|\, (A_{g(\boldsymbol{u})}^{\boldsymbol{u}} - A_{g(\boldsymbol{v})}^{\boldsymbol{v}})^2 \otimes G_g |\psi\rangle\,.$$

(50)

For each $g \in \mathcal{P}(m,q,d)$, define

$$\forall u \in \mathbb{F}_q^m,\ A(g)^u := A_{g(u)}^u, \qquad \text{and} \qquad |\psi_g\rangle := I \otimes (G_g)^{1/2} |\psi\rangle\,.$$

Then

$$
\begin{aligned}
(50) &= \sum_{g\in\mathcal{P}(m,q,d)} \underset{\boldsymbol{u},\boldsymbol{v}\sim\mathbb{F}_q^n}{\mathbf{E}} \langle\psi_g|\, (A(g)^{\boldsymbol{u}} - A(g)^{\boldsymbol{v}})^2 \otimes I\, |\psi_g\rangle \\
&= \sum_{g\in\mathcal{P}(m,q,d)} 2\cdot \mathbf{Var}_{\text{global}}(A(g),\psi_g) \\
&\leq \sum_{g\in\mathcal{P}(m,q,d)} 2m\cdot \mathbf{Var}_{\text{local}}(A(g),\psi_g) \qquad\qquad\text{(by Lemma 7.7)}\\
&= m\cdot \sum_{g\in\mathcal{P}(m,q,d)} \underset{(\boldsymbol{u},\boldsymbol{v})\sim C}{\mathbf{E}} \langle\psi_g|\, (A(g)^{\boldsymbol{u}} - A(g)^{\boldsymbol{v}})^2 \otimes I\, |\psi_g\rangle \\
&= m\cdot \sum_{g\in\mathcal{P}(m,q,d)} \underset{(\boldsymbol{u},\boldsymbol{v})\sim C}{\mathbf{E}} \langle\psi|\, (A_{g(\boldsymbol{u})}^{\boldsymbol{u}} - A_{g(\boldsymbol{v})}^{\boldsymbol{v}})^2 \otimes G_g |\psi\rangle \\
&\leq m\cdot 24\left(\epsilon + \delta + \frac{md}{q}\right). \qquad\qquad\qquad\text{(by Equation (48))}
\end{aligned}
$$

This concludes the proof. □

9   Self-improvement

Throughout this section, (ψ, A, B, L) will denote a fixed (ϵ,δ,γ)-good symmetric strategy for the (m,q,d) low individual degree test. The majority of this section will be devoted to proving Lemma 9.1 below, which is a slightly weaker form of Theorem 6.2. The key difference is that the measurement H it outputs is allowed to be non-projective, rather than the projective measurement given by Theorem 6.2. Having proven this, we can apply Theorem 5.2 to produce a projective measurement; this is done in Section 9.3 below, completing the proof of Theorem 6.2.

We now highlight other differences between Lemma 9.1 and Theorem 6.2. Since H is non-projective, we have stated its strong self-consistency in Item 3 in terms of Definition 4.34; see Proposition 4.36 for a proof that these conditions are equivalent for projective sub-measurements. The other key difference is that the boundedness condition is modified slightly in Item 4. Finally, the error ζ is substantially smaller.

Lemma 9.1 (Self-improvement with non-projective output)**.** *Let $G \in \text{PolyMeas}(m,q,d)$ be a measurement with the following property:*

1. *(Consistency with A):   On average over $\boldsymbol{u} \sim \mathbb{F}_q^m$,*

$$A_a^u \otimes I \simeq_\nu I \otimes G_{[g(u)=a]}\,.$$

Let

$$\zeta = 100m \cdot \left(\epsilon^{1/2} + \delta^{1/2} + (d/q)^{1/2}\right).$$

Then there exists $H \in \text{PolySub}(m,q,d)$ with the following properties:

1. *(Completeness):* If $H = \sum_h H_h$, then

$$\langle \psi | H \otimes I | \psi \rangle \geq (1 - \nu) - \zeta.$$

2. *(Consistency with A):* On average over $\boldsymbol{u} \sim \mathbb{F}_q^m$,

$$A_a^u \otimes I \simeq_\zeta I \otimes H_{[h(u)=a]}.$$

3. *(Strong self-consistency):*

$$\sum_h \langle \psi | H_h \otimes H_h | \psi \rangle \geq \langle \psi | H \otimes I | \psi \rangle - \zeta.$$

4. *(Boundedness):* There exists a positive-semidefinite matrix Z such that

$$\langle \psi | Z \otimes I | \psi \rangle - \mathbf{E}_{\boldsymbol{u}} \sum_a \langle \psi | A_a^{\boldsymbol{u}} \otimes H_{[h(\boldsymbol{u})=a]} | \psi \rangle \leq \zeta$$

and for each $h \in \mathcal{P}(m, q, d)$,

$$Z \geq \left(\mathbf{E}_{\boldsymbol{u}} A_{h(\boldsymbol{u})}^{\boldsymbol{u}} \right).$$

9.1 A semidefinite program

A key element in the proof of Lemma 9.1 will be a pair of primal and dual semidefinite programs. To define them, it will be convenient to introduce the notational shorthand

$$A_g = \mathbf{E}_{\boldsymbol{u} \sim \mathbb{F}_q^m} A_{g(\boldsymbol{u})}^{\boldsymbol{u}}.$$

Then the primal is

$$\sup \quad \sum_g \mathrm{tr}(T_g \cdot A_g) \tag{51}$$
$$\text{s.t.} \quad T_g \geq 0 \qquad \forall g \in \mathcal{P}(m, q, d) \;,$$
$$\sum_g T_g \leq I,$$

and the dual is

$$\inf \quad \mathrm{tr}(Z) \tag{52}$$
$$\text{s.t.} \quad Z \geq A_g. \tag{53}$$

We will prove that these two program are indeed dual to each other in Lemma 9.2 below.

Lemma 9.2. *The semidefinite programs* (51) *and* (52) *are dual to each other. Moreover there is an optimal pair of solutions* $\{T_g\}$ *to* (51) *and* Z *to* (52) *such that* $\sum_g T_g = I$ *and*

$$T_g Z = T_g A_g, \qquad \forall g \in \mathcal{P}(m, q, d). \tag{54}$$

Proof. To show that (51) and (52) are dual to each other, we begin by rewriting the primal (51) in canonical form. Let r be the dimension of the space on which A acts. Let $M = |\mathcal{P}(m, q, d)|$ be the number of polynomials with individual degree d. We will assume some ordering of these polynomials $g_1, \ldots, g_M \in \mathcal{P}(m, q, d)$ which is allowed to be arbitrary. Consider the following semidefinite program:

$$\begin{aligned}
\sup \quad & \operatorname{tr}(C^\dagger X) & (55) \\
\text{s.t.} \quad & \operatorname{tr}(D_{ij}^\dagger X) = b_{ij} & \forall i, j \in \{1, \ldots, r\}, \\
& X \geq 0,
\end{aligned}$$

where the variables C, D_{ij}, and b_{ij} are defined as follows:

$$C = \sum_{i=1}^{M} |i\rangle \langle i| \otimes A_{g_i}$$

$$\forall i, j \in [r], \quad D_{ij} = \sum_{k=1}^{M+1} |k\rangle \langle k| \otimes |i\rangle \langle j|, \qquad b_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

We claim that (55) is equivalent to (51). To see this, let

$$X = \sum_{i,j=1}^{M+1} |i\rangle \langle j| \otimes X_{ij}$$

be a feasible solution to (55). Then because $X \geq 0$, $X_{ii} \geq 0$ for each $i \in [M+1]$. In addition

$$\begin{aligned}
\operatorname{tr}(D_{i_1 i_2}^\dagger X) &= \sum_k \sum_{j_1 j_2} \operatorname{tr}((|k\rangle \langle k| \otimes |i_1\rangle \langle i_2|) \cdot (|j_1\rangle \langle j_2| \otimes X_{j_1 j_2})) \\
&= \sum_k \langle i_2| X_{kk} |i_1\rangle \\
&= \begin{cases} 1 & \text{if } i_1 = i_2, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

This is equivalent to the statement $\sum_{i=1}^{M+1} X_{ii} = I$, which implies that $\sum_{i=1}^{M} X_{ii} \leq I$. Finally, the objective value (55) is

$$\begin{aligned}
\operatorname{tr}(C^\dagger X) &= \sum_{i=1}^{M} \sum_{j,k=1}^{M+1} \operatorname{tr}((|i\rangle \langle i| \otimes A_{g_i}^\dagger) \cdot (|j\rangle \langle k| \otimes X_{j,k})) \\
&= \sum_{i=1}^{M} \operatorname{tr}(A_{g_i}^\dagger \cdot X_{i,i}) \\
&= \sum_{i=1}^{M} \operatorname{tr}(A_{g_i} \cdot X_{i,i}). & \text{(because } A \text{ is Hermitian)}
\end{aligned}$$

As a result, setting $T_{g_i} = X_{ii}$ for each $i \in [M]$ gives a feasible solution to the original semidefinite program (51) with a matching objective value. A similar transformation allows us to convert solutions of (51) to (55).

The dual of (55) is

$$\inf \quad \sum_i z_{ij} b_{ij} \tag{56}$$

$$\text{s.t.} \quad \sum_{i,j} z_{ij} D_{ij} \geq C. \tag{57}$$

We claim that (56) is equivalent to (52). To see this, we first calculate

$$\sum_{i,j=1}^{r} z_{ij} D_{ij} = \sum_{i,j=1}^{r} z_{ij} \cdot \sum_{k=1}^{M+1} |k\rangle\langle k| \otimes |i\rangle\langle j| = \sum_{k=1}^{M+1} |k\rangle\langle k| \otimes \left(\sum_{i,j=1}^{r} z_{ij} |i\rangle\langle j| \right)$$

$$=: \sum_{k=1}^{M+1} |k\rangle\langle k| \otimes Z.$$

Then the constraint (57) states that

$$\sum_{i=1}^{M+1} |i\rangle\langle i| \otimes Z \geq C = \sum_{i=1}^{M} |i\rangle\langle i| \otimes A_{g_i},$$

which is equivalent to the statement that $Z \geq A_{g_i}$ for all $i \in [M]$ and $Z \geq 0$. In other words, Z is a feasible solution to the original dual SDP (52), with value

$$\text{tr}(Z) = \sum_{i=1}^{r} Z_{ii} = \sum_{i,j=1}^{r} b_{ij} Z_{ij} = \sum_{i,j=1}^{r} b_{ij} z_{ij},$$

the same value as in (56). Hence, the two dual programs are the same as well, which implies that (51) and (52) form a primal/dual pair.

To show that a primal/dual pair satisfies *strong duality*, i.e. that their optimum values are the same, we use Slater's condition [BV04, Section 5.2.3] and show that they satisfy *strict feasibility*, which means both have a feasible solution which is positive *definite* that satisfies all constraints with a strict inequality. It can be checked that the following two solutions to (51) and (52) satisfy this property:

$$\forall g \in \mathcal{P}(m,q,d), \quad T_g = \frac{1}{2M} \cdot I, \qquad\qquad Z = 2I.$$

Because they satisfy strong duality, their optimal solutions satisfy the *complementary slackness* condition (see [AHO97]). If X and (z_{ij}) are an optimal pair of solutions to (55) and (56) respectively, this implies that

$$X \left(\sum_{i,j} z_{ij} D_{ij} - C \right) = 0 . \tag{58}$$

Clearly for any optimal pair (X, z_{ij}) we can assume without loss of generality that X is block-diagonal. Then if we translate (58) back to the variables $\{T_g\}$ and Z, we get

$$0 = X \left(\sum_{i,j} z_{ij} D_{ij} - C \right) = \sum_{i=1}^{M+1} |i\rangle\langle i| \otimes X_{ii} \cdot \left(\sum_{i=1}^{M+1} |i\rangle\langle i| \otimes Z - \sum_{i=1}^{M} |i\rangle\langle i| \otimes A_{g_i} \right)$$

$$= \sum_{i=1}^{M} |i\rangle\langle i| \otimes (X_{ii} \cdot (Z - A_{g_i})) + |M+1\rangle\langle M+1| \otimes (X_{M+1,M+1} \cdot Z).$$

This implies that $X_{ii} \cdot (Z - A_{g_i}) = 0$ for $i \in [M]$ and $X_{M+1,M+1} = 0$. Translating back to the variables $\{T_g\}$ and Z, this gives $T_g(Z - A_g) = 0$ for all $g \in \mathcal{P}(m,q,d)$ and $\sum_g T_g = I$. $\qquad\square$

9.2 Proof of Lemma 9.1

We let $T = \{T_g\}$ and Z be the optimal solutions to the SDPs (51) and (52) respectively given by Lemma 9.2. Then T is a measurement, and

$$\forall g \in \mathcal{P}(m,q,d): \qquad Z \geq (\underset{\boldsymbol{u}}{\mathbf{E}}\, A^{\boldsymbol{u}}_{g(\boldsymbol{u})}), \tag{59}$$

$$T_g \cdot Z = T_g \cdot (\underset{\boldsymbol{u}}{\mathbf{E}}\, A^{\boldsymbol{u}}_{g(\boldsymbol{u})}). \tag{60}$$

For each $u \in \mathbb{F}_q^m$, define $H^u = \{H^u_h\}_{h \in \mathcal{P}(m,q,d)}$ as

$$H^u_h := A^u_{h(u)} \cdot T_h \cdot A^u_{h(u)}.$$

Let $u \in \mathbb{F}_q^m$. Then

$$\sum_{h \in \mathcal{P}(m,q,d)} H^u_h = \sum_{h \in \mathcal{P}(m,q,d)} A^u_{h(u)} \cdot T_h \cdot A^u_{h(u)}$$

$$= \sum_{a \in \mathbb{F}_q} A^u_a \cdot \left(\sum_{h:h(u)=a} T_h \right) \cdot A^u_a$$

$$\leq \sum_{a \in \mathbb{F}_q} (A^u_a)^2 \qquad \text{(because } T \text{ is a measurement)}$$

$$= \sum_{a \in \mathbb{F}_q} A^u_a \qquad \text{(}A \text{ is projective)}$$

$$= I.$$

Hence, H^u is a sub-measurement, and therefore $H^u \in \mathrm{PolySub}(m,q,d)$. Next, define $H = \{H_h\}_{h \in \mathcal{P}(m,q,d)}$ as

$$H_h := \underset{\boldsymbol{u} \sim \mathbb{F}_q^m}{\mathbf{E}}\, H^{\boldsymbol{u}}_h.$$

Then

$$\sum_{h \in \mathcal{P}(m,q,d)} H_h = \underset{\boldsymbol{u} \sim \mathbb{F}_q^m}{\mathbf{E}} \sum_{h \in \mathcal{P}(m,q,d)} H^{\boldsymbol{u}}_h \leq I. \qquad \text{(}H^u \text{ is a sub-measurement)}$$

Hence, H is a sub-measurement, and therefore $H \in \mathrm{PolySub}(m,q,d)$.

Set

$$\zeta_{\text{variance}} = 24m \cdot \left(\epsilon + \delta + \frac{md}{q} \right)$$

to be the error in Equation (47). Prior to showing that H satisfies Items 1 to 4, we will prove the following technical lemma.

Lemma 9.3. *Suppose $M = \{M^u_o\}$ is a sub-measurement with outcomes in some set $\mathcal{O}$. For each $u \in \mathbb{F}_q^m$, let S_u be a subset of $\mathcal{O} \otimes \mathcal{P}(m,q,d)$. Then*

$$\underset{\boldsymbol{u} \sim \mathbb{F}_q^m}{\mathbf{E}} \sum_{(o,h) \in S_{\boldsymbol{u}}} \langle \psi | M^{\boldsymbol{u}}_o \otimes H_h | \psi \rangle \approx_{4\sqrt{\zeta_{\text{variance}}}} \underset{\boldsymbol{u} \sim \mathbb{F}_q^m}{\mathbf{E}} \sum_{(o,h) \in S_{\boldsymbol{u}}} \langle \psi | (A^{\boldsymbol{u}}_{h(\boldsymbol{u})} \cdot M^{\boldsymbol{u}}_o \cdot A^{\boldsymbol{u}}_{h(\boldsymbol{u})}) \otimes T_h | \psi \rangle .$$

Proof. We begin by expanding

$$\underset{\boldsymbol{u} \sim \mathbb{F}_q^m}{\mathbf{E}} \sum_{(o,h) \in S_{\boldsymbol{u}}} \langle \psi | M^{\boldsymbol{u}}_o \otimes H_h | \psi \rangle = \underset{\boldsymbol{u},\boldsymbol{v}}{\mathbf{E}} \sum_{(o,h) \in S_{\boldsymbol{u}}} \langle \psi | M^{\boldsymbol{u}}_o \otimes H^{\boldsymbol{v}}_h | \psi \rangle$$

$$= \underset{\boldsymbol{u},\boldsymbol{v}}{\mathbf{E}} \sum_{(o,h) \in S_{\boldsymbol{u}}} \langle \psi | M^{\boldsymbol{u}}_o \otimes (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \cdot T_h \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) | \psi \rangle . \tag{61}$$

We claim that

$$(61) \approx_{\sqrt{2\delta}} \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \cdot M^{\boldsymbol{u}}_o) \otimes (T_h \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) |\psi\rangle . \tag{62}$$

To show this, we bound the magnitude of the difference.

$$\left| \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \otimes I - I \otimes A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \cdot (M^{\boldsymbol{u}}_o \otimes (T_h \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})})) |\psi\rangle \right|$$

$$\leq \sqrt{ \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \otimes I - I \otimes A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \cdot (M^{\boldsymbol{u}}_o \otimes T_h) \cdot (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \otimes I - I \otimes A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) |\psi\rangle }$$

$$\cdot \sqrt{ \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| M^{\boldsymbol{u}}_o \otimes (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \cdot T_h \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) |\psi\rangle } . \tag{63}$$

The term inside the first square root is

$$\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{a\in\mathbb{F}_q} \langle\psi| (A^{\boldsymbol{v}}_a \otimes I - I \otimes A^{\boldsymbol{v}}_a) \cdot \left(\sum_{(o,h)\in S_{\boldsymbol{u}}, h(\boldsymbol{v})=a} M^{\boldsymbol{u}}_o \otimes T_h \right) \cdot (A^{\boldsymbol{v}}_a \otimes I - I \otimes A^{\boldsymbol{v}}_a) |\psi\rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{a\in\mathbb{F}_q} \langle\psi| (A^{\boldsymbol{v}}_a \otimes I - I \otimes A^{\boldsymbol{v}}_a)^2 |\psi\rangle , \qquad \text{(because } M^{\boldsymbol{u}} \text{ and } T_h \text{ are sub-measurements)}$$

which is at most 2δ by Proposition 4.22 and because A is δ-self-consistent. The term inside the second square root is

$$\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| M^{\boldsymbol{u}}_o \otimes H^{\boldsymbol{v}}_h |\psi\rangle ,$$

which is at most 1 because $M^{\boldsymbol{u}}$ and $H^{\boldsymbol{v}}$ are sub-measurements. Next, we claim that

$$(62) \approx_{\sqrt{2\delta}} \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \cdot M^{\boldsymbol{u}}_o \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \otimes T_h |\psi\rangle . \tag{64}$$

To show this, we bound the magnitude of the difference.

$$\left| \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| ((A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \cdot M^{\boldsymbol{u}}_o) \otimes T_h) \cdot (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \otimes I - I \otimes A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) |\psi\rangle \right|$$

$$\leq \sqrt{ \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \cdot M^{\boldsymbol{u}}_o \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \otimes T_h |\psi\rangle }$$

$$\cdot \sqrt{ \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \otimes I - I \otimes A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \cdot (M^{\boldsymbol{u}}_o \otimes T_h) \cdot (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \otimes I - I \otimes A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) |\psi\rangle } . \tag{65}$$

The term inside the first square root is

$$\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{h} \langle\psi| (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \cdot \left(\sum_{o:(o,h)\in S_{\boldsymbol{u}}} M^{\boldsymbol{u}}_o \right) \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \otimes T_h |\psi\rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{h} \langle\psi| (A^{\boldsymbol{v}}_{h(\boldsymbol{v})})^2 \otimes T_h |\psi\rangle \qquad \text{(because } M \text{ is a sub-measurement)}$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{h} \langle\psi| I \otimes T_h |\psi\rangle \qquad \text{(because } A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \leq I)$$

$$= 1. \qquad \text{(because } T \text{ is a measurement)}$$

As for the term inside the second square root, it is equal to the term inside the first square root in Equation (63), which we showed was at most 2δ.

Having moved both A's to the left-hand side, we want to show that

$$(64) \approx_{\sqrt{\zeta_{\text{variance}}}} \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| (A^{\boldsymbol{u}}_{h(\boldsymbol{u})} \cdot M^{\boldsymbol{u}}_o \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \otimes T_h |\psi\rangle. \tag{66}$$

To show this, we bound the magnitude of the difference.

$$\left| \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| ((A^{\boldsymbol{v}}_{h(\boldsymbol{v})} - A^{\boldsymbol{u}}_{h(\boldsymbol{u})}) \cdot M^{\boldsymbol{u}}_o \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \otimes T_h |\psi\rangle \right|$$

$$\leq \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| ((A^{\boldsymbol{v}}_{h(\boldsymbol{v})} - A^{\boldsymbol{u}}_{h(\boldsymbol{u})}) \cdot M^{\boldsymbol{u}}_o \cdot (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} - A^{\boldsymbol{u}}_{h(\boldsymbol{u})})) \otimes T_h |\psi\rangle}$$

$$\cdot \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \cdot M^{\boldsymbol{u}}_o \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \otimes T_h |\psi\rangle}. \tag{67}$$

The term inside the first square root is

$$\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_h \langle\psi| ((A^{\boldsymbol{v}}_{h(\boldsymbol{v})} - A^{\boldsymbol{u}}_{h(\boldsymbol{u})}) \cdot \left(\sum_{o:(o,h)\in S_{\boldsymbol{u}}} M^{\boldsymbol{u}}_o \right) \cdot (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} - A^{\boldsymbol{u}}_{h(\boldsymbol{u})})) \otimes T_h |\psi\rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_h \langle\psi| (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} - A^{\boldsymbol{u}}_{h(\boldsymbol{u})})^2 \otimes T_h |\psi\rangle. \qquad \text{(because $M^{\boldsymbol{u}}$ is a sub-measurement)}$$

But $T \in \text{PolySub}(m, q, d)$, and so by Lemma 8.3 this expression is at most ζ_{variance}. As for the term inside the second square root, it is equal to the term inside the first square root in Equation (65), which we showed was at most 1. Finally, we want to show that

$$(66) \approx_{\sqrt{\zeta_{\text{variance}}}} \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| (A^{\boldsymbol{u}}_{h(\boldsymbol{u})} \cdot M^{\boldsymbol{u}}_o \cdot A^{\boldsymbol{u}}_{h(\boldsymbol{u})}) \otimes T_h |\psi\rangle. \tag{68}$$

To show this, we bound the magnitude of the difference.

$$\left| \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| (A^{\boldsymbol{u}}_{h(\boldsymbol{u})} \cdot M^{\boldsymbol{u}}_o \cdot (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} - A^{\boldsymbol{u}}_{h(\boldsymbol{u})})) \otimes T_h |\psi\rangle \right|$$

$$\leq \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| (A^{\boldsymbol{u}}_{h(\boldsymbol{u})} \cdot M^{\boldsymbol{u}}_o \cdot A^{\boldsymbol{u}}_{h(\boldsymbol{u})}) \otimes T_h |\psi\rangle}$$

$$\cdot \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{(o,h)\in S_{\boldsymbol{u}}} \langle\psi| ((A^{\boldsymbol{v}}_{h(\boldsymbol{v})} - A^{\boldsymbol{u}}_{h(\boldsymbol{u})}) \cdot M^{\boldsymbol{u}}_o \cdot (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} - A^{\boldsymbol{u}}_{h(\boldsymbol{u})})) \otimes T_h |\psi\rangle}.$$

The term inside the first square root is

$$\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_h \langle\psi| (A^{\boldsymbol{u}}_{h(\boldsymbol{u})} \cdot \left(\sum_{o:(o,h)\in S_{\boldsymbol{u}}} M^{\boldsymbol{u}}_o \right) \cdot A^{\boldsymbol{u}}_{h(\boldsymbol{u})}) \otimes T_h |\psi\rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_h \langle\psi| (A^{\boldsymbol{u}}_{h(\boldsymbol{u})})^2 \otimes T_h |\psi\rangle \qquad \text{(because M is a sub-measurement)}$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_h \langle\psi| I \otimes T_h |\psi\rangle \qquad \text{(because $A^{\boldsymbol{u}}_{h(\boldsymbol{u})} \leq I$)}$$

$$= 1. \qquad \text{(because T is a measurement)}$$

73

As for the term inside the second square root, it is equal to the term inside the first square root in Equation (67), which we showed was at most ζ_{variance}. This concludes the proof with an error of $2\sqrt{2\delta} + 2\sqrt{\zeta_{\text{variance}}}$. The lemma now follows by observing that $2\delta \leq \zeta_{\text{variance}}$. $\qquad\square$

We now show that H satisfies Items 1 to 4.

Proof of Item 1 (Completeness). The completeness of H is

$$
\begin{aligned}
\sum_h \langle\psi| H_h \otimes I |\psi\rangle &= \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_h \langle\psi| H_h^{\boldsymbol{u}} \otimes I |\psi\rangle \\
&= \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_h \langle\psi| (A_{h(\boldsymbol{u})}^{\boldsymbol{u}} \cdot T_h \cdot A_{h(\boldsymbol{u})}^{\boldsymbol{u}}) \otimes I |\psi\rangle \\
&= \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_a \langle\psi| (A_a^{\boldsymbol{u}} \cdot T_{[h(\boldsymbol{u})=a]} \cdot A_a^{\boldsymbol{u}}) \otimes I |\psi\rangle .
\end{aligned}
\tag{69}
$$

We claim that

$$
(69) \approx_{2\sqrt{\delta}} \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_a \langle\psi| (T_{[h(\boldsymbol{u})=a]} \cdot A_a^{\boldsymbol{u}}) \otimes A_a^{\boldsymbol{u}} |\psi\rangle .
\tag{70}
$$

To show this, we bound the magnitude of the difference

$$
\left| \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_a \langle\psi| (A_a^{\boldsymbol{u}} \otimes I - I \otimes A_a^{\boldsymbol{u}}) \cdot ((T_{[h(\boldsymbol{u})=a]} \cdot A_a^{\boldsymbol{u}}) \otimes I) |\psi\rangle \right|
$$

$$
\leq \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_a \langle\psi| (A_a^{\boldsymbol{u}} \otimes I - I \otimes A_a^{\boldsymbol{u}})^2 |\psi\rangle} \cdot \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_a \langle\psi| (A_a^{\boldsymbol{u}} \cdot T_{[h(\boldsymbol{u})=a]}^2 \cdot A_a^{\boldsymbol{u}}) \otimes I |\psi\rangle}.
$$

The expression inside the first square root is at most $2\delta \leq 4\delta$ by Proposition 4.22 and the self-consistency of A, and the expression inside the second square root is at most 1 because $T_{[h(\boldsymbol{u})=a]} \leq I$. Next, we claim that

$$
(70) \approx_{\sqrt{\delta}} \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_a \langle\psi| (T_{[h(\boldsymbol{u})=a]} \cdot A_a^{\boldsymbol{u}}) \otimes I |\psi\rangle .
\tag{71}
$$

To show this, we bound the magnitude of the difference.

$$
\left| \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_a \langle\psi| (T_{[h(\boldsymbol{u})=a]} \cdot A_a^{\boldsymbol{u}}) \otimes (I - A_a^{\boldsymbol{u}}) |\psi\rangle \right|
$$

$$
\leq \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_a \langle\psi| (T_{[h(\boldsymbol{u})=a]})^2 \otimes I |\psi\rangle} \cdot \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_a \langle\psi| (A_a^{\boldsymbol{u}})^2 \otimes (I - A_a^{\boldsymbol{u}})^2 |\psi\rangle}.
$$

The expression inside the first square root is at most 1 because T is a sub-measurement. By the projectivity of A, the expression inside the second square root is equal to

$$
\mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_a \langle\psi| (A_a^{\boldsymbol{u}}) \otimes (I - A_a^{\boldsymbol{u}}) |\psi\rangle = \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_{a \neq b} \langle\psi| A_a^{\boldsymbol{u}} \otimes A_b^{\boldsymbol{u}} |\psi\rangle ,
$$

which is at most δ by the self-consistency of A. We can rewrite Equation (71) as

$$\mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_h \langle\psi| (T_h \cdot A^{\boldsymbol{u}}_{h(\boldsymbol{u})}) \otimes I |\psi\rangle$$

$$= \sum_h \langle\psi| (T_h \cdot \mathop{\mathbf{E}}_{\boldsymbol{u}} A^{\boldsymbol{u}}_{h(\boldsymbol{u})}) \otimes I |\psi\rangle$$

$$= \sum_h \langle\psi| (T_h \cdot Z) \otimes I |\psi\rangle \qquad \text{(by Equation (60))}$$

$$= \langle\psi| Z \otimes I |\psi\rangle. \qquad \text{(because } T \text{ is a measurement)}$$

We pause and record what we have shown so far:

$$\sum_h \langle\psi| H_h \otimes I |\psi\rangle \geq \langle\psi| Z \otimes I |\psi\rangle - 3\sqrt{\delta}. \tag{72}$$

At this point, we can lower-bound

$$\langle\psi| Z \otimes I |\psi\rangle \geq \sum_g \langle\psi| Z \otimes G_g |\psi\rangle \qquad \text{(because } G \text{ is a sub-measurement)}$$

$$\geq \sum_g \langle\psi| (\mathop{\mathbf{E}}_{\boldsymbol{u}} A^{\boldsymbol{u}}_{g(\boldsymbol{u})}) \otimes G_g |\psi\rangle \qquad \text{(by Equation (59))}$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_a \langle\psi| A^{\boldsymbol{u}}_a \otimes G_{[g(\boldsymbol{u})=a]} |\psi\rangle$$

$$\geq 1 - \nu. \qquad \text{(by Item 1 and Proposition 4.19)}$$

Thus, the completeness is at least $1 - \nu - 3\sqrt{\delta}$. The proof follows from noting that $3\sqrt{\delta} \leq \zeta$.

Proof of Item 2 (Consistency with A). The inconsistency of H with A is

$$\mathop{\mathbf{E}}_{\boldsymbol{u}\sim\mathbb{F}_q^m} \sum_{a\neq b} \langle\psi| A^{\boldsymbol{u}}_a \otimes H_{[h(\boldsymbol{u})=b]} |\psi\rangle = \mathop{\mathbf{E}}_{\boldsymbol{u}\sim\mathbb{F}_q^m} \sum_{a,h:h(\boldsymbol{u})\neq a} \langle\psi| A^{\boldsymbol{u}}_a \otimes H_h |\psi\rangle. \tag{73}$$

We now apply Lemma 9.3 to the right-hand side of Equation (73). To do so, we set $\mathcal{O} = \mathbb{F}_q^m$, $M = A$, and $S_u = \{(a,h) : h(u) \neq a\}$. Then Lemma 9.3 implies that

$$(73) \approx_{4\sqrt{\zeta_{\text{variance}}}} \mathop{\mathbf{E}}_{\boldsymbol{u}\sim\mathbb{F}_q^m} \sum_{a,h:h(\boldsymbol{u})\neq a} \langle\psi| (A^{\boldsymbol{u}}_{h(\boldsymbol{u})} \cdot A^{\boldsymbol{u}}_a \cdot A^{\boldsymbol{u}}_{h(\boldsymbol{u})}) \otimes T_h |\psi\rangle,$$

which is equal to 0 because A is projective. This implies that

$$\mathop{\mathbf{E}}_{\boldsymbol{u}\sim\mathbb{F}_q^m} \sum_{a\neq b} \langle\psi| A^{\boldsymbol{u}}_a \otimes H_{[h(\boldsymbol{u})=b]} |\psi\rangle \leq 4\sqrt{\zeta_{\text{variance}}}. \tag{74}$$

The proof follows from noting that

$$4\sqrt{\zeta_{\text{variance}}} = 4 \cdot \sqrt{24m \cdot \left(\epsilon + \delta + \frac{md}{q}\right)}$$

$$\leq 20m \cdot \left(\epsilon^{1/2} + \delta^{1/2} + (d/q)^{1/2}\right)$$

$$\leq \zeta.$$

Proof of Item 3 (Strong self-consistency). We begin by recording the following facts which follow from the definition of H_h^u and the projectivity of A:

$$H_h^u = A_{h(u)}^u \cdot T_h \cdot A_{h(u)}^u = A_{h(u)}^u \cdot H_h^u \cdot A_{h(u)}^u, \tag{75}$$

$$A_{h(u)}^u \cdot H_{h'}^u \cdot A_{h(u)}^u = H_{h'}^u \cdot A_{h(u)}^u = (A_{h(u)}^u \cdot T_{h'} \cdot A_{h(u)}^u) \cdot \mathbf{1}[h(u) = h'(u)]. \tag{76}$$

The strong self-consistency of H is

$$\sum_{h \in \mathcal{P}(m,q,d)} \langle\psi| \, H_h \otimes H_h \, |\psi\rangle = \mathop{\mathbf{E}}_{\boldsymbol{u} \sim \mathbb{F}_q^m} \sum_{h \in \mathcal{P}(m,q,d)} \langle\psi| \, H_h^{\boldsymbol{u}} \otimes H_h \, |\psi\rangle \,. \tag{77}$$

We now apply Lemma 9.3 to the right-hand side of Equation (77). To do so, we set

$$\mathcal{O} = \mathcal{P}(m,q,d), \ M = H, \text{ and } S_u = \{(h,h) : h \in \mathcal{P}(m,q,d)\}.$$

Then Lemma 9.3 implies that

$$(77) \approx_{4\sqrt{\zeta_{\text{variance}}}} \mathop{\mathbf{E}}_{\boldsymbol{u} \sim \mathbb{F}_q^m} \sum_{h \in \mathcal{P}(m,q,d)} \langle\psi| \, (A_{h(\boldsymbol{u})}^{\boldsymbol{u}} \cdot H_h^{\boldsymbol{u}} \cdot A_{h(\boldsymbol{u})}^{\boldsymbol{u}}) \otimes T_h \, |\psi\rangle \,. \tag{78}$$

Having placed two A's on the left-hand side, we want to show that

$$(78) \approx_{2\sqrt{\zeta_{\text{variance}}} + \frac{md}{q}} \mathop{\mathbf{E}}_{\boldsymbol{u} \sim \mathbb{F}_q^m} \sum_{h,h' \in \mathcal{P}(m,q,d)} \langle\psi| \, (A_{h(\boldsymbol{u})}^{\boldsymbol{u}} \cdot H_{h'}^{\boldsymbol{u}} \cdot A_{h(\boldsymbol{u})}^{\boldsymbol{u}}) \otimes T_h \, |\psi\rangle \,. \tag{79}$$

We note that (79) is at least as big as (78). Thus, we want to upper-bound

$$
\begin{aligned}
(79) - (78) &= \mathop{\mathbf{E}}_{\boldsymbol{u} \sim \mathbb{F}_q^m} \sum_{h \neq h'} \langle\psi| \, (A_{h(\boldsymbol{u})}^{\boldsymbol{u}} \cdot H_{h'}^{\boldsymbol{u}} \cdot A_{h(\boldsymbol{u})}^{\boldsymbol{u}}) \otimes T_h \, |\psi\rangle \\
&= \mathop{\mathbf{E}}_{\boldsymbol{u} \sim \mathbb{F}_q^m} \sum_{h \neq h'} \langle\psi| \, (A_{h(\boldsymbol{u})}^{\boldsymbol{u}} \cdot T_{h'} \cdot A_{h(\boldsymbol{u})}^{\boldsymbol{u}}) \otimes T_h \, |\psi\rangle \cdot \mathbf{1}[h(\boldsymbol{u}) = h'(\boldsymbol{u})],
\end{aligned} \tag{80}
$$

where the second equality is by Equation (76). To do this, we first show that

$$(80) \approx_{\sqrt{\zeta_{\text{variance}}}} \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{h \neq h'} \langle\psi| \, (A_{h(\boldsymbol{v})}^{\boldsymbol{v}} \cdot T_{h'} \cdot A_{h(\boldsymbol{u})}^{\boldsymbol{u}}) \otimes T_h \, |\psi\rangle \cdot \mathbf{1}[h(\boldsymbol{u}) = h'(\boldsymbol{u})]. \tag{81}$$

To show this, we bound the magnitude of the difference.

$$
\left| \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{h \neq h'} \langle\psi| \, ((A_{h(\boldsymbol{u})}^{\boldsymbol{u}} - A_{h(\boldsymbol{v})}^{\boldsymbol{v}}) \cdot T_{h'} \cdot A_{h(\boldsymbol{u})}^{\boldsymbol{u}}) \otimes T_h \, |\psi\rangle \cdot \mathbf{1}[h(\boldsymbol{u}) = h'(\boldsymbol{u})] \right|
$$

$$
\leq \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{h \neq h'} \langle\psi| \, ((A_{h(\boldsymbol{u})}^{\boldsymbol{u}} - A_{h(\boldsymbol{v})}^{\boldsymbol{v}}) \cdot T_{h'} \cdot (A_{h(\boldsymbol{u})}^{\boldsymbol{u}} - A_{h(\boldsymbol{v})}^{\boldsymbol{v}})) \otimes T_h \, |\psi\rangle}
$$

$$
\cdot \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{h \neq h'} \langle\psi| \, (A_{h(\boldsymbol{u})}^{\boldsymbol{u}} \cdot T_{h'} \cdot A_{h(\boldsymbol{u})}^{\boldsymbol{u}}) \otimes T_h \, |\psi\rangle \cdot \mathbf{1}[h(\boldsymbol{u}) = h'(\boldsymbol{u})]} \,. \tag{82}
$$

The term inside the first square root is

$$
\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_h \langle\psi| \, ((A_{h(\boldsymbol{u})}^{\boldsymbol{u}} - A_{h(\boldsymbol{v})}^{\boldsymbol{v}}) \cdot \left(\sum_{h' \neq h} T_{h'} \right) \cdot (A_{h(\boldsymbol{u})}^{\boldsymbol{u}} - A_{h(\boldsymbol{v})}^{\boldsymbol{v}})) \otimes T_h \, |\psi\rangle
$$

$$
\leq \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_h \langle\psi| \, (A_{h(\boldsymbol{u})}^{\boldsymbol{u}} - A_{h(\boldsymbol{v})}^{\boldsymbol{v}})^2 \otimes T_h \, |\psi\rangle \,, \qquad \text{(because } T \text{ is a sub-measurement)}
$$

76

But $T \in \mathcal{P}(m, q, d)$, and so by Lemma 8.3 this expression is at most ζ_{variance}. The term inside the second square root is equal to

$$\mathop{\mathbf{E}}_{\boldsymbol{u}, \boldsymbol{v}} \sum_{h \neq h'} \langle \psi | H^{\boldsymbol{u}}_{h'} \otimes T_h |\psi\rangle \cdot \mathbf{1}[h(\boldsymbol{u}) = h'(\boldsymbol{u})],$$

which is at most 1 because T and $H^{\boldsymbol{u}}$ are sub-measurements. Next, we show that

$$(81) \approx_{\sqrt{\zeta_{\text{variance}}}} \mathop{\mathbf{E}}_{\boldsymbol{u}, \boldsymbol{v}} \sum_{h \neq h'} \langle \psi | (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \cdot T_{h'} \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \otimes T_h |\psi\rangle \cdot \mathbf{1}[h(\boldsymbol{u}) = h'(\boldsymbol{u})]. \tag{83}$$

To show this, we bound the magnitude of the difference.

$$\left| \mathop{\mathbf{E}}_{\boldsymbol{u}, \boldsymbol{v}} \sum_{h \neq h'} \langle \psi | (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \cdot T_{h'} \cdot (A^{\boldsymbol{u}}_{h(\boldsymbol{u})} - A^{\boldsymbol{v}}_{h(\boldsymbol{v})})) \otimes T_h |\psi\rangle \cdot \mathbf{1}[h(\boldsymbol{u}) = h'(\boldsymbol{u})] \right|$$

$$\leq \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u}, \boldsymbol{v}} \sum_{h \neq h'} \langle \psi | (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \cdot T_{h'} \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \otimes T_h |\psi\rangle \cdot \mathbf{1}[h(\boldsymbol{u}) = h'(\boldsymbol{u})]}$$

$$\cdot \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u}, \boldsymbol{v}} \sum_{h \neq h'} \langle \psi | ((A^{\boldsymbol{u}}_{h(\boldsymbol{u})} - A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \cdot T_{h'} \cdot (A^{\boldsymbol{u}}_{h(\boldsymbol{u})} - A^{\boldsymbol{v}}_{h(\boldsymbol{v})})) \otimes T_h |\psi\rangle}.$$

The term inside the first square root is at most

$$\mathop{\mathbf{E}}_{\boldsymbol{u}, \boldsymbol{v}} \sum_{h \neq h'} \langle \psi | (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \cdot T_{h'} \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \otimes T_h |\psi\rangle = \mathop{\mathbf{E}}_{\boldsymbol{u}, \boldsymbol{v}} \sum_{h} \langle \psi | (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \cdot \left(\sum_{h' \neq h} T_{h'} \right) \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \otimes T_h |\psi\rangle \tag{84}$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{u}, \boldsymbol{v}} \sum_{h} \langle \psi | (A^{\boldsymbol{v}}_{h(\boldsymbol{v})})^2 \otimes T_h |\psi\rangle \quad (T \text{ is a sub-measurement})$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{u}, \boldsymbol{v}} \sum_{h} \langle \psi | I \otimes T_h |\psi\rangle \quad (\text{because } A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \leq I)$$

$$\leq 1, \tag{85}$$

where the last step again uses the fact that T is a sub-measurement. As for the term inside the second square root, it is equal to the term inside the first square root in Equation (82), which we showed was at most ζ_{variance}. Finally, Equation (83) is equal to

$$\mathop{\mathbf{E}}_{\boldsymbol{v}} \sum_{h \neq h'} \langle \psi | (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \cdot T_{h'} \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \otimes T_h |\psi\rangle \cdot \mathop{\mathbf{E}}_{\boldsymbol{u}} \mathbf{1}[h(\boldsymbol{u}) = h'(\boldsymbol{u})]$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{v}} \sum_{h \neq h'} \langle \psi | (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \cdot T_{h'} \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \otimes T_h |\psi\rangle \cdot \frac{md}{q} \quad (\text{by Schwartz-Zippel})$$

$$\leq \frac{md}{q}. \quad (\text{by Equation (85)})$$

By Equation (76), Equation (79) is equal to

$$\mathop{\mathbf{E}}_{\boldsymbol{u} \sim \mathbb{F}_q^m} \sum_{h, h' \in \mathcal{P}(m, q, d)} \langle \psi | (H^{\boldsymbol{u}}_{h'} \cdot A^{\boldsymbol{u}}_{h(\boldsymbol{u})}) \otimes T_h |\psi\rangle . \tag{86}$$

Now, we show that

$$(86) \approx_{\sqrt{\zeta_{\text{variance}}}} \mathop{\mathbf{E}}_{\boldsymbol{u}, \boldsymbol{v}} \sum_{h, h' \in \mathcal{P}(m, q, d)} \langle \psi | (H^{\boldsymbol{u}}_{h'} \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \otimes T_h |\psi\rangle . \tag{87}$$

To show this, we bound the magnitude of the difference.

$$
\left| \underset{\boldsymbol{u},\boldsymbol{v}}{\mathbf{E}} \sum_{h,h' \in \mathcal{P}(m,q,d)} \langle \psi | \, (H^{\boldsymbol{u}}_{h'} \cdot (A^{\boldsymbol{u}}_{h(\boldsymbol{u})} - A^{\boldsymbol{v}}_{h(\boldsymbol{v})})) \otimes T_h \, | \psi \rangle \, \right|
$$

$$
\leq \sqrt{ \underset{\boldsymbol{u},\boldsymbol{v}}{\mathbf{E}} \sum_{h,h' \in \mathcal{P}(m,q,d)} \langle \psi | \, H^{\boldsymbol{u}}_{h'} \otimes T_h \, | \psi \rangle }
$$

$$
\cdot \sqrt{ \underset{\boldsymbol{u},\boldsymbol{v}}{\mathbf{E}} \sum_{h,h' \in \mathcal{P}(m,q,d)} \langle \psi | \, ((A^{\boldsymbol{u}}_{h(\boldsymbol{u})} - A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \cdot H^{\boldsymbol{u}}_{h'} \cdot (A^{\boldsymbol{u}}_{h(\boldsymbol{u})} - A^{\boldsymbol{v}}_{h(\boldsymbol{v})})) \otimes T_h \, | \psi \rangle } .
$$

The expression inside the first square root is at most 1 because T and $H^{\boldsymbol{u}}$ are sub-measurements. The term inside the second square root is

$$
\underset{\boldsymbol{u},\boldsymbol{v}}{\mathbf{E}} \sum_{h} \langle \psi | \, ((A^{\boldsymbol{u}}_{h(\boldsymbol{u})} - A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \cdot \left(\sum_{h'} H^{\boldsymbol{u}}_{h'} \right) \cdot (A^{\boldsymbol{u}}_{h(\boldsymbol{u})} - A^{\boldsymbol{v}}_{h(\boldsymbol{v})})) \otimes T_h \, | \psi \rangle
$$

$$
\leq \underset{\boldsymbol{u},\boldsymbol{v}}{\mathbf{E}} \sum_{h} \langle \psi | \, (A^{\boldsymbol{u}}_{h(\boldsymbol{u})} - A^{\boldsymbol{v}}_{h(\boldsymbol{v})})^2 \otimes T_h \, | \psi \rangle . \qquad \text{(because } H^{\boldsymbol{u}} \text{ is a sub-measurement)}
$$

But $T \in \mathrm{PolySub}(m,q,d)$, and so by Lemma 8.3 this expression is at most $\zeta_{\mathrm{variance}}$. Next, we show that

$$
(87) \approx_{\sqrt{2\delta}} \underset{\boldsymbol{u},\boldsymbol{v}}{\mathbf{E}} \sum_{h,h' \in \mathcal{P}(m,q,d)} \langle \psi | \, H^{\boldsymbol{u}}_{h'} \otimes (T_h \cdot A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \, | \psi \rangle . \tag{88}
$$

To show this, we bound the magnitude of the difference.

$$
\left| \underset{\boldsymbol{u},\boldsymbol{v}}{\mathbf{E}} \sum_{h,h' \in \mathcal{P}(m,q,d)} \langle \psi | \, (H^{\boldsymbol{u}}_{h'} \otimes T_h) \cdot (A^{\boldsymbol{v}}_{h(\boldsymbol{v})} \otimes I - I \otimes A^{\boldsymbol{v}}_{h(\boldsymbol{v})}) \, | \psi \rangle \, \right|
$$

$$
= \left| \underset{\boldsymbol{u},\boldsymbol{v}}{\mathbf{E}} \sum_{a \in \mathbb{F}_q} \langle \psi | \, \left(\sum_{h'} H^{\boldsymbol{u}}_{h'} \otimes \sum_{h:h(\boldsymbol{v})=a} T_h \right) \cdot (A^{\boldsymbol{v}}_a \otimes I - I \otimes A^{\boldsymbol{v}}_a) \, | \psi \rangle \, \right|
$$

$$
\leq \sqrt{ \underset{\boldsymbol{u},\boldsymbol{v}}{\mathbf{E}} \sum_{a \in \mathbb{F}_q} \langle \psi | \, \left(\sum_{h'} H^{\boldsymbol{u}}_{h'} \otimes \sum_{h:h(\boldsymbol{v})=a} T_h \right)^2 | \psi \rangle } \cdot \sqrt{ \underset{\boldsymbol{u},\boldsymbol{v}}{\mathbf{E}} \sum_{a \in \mathbb{F}_q} \langle \psi | \, (A^{\boldsymbol{v}}_a \otimes I - I \otimes A^{\boldsymbol{v}}_a)^2 | \psi \rangle } .
$$

The expression inside the first square root is at most 1 because T and $H^{\boldsymbol{u}}$ are sub-measurements, and the expression inside the second square root is at most 2δ by Proposition 4.22 and the self-

78

consistency of A. Now, Equation (88) is equal to

$$\mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_{h,h'} \langle\psi| H_{h'}^{\boldsymbol{u}} \otimes (T_h \cdot \mathop{\mathbf{E}}_{\boldsymbol{v}} A_{h(\boldsymbol{v})}^{\boldsymbol{v}}) |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_{h,h'} \langle\psi| H_{h'}^{\boldsymbol{u}} \otimes (T_h \cdot Z) |\psi\rangle \qquad \text{(by Equation (60))}$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_{h'} \langle\psi| H_{h'}^{\boldsymbol{u}} \otimes Z |\psi\rangle \qquad \text{(because } T \text{ is a measurement)}$$

$$\geq \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_{h'} \langle\psi| H_{h'}^{\boldsymbol{u}} \otimes \mathop{\mathbf{E}}_{\boldsymbol{v}} A_{h'(\boldsymbol{v})}^{\boldsymbol{v}} |\psi\rangle \qquad \text{(by Equation (59))}$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v}} \sum_{h'} \langle\psi| H_{h'}^{\boldsymbol{u}} \otimes A_{h'(\boldsymbol{v})}^{\boldsymbol{v}} |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{v}} \sum_{a} \langle\psi| H_{[h(\boldsymbol{v})=a]} \otimes A_a^{\boldsymbol{v}} |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{v}} \sum_{a} \langle\psi| H_{[h(\boldsymbol{v})=a]} \otimes I |\psi\rangle - \mathop{\mathbf{E}}_{\boldsymbol{v}} \sum_{a \neq b} \langle\psi| H_{[h(\boldsymbol{v})=a]} \otimes A_b^{\boldsymbol{v}} |\psi\rangle \qquad \text{(because } A \text{ is a measurement)}$$

$$\geq \mathop{\mathbf{E}}_{\boldsymbol{v}} \sum_{a} \langle\psi| H_{[h(\boldsymbol{v})=a]} \otimes I |\psi\rangle - 4\sqrt{\zeta_{\text{variance}}} \qquad \text{(by Equation (74))}$$

$$= \sum_{h} \langle\psi| H_h \otimes I |\psi\rangle - 4\sqrt{\zeta_{\text{variance}}}.$$

In total, we have shown that

$$\sum_{h\in\mathcal{P}(m,q,d)} \langle\psi| H_h \otimes H_h |\psi\rangle \geq \sum_{h} \langle\psi| H_h \otimes I |\psi\rangle - 7\sqrt{\zeta_{\text{variance}}} - \sqrt{2\delta} - \frac{md}{q} - 4\sqrt{\zeta_{\text{variance}}}.$$

Because

$$11\sqrt{\zeta_{\text{variance}}} + \sqrt{2\delta} + \frac{md}{q} = 11\sqrt{24m \cdot \left(\epsilon + \delta + \frac{md}{q}\right)} + \sqrt{2\delta} + \frac{md}{q}$$

$$\leq 55m \cdot \left(\epsilon^{1/2} + \delta^{1/2} + (d/q)^{1/2}\right) + 2\sqrt{\delta} + m \cdot (d/q)^{1/2}$$

$$\leq 57m \cdot \left(\epsilon^{1/2} + \delta^{1/2} + (d/q)^{1/2}\right)$$

$$\leq \zeta,$$

this concludes the proof of Item 3.

Proof of Item 4 (Boundedness). The boundedness of H is

$$\langle\psi| Z \otimes I |\psi\rangle - \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_a \langle\psi| A_a^{\boldsymbol{u}} \otimes H_{[h(\boldsymbol{u})=a]} |\psi\rangle$$

$$= \langle\psi| Z \otimes I |\psi\rangle - \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_h \langle\psi| A_{h(\boldsymbol{u})}^{\boldsymbol{u}} \otimes H_h |\psi\rangle$$

$$= \langle\psi| Z \otimes I |\psi\rangle - \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_h \langle\psi| I \otimes H_h |\psi\rangle + \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_{a,h:h(\boldsymbol{u})\neq a} \langle\psi| A_a^{\boldsymbol{u}} \otimes H_h |\psi\rangle$$

$$\hspace{8cm}\text{(because } A \text{ is a measurement)}$$

$$\leq \langle\psi| Z \otimes I |\psi\rangle - \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_h \langle\psi| I \otimes H_h |\psi\rangle + 4\sqrt{\zeta_{\text{variance}}} \hspace{2cm}\text{(by Equation (74))}$$

$$\leq 3\sqrt{\delta} + 4\sqrt{\zeta_{\text{variance}}}. \hspace{5cm}\text{(by Equation (72))}$$

We can bound the error by

$$3\sqrt{\delta} + 4\sqrt{\zeta_{\text{variance}}} = 3\sqrt{\delta} + 4\sqrt{24m \cdot \left(\epsilon + \delta + \frac{md}{q}\right)}$$

$$\leq 3\sqrt{\delta} + 20m \cdot \left(\epsilon^{1/2} + \delta^{1/2} + (d/q)^{1/2}\right)$$

$$\leq 23m \cdot \left(\epsilon^{1/2} + \delta^{1/2} + (d/q)^{1/2}\right)$$

$$\leq \zeta.$$

This completes the proof. $\square$

9.3 Self-improving to a projective measurement

We now prove the full self-improvement theorem, i.e. Theorem 6.2. To do so, we will apply the orthonormalization lemma Theorem 5.2 to the output of Lemma 9.1 and argue that it maintains the four properties of H.

Theorem 9.4 (Self-improvement; Theorem 6.2 restated)**.** *Let $G \in \text{PolyMeas}(m,q,d)$ be a measurement with the following properties:*

1. *(Consistency with A): On average over $\boldsymbol{u} \sim \mathbb{F}_q^m$,*

$$A_a^u \otimes I \simeq_\nu I \otimes G_{[g(u)=a]}.$$

Let

$$\zeta = 3000m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + (d/q)^{1/32}\right).$$

Then there exists a projective sub-measurement $H \in \text{PolySub}(m,q,d)$ with the following properties:

1. *(Completeness): If $H = \sum_h H_h$, then*

$$\langle\psi| H \otimes I |\psi\rangle \geq (1-\nu) - \zeta.$$

2. *(Consistency with A): On average over $\boldsymbol{u} \sim \mathbb{F}_q^m$,*

$$A_a^u \otimes I \simeq_\zeta I \otimes H_{[h(u)=a]}.$$

80

3. (*Strong self-consistency*):
$$H_h \otimes I \approx_\zeta I \otimes H_h.$$

4. (*Boundedness*): *There exists a positive-semidefinite matrix Z such that*
$$\langle\psi| Z \otimes (I - H) |\psi\rangle \le \zeta$$

and for each $h \in \mathcal{P}(m, q, d)$,
$$Z \ge \left(\mathop{\mathbf{E}}_{\boldsymbol{u}} A^{\boldsymbol{u}}_{h(\boldsymbol{u})}\right).$$

Proof. We note that the bound we are proving is trivial when at least one of ϵ, δ, or d/q is ≥ 1. In this case, $\zeta \ge 3000$. Hence, we may assume that $\gamma, \zeta, d/q \le 1$. This will aid us when carrying out the error calculations near the end of the proof, as it allows us to bound terms like $\epsilon^{1/2}$ by terms like $\epsilon^{1/4}$.

To begin, apply Lemma 9.1 to G. Let $\widehat{H} \in \mathrm{PolySub}(m, q, d)$ be the sub-measurement it outputs and let $\widehat{\zeta}$ be the error
$$\widehat{\zeta} = 100m \cdot \left(\epsilon^{1/2} + \delta^{1/2} + (d/q)^{1/2}\right).$$

By Item 3 of Lemma 9.1,

$$\sum_h \langle\psi| \widehat{H}_h \otimes \widehat{H}_h |\psi\rangle \ge \langle\psi| \widehat{H} \otimes I |\psi\rangle - \widehat{\zeta}.$$

Thus, Theorem 5.2 implies the existence of a projective sub-measurement $H \in \mathrm{PolySub}(m, q, d)$ such that
$$\widehat{H}_h \otimes I \approx_{\widehat{\zeta}_{\mathrm{ortho}}} H_h \otimes I, \tag{89}$$

where
$$\widehat{\zeta}_{\mathrm{ortho}} = 100\widehat{\zeta}^{1/4}.$$

In addition, Proposition 4.39 implies that
$$\widehat{H}_{[h(u)=a]} \otimes I \approx_{\widehat{\zeta}_{\mathrm{dataprocess}}} H_{[h(u)=a]} \otimes I, \tag{90}$$

where
$$\widehat{\zeta}_{\mathrm{dataprocess}} = 8\widehat{\zeta} + 8\sqrt{\widehat{\zeta}_{\mathrm{ortho}}}.$$

Now we prove the four properties of this theorem. We will show that each quantity is bounded by some error, and then at the end of this proof we will show that all four errors are bounded by ζ.

1. (*Completeness*): Proposition 4.38 implies that

$$\begin{aligned}
\langle\psi| H \otimes I |\psi\rangle &\ge \langle\psi| \widehat{H} \otimes I |\psi\rangle - \widehat{\zeta} - 2\sqrt{\widehat{\zeta}_{\mathrm{ortho}}} \\
&\ge (1 - \nu) - 2\widehat{\zeta} - 2\sqrt{\widehat{\zeta}_{\mathrm{ortho}}}. \qquad \text{(by Item 1 of Lemma 9.1)}
\end{aligned}$$

2. (*Consistency with A*): Proposition 4.21 applied to Item 2 of Lemma 9.1 implies that

$$A^u_a \otimes I \simeq_{\widehat{\zeta}+\sqrt{\widehat{\zeta}_{\mathrm{dataprocess}}}} I \otimes H_{[h(u)=a]}.$$

81

3. (Strong self-consistency): [Proposition 4.36] implies that

$$\widehat{H}_h \otimes I \approx_{2\widehat{\zeta}} I \otimes \widehat{H}_h.$$

Thus,

$$H_h \otimes I \approx_{\widehat{\zeta}_{\mathrm{ortho}}} \widehat{H}_h \otimes I \approx_{2\widehat{\zeta}} I \otimes \widehat{H}_h \approx_{\widehat{\zeta}_{\mathrm{ortho}}} I \otimes H_h.$$

Hence, by [Proposition 4.28],

$$H_h \otimes I \approx_{6\widehat{\zeta}+6\widehat{\zeta}_{\mathrm{ortho}}} I \otimes H_h.$$

4. (Boundedness): The boundedness of H is

$$
\begin{aligned}
& \langle\psi|\, Z \otimes (I - H)\, |\psi\rangle \\
=\ & \langle\psi|\, Z \otimes I\, |\psi\rangle - \sum_h \langle\psi|\, Z \otimes H_h\, |\psi\rangle \\
\leq\ & \langle\psi|\, Z \otimes I\, |\psi\rangle - \sum_h \langle\psi|\, \mathbf{E}_{\boldsymbol{u}} A^{\boldsymbol{u}}_{h(\boldsymbol{u})} \otimes H_h\, |\psi\rangle && \text{(by Item 4 of Lemma 9.1)} \\
=\ & \langle\psi|\, Z \otimes I\, |\psi\rangle - \mathbf{E}_{\boldsymbol{u}} \sum_h \langle\psi|\, A^{\boldsymbol{u}}_{h(\boldsymbol{u})} \otimes H_h\, |\psi\rangle \\
=\ & \langle\psi|\, Z \otimes I\, |\psi\rangle - \mathbf{E}_{\boldsymbol{u}} \sum_a \langle\psi|\, A^{\boldsymbol{u}}_a \otimes H_{[h(\boldsymbol{u})=a]}\, |\psi\rangle . && (91)
\end{aligned}
$$

By [Proposition 4.25],

$$\mathbf{E}_{\boldsymbol{u}} \sum_a \langle\psi|\, A^{\boldsymbol{u}}_a \otimes H_{[h(\boldsymbol{u})=a]}\, |\psi\rangle \approx_{\sqrt{\widehat{\zeta}_{\mathrm{dataprocess}}}} \mathbf{E}_{\boldsymbol{u}} \sum_a \langle\psi|\, A^{\boldsymbol{u}}_a \otimes \widehat{H}_{[h(\boldsymbol{u})=a]}\, |\psi\rangle .$$

Hence,

$$
\begin{aligned}
(91) \leq\ & \langle\psi|\, Z \otimes I\, |\psi\rangle - \mathbf{E}_{\boldsymbol{u}} \sum_a \langle\psi|\, A^{\boldsymbol{u}}_a \otimes \widehat{H}_{[h(\boldsymbol{u})=a]}\, |\psi\rangle + \sqrt{\widehat{\zeta}_{\mathrm{dataprocess}}} \\
\leq\ & \widehat{\zeta} + \sqrt{\widehat{\zeta}_{\mathrm{dataprocess}}}. && \text{(by Item 4 of Lemma 9.1)}
\end{aligned}
$$

This shows the four properties hold with errors

$$2\widehat{\zeta} + 2\sqrt{\widehat{\zeta}_{\mathrm{ortho}}}, \quad \widehat{\zeta} + \sqrt{\widehat{\zeta}_{\mathrm{dataprocess}}}, \quad 6\widehat{\zeta} + 6\widehat{\zeta}_{\mathrm{ortho}}, \quad \text{and} \quad \widehat{\zeta} + \sqrt{\widehat{\zeta}_{\mathrm{dataprocess}}},$$

respectively. We now show that these four are bounded by ζ. First, using the fact that $100^{1/4} \leq 4$, we note that

$$
\begin{aligned}
\widehat{\zeta}_{\mathrm{ortho}} = 100\widehat{\zeta}^{1/4} &= 100\Big(100m \cdot \big(\epsilon^{1/2} + \delta^{1/2} + (d/q)^{1/2}\big)\Big)^{1/4} \\
&\leq 400m \cdot \big(\epsilon^{1/8} + \delta^{1/8} + (d/q)^{1/8}\big).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
6\widehat{\zeta} + 6\widehat{\zeta}_{\mathrm{ortho}} &\leq 6\Big(100m \cdot \big(\epsilon^{1/2} + \delta^{1/2} + (d/q)^{1/2}\big)\Big) + 6\Big(400m \cdot \big(\epsilon^{1/8} + \delta^{1/8} + (d/q)^{1/8}\big)\Big) \\
&\leq 600m \cdot \big(\epsilon^{1/8} + \delta^{1/8} + (d/q)^{1/8}\big) + 2400m \cdot \big(\epsilon^{1/8} + \delta^{1/8} + (d/q)^{1/8}\big) \\
&= 3000m \cdot \big(\epsilon^{1/8} + \delta^{1/8} + (d/q)^{1/8}\big),
\end{aligned}
$$

which is less than ζ. In addition, using the fact that $\sqrt{400} = 20$, we also note that

$$
\begin{aligned}
\widehat{\zeta}_{\text{dataprocess}} &= 8\widehat{\zeta} + 8\sqrt{\widehat{\zeta}_{\text{ortho}}} \\
&\leq 8\left(100m \cdot \left(\epsilon^{1/2} + \delta^{1/2} + (d/q)^{1/2}\right)\right) + 8\sqrt{400m \cdot \left(\epsilon^{1/8} + \delta^{1/8} + (d/q)^{1/8}\right)} \\
&\leq 800m \cdot \left(\epsilon^{1/16} + \delta^{1/16} + (d/q)^{1/16}\right) + 160m \cdot \left(\epsilon^{1/16} + \delta^{1/16} + (d/q)^{1/16}\right) \\
&= 960m \cdot \left(\epsilon^{1/16} + \delta^{1/16} + (d/q)^{1/16}\right),
\end{aligned}
$$

which is clearly less than ζ. Thus, $2\widehat{\zeta} + 2\sqrt{\widehat{\zeta}_{\text{ortho}}} \leq \widehat{\zeta}_{\text{dataprocess}} \leq \zeta$. Finally, using $\sqrt{960} \leq 31$,

$$
\begin{aligned}
\widehat{\zeta} + \sqrt{\widehat{\zeta}_{\text{dataprocess}}} &\leq 100m \cdot \left(\epsilon^{1/2} + \delta^{1/2} + (d/q)^{1/2}\right) + \sqrt{960m \cdot \left(\epsilon^{1/16} + \delta^{1/16} + (d/q)^{1/16}\right)} \\
&\leq 100m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + (d/q)^{1/32}\right) + 31m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + (d/q)^{1/32}\right) \\
&= 131m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + (d/q)^{1/32}\right),
\end{aligned}
$$

which is also less than ζ. This completes the proof. $\qquad\square$

10   Commutativity of the points measurements

Theorem 10.1. *Let (ψ, A, B, L) be an $(\epsilon, \delta, \gamma)$-good symmetric strategy for the (m, q, d) low individual degree test. On average over independent and uniformly random $\boldsymbol{u}, \boldsymbol{v} \sim \mathbb{F}_q^m$,*

$$
(A_a^u \cdot A_b^v) \otimes I \approx_{32\gamma m} (A_b^v \cdot A_a^u) \otimes I.
$$

Proof. The strategy passes the diagonal lines test with probability $1 - \gamma$. Therefore, it passes the m-restricted diagonal lines test with probability $1 - \gamma \cdot m$. This means that

$$
A_a^u \otimes I \simeq_{\gamma \cdot m} I \otimes L_{[f(u)=a]}^\ell,
$$

on average over a uniformly random $\boldsymbol{u} \sim \mathbb{F}_q^m$ and a uniformly random line $\boldsymbol{\ell}$ in $\mathbb{F}_q^m$ containing $\boldsymbol{u}$. By Proposition 4.22, this implies that

$$
A_a^u \otimes I \approx_{2 \cdot \gamma \cdot m} I \otimes L_{[f(u)=a]}^\ell. \tag{92}
$$

Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be independent and uniformly random points in $\mathbb{F}_q^m$. Let $\boldsymbol{\ell}$ be a uniformly random line in $\mathbb{F}_q^m$ containing both points. (If $\boldsymbol{u}$ and $\boldsymbol{v}$ are distinct, then $\boldsymbol{\ell}$ is just the unique line that passes through both of them. Otherwise, $\boldsymbol{u} = \boldsymbol{v}$, and $\boldsymbol{\ell}$ is a uniformly random line passing through $\boldsymbol{u}$.) Then the marginal distribution on $\boldsymbol{u}$ and $\boldsymbol{\ell}$ is as in Equation (92), as is the marginal distribution on $\boldsymbol{v}$ and $\boldsymbol{\ell}$. As a result,

$$
\begin{aligned}
(A_a^u \cdot A_b^v) \otimes I &\approx_{2 \cdot \gamma \cdot m} A_a^u \otimes L_{[f(v)=b]}^\ell && \text{(by Equation (92))} \\
&\approx_{2 \cdot \gamma \cdot m} I \otimes (L_{[f(v)=b]}^\ell \cdot L_{[f'(u)=a]}^\ell) && \text{(by Equation (92))} \\
&= I \otimes (L_{[f'(u)=a]}^\ell \cdot L_{[f(v)=b]}^\ell) && \text{(because L is projective)} \\
&\approx_{2 \cdot \gamma \cdot m} A_b^v \otimes L_{[f'(u)=a]}^\ell && \text{(by Equation (92))} \\
&\approx_{2 \cdot \gamma \cdot m} (A_b^v \cdot A_a^u) \otimes I. && \text{(by Equation (92))}
\end{aligned}
$$

The theorem now follows from Proposition 4.28. $\qquad\square$

11 Commutativity

Let $(\boldsymbol{u}, \boldsymbol{x})$ and $(\boldsymbol{v}, \boldsymbol{y})$ be sampled independently and uniformly at random from $\mathbb{F}_q^{m+1}$. In Section 11.1, we will show that the G measurements approximately commute "after evaluation"; namely, that $G^{\boldsymbol{x}}_{[g(\boldsymbol{u})=a]}$ commutes with $G^{\boldsymbol{y}}_{[g(\boldsymbol{v})=b]}$. Then, in Section 11.2, we will use this to show that $G^{\boldsymbol{x}}_g$ approximately commutes with $G^{\boldsymbol{y}}_h$. This is necessary if we wish to "paste" together several $G^{\tilde{x}}$ measurements at different points $x \in \mathbb{F}_q$ to produce a single global measurement $H \in \text{PolySub}(m+1, q, d)$, as we do in Section 12 below.

11.1 Commutativity of G after evaluation

Lemma 11.1 (Commutativity of G after evaluation). *Let (ψ, A, B, L) be an $(\epsilon, \delta, \gamma)$-good symmetric strategy for the $(m+1, q, d)$ low individual degree test. Let $\{G^x\} \in \text{PolySub}(m, q, d)$ be a collection of projective sub-measurements indexed by $x \in \mathbb{F}_q$ with the following properties:*

1. *(Consistency with A):* *On average over $(\boldsymbol{u}, \boldsymbol{x}) \sim \mathbb{F}_q^{m+1}$,*

$$A_a^{u,x} \otimes I \simeq_\zeta I \otimes G^x_{[g(u)=a]}.$$

2. *(Strong self-consistency):* *On average over $\boldsymbol{x} \sim \mathbb{F}_q$,*

$$G^x_g \otimes I \approx_\zeta I \otimes G^x_g.$$

3. *(Boundedness):* *There exists a positive-semidefinite matrix Z^x for each $x \in \mathbb{F}_q$ such that*

$$\mathbf{E}_{\boldsymbol{x}} \langle \psi | \, (I - G^{\boldsymbol{x}}) \otimes Z^{\boldsymbol{x}} \, |\psi\rangle \leq \zeta$$

and for each $x \in \mathbb{F}_q$ and $g \in \mathcal{P}(m, q, d)$,

$$Z^x \geq \left(\mathbf{E}_{\boldsymbol{u}} A_{g(\boldsymbol{u})}^{\boldsymbol{u}, x} \right).$$

Let

$$\nu = 48m \cdot (\gamma^{1/2} + \zeta^{1/2}).$$

Then on average over independent and uniformly random $(\boldsymbol{u}, \boldsymbol{x}), (\boldsymbol{v}, \boldsymbol{y}) \sim \mathbb{F}_q^{m+1}$,

$$G^x_{[g(u)=a]} G^y_{[h(v)=b]} \otimes I \approx_\nu G^y_{[h(v)=b]} G^x_{[g(u)=a]} \otimes I.$$

Proof. For notational convenience we use the abbreviation $G_a^{u,x} = G^x_{[g(u)=a]}$ for all $(u, x) \in \mathbb{F}_q^{m+1}$. We expand the square:

$$
\begin{aligned}
&\mathbf{E}_{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{x}, \boldsymbol{y}} \sum_{a,b} \langle \psi | \, (G_a^{\boldsymbol{u}, \boldsymbol{x}} G_b^{\boldsymbol{v}, \boldsymbol{y}} - G_b^{\boldsymbol{v}, \boldsymbol{y}} G_a^{\boldsymbol{u}, \boldsymbol{x}})^\dagger \cdot (G_a^{\boldsymbol{u}, \boldsymbol{x}} G_b^{\boldsymbol{v}, \boldsymbol{y}} - G_b^{\boldsymbol{v}, \boldsymbol{y}} G_a^{\boldsymbol{u}, \boldsymbol{x}}) \otimes I \, |\psi\rangle \\
= \ &\mathbf{E}_{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{x}, \boldsymbol{y}} \sum_{a,b} \langle \psi | \, (G_b^{\boldsymbol{v}, \boldsymbol{y}} G_a^{\boldsymbol{u}, \boldsymbol{x}} - G_a^{\boldsymbol{u}, \boldsymbol{x}} G_b^{\boldsymbol{v}, \boldsymbol{y}}) \cdot (G_a^{\boldsymbol{u}, \boldsymbol{x}} G_b^{\boldsymbol{v}, \boldsymbol{y}} - G_b^{\boldsymbol{v}, \boldsymbol{y}} G_a^{\boldsymbol{u}, \boldsymbol{x}}) \otimes I \, |\psi\rangle \\
= \ &2 \cdot \mathbf{E}_{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{x}, \boldsymbol{y}} \sum_{a,b} \Big(\langle \psi | G_b^{\boldsymbol{v}, \boldsymbol{y}} G_a^{\boldsymbol{u}, \boldsymbol{x}} G_b^{\boldsymbol{v}, \boldsymbol{y}} \otimes I \, |\psi\rangle - \langle \psi | G_a^{\boldsymbol{u}, \boldsymbol{x}} G_b^{\boldsymbol{v}, \boldsymbol{y}} G_a^{\boldsymbol{u}, \boldsymbol{x}} G_b^{\boldsymbol{v}, \boldsymbol{y}} \otimes I \, |\psi\rangle \Big),
\end{aligned}
\tag{93}
$$

where the last step uses the projectivity of G.

We will show that the second term of Equation (93) is close to the first term. To begin, we note that for each $(u, x) \in \mathbb{F}_q^{m+1}$,

$$G^{u,x} = \sum_a G_a^{u,x} = \sum_a G_{[g(u)=a]}^x = G^x. \tag{94}$$

As a result, Item 1 and Proposition 4.31 imply that

$$G_a^{u,x} \otimes I \approx_{4\zeta} G^{u,x} \otimes A_a^{u,x}$$
$$= G^x \otimes A_a^{u,x}, \tag{95}$$

where the second step is by Equation (94). We can therefore approximate the second term of Equation (93) as

$$\mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \langle \psi | \, G_a^{\boldsymbol{u,x}} G_b^{\boldsymbol{v,y}} G_a^{\boldsymbol{u,x}} G_b^{\boldsymbol{v,y}} \otimes I \, | \psi \rangle$$
$$\approx_{2\sqrt{\zeta}} \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \langle \psi | \, G_a^{\boldsymbol{u,x}} G_b^{\boldsymbol{v,y}} G_a^{\boldsymbol{u,x}} G^{\boldsymbol{y}} \otimes A_b^{\boldsymbol{v,y}} \, | \psi \rangle . \tag{96}$$

using Proposition 4.24 and Equation (95). Next, we claim that

$$(96) \approx_{\sqrt{\zeta}} \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \langle \psi | \, G_a^{\boldsymbol{u,x}} G_b^{\boldsymbol{v,y}} G_a^{\boldsymbol{u,x}} \otimes A_b^{\boldsymbol{v,y}} \, | \psi \rangle . \tag{97}$$

This is proved in Claim 11.2 below. Continuing, we have

$$(97) \approx_{2\sqrt{\zeta}} \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \langle \psi | \, G_a^{\boldsymbol{u,x}} G_b^{\boldsymbol{v,y}} G^{\boldsymbol{x}} \otimes A_b^{\boldsymbol{v,y}} A_a^{\boldsymbol{u,x}} \, | \psi \rangle$$
$$\approx_{6\sqrt{\gamma(m+1)}} \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \langle \psi | \, G_a^{\boldsymbol{u,x}} G_b^{\boldsymbol{v,y}} G^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{u,x}} A_b^{\boldsymbol{v,y}} \, | \psi \rangle . \tag{98}$$

The first approximation again uses Proposition 4.24 and Equation (95). The second approximation follows from Proposition 4.24 and Theorem 10.1. Next, we claim that

$$(98) \approx_{\sqrt{\zeta} + 6\sqrt{\gamma(m+1)}} \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \langle \psi | \, G_a^{\boldsymbol{u,x}} G_b^{\boldsymbol{v,y}} \otimes A_a^{\boldsymbol{u,x}} A_b^{\boldsymbol{v,y}} \, | \psi \rangle . \tag{99}$$

This is proved in Claim 11.3 below. We now apply Equation (95) twice with the help of Proposition 4.24.

$$(99) = \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \langle \psi | \, G^{\boldsymbol{x}} G_a^{\boldsymbol{u,x}} G_b^{\boldsymbol{v,y}} \otimes A_a^{\boldsymbol{u,x}} A_b^{\boldsymbol{v,y}} \, | \psi \rangle \qquad \text{(because } G \text{ is projective)}$$

$$\approx_{2\sqrt{\zeta}} \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \langle \psi | \, G_a^{\boldsymbol{u,x}} G_a^{\boldsymbol{u,x}} G_b^{\boldsymbol{v,y}} \otimes A_b^{\boldsymbol{v,y}} \, | \psi \rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \langle \psi | \, G_a^{\boldsymbol{u,x}} G_b^{\boldsymbol{v,y}} \otimes A_b^{\boldsymbol{v,y}} \, | \psi \rangle \qquad \text{(because } G \text{ is projective)}$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \langle \psi | \, G_a^{\boldsymbol{u,x}} G_b^{\boldsymbol{v,y}} G^{\boldsymbol{y}} \otimes A_b^{\boldsymbol{v,y}} \, | \psi \rangle \qquad \text{(because } G \text{ is projective)}$$

$$\approx_{2\sqrt{\zeta}} \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \langle \psi | \, G_a^{\boldsymbol{u,x}} G_b^{\boldsymbol{v,y}} G_b^{\boldsymbol{v,y}} \otimes I \, | \psi \rangle . \tag{100}$$

Now, Item 2 and the fact that G is projective allows us to apply Proposition 4.36, which states that G is $\zeta/2$-strongly self-consistent. Hence, Proposition 4.37 says that we can "post-process" its measurement outcomes:

$$G^x_{[g(u)=a]} \otimes I \approx_\zeta I \otimes G^x_{[g(u)=a]}.$$

In other words, using our abbreviation,

$$G^{u,x}_a \otimes I \approx_\zeta I \otimes G^{u,x}_a. \tag{101}$$

Applying Equation (101) twice with the help of Proposition 4.24, we conclude that

$$(100) \approx_{\sqrt\zeta} \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \langle\psi|\, G^{\boldsymbol{u,x}}_a G^{\boldsymbol{v,y}}_b \otimes G^{\boldsymbol{v,y}}_b \,|\psi\rangle$$

$$\approx_{\sqrt\zeta} \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \langle\psi|\, G^{\boldsymbol{v,y}}_b G^{\boldsymbol{u,x}}_a G^{\boldsymbol{v,y}}_b \otimes I \,|\psi\rangle.$$

Hence we have shown that the second term of Equation (93) is close to the first term.

Putting everything together, this shows that Equation (93) is bounded by

$$2 \cdot \left(2\sqrt\zeta + \sqrt\zeta + 2\sqrt\zeta + 6\sqrt{\gamma(m+1)} + \sqrt\zeta + 6\sqrt{\gamma(m+1)} + 2\sqrt\zeta + 2\sqrt\zeta + \sqrt\zeta + \sqrt\zeta\right)$$

$$= 24 \cdot \left(\sqrt{\gamma(m+1)} + \sqrt\zeta\right)$$

$$\leq 24\sqrt{m+1} \cdot \left(\sqrt\gamma + \sqrt\zeta\right)$$

$$\leq 48m \cdot \left(\sqrt\gamma + \sqrt\zeta\right),$$

and this completes the proof of the lemma, modulo the proofs of Claim 11.2 and Claim 11.3. We now prove these claims.

Claim 11.2.

$$\mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \langle\psi|\, G^{\boldsymbol{u,x}}_a G^{\boldsymbol{v,y}}_b G^{\boldsymbol{u,x}}_a G^{\boldsymbol{y}} \otimes A^{\boldsymbol{v,y}}_b \,|\psi\rangle$$

$$\approx_{\sqrt\zeta} \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \langle\psi|\, G^{\boldsymbol{u,x}}_a G^{\boldsymbol{v,y}}_b G^{\boldsymbol{u,x}}_a \otimes A^{\boldsymbol{v,y}}_b \,|\psi\rangle.$$

Proof. Recall that $G^{\boldsymbol{v,y}}_b = G^{\boldsymbol{y}}_{[g(v)=b]} = \sum_{g:g(v)=b} G^{\boldsymbol{y}}_g$. For all $y \in \mathbb{F}_q$ and $g \in \mathcal{P}(m,q,d)$, define the matrix

$$R^y_g = \mathop{\mathbf{E}}_{\boldsymbol{u,x}} \sum_a G^{\boldsymbol{u,x}}_a G^{\boldsymbol{y}}_g G^{\boldsymbol{u,x}}_a.$$

Then because G is a sub-measurement,

$$\sum_g R^y_g = \mathop{\mathbf{E}}_{\boldsymbol{u,x}} \sum_a G^{\boldsymbol{u,x}}_a \cdot \left(\sum_g G^{\boldsymbol{y}}_g\right) \cdot G^{\boldsymbol{u,x}}_a \leq \mathop{\mathbf{E}}_{\boldsymbol{u,x}} \sum_a G^{\boldsymbol{u,x}}_a \leq I.$$

As a result, R^y is a sub-measurement in $\mathrm{PolySub}(m,q,d)$.

Our goal is to bound the magnitude of

$$\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}} \sum_{a,b} \langle\psi|\, G_a^{\boldsymbol{u},\boldsymbol{x}} G_b^{\boldsymbol{v},\boldsymbol{y}} G_a^{\boldsymbol{u},\boldsymbol{x}} (I - G^{\boldsymbol{y}}) \otimes A_b^{\boldsymbol{v},\boldsymbol{y}} \,|\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}} \sum_{a,b} \sum_{g:g(\boldsymbol{v})=b} \langle\psi|\, G_a^{\boldsymbol{u},\boldsymbol{x}} G_g^{\boldsymbol{y}} G_a^{\boldsymbol{u},\boldsymbol{x}} (I - G^{\boldsymbol{y}}) \otimes A_b^{\boldsymbol{v},\boldsymbol{y}} \,|\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}} \sum_{a,g} \langle\psi|\, G_a^{\boldsymbol{u},\boldsymbol{x}} G_g^{\boldsymbol{y}} G_a^{\boldsymbol{u},\boldsymbol{x}} (I - G^{\boldsymbol{y}}) \otimes A_{g(\boldsymbol{v})}^{\boldsymbol{v},\boldsymbol{y}} \,|\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{v},\boldsymbol{y}} \sum_{g} \langle\psi|\, R_g^{\boldsymbol{y}} (I - G^{\boldsymbol{y}}) \otimes A_{g(\boldsymbol{v})}^{\boldsymbol{v},\boldsymbol{y}} \,|\psi\rangle . \tag{102}$$

We can now bound the magnitude as follows.

$$|(102)| = \left|\, \mathop{\mathbf{E}}_{\boldsymbol{v},\boldsymbol{y}} \sum_{g} \langle\psi|\, \left(\sqrt{R_g^{\boldsymbol{y}}} \otimes I\right) \cdot \left(\sqrt{R_g^{\boldsymbol{y}}}(I - G^{\boldsymbol{y}}) \otimes A_{g(\boldsymbol{v})}^{\boldsymbol{v},\boldsymbol{y}}\right) |\psi\rangle \,\right|$$

$$\leq \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{v},\boldsymbol{y}} \sum_{g} \langle\psi|\, R_g^{\boldsymbol{y}} \otimes I \,|\psi\rangle} \cdot \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{v},\boldsymbol{y}} \sum_{g} \langle\psi|\, (I - G^{\boldsymbol{y}}) R_g^{\boldsymbol{y}} (I - G^{\boldsymbol{y}}) \otimes A_{g(\boldsymbol{v})}^{\boldsymbol{v},\boldsymbol{y}} \,|\psi\rangle} .$$

The term inside the first square root is at most 1 because $R^{\boldsymbol{y}}$ is a sub-measurement. The term inside the second square root is

$$\mathop{\mathbf{E}}_{\boldsymbol{v},\boldsymbol{y}} \sum_{g} \langle\psi|\, (I - G^{\boldsymbol{y}}) R_g^{\boldsymbol{y}} (I - G^{\boldsymbol{y}}) \otimes A_{g(\boldsymbol{v})}^{\boldsymbol{v},\boldsymbol{y}} \,|\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{y}} \sum_{g} \langle\psi|\, (I - G^{\boldsymbol{y}}) R_g^{\boldsymbol{y}} (I - G^{\boldsymbol{y}}) \otimes \left(\mathop{\mathbf{E}}_{\boldsymbol{v}} A_{g(\boldsymbol{v})}^{\boldsymbol{v},\boldsymbol{y}}\right) |\psi\rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{y}} \sum_{g} \langle\psi|\, (I - G^{\boldsymbol{y}}) R_g^{\boldsymbol{y}} (I - G^{\boldsymbol{y}}) \otimes Z^{\boldsymbol{y}} \,|\psi\rangle \qquad\qquad \text{(by Item 3)}$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{y}} \langle\psi|\, (I - G^{\boldsymbol{y}}) \otimes Z^{\boldsymbol{y}} \,|\psi\rangle$$

$$\leq \zeta . \qquad\qquad \text{(by Item 3)}$$

This concludes the proof. □

Claim 11.3.

$$\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}} \sum_{a,b} \langle\psi|\, G_a^{\boldsymbol{u},\boldsymbol{x}} G_b^{\boldsymbol{v},\boldsymbol{y}} G^{\boldsymbol{x}} \otimes A_a^{\boldsymbol{u},\boldsymbol{x}} A_b^{\boldsymbol{v},\boldsymbol{y}} \,|\psi\rangle$$

$$\approx_{\sqrt{\zeta}+6\sqrt{\gamma(m+1)}} \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}} \sum_{a,b} \langle\psi|\, G_a^{\boldsymbol{u},\boldsymbol{x}} G_b^{\boldsymbol{v},\boldsymbol{y}} \otimes A_a^{\boldsymbol{u},\boldsymbol{x}} A_b^{\boldsymbol{v},\boldsymbol{y}} \,|\psi\rangle .$$

Proof. Our goal is to bound the magnitude of

$$\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}} \sum_{a,b} \langle\psi|\, G_a^{\boldsymbol{u},\boldsymbol{x}} G_b^{\boldsymbol{v},\boldsymbol{y}} (I - G^{\boldsymbol{x}}) \otimes A_a^{\boldsymbol{u},\boldsymbol{x}} A_b^{\boldsymbol{v},\boldsymbol{y}} \,|\psi\rangle$$

$$\approx_{6\sqrt{\gamma(m+1)}} \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}} \sum_{a,b} \langle\psi|\, G_a^{\boldsymbol{u},\boldsymbol{x}} G_b^{\boldsymbol{v},\boldsymbol{y}} (I - G^{\boldsymbol{x}}) \otimes A_b^{\boldsymbol{v},\boldsymbol{y}} A_a^{\boldsymbol{u},\boldsymbol{x}} \,|\psi\rangle . \tag{103}$$

where the second line follows from Proposition 4.24 and Theorem 10.1. We now proceed nearly identically to Claim 11.2.

$$(103) = \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{a,b} \sum_{g:g(\boldsymbol{u})=a} \langle\psi|\, G_g^{\boldsymbol{x}} G_b^{\boldsymbol{v,y}}(I - G^{\boldsymbol{x}}) \otimes A_b^{\boldsymbol{v,y}} A_a^{\boldsymbol{u,x}} |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{g,b} \langle\psi|\, G_g^{\boldsymbol{x}} G_b^{\boldsymbol{v,y}}(I - G^{\boldsymbol{x}}) \otimes A_b^{\boldsymbol{v,y}} A_{g(\boldsymbol{u})}^{\boldsymbol{u,x}} |\psi\rangle . \qquad (104)$$

Then we can bound the magnitude as follows.

$$|(104)| = \Big| \mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{g,b} \langle\psi| \left(\sqrt{G_g^{\boldsymbol{x}}} \otimes A_b^{\boldsymbol{v,y}} \right) \cdot \left(\sqrt{G_g^{\boldsymbol{x}}} G_b^{\boldsymbol{v,y}}(I - G^{\boldsymbol{x}}) \otimes A_{g(\boldsymbol{u})}^{\boldsymbol{u,x}} \right) |\psi\rangle \Big|$$

$$\leq \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{g,b} \langle\psi|\, G_g^{\boldsymbol{x}} \otimes A_b^{\boldsymbol{v,y}} |\psi\rangle} \cdot \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{g,b} \langle\psi|\, (I - G^{\boldsymbol{x}}) G_b^{\boldsymbol{v,y}} G_g^{\boldsymbol{x}} G_b^{\boldsymbol{v,y}}(I - G^{\boldsymbol{x}}) \otimes A_{g(\boldsymbol{u})}^{\boldsymbol{u,x}} |\psi\rangle}.$$

The term inside the first square root is at most 1 because G and A are sub-measurements. The term inside the second square root is

$$\mathop{\mathbf{E}}_{\boldsymbol{u,v,x,y}} \sum_{g,b} \langle\psi|\, (I - G^{\boldsymbol{x}}) G_b^{\boldsymbol{v,y}} G_g^{\boldsymbol{x}} G_b^{\boldsymbol{v,y}}(I - G^{\boldsymbol{x}}) \otimes A_{g(\boldsymbol{u})}^{\boldsymbol{u,x}} |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{v,x,y}} \sum_{g,b} \langle\psi|\, (I - G^{\boldsymbol{x}}) G_b^{\boldsymbol{v,y}} G_g^{\boldsymbol{x}} G_b^{\boldsymbol{v,y}}(I - G^{\boldsymbol{x}}) \otimes \left(\mathop{\mathbf{E}}_{\boldsymbol{u}} A_{g(\boldsymbol{u})}^{\boldsymbol{u,x}} \right) |\psi\rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{v,x,y}} \sum_{g,b} \langle\psi|\, (I - G^{\boldsymbol{x}}) G_b^{\boldsymbol{v,y}} G_g^{\boldsymbol{x}} G_b^{\boldsymbol{v,y}}(I - G^{\boldsymbol{x}}) \otimes Z^{\boldsymbol{x}} |\psi\rangle \qquad \text{(by Item 3)}$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{v,x,y}} \sum_{b} \langle\psi|\, (I - G^{\boldsymbol{x}}) G_b^{\boldsymbol{v,y}}(I - G^{\boldsymbol{x}}) \otimes Z^{\boldsymbol{x}} |\psi\rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{v,x,y}} \langle\psi|\, (I - G^{\boldsymbol{x}}) \otimes Z^{\boldsymbol{x}} |\psi\rangle$$

$$\leq \zeta. \qquad \text{(by Item 3)}$$

This concludes the proof. $\qquad\square$

Having proved Claim 11.2 and Claim 11.3, we conclude the proof of Lemma 11.1. $\qquad\square$

11.2 Commutativity of G

Theorem 11.4 (Commutativity of G). *Let (ψ, A, B, L) be an $(\epsilon, \delta, \gamma)$-good symmetric strategy for the (m, q, d) low individual degree test. Let $\{G^x\}_{x \in \mathbb{F}_q}$ denote a set of projective sub-measurements in $\mathrm{PolySub}(m, q, d)$ with the following properties:*

1. *(Consistency with A): On average over $(\boldsymbol{u}, \boldsymbol{x}) \sim \mathbb{F}_q^{m+1}$,*

$$A_a^{u,x} \otimes I \simeq_\zeta I \otimes G_{[g(u)=a]}^x.$$

2. *(Strong self-consistency): On average over $\boldsymbol{x} \sim \mathbb{F}_q$,*

$$G_g^x \otimes I \approx_\zeta I \otimes G_g^x.$$

3. (*Boundedness*): *There exists a positive-semidefinite matrix Z^x for each $x \in \mathbb{F}_q$ such that*

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \langle \psi | \, (I - G^{\boldsymbol{x}}) \otimes Z^{\boldsymbol{x}} \, | \psi \rangle \leq \zeta$$

and for each $x \in \mathbb{F}_q$ and $g \in \mathcal{P}(m, q, d)$,

$$Z^x \geq \left(\mathop{\mathbf{E}}_{\boldsymbol{u}} A^{\boldsymbol{u}, x}_{g(\boldsymbol{u})} \right).$$

Let

$$\nu = 30m \cdot \left(\gamma^{1/4} + \zeta^{1/4} + (d/q)^{1/4} \right).$$

Then on average over independent and uniformly random $(\boldsymbol{u}, \boldsymbol{x}), (\boldsymbol{v}, \boldsymbol{y}) \sim \mathbb{F}_q^{m+1}$,

$$G^x_g G^y_h \otimes I \approx_\nu G^y_h G^x_g \otimes I$$

Proof. We note that the bound we are proving is trivial when at least one of γ, ζ, or d/q is ≥ 1. In this case, $\nu \geq 30$. On the other hand,

$$\mathop{\mathbf{E}}_{\boldsymbol{x}, \boldsymbol{y}} \sum_{g, h} \| (G^{\boldsymbol{x}}_g G^{\boldsymbol{y}}_h \otimes I - G^{\boldsymbol{y}}_h G^{\boldsymbol{x}}_g \otimes I) \, | \psi \rangle \|^2$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}, \boldsymbol{y}} \sum_{g, h} \| (G^{\boldsymbol{x}}_g G^{\boldsymbol{y}}_h \otimes I) \, | \psi \rangle + (-G^{\boldsymbol{y}}_h G^{\boldsymbol{x}}_g \otimes I) \, | \psi \rangle \|^2$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{x}, \boldsymbol{y}} \sum_{g, h} 2 \cdot \left(\| (G^{\boldsymbol{x}}_g G^{\boldsymbol{y}}_h \otimes I) \, | \psi \rangle \|^2 + \| (G^{\boldsymbol{y}}_h G^{\boldsymbol{x}}_g \otimes I) \, | \psi \rangle \|^2 \right) \qquad \text{(by Proposition 4.27)}$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}, \boldsymbol{y}} \sum_{g, h} 4 \cdot \| (G^{\boldsymbol{x}}_g G^{\boldsymbol{y}}_h \otimes I) \, | \psi \rangle \|^2 \qquad \text{(by symmetry of the terms)}$$

$$= 4 \cdot \mathop{\mathbf{E}}_{\boldsymbol{x}, \boldsymbol{y}} \sum_{g, y} \langle \psi | \, (G^{\boldsymbol{y}}_h G^{\boldsymbol{x}}_g G^{\boldsymbol{y}}_h \otimes I) \, | \psi \rangle$$

$$\leq 4, \qquad \text{(because G is a sub-measurement)}$$

which is therefore less than ν. Hence, we may assume that $\gamma, \zeta, d/q \leq 1$. This will aid us when carrying out the error calculations near the end of the proof, as it allows us to bound terms like $\zeta^{1/2}$ by terms like $\zeta^{1/4}$.

Our goal is to bound

$$\mathop{\mathbf{E}}_{\boldsymbol{x}, \boldsymbol{y}} \sum_{g, h} \langle \psi | \, (G^{\boldsymbol{x}}_g G^{\boldsymbol{y}}_h - G^{\boldsymbol{y}}_h G^{\boldsymbol{x}}_g)^\dagger \cdot (G^{\boldsymbol{x}}_g G^{\boldsymbol{y}}_h - G^{\boldsymbol{y}}_h G^{\boldsymbol{x}}_g) \otimes I \, | \psi \rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}, \boldsymbol{y}} \sum_{g, h} \langle \psi | \, (G^{\boldsymbol{y}}_h G^{\boldsymbol{x}}_g - G^{\boldsymbol{x}}_g G^{\boldsymbol{y}}_h) \cdot (G^{\boldsymbol{x}}_g G^{\boldsymbol{y}}_h - G^{\boldsymbol{y}}_h G^{\boldsymbol{x}}_g) \otimes I \, | \psi \rangle$$

$$= 2 \cdot \mathop{\mathbf{E}}_{\boldsymbol{x}, \boldsymbol{y}} \sum_{g, h} \langle \psi | \, G^{\boldsymbol{x}}_g G^{\boldsymbol{y}}_h G^{\boldsymbol{x}}_g \otimes I \, | \psi \rangle - 2 \cdot \mathop{\mathbf{E}}_{\boldsymbol{x}, \boldsymbol{y}} \sum_{g, h} \langle \psi | \, G^{\boldsymbol{x}}_g G^{\boldsymbol{y}}_h G^{\boldsymbol{x}}_g G^{\boldsymbol{y}}_h \otimes I \, | \psi \rangle. \qquad (105)$$

We will show that both terms in Equation (105) are close to $\langle \psi | \, G \otimes G \, | \psi \rangle$, where we recall that $G = \mathbf{E}_{\boldsymbol{x}} \, G^{\boldsymbol{x}}$. With the errors we derive, this will imply Theorem 11.4.

For the first term in Equation (105), we have

$$\mathop{\mathbf{E}}_{\boldsymbol{x}, \boldsymbol{y}} \sum_{g, h} \langle \psi | \, G^{\boldsymbol{x}}_g G^{\boldsymbol{y}}_h G^{\boldsymbol{x}}_g \otimes I \, | \psi \rangle = \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{g} \langle \psi | \, (G^{\boldsymbol{x}}_g \cdot G \cdot G^{\boldsymbol{x}}_g) \otimes I \, | \psi \rangle$$

$$\approx_{2\sqrt{\zeta}} \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{g} \langle \psi | \, G \otimes G^{\boldsymbol{x}}_g \, | \psi \rangle$$

$$= \langle \psi | \, G \otimes G \, | \psi \rangle,$$

where the approximation is by Proposition 4.32 and Item 2.

For the second term in Equation (105), we begin with the following lemma.

Lemma 11.5. *Let $P = \{P_a\}$ be a sub-measurement and $Q = \{Q_b\}$ be a projective sub-measurement. Define $C_{a,b} = Q_b \cdot P_a \cdot Q_b$. Then $\sum_a (\sum_b C_{a,b})^\dagger (\sum_b C_{a,b}) = \sum_a (\sum_b C_{a,b})(\sum_b C_{a,b})^\dagger \leq I$.*

Proof. The first equality follows from the fact that $C_{a,b}$ is Hermitian. As for the second equality,

$$
\begin{aligned}
\sum_a \Big(\sum_b C_{a,b}\Big)\Big(\sum_b C_{a,b}\Big)^\dagger &= \sum_a \sum_{b,b'} C_{a,b} \cdot C_{a,b'} && (C_{a,b'} \text{ is Hermitian}) \\
&= \sum_a \sum_{b,b'} (Q_b \cdot P_a \cdot Q_b) \cdot (Q_{b'} \cdot P_a \cdot Q_{b'}) \\
&= \sum_a \sum_b Q_b \cdot P_a \cdot Q_b \cdot P_a \cdot Q_b && (\text{because } Q \text{ is projective}) \\
&\leq \sum_a \sum_b Q_b \cdot P_a \cdot Q_b && (\text{because } Q_b \leq I) \\
&\leq I,
\end{aligned}
$$

because P and Q are sub-measurements. $\qquad\square$

First, we show that

$$
\mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{g,h} \langle\psi| G_g^{\boldsymbol{x}} G_h^{\boldsymbol{y}} G_g^{\boldsymbol{x}} G_h^{\boldsymbol{y}} \otimes I |\psi\rangle \approx_{\sqrt{\zeta}} \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{g,h} \langle\psi| G_h^{\boldsymbol{y}} G_g^{\boldsymbol{x}} G_h^{\boldsymbol{y}} \otimes G_g^{\boldsymbol{x}} |\psi\rangle \tag{106}
$$

This follows from Proposition 4.24 where we let "A_a^x" and "B_a^x" in the Proposition be $G_g^x \otimes I$ and $I \otimes G_g^x$, respectively, and let "$C_{a,b}^x$" denote $G_h^y G_g^x G_h^y$. The closeness between "A_a^x" and "B_a^x" follows from Item 2, and the normalization condition on "$C_{a,b}^x$" follows from the projectivity of the $\{G_h^y\}$ measurements and Lemma 11.5. Next, we show that

$$
(106) = \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{g,h} \langle\psi| G_h^{\boldsymbol{y}} G_g^{\boldsymbol{x}} G_h^{\boldsymbol{y}} \otimes G_g^{\boldsymbol{x}} |\psi\rangle \approx_{\frac{dm}{q}} \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}} \sum_{a,h} \langle\psi| G_h^{\boldsymbol{y}} G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} G_h^{\boldsymbol{y}} \otimes G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} |\psi\rangle. \tag{107}
$$

To do so, we first compute the difference as

$$
\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}} \sum_{g \neq g', h} \mathbf{1}[g(\boldsymbol{u}) = g'(\boldsymbol{u})] \cdot \langle\psi| G_h^{\boldsymbol{y}} G_g^{\boldsymbol{x}} G_h^{\boldsymbol{y}} \otimes G_{g'}^{\boldsymbol{x}} |\psi\rangle. \tag{108}
$$

This is nonnegative and real, so it suffices to upper bound it, which we do as follows.

$$
\begin{aligned}
(108) &= \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{g \neq g', h} \langle\psi| G_h^{\boldsymbol{y}} G_g^{\boldsymbol{x}} G_h^{\boldsymbol{y}} \otimes G_{g'}^{\boldsymbol{x}} |\psi\rangle \cdot \mathop{\mathbf{E}}_{\boldsymbol{u}} \mathbf{1}[g(\boldsymbol{u}) = g'(\boldsymbol{u})] \\
&\leq \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{g \neq g', h} \langle\psi| G_h^{\boldsymbol{y}} G_g^{\boldsymbol{x}} G_h^{\boldsymbol{y}} \otimes G_{g'}^{\boldsymbol{x}} |\psi\rangle \cdot \frac{dm}{q} && (\text{by Schwartz-Zippel}) \\
&\leq \frac{dm}{q}. && (\text{because } G \text{ is a sub-measurement})
\end{aligned}
$$

90

Next, we show that

$$(107) = \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}} \sum_{a,h} \langle\psi|\, G_h^{\boldsymbol{y}} G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} G_h^{\boldsymbol{y}} \otimes G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}}\, |\psi\rangle \qquad \text{(Equation (107) rewriten)}$$

$$\approx_{\sqrt{\zeta}} \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}} \sum_{a,h} \langle\psi|\, G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} G_h^{\boldsymbol{y}} G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} G_h^{\boldsymbol{y}} \otimes I\, |\psi\rangle$$

$$\approx_{\sqrt{\zeta}} \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}} \sum_{a,h} \langle\psi|\, G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} G_h^{\boldsymbol{y}} G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} \otimes G_h^{\boldsymbol{y}}\, |\psi\rangle\,. \qquad (109)$$

These two approximations are derived as follows.

1. In the first approximation we used Proposition 4.24 where we let "A_a^x" and "B_a^x" in the Proposition be $G_{[g(u)=a]}^x \otimes I$ and $I \otimes G_{[g(u)=a]}^x$, respectively, and let "$C_{a,b}^x$" denote $G_h^y G_{[g(u)=a]}^x G_h^y$. The closeness between "A_a^x" and "B_a^x" follows from Item 2, and the normalization condition on "$C_{a,b}^x$" follows from the projectivity of the $\{G_h^y\}$ measurements and Lemma 11.5.

2. In the second approximation we used Proposition 4.24 where we let "A_a^x" and "B_a^x" in the Proposition be $G_h^y \otimes I$ and $I \otimes G_h^y$, respectively, and let "$C_{a,b}^x$" denote $G_{[g(u)=a]}^x G_h^y G_{[g(u)=a]}^x$. The closeness between "A_a^x" and "B_a^x" follows from Item 2, and the normalization condition on "$C_{a,b}^x$" follows from the projectivity of the $\{G_{[g(u)=a]}^x\}$ measurements and Lemma 11.5.

Analogously to the derivation of Equation (107), we now show that

$$(109) = \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}} \sum_{a,h} \langle\psi|\, G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} G_h^{\boldsymbol{y}} G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} \otimes G_h^{\boldsymbol{y}}\, |\psi\rangle \qquad \text{(Equation (109) restated)}$$

$$\approx_{\frac{dm}{q}} \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}} \sum_{a,b} \langle\psi|\, G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} G_{[h(\boldsymbol{v})=b]}^{\boldsymbol{y}} G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} \otimes G_{[h(\boldsymbol{v})=b]}^{\boldsymbol{y}}\, |\psi\rangle\,. \qquad (110)$$

To do so, we first compute the difference as

$$\mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}} \sum_{a,h\neq h'} \mathbf{1}[h(\boldsymbol{v}) = h'(\boldsymbol{v})] \cdot \langle\psi|\, G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} G_h^{\boldsymbol{y}} G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} \otimes G_{h'}^{\boldsymbol{y}}\, |\psi\rangle\,. \qquad (111)$$

This is nonnegative and real, so it suffices to upper bound it, which we do as follows.

$$(111) = \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}} \sum_{a,h\neq h'} \langle\psi|\, G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} G_h^{\boldsymbol{y}} G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} \otimes G_{h'}^{\boldsymbol{y}}\, |\psi\rangle \cdot \mathop{\mathbf{E}}_{\boldsymbol{v}} \mathbf{1}[h(\boldsymbol{v}) = h'(\boldsymbol{v})]$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}} \sum_{a,h\neq h'} \langle\psi|\, G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} G_h^{\boldsymbol{y}} G_{[g(\boldsymbol{u})=a]}^{\boldsymbol{x}} \otimes G_{h'}^{\boldsymbol{y}}\, |\psi\rangle \cdot \frac{dm}{q} \qquad \text{(by Schwartz-Zippel)}$$

$$\leq \frac{dm}{q}\,. \qquad \text{(because } G \text{ is a sub-measurement)}$$

Now, we set

$$\nu_{\text{evaluation}} = 48m \cdot (\gamma^{1/2} + \zeta^{1/2})$$

to be the approximation error from Lemma 11.1. We conclude by showing that

$$(110) = \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}} \sum_{a,b} \langle\psi|\, G^{\boldsymbol{x}}_{[g(\boldsymbol{u})=a]} G^{\boldsymbol{y}}_{[h(\boldsymbol{v})=b]} G^{\boldsymbol{x}}_{[g(\boldsymbol{u})=a]} \otimes G^{\boldsymbol{y}}_{[h(\boldsymbol{v})=b]}\, |\psi\rangle \qquad \text{(Equation (110) restated)}$$

$$\approx_{\sqrt{\nu_{\text{evaluation}}}} \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}} \sum_{a,b} \langle\psi|\, G^{\boldsymbol{x}}_{[g(\boldsymbol{u})=a]} G^{\boldsymbol{x}}_{[g(\boldsymbol{u})=a]} G^{\boldsymbol{y}}_{[h(\boldsymbol{v})=b]} \otimes G^{\boldsymbol{y}}_{[h(\boldsymbol{v})=b]}\, |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}} \sum_{a,b} \langle\psi|\, G^{\boldsymbol{x}}_{[g(\boldsymbol{u})=a]} G^{\boldsymbol{y}}_{[h(\boldsymbol{v})=b]} \otimes G^{\boldsymbol{y}}_{[h(\boldsymbol{v})=b]}\, |\psi\rangle$$

$$\approx_{\sqrt{\zeta}} \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}} \sum_{a,b} \langle\psi|\, G^{\boldsymbol{x}}_{[g(\boldsymbol{u})=a]} \otimes G^{\boldsymbol{y}}_{[h(\boldsymbol{v})=b]} G^{\boldsymbol{y}}_{[h(\boldsymbol{v})=b]}\, |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}} \sum_{a,b} \langle\psi|\, G^{\boldsymbol{x}}_{[g(\boldsymbol{u})=a]} \otimes G^{\boldsymbol{y}}_{[h(\boldsymbol{v})=b]}\, |\psi\rangle$$

$$= \langle\psi|\, G \otimes G\, |\psi\rangle .$$

The third and fifth lines follow from the projectivity of the G measurements. The two approximations are derived as follows.

1. In the first approximation we used Proposition 4.24 where we let "A_a^x" and "B_a^x" in the Proposition be $G^x_{[g(u)=a]} G^y_{[h(v)=b]} \otimes I$ and $G^y_{[h(v)=b]} G^x_{[g(u)=a]} \otimes I$, respectively, and let "$C_{a,b}^x$" denote $G^x_{[g(u)=a]} \otimes G^y_{[h(v)=b]}$. The closeness between "A_a^x" and "B_a^x" follows from Lemma 11.1, and the normalization condition on "$C_{a,b}^x$" follows from it being a sub-measurement.

2. In the second approximation we used Proposition 4.24 where we let "A_a^x" and "B_a^x" in the Proposition be $G^y_{[h(v)=b]} \otimes I$ and $I \otimes G^y_{[h(v)=b]}$, respectively, and let "$C_{a,b}^x$" denote $G^x_{[g(u)=a]} \otimes G^y_{[h(v)=b]}$. The closeness between "A_a^x" and "B_a^x" follows from Item 2, and the normalization condition on "$C_{a,b}^x$" follows from it being a sub-measurement.

This shows that the second term in Equation (105) is approximately $\langle\psi|\, G \otimes G\, |\psi\rangle$, as desired. In total, we have incurred an error of

$$2 \cdot \left(2\sqrt{\zeta} + \sqrt{\zeta} + \frac{dm}{q} + \sqrt{\zeta} + \sqrt{\zeta} + \frac{dm}{q} + \sqrt{\nu_{\text{evaluation}}} + \sqrt{\zeta} \right)$$

$$= 12\zeta^{1/2} + 2m \cdot (d/q) + 2 \cdot \sqrt{48m \cdot (\gamma^{1/2} + \zeta^{1/2})}$$

$$\leq 12\zeta^{1/2} + 2m \cdot (d/q) + 14m \cdot \left(\gamma^{1/4} + \zeta^{1/4} \right)$$

$$\leq 12\zeta^{1/4} + 2m \cdot (d/q)^{1/4} + 14m \cdot \left(\gamma^{1/4} + \zeta^{1/4} \right)$$

$$\leq 30m \cdot \left(\gamma^{1/4} + \zeta^{1/4} + (d/q)^{1/4} \right) .$$

This concludes the proof of Theorem 11.4. □

12 Pasting

Theorem 12.1 (Pasting). *Let (ψ, A, B, L) be an $(\epsilon, \delta, \gamma)$-good symmetric strategy for the $(m + 1, q, d)$ low individual degree test. Let $\{G^x\}_{x \in \mathbb{F}_q}$ denote a set of projective sub-measurements in PolySub(m, q, d) with the following properties:*

1. *(Completeness):* If $G = \mathbf{E_x} \sum_g G_g^{\boldsymbol{x}}$, then

$$\langle \psi | \, G \otimes I \, | \psi \rangle \geq 1 - \kappa.$$

2. *(Consistency with A):* On average over $(\boldsymbol{u}, \boldsymbol{x}) \sim \mathbb{F}_q^{m+1}$,

$$A_a^{u,x} \otimes I \simeq_\zeta I \otimes G_{[g(u)=a]}^x.$$

3. *(Strong self-consistency):* On average over $\boldsymbol{x} \sim \mathbb{F}_q$,

$$G_g^x \otimes I \approx_\zeta I \otimes G_g^x.$$

4. *(Boundedness):* There exists a positive-semidefinite matrix Z^x for each $x \in \mathbb{F}_q$ such that

$$\mathbf{E}_{\boldsymbol{x}} \langle \psi | \, (I - G^{\boldsymbol{x}}) \otimes Z^{\boldsymbol{x}} \, | \psi \rangle \leq \zeta$$

and for each $x \in \mathbb{F}_q$ and $g \in \mathcal{P}(m, q, d)$,

$$Z^x \geq \left(\mathbf{E}_{\boldsymbol{u}} A_{g(\boldsymbol{u})}^{\boldsymbol{u},x} \right).$$

Let $k \geq 400md$ be an integer. Let

$$\nu = 100 k^2 m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32} \right),$$

$$\sigma = \kappa \cdot \left(1 + \frac{1}{100m} \right) + 2\nu + e^{-k/(80000m^2)}.$$

Then there exists a "pasted" measurement $H \in \mathrm{PolyMeas}(m+1, q, d)$ which satisfies the following property.

1. *(Consistency with A):* On average over $\boldsymbol{u} \sim \mathbb{F}_q^{m+1}$,

$$A_a^u \otimes I \simeq_\sigma I \otimes H_{[h(u)=a]}.$$

We note that the bound we are proving is trivial when at least one of ϵ, δ, γ, ζ, or d/q is ≥ 1, as ν is at least 1 in that case. Hence, we may assume that $\epsilon, \delta, \gamma, \zeta, d/q \leq 1$. This will aid us when carrying out the error calculations, as it allows us to bound terms like $\zeta^{1/2}$ by terms like $\zeta^{1/4}$.

It will be convenient to state certain consistency relations that follow easily from the hypotheses of the theorem. In this section, we will only use the axis-parallel lines test in the $(m+1)$-st direction, i.e. in the case when the line is of the form $\ell = \{(u, x) \mid x \in \mathbb{F}_q\}$. Such a line is specified by a point $u \in \mathbb{F}_q^m$. As a result, we can use the shorthand introduced in Notation 3.8, where instead of writing B_f^ℓ for such a line, we write B_f^u. We will also make use of the other shorthand from Notation 3.8 in which for a function $f : \ell \to \mathbb{F}_q$, we will sometimes write $f(x)$ instead of $f(u, x)$.

We know that (ψ, A, B, L) passes the axis-parallel lines test with probability $1 - \epsilon$, and so it passes this test with probability $1 - (m+1) \cdot \epsilon$ conditioned on the random direction $\boldsymbol{i} \in \{1, \dots, m+1\}$ from the test being equal to $m+1$. This means the following: if $(\boldsymbol{u}, \boldsymbol{x}) \sim \mathbb{F}_q^{m+1}$ is drawn uniformly at random and $\boldsymbol{\ell} = \{(\boldsymbol{u}, y) \mid y \in \mathbb{F}_q\}$ is the line through it in the $(m+1)$-st direction, then

$$A_a^{u,x} \otimes I \simeq_{(m+1)\epsilon} I \otimes B_{[f(x)=a]}^\ell = I \otimes B_{[f(x)=a]}^u.$$

93

Hence, Proposition 4.22 and the fact that $(m + 1) \leq 2m$ imply that

$$A_a^{u,x} \otimes I \approx_{4m\epsilon} I \otimes B_{[f(x)=a]}^u.$$

Next, from the assumption that the given strategy is $(\epsilon, \delta, \gamma)$-good, Proposition 4.22 implies that

$$I \otimes A_a^{u,x} \approx_{2\delta} A_a^{u,x} \otimes I \approx_{4m\epsilon} I \otimes B_{[f(x)=a]}^u.$$

As a result, Proposition 4.28 implies that

$$I \otimes A_a^{u,x} \approx_{8m\epsilon+4\delta} I \otimes B_{[f(x)=a]}^u. \tag{112}$$

Proposition 4.21 applied to Item 2 and Equation (112) implies

$$G_{[g(u)=a]}^x \otimes I \simeq_{\nu_1} I \otimes B_{[f(x)=a]}^u, \tag{113}$$

where

$$\nu_1 = \zeta + \sqrt{8m\epsilon + 4\delta}. \tag{114}$$

Finally, Theorem 11.4 implies that

$$G_g^x G_h^y \otimes I \approx_{\nu_{\text{commute}}} G_h^y G_g^x \otimes I, \tag{115}$$

where

$$\nu_{\text{commute}} = 30m \cdot \left(\gamma^{1/4} + \zeta^{1/4} + (d/q)^{1/4} \right).$$

12.1 From measurements to sub-measurements

Rather than designing the pasted measurement guaranteed by Theorem 12.1, it is convenient to first design a pasted *sub*-measurement, and then convert it to a measurement. The next lemma shows the bounds we achieve for this sub-measurement.

Lemma 12.2. *There exists a "pasted" sub-measurement $H \in \mathrm{PolySub}(m + 1, q, d)$ which satisfies the following properties.*

1. *(Consistency with A):   On average over $\boldsymbol{u} \sim \mathbb{F}_q^{m+1}$,*

   $$A_a^u \otimes I \simeq_\nu I \otimes H_{[h(u)=a]}.$$

2. *(Completeness):   If $H = \sum_h H_h$, then*

   $$\langle \psi | H \otimes I | \psi \rangle \geq 1 - \kappa \cdot \left(1 + \frac{1}{100m} \right) - \nu - e^{-k/(80000m^2)}.$$

Proof of Theorem 12.1 assuming Lemma 12.2. Let H be the sub-measurement in $\mathrm{PolySub}(m + 1, q, d)$ guaranteed by Lemma 12.2. Let h^* be an arbitrary polynomial in $\mathcal{P}(m + 1, q, d)$. We define the measurement $H_{\text{meas}} \in \mathrm{PolyMeas}(m + 1, q, d)$ as follows.

$$(H_{\text{meas}})_h = \begin{cases} H_{h^*} + (I - H) & \text{if } h = h^*, \\ H_h & \text{otherwise.} \end{cases}$$

This is clearly a measurement, as the sum of its POVM elements is $H + (I - H) = H$. In addition,

$$\mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_{a \neq b} \langle \psi | A_a^{\boldsymbol{u}} \otimes (H_{\mathrm{meas}})_{[h(\boldsymbol{u})=b]} | \psi \rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_a \sum_{h:h(\boldsymbol{u})\neq a} \langle \psi | A_a^{\boldsymbol{u}} \otimes (H_{\mathrm{meas}})_h | \psi \rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_a \sum_{h:h(\boldsymbol{u})\neq a} \langle \psi | A_a^{\boldsymbol{u}} \otimes H_h | \psi \rangle + \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_{a:h^*(\boldsymbol{u})\neq a} \langle \psi | A_a^{\boldsymbol{u}} \otimes (I - H) | \psi \rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{u}} \sum_a \sum_{h:h(\boldsymbol{u})\neq a} \langle \psi | A_a^{\boldsymbol{u}} \otimes H_h | \psi \rangle + \mathop{\mathbf{E}}_{\boldsymbol{u}} \langle \psi | I \otimes (I - H) | \psi \rangle$$

$$\leq (\nu) + \left(\kappa \cdot \left(1 + \frac{1}{100m} \right) + \nu + e^{-k/(80000m^2)} \right) \qquad \text{(by Items 1 and 2)}$$

$$= \sigma.$$

Thus, $A_a^u \otimes I \simeq_\sigma I \otimes (H_{\mathrm{meas}})_{[h(u)=a]}$. This completes the proof. $\qquad \square$

We will now spend the rest of this section proving Lemma 12.2, i.e. designing the pasted sub-measurement H.

12.2   The pasted sub-measurement

Definition 12.3. For $k \geq 1$, we let $\mathsf{Distinct}_k$ be the set of tuples $(x_1, \ldots, x_k) \in \mathbb{F}_q^k$ such that $x_i \neq x_j$ for all $i \neq j$. We write $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) \sim \mathsf{Distinct}_k$ for a uniformly random element of this set.

In manipulations involving our pasted measurement, it will often be useful to switch the expectation over $\mathsf{Distinct}_k$ with an expectation over uniformly random k-tuples of elements of $\mathbb{F}_q$. The following proposition bounds the distance between these two distributions.

Proposition 12.4. Let $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) \sim \mathbb{F}_q^k$ be sampled uniformly at random and let $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k) \sim \mathsf{Distinct}_k$. Then

$$d_{\mathrm{TV}}(\boldsymbol{x}, \boldsymbol{y}) \leq \frac{k^2}{q}.$$

Proof. For any $z = (z_1, \ldots, z_k) \in \mathbb{F}_q^k$,

$$\mathbf{Pr}[\boldsymbol{x} = z] = \frac{1}{q^k},$$

$$\mathbf{Pr}[\boldsymbol{y} = z] = \begin{cases} \frac{1}{\binom{q}{k}k!} & \text{if } z \in \mathsf{Distinct}_k, \\ 0 & \text{otherwise.} \end{cases}$$

Because $\frac{1}{\binom{q}{k}k!} \geq \frac{1}{q^k}$, $\mathbf{Pr}[\boldsymbol{x} = z] \geq \mathbf{Pr}[\boldsymbol{y} = z]$ if and only if $z \notin \mathsf{Distinct}_k$. Hence,

$$d_{\mathrm{TV}}(\boldsymbol{x}, \boldsymbol{y}) = \max_{S \subseteq \mathbb{F}_q^k} \{ \mathbf{Pr}[\boldsymbol{x} \in S] - \mathbf{Pr}[\boldsymbol{y} \in S] \}$$

$$= \mathbf{Pr}[\boldsymbol{x} \in \overline{\mathsf{Distinct}_k}] - \mathbf{Pr}[\boldsymbol{y} \in \overline{\mathsf{Distinct}_k}] = \mathbf{Pr}[\boldsymbol{x} \in \overline{\mathsf{Distinct}_k}].$$

We can upper-bound this probability as follows.

$$\mathbf{Pr}[\boldsymbol{x} \in \overline{\mathsf{Distinct}_k}] = \mathbf{Pr}[\exists i : \boldsymbol{x}_i \in \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1}\}]$$

$$\leq \sum_{i=2}^{k} \mathbf{Pr}[\boldsymbol{x}_i \in \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1}\}] \leq \sum_{i=2}^{k} \left(\frac{i-1}{q} \right) = \frac{k(k-1)}{2q}.$$

This concludes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

To begin, we will describe our construction of the global pasted measurement. In fact, we will consider *two* separate constructions of the global pasted measurement. The first is given in Section 12.2.1 below. It is the more natural of the two, but unfortunately we do not know how to prove that it works correctly. This motivates our second construction, given in Section 12.2.2 below, which is designed to circumvent the problems in the first construction.

12.2.1 The first construction

The first construction of $H = \{H_h\}$ is conceptually simple: we perform the G sub-measurement $d+1$ times to produce $d+1$ polynomials $g_1, \ldots, g_{d+1} \in \mathcal{P}(m, q, d)$. We then perform polynomial interpolation to produce a single global polynomial $h \in \mathcal{P}(m+1, q, d)$. In more detail, this construction involves three steps.

1. (Pasting): Let $x_1, \ldots, x_{d+1} \in \mathbb{F}_q$. We will define an initial "sandwiched" measurement as follows:

$$\widehat{H}_{g_1,\ldots,g_{d+1}}^{x_1,\ldots,x_{d+1}} = G_{g_1}^{x_1} \cdot G_{g_2}^{x_2} \cdots G_{g_{d+1}}^{x_{d+1}} \cdots G_{g_2}^{x_2} \cdot G_{g_1}^{x_1}.$$

$\widehat{H}^{x_1,\ldots,x_{d+1}}$ has a natural interpretation as the sub-measurement in which one performs the sub-measurements $G^{x_1}, G^{x_2}, \ldots, G^{x_{d+1}}$ one after another and outputs their results.

2. (Interpolation) Next, let $(x_1, \ldots, x_{d+1}) \in \mathsf{Distinct}_{d+1}$. We define the interpolated measurement

$$H_h^{x_1,\ldots,x_{d+1}} = \widehat{H}_{h|_{x_1},\ldots,h|_{x_{d+1}}}^{x_1,\ldots,x_{d+1}}.$$

This performs the $\widehat{H}^{x_1,\ldots,x_{d+1}}$ measurement and outputs the h which is consistent with the outcomes $g_1, \ldots, g_{d+1}$ if one exists. (We note that this is not necessarily a sub-measurement if $(x_1, \ldots, x_{d+1}) \notin \mathsf{Distinct}_{d+1}$. This is because there may exist $h \neq h'$ for which $h|_{x_i} = (h')|_{x_i}$ for all $1 \leq i \leq d+1$.)

3. (Averaging): Finally, we randomize over the choice of $(x_1, \ldots, x_{d+1})$. In other words, we define

$$H_h = \mathop{\mathbf{E}}_{(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{d+1}) \sim \mathsf{Distinct}_{d+1}} H_h^{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{d+1}}.$$

To show that this construction works, we need to show two things: (i) that H_h has good agreement with A, and (ii) that H's completeness is close to G's. Showing (i) is simple and follows from the approximate commutativity of G^x and G^y established in Theorem 11.4. What we do not know how to do is to show (ii), at least for general d. In fact, at first glance it even looks like it should be false! To see why, we note that it is possible to show that the completeness of H can be approximated as

$$\langle \psi | H \otimes I | \psi \rangle \approx \langle \psi | G^{d+1} \otimes I | \psi \rangle ,$$

because H is performing the G measurement $d + 1$ times. We would therefore like to show that

$$\langle \psi | \, G^{d+1} \otimes I \, | \psi \rangle \approx \langle \psi | \, G \otimes I \, | \psi \rangle . \tag{116}$$

But if $\langle \psi | \, G \otimes I \, | \psi \rangle = 1 - \kappa$, a "naive analysis" would lead to the conclusion that

$$\langle \psi | \, G^{d+1} \otimes I \, | \psi \rangle \approx 1 - (d + 1) \cdot \kappa, \tag{117}$$

which would be too great of a loss in completeness for our proof strategy to work.

However, we can show Equation (116) is actually correct and therefore this "naive analysis" is incorrect, at least in the case of constant d. To see how this is possible, note that Equation (117) is in fact correct when all of G's eigenvalues are equal to $1 - \kappa$, in which case $G^{d+1} = (1 - \kappa)^d \cdot G$. So we would like to show that, on the contrary, the fact that G's incompleteness is equal to $1 - \kappa$ is because a "$(1 - \kappa)$ fraction" of its eigenvalues are equal to 1, and the remaining "κ fraction" of its eigenvalues are equal to 0. In this case, $G^{d+1} = G$, and so Equation (116) holds.

We now sketch the argument that shows this holds, at least for constant d. To do this, it is first possible to show that

$$\langle \psi | \, G^{d+2} \otimes I \, | \psi \rangle \approx_\Delta \langle \psi | \, G^{d+1} \otimes I \, | \psi \rangle , \tag{118}$$

where $\Delta = \text{poly}(m) \cdot \text{poly}(\epsilon, \delta, \gamma, \zeta, d/q)$ is small. Intuitively, this is because after performing $(d + 1)$ G measurements, the outcome of another G measurement is essentially determined due to interpolation. (The proof is of this fact is slightly subtle and involves the Z-boundedness condition.)

From here, it is possible to derive Equation (116) as follows. Write the eigendecomposition $G = \sum_i \lambda_i | v_i \rangle \langle v_i |$ of G. This defines a probability distribution μ over eigenvectors, where eigenvector i occurs with probability $\mu(i) = \langle \psi | \, (| v_i \rangle \langle v_i | \otimes I) \, | \psi \rangle$. In this language, for any power k we can write

$$\langle \psi | \, G^k \otimes I \, | \psi \rangle = \langle \psi | \left(\sum_i \lambda_i^k | v_i \rangle \langle v_i | \right) \otimes I \, | \psi \rangle = \sum_i \lambda_i^k \cdot \mu(i) = \mathop{\mathbf{E}}_{i \sim \mu} [\lambda_i^k].$$

Thus, Equation (118) is equivalent to the statement that

$$\Delta \geq \langle \psi | \, G^{d+1} \otimes I \, | \psi \rangle - \langle \psi | \, G^{d+2} \otimes I \, | \psi \rangle = \mathop{\mathbf{E}}_{i \sim \mu} [\lambda_i^{d+1}] - \mathop{\mathbf{E}}_{i \sim \mu} [\lambda_i^{d+2}] = \mathop{\mathbf{E}}_{i \sim \mu} [\lambda_i^{d+1} (1 - \lambda_i)]. \tag{119}$$

By Equation (116), our goal is to bound

$$\langle \psi | \, G \otimes I \, | \psi \rangle - \langle \psi | \, G^{d+1} \otimes I \, | \psi \rangle = \mathop{\mathbf{E}}_{i \sim \mu} [\lambda_i] - \mathop{\mathbf{E}}_{i \sim \mu} [\lambda_i^{d+1}] = \mathop{\mathbf{E}}_{i \sim \mu} [\lambda_i (1 - \lambda_i^d)].$$

To do so, we will use the following lemma.

Lemma 12.5. *For any real number $0 \leq \lambda \leq 1$,*

$$\lambda (1 - \lambda^d) \leq 2 \cdot \left(\lambda^{d+1} (1 - \lambda) \right)^{1/(d+1)}.$$

Proof. The lemma is trivial for $\lambda = 1$, and so we will assume that $\lambda \neq 1$. We note that

$$\lambda^{d+1} (1 - \lambda^d)^{d+1} \leq \lambda^{d+1} (1 - \lambda^d) = \lambda^{d+1} (1 - \lambda) \cdot \left(\frac{1 - \lambda^d}{1 - \lambda} \right)$$

$$= \lambda^{d+1} (1 - \lambda) \cdot (1 + \lambda + \cdots + \lambda^{d-1}) \leq d \cdot \lambda^{d+1} (1 - \lambda).$$

The lemma now follows by taking the $(d + 1)$-st root of both sides and noting that $d^{1/(d+1)} \leq 2$ for all integers $d \geq 1$. $\qquad \square$

Then Lemma 12.5 implies that

$$\mathop{\mathbf{E}}_{i\sim\mu}[\lambda_i(1-\lambda_i^d)] \le 2 \cdot \mathop{\mathbf{E}}_{i\sim\mu}\left[\left(\lambda_i^{d+1}(1-\lambda_i)\right)^{1/(d+1)}\right]$$

$$\le 2 \cdot \left(\mathop{\mathbf{E}}_{i\sim\mu}[\lambda_i^{d+1}(1-\lambda_i)]\right)^{1/(d+1)} \qquad \text{(because } a \mapsto a^{1/(d+1)} \text{ is concave)}$$

$$\le 2 \cdot \Delta^{1/(d+1)}. \qquad\qquad\qquad\qquad\qquad \text{(by Equation (119))}$$

Thus, we have established the bound

$$\langle\psi| \, G \otimes I \, |\psi\rangle - \langle\psi| \, G^{d+1} \otimes I \, |\psi\rangle \le 2 \cdot \Delta^{1/d+1} = 2 \cdot (\mathrm{poly}(m) \cdot \mathrm{poly}(\epsilon, \delta, \gamma, \zeta, d/q))^{1/d+1}.$$

For constant d, this bound is suitable for our proof, as the right-hand side is still a polynomial in the relevant parameters. However, when d is larger, for this bound to be meaningful, we need ϵ, δ, etc. to be exponentially small in d, which is a more stringent condition than we generally allow (unless, again, d is a constant). That said, we believe this large error may be an artifact of the proof strategy rather than something intrinsic to the construction. We leave this to future work.

12.2.2 The second construction

The second construction of $H = \{H_h\}$ is designed to circumvent the problem of the first construction, which is that its completeness was difficult to analyze. Instead, we will design a pasted measurement in which the "naive analysis" actually gets us the bound we want, which is that H's completeness is close to G's completeness. Before describing the construction, we need the following definitions.

Definition 12.6 (G's incomplete part)**.** For each $x \in \mathbb{F}_q$, we write $G^x = \sum_g G_g^x$ and $G_\perp^x = I - G^x$ for the "complete" and "incomplete" parts of G^x, respectively.

It will be convenient to sometimes regard G^x as a complete measurement by throwing in the additional measurement outcome "$\perp$". To distinguish this from G^x as a sub-measurement, we will use the notation "$\widehat{G}^x$". In other words, we let $\widehat{G} = \{\widehat{G}_g^x\}$ be the projective measurement defined as

$$\widehat{G}_g^x = \begin{cases} G_g^x & \text{if } g \in \mathcal{P}(m,q,d), \\ G_\perp^x & \text{if } g = \perp. \end{cases}$$

This measurement has outcomes ranging over the set $\mathcal{P}^+(m,q,d) := \mathcal{P}(m,q,d) \cup \{\perp\}$.

Definition 12.7 (Types)**.** A type τ is an element of $\{0,1\}^k$ for some integer k. We write $|\tau| = \tau_1 + \cdots + \tau_k$ for the Hamming weight of τ. We will also associate τ with the set $\{i \mid \tau_i = 1\}$ and write $i \in \tau$ if $\tau_i = 1$.

Suppose we perform the $\widehat{G}$ measurement k times in succession, generating the random outcomes $g_1, \ldots, g_k$. Let us write $\tau \in \{0,1\}^k$ for the "type" of these outcomes, where

$$\tau_i = \begin{cases} 1 & \text{if } g_i \in \mathcal{P}(m,q,d), \\ 0 & \text{if } g_i = \perp. \end{cases}$$

Assuming the g_i's are not inconsistent, then we can interpolate them to produce a global polynomial h whenever $|\tau| \ge d+1$. Hence, we would like to understand the probability that $|\tau| \ge d+1$ and ensure that it is as large as possible. The probability that the measurement $\widehat{G}$ returns a polynomial $g \in \mathcal{P}(m,q,d)$ is equal to the completeness of G, which is $1 - \kappa$. This tells us that the probability

that $\boldsymbol{\tau}_1 = 1$ is $1 - \kappa$. We might naively expect that the same holds for the other $\boldsymbol{\tau}_i$'s as well. We might also naively expect that the $\boldsymbol{\tau}_i$'s are independent. These two assumptions should not be expected to hold in general, as they ignore correlations between the measurements and the fact that each measurement perturbs the state $|\psi\rangle$ for subsequent measurements to use. However, if we make these assumptions, then we at least have a simple toy model for the measurement outcomes: $\boldsymbol{\tau} \sim \mathsf{Binomial}(k, 1 - \kappa)$.

In this toy model, we expect $|\boldsymbol{\tau}| \approx k \cdot (1 - \kappa)$ on average. This was the problem with the "naive analysis" from the first construction: if $k = d + 1$ and κ is reasonably large (say, on the order of $1/d$), then we don't expect $|\boldsymbol{\tau}|$ to be $\geq d + 1$ with high probability, and so we can't interpolate to produce a global polynomial. This suggests an alternative strategy: simply choose k large enough so that $k \cdot (1 - \kappa) \gg d + 1$. In fact, as we are aiming for H to have completeness close to $1 - \kappa$, we should choose k so large that $|\boldsymbol{\tau}| \geq d + 1$ with probability roughly $1 - \kappa$. This is easily done with a Chernoff bound, which is responsible for the exponential error term in Item 2 of Lemma 12.2. On the other hand, if we set k *too* large, then we increase the risk that our k outcomes $g_1, \ldots, g_k$ are inconsistent with each other, which is an additional source of error. This is responsible for the tradeoff between "large" and "small" k discussed in Section 6.1 above.

Although, this "naive analysis" only holds in this toy model, it still motivates our second construction of H, which we state below. We will show that the naive analysis, in which we treat $\boldsymbol{\tau}$ as a binomial random variable and bound $|\boldsymbol{\tau}|$ using a Chernoff bound, can actually be made formal.

Definition 12.8 (The pasted measurement). Let $k \geq d + 1$ be an integer.

1. (Pasting): Let $x_1, \ldots, x_k \in \mathbb{F}_q$. We will define an initial "sandwiched" measurement as follows:
$$\widehat{H}^{x_1, \ldots, x_k}_{g_1, \ldots, g_k} = \widehat{G}^{x_1}_{g_1} \cdot \widehat{G}^{x_2}_{g_2} \cdots \widehat{G}^{x_k}_{g_k} \cdots \widehat{G}^{x_2}_{g_2} \cdot \widehat{G}^{x_1}_{g_1}.$$

2. (Interpolation): Next, let $(x_1, \ldots, x_k) \in \mathsf{Distinct}_k$. For any string $w \in \{0, 1\}^k$ and polynomial $h \in \mathcal{P}(m + 1, q, d)$, we define h_w to be the tuple $(g_1, \ldots, g_k) \in \mathcal{P}^+(m, q, d)^k$ where $g_i = \perp$ if $w_i = 0$ and $g_i = h|_{x_i}$ otherwise. We define the interpolated measurement
$$H^{x_1, \ldots, x_k}_h = \sum_{w : |w| \geq d+1} \widehat{H}^{x_1, \ldots, x_k}_{h_w}.$$

3. (Averaging): Finally, we randomize over the choice of $(x_1, \ldots, x_k)$. In other words, we define
$$H_h = \mathop{\mathbf{E}}_{(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) \sim \mathsf{Distinct}_k} H^{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k}_h.$$

To analyze the second construction, we first need to show that the $\widehat{G}$ measurement satisfies some basic properties, like commutation with itself. These are shown in Section 12.3, where they follow from the fact that similar properties hold for G. Using this, we prove that $\widehat{H}$ is consistent with B in Section 12.4, which we use to prove that H is consistent with A in Section 12.5. Finally, we analyze the completeness of H in Section 12.6.

12.3 Strong self-consistency and commutation of $\widehat{G}$

In this section, we show that $\widehat{G}$ is strongly self-consistent and commutes with itself. As we already know this holds for the sub-measurement G, our task essentially reduces to showing that these properties also hold for G's incomplete part, i.e. $G_\perp$. As it is more convenient to work with $G = I - G_\perp$ rather than $G_\perp$, we will first show that these properties hold for G; the fact that they also hold for $G_\perp$ will then follow as an immediate corollary.

12.3.1 Strong self-consistency of $G_\perp$

Lemma 12.9 (Strong self-consistency of G's complete part)**.**

$$G^x \otimes I \approx_\zeta I \otimes G^x.$$

Proof. Because G is a projective measurement, Proposition 4.36 implies that the strong self-consistency of G from Item 3 is equivalent to

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_g \langle \psi | G_g^{\boldsymbol{x}} \otimes G_g^{\boldsymbol{x}} | \psi \rangle \geq \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_g \langle \psi | G_g^{\boldsymbol{x}} \otimes I | \psi \rangle - \frac{1}{2} \cdot \zeta. \tag{120}$$

Our goal is to bound

$$
\begin{aligned}
&\mathop{\mathbf{E}}_{\boldsymbol{x}} \| (G^{\boldsymbol{x}} \otimes I - I \otimes G^{\boldsymbol{x}}) | \psi \rangle \|^2 \\
=~& 2 \cdot \mathop{\mathbf{E}}_{\boldsymbol{x}} \langle \psi | G^{\boldsymbol{x}} \otimes I | \psi \rangle - 2 \cdot \mathop{\mathbf{E}}_{\boldsymbol{x}} \langle \psi | G^{\boldsymbol{x}} \otimes G^{\boldsymbol{x}} | \psi \rangle \\
=~& 2 \cdot \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{g_1} \langle \psi | G_g^{\boldsymbol{x}} \otimes I | \psi \rangle - 2 \cdot \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_{g,h} \langle \psi | G_g^{\boldsymbol{x}} \otimes G_h^{\boldsymbol{x}} | \psi \rangle \\
\leq~& 2 \cdot \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_g \langle \psi | G_g^{\boldsymbol{x}} \otimes I | \psi \rangle - 2 \cdot \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_g \langle \psi | G_g^{\boldsymbol{x}} \otimes G_g^{\boldsymbol{x}} | \psi \rangle .
\end{aligned}
$$

But this is at most ζ by Equation (120). $\qquad\square$

Corollary 12.10 (Strong self-consistency of G's incomplete part)**.**

$$G_\perp^x \otimes I \approx_\zeta I \otimes G_\perp^x.$$

Proof. For any x,

$$G_\perp^x \otimes I - I \otimes G_\perp^x = (I - G^x) \otimes I - I \otimes (I - G^x) = I \otimes G^x - G^x \otimes I.$$

Thus,

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \| (G_\perp^{\boldsymbol{x}} \otimes I - I \otimes G_\perp^{\boldsymbol{x}}) | \psi \rangle \|^2 = \mathop{\mathbf{E}}_{\boldsymbol{x}} \| (I \otimes G^{\boldsymbol{x}} - G^{\boldsymbol{x}} \otimes I) | \psi \rangle \|^2,$$

which is at most ζ by Lemma 12.9. $\qquad\square$

12.3.2 Commutativity of $G_\perp$

Lemma 12.11 (Commutativity with G_g^x implies commutativity with G^x)**.** *Let $M = \{M_o^x\}$ be a projective sub-measurement with outcomes in some set $\mathcal{O}$. Suppose that*

$$M_o^x \otimes I \approx_\omega I \otimes M_o^x, \tag{121}$$

and

$$G_g^x M_o^y \otimes I \approx_\chi M_o^y G_g^x \otimes I \tag{122}$$

over independent and uniformly random $\boldsymbol{x}, \boldsymbol{y} \sim \mathbb{F}_q$. Then

$$G^x M_o^y \otimes I \approx_{6\sqrt{\zeta} + 6\sqrt{\omega} + 4\sqrt{\chi}} M_o^y G^x \otimes I.$$

Proof. The error we wish to bound is

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_o \| (G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} - M_o^{\boldsymbol{y}} G^{\boldsymbol{x}}) \otimes I \, |\psi\rangle \|^2$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_o \langle\psi| \, M_o^{\boldsymbol{y}} (G^{\boldsymbol{x}})^2 M_o^{\boldsymbol{y}} \otimes I \, |\psi\rangle + \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_o \langle\psi| \, G^{\boldsymbol{x}} (M_o^{\boldsymbol{y}})^2 G^{\boldsymbol{x}} \otimes I \, |\psi\rangle$$

$$- \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_o \langle\psi| \, G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} \otimes I \, |\psi\rangle - \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_g \langle\psi| \, M_o^{\boldsymbol{y}} G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G^{\boldsymbol{x}} \otimes I \, |\psi\rangle. \qquad (123)$$

We will show that all four terms in Equation (123) are close to $\langle\psi| \, G \otimes M \, |\psi\rangle$, where $M = \mathbf{E}_{\boldsymbol{y}} \sum_o M_o^{\boldsymbol{y}}$.

For the first term in Equation (123), we have

$$\mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_o \langle\psi| \, M_o^{\boldsymbol{y}} (G^{\boldsymbol{x}})^2 M_o^{\boldsymbol{y}} \otimes I \, |\psi\rangle = \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_o \langle\psi| \, M_o^{\boldsymbol{y}} G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} \otimes I \, |\psi\rangle \qquad \text{(because } G \text{ is projective)}$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{y}} \sum_o \langle\psi| \, M_o^{\boldsymbol{y}} G M_o^{\boldsymbol{y}} \otimes I \, |\psi\rangle$$

$$\approx_{2\sqrt{\omega}} \mathop{\mathbf{E}}_{\boldsymbol{y}} \sum_o \langle\psi| \, G \otimes M_o^{\boldsymbol{y}} \, |\psi\rangle$$

$$\text{(by Proposition 4.32 and Equation (121))}$$

$$= \langle\psi| \, G \otimes M \, |\psi\rangle.$$

Similarly, for the second term in Equation (123), we have

$$\mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_o \langle\psi| \, G^{\boldsymbol{x}} (M_o^{\boldsymbol{y}})^2 G^{\boldsymbol{x}} \otimes I \, |\psi\rangle = \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_o \langle\psi| \, G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G^{\boldsymbol{x}} \otimes I \, |\psi\rangle \qquad \text{(because } M \text{ is projective)}$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}} \langle\psi| \, G^{\boldsymbol{x}} M G^{\boldsymbol{x}} \otimes I \, |\psi\rangle$$

$$\approx_{2\sqrt{\zeta}} \mathop{\mathbf{E}}_{\boldsymbol{x}} \langle\psi| \, M \otimes G^{\boldsymbol{x}} \, |\psi\rangle \quad \text{(by Proposition 4.32 and Lemma 12.9)}$$

$$= \langle\psi| \, M \otimes G \, |\psi\rangle.$$

For the third term in Equation (123), we begin by claiming that

$$\mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_o \langle\psi| \, G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} \otimes I \, |\psi\rangle = \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} \otimes I \, |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} \otimes I \, |\psi\rangle \quad \text{(because } G \text{ is projective)}$$

$$\approx_{\sqrt{\chi}} \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} \otimes I \, |\psi\rangle. \qquad (124)$$

To show this, we bound the magnitude of the difference.

$$\left| \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, (G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} \otimes I) \cdot ((G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} - M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}}) \otimes I) \, |\psi\rangle \right|$$

$$\leq \sqrt{ \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, (G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G^{\boldsymbol{x}}) \otimes I \, |\psi\rangle }$$

$$\cdot \sqrt{ \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, ((M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} - G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}}) \cdot (G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} - M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}}) \otimes I) \, |\psi\rangle}.$$

The expression inside the first square root is at most 1 because G and M are sub-measurements. The expression inside the second square root is at most χ by Equation (122). Next, we claim that

$$(124) \approx_{\sqrt{\zeta}} \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} \otimes G_g^{\boldsymbol{x}} \, |\psi\rangle. \tag{125}$$

To show this, we bound the magnitude of the difference.

$$\left| \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, (G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} \otimes I) \cdot ((M_o^{\boldsymbol{y}} \otimes I) \cdot (G_g^{\boldsymbol{x}} \otimes I - I \otimes G_g^{\boldsymbol{x}})) \, |\psi\rangle \right|$$

$$\leq \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, (G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G^{\boldsymbol{x}}) \otimes I \, |\psi\rangle}$$

$$\cdot \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, ((G_g^{\boldsymbol{x}} \otimes I - I \otimes G_g^{\boldsymbol{x}}) \cdot (M_o^{\boldsymbol{y}} \otimes I) \cdot (G_g^{\boldsymbol{x}} \otimes I - I \otimes G_g^{\boldsymbol{x}})) \, |\psi\rangle}.$$

The expression inside the first square root is at most 1 because G and M are sub-measurements. The expression inside the second square root is

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_g \langle\psi| \, (G_g^{\boldsymbol{x}} \otimes I - I \otimes G_g^{\boldsymbol{x}}) \cdot \left(\mathop{\mathbf{E}}_{\boldsymbol{y}} \sum_o M_o^{\boldsymbol{y}} \otimes I \right) \cdot (G_g^{\boldsymbol{x}} \otimes I - I \otimes G_g^{\boldsymbol{x}}) \, |\psi\rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_g \langle\psi| \, (G_g^{\boldsymbol{x}} \otimes I - I \otimes G_g^{\boldsymbol{x}})^2 \, |\psi\rangle. \qquad \text{(because } M \text{ is a sub-measurement)}$$

This is at most ζ by Item 3. Next, we claim that

$$(125) \approx_{\sqrt{\zeta}} \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, G_g^{\boldsymbol{x}} G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} \otimes I \, |\psi\rangle. \tag{126}$$

To show this, we bound the magnitude of the difference.

$$\left| \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, ((G_g^{\boldsymbol{x}} \otimes I - I \otimes G_g^{\boldsymbol{x}}) \cdot (G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} \otimes I)) \cdot (G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} \otimes I) \, |\psi\rangle \right|$$

$$\leq \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, ((G_g^{\boldsymbol{x}} \otimes I - I \otimes G_g^{\boldsymbol{x}}) \cdot (G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G^{\boldsymbol{x}} \otimes I) \cdot (G_g^{\boldsymbol{x}} \otimes I - I \otimes G_g^{\boldsymbol{x}})) \, |\psi\rangle}$$

$$\cdot \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, (M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}}) \otimes I \, |\psi\rangle}.$$

The expression inside the first square root is

$$\mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_g \langle\psi| \, ((G_g^{\boldsymbol{x}} \otimes I - I \otimes G_g^{\boldsymbol{x}}) \cdot \left(\mathop{\mathbf{E}}_{\boldsymbol{y}} \sum_o G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G^{\boldsymbol{x}} \otimes I \right) \cdot (G_g^{\boldsymbol{x}} \otimes I - I \otimes G_g^{\boldsymbol{x}})) \, |\psi\rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{x}} \sum_g \langle\psi| \, (G_g^{\boldsymbol{x}} \otimes I - I \otimes G_g^{\boldsymbol{x}})^2 \, |\psi\rangle. \qquad \text{(because } G \text{ and } M \text{ are sub-measurements)}$$

This is at most ζ by Item 3. The expression inside the second square root is at most 1 because G and M are sub-measurements. Next, we claim that

$$(126) = \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} \otimes I \, |\psi\rangle \qquad \text{(because } G \text{ is projective)}$$

$$\approx_{\sqrt{\chi}} \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \, M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} \otimes I \, |\psi\rangle. \tag{127}$$

To show this, we bound the magnitude of the difference.

$$\left| \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \left((G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} - M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}}) \otimes I \right) \cdot (G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} \otimes I) |\psi\rangle \right|$$

$$\leq \sqrt{ \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \left((G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} - M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}}) \cdot (M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} - G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}}) \otimes I \right) |\psi\rangle }$$

$$\cdot \sqrt{ \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| \left(M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} \right) \otimes I |\psi\rangle }.$$

The expression inside the first square root is at most χ by Equation (122). The expression inside the second square root is at most 1 because G and M are sub-measurements. Finally,

$$(127) = \mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_{o,g} \langle\psi| M_o^{\boldsymbol{y}} G_g^{\boldsymbol{x}} M_o^{\boldsymbol{y}} \otimes I |\psi\rangle \qquad\qquad \text{(because } G \text{ is projective)}$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{y}} \sum_o \langle\psi| M_o^{\boldsymbol{y}} G M_o^{\boldsymbol{y}} \otimes I |\psi\rangle$$

$$\approx_{2\sqrt{\omega}} \mathop{\mathbf{E}}_{\boldsymbol{y}} \sum_o \langle\psi| G \otimes M_o^{\boldsymbol{y}} |\psi\rangle \qquad\qquad \text{(by Proposition 4.32 and Equation (121))}$$

$$= \langle\psi| G \otimes M |\psi\rangle .$$

In total, this shows that

$$\mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_o \langle\psi| G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} \otimes I |\psi\rangle \approx_{2\sqrt{\zeta}+2\sqrt{\omega}+2\sqrt{\chi}} \langle\psi| G \otimes M |\psi\rangle . \tag{128}$$

The fourth term in Equation (123) is the Hermitian conjugate of the third term. As a result, Equation (128) implies that

$$\mathop{\mathbf{E}}_{\boldsymbol{x},\boldsymbol{y}} \sum_o \langle\psi| M_o^{\boldsymbol{y}} G^{\boldsymbol{x}} M_o^{\boldsymbol{y}} G^{\boldsymbol{x}} \otimes I |\psi\rangle \approx_{2\sqrt{\zeta}+2\sqrt{\omega}+2\sqrt{\chi}} \langle\psi| G \otimes M |\psi\rangle$$

as well.

In total, this gives an error of

$$2\sqrt{\omega} + 2\sqrt{\zeta} + 2 \cdot \left(2\sqrt{\zeta} + 2\sqrt{\omega} + 2\sqrt{\chi} \right) = 6\sqrt{\zeta} + 6\sqrt{\omega} + 4\sqrt{\chi}.$$

This proves the claimed bound. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \square$

Corollary 12.12 (Commutativity of G's complete part)**.** *The following commutation relations hold.*

$$G_g^{\boldsymbol{x}} G^{\boldsymbol{y}} \otimes I \approx_{\nu_2} G^{\boldsymbol{y}} G_g^{\boldsymbol{x}} \otimes I$$
$$G^{\boldsymbol{x}} G^{\boldsymbol{y}} \otimes I \approx_{\nu_2} G^{\boldsymbol{y}} G^{\boldsymbol{x}} \otimes I,$$

where

$$\nu_2 = 36m \cdot \left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16} \right).$$

Proof. By Equation (115),
$$G_g^x G_h^y \otimes I \approx_{\nu_{\text{commute}}} G_h^y G_g^x \otimes I. \tag{129}$$
Now we apply Lemma 12.11 to Equation (129). To do so, we set the "$\{M_o^y\}$" sub-measurement to be $\{G_g^x\}$, and therefore $\mathcal{O} = \mathcal{P}(m, q, d)$. This implies that
$$G_g^x G^y \otimes I \approx_{\theta_1} G^y G_g^x \otimes I \tag{130}$$
for $\theta_1 = 12\sqrt{\zeta} + 4\sqrt{\nu_{\text{commute}}}$.

Next, we apply Lemma 12.11 to Equation (130). This time, we let $\mathcal{O}$ be a set containing a single outcome, and for this outcome o, we set $M_o^y = G^y$. This implies that
$$G^x G^y \otimes I \approx_{\theta_2} G^y G^x \otimes I$$
for $\theta_2 = 12\sqrt{\zeta} + 4\sqrt{\theta_1}$. This uses Lemma 12.9 for the strong self-consistency of $\{M_o^y\}$.

We now show that ν_2 bounds θ_1 and θ_2. First, using $\sqrt{30} \leq 6$, we have
$$\begin{aligned}
\theta_1 = 12\sqrt{\zeta} + 4\sqrt{\nu_{\text{commute}}} &= 12\sqrt{\zeta} + 4\sqrt{30m \cdot \left(\gamma^{1/4} + \zeta^{1/4} + (d/q)^{1/4}\right)} \\
&\leq 12\zeta^{1/8} + 24m \cdot \left(\gamma^{1/8} + \zeta^{1/8} + (d/q)^{1/8}\right) \\
&\leq 36m \cdot \left(\gamma^{1/8} + \zeta^{1/8} + (d/q)^{1/8}\right).
\end{aligned}$$
This is clearly less than ν_2. Next, we have
$$\begin{aligned}
\theta_2 = 12\sqrt{\zeta} + 4\sqrt{\theta_1} &\leq 12\sqrt{\zeta} + 4\sqrt{36m \cdot \left(\gamma^{1/8} + \zeta^{1/8} + (d/q)^{1/8}\right)} \\
&\leq 12\zeta^{1/16} + 24m \cdot \left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16}\right) \\
&\leq 36m \cdot \left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16}\right).
\end{aligned}$$
This is equal to ν_2, which completes the proof. $\qquad \square$

Corollary 12.13 (Commutativity of G's incomplete part). *The following commutation relations hold.*
$$\begin{aligned}
G_g^x G_\perp^y \otimes I &\approx_{\nu_2} G_\perp^y G_g^x \otimes I \\
G_\perp^x G_\perp^y \otimes I &\approx_{\nu_2} G_\perp^y G_\perp^x \otimes I,
\end{aligned}$$
where ν_2 is as in Corollary 12.12.

Proof. First, we note that
$$G_g^x G_\perp^y - G_\perp^y G_g^x = G_g^x \cdot (I - G^y) - (I - G^y) \cdot G_g^x = G^y G_g^x - G_g^x G^y.$$
Hence, by Corollary 12.12,
$$\mathop{\mathbf{E}}_{x,y} \sum_g \|(G_g^x G_\perp^y - G_\perp^y G_g^x) \otimes I \, |\psi\rangle\|^2 = \mathop{\mathbf{E}}_{x,y} \sum_g \|(G^y G_g^x - G_g^x G^y) \otimes I \, |\psi\rangle\|^2 \leq \nu_2.$$
Next, we note that
$$\begin{aligned}
G_\perp^x G_\perp^y - G_\perp^y G_\perp^x &= (I - G^x) \cdot (I - G^y) - (I - G^y) \cdot (I - G^x) \\
&= (I - G^x - G^y + G^x G^y) - (I - G^y - G^x + G^y G^x) \\
&= G^x G^y - G^y G^x.
\end{aligned}$$
As a result, by Corollary 12.12,
$$\mathop{\mathbf{E}}_{x,y} \|(G_\perp^x G_\perp^y - G_\perp^y G_\perp^x) \otimes I \, |\psi\rangle\|^2 = \mathop{\mathbf{E}}_{x,y} \|(G^y G^x - G^x G^y) \otimes I \, |\psi\rangle\|^2 \leq \nu_2. \qquad \square$$

12.3.3   Putting everything together

Now we combine the results of the previous two sections to show our strong self-consistency and commutation results for $\widehat{G}$.

Corollary 12.14 (Strong self-consistency and commutation of $\widehat{G}$). *$\widehat{G}$ obeys the following strong self-consistency and commutation properties.*

$$\widehat{G}_g^x \otimes I \approx_{2\zeta} I \otimes \widehat{G}_g^x, \tag{131}$$

$$\widehat{G}_g^x \widehat{G}_h^y \otimes I \approx_{\nu_3} \widehat{G}_h^y \widehat{G}_g^x \otimes I, \tag{132}$$

where

$$\nu_3 = 138m \cdot \left(\zeta^{1/16} + \gamma^{1/16} + (d/q)^{1/16} \right).$$

Proof. We begin with Equation (131). To prove this, we wish to bound

$$
\mathbf{E}_{\boldsymbol{x}} \sum_g \| (\widehat{G}_g^{\boldsymbol{x}} \otimes I - I \otimes \widehat{G}_g^{\boldsymbol{x}}) |\psi\rangle \|^2
$$

$$
= \mathbf{E}_{\boldsymbol{x}} \sum_{g \in \mathcal{P}(m,q,d)} \| (G_g^{\boldsymbol{x}} \otimes I - I \otimes G_g^{\boldsymbol{x}}) |\psi\rangle \|^2 + \mathbf{E}_{\boldsymbol{x}} \| (G_\perp^{\boldsymbol{x}} \otimes I - I \otimes G_\perp^{\boldsymbol{x}}) |\psi\rangle \|^2
$$

$$
\leq \zeta + \zeta = 2\zeta,
$$

by Item 3 and Corollary 12.10.

Next, we show Equation (132). To prove this, we wish to bound

$$
\mathbf{E}_{\boldsymbol{x},\boldsymbol{y}} \sum_{g,h} \| (\widehat{G}_g^{\boldsymbol{x}} \widehat{G}_h^{\boldsymbol{y}} - \widehat{G}_h^{\boldsymbol{y}} \widehat{G}_g^{\boldsymbol{x}}) \otimes I |\psi\rangle \|^2
$$

$$
= \mathbf{E}_{\boldsymbol{x},\boldsymbol{y}} \sum_{g,h \in \mathcal{P}(m,q,d)} \| (G_g^{\boldsymbol{x}} G_h^{\boldsymbol{y}} - G_h^{\boldsymbol{y}} G_g^{\boldsymbol{x}}) \otimes I |\psi\rangle \|^2 + \mathbf{E}_{\boldsymbol{x},\boldsymbol{y}} \| (G_\perp^{\boldsymbol{x}} G_\perp^{\boldsymbol{y}} - G_\perp^{\boldsymbol{y}} G_\perp^{\boldsymbol{x}}) \otimes I |\psi\rangle \|^2
$$

$$
+ \mathbf{E}_{\boldsymbol{x},\boldsymbol{y}} \sum_{g \in \mathcal{P}(m,q,d)} \| (G_g^{\boldsymbol{x}} G_\perp^{\boldsymbol{y}} - G_\perp^{\boldsymbol{y}} G_g^{\boldsymbol{x}}) \otimes I |\psi\rangle \|^2 + \mathbf{E}_{\boldsymbol{x},\boldsymbol{y}} \sum_{h \in \mathcal{P}(m,q,d)} \| (G_\perp^{\boldsymbol{x}} G_h^{\boldsymbol{y}} - G_h^{\boldsymbol{y}} G_\perp^{\boldsymbol{x}}) \otimes I |\psi\rangle \|^2
$$

$$
\leq \nu_{\text{commute}} + 3\nu_2,
$$

by Equation (115) and Corollary 12.13. We can therefore bound this by

$$
\nu_{\text{commute}} + 3\nu_2 = 30m \cdot \left(\gamma^{1/4} + \zeta^{1/4} + (d/q)^{1/4} \right) + 3 \cdot 36m \cdot \left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16} \right)
$$

$$
\leq 30m \cdot \left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16} \right) + 3 \cdot 36m \cdot \left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16} \right)
$$

$$
= 138m \cdot \left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16} \right).
$$

This completes the proof. $\qquad \square$

12.4   Consistency of the sandwich $\widehat{H}$ with B

In the next lemmas, we will show that $\widehat{H}$ is consistent with the lines measurement B. To start, we show some self-consistency and commutativity properties of $\widehat{H}$.

Lemma 12.15 (Commuting past multiple $\widehat{G}$'s). *For all $k \geq 2$,*

$$\widehat{G}_{g_1}^{x_1}\widehat{G}_{g_2}^{x_2}\cdots\widehat{G}_{g_k}^{x_k} \otimes I \approx_{\nu_4} \widehat{G}_{g_2}^{x_2}\cdots\widehat{G}_{g_k}^{x_k}\widehat{G}_{g_1}^{x_1} \otimes I,$$

where

$$\nu_4 = 426k^2m \cdot \left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16}\right).$$

Proof. This proof will consist of multiple applications of Equations (131) and (132); for each line, we will specify which equation to apply. Each line will also involve an application of Proposition 4.26, which we will specify only implicitly.

$$\widehat{G}_{g_1}^{x_1}\widehat{G}_{g_2}^{x_2}\cdots\widehat{G}_{g_k}^{x_k} \otimes I$$

$$\approx_{2\zeta} \widehat{G}_{g_1}^{x_1}\widehat{G}_{g_2}^{x_2}\cdots\widehat{G}_{g_{k-1}}^{x_{k-1}} \otimes \widehat{G}_{g_k}^{x_k} \qquad \text{(by Equation (131))}$$

$$\cdots$$

$$\approx_{2\zeta} \widehat{G}_{g_1}^{x_1}\widehat{G}_{g_2}^{x_2} \otimes \widehat{G}_{g_k}^{x_k}\cdots\widehat{G}_{g_3}^{x_3} \qquad \text{(by Equation (131))}$$

$$\approx_{\nu_3} \widehat{G}_{g_2}^{x_2}G_{g_1}^{x_1} \otimes \widehat{G}_{g_k}^{x_k}\cdots\widehat{G}_{g_3}^{x_3} \qquad \text{(by Equation (132))}$$

$$\approx_{2\zeta} \widehat{G}_{g_2}^{x_2}G_{g_1}^{x_1}\widehat{G}_{g_3}^{x_3} \otimes \widehat{G}_{g_k}^{x_k}\cdots\widehat{G}_{g_4}^{x_4} \qquad \text{(by Equation (131))}$$

$$\approx_{\nu_3} \widehat{G}_{g_2}^{x_2}\widehat{G}_{g_3}^{x_3}\widehat{G}_{g_1}^{x_1} \otimes \widehat{G}_{g_k}^{x_k}\cdots\widehat{G}_{g_4}^{x_4} \qquad \text{(by Equation (132))}$$

$$\cdots$$

$$\approx_{2\zeta} \widehat{G}_{g_2}^{x_2}\cdots\widehat{G}_{g_{k-1}}^{x_{k-1}}\widehat{G}_{g_1}^{x_1}\widehat{G}_{g_k}^{x_k} \otimes I \qquad \text{(by Equation (131))}$$

$$\approx_{\nu_3} \widehat{G}_{g_2}^{x_2}\cdots\widehat{G}_{g_{k-1}}^{x_{k-1}}\widehat{G}_{g_k}^{x_k}\widehat{G}_{g_1}^{x_1} \otimes I. \qquad \text{(by Equation (132))}$$

In total, we have $(k-2) + (k-2) \leq 2k$ applications of Equation (131) with error 2ζ each and $(k-1) \leq k$ applications of Equation (132) with error ν_3 each. By Proposition 4.28, this implies that

$$G_{g_1}^{x_1}G_{g_2}^{x_2}\cdots G_{g_k}^{x_k} \otimes I \approx_{3k \cdot (4k\zeta + k\nu_3)} G_{g_2}^{x_2}\cdots G_{g_{k-1}}^{x_{k-1}}G_{g_k}^{x_k}G_{g_1}^{x_1} \otimes I,$$

as claimed. We can bound this as follows.

$$3k \cdot (4k\zeta + k\nu_3) = 12k^2\zeta + 3k^2\nu_3$$

$$= 12k^2\zeta + 3k^2 \cdot 138m \cdot \left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16}\right)$$

$$\leq 12k^2\zeta^{1/16} + 414k^2m \cdot \left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16}\right)$$

$$\leq 426k^2m \cdot \left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16}\right).$$

This completes the proof. □

Lemma 12.16 (Consistency of $\widehat{H}$ with B). *For any $1 \leq i \leq k$,*

$$\mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_k} \sum_{g_1,\ldots,g_k : g_i \neq \perp} \sum_{a \neq g_i(\boldsymbol{u})} \langle\psi| \widehat{H}_{g_1,\ldots,g_k}^{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_k} \otimes B_{[f(\boldsymbol{x}_i)=a]}^{\boldsymbol{u}} |\psi\rangle \leq \nu_5,$$

where

$$\nu_5 = 43km \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right).$$

Proof. To begin, we note that

$$\sum_{g_{i+1},\ldots,g_k} \widehat{H}^{x_1,\ldots,x_k}_{g_1,\ldots,g_k} = \sum_{g_{i+1},\ldots,g_k} \widehat{G}^{x_1}_{g_1} \cdots \widehat{G}^{x_k}_{g_k} \cdots \widehat{G}^{x_1}_{g_1}$$

$$= \sum_{g_{i+1},\ldots,g_{k-1}} \widehat{G}^{x_1}_{g_1} \cdots \Big(\sum_{g_k} \widehat{G}^{x_k}_{g_k}\Big) \cdots \widehat{G}^{x_1}_{g_1}$$

$$= \sum_{g_{i+1},\ldots,g_{k-1}} \widehat{G}^{x_1}_{g_1} \cdots \widehat{G}^{x_{k-1}}_{g_{k-1}} \cdot I \cdot \widehat{G}^{x_{k-1}}_{g_{k-1}} \cdots \widehat{G}^{x_1}_{g_1} \quad \text{(because } \widehat{G}^{x_k} \text{ is a measurement)}$$

$$= \sum_{g_{i+1},\ldots,g_{k-1}} \widehat{G}^{x_1}_{g_1} \cdots \widehat{G}^{x_{k-1}}_{g_{k-1}} \cdots \widehat{G}^{x_1}_{g_1} \quad\quad\quad \text{(because } \widehat{G}^{x_{k-1}} \text{ is projective)}$$

$$\ldots$$

$$= \widehat{G}^{x_1}_{g_1} \cdots \widehat{G}^{x_i}_{g_i} \cdots \widehat{G}^{x_1}_{g_1}$$

$$= \widehat{H}^{x_1,\ldots,x_i}_{g_1,\ldots,g_i}.$$

As a result,

$$\operatorname*{\mathbf{E}}_{\boldsymbol{u}} \operatorname*{\mathbf{E}}_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_k} \sum_{g_1,\ldots,g_k:g_i\neq\perp} \sum_{a\neq g_i(\boldsymbol{u})} \langle\psi| \widehat{H}^{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_k}_{g_1,\ldots,g_k} \otimes B^{\boldsymbol{u}}_{[f(\boldsymbol{x}_i)=a]} |\psi\rangle$$

$$= \operatorname*{\mathbf{E}}_{\boldsymbol{u}} \operatorname*{\mathbf{E}}_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_i} \sum_{g_1,\ldots,g_i:g_i\neq\perp} \sum_{a\neq g_i(\boldsymbol{u})} \langle\psi| \widehat{H}^{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_i}_{g_1,\ldots,g_i} \otimes B^{\boldsymbol{u}}_{[f(\boldsymbol{x}_i)=a]} |\psi\rangle. \tag{133}$$

For shorthand, we write

$$\widehat{G}^{x_{<i}}_{g_{<i}} = \widehat{G}^{x_1}_{g_1} \cdots \widehat{G}^{x_{i-1}}_{g_{i-1}}.$$

Then we claim that

$$(133) = \operatorname*{\mathbf{E}}_{\boldsymbol{u}} \operatorname*{\mathbf{E}}_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_i} \sum_{g_1,\ldots,g_i:g_i\neq\perp} \langle\psi| \widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}} \cdot \widehat{G}^{\boldsymbol{x}_i}_{g_i} \cdot \widehat{G}^{\boldsymbol{x}_i}_{g_i} \cdot (\widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}})^{\dagger} \otimes (I - B^{\boldsymbol{u}}_{[f(\boldsymbol{x}_i)=g_i(\boldsymbol{u})]}) |\psi\rangle$$

$$\approx_{\sqrt{\nu_4}} \operatorname*{\mathbf{E}}_{\boldsymbol{u}} \operatorname*{\mathbf{E}}_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_i} \sum_{g_1,\ldots,g_i:g_i\neq\perp} \langle\psi| \widehat{G}^{\boldsymbol{x}_i}_{g_i} \cdot \widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}} \cdot \widehat{G}^{\boldsymbol{x}_i}_{g_i} \cdot (\widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}})^{\dagger} \otimes (I - B^{\boldsymbol{u}}_{[f(\boldsymbol{x}_i)=g_i(\boldsymbol{u})]}) |\psi\rangle. \tag{134}$$

To show this, we bound the magnitude of the difference.

$$\operatorname*{\mathbf{E}}_{\boldsymbol{u}} \operatorname*{\mathbf{E}}_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_i} \sum_{g_1,\ldots,g_i:g_i\neq\perp} \langle\psi| ((\widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}} \cdot \widehat{G}^{\boldsymbol{x}_i}_{g_i} - \widehat{G}^{\boldsymbol{x}_i}_{g_i} \cdot \widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}}) \otimes I) \cdot (\widehat{G}^{\boldsymbol{x}_i}_{g_i} \cdot (\widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}})^{\dagger} \otimes (I - B^{\boldsymbol{u}}_{[f(\boldsymbol{x}_i)=g_i(\boldsymbol{u})]})) |\psi\rangle$$

$$\leq \sqrt{\operatorname*{\mathbf{E}}_{\boldsymbol{u}} \operatorname*{\mathbf{E}}_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_i} \sum_{g_1,\ldots,g_i:g_i\neq\perp} \langle\psi| ((\widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}} \cdot \widehat{G}^{\boldsymbol{x}_i}_{g_i} - \widehat{G}^{\boldsymbol{x}_i}_{g_i} \cdot \widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}}) \cdot (\widehat{G}^{\boldsymbol{x}_i}_{g_i} \cdot (\widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}})^{\dagger} - (\widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}})^{\dagger} \cdot \widehat{G}^{\boldsymbol{x}_i}_{g_i})) \otimes I |\psi\rangle}$$

$$\cdot \sqrt{\operatorname*{\mathbf{E}}_{\boldsymbol{u}} \operatorname*{\mathbf{E}}_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_i} \sum_{g_1,\ldots,g_i:g_i\neq\perp} \langle\psi| (\widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}} \cdot \widehat{G}^{\boldsymbol{x}_i}_{g_i} \cdot (\widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}})^{\dagger}) \otimes (I - B^{\boldsymbol{u}}_{[f(\boldsymbol{x}_i)=g_i(\boldsymbol{u})]})^2 |\psi\rangle}.$$

The term inside the first square root is at most

$$\operatorname*{\mathbf{E}}_{\boldsymbol{u}} \operatorname*{\mathbf{E}}_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_i} \sum_{g_1,\ldots,g_i} \langle\psi| ((\widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}} \cdot \widehat{G}^{\boldsymbol{x}_i}_{g_i} - \widehat{G}^{\boldsymbol{x}_i}_{g_i} \cdot \widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}}) \cdot (\widehat{G}^{\boldsymbol{x}_i}_{g_i} \cdot (\widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}})^{\dagger} - (\widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}})^{\dagger} \cdot \widehat{G}^{\boldsymbol{x}_i}_{g_i})) \otimes I |\psi\rangle, \tag{135}$$

which is at most ν_4 by Lemma 12.15. The term inside the second square root is at most 1 because B and $\widehat{G}$ are measurements. Next, we claim that

$$(134) \approx_{\sqrt{\nu_4}} \operatorname*{\mathbf{E}}_{\boldsymbol{u}} \operatorname*{\mathbf{E}}_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_i} \sum_{g_1,\ldots,g_i:g_i\neq\perp} \langle\psi| \widehat{G}^{\boldsymbol{x}_i}_{g_i} \cdot \widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}} \cdot (\widehat{G}^{\boldsymbol{x}_{<i}}_{g_{<i}})^{\dagger} \cdot \widehat{G}^{\boldsymbol{x}_i}_{g_i} \otimes (I - B^{\boldsymbol{u}}_{[f(\boldsymbol{x}_i)=g_i(\boldsymbol{u})]}) |\psi\rangle. \tag{136}$$

To show this, we bound the magnitude of the difference.

$$\mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{x}_1,\dots,\boldsymbol{x}_i} \sum_{g_1,\dots,g_i:g_i\neq\perp} \langle\psi| (\widehat{G}_{g_i}^{\boldsymbol{x}_i} \cdot \widehat{G}_{g_{<i}}^{\boldsymbol{x}_{<i}} \otimes (I - B_{[f(\boldsymbol{x}_i)=g_i(\boldsymbol{u})]}^{\boldsymbol{u}})) \cdot (((\widehat{G}_{g_{<i}}^{\boldsymbol{x}_{<i}})^\dagger \cdot \widehat{G}_{g_i}^{\boldsymbol{x}_i} - \widehat{G}_{g_i}^{\boldsymbol{x}_i} \cdot (\widehat{G}_{g_{<i}}^{\boldsymbol{x}_{<i}})^\dagger) \otimes I) |\psi\rangle$$

$$\leq \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{x}_1,\dots,\boldsymbol{x}_i} \sum_{g_1,\dots,g_i:g_i\neq\perp} \langle\psi| (\widehat{G}_{g_i}^{\boldsymbol{x}_i} \cdot \widehat{G}_{g_{<i}}^{\boldsymbol{x}_{<i}} \cdot (\widehat{G}_{g_{<i}}^{\boldsymbol{x}_{<i}})^\dagger \cdot \widehat{G}_{g_i}^{\boldsymbol{x}_i}) \otimes (I - B_{[f(\boldsymbol{x}_i)=g_i(\boldsymbol{u})]}^{\boldsymbol{u}})^2 |\psi\rangle}$$

$$\cdot \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{x}_1,\dots,\boldsymbol{x}_i} \sum_{g_1,\dots,g_i:g_i\neq\perp} \langle\psi| ((\widehat{G}_{g_{<i}}^{\boldsymbol{x}_{<i}} \cdot \widehat{G}_{g_i}^{\boldsymbol{x}_i} - \widehat{G}_{g_i}^{\boldsymbol{x}_i} \cdot \widehat{G}_{g_{<i}}^{\boldsymbol{x}_{<i}}) \cdot ((\widehat{G}_{g_{<i}}^{\boldsymbol{x}_{<i}})^\dagger \cdot \widehat{G}_{g_i}^{\boldsymbol{x}_i} - \widehat{G}_{g_i}^{\boldsymbol{x}_i} \cdot (\widehat{G}_{g_{<i}}^{\boldsymbol{x}_{<i}}))^\dagger) \otimes I |\psi\rangle}.$$

The term inside the first square root is at most 1 because B and $\widehat{G}$ are measurements. The term inside the second square root is at most Equation (135), which is at most ν_4 by Lemma 12.15. But

$$\sum_{g_1,\dots,g_{i-1}} \widehat{G}_{g_{<i}}^{x_{<i}} \cdot (\widehat{G}_{g_{<i}}^{x_{<i}})^\dagger = \sum_{g_1,\dots,g_{i-1}} \widehat{G}_{g_1}^{x_1} \cdots \widehat{G}_{g_{i-1}}^{x_{i-1}} \cdot \widehat{G}_{g_{i-1}}^{x_{i-1}} \cdots \widehat{G}_{g_1}^{x_1}$$

$$= \sum_{g_1,\dots,g_{i-2}} \widehat{G}_{g_1}^{x_1} \cdots \left(\sum_{g_{i-1}} \widehat{G}_{g_{i-1}}^{x_{i-1}} \right) \cdots \widehat{G}_{g_1}^{x_1}$$

$$= \sum_{g_1,\dots,g_{i-2}} \widehat{G}_{g_1}^{x_1} \cdots I \cdots \widehat{G}_{g_1}^{x_1}$$

$$\cdots$$

$$= I.$$

Thus,

$$(136) = \mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{x}_1,\dots,\boldsymbol{x}_i} \sum_{g_i:g_i\neq\perp} \langle\psi| \widehat{G}_{g_i}^{\boldsymbol{x}_i} \otimes (I - B_{[f(\boldsymbol{x}_i)=g_i(\boldsymbol{u})]}^{\boldsymbol{u}}) |\psi\rangle \leq \nu_1,$$

by Equation (113). In total, using $\sqrt{426} \leq 21$, this gives an error of

$$\nu_1 + 2\sqrt{\nu_4} = \zeta + \sqrt{8m\epsilon + 4\delta} + 2 \cdot \sqrt{426k^2m \cdot \left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16}\right)}$$

$$\leq \zeta^{1/32} + 3m\epsilon^{1/32} + 2\delta^{1/32} + 42km \cdot \left(\gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right)$$

$$\leq 43km \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right).$$

This completes the proof. $\qquad\qquad\square$

12.5   Consistency of H with A

Lemma 12.17 (Consistency of H with B).

$$H_{[h|_u=f]} \otimes I \simeq_{\nu_6} I \otimes B_f^u,$$

where

$$\nu_6 = 44k^2m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right).$$

Proof. Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \sim \mathbb{F}_q$ be independent and uniformly random. Our goal is to bound

$$\mathbf{E}_{\boldsymbol{u}} \sum_{f \neq f'} \langle \psi | H_{[h|_{\boldsymbol{u}} = f']} \otimes B_f^{\boldsymbol{u}} | \psi \rangle$$

$$= \mathbf{E}_{\boldsymbol{u}} \sum_h \sum_{f \neq h|_{\boldsymbol{u}}} \langle \psi | H_h \otimes B_f^{\boldsymbol{u}} | \psi \rangle$$

$$= \mathbf{E}_{\boldsymbol{u}} \mathbf{E}_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \sum_h \sum_{f \neq h|_{\boldsymbol{u}}} \langle \psi | H_h^{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \otimes B_f^{\boldsymbol{u}} | \psi \rangle$$

$$= \mathbf{E}_{\boldsymbol{u}} \mathbf{E}_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \sum_h \sum_{w : |w| \geq d+1} \sum_{f \neq h|_{\boldsymbol{u}}} \langle \psi | \widehat{H}_{h_w}^{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \otimes B_f^{\boldsymbol{u}} | \psi \rangle$$

$$= \mathbf{E}_{\boldsymbol{u}} \mathbf{E}_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \sum_h \sum_{w : |w| \geq d+1} \sum_{(g_1, \ldots, g_k) = h_w} \sum_{f \neq h|_{\boldsymbol{u}}} \langle \psi | \widehat{H}_{g_1, \ldots, g_k}^{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \otimes B_f^{\boldsymbol{u}} | \psi \rangle. \qquad (137)$$

We note that the sum over $(g_1, \ldots, g_k) = h_w$ in the final step is trivial because there is only ever one tuple $(g_1, \ldots, g_k)$ which is equal to h_w. Because $|w| \geq d+1$, there exist at least $(d+1)$ coordinates i such that $g_i \neq \perp$ and hence $g_i = h|_{\boldsymbol{x}_i}$. Since f is degree-d, if it is not equal to $h|_{\boldsymbol{u}}$, then there must exist an i such that $g_i \neq \perp$ and $g_i(\boldsymbol{u}) \neq f(\boldsymbol{x}_i)$. Thus,

$$(137) = \mathbf{E}_{\boldsymbol{u}} \mathbf{E}_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \sum_h \sum_{w : |w| \geq d+1} \sum_{(g_1, \ldots, g_k) = h_w} \sum_{\substack{f : \exists i : g_i \neq \perp, \\ g_i(\boldsymbol{u}) \neq f(\boldsymbol{x}_i)}} \langle \psi | \widehat{H}_{g_1, \ldots, g_k}^{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \otimes B_f^{\boldsymbol{u}} | \psi \rangle$$

$$\leq \mathbf{E}_{\boldsymbol{u}} \mathbf{E}_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \sum_{g_1, \ldots, g_k} \sum_{\substack{f : \exists i : g_i \neq \perp, \\ g_i(\boldsymbol{u}) \neq f(\boldsymbol{x}_i)}} \langle \psi | \widehat{H}_{g_1, \ldots, g_k}^{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \otimes B_f^{\boldsymbol{u}} | \psi \rangle. \qquad (138)$$

Let $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k) \sim \mathsf{Distinct}_k$. Then by Proposition 12.4, Equation (138) is (k^2/q)-close to

$$\mathbf{E}_{\boldsymbol{u}} \mathbf{E}_{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k} \sum_{g_1, \ldots, g_k} \sum_{\substack{f : \exists i : g_i \neq \perp, \\ g_i(\boldsymbol{u}) \neq f(\boldsymbol{y}_i)}} \langle \psi | \widehat{H}_{g_1, \ldots, g_k}^{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k} \otimes B_f^{\boldsymbol{u}} | \psi \rangle$$

$$\leq \sum_i \mathbf{E}_{\boldsymbol{u}} \mathbf{E}_{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k} \sum_{g_1, \ldots, g_k} \sum_{\substack{f : g_i \neq \perp, \\ g_i(\boldsymbol{u}) \neq f(\boldsymbol{y}_i)}} \langle \psi | \widehat{H}_{g_1, \ldots, g_k}^{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k} \otimes B_f^{\boldsymbol{u}} | \psi \rangle \qquad \text{(by the union bound)}$$

$$= \sum_i \mathbf{E}_{\boldsymbol{u}} \mathbf{E}_{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k} \sum_{g_1, \ldots, g_k : g_i \neq \perp} \sum_{a \neq g_i(\boldsymbol{u})} \langle \psi | \widehat{H}_{g_1, \ldots, g_k}^{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k} \otimes B_{[f(\boldsymbol{y}_i) = a]}^{\boldsymbol{u}} | \psi \rangle$$

$$\leq \sum_i \nu_5 \qquad \text{(by Lemma 12.16)}$$

$$= k \cdot \nu_5.$$

In total, using $1/q \leq (d/q)^{1/32}$, this gives an error of

$$\frac{k^2}{q} + k \cdot \nu_5 = \frac{k^2}{q} + k \cdot 43km \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32} \right)$$

$$= 44k^2m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32} \right).$$

This completes the proof. $\qquad \square$

Corollary 12.18 (Consistency of H with A; Proof of Item 1 in Lemma 12.2)**.**

$$H_{[h(u,x) = a]} \otimes I \simeq_\nu I \otimes A_a^{u,x}.$$

Proof. Lemma 12.17 implies that

$$H_{[h(u,x)=a]} \otimes I \simeq_{\nu_6} I \otimes B^u_{[f(x)=a]}. \tag{139}$$

Proposition 4.21 applied to Equation (139) and Equation (112) implies that

$$H_{[h(u,x)=a]} \otimes I \simeq_{\nu_6+\sqrt{8m\epsilon+4\delta}} I \otimes A^{u,x}_a.$$

We can bound this error by

$$
\begin{aligned}
\sqrt{8m\epsilon+4\delta} + \nu_6 &= \sqrt{8m\epsilon+4\delta} + 44k^2 m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32} \right) \\
&\leq 3m\epsilon^{1/32} + 2\delta^{1/32} + 44k^2 m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32} \right) \\
&\leq 47k^2 m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32} \right).
\end{aligned}
$$

This is clearly at most ν, which completes the proof. $\qquad\square$

12.6 Completeness of H

Definition 12.19. Let $\tau \in \{0,1\}^k$ be a type. We define the following two subsets of $\mathcal{P}^+(m,q,d)^k$.

- We define $\mathsf{Outcomes}_\tau$ to be the set of tuples $(g_1,\ldots,g_k)$ such that $g_i \in \mathcal{P}(m,q,d)$ for each $i \in \tau$ and $g_i = \perp$ for each $i \notin \tau$. This is the set of possible outcomes of the $\widehat{H}$ measurement of type τ.

- Let $x_1,\ldots,x_k \in \mathbb{F}_q$. We define $\mathsf{Global}_\tau(x)$ to be the subset of $\mathsf{Outcomes}_\tau$ containing only those tuples which are consistent with a global polynomial. In other words, $(g_1,\ldots,g_k) \in \mathsf{Global}_\tau(x)$ if there exists an $h \in \mathcal{P}(m+1,q,d)$ such that $g_i = h|_{x_i}$ for each $i \in \tau$. Next, we define

$$\overline{\mathsf{Global}_\tau(x)} = \mathsf{Outcomes}_\tau \setminus \mathsf{Global}_\tau(x).$$

This contains those tuples of type τ with no consistent global polynomial.

Lemma 12.20. *Let $x_1,\ldots,x_k \sim \mathbb{F}_q$ be sampled uniformly at random. Then*

$$\langle\psi| H \otimes I |\psi\rangle \approx_{\nu_7} \mathop{\mathbf{E}}_{x_1,\ldots,x_k} \sum_{\tau:|\tau|\geq d+1} \sum_{(g_1,\ldots,g_k)\in\mathsf{Outcomes}_\tau} \langle\psi| \widehat{H}^{x_1,\ldots,x_k}_{g_1,\ldots,g_k} \otimes I |\psi\rangle,$$

where

$$\nu_7 = 46k^2 m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32} \right).$$

Proof. Let $(y_1,\ldots,y_k) \sim \mathsf{Distinct}_k$. By definition,

$$
\begin{aligned}
\langle\psi| H \otimes I |\psi\rangle &= \sum_h \langle\psi| H_h \otimes I |\psi\rangle \\
&= \mathop{\mathbf{E}}_{y_1,\ldots,y_k} \sum_h \langle\psi| H^{y_1,\ldots,y_k}_h \otimes I |\psi\rangle \\
&= \mathop{\mathbf{E}}_{y_1,\ldots,y_k} \sum_h \sum_{\tau:|\tau|\geq d+1} \langle\psi| \widehat{H}^{y_1,\ldots,y_k}_{h_\tau} \otimes I |\psi\rangle \\
&= \mathop{\mathbf{E}}_{y_1,\ldots,y_k} \sum_h \sum_{\tau:|\tau|\geq d+1} \sum_{(g_1,\ldots,g_k)=h_\tau} \langle\psi| \widehat{H}^{y_1,\ldots,y_k}_{g_1,\ldots,g_k} \otimes I |\psi\rangle \\
&= \mathop{\mathbf{E}}_{y_1,\ldots,y_k} \sum_{\tau:|\tau|\geq d+1} \sum_{(g_1,\ldots,g_k)\in\mathsf{Global}_\tau(y)} \langle\psi| \widehat{H}^{y_1,\ldots,y_k}_{g_1,\ldots,g_k} \otimes I |\psi\rangle. \tag{140}
\end{aligned}
$$

110

The sum in Equation (140) is over $g_1, \ldots, g_k$ which are consistent with a global polynomial. We will now show that the value of this sum remains largely unchanged if we drop this condition. In particular, we claim that

$$(140) \approx_{\frac{k^2}{q} + k \cdot \nu_5 + \frac{md}{q}} \mathop{\mathbf{E}}_{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k} \sum_{\tau : |\tau| \geq d+1} \sum_{(g_1, \ldots, g_k) \in \mathsf{Outcomes}_\tau} \langle \psi | \widehat{H}^{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k}_{g_1, \ldots, g_k} \otimes I \, |\psi \rangle. \tag{141}$$

To show this, we note that $(141) \geq (140)$. Hence, it suffices to upper bound their difference.

$$(141) - (140) = \mathop{\mathbf{E}}_{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k} \sum_{\tau : |\tau| \geq d+1} \sum_{(g_1, \ldots, g_k) \in \overline{\mathsf{Global}}_\tau(\boldsymbol{y})} \langle \psi | \widehat{H}^{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k}_{g_1, \ldots, g_k} \otimes I \, |\psi \rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k} \sum_{\tau : |\tau| \geq d+1} \sum_f \sum_{(g_1, \ldots, g_k) \in \overline{\mathsf{Global}}_\tau(\boldsymbol{y})} \langle \psi | \widehat{H}^{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k}_{g_1, \ldots, g_k} \otimes B^{\boldsymbol{u}}_f \, |\psi \rangle, \tag{142}$$

because B is a measurement. Next, we claim that

$$(142) \approx_{\frac{k^2}{q} + k \cdot \nu_5} \mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k} \sum_{\tau : |\tau| \geq d+1} \sum_f \sum_{(g_1, \ldots, g_k) \in \overline{\mathsf{Global}}_\tau(\boldsymbol{y})} \langle \psi | \widehat{H}^{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k}_{g_1, \ldots, g_k} \otimes B^{\boldsymbol{u}}_f \, |\psi \rangle \cdot \mathbf{1}[\forall i \in \tau, f(\boldsymbol{y}_i) = g_i(\boldsymbol{u})].$$
$$\tag{143}$$

To show this, we note that $(142) \geq (143)$. Thus, it suffices to upper bound their difference $(142) - (143)$. This is given by

$$\mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k} \sum_{\tau : |\tau| \geq d+1} \sum_f \sum_{(g_1, \ldots, g_k) \in \overline{\mathsf{Global}}_\tau(\boldsymbol{y})} \langle \psi | \widehat{H}^{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k}_{g_1, \ldots, g_k} \otimes B^{\boldsymbol{u}}_f \, |\psi \rangle \cdot \mathbf{1}[\exists i \in \tau, f(\boldsymbol{y}_i) \neq g_i(\boldsymbol{u})]. \tag{144}$$

Recall that $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \sim \mathbb{F}_q$ are sampled independently and uniformly at random. Then Proposition 12.4 implies that Equation (144) is (k^2/q)-close to

$$\mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \sum_{\tau : |\tau| \geq d+1} \sum_f \sum_{(g_1, \ldots, g_k) \in \overline{\mathsf{Global}}_\tau(\boldsymbol{x})} \langle \psi | \widehat{H}^{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k}_{g_1, \ldots, g_k} \otimes B^{\boldsymbol{u}}_f \, |\psi \rangle \cdot \mathbf{1}[\exists i \in \tau, f(\boldsymbol{x}_i) \neq g_i(\boldsymbol{u})]$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \sum_{\tau : |\tau| \geq d+1} \sum_f \sum_{(g_1, \ldots, g_k) \in \mathsf{Outcomes}_\tau} \langle \psi | \widehat{H}^{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k}_{g_1, \ldots, g_k} \otimes B^{\boldsymbol{u}}_f \, |\psi \rangle \cdot \mathbf{1}[\exists i \in \tau, f(\boldsymbol{x}_i) \neq g_i(\boldsymbol{u})]$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \sum_\tau \sum_f \sum_{(g_1, \ldots, g_k) \in \mathsf{Outcomes}_\tau} \langle \psi | \widehat{H}^{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k}_{g_1, \ldots, g_k} \otimes B^{\boldsymbol{u}}_f \, |\psi \rangle \cdot \mathbf{1}[\exists i \in \tau, f(\boldsymbol{x}_i) \neq g_i(\boldsymbol{u})]$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \sum_\tau \sum_f \sum_{(g_1, \ldots, g_k) \in \mathsf{Outcomes}_\tau} \langle \psi | \widehat{H}^{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k}_{g_1, \ldots, g_k} \otimes B^{\boldsymbol{u}}_f \, |\psi \rangle \cdot \left(\sum_{i \in \tau} \mathbf{1}[f(\boldsymbol{x}_i) \neq g_i(\boldsymbol{u})] \right)$$

$$= \sum_i \mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \sum_{g_1, \ldots, g_k : g_i \neq \perp} \sum_{f : f(\boldsymbol{x}_i) \neq g_i(\boldsymbol{u})} \langle \psi | \widehat{H}^{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k}_{g_1, \ldots, g_k} \otimes B^{\boldsymbol{u}}_f \, |\psi \rangle$$

$$= \sum_i \mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k} \sum_{g_1, \ldots, g_k : g_i \neq \perp} \sum_{a \neq g_i(\boldsymbol{u})} \langle \psi | \widehat{H}^{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k}_{g_1, \ldots, g_k} \otimes B^{\boldsymbol{u}}_{[f(\boldsymbol{x}_i)=a]} \, |\psi \rangle$$

$$\leq \sum_i (\nu_5) \qquad\qquad \text{(by Lemma 12.16)}$$

$$= k \cdot \nu_5.$$

Returning to Equation (143), we introduce the notation $\mathsf{Consistent}_\tau(g, \boldsymbol{y}, \boldsymbol{u})$ to indicate whether there is a consistent degree-d polynomial interpolating the g_i's along the line in direction u. In other words,

$$\mathsf{Consistent}_\tau(g, \boldsymbol{y}, \boldsymbol{u}) = \begin{cases} 1 & \text{if } \exists f \text{ such that } f(\boldsymbol{y}_i) = g_i(\boldsymbol{u}) \text{ for all } i \in \tau, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, for any f,

$$\mathbf{1}[\forall i \in \tau, f(\boldsymbol{y}_i) = g_i(\boldsymbol{u})] \leq \mathbf{1}[\mathsf{Consistent}_\tau(g, \boldsymbol{y}, \boldsymbol{u})]$$

because the left-hand side indicates whether f is the consistent degree-d polynomial interpolating along direction u. As a result,

$$
\begin{aligned}
(143) \leq\ & \mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{y}_1,\dots,\boldsymbol{y}_k} \sum_{\tau:|\tau|\geq d+1} \sum_{f} \sum_{(g_1,\dots,g_k)\in\overline{\mathsf{Global}_\tau(\boldsymbol{y})}} \langle\psi|\, \widehat{H}^{\boldsymbol{y}_1,\dots,\boldsymbol{y}_k}_{g_1,\dots,g_k} \otimes B^{\boldsymbol{u}}_f \,|\psi\rangle \cdot \mathbf{1}[\mathsf{Consistent}_\tau(g,\boldsymbol{y},\boldsymbol{u})] \\
=\ & \mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{y}_1,\dots,\boldsymbol{y}_k} \sum_{\tau:|\tau|\geq d+1} \sum_{(g_1,\dots,g_k)\in\overline{\mathsf{Global}_\tau(\boldsymbol{y})}} \langle\psi|\, \widehat{H}^{\boldsymbol{y}_1,\dots,\boldsymbol{y}_k}_{g_1,\dots,g_k} \otimes \Big(\sum_f B^{\boldsymbol{u}}_f\Big) \,|\psi\rangle \cdot \mathbf{1}[\mathsf{Consistent}_\tau(g,\boldsymbol{y},\boldsymbol{u})] \\
=\ & \mathop{\mathbf{E}}_{\boldsymbol{u}} \mathop{\mathbf{E}}_{\boldsymbol{y}_1,\dots,\boldsymbol{y}_k} \sum_{\tau:|\tau|\geq d+1} \sum_{(g_1,\dots,g_k)\in\overline{\mathsf{Global}_\tau(\boldsymbol{y})}} \langle\psi|\, \widehat{H}^{\boldsymbol{y}_1,\dots,\boldsymbol{y}_k}_{g_1,\dots,g_k} \otimes I \,|\psi\rangle \cdot \mathbf{1}[\mathsf{Consistent}_\tau(g,\boldsymbol{y},\boldsymbol{u})] \\
=\ & \mathop{\mathbf{E}}_{\boldsymbol{y}_1,\dots,\boldsymbol{y}_k} \sum_{\tau:|\tau|\geq d+1} \sum_{(g_1,\dots,g_k)\in\overline{\mathsf{Global}_\tau(\boldsymbol{y})}} \langle\psi|\, \widehat{H}^{\boldsymbol{y}_1,\dots,\boldsymbol{y}_k}_{g_1,\dots,g_k} \otimes I \,|\psi\rangle \cdot \mathop{\mathbf{Pr}}_{\boldsymbol{u}}[\mathsf{Consistent}_\tau(g,\boldsymbol{y},\boldsymbol{u})]. \qquad (145)
\end{aligned}
$$

Let us now fix $(y_1,\dots,y_k) \in \mathsf{Distinct}_k$, a type τ such that $|\tau| \geq d+1$, and $(g_1,\dots,g_k) \in \overline{\mathsf{Global}_\tau(y)}$. Suppose without loss of generality that $\tau_1 = \cdots = \tau_{d+1} = 1$, so that $g_1,\dots,g_{d+1} \in \mathcal{P}(m,q,d)$. Then there is a unique polynomial $h^* \in \mathcal{P}(m+1,q,d)$ which interpolates $g_1,\dots,g_{d+1}$. In other words, for all $1 \leq i \leq d+1$,

$$h^*(u, y_i) = g_i(u).$$

In addition, for any u, $(h^*)|_u$ is the unique degree-d polynomial which interpolates $g_1,\dots,g_{d+1}$ along the line in direction u. Thus, if there is a consistent degree-d polynomial interpolating all the g_i's along the line in direction u, then it is $(h^*)|_u$. In math,

$$\mathsf{Consistent}_\tau(g,y,u) = 1 \qquad \text{if and only if} \qquad \forall i \in \tau, g_i(u) = h^*(u, y_i).$$

On the other hand, because $(g_1,\dots,g_k) \in \overline{\mathsf{Global}_\tau(y)}$, there exists an $i^* \in \tau$ such that $g_{i^*} \neq (h^*)|_{y_{i^*}}$. Hence,

$$\mathop{\mathbf{Pr}}_{\boldsymbol{u}}[\mathsf{Consistent}_\tau(g,y,\boldsymbol{u})] \leq \mathop{\mathbf{Pr}}_{\boldsymbol{u}}[g_{i^*}(\boldsymbol{u}) = h^*(\boldsymbol{u}, y_{i^*})] = \mathop{\mathbf{Pr}}_{\boldsymbol{u}}[g_{i^*}(\boldsymbol{u}) = (h^*)|_{y_{i^*}}(\boldsymbol{u})] \leq \frac{md}{q},$$

by Schwartz-Zippel. As a result,

$$
\begin{aligned}
(145) \leq\ & \mathop{\mathbf{E}}_{\boldsymbol{y}_1,\dots,\boldsymbol{y}_k} \sum_{\tau:|\tau|\geq d+1} \sum_{(g_1,\dots,g_k)\in\overline{\mathsf{Global}_\tau(\boldsymbol{y})}} \langle\psi|\, \widehat{H}^{\boldsymbol{y}_1,\dots,\boldsymbol{y}_k}_{g_1,\dots,g_k} \otimes I \,|\psi\rangle \cdot \frac{md}{q} \\
\leq\ & \mathop{\mathbf{E}}_{\boldsymbol{y}_1,\dots,\boldsymbol{y}_k} \sum_{g_1,\dots,g_k} \langle\psi|\, \widehat{H}^{\boldsymbol{y}_1,\dots,\boldsymbol{y}_k}_{g_1,\dots,g_k} \otimes I \,|\psi\rangle \cdot \frac{md}{q} \\
=\ & \frac{md}{q}.
\end{aligned}
$$

This establishes Equation (141). Finally, Proposition 12.4 implies that Equation (141) is (k^2/q)-close to

$$\mathop{\mathbf{E}}_{\boldsymbol{x}_1,\dots,\boldsymbol{x}_k} \sum_{\tau:|\tau|\geq d+1} \sum_{(g_1,\dots,g_k)\in\mathsf{Outcomes}_\tau} \langle\psi|\, \widehat{H}^{\boldsymbol{x}_1,\dots,\boldsymbol{x}_k}_{g_1,\dots,g_k} \otimes I \,|\psi\rangle.$$

In total, this gives an error of

$$2\frac{k^2}{q} + \frac{md}{q} + k \cdot \nu_5 = 2\frac{k^2}{q} + \frac{md}{q} + k \cdot 43km \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right)$$

$$\leq 3\frac{k^2md}{q} + 43k^2m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right)$$

$$\leq 3k^2m(d/q)^{1/32} + 43k^2m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right)$$

$$\leq 46k^2m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right).$$

This completes the proof of the lemma. $\qquad\square$

Lemma 12.21. *Let $x_1, \ldots, x_k \sim \mathbb{F}_q$ be sampled uniformly at random. Then*

$$\mathop{\mathbf{E}}_{x_1,\ldots,x_k} \sum_{\tau:|\tau|\geq d+1} \sum_{(g_1,\ldots,g_k)\in\mathsf{Outcomes}_\tau} \langle\psi| \widehat{H}^{x_1,\ldots,x_k}_{g_1,\ldots,g_k} \otimes I |\psi\rangle \approx_{\nu_8} \sum_{i=d+1}^{k} \binom{k}{i} \langle\psi| G^i(I-G)^{k-i} \otimes I |\psi\rangle,$$

where

$$\nu_8 = 46km \cdot \left(\gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right).$$

Proof. We begin by introducing some notation that we will use throughout the proof. Let $\tau \in \{0,1\}^k$ be a type. Then we define

$$\tau_{<\ell} = (\tau_1, \ldots, \tau_{\ell-1}) \in \{0,1\}^{\ell-1}, \quad \tau_{>\ell} = (\tau_{\ell+1}, \ldots, \tau_k) \in \{0,1\}^{k-\ell},$$

and we define $\tau_{\leq\ell}$ and $\tau_{\geq\ell}$ similarly. In addition, given $(g_1, \ldots, g_k) \in \mathcal{P}^+(m,q,d)^k$, we define

$$g_{<\ell} = (g_1, \ldots, g_{\ell-1}) \in \mathcal{P}^+(m,q,d)^{\ell-1}, \quad g_{>\ell} = (g_{\ell+1}, \ldots, g_k) \in \mathcal{P}^+(m,q,d)^{k-\ell}$$

and we define $g_{\leq\ell}$ and $g_{\geq\ell}$ similarly. Using this notation, we can write

$$\widehat{H}^{x_{\geq\ell}}_{g_{\geq\ell}} = \widehat{H}^{x_\ell,\ldots,x_k}_{g_\ell,\ldots,g_k}.$$

Next, we introduce the notation

$$\widehat{G}^{x_{\geq\ell}}_{g_{\geq\ell}} = \widehat{G}^{x_\ell}_{g_\ell} \cdots \widehat{G}^{x_k}_{g_k}.$$

This satisfies the recurrence relation

$$\widehat{G}^{x_{\geq\ell}}_{g_{\geq\ell}} = \widehat{G}^{x_\ell}_{g_\ell} \cdot \widehat{G}^{x_{>\ell}}_{g_{>\ell}}. \tag{146}$$

Furthermore, we can write

$$\widehat{H}^{x_{\geq\ell}}_{g_{\geq\ell}} = (\widehat{G}^{x_{\geq\ell}}_{g_{\geq\ell}}) \cdot (\widehat{G}^{x_{\geq\ell}}_{g_{\geq\ell}})^\dagger. \tag{147}$$

Finally, we will write O_τ as shorthand for $\mathsf{Outcomes}_\tau$.

To prove the lemma, we will show that for each $1 \leq \ell \leq k$,

$$\mathop{\mathbf{E}}_{x_{\geq\ell}} \sum_{\tau:|\tau|\geq d+1} \sum_{g_{\geq\ell}\in\mathsf{O}_{\tau_{\geq\ell}}} \langle\psi| \widehat{H}^{x_{\geq\ell}}_{g_{\geq\ell}} \otimes (G^{|\tau_{<\ell}|} \cdot (I-G)^{(\ell-1)-|\tau_{<\ell}|}) |\psi\rangle$$

$$\approx_{2\sqrt{2\zeta}+2\sqrt{\nu_4}} \mathop{\mathbf{E}}_{x_{>\ell}} \sum_{\tau:|\tau|\geq d+1} \sum_{g_{>\ell}\in\mathsf{O}_{\tau_{>\ell}}} \langle\psi| \widehat{H}^{x_{>\ell}}_{g_{>\ell}} \otimes (G^{|\tau_{\leq\ell}|} \cdot (I-G)^{\ell-|\tau_{\leq\ell}|}) |\psi\rangle. \tag{148}$$

If we then repeatedly apply Equation (148) for $\ell = 1, \ldots, k$, we derive

$$\underset{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_k}{\mathbf{E}} \sum_{\tau:|\tau|\geq d+1} \sum_{(g_1,\ldots,g_k)\in\mathsf{Outcomes}_\tau} \langle\psi| \widehat{H}_{g_1,\ldots,g_k}^{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_k} \otimes I |\psi\rangle$$

$$= \underset{\boldsymbol{x}_{\geq 1}}{\mathbf{E}} \sum_{\tau:|\tau|\geq d+1} \sum_{g_{\geq 1}\in\mathsf{O}_\tau} \langle\psi| \widehat{H}_{g_{\geq 1}}^{\boldsymbol{x}_{\geq 1}} \otimes I |\psi\rangle$$

$$\approx_{2\sqrt{2\zeta}+2\sqrt{\nu_4}} \underset{\boldsymbol{x}_{\geq 2}}{\mathbf{E}} \sum_{\tau:|\tau|\geq d+1} \sum_{g_{\geq 2}\in\mathsf{O}_{\tau_{\geq 2}}} \langle\psi| \widehat{H}_{g_{\geq 2}}^{\boldsymbol{x}_{\geq 2}} \otimes (G^{|\tau_{\leq 1}|} \cdot (I-G)^{1-|\tau_{\leq 1}|}) |\psi\rangle$$

$$\approx_{2\sqrt{2\zeta}+2\sqrt{\nu_4}} \underset{\boldsymbol{x}_{\geq 3}}{\mathbf{E}} \sum_{\tau:|\tau|\geq d+1} \sum_{g_{\geq 3}\in\mathsf{O}_{\tau_{\geq 3}}} \langle\psi| \widehat{H}_{g_{\geq 3}}^{\boldsymbol{x}_{\geq 3}} \otimes (G^{|\tau_{\leq 2}|} \cdot (I-G)^{2-|\tau_{\leq 2}|}) |\psi\rangle$$

$$\cdots$$

$$\approx_{2\sqrt{2\zeta}+2\sqrt{\nu_4}} \sum_{\tau:|\tau|\geq d+1} \langle\psi| I \otimes (G^{|\tau|} \cdot (I-G)^{k-|\tau|}) |\psi\rangle$$

$$= \sum_{i=d+1}^{k} \binom{k}{i} \langle\psi| I \otimes (G^i (I-G)^{k-i}) |\psi\rangle .$$

In total, using $\sqrt{426} \leq 21$, this gives an error of

$$k \cdot (2\sqrt{2\zeta} + 2\sqrt{\nu_4}) = k \cdot 2\sqrt{2\zeta} + 2 \cdot \sqrt{426 k^2 m \cdot \left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16}\right)}$$

$$\leq 4k \cdot \zeta^{1/32} + 42km \cdot \left(\gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right)$$

$$\leq 46km \cdot \left(\gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right) .$$

This proves the lemma.

We now prove Equation (148). To begin, for each $1 \leq \ell \leq k+1$ and $\tau_{\geq\ell} \in \{0,1\}^{k-\ell+1}$, we define the matrix

$$S_{\tau_{\geq\ell}} = \sum_{\tau_{<\ell}:|\tau|\geq d+1} G^{|\tau_{<\ell}|} \cdot (I-G)^{(\ell-1)-|\tau_{<\ell}|}. \tag{149}$$

Then the statement in Equation (148) can be rewritten as

$$\underset{\boldsymbol{x}_{\geq\ell}}{\mathbf{E}} \sum_{\tau_{\geq\ell}} \sum_{g_{\geq\ell}\in\mathsf{O}_{\tau_{\geq\ell}}} \langle\psi| \widehat{H}_{g_{\geq\ell}}^{\boldsymbol{x}_{\geq\ell}} \otimes S_{\tau_{\geq\ell}} |\psi\rangle$$

$$\approx_{2\sqrt{2\zeta}+2\sqrt{\nu_4}} \underset{\boldsymbol{x}_{>\ell}}{\mathbf{E}} \sum_{\tau_{>\ell}} \sum_{g_{>\ell}\in\mathsf{O}_{\tau_{>\ell}}} \langle\psi| \widehat{H}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \otimes S_{\tau_{>\ell}} |\psi\rangle . \tag{150}$$

To prove this, we will use several facts about $S_{\tau_{\geq\ell}}$. First, S is Hermitian and positive semidefinite. This is because each term in Equation (149) is a product of G and $(I-G)$. These matrices commute with each other, and both are Hermitian and positive semidefinite. Next, S is bounded:

$$S_{\tau_{\geq\ell}} = \sum_{\tau_{<\ell}:|\tau|\geq d+1} G^{|\tau_{<\ell}|} \cdot (I-G)^{(\ell-1)-|\tau_{<\ell}|}$$

$$\leq \sum_{\tau_{<\ell}} G^{|\tau_{<\ell}|} \cdot (I-G)^{(\ell-1)-|\tau_{<\ell}|}$$

$$= (G + (I-G))^{\ell-1}$$

$$= I. \tag{151}$$

In addition, for any $\tau_\ell \in \{0,1\}$,

$$\left(\mathop{\mathbf{E}}_{\boldsymbol{x}_\ell} \sum_{g_\ell \in \mathsf{O}_{\tau_\ell}} \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell}\right) = \begin{cases} G & \text{if } \tau_\ell = 1, \\ (I-G) & \text{if } \tau_\ell = 0, \end{cases}$$

$$= G^{\tau_\ell} \cdot (I-G)^{1-\tau_\ell}. \tag{152}$$

Thus, for any $\tau_{>\ell}$,

$$\sum_{\tau_\ell} S_{\tau_{\geq\ell}} \cdot \left(\mathop{\mathbf{E}}_{\boldsymbol{x}_\ell} \sum_{g_\ell \in \mathsf{O}_{\tau_\ell}} \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell}\right) = \sum_{\tau_\ell} S_{\tau_{\geq\ell}} \cdot (G^{\tau_\ell} \cdot (I-G)^{1-\tau_\ell})$$

$$= \sum_{\tau_\ell} \sum_{\tau_{<\ell}:|\tau|\geq d+1} G^{|\tau_{<\ell}|} \cdot (I-G)^{(\ell-1)-|\tau_{<\ell}|} \cdot (G^{\tau_\ell} \cdot (I-G)^{1-\tau_\ell})$$

$$= \sum_{\tau_\ell} \sum_{\tau_{<\ell}:|\tau|\geq d+1} G^{|\tau_{\leq\ell}|} \cdot (I-G)^{\ell-|\tau_{\leq\ell}|}$$

$$= \sum_{\tau_{\leq\ell}:|\tau|\geq d+1} G^{|\tau_{\leq\ell}|} \cdot (I-G)^{\ell-|\tau_{\leq\ell}|}$$

$$= S_{\tau_{>\ell}}. \tag{153}$$

Finally, for any $\tau_{\geq\ell}$,

$$S_{\tau_{\geq\ell}} \cdot \left(\mathop{\mathbf{E}}_{\boldsymbol{x}_\ell} \sum_{g_\ell \in \mathsf{O}_{\tau_\ell}} \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell}\right) \cdot S_{\tau_{\geq\ell}}$$

$$= S_{\tau_{\geq\ell}} \cdot (G^{\tau_\ell} \cdot (I-G)^{1-\tau_\ell}) \cdot S_{\tau_{\geq\ell}} \qquad \text{(by Equation (152))}$$

$$= \sqrt{G^{\tau_\ell} \cdot (I-G)^{1-\tau_\ell}} \cdot (S_{\tau_{\geq\ell}})^2 \cdot \sqrt{G^{\tau_\ell} \cdot (I-G)^{1-\tau_\ell}}$$

$$\qquad\qquad \text{(because } S_{\tau_{\geq\ell}} \text{ commutes with } G \text{ and } (I-G))$$

$$\leq \sqrt{G^{\tau_\ell} \cdot (I-G)^{1-\tau_\ell}} \cdot I \cdot \sqrt{G^{\tau_\ell} \cdot (I-G)^{1-\tau_\ell}} \qquad \text{(by Equation (151))}$$

$$= G^{\tau_\ell} \cdot (I-G)^{1-\tau_\ell}$$

$$= \left(\mathop{\mathbf{E}}_{\boldsymbol{x}_\ell} \sum_{g_\ell \in \mathsf{O}_{\tau_\ell}} \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell}\right), \tag{154}$$

where the last step uses Equation (152) again. This concludes the set of facts we will need about $S_{\tau_{\geq\ell}}$.

Now we prove Equation (150). To start, we write $\widehat{H}$ as a sandwich of $\widehat{G}$ operators, and move the rightmost $\widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell}$ to the second tensor factor.

$$\mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq\ell}} \sum_{\tau_{\geq\ell}} \sum_{(g_\ell,\ldots,g_k)\in\mathsf{O}_{\tau_{\geq\ell}}} \langle\psi| \, \widehat{H}^{\boldsymbol{x}_\ell,\ldots,\boldsymbol{x}_k}_{g_\ell,\ldots,g_k} \otimes S_{\tau_{\geq\ell}} \, |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq\ell}} \sum_{\tau_{\geq\ell}} \sum_{(g_\ell,\ldots,g_k)\in\mathsf{O}_{\tau_{\geq\ell}}} \langle\psi| \, (\widehat{G}^{\boldsymbol{x}_{\geq\ell}}_{g_{\geq\ell}} \cdot (\widehat{G}^{\boldsymbol{x}_{\geq\ell}}_{g_{\geq\ell}})^\dagger) \otimes S_{\tau_{\geq\ell}} \, |\psi\rangle \qquad \text{(by Equation (147))}$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq\ell}} \sum_{\tau_{\geq\ell}} \sum_{(g_\ell,\ldots,g_k)\in\mathsf{O}_{\tau_{\geq\ell}}} \langle\psi| \, (\widehat{G}^{\boldsymbol{x}_{\geq\ell}}_{g_{\geq\ell}} \cdot (\widehat{G}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}})^\dagger \cdot \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell}) \otimes S_{\tau_{\geq\ell}} \, |\psi\rangle \qquad \text{(by Equation (146))}$$

$$\approx_{\sqrt{2\zeta}} \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq\ell}} \sum_{\tau_{\geq\ell}} \sum_{(g_\ell,\ldots,g_k)\in\mathsf{O}_{\tau_{\geq\ell}}} \langle\psi| \, (\widehat{G}^{\boldsymbol{x}_{\geq\ell}}_{g_{\geq\ell}} \cdot (\widehat{G}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}})^\dagger) \otimes (S_{\tau_{\geq\ell}} \cdot \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell}) \, |\psi\rangle \, . \tag{155}$$

To justify the approximation, we bound the error.

$$\left| \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | \, (\widehat{G}_{g_{\geq \ell}}^{\boldsymbol{x}_{\geq \ell}} \otimes S_{\tau_{\geq \ell}}) \cdot (((\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}})^\dagger \otimes I) \cdot (\widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \otimes I - I \otimes \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell})) \, | \psi \rangle \, \right|$$

$$\leq \sqrt{ \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | \, (\widehat{G}_{g_{\geq \ell}}^{\boldsymbol{x}_{\geq \ell}} \cdot (\widehat{G}_{g_{\geq \ell}}^{\boldsymbol{x}_{\geq \ell}})^\dagger) \otimes (S_{\tau_{\geq \ell}})^2 \, | \psi \rangle }$$

$$\cdot \sqrt{ \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | \, ((\widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \otimes I - I \otimes \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}) \cdot (\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \cdot (\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}})^\dagger \otimes I) \cdot (\widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \otimes I - I \otimes \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell})) \, | \psi \rangle }. $$

$$(156)$$

The expression inside the first square root is equal to

$$\mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | \, \widehat{H}_{g_{\geq \ell}}^{\boldsymbol{x}_{\geq \ell}} \otimes (S_{\tau_{\geq \ell}})^2 \, | \psi \rangle \, ,$$

which is at most 1 because $S_{\tau_{\geq \ell}} \leq I$ and $\widehat{H}$ is a sub-measurement. The expression inside the second square root is equal to

$$\mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | \, ((\widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \otimes I - I \otimes \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}) \cdot (\widehat{H}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \otimes I) \cdot (\widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \otimes I - I \otimes \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell})) \, | \psi \rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{g_\ell, \ldots, g_k} \langle \psi | \, ((\widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \otimes I - I \otimes \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}) \cdot (\widehat{H}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \otimes I) \cdot (\widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \otimes I - I \otimes \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell})) \, | \psi \rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{g_\ell} \langle \psi | \, (\widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \otimes I - I \otimes \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell})^2 \, | \psi \rangle \qquad \text{(because } \widehat{H} \text{ is a sub-measurement)}$$

$$\leq 2\zeta. \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(by Corollary 12.14)}$$

We will now commute the leftmost $\widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}$ in Equation (155) to the right in two stages. In the first stage, we have:

$$(155) = \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | \, (\widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \cdot \widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \cdot (\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}})^\dagger) \otimes (S_{\tau_{\geq \ell}} \cdot \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}) \, | \psi \rangle$$

$$\approx_{\sqrt{\nu_4}} \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | \, (\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \cdot \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \cdot (\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}})^\dagger) \otimes (S_{\tau_{\geq \ell}} \cdot \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}) \, | \psi \rangle \, . \qquad (157)$$

To justify the approximation, we bound the magnitude of the difference.

$$\left| \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | \, ([\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}}, \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}] \otimes I) \cdot ((\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}})^\dagger \otimes (S_{\tau_{\geq \ell}} \cdot \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell})) \, | \psi \rangle \, \right|$$

$$\leq \sqrt{ \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | \, ([\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}}, \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}]) \cdot ([\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}}, \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}])^\dagger \otimes I \, | \psi \rangle }$$

$$\cdot \sqrt{ \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | \, (\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \cdot (\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}})^\dagger) \otimes (\widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \cdot (S_{\tau_{\geq \ell}})^2 \cdot \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}) \, | \psi \rangle }. \quad (158)$$

The quantity inside the first square root is at most

$$\mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{g_\ell,\dots,g_k} \langle\psi| \left([\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}}, \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}]\right) \cdot \left([\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}}, \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}]\right)^\dagger \otimes I \, |\psi\rangle\,,$$

which is at most ν_4 by Lemma 12.15. The quantity inside the second square root is equal to

$$\mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle\psi| \, \widehat{H}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \otimes (\widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \cdot (S_{\tau_{\geq \ell}})^2 \cdot \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}) \, |\psi\rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle\psi| \, \widehat{H}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \otimes (\widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \cdot I \cdot \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}) \, |\psi\rangle \qquad \text{(because } S_{\tau_{\geq \ell}} \leq I)$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle\psi| \, \widehat{H}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \otimes \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \, |\psi\rangle$$

$$\leq 1. \qquad \text{(because } \widehat{G} \text{ and } \widehat{H} \text{ are sub-measurements)}$$

We continue commuting the leftmost $\widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}$ to the right.

$$(157) = \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle\psi| \, (\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \cdot \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \cdot (\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}})^\dagger) \otimes (S_{\tau_{\geq \ell}} \cdot \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}) \, |\psi\rangle$$

$$\approx_{\sqrt{\nu_4}} \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle\psi| \, (\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \cdot (\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}})^\dagger \cdot \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}) \otimes (S_{\tau_{\geq \ell}} \cdot \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}) \, |\psi\rangle\,. \qquad (159)$$

To justify the approximation, we will need to be slightly more clever this time. First, as always, we bound the magnitude of the difference:

$$\left| \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle\psi| \, (\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \otimes (S_{\tau_{\geq \ell}} \cdot \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell})) \cdot ([(\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}})^\dagger, \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}] \otimes I) \, |\psi\rangle \right|$$

$$\leq \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle\psi| \, (\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}}) \cdot (\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}})^\dagger \otimes (S_{\tau_{\geq \ell}} \cdot \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \cdot S_{\tau_{\geq \ell}}) \, |\psi\rangle}$$

$$\cdot \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle\psi| \, ([(\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}})^\dagger, \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}])^\dagger \cdot ([(\widehat{G}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}})^\dagger, \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell}]) \otimes I \, |\psi\rangle}. \quad (160)$$

The term inside the first square root is equal to

$$\mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle\psi| \, \widehat{H}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \otimes (S_{\tau_{\geq \ell}} \cdot \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \cdot S_{\tau_{\geq \ell}}) \, |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}_{> \ell}} \sum_{\tau_{\geq \ell}} \sum_{g_{>\ell} \in \mathsf{O}_{\tau_{>\ell}}} \langle\psi| \, \widehat{H}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \otimes (S_{\tau_{\geq \ell}} \cdot \Big(\mathop{\mathbf{E}}_{\boldsymbol{x}_\ell} \sum_{g_\ell \in \mathsf{O}_{\tau_\ell}} \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \Big) \cdot S_{\tau_{\geq \ell}}) \, |\psi\rangle$$

$$\leq \mathop{\mathbf{E}}_{\boldsymbol{x}_{> \ell}} \sum_{\tau_{\geq \ell}} \sum_{g_{>\ell} \in \mathsf{O}_{\tau_{>\ell}}} \langle\psi| \, \widehat{H}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \otimes \Big(\mathop{\mathbf{E}}_{\boldsymbol{x}_\ell} \sum_{g_\ell \in \mathsf{O}_{\tau_\ell}} \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \Big) \, |\psi\rangle \qquad \text{(by Equation (154))}$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle\psi| \, \widehat{H}_{g_{>\ell}}^{\boldsymbol{x}_{>\ell}} \otimes \widehat{G}_{g_\ell}^{\boldsymbol{x}_\ell} \, |\psi\rangle$$

$$\leq 1. \qquad \text{(because } \widehat{G} \text{ and } \widehat{H} \text{ are sub-measurements)}$$

The term inside the second square root is equal to the term inside the first square root of Equation (158), which we bounded by ν_4. We are now ready for the final step, which is to bring the leftmost $\widehat{G}^{x_\ell}_{g_\ell}$ over to the second tensor factor.

$$(159) = \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq\ell}} \sum_{\tau_{\geq\ell}} \sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}} \langle\psi|\,(\widehat{G}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}} \cdot (\widehat{G}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}})^\dagger \cdot \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell}) \otimes (S_{\tau_{\geq\ell}} \cdot \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell})\,|\psi\rangle$$

$$\approx_{\sqrt{2\zeta}} \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq\ell}} \sum_{\tau_{\geq\ell}} \sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}} \langle\psi|\,(\widehat{G}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}} \cdot (\widehat{G}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}})^\dagger) \otimes (S_{\tau_{\geq\ell}} \cdot \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell} \cdot \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell})\,|\psi\rangle. \tag{161}$$

To justify the approximation, we bound the magnitude of the difference.

$$\left| \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq\ell}} \sum_{\tau_{\geq\ell}} \sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}} \langle\psi|\,(\widehat{G}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}} \otimes (S_{\tau_{\geq\ell}} \cdot \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell})) \cdot (((\widehat{G}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}})^\dagger \otimes I) \cdot (\widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell} \otimes I - I \otimes \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell}))\,|\psi\rangle \right|$$

$$\leq \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq\ell}} \sum_{\tau_{\geq\ell}} \sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}} \langle\psi|\,(\widehat{G}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}} \cdot (\widehat{G}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}})^\dagger) \otimes (S_{\tau_{\geq\ell}} \cdot \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell} \cdot S_{\tau_{\geq\ell}})\,|\psi\rangle}$$

$$\cdot \sqrt{\mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq\ell}} \sum_{\tau_{\geq\ell}} \sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}} \langle\psi|\,((\widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell} \otimes I - I \otimes \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell}) \cdot (\widehat{G}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}} \cdot (\widehat{G}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}})^\dagger \otimes I) \cdot (\widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell} \otimes I - I \otimes \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell})).\,|\psi\rangle}$$

The expression inside the first is equal to the expression inside the first square root in Equation (160), which we bounded by 1. The expression inside the second square root is equal to the expression inside the second square root in Equation (156), which we bounded by 2ζ. We end by noting that

$$(161) = \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq\ell}} \sum_{\tau_{\geq\ell}} \sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}} \langle\psi|\,(\widehat{G}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}} \cdot (\widehat{G}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}})^\dagger) \otimes (S_{\tau_{\geq\ell}} \cdot \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell})\,|\psi\rangle \qquad \text{(because } \widehat{G} \text{ is projective)}$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}_{\geq\ell}} \sum_{\tau_{\geq\ell}} \sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}} \langle\psi|\,\widehat{H}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}} \otimes (S_{\tau_{\geq\ell}} \cdot \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell})\,|\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}_{>\ell}} \sum_{\tau_{>\ell}} \sum_{g_{>\ell}\in\mathsf{O}_{\tau_{>\ell}}} \langle\psi|\,\widehat{H}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}} \otimes \left(\sum_{\tau_\ell} S_{\tau_{\geq\ell}} \cdot \left(\mathop{\mathbf{E}}_{\boldsymbol{x}_\ell} \sum_{g_\ell\in\mathsf{O}_{\tau_\ell}} \widehat{G}^{\boldsymbol{x}_\ell}_{g_\ell} \right) \right) |\psi\rangle$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x}_{>\ell}} \sum_{\tau_{>\ell}} \sum_{g_{>\ell}\in\mathsf{O}_{\tau_{>\ell}}} \langle\psi|\,\widehat{H}^{\boldsymbol{x}_{>\ell}}_{g_{>\ell}} \otimes S_{\tau_{>\ell}}\,|\psi\rangle. \qquad \text{(by Equation (153))}$$

This concludes the proof of Equation (150) and therefore proves the lemma. $\qquad\square$

Lemma 12.22. *Let* $0 < \theta < 1$ *and let* $k, d > 0$ *be integers such that* $k \geq 2d/\theta$*. Define the matrix-valued function* F *by*

$$F(X) = \sum_{r=d+1}^{k} \binom{k}{r} X^r (I - X)^{r-k}.$$

Then for any Hermitian matrix X *such that* $0 \leq X \leq I$ *and* $\langle\psi|\,X \otimes I\,|\psi\rangle \geq 1 - \kappa$*, it holds that*

$$\langle\psi|\,F(X) \otimes I\,|\psi\rangle \geq 1 - \frac{\kappa}{1-\theta} - e^{-\theta^2 k/2}.$$

Proof. Let ρ be the reduced state of $|\psi\rangle$ on one prover's subsystem (since the state is assumed to be symmetric, it does not matter which prover we take). Write the eigendecomposition $X = \sum_i \lambda_i |v_i\rangle\langle v_i|$ of X, where we allow some of the eigenvalues to be 0 so that the set of eigenvectors $\{|v_i\rangle\}$ forms an orthonormal basis of the space. This defines a probability distribution μ over eigenvectors, where eigenvector i occurs with probability $\mu(i) = \langle v_i| \rho |v_i\rangle$. The given condition $\langle\psi| X \otimes I |\psi\rangle = \text{tr}(X\rho) \geq 1 - \kappa$ implies that

$$\mathop{\mathbf{E}}_{i\sim\mu} \lambda_i = \sum_i \lambda_i \cdot \mu(i) = \sum_i \lambda_i \cdot \langle v_i| \rho |v_i\rangle$$
$$= \sum_i \lambda_i \cdot \text{tr}(|v_i\rangle\langle v_i| \cdot \rho)$$
$$= \text{tr}\Big(\sum_i \lambda_i |v_i\rangle\langle v_i| \cdot \rho \Big) = \text{tr}(X\rho) \geq 1 - \kappa.$$

Or, equivalently,

$$\mathop{\mathbf{E}}_{i\sim\mu} (1 - \lambda_i) \leq \kappa.$$

By Markov's inequality, for any $0 < \theta < 1$, we have

$$\mathop{\mathbf{Pr}}_{i\sim\mu}[(1 - \lambda_i) \geq (1 - \theta)] \leq \frac{\mathbf{E}_{i\sim\mu}(1 - \lambda_i)}{1 - \theta} \leq \frac{\kappa}{1 - \theta}.$$

In other words,

$$\mathop{\mathbf{Pr}}_{i\sim\mu}[\lambda_i > \theta] = \mathop{\mathbf{Pr}}_{i\sim\mu}[(1 - \lambda_i) < (1 - \theta)] = 1 - \mathop{\mathbf{Pr}}_{i\sim\mu}[(1 - \lambda_i) \geq (1 - \theta)] \geq 1 - \frac{\kappa}{1 - \theta}. \tag{162}$$

Thus, it holds that with probability at least $1 - \frac{\kappa}{1-\theta}$ over $i \sim \mu$, $\lambda_i > \theta$.

We now evaluate $\langle\psi| F(X) \otimes I |\psi\rangle$. We will essentially do this eigenvalue by eigenvalue. To begin, we consider a hypothetical eigenvalue $d/k \leq p \leq 1$. Observe that $F(p)$ is precisely the probability probability of observing at least $d+1$ successes out of k i.i.d. Bernoulli trials, each of which succeeds with probability p. In other words, it is the probability that $Y := Y_1 + \cdots + Y_k \geq d + 1$, where $Y_1, \ldots, Y_k \sim \text{Bernoulli}(p)$. We can bound this probability by the additive Chernoff bound (see the second additive bound in [Blu11]):

$$\mathbf{Pr}[Y \leq d] = \mathbf{Pr}[Y \leq pk - (pk - d)]$$
$$= \mathbf{Pr}\Big[Y \leq pk - \Big(p - \frac{d}{k}\Big) \cdot k\Big]$$
$$\leq \exp\Big(-2\Big(p - \frac{d}{k}\Big)^2 \cdot k\Big).$$

Thus,

$$F(p) = \mathbf{Pr}[Y \geq d + 1] = 1 - \mathbf{Pr}[Y \leq d] \geq 1 - \exp\Big(-2\Big(p - \frac{d}{k}\Big)^2 \cdot k\Big). \tag{163}$$

Putting the pieces together, we compute $\langle\psi|\,f(X)\otimes I\,|\psi\rangle$:

$$
\begin{aligned}
\langle\psi|\,F(X)\otimes I\,|\psi\rangle &= \mathrm{tr}(F(X)\rho) \\
&= \sum_i F(\lambda_i)\,\langle v_i|\,\rho\,|v_i\rangle \\
&= \mathop{\mathbf{E}}_{\boldsymbol{i}\sim\mu} F(\lambda_{\boldsymbol{i}}) \\
&\geq \mathop{\mathbf{Pr}}_{\boldsymbol{i}\sim\mu}[\lambda_{\boldsymbol{i}}\geq\theta]\cdot F(\theta) \\
&\geq \Big(1-\frac{\kappa}{1-\theta}\Big)\cdot F(\theta) && \text{(by Equation (162))} \\
&\geq \Big(1-\frac{\kappa}{1-\theta}\Big)\cdot\Big(1-\exp\Big(-2\Big(\theta-\frac{d}{k}\Big)^2\cdot k\Big)\Big), && (164)
\end{aligned}
$$

where the last step uses Equation (163). Next, we claim that if $a,b,c\geq0$ satisfy $a\geq(1-b)\cdot(1-c)$, then $a\geq1-b-c$. This is because if either b or c is at least 1, then the conclusion is trivially true, and if both are less than 1, then

$$
(1-b)\cdot(1-c)=1\cdot(1-c)-b\cdot(1-c)\geq1\cdot(1-c)-b\cdot1=1-c-b.
$$

Since $\langle\psi|\,F(X)\otimes I\,|\psi\rangle$ is manifestly positive, we can apply this to Equation (164), yielding

$$
\langle\psi|\,F(X)\otimes I\,|\psi\rangle\geq1-\frac{\kappa}{1-\theta}-\exp\Big(-2\Big(\theta-\frac{d}{k}\Big)^2\cdot k\Big).
$$

Finally, we note that because $k\geq2d/\theta$, we have $\theta/2\geq d/k$. This implies that $\theta-d/k\geq\theta-\theta/2=\theta/2$, and so $(\theta-d/k)^2\geq(\theta/2)^2=\theta^2/4$. As a result,

$$
\exp\Big(2\Big(\theta-\frac{d}{k}\Big)^2\cdot k\Big)\geq\exp\Big(\theta^2 k/2\Big).
$$

Equivalently,

$$
\exp\Big(-2\Big(\theta-\frac{d}{k}\Big)^2\cdot k\Big)\leq\exp\Big(-\theta^2 k/2\Big).
$$

Thus, we conclude

$$
\langle\psi|\,F(X)\otimes I\,|\psi\rangle\geq1-\frac{\kappa}{1-\theta}-\exp\Big(-\theta^2 k/2\Big). \qquad\square
$$

Corollary 12.23 (Completeness of H; Proof of Item 2 of Lemma 12.2)**.** *Let $k\geq400md$. Then*

$$
\langle\psi|\,H\otimes I\,|\psi\rangle\geq1-\kappa\cdot\Big(1+\frac{1}{100m}\Big)-\nu-e^{-k/(80000m^2)}.
$$

Proof. We begin by approximating the completeness as follows.

$$
\begin{aligned}
\langle\psi|\,H\otimes I\,|\psi\rangle &\approx_{\nu_7} \mathop{\mathbf{E}}_{\boldsymbol{x}_1,\dots,\boldsymbol{x}_k}\sum_{\tau:|\tau|\geq d+1}\sum_{(g_1,\dots,g_k)\in\mathsf{Outcomes}_\tau}\langle\psi|\,\widehat{H}^{\boldsymbol{x}_1,\dots,\boldsymbol{x}_k}_{g_1,\dots,g_k}\otimes I\,|\psi\rangle && \text{(by Lemma 12.20)} \\
&\approx_{\nu_8} \sum_{i=d+1}^{k}\binom{k}{i}\langle\psi|\,G^i(I-G)^{k-i}\otimes I\,|\psi\rangle && \text{(by Lemma 12.21)} \\
&= \langle\psi|\,F(G)\otimes I\,|\psi\rangle.
\end{aligned}
$$

We can bound the error incurred here by

$$\nu_7 + \nu_8 = 46k^2 m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32} \right) + 46km \cdot \left(\gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32} \right)$$
$$\leq 100k^2 m \cdot \left(\epsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32} \right)$$
$$= \nu.$$

Let $\theta = 1/(200m)$. Note that

$$\frac{1}{1-\theta} = \frac{1}{1 - 1/(200m)} = \frac{200m}{200m - 1} = 1 + \frac{1}{200m - 1} \leq 1 + \frac{1}{100m}.$$

Then $k \geq 2d/\theta$ and, therefore, $k \geq d+1$. As a result, k and θ satisfy the hypothesis of Lemma 12.22, namely that $k \geq \max\{d + 1, 2d/\theta\}$. Thus, Lemma 12.22 implies that

$$\langle \psi | \, F(G) \otimes I \, | \psi \rangle \geq 1 - \frac{\kappa}{1 - \theta} - e^{-\theta^2 k/2}$$
$$= 1 - \frac{\kappa}{1 - \theta} - e^{-k/(80000m^2)}$$
$$\geq 1 - \kappa \cdot \left(1 + \frac{1}{100m} \right) - e^{-k/(80000m^2)}.$$

In total, we have

$$\langle \psi | \, H \otimes I \, | \psi \rangle \geq 1 - \kappa \cdot \left(1 + \frac{1}{100m} \right) - \nu - e^{-k/(80000m^2)}.$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

References

[AHO97] Farid Alizadeh, Jean-Pierre Haeberly, and Michael Overton. Complementarity and nondegeneracy in semidefinite programming. *Mathematical programming*, 77(1):111–128, 1997. 9.1

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, 1998. (document), 1

[Ara02] PK Aravind. A simple demonstration of Bell's theorem involving two observers and no probabilities or inequalities. *arXiv preprint quant-ph/0206070*, 2002. 2

[AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: a new characterization of NP. *Journal of the ACM*, 45(1):70–122, 1998. (document), 1, 1

[BFL91] László Babai, Lance Fortnow, and Carsten Lund. Non-deterministic exponential time has two-prover interactive protocols. *Computational complexity*, 1(1):3–40, 1991. (document), 1, 1, 2, 2

[BLR93] Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. *Journal of computer and system sciences*, 47(3):549–595, 1993. 4

[Blu11]   Avrim Blum. Lecture 5 from 15-859(M): Randomized Algorithms. Found at http://www.cs.cmu.edu/~avrim/Randalgs11/lectures/lect0124.pdf, 2011. 12.6

[BV04]   Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. 2004. 9.1

[CMS17]   Alessandro Chiesa, Peter Manohar, and Igor Shinkar. On axis-parallel tests for tensor product codes. In *Proceedings of the 21st Annual International Workshop on Randomization and Computation*, 2017. 3

[FHS94]   Katalin Friedl, Zsolt Hatsagi, and Alexander Shen. Low-degree tests. In *Proceedings of the 5rd Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 57–64, 1994. 1

[IV12]   Tsuyoshi Ito and Thomas Vidick. A multi-prover interactive proof for NEXP sound against entangled provers. In *Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science*, pages 243–252, 2012. 1.1, 1.1, 4, 2

[JNV+20]   Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. MIP* = RE. Technical report, arXiv:2001.04383, 2020. (document), 1.1, 1.1, 1.1, 2, 4.11

[KS08]   Tali Kaufman and Madhu Sudan. Algebraic property testing: the role of invariance. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, pages 403–412, 2008. 3

[KV11]   Julia Kempe and Thomas Vidick. Parallel repetition of entangled games. In *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing*, pages 353–362, 2011. 5

[Mer90]   David Mermin. Simple unified form for the major no-hidden-variables theorems. *Physical Review Letters*, 65(27):3373, 1990. 2

[NV18a]   Anand Natarajan and Thomas Vidick. Low-degree testing for quantum states, and a quantum entangled games PCP. In *Proceedings of the 59th Annual IEEE Symposium on Foundations of Computer Science*, 2018. 1, 1.1, 1.1

[NV18b]   Anand Natarajan and Thomas Vidick. Two-player entangled games are NP-hard. In *Proceedings of the 33rd Annual IEEE Conference on Computational Complexity*, 2018. 1.1, 1, 1.1, 1.1, 1.2, 2

[NW19]   Anand Natarajan and John Wright. NEEXP ⊆ MIP*. In *Proceedings of the 60th Annual IEEE Symposium on Foundations of Computer Science*, pages 510–518, 2019. 1.1, 1.1, 4.11, 4.4, 4.4.1, 4.4.3, 4.4.4, 5, 5

[Per90]   Asher Peres. Incompatible results of quantum measurements. *Physics Letters A*, 151(3-4):107–108, 1990. 2

[PS94]   Alexander Polishchuk and Daniel Spielman. Nearly-linear size holographic proofs. In *Proceedings of the 26th Annual ACM Symposium on Theory of Computing*, pages 194–203, 1994. 1, 1.2

[RS97]   Ran Raz and Shmuel Safra. A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In *Proceedings of the 29th Annual ACM Symposium on Theory of Computing*, pages 475–484, 1997. 1.1

[Sch80]   Jacob Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM*, 27(4):701–717, 1980. 4.6

[Sud11]   Madhu Sudan. Guest column: testing linear properties: some general theme. *ACM SIGACT News*, 42(1):59–80, 2011. 3

[Vid11]   Thomas Vidick. *The complexity of entangled games.* PhD thesis, University of California, Berkeley, 2011. 4

[Vid16]   Thomas Vidick. Three-player entangled XOR games are NP-hard to approximate. *SIAM Journal on Computing*, 45(3):1007–1063, 2016. (document), 1.1, 1.1, 1.1, 1.2, 2

[Zip79]   Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Proceedings of the 2nd International Symposium on Symbolic and Algebraic Manipulation*, pages 216–226, 1979. 4.6