Blueprint for arXiv:2009.12982

Quantum Soundness of the Classical Low Individual Degree Test

MIPStarRE Project

August 24, 2026

Last updated: August 24, 2026

Contents

Chapter 1

Overview

This blueprint records the formalization plan for the low individual degree test paper of Ji, Natarajan, Vidick, Wright, and Yuen [JNV+20]. The paper is available as arXiv:2009.12982.

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

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

Here a *denotes the prover's point answer.*

Remark 1.2. The paper theorem is an external classical soundness input and is not used by the current low individual degree test formalization route.

Proof. See Raz–Safra [RS97]. □

Theorem 1.3 (Polishchuk–Spielman)**.** *Suppose Provers* A *and* B *pass the low individual degree test with probability* $1 - \varepsilon$*. Then there exists a polynomial* $g\colon \mathbb{F}_q^m \to \mathbb{F}_q$ *with individual degree* d *such that*

$$\mathbb{P}_{u\sim\mathbb{F}_q^m}[g(u) = a] \geq 1 - \mathrm{poly}(m) \cdot \big(\mathrm{poly}(\varepsilon) + \mathrm{poly}(d/q)\big).$$

Here a *denotes the prover's point answer.*

Remark 1.4. The paper theorem is an external classical soundness input. The nearby Lean corollary below records the specialized conditional interface currently present in the repository; it is not used to mark this source theorem as formalized.

Proof. See Polishchuk–Spielman [PS94]. □

Proposition 1.5 (Conditional Polishchuk–Spielman corollary)**.** *The Lean theorem* `MIPStarRE.LDT.Test.classicalT` *is a conditional corollary of the quoted Polishchuk–Spielman theorem. It assumes the modeled deterministic low individual degree test pass condition* `MIPStarRE.LDT.Test.TwoProverClassicalLIDPassCondition` *and a specialized external hypothesis* `MIPStarRE.LDT.Test.PolishchukSpielmanClassicalSoundnessStatement` *at the caller's chosen slack bound. Under these hypotheses it returns a repository polynomial whose evaluations agree with the point answers up to that slack.*

Proof. The Lean proof applies the specialized external Polishchuk–Spielman hypothesis to the modeled low individual degree test pass condition. □

Remark 1.6 (A degree-$(d+1)$ strategy that passes with probability $1-\frac{1}{m}$). Fix the polynomial $h(x_1,\dots,x_m)=x_1^{d+1}$. Consider the following classical strategy for the degree-d low individual degree test. On a point query $u\in\mathbb{F}_q^m$, Prover A answers $a=h(u)$. On an axis-parallel line query $\ell=\{u+xe_i : x\in\mathbb{F}_q\}$, Prover B returns the restriction $h|_\ell$ when $i\neq 1$, since h is then constant along ℓ, and returns the zero polynomial when $i=1$.

The verifier certainly accepts whenever $i\neq 1$, which happens with probability $1-\frac{1}{m}$. Thus the strategy passes the degree-d test with probability at least $1-\varepsilon$ for $\varepsilon=\frac{1}{m}$. However, Prover A is answering according to a polynomial whose individual degree in the first variable is $d+1$. By Schwartz–Zippel, any polynomial g of individual degree at most d agrees with h on at most a $\frac{d+1}{q}$ fraction of all inputs, so

$$\mathbb{P}_{u\sim\mathbb{F}_q^m}[g(u)=a]\leq 1-m\varepsilon+\frac{d+1}{q}.$$

This shows that the m-dependent slack in Theorems 1.3 and 1.7 is unavoidable up to polynomial factors.

Theorem 1.7 (Main theorem, informal)**.** *If $d>0$ and a two-prover projective strategy passes the (m,q,d) low individual degree test with probability $1-\varepsilon$, then there are global polynomial measurements whose evaluations agree with the point answers except with error* $\mathrm{poly}(m)(\mathrm{poly}(\varepsilon)+\mathrm{poly}(d/q))$.

Proof. This is the informal form of Theorem 2.14. ☐

Chapter 2

The low individual degree test

2.1 The game and its strategies

Definition 2.1 (Roles)**.** A role is an element of $\{\mathrm{A}, \mathrm{B}\}$. For a role $r \in \{\mathrm{A}, \mathrm{B}\}$, write $\bar{r}$ for the other element of $\{\mathrm{A}, \mathrm{B}\}$.

Definition 2.2 (Low individual degree test)**.** Fix integers $m, d \geq 0$ and a prime power q. The (m, q, d) low individual degree test has three equally likely subtests. In the axis-parallel lines test, choose a uniformly random role $r \in \{\mathrm{A}, \mathrm{B}\}$, a uniformly random point $u \in \mathbb{F}_q^m$, and a uniformly random coordinate $i \in \{1, \dots, m\}$. Let $\ell = \{u + te_i : t \in \mathbb{F}_q\}$. Player r receives ℓ and returns a degree-d univariate polynomial $f\colon \ell \to \mathbb{F}_q$; Player $\bar{r}$ receives u and returns a field element $a \in \mathbb{F}_q$. The verifier accepts when $f(u) = a$. In the self-consistency test, both provers receive the same uniformly random point $u \in \mathbb{F}_q^m$ and must return the same field element. In the diagonal lines test, choose a uniformly random role $r \in \{\mathrm{A}, \mathrm{B}\}$, a uniformly random point $u \in \mathbb{F}_q^m$, a uniformly random index $i \in \{1, \dots, m\}$, and a uniformly random direction vector $v \in \mathbb{F}_q^m$ whose last $m - i$ coordinates are 0. Let $\ell = \{u + tv : t \in \mathbb{F}_q\}$. Player r receives ℓ and returns a degree-md univariate polynomial $f\colon \ell \to \mathbb{F}_q$; Player $\bar{r}$ receives u and returns a field element $a \in \mathbb{F}_q$. The verifier accepts when $f(u) = a$.

Lemma 2.3 (Branch-average form of the classical test)**.** *For every classical strategy, the acceptance probability of the low individual degree test is*

$$\frac{1}{3}\left(\frac{P_{\text{axis}}^{\text{line,point}} + P_{\text{axis}}^{\text{point,line}}}{2} + P_{\text{self}} + \frac{P_{\text{diag}}^{\text{line,point}} + P_{\text{diag}}^{\text{point,line}}}{2}\right),$$

where the five terms are the acceptance probabilities of the two axis-parallel role orderings, the self-consistency branch, and the two diagonal role orderings.

Proof. Expand the definition of the total acceptance probability as the average of the three subtests. The axis-parallel and diagonal branches each first choose one of the two role orderings uniformly, so their branch acceptance probabilities are the corresponding averages displayed above. □

Definition 2.4 (Symmetric projective strategy)**.** A symmetric projective strategy for the (m, q, d) low individual degree test consists of a permutation-invariant bipartite state $|\psi\rangle \in \mathcal{H} \otimes \mathcal{H}$, a projective point measurement $A^u = \{A_a^u\}_{a \in \mathbb{F}_q}$ for each $u \in \mathbb{F}_q^m$, a projective axis-parallel line measurement $B^\ell = \{B_f^\ell\}$ for each axis-parallel line ℓ, and a projective diagonal-line measurement $L^\ell = \{L_f^\ell\}$ for each line ℓ.

Definition 2.5 (General projective strategy)**.** A general projective strategy for the (m, q, d) low individual degree test consists of a bipartite state $|\psi\rangle \in \mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}$ together with point, axis-parallel line, and diagonal-line projective measurements for each prover separately. We write these measurements as $A^{\mathrm{A}}, B^{\mathrm{A}}, L^{\mathrm{A}}$ on the first prover and $A^{\mathrm{B}}, B^{\mathrm{B}}, L^{\mathrm{B}}$ on the second.

Lemma 2.6 (Bipartite strategy block identities)**.** *The block operators obtained from a bipartite projective strategy are positive semidefinite, and the role-indexed blocks multiply according to the Kronecker delta on roles.*

Proof. The assertions are the block-diagonal matrix calculations for the direct-sum and role-register operators. Positivity follows by writing each positive operator as C^*C and placing the corresponding square root in the same block decomposition. $\square$

Lemma 2.7 (Heterogeneous role-register unsymmetrization)**.** *Let ψ be a normalized bipartite state on $\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}$, let M^{A} and M^{B} be projective measurements on the two prover spaces with a common outcome type, and let G be an arbitrary measurement on the heterogeneous role-register space*

$$\{\mathrm{A}, \mathrm{B}\} \times (\mathcal{H}_{\mathrm{A}} \oplus \mathcal{H}_{\mathrm{B}}).$$

After applying the direct-sum role-register symmetrization, the consistency defect between the symmetrized point measurement and G is the average of the two defects obtained from the occupied principal blocks of G. In particular, each extracted defect is at most twice the symmetrized defect.

Proof. The proof is a trace-compression calculation. The AB-supported component of the symmetrized density pairs with an arbitrary role-register observable only through the principal block indexed by $(\mathrm{A}, \mathrm{inl})$ on the left and $(\mathrm{B}, \mathrm{inr})$ on the right, and similarly for the BA component. Summing over outcomes gives the average-of-two-defects identity, and nonnegativity of consistency defects gives the two factor-two consequences. $\square$

Lemma 2.8 (Heterogeneous role-register symmetrization)**.** *Let*

$$(\psi, A^{\mathrm{A}}, B^{\mathrm{A}}, L^{\mathrm{A}}, A^{\mathrm{B}}, B^{\mathrm{B}}, L^{\mathrm{B}})$$

be a two-space projective strategy passing the low individual degree test with error at most ε. The heterogeneous role-register construction on

$$\{\mathrm{A}, \mathrm{B}\} \times (\mathcal{H}_{\mathrm{A}} \oplus \mathcal{H}_{\mathrm{B}})$$

produces a symmetric strategy that is $(3\varepsilon, 3\varepsilon, 3\varepsilon)$-good.

Proof. The role-register state is the direct-sum version of the symmetrized state in the paper proof. The axis-parallel and diagonal branches are exactly the role averages of the original strategy, and the self-consistency branch is the original point-agreement branch. Since the test chooses among its three branches with equal probability, each branch is bounded by 3ε. $\square$

Remark 2.9. Throughout this work, the term strategy refers to a projective strategy, and the term symmetric strategy refers to a symmetric projective strategy.

Definition 2.10 (Good strategy)**.** A strategy is $(\varepsilon, \delta, \gamma)$-good if it passes the axis-parallel lines test with probability at least $1 - \varepsilon$, the self-consistency test with probability at least $1 - \delta$, and the diagonal lines test with probability at least $1 - \gamma$.

Remark 2.11 (Good-strategy characterization)**.** Using notation which will be introduced in Chapter 3 below, a symmetric strategy is $(\varepsilon, \delta, \gamma)$-good if and only if it satisfies the following three conditions. For ℓ and u as in the axis-parallel lines test,

$$A_a^u \otimes I \simeq_\varepsilon I \otimes B^\ell_{[f(u)=a]}, \quad \text{and} \quad A_a^u \otimes I \simeq_\delta I \otimes A_a^u.$$

And for ℓ and u as in the diagonal lines test,

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

Definition 2.12 (Last-direction line notation)**.** For $u \in \mathbb{F}_q^m$, the last-direction axis-parallel line in $\mathbb{F}_q^{m+1}$ is

$$\ell_u = \{(u, x) \mid x \in \mathbb{F}_q\}.$$

If (ψ, A, B, L) is a symmetric strategy for the $(m+1, q, d)$ low individual degree test, we write B_f^u for the operator $B_f^{\ell_u}$. For a function $f \colon \ell_u \to \mathbb{F}_q$, we also write $f(x)$ for $f(u, x)$.

Definition 2.13 (Restricted diagonal lines test)**.** Consider the (m, q, d) low individual degree test. For $j \in \{1, \dots, m\}$, the j-restricted diagonal lines test is the diagonal lines test conditioned on $i = j$. In particular, the m-restricted diagonal lines test is the diagonal lines test in which the sampled line is a uniformly random line in $\mathbb{F}_q^m$.

2.2 Formal soundness target

Theorem 2.14 (Quantum soundness of the low individual degree test)**.** *Let $(\psi, A^{\mathrm{A}}, B^{\mathrm{A}}, L^{\mathrm{A}}, A^{\mathrm{B}}, B^{\mathrm{B}}, L^{\mathrm{B}})$ be a projective strategy that passes the (m, q, d) low individual degree test with probability at least $1 - \varepsilon$. Let $k > 0$ be an integer with $k \geq 400md$, and set*

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

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

1. *on average over $u \sim \mathbb{F}_q^m$,*

$$A_a^{\mathrm{A},u} \otimes I \simeq_\nu I \otimes G^{\mathrm{B}}_{[g(u)=a]}, \qquad I \otimes A_a^{\mathrm{B},u} \simeq_\nu G^{\mathrm{A}}_{[g(u)=a]} \otimes I;$$

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

Proof. The saturated-error branch is formalized by `mainFormal_trivial_witness`. The remaining branch is the small-error source-boundary proposition Proposition 2.16. The factor 400 and the nonzero sampling condition are the corrected source boundaries recorded in [con26b] and [con26c]. □

Proposition 2.15 (Source-boundary reduction for the final theorem)**.** *This proposition records the saturated-error reduction for the corrected two-space theorem Theorem 2.14. It proves the saturated-error branch by the trivial consistency bound and reduces the remaining non-vacuous branch to Proposition 2.16. It is not an additional hypothesis of the paper theorem.*

Proof. The saturated-error branch is already proved. The small-error branch is Proposition 2.16. □

Proposition 2.16 (Small-error source-boundary conclusion for the final theorem)**.** *This is the non-vacuous branch of the corrected two-space theorem. It proves the conclusions of Theorem 2.14 under $k \geq 400md$, $k > 0$, and $\nu < 1$. It is not a source assumption. Its proof uses the two-space role-register reduction and the scalar absorption available after excluding the zero-sampling boundary.*

Proof. The proof applies the checked two-space role-register construction and then absorbs its explicit errors into ν at the nonzero sampling boundary. □

Proposition 2.17 (Two-space role-register reduction for the source theorem)**.** *The printed final theorem starts from a projective strategy on two possibly different local Hilbert spaces. The source proof uses the role-register symmetrization and then unsymmetrizes the polynomial measurement. Lean formalizes this reduction for the general two-space source strategy. The trace-level factor-two unsymmetrization estimate for arbitrary heterogeneous role-register measurements is formalized in Lemma 2.7. The heterogeneous role-register symmetrization may be fed into the source-shaped main-induction theorem under the corrected large-k hypothesis $k \geq 400md$, and the resulting role-register consistency estimate can be unsymmetrized into the two point-consistency estimates on the original two-space strategy with the expected factor 2.*

The remaining steps of the source-boundary passage are also checked. The point-agreement branch is available for the paper-faithful two-space strategy, and the Schwartz–Zippel Step 5 theorem has been generalized from the same-space carrier to a bipartite state on $H_A \otimes H_B$. The heterogeneous triangle/SDD comparison theorem gives the complete-measurement full-polynomial consistency statement at the end of the paper's Step 5 calculation. The two heterogeneous orthonormalization applications produce projective submeasurements on the two local Hilbert spaces, and the completion step widens the resulting left- and right-factor state-dependent-distance estimates to the orthonormalize-and-complete error appearing in the paper. The repaired polynomial line-169 consistency relations use the Cauchy–Schwarz loss from the pre-completion orthonormalization estimates, not a new hypothesis. The final point-evaluation triangle postprocesses these relations and the completed Q_A, Q_B projective consistency estimate by evaluation at a point, combines them by the heterogeneous triangle inequality, and absorbs the explicit pre-absorption errors into the final $\nu =$ `mainFormalError` bound under the nonzero scalar-cascade boundary $0 < k$. This is the small-error branch used in the corrected two-space source theorem, where $k > 0$ is part of the theorem statement.

Proof. The proof is the linked heterogeneous role-register chain, ending in `mainFormalConclusion_ofRoleRegisterSc`
It combines the source-shaped main-induction theorem, the heterogeneous unsymmetrization estimates, the step making measurements projective on both local Hilbert spaces, the completion and line-169 consistency steps, and the final scalar absorption under the nonzero sampling boundary. □

Remark 2.18 (Final-theorem zero-sampling boundary). The paper states Theorem 2.14 for $k \geq md$, while the corrected formal route uses the confirmed large-k condition $k \geq 400md$ and a nonzero k boundary for the scalar cascade. The factor 400 is now treated as a statement correction, following Theorem 10.14. The other corrected final-theorem boundary is the zero-sampling case allowed when $d = 0$, since $400md \leq k$ then permits $k = 0$. At this excluded boundary the displayed final error itself collapses to 0; the formal calculation of this equality is linked in Lean, and the obstruction is documented in [con26c].

Remark 2.19 (k-bound boundary for the public theorem). The paper statement uses the weaker hypothesis $k \geq md$, but the proof invokes the Section 6 pasting theorem with side condition $k \geq 400\,md$. This side condition is not implied by $k \geq md$. The Lean theorem therefore records the stronger large-k assumption required by the pasting step, together with the scalar-cascade boundary $k > 0$, rather than hiding the side-condition gap. The helper `mainFormal_trivial_witness` remains only for already-saturated branches where $\nu \geq 1$ makes the three $\simeq_\nu$ conclusions vacuous; it is not a replacement for the zero-sampling boundary $k = 0$ when $md = 0$. The source-boundary reduction proposition proves the saturated-error branch and delegates the non-vacuous branch to the checked small-error proposition; neither declaration is an extra hypothesis of Theorem 2.14.

Chapter 3

Preliminaries

3.1 Polynomials and measurements

For complex numbers $\alpha, \beta \in \mathbb{C}$, write $\alpha \approx_\varepsilon \beta$ when $|\alpha - \beta| \leq \varepsilon$. In particular,

$$\text{if } \alpha \approx_\varepsilon \beta \text{ and } \beta \approx_\delta \gamma, \text{ then } \alpha \approx_{\varepsilon+\delta} \gamma. \tag{3.1}$$

Fix a prime power $q = p^t$. Write $\omega = e^{2\pi i/p}$, and let $\mathrm{tr} : \mathbb{F}_q \to \mathbb{F}_p$ denote the finite-field trace.

Definition 3.1 (Finite field trace)**.** The finite-field trace is the map $\mathrm{tr} : \mathbb{F}_q \to \mathbb{F}_p$ defined by

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

Lemma 3.2. *Let $a \in \mathbb{F}_q$. Then*

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

Proof. If $a = 0$, then $\mathrm{tr}[x \cdot a] = 0$ for every x, so the average is 1. If $a \neq 0$, choose $y \in \mathbb{F}_q$ such that $\mathrm{tr}[a \cdot y] \neq 0$, and set

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

Translation invariance of the uniform distribution gives

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

Since $\omega^{\mathrm{tr}[y\cdot a]} \neq 1$, it follows that $C = 0$. □

Lemma 3.3. *Let $v \in \mathbb{F}_q^m$. Then*

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

Proof. By linearity of the trace and independence of the coordinates,

$$\mathbb{E}_{u\sim\mathbb{F}_q^m} \omega^{\mathrm{tr}[u\cdot v]} = \prod_{i=1}^{m} \mathbb{E}_{u_i\sim\mathbb{F}_q} \omega^{\mathrm{tr}[u_i\cdot v_i]}.$$

Proposition 3.2 shows that the i-th factor is 1 when $v_i = 0$ and 0 otherwise, so the product is 1 exactly when $v = 0$. □

Definition 3.4 (Low individual degree polynomials). Let $\mathcal{P}(m, q, d)$ be the set of polynomials $g \in \mathbb{F}_q[x_1, \ldots, x_m]$ whose degree in each variable is at most d, viewed as functions $\mathbb{F}_q^m \to \mathbb{F}_q$ via evaluation. (Over finite fields, distinct low-degree polynomials can induce the same function; we identify elements of $\mathcal{P}(m, q, d)$ with their polynomial representatives, not their functional equivalence classes.)

Remark 3.5 (Individual-degree convention). Here, "individual degree d" means degree at most d in each coordinate. In particular,

$$\mathcal{P}(m, q, d) \subseteq \mathcal{P}(m, q, d + 1).$$

Lemma 3.6 (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*

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

Proof. Apply the Schwartz–Zippel lemma to the nonzero polynomial $g - h$, which has total degree at most d. ☐

Lemma 3.7 (Schwartz–Zippel for individual degree). *If $g, h \in \mathcal{P}(m, q, d)$ are distinct, then*

$$\mathbb{P}_{u \sim \mathbb{F}_q^m}[g(u) = h(u)] \leq \frac{md}{q}.$$

Proof. The difference $g - h$ is a nonzero polynomial of total degree at most md. Apply Lemma 3.6. ☐

Lemma 3.8 (Polynomial agreement bound on coded points). *Let g, g' : Polynomial params be two distinct full polynomial outcomes with individual degrees at most d. Then*

$$\mathbb{E}_{u \sim \text{Point params}}[\mathbf{1}[g(u) = g'(u)]] \leq \frac{md}{q}.$$

This packages the md/q loss term used in the `mainFormal` self-consistency cascade (`inductive_step.tex`, lines 119–133) and in `comMain` (`commutativity-G.tex`).

Proof. Transport along the equivalence between coded $\mathbb{F}_q$ points and the underlying scalar function space, then apply Lemma 3.7. ☐

Lemma 3.9 (Axis-line polynomial agreement bound). *Let f, h : AxisLinePolynomial params be two line-polynomial outcomes with distinct underlying polynomials. Then*

$$\mathbb{E}_{t \sim \mathbb{F}_q}[\mathbf{1}[f(t) = h(t)]] \leq \frac{md}{q}.$$

The native univariate root-counting bound is d/q; the statement deliberately pads this to the paper's ambient md/q loss for Lemma 6.1.

Proof. Reindex coded line parameters to the scalar field, count the roots of the nonzero univariate polynomial $f - h$, and use $m \geq 1$ to pad d/q to md/q. ☐

Definition 3.10 (Measurements and submeasurements). Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{A}$ be a set of outcomes. A submeasurement on $\mathcal{A}$ is a family $A = \{A_a\}_{a \in \mathcal{A}}$ of Hermitian positive semidefinite operators on $\mathcal{H}$ such that $\sum_a A_a \leq I$. It is a measurement when $\sum_a A_a = I$, and it is projective when each A_a is an idempotent projection. (In the Lean formalization, outcome sets are assumed finite throughout.)

Definition 3.11 (Low-degree polynomial measurements)**.** We write $\mathrm{PolySub}(m,q,d)$ for the submeasurements indexed by $\mathcal{P}(m,q,d)$, and $\mathrm{PolyMeas}(m,q,d)$ for the corresponding measurements.

Definition 3.12 (Post-processing)**.** Let $A = \{A_a\}_{a\in\mathcal{A}}$ be a family of operators and let $f\colon \mathcal{A}\to\mathcal{B}$. The post-processed family $A_{[f(a)=b]}$ is defined by

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

Proposition 3.13 (Post-processing preserves measurements)**.** *Let $A = \{A_a\}_{a\in\mathcal{A}}$ be a family of operators and let $f\colon \mathcal{A}\to\mathcal{B}$. Then*

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

Consequently, if $\{A_a\}$ is a submeasurement, respectively a measurement, then $\{A_{[f(a)=b]}\}$ is again a submeasurement, respectively a measurement.

Proof. Each $a\in\mathcal{A}$ contributes exactly once to the right-hand side, namely in the summand indexed by $b = f(a)$. □

Remark 3.14 (Post-processing notation)**.** In the notation $A_{[f(a)=b]}$, the measurement outcome is the variable to which the function is applied. For example, if $G = \{G_g\}\in \mathrm{PolySub}(m,q,d)$ and $u\in\mathbb{F}_q^m$, then

$$G_{[g(u)=a]} = \sum_{g\,:\,g(u)=a} G_g.$$

Here g is the measurement outcome, and u is the evaluation point.

Definition 3.15 (Completion of a submeasurement)**.** Let $A = \{A_a^x\}_{a\in\mathcal{A}}$ be a submeasurement. Its completion is the measurement $\widehat{A}$ with outcome set $\widehat{\mathcal{A}} = \mathcal{A}\cup\{\perp\}$ given by $\widehat{A}_a^x = A_a^x$ for $a\in\mathcal{A}$ and $\widehat{A}_\perp^x = I - \sum_a A_a^x$.

3.2 Consistency and state-dependent distance

Definition 3.16 (Consistency)**.** Let $|\psi\rangle \in \mathcal{H}_\mathrm{A}\otimes\mathcal{H}_\mathrm{B}$. Let $A = \{A_a^x\}$ and $B = \{B_a^x\}$ be submeasurements with the same answer set, and let $\mathcal{D}$ be a distribution on the question set. We write

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

when

$$\mathbb{E}_{x\sim\mathcal{D}}\sum_{a\neq b}\langle\psi|A_a^x\otimes B_b^x|\psi\rangle \le \delta. \tag{3.2}$$

Lemma 3.17 (Characterization of a good symmetric strategy)**.** *A symmetric projective strategy (ψ,A,B,L) is $(\varepsilon,\delta,\gamma)$-good if and only if the following hold on the relevant test distributions:*

$$A_a^u\otimes I\simeq_\varepsilon I\otimes B^\ell_{[f(u)=a]},\qquad A_a^u\otimes I\simeq_\delta I\otimes A_a^u,\qquad A_a^u\otimes I\simeq_\gamma I\otimes L^\ell_{[f(u)=a]}.$$

Proof. Each subtest accepts exactly when the two measurement outcomes agree. Because the relevant families are measurements, Lemma 3.18 rewrites those acceptance probabilities as consistency bounds. □

Lemma 3.18 (Consistency for measurements)**.** *If $A = \{A^x_a\}$ and $B = \{B^x_a\}$ are measurements, then*

$$A^x_a \otimes I \simeq_\delta I \otimes B^x_a \qquad \textit{if and only if} \qquad \mathbb{E}_x \sum_a \langle\psi| A^x_a \otimes B^x_a |\psi\rangle \geq 1 - \delta.$$

Proof. Expand the off-diagonal sum by writing $\sum_{a \neq b} A^x_a \otimes B^x_b = \sum_b (I - A^x_b) \otimes B^x_b$. $\square$

Definition 3.19 (State-dependent distance)**.** Let $|\psi\rangle \in \mathcal{H}$, let $A = \{A^x_a\}$ and $B = \{B^x_a\}$ be families of operators on $\mathcal{H}$, and let $\mathcal{D}$ be a distribution on the question set. We write

$$A^x_a \approx_\delta B^x_a$$

when

$$\mathbb{E}_{x \sim \mathcal{D}} \sum_a \|(A^x_a - B^x_a)|\psi\rangle\|^2 \leq \delta.$$

In particular, if $A^x_a \otimes I \simeq_\varepsilon I \otimes C^x_a$, the question is what hypothesis on A and B allows one to conclude that

$$B^x_a \otimes I \simeq_\varepsilon I \otimes C^x_a. \tag{3.3}$$

Lemma 3.20 (Transferring consistency through state-dependent distance)**.** *Let $A = \{A^x_a\}$ and $B = \{B^x_a\}$ be measurements and let $C = \{C^x_a\}$ be a submeasurement. If $A^x_a \otimes I \simeq_\delta I \otimes C^x_a$ and $A^x_a \otimes I \approx_\varepsilon B^x_a \otimes I$, then*

$$B^x_a \otimes I \simeq_{\delta + \sqrt{\varepsilon}} I \otimes C^x_a.$$

Proof. Rewrite the inconsistency with C as total mass minus diagonal overlap. The total mass is unchanged because A and B are measurements, and the diagonal overlap changes by at most $\sqrt{\varepsilon}$ by Cauchy–Schwarz. $\square$

Lemma 3.21 (Consistency implies state-dependent distance for measurements)**.** *If A and B are measurements and $A^x_a \otimes I \simeq_\delta I \otimes B^x_a$, then*

$$A^x_a \otimes I \approx_{2\delta} I \otimes B^x_a.$$

If both measurements are projective, the converse also holds: if $A^x_a \otimes I \approx_{2\delta} I \otimes B^x_a$ then $A^x_a \otimes I \simeq_\delta I \otimes B^x_a$.

Proof. Expand the squared norm of $(A^x_a \otimes I - I \otimes B^x_a)|\psi\rangle$ and use the diagonal-overlap formula from Lemma 3.18; projectivity turns the diagonal inequality into an equality, giving the converse. $\square$

Remark 3.22. The implication above does not extend to submeasurements. For example, if $A^x_a = 0$ for all a, then $A^x_a \otimes I \simeq_0 I \otimes B^x_a$, but

$$\mathbb{E}_x \sum_a \|(A^x_a \otimes I - I \otimes B^x_a)|\psi\rangle\|^2 = \mathbb{E}_x \sum_a \|(I \otimes B^x_a)|\psi\rangle\|^2,$$

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

3.3 Consistency from state-dependent distance

Theorem 3.23 (Closeness of inner products)**.** *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,*

$$\sum_a \Big(\sum_b C_{a,b}^x\Big)\Big(\sum_b C_{a,b}^x\Big)^\dagger \le I.$$

Then

$$\mathbb{E}_x \sum_{a,b} \langle\psi|C_{a,b}^x A_a^x|\psi\rangle \approx_{\sqrt{\gamma}} \mathbb{E}_x \sum_{a,b} \langle\psi|C_{a,b}^x B_a^x|\psi\rangle. \tag{3.4}$$

Similarly, suppose that $(A_a^x)^\dagger \approx_\gamma (B_a^x)^\dagger$ and that for all x,

$$\sum_a \Big(\sum_b C_{a,b}^x\Big)^\dagger \Big(\sum_b C_{a,b}^x\Big) \le I.$$

Then

$$\mathbb{E}_x \sum_{a,b} \langle\psi|A_a^x C_{a,b}^x|\psi\rangle \approx_{\sqrt{\gamma}} \mathbb{E}_x \sum_{a,b} \langle\psi|B_a^x C_{a,b}^x|\psi\rangle. \tag{3.5}$$

Proof. To prove (3.4), write the difference as

$$\mathbb{E}_x \sum_a \langle\psi|\Big(\sum_b C_{a,b}^x\Big)(A_a^x - B_a^x)|\psi\rangle.$$

Cauchy–Schwarz bounds its magnitude by

$$\Big(\mathbb{E}_x \sum_a \langle\psi|\Big(\sum_b C_{a,b}^x\Big)\Big(\sum_b C_{a,b}^x\Big)^\dagger|\psi\rangle\Big)^{1/2}\Big(\mathbb{E}_x \sum_a \langle\psi|(A_a^x - B_a^x)^\dagger(A_a^x - B_a^x)|\psi\rangle\Big)^{1/2},$$

and the two hypotheses bound these factors by 1 and $\sqrt{\gamma}$ respectively. For (3.5), rewrite the difference as

$$\mathbb{E}_x \sum_{a,b} \langle\psi|(C_{a,b}^x)^\dagger((A_a^x)^\dagger - (B_a^x)^\dagger)|\psi\rangle,$$

and apply (3.4) to the adjoint families. □

Theorem 3.24. *Let $A = \{A_a^x\}$, $B = \{B_a^x\}$, and $C = \{C_a^x\}$ be submeasurements such that $A_a^x \approx_\delta B_a^x$. Then*

$$\mathbb{E}_x \sum_a \langle\psi|A_a^x C_a^x|\psi\rangle \approx_{\sqrt{\delta}} \mathbb{E}_x \sum_a \langle\psi|B_a^x C_a^x|\psi\rangle.$$

Proof. The difference has magnitude

$$\Big|\mathbb{E}_x \sum_a \langle\psi|(A_a^x - B_a^x)C_a^x|\psi\rangle\Big|.$$

By Cauchy–Schwarz this is at most

$$\Big(\mathbb{E}_x \sum_a \langle\psi|(A_a^x - B_a^x)^2|\psi\rangle\Big)^{1/2}\Big(\mathbb{E}_x \sum_a \langle\psi|(C_a^x)^2|\psi\rangle\Big)^{1/2}.$$

The first factor is at most $\sqrt{\delta}$ by hypothesis, and the second is at most 1 because C is a submeasurement. □

Theorem 3.25. *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 term is

$$\mathbb{E}_x \sum_{a,b} \langle \psi | (A_a^x - B_a^x)^\dagger (C_{a,b}^x)^\dagger C_{a,b}^x (A_a^x - B_a^x) | \psi \rangle.$$

Using $\sum_b (C_{a,b}^x)^\dagger (C_{a,b}^x) \leq I$, this is bounded by

$$\mathbb{E}_x \sum_a \langle \psi | (A_a^x - B_a^x)^\dagger (A_a^x - B_a^x) | \psi \rangle \leq \delta.$$

□

Lemma 3.26 (Triangle inequality for vectors squared)**.** *For vectors $|\psi_1\rangle, \ldots, |\psi_k\rangle$,*

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

Proof. First,

$$(x_1 + \cdots + x_k)^2 \leq k(x_1^2 + \cdots + x_k^2) \tag{3.6}$$

for all real numbers $x_1, \ldots, x_k$. Apply (3.6) to $x_i = \| |\psi_i\rangle \|$ and combine it with the triangle inequality. □

Lemma 3.27 (Triangle inequality for $\approx_\delta$)**.** *Suppose $(A_i)_a^x \approx_{\delta_i} (A_{i+1})_a^x$ for $i = 1, \ldots, k$. Then*

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

Proof. Expand $(A_1 - A_{k+1})|\psi\rangle$ as a telescoping sum and apply Lemma 3.26. □

Remark 3.28 (Lean finite-step triangle helpers). The general operator-family theorem above is complemented in Lean by reusable binary and three-step helper lemmas for operator expectations, ‘qSDD‘, and ‘SDDRel‘. The three-step ‘SDDRel‘ helper is the specialization used in the Step 6 projectivization chain in Theorem 2.14; this remark is only a cross-reference and makes no separate proof-level completeness claim.

Lemma 3.29 (Triangle inequality for $\simeq_\delta$)**.** *Suppose $A_a^x \otimes I \simeq_\varepsilon I \otimes B_a^x$, $C_a^x \otimes I \simeq_\delta I \otimes B_a^x$, and $C_a^x \otimes I \simeq_\gamma I \otimes D_a^x$, where all four families are measurements. Then*

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

Proof. Convert the two consistencies involving C to state-dependent distance, compose them by the triangle inequality, and transfer the result back to consistency by Lemma 3.20. □

Lemma 3.30 (Data processing for consistency)**.** *If $A = \{A_a^x\}$ and $B = \{B_a^x\}$ are measurements such that $A_a^x \otimes I \simeq_\delta I \otimes B_a^x$, and f is a map on the answer set, then*

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

Proof. Merging outcome classes can only decrease the off-diagonal mass. □

Lemma 3.31 (Consistency controls a submeasurement). *Let $A = \{A^x_a\}$ be a submeasurement and $B = \{B^x_a\}$ a measurement such that $A^x_a \otimes I \simeq_\gamma I \otimes B^x_a$. Then*

$$A^x_a \otimes I \approx_\gamma A^x_a \otimes B^x_a \approx_\gamma A^x \otimes B^x_a, \tag{3.7}$$

where $A^x = \sum_a A^x_a$. In particular,

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

Proof. For the first approximation in (3.7), bound

$$\mathbb{E}_x \sum_a \langle\psi|\big(A^x_a \otimes (I - B^x_a)\big)^2|\psi\rangle$$

by the inconsistency between A and B. The second approximation is identical, applied to

$$\mathbb{E}_x \sum_a \langle\psi|\big((A^x - A^x_a) \otimes B^x_a\big)^2|\psi\rangle.$$

Then compose the two bounds with Lemma 3.27. □

Lemma 3.32 (Switch-sandwich). *Suppose $A = \{A^x_a\}$ is a projective submeasurement satisfying*

$$A^x_a \otimes I \approx_\delta I \otimes A^x_a. \tag{3.8}$$

Then for every operator B with $0 \le B \le I$,

$$\mathbb{E}_x \sum_a \langle\psi|A^x_a B A^x_a \otimes I|\psi\rangle \approx_{2\sqrt{\delta}} \mathbb{E}_x \sum_a \langle\psi|B \otimes A^x_a|\psi\rangle \approx_{\sqrt{\delta}} \mathbb{E}_x \sum_a \langle\psi|B A^x_a \otimes I|\psi\rangle. \tag{3.9}$$

Proof. We prove the first approximation in (3.9) in two steps. First,

$$\mathbb{E}_x \sum_a \langle\psi|A^x_a B A^x_a \otimes I|\psi\rangle \approx_{\sqrt{\delta}} \mathbb{E}_x \sum_a \langle\psi|A^x_a B \otimes A^x_a|\psi\rangle. \tag{3.10}$$

This is obtained by applying Cauchy–Schwarz to the difference and using $B^2 \le B \le I$ together with (3.8). For the second step, move the remaining left copy of A^x_a across the bipartition in the same way; projectivity then collapses the sandwich and yields the first approximation in (3.9). The second approximation in (3.9) is the same argument with the rightmost factor A^x_a omitted. □

3.4 Strong self-consistency

Lemma 3.33 (Completeness transfer to a projective approximation). *Let $A = \{A^x_a\}$ be a submeasurement and $P = \{P^x_a\}$ a projective submeasurement such that $A^x_a \otimes I \approx_\varepsilon P^x_a \otimes I$. Then*

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

Proof. Since P is projective,

$$\langle\psi|P \otimes I|\psi\rangle = \mathbb{E}_x \sum_a \langle\psi|(P^x_a)^2 \otimes I|\psi\rangle.$$

Apply Proposition 3.24 twice to replace $(P^x_a)^2$ first by $A^x_a P^x_a$ and then by $(A^x_a)^2$, and conclude with $(A^x_a)^2 \le A^x_a$. □

Definition 3.34 (Strong self-consistency)**.** Let $|\psi\rangle$ be permutation-invariant and let $A = \{A_a^x\}$ be a submeasurement. We say that A is δ-strongly self-consistent when

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

Lemma 3.35 (Strong self-consistency implies ordinary self-consistency)**.** *If A is δ-strongly self-consistent, then $A_a^x \otimes I \simeq_\delta I \otimes A_a^x$. If A is a measurement, the converse also holds.*

Proof. Rewrite the off-diagonal mass as the difference between the total mass of A and the diagonal overlap in the strong self-consistency inequality. When A is a full measurement, the total mass is the identity, so this inequality becomes an equality; hence the ordinary self-consistency defect and the strong self-consistency defect coincide, giving the converse. □

Lemma 3.36 (Strong self-consistency and state-dependent distance)**.** *If A is δ-strongly self-consistent, then*

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

If A is projective, the converse also holds.

Proof. Expanding the squared norm gives

$$\begin{aligned}
&\mathbb{E}_x \sum_a \|(A_a^x \otimes I - I \otimes A_a^x)|\psi\rangle\|^2 \\
&= 2 \cdot \Big(\mathbb{E}_x \sum_a \langle\psi|(A_a^x)^2 \otimes I|\psi\rangle - \mathbb{E}_x \sum_a \langle\psi|A_a^x \otimes A_a^x|\psi\rangle\Big) \\
&\leq 2 \cdot \Big(\mathbb{E}_x \sum_a \langle\psi|A_a^x \otimes I|\psi\rangle - \mathbb{E}_x \sum_a \langle\psi|A_a^x \otimes A_a^x|\psi\rangle\Big). \qquad (3.11)
\end{aligned}$$

The strong self-consistency bound controls the last line by 2δ. If A is projective, then (3.11) is an equality, which gives the converse. □

Lemma 3.37 (Post-processing preserves strong self-consistency)**.** *If A is δ-strongly self-consistent and f is a function on its answer set, then*

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

Proof. The same computation as in Lemma 3.36 gives

$$\begin{aligned}
&\mathbb{E}_x \sum_b \|(A^x_{[f(a)=b]} \otimes I - I \otimes A^x_{[f(a)=b]})|\psi\rangle\|^2 \\
&\leq 2 \cdot \Big(\langle\psi|A \otimes I|\psi\rangle - \mathbb{E}_x \sum_b \langle\psi|A^x_{[f(a)=b]} \otimes A^x_{[f(a)=b]}|\psi\rangle\Big). \qquad (3.12)
\end{aligned}$$

The post-processed diagonal term dominates $\mathbb{E}_x \sum_a \langle\psi|A_a^x \otimes A_a^x|\psi\rangle$, so strong self-consistency bounds (3.12) by 2δ. □

Lemma 3.38 (Completeness transfer for a strongly self-consistent approximation)**.** *Let A be a δ-strongly self-consistent submeasurement and let B be a submeasurement such that $A_a^x \otimes I \approx_\varepsilon B_a^x \otimes I$. Then*

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

Proof. Bound $\langle\psi|B \otimes I|\psi\rangle$ from below by $\mathbb{E}_x \sum_a \langle\psi|B_a^x \otimes B_a^x|\psi\rangle$. Then apply Proposition 3.24 twice to replace this by $\mathbb{E}_x \sum_a \langle\psi|A_a^x \otimes A_a^x|\psi\rangle$, and finish with strong self-consistency of A. □

Lemma 3.39 (Self-consistency implies data-processing closeness)**.** *Let A be a δ-strongly self-consistent submeasurement and let P be a projective submeasurement such that $P^x_a \otimes I \approx_\varepsilon A^x_a \otimes I$. Then for every function f,*

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

Proof. First prove

$$P^x_{[f(a)=b]} \otimes I \approx_{2\delta+4\sqrt{\varepsilon}} I \otimes A^x_{[f(a)=b]}. \tag{3.13}$$

Expanding the squared norm and using projectivity gives

$$\begin{aligned}&\mathbb{E}_x \sum_b \|(P^x_{[f(a)=b]} \otimes I - I \otimes A^x_{[f(a)=b]})|\psi\rangle\|^2\\ &\leq \langle\psi|P \otimes I|\psi\rangle + \langle\psi|I \otimes A|\psi\rangle - 2 \cdot \mathbb{E}_x \sum_b \langle\psi|P^x_{[f(a)=b]} \otimes A^x_{[f(a)=b]}|\psi\rangle. \end{aligned} \tag{3.14}$$

Proposition 3.33 bounds the first term by $\langle\psi|A \otimes I|\psi\rangle + 2\sqrt{\varepsilon}$, and Proposition 3.24 together with strong self-consistency bounds the third term below by $\langle\psi|A \otimes I|\psi\rangle - \delta - \sqrt{\varepsilon}$. Substituting into (3.14) yields (3.13). Now combine (3.13) with

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

from Lemma 3.37, and conclude by Lemma 3.27. □

Lemma 3.40 (Self-consistency moves one copy to the same side)**.** *Let $A = \{A_a\}$ be a ζ-strongly self-consistent submeasurement. 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. Apply Cauchy–Schwarz to the two-sided overlap $\sum_a \langle\psi|A_a \otimes A_a|\psi\rangle$ and compare it with the strong self-consistency lower bound. □

Lemma 3.41 (Bounding the missing mass of a close submeasurement)**.** *Let $A = \{A_a\}$ be a measurement that is ζ-strongly self-consistent, and let $B = \{B_a\}$ be a submeasurement such that $A_a \otimes I \approx_\delta B_a \otimes I$. Writing $B = \sum_a B_a$, one has*

$$\|(I - B) \otimes I|\psi\rangle\|^2 \leq 2\sqrt{\delta} + \zeta.$$

Proof. Since $0 \leq I - B \leq I$, the left-hand side is at most $1 - \langle\psi|B \otimes I|\psi\rangle$. Compare $\sum_a \langle\psi|B_a^2 \otimes I|\psi\rangle$ with $\sum_a \langle\psi|A_a B_a \otimes I|\psi\rangle$ and then with $\sum_a \langle\psi|A_a^2 \otimes I|\psi\rangle$ using the $\approx_\delta$ hypothesis, and finish with Lemma 3.40. □

Lemma 3.42 (Completing a close submeasurement to a full measurement)**.** *Let $A = \{A_a\}$ be a measurement that is ζ-strongly self-consistent, let $B = \{B_a\}$ be a submeasurement such that $A_a \otimes I \approx_\delta B_a \otimes I$, and let C be the measurement obtained by adding the missing mass $I - B$ to one distinguished answer a^*. Then*

$$A_a \otimes I \approx_{2\delta+4\sqrt{\delta}+2\zeta} C_a \otimes I.$$

Proof of 3.42. By Lemma 3.26,

$$\sum_a \|(A_a - C_a) \otimes I|\psi\rangle\|^2 \leq 2\delta + 2 \cdot \|(I - B) \otimes I|\psi\rangle\|^2.$$

It remains to bound the residual term. Since $0 \leq I - B \leq I$,

$$\|(I - B) \otimes I|\psi\rangle\|^2 \leq 1 - \sum_a \langle\psi|B_a^2 \otimes I|\psi\rangle. \tag{3.15}$$

By Proposition 3.24,

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

Proposition 3.43 gives

$$\sum_a \langle\psi|A_a^2 \otimes I|\psi\rangle \geq \langle\psi|A \otimes I|\psi\rangle - \zeta = 1 - \zeta.$$

Hence (3.15) implies

$$\|(I - B) \otimes I|\psi\rangle\|^2 \leq 2\sqrt{\delta} + \zeta.$$

Substituting this into the first bound proves the claim. □

Theorem 3.43. *This is the same statement as Lemma 3.40: for a ζ-strongly self-consistent submeasurement A,*

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

Proof. Immediate from Lemma 3.40. □

Chapter 4

Making measurements projective

Theorem 4.1 (Naimark dilation)**.** *Let $|\psi\rangle$ be a state in $\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}$. Let $A = \{A^x_a\}$ be a sub-measurement acting on $\mathcal{H}_{\mathrm{A}}$ and $B = \{B^y_b\}$ be a sub-measurement acting on $\mathcal{H}_{\mathrm{B}}$. Then there exist Hilbert spaces $\mathcal{H}_{\mathrm{A}_{\mathsf{aux}}}$ and $\mathcal{H}_{\mathrm{B}_{\mathsf{aux}}}$, a state $|\mathsf{aux}\rangle \in \mathcal{H}_{\mathrm{A}_{\mathsf{aux}}} \otimes \mathcal{H}_{\mathrm{B}_{\mathsf{aux}}}$, and projective sub-measurements $\widehat{A} = \{\widehat{A}^x_a\}$ and $\widehat{B} = \{\widehat{B}^y_b\}$ acting on $\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{A}_{\mathsf{aux}}}$ and $\mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{B}_{\mathsf{aux}}}$, respectively, such that, if $|\widehat{\psi}\rangle = |\psi\rangle \otimes |\mathsf{aux}\rangle$, then for all x, y, a, b,*

$$\langle\psi|A^x_a \otimes B^y_b|\psi\rangle = \langle\widehat{\psi}|\widehat{A}^x_a \otimes \widehat{B}^y_b|\widehat{\psi}\rangle.$$

In addition, $|\mathsf{aux}\rangle$ is a product state. The word "measurement" in the source statement is read here in the projective-submeasurement sense used by Lemma 4.7: for a general sub-measurement, the residual mass is carried by the additional $\perp$ outcome in the auxiliary construction, so the original outcomes need not sum to the identity.

Proof.

For each question x, choose a local auxiliary state $|\mathsf{aux}_{\mathrm{A},x}\rangle$ of dimension one larger than the number of outcomes of A^x, and apply Lemma 4.7 to obtain a projective sub-measurement $\widetilde{A}^x$. Do the same for each question y on Bob's side, obtaining auxiliary states $|\mathsf{aux}_{\mathrm{B},y}\rangle$ and projective sub-measurements $\widetilde{B}^y$.

Let

$$|\mathsf{aux}\rangle = \Big(\bigotimes_x |\mathsf{aux}_{\mathrm{A},x}\rangle\Big) \otimes \Big(\bigotimes_y |\mathsf{aux}_{\mathrm{B},y}\rangle\Big), \qquad |\widehat{\psi}\rangle = |\psi\rangle \otimes |\mathsf{aux}\rangle,$$

and let $\widehat{A}^x_a$ and $\widehat{B}^y_b$ be obtained by letting $\widetilde{A}^x_a$ or $\widetilde{B}^y_b$ act on the relevant auxiliary factor and the identity act on every other auxiliary factor. Because the auxiliary state is a product state and the two dilations act on disjoint tensor factors, the correlation for fixed (x, y) reduces to the compression identity from Lemma 4.7 on the x- and y-registers. Hence

$$\langle\widehat{\psi}|\widehat{A}^x_a \otimes \widehat{B}^y_b|\widehat{\psi}\rangle = \langle\psi|A^x_a \otimes B^y_b|\psi\rangle.$$

□

Remark 4.2 (Lean questionwise Naimark interface). In addition to Theorem 4.1, Lean records the questionwise one-measurement interface obtained from Lemma 4.7. For every indexed family A^x and B^y, it produces local dilation data for each question and proves the corresponding single-outcome marginal-preservation identities. This is a useful Lean-only auxiliary interface for downstream arguments, not a substitute for the full tensor-product correlation theorem, which is proved separately by `naimarkTensorProductCorrelation`.

Remark 4.3 (Lean auxiliary declarations for Naimark). The checked Lean proof of Theorem 4.1 uses the standard one-measurement auxiliary register

$$\mathrm{Option}(\mathcal{A}),$$

with the all-$\perp$ vector as the initial auxiliary state. A separate Naimark unitary is chosen for each question, so the proof does not need a distinct auxiliary tensor factor for every question. This is equivalent for the theorem's conclusion, since no commutation relation between different questions' dilated measurements is required. The declaration `OneMeasNaimarkData.twoSidedCorrelationPreservation` proves the four-register trace identity which turns the two local compression identities into the bipartite correlation-preservation identity used by `naimarkTensorProductCorrelation`.

Theorem 4.4 (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

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

Proof. Complete the sub-measurement by adding the missing outcome and use permutation invariance to convert strong self-consistency into the cross-consistency hypothesis of Lemma 4.9. Applying the locality-preserving projectivization repair gives the required projective sub-measurement with the paper constant $100\zeta^{1/4}$. □

Remark 4.5 (Lean right-register completion helpers). The inductive-step application of the orthogonalization lemma uses completion on both tensor factors. The formal development records the analytic completion statement on the left register; the shared preliminary helper compares right- and left-tensor placements at the level of the underlying state-dependent-distance form. Its submeasurement and indexed lemmas transport state-dependent distance to the right register under the permutation-invariance hypothesis on the state. The last linked declaration is a Lean-only construction from two already constructed orthonormalize-and-complete statements and a pre-projective consistency proof. It is not a source theorem and it is not an additional hypothesis of Theorem 2.14; it records only the tensor-factor bookkeeping needed to form the later line-156 projectivization transition.

Remark 4.6 (Lean line-169 projectivization match-mass declarations). Lean exposes the quantitative obstruction at the projectivization transition used in the inductive step. The existing line-156 transition gives only the generic consequence of Proposition 3.20: from $G^{\mathrm{A}} \simeq_{\zeta_1} G^{\mathrm{B}}$ and $G^{\mathrm{A}} \approx_{\zeta_2} Q^{\mathrm{A}}$ one obtains $Q^{\mathrm{A}} \simeq_{\zeta_1+\sqrt{\zeta_2}} G^{\mathrm{B}}$, and similarly on the other tensor factor. This direct square-root-loss route is an immediate specialization of the general Lean declarations `MIPStarRE.LDT.Preliminaries.triangleSub` and `MIPStarRE.LDT.Preliminaries.triangleSub_right` to the handoff fields, and is no longer maintained as separate named projectivization API.

Lean also proves a more local repaired transport that keeps the comparison at the pre-completion stage. If P^{A} is the orthonormalized projective sub-measurement with $G^{\mathrm{A}} \approx_\epsilon P^{\mathrm{A}}$, then Proposition 3.24 bounds the diagonal match-mass loss by $\sqrt{\epsilon}$, and canonical completion can only add nonnegative diagonal match mass. Thus the declarations `leftConsistency_of_completion_and_sdd` and `rightConsistency_of_completion_and_sdd` prove $Q^{\mathrm{A}} \simeq_{\zeta_1+\sqrt{\epsilon}} G^{\mathrm{B}}$ and its Bob-side mirror from the pre-projective ζ_1 consistency plus the pre-completion $\approx_\epsilon$ comparison. In particular, substituting the checked orthonormalization error $\epsilon = 100\zeta_1^{1/4}$ yields the repaired line-169 bound

$\zeta_1 + 10\zeta_1^{1/8}$ recorded by `leftConsistency_with_orthonormalization_loss` and its mirror. This is much sharper than the direct completion-level $\zeta_1 + \sqrt{\zeta_2}$ route, but it is still not the paper's exact ζ_1.

Lean also isolates the completion fact used inside this sharper repair: the canonical completion of a projective submeasurement can only increase its diagonal match mass against a fixed partner submeasurement. This is the formal declaration `ProjectivizationMatchMassMonotonicity.completeAtOutcomePr` The sharper route therefore pays its square-root loss before completion and then uses completion only in this monotone direction, with no additional match-mass penalty.

The older exact-ζ_1 construction-level monotonicity statement has been retired. The current Section 5 API proves the checked repaired transport above, and the active orthonormalization route now uses those repaired completion-transport estimates rather than a separate exact match-mass witness.

4.0.1 Naimark dilation

Lemma 4.7. *Let $A = \{A_a\}$ be a sub-measurement with k distinct outcomes $a \in \mathcal{A}$, and let $|\mathsf{aux}\rangle \in \mathbb{C}^{k+1}$ be any state. Then there exists a projective sub-measurement $\widehat{A} = \{\widehat{A}_a\}$ such that for each outcome a,*

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

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

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

where $A = \sum_a A_a$. Then

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

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

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

Define

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

Each $\widehat{A}_a$ is a projector, and

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

□

Remark 4.8 (Naimark dilation does not preserve $\approx_\delta$). Let $|\psi\rangle$ be any state in $\mathbb{C}^d \otimes \mathbb{C}^d$, and consider the two-outcome measurements $A = \{A_0, A_1\}$ and $B = \{B_0, B_1\}$ with

$$A_0 = A_1 = B_0 = B_1 = \tfrac{1}{2} I.$$

For each outcome $a \in \{0, 1\}$, the operators $A_a \otimes I$ and $I \otimes B_a$ are both equal to $\frac{1}{2} I \otimes I$, so $A_a \otimes I \approx_0 I \otimes B_a$ on the state $|\psi\rangle$.

Now apply the Naimark construction with a qubit auxiliary space on each side and auxiliary state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Since A and B are genuine measurements rather than sub-measurements, the extra $|\perp\rangle$ direction from 4.7 is unnecessary in this example. The unitary from 4.7 can then be taken to be the identity, so the dilated measurements are

$$\widehat{A}_a = I \otimes |a\rangle\langle a|, \qquad \widehat{B}_a = I \otimes |a\rangle\langle a|.$$

If $|\widehat{\psi}\rangle = |\psi\rangle \otimes |+\rangle \otimes |+\rangle$, then a direct calculation gives

$$\sum_{a\in\{0,1\}} \left\|(\widehat{A}_a \otimes I - I \otimes \widehat{B}_a)|\widehat{\psi}\rangle\right\|^2 = 1.$$

Thus on the state $|\widehat{\psi}\rangle$ the dilated family satisfies $\widehat{A}_a \otimes I \approx_\delta I \otimes \widehat{B}_a$ only for $\delta \geq 1$, even though the original family satisfied the same relation with $\delta = 0$. This is why we use Theorem 4.4 rather than simply applying Theorem 4.1: Naimark preserves the outcome probabilities, but it need not preserve the state-dependent $\approx_\delta$ estimates needed later.

4.0.2 Orthogonalization lemma

Lemma 4.9 (Orthogonalization lemma for measurements)**.** *Let $|\psi\rangle$ be a state, 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 $|\psi\rangle$. Here $|\psi\rangle$ is not assumed to be permutation-invariant. Then there exists a projective sub-measurement $P = \{P_a\}$ such that

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

Lemma 4.10 (Formalized orthogonalization envelope for measurements)**.** *This same-space corollary weakens Lemma 4.9 to the public orthonormalization envelope $100\zeta^{1/4}$. Thus, if $A = \{A_a\}$ and $B = \{B_a\}$ are measurements on the same finite-dimensional space satisfying*

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

on a normalized bipartite state $|\psi\rangle$, then there exists a projective sub-measurement $P = \{P_a\}$ such that

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

This is the corrected Lean theorem proved from the Section 5 projectivization repair construction. The sharper heterogeneous source statement $84\zeta^{1/4}$ is now formalized as Lemma 4.9.

Proof. The consistency hypothesis first gives the source almost-projective estimate for $A_a \otimes I$. The locality-preserving projectivization repair then produces a projective sub-measurement on the original space, and the scalar bounds convert the 84-constant repair estimate into the public $100\zeta^{1/4}$ envelope. □

Remark 4.11 (Lean auxiliary declarations for the orthogonalization lemma). The Lean declaration `MIPStarRE.LDT.MakingMeasurementsProjective.orthonormalizationMainLemma` proves Lemma 4.9 without adding spectral-truncation or repair hypotheses. The proof first converts the cross-consistency hypothesis into the source almost-projective estimate for the measurement

$A_a \otimes I$, and then applies the locality-preserving $Q/X/\widehat{X}/P$ construction in its heterogeneous left-register form. Lemma 4.10 remains as the same-space corollary with the weaker public $100\zeta^{1/4}$ envelope. The measurement-level declarations `orthonormalizationMeasurement_of_consistency` and `orthonormalizationMeasurement` are proved corollaries, not conditional wrappers: they derive the projective submeasurement from the displayed consistency or self-consistency hypotheses.

Lemma 4.12 (Locality-preserving projectivization)**.** *Paper origin:* `references/ldt-paper/orthonormalization.t` *lines 534–860 (rank reduction and the Q-side setup, including 4.28 and 4.29) and 862–1194 (the $X/\widehat{X}/P$ algebra: 4.30, 4.33, 4.34, 4.36, 4.37, 4.38, plus the final triangle-inequality calculation producing the $84\zeta^{1/4}$ bound). Together these are the late repair stage of the orthogonalization-lemma proof.*

Given a normalized bipartite state $|\psi\rangle$, a measurement $A = \{A_a\}$, and the source almost-projective estimate for the left-lifted family $\{A_a \otimes I\}$, there exists a locality-preserving projective sub-measurement with error $84\zeta^{1/4}$.

More precisely, if $|\psi\rangle$ is a normalized bipartite state, if $A = \{A_a\}$ is a measurement, and if

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

then there exists a projective sub-measurement $P = \{P_a\}$ on the underlying space such that

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

The locality-preserving form – output $P_a \otimes I$ rather than an arbitrary lifted family – is recorded in Remark 4.11 and is the form required by the unconditional orthonormalization theorem (4.4).

Proof. Pass to the left marginal state, apply the rank-reduction and $Q/X/\widehat{X}/P$ construction there, and transport the resulting $\approx_{84\zeta^{1/4}}$ estimate back to left lifts via the left-marginal expectation formula. □

Proposition 4.13 (Measurement-level orthogonalization from consistency)**.** *This is the same-space measurement corollary recorded in Remark 4.11. Let $|\psi\rangle$ be a normalized bipartite state, let $A = \{A_a\}$ and $B = \{B_a\}$ be measurements on the same finite-dimensional space, and let $0 \le \zeta$. Assume the bipartite consistency estimate*

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

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

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

Proof. The proof applies Lemma 4.9 and weakens the $84\zeta^{1/4}$ estimate to the public orthonormalization envelope. The spectral-truncation and locality-preserving repair steps are internal to that proof; they are not hypotheses of this declaration. □

Proposition 4.14 (Measurement-level orthogonalization from self-consistency)**.** *This is the self-consistency specialization of Proposition 4.13. Let $|\psi\rangle$ be a normalized bipartite state, let $A = \{A_a\}$ be a measurement, and let $0 \le \zeta$. If A is bipartite strongly self-consistent with error ζ, then there exists a projective sub-measurement $P = \{P_a\}$ such that*

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

Proof. Since A is complete, the self-consistency defect is the consistency defect of A with itself. The proof therefore applies the consistency corollary with $B = A$. No spectral-truncation witness or repair witness is assumed at this interface. $\square$

Proposition 4.15 (Completion-route orthonormalization)**.**

The paper theorem 4.4 is proved in Lean with the $100\zeta^{1/4}$ constant. The declaration checked here is the completion-route theorem with the weaker bound $120\zeta^{1/4}$, recorded in `orthonormalizationCompletionRout`
See [con26a].

Let $|\psi\rangle$ be a normalized permutation-invariant state, and let $A = \{A_a\}_{a\in\mathcal{A}}$ be a sub-measurement which is bipartite strongly self-consistent with error ζ. Then there exists a projective sub-measurement $P = \{P_a\}_{a\in\mathcal{A}}$ such that

$$A_a \otimes I \approx_{120\zeta^{1/4}} P_a \otimes I.$$

Proof. Write $A = \sum_a A_a$. The strong self-consistency assumption implies

$$\langle\psi|A \otimes A|\psi\rangle \geq \langle\psi|A \otimes I|\psi\rangle - \zeta,$$

so

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

Rearranging once more gives

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

Complete A to the measurement $\widehat{A} = \{\widehat{A}_a\}_{a\in\widehat{\mathcal{A}}}$ on $\widehat{\mathcal{A}} = \mathcal{A} \cup \{\perp\}$ by setting $\widehat{A}_a = A_a$ for $a \in \mathcal{A}$ and $\widehat{A}_\perp = I - A$. Adding the two preceding inequalities shows that $\widehat{A}$ is 2ζ-strongly self-consistent. By 3.35, this gives the source almost-projective estimate

$$\sum_a \langle\psi|((\widehat{A}_a \otimes I) - (\widehat{A}_a \otimes I)^2)|\psi\rangle \leq 4\zeta.$$

Apply 4.12 to obtain a projective sub-measurement $\widehat{P}$ such that

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

Discard the extra outcome and set $P_a = \widehat{P}_a$ for $a \in \mathcal{A}$. Then

$$A_a \otimes I \approx_{84(4\zeta)^{1/4}} P_a \otimes I,$$

and $84 \cdot 4^{1/4} \leq 120$, so the completion-route bound follows. $\square$

Proof of 4.9. The statement is trivial when $\zeta > 1/4$, so assume

$$\zeta \leq 1/4. \tag{4.1}$$

By 3.18,

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

Apply Cauchy–Schwarz to

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

to obtain

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

Because A is a measurement, this rewrites as

$$\sum_a \langle\psi|(A_a - (A_a)^2) \otimes I|\psi\rangle \leq 2\zeta. \tag{4.2}$$

Lemma 4.16. *For any $x \in [0,1]$,*

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

Proof. If $x \geq 1-\delta$, then $\mathsf{trunc}_\delta(x) = 1$. Since $\delta \leq 1/2$ below, one also has $x \geq \delta$, so

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

If $x < 1-\delta$, then $\mathsf{trunc}_\delta(x) = 0$, and similarly

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

$\square$

Lemma 4.17 (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. Write the eigendecomposition of A_a as

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

Choose

$$0 < \delta \leq 1/2, \tag{4.3}$$

define $\mathsf{trunc}_\delta : [0,1] \to \{0,1\}$ by $\mathsf{trunc}_\delta(x) = 1$ when $x \geq 1-\delta$ and 0 otherwise, and set

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

By 4.16,

$$(A_a - R_a)^2 \leq \frac{1}{\delta}(A_a - (A_a)^2).$$

Summing expectations and using 4.2 gives

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

so $A_a \otimes I \approx_{2\zeta/\delta} R_a \otimes I$.

For every $x \in [0,1]$, one has $\mathtt{trunc}_\delta(x) \le x/(1-\delta)$, hence

$$R_a \le \frac{1}{1-\delta} A_a.$$

Summing over a and using that A is a measurement,

$$R = \sum_a R_a \le \frac{1}{1-\delta} I \le (1+2\delta) I,$$

where the last estimate uses $\delta \le 1/2$. Setting $\delta = \sqrt{\zeta}$ gives the claim, and 4.1 ensures $\delta \le 1/2$. $\square$

Remark 4.18 (Lean auxiliary declarations for rounding to projectors). The source-facing Lean declarations for Lemma 4.17 are the proposition `projectiveNonMeasurement` and the construction theorem `projectiveNonMeasurement_of_sourceAlmostProjective_full`. The declarations listed here are the endpoint and regime splits, the `AlmostProjMeasStatement` witness forms, and the spectral-truncation statement conversions used by later Section 5 arguments. They are not additional hypotheses in the paper statement of Lemma 4.17.

Lemma 4.19 (Combinatorial averaging for the $r > d$ branch). *Let $f : \alpha \to \mathbb{R}$ be a function on a finite index set α with $|\alpha| = r$, and let $L \subseteq \alpha$ with $|L| = d$ and $d < r$. Write $S = L^\complement$ and assume $f(s) \le f(l)$ for every $s \in S$ and $l \in L$. Then the pairwise double-counting inequality*

$$|L| \sum_{s \in S} f(s) \le |S| \sum_{l \in L} f(l)$$

holds, and hence also

$$r \sum_{s \in S} f(s) \le |S| \sum_{x \in \alpha} f(x).$$

If in addition $r \le (1 + 2\sqrt{\zeta})\, d$, $\sum_{x\in\alpha} f(x) \le 1 + 2\sqrt{\zeta}$, and $0 \le \zeta \le 1/4$, then

$$\sum_{s \in S} f(s) \le 4\sqrt{\zeta}.$$

Proof. Choose L to be a set of d indices on which the sum of f is maximal. If some $s \in S$ and $l \in L$ satisfied $f(l) < f(s)$, replacing l by s would increase this sum, so every small index is bounded by every large index. Summing the inequalities $f(s) \le f(l)$ over all pairs $(s,l) \in S \times L$ gives

$$|L| \sum_{s \in S} f(s) \le |S| \sum_{l \in L} f(l).$$

Since α is the disjoint union of L and S, adding $|S| \sum_{s \in S} f(s)$ to both sides yields

$$r \sum_{s \in S} f(s) \le |S| \sum_{x \in \alpha} f(x).$$

The hypotheses $|L| = d$ and $r \le (1+2\sqrt{\zeta})d$ imply $|S|/r \le 2\sqrt{\zeta}$. Dividing by r and using $\sum_{x \in \alpha} f(x) \le 1 + 2\sqrt{\zeta}$ gives

$$\sum_{s \in S} f(s) \le 2\sqrt{\zeta}(1 + 2\sqrt{\zeta}).$$

Finally, $\zeta \le 1/4$ implies $\sqrt{\zeta} \le 1/2$, and hence $2\sqrt{\zeta}(1+2\sqrt{\zeta}) \le 4\sqrt{\zeta}$. $\square$

Lemma 4.20 (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,

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

Proof. Let $\{R_a\}$ be given by 4.17. Write $r_a = \operatorname{rank}(R_a)$ and $r = \sum_a r_a$. If $r \leq d$, take $Q_a = R_a$. Otherwise,

$$r = \sum_a r_a = \sum_a \operatorname{Tr}(R_a) = \operatorname{Tr}(R) \leq (1 + 2\sqrt{\zeta})d. \tag{4.4}$$

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

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

Define the overlaps

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

Their total satisfies

$$\sum_a \sum_{i=1}^{r_a} o_{a,i} = \langle\psi|R \otimes I|\psi\rangle \leq 1 + 2\sqrt{\zeta}. \tag{4.5}$$

Let Large be the set of the d largest overlaps $o_{a,i}$, breaking ties arbitrarily, and let Small be the complement. Then

$$|\mathsf{Small}| = r - d \leq 2\sqrt{\zeta}\, d \leq 2\sqrt{\zeta}\, r, \tag{4.6}$$

by 4.4. Averaging the overlaps over the r indices gives

$$\sum_{(a,i)\in\mathsf{Small}} o_{a,i} \leq \frac{|\mathsf{Small}|}{r} \sum_{a,i} o_{a,i} \leq 4\sqrt{\zeta}, \tag{4.7}$$

using 4.6, 4.5, and 4.1.

For each a, let $\mathsf{Large}_a = \{i : (a,i) \in \mathsf{Large}\}$ and define

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

Then $\sum_a \operatorname{rank}(Q_a) = |\mathsf{Large}| \leq d$, and $Q_a \leq R_a$ for every a, so

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

Also,

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

is a projector, hence

$$\sum_a \|(R_a - Q_a) \otimes I|\psi\rangle\|^2 = \sum_{(a,i)\in\mathsf{Small}} o_{a,i} \leq 4\sqrt{\zeta}$$

by 4.7. Therefore $R_a \otimes I \approx_{4\sqrt{\zeta}} Q_a \otimes I$. Combining this with 4.17 and the triangle inequality for $\approx_\delta$ gives

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

□

Henceforth let $Q = \{Q_a\}$ be the family given by 4.20.

Lemma 4.21 (Completeness of Q)**.**

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

Proof. Since each Q_a is projective,

$$\langle\psi|Q \otimes I|\psi\rangle = \sum_a \langle\psi|(Q_a)^2 \otimes I|\psi\rangle \approx_{5\zeta^{1/4}} \sum_a \langle\psi|(Q_a \cdot A_a) \otimes I|\psi\rangle. \tag{4.8}$$

By Cauchy–Schwarz,

$$\sum_a \langle\psi|(Q_a)^2 \otimes I|\psi\rangle \leq 1 + 2\sqrt{\zeta} \quad \text{and} \quad \sum_a \langle\psi|(Q_a - A_a)^2 \otimes I|\psi\rangle \leq 12\sqrt{\zeta}$$

from 4.20. A second Cauchy–Schwarz estimate yields

$$\sum_a \langle\psi|(Q_a \cdot A_a) \otimes I|\psi\rangle \approx_{4\zeta^{1/4}} \sum_a \langle\psi|(A_a)^2 \otimes I|\psi\rangle. \tag{4.9}$$

Combining 4.8 and 4.9 with 4.2 gives

$$\langle\psi|Q \otimes I|\psi\rangle \geq \sum_a \langle\psi|A_a \otimes I|\psi\rangle - 2\zeta - 9\zeta^{1/4} = 1 - 11\zeta^{1/4},$$

since $\zeta \leq \zeta^{1/4}$ under 4.1. □

Lemma 4.22 (Completeness of $\sqrt{Q}$)**.**

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

Proof. Diagonalize $Q = \sum_i \nu_i |u_i\rangle\langle u_i|$. Because $Q \leq (1 + 2\sqrt{\zeta})I$ by 4.20, each ν_i is at most $1 + 2\sqrt{\zeta}$. Therefore

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

The scalar estimate

$$\frac{1}{\sqrt{1 + 2\sqrt{\zeta}}} \geq 1 - \sqrt{\zeta}$$

gives $\sqrt{Q} \geq (1 - \sqrt{\zeta})Q$. Hence 4.21 implies

$$\langle\psi|\sqrt{Q} \otimes I|\psi\rangle \geq (1 - \sqrt{\zeta})(1 - 11\zeta^{1/4}) \geq 1 - 12\zeta^{1/4}.$$

□

Lemma 4.23 (Q is almost projective)**.**

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

Proof. Since $Q \le (1+2\sqrt{\zeta})I$ by 4.20,

$$\begin{aligned}\sum_a Q_a \cdot Q \cdot Q_a - \sum_a Q_a &\le (1+2\sqrt{\zeta}) \sum_a Q_a \cdot I \cdot Q_a - \sum_a Q_a \\ &= (1+2\sqrt{\zeta})Q - Q \\ &= 2\sqrt{\zeta}\, Q \\ &\le 2\sqrt{\zeta}(1+2\sqrt{\zeta})I \\ &\le 4\sqrt{\zeta}\, I,\end{aligned}$$

using 4.1 in the last step. □

Definition 4.24 (Matrix decomposition of Q_a)**.** For each a, let

$$m_a = \mathrm{rank}(Q_a),$$

and let $|v_{a,1}\rangle, \dots, |v_{a,m_a}\rangle$ 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 let $\{|a,i\rangle\}$ be an orthonormal basis of $\mathbb{C}^m$, indexed by all a and all $1 \le i \le m_a$. For each a, define

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

and define

$$X = \sum_a X_a = \sum_a \sum_{i=1}^{m_a} |a,i\rangle\langle v_{a,i}|. \tag{4.10}$$

Let $T = \{T_a\}$. Then T is a projective measurement on $\mathbb{C}^m$.

Proof. Choose an orthonormal basis of the range of each projector Q_a. The rank-one expansion gives the displayed decomposition, and the orthogonality of the labelled basis vectors makes the family T a projective measurement. □

Definition 4.25 (SVD of X)**.** The paper presents this step by choosing a singular value decomposition

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

The formalization has two closely related auxiliary declarations. The reduced datum records the downstream consequences of the choice: the matrices X and $\widehat{X}$, the identities $X^\dagger X = Q$, $\widehat{X}\widehat{X}^\dagger = I$, and $X^\dagger \widehat{X} = \sqrt{Q}$, together with the $Q_a = X^\dagger T_a X$ restatement. The rectangular-SVD lemmas take explicit matrices $U, V, \Sigma, I_{m\times d}$ and derive the two stored $\widehat{X}$ identities from the matrix calculation below.

Proof. Apply the finite-dimensional singular value decomposition to X and record exactly the consequences used later: the identities for $X^\dagger X$, $\widehat{X}\widehat{X}^\dagger$, $X^\dagger \widehat{X}$, and the restatement of the Q_a's. The unitary-group constructors are the same rectangular-SVD calculation, with the square factors taken as elements of the relevant unitary groups rather than as matrices accompanied by separate unitarity equations. In the sigma-range specialization, the auxiliary row space is the lifted finite-enumeration carrier associated to the ranks of the Q_a's, which is definitionally the carrier of the canonical sigma-range Q-layer. □

Remark 4.26 (Formalization support for orthogonal projector families). If the family $\{Q_a\}$ were a projective measurement, then the vectors $|v_{a,i}\rangle$, over all a and all $1 \le i \le m_a$, would be orthonormal. Then (4.10) would already be a singular value decomposition of X, with $\Sigma = I_{m\times d}$. Equivalently,

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

This motivates replacing Σ by I in the definition of $\widehat{X}$ and P.

Definition 4.27 (Definition of P)**.** The paper defines

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

We take $\widehat{X}$ to be the reduced matrix supplied by the stored datum. For each a, define

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

Proof. The definition is the formal $Q/X/\widehat{X}/P$ construction: once $\widehat{X}$ is chosen from the stored singular-value data, each P_a is defined as the positive operator $\widehat{X}_a^\dagger \widehat{X}_a$. □

Lemma 4.28. *For each a,*

$$X_a = T_a \cdot X.$$

Proof. Multiply the definition of X by T_a on the left and use that T_a selects exactly the basis vectors $|a, i\rangle$:

$$T_a \cdot X = \Big(\sum_{i=1}^{m_a} |a,i\rangle\langle a,i|\Big)\Big(\sum_b \sum_{j=1}^{m_b} |b,j\rangle\langle v_{b,j}|\Big) = \sum_{i=1}^{m_a} |a,i\rangle\langle v_{a,i}| = X_a.$$

□

Lemma 4.29 (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 identity $X_a^\dagger X_a = Q_a$ follows by expanding the basis sums. The remaining equalities come from 4.28 and the projectivity of T:

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

□

Lemma 4.30 (X squared)**.**

$$X^\dagger \cdot X = Q.$$

Proof. Use that T is a measurement and then 4.29:

$$X^\dagger \cdot X = \sum_a X^\dagger \cdot T_a \cdot X = \sum_a Q_a = Q.$$

The additional paper identities involving U, V, Σ are the usual SVD computations. The Lean development also records the spectral calculation which will identify the mixed positive-Gram-image product with $\sqrt{Q}$: the positive eigenvalues give the nonzero columns, while the zero eigenspace has zero coefficient in the square-root expansion and is killed by X. The conjugate right-eigenvector rows over the positive spectrum are also orthonormal. Equivalently, any row-space vector orthogonal to the positive Gram images is killed by $X^\dagger$. □

Lemma 4.31. *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. Expand the square and apply 4.30 and 4.29 term by term:

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

while

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

and $X_a^\dagger I_{m\times m} X_a = X_a^\dagger X_a = Q_a$. Substituting these identities into

$$(X \cdot X^\dagger - I_{m\times m})^2 = X \cdot X^\dagger \cdot X \cdot X^\dagger - 2X \cdot X^\dagger + I_{m\times m}$$

gives the claim. □

Lemma 4.32 (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. Since T is projective,

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

□

Lemma 4.33 ($\widehat{X}$ squared)**.**

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

Proof. The paper proves this by multiplying out $\widehat{X} = U \cdot I_{m\times d} \cdot V^\dagger$ and using the unitarity of U and V. In Lean, the dimension hypothesis $m \le d$ already gives a rectangular matrix with orthonormal rows on the sigma auxiliary space; the chosen matrix `sigmaFinXHatCoisometry` and its specification record this existence result as a hypothesis for the future construction. The formal lemma `orthonormal_normalized_image_of_adjoint_comp_eigenvectors` records the singular-vector normalization calculation, while `normalized_matrix_image_rows_mul_conjTranspose` combines the corresponding row-coisometry identity, and `positive_gram_spectrum_image_rows_transpose_mixed` records the adjoint action $X^\dagger(\lambda^{-1/2} X v) = \lambda^{1/2} v$ on the positive spectral part of the Gram operator. The formal lemma `exists_unitaryGroup_with_positive_gram_spectrum_rows` records the next finite-dimensional step: after choosing distinct row positions for the positive spectrum, these normalized image rows extend to a square unitary group element on the row space. The companion lemma `exists_rectangular_coisometry_extending_orthonormal_rows` records the analogous rectangular completion for prescribed right rows when the dimension bound permits it. The formal lemma `unitaryGroup_mul_rectangular_coisometry` and its right-multiplication companion `rectangular_coisometry_mul_conjTranspose_unitaryGroup` record the two unitary-preservation steps in the paper's multiplication. The formal lemma `rectangularSvd_xHat_coisometry_unitaryGrou` applies these preservation steps to the rectangular-SVD expression. The unconditional positive-Gram route is `exists_xHat_of_sigmaFinRangeEmbedding_positiveGram`, which constructs $\widehat{X}$ from the rank-reduction witness and proves $\widehat{X}\widehat{X}^\dagger = I$ together with the mixed square-root identity below. □

Lemma 4.34 (X times $\widehat{X}$)**.**

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

Proof. The paper obtains this from the SVD formula for X and the definition of $\widehat{X}$. In Lean, the unconditional QXP data first proves that $X\widehat{X}^\dagger$ is Hermitian and positive. The formal lemma `x_mul_xHat_adjoint_spectral_theorem` then applies Mathlib's spectral theorem for Hermitian matrices, giving the displayed $U\Sigma U^\dagger$ form with the Mathlib eigenvector unitary and the diagonal matrix of real eigenvalues. The auxiliary lemma `rectangularSvd_x_mul_xHat_conjTranspose_raw` records the same calculation under explicit rectangular-SVD data, where the middle factor $SI^\dagger_{m\times d}$ is the formal square matrix $\Sigma_{m\times m}$. The formal lemma `rectangularSvd_xHat_mixed_of_sqrtQ_unitaryGroup` records the second matrix calculation in the form whose square-root target is the operator Q supplied by the QXP construction. The formal lemma `rectangularSvd_middle_eq_sqrt_of_square` records the spectral step: a positive middle factor whose square is Q is the positive square root of Q. The reduced datum then records this mixed-product identity as a primitive consequence, which is exactly the part needed in 4.36 and 4.38. □

Remark 4.35 (Lean status for the QXP algebra). The Lean declarations linked in Lemmas 4.30, 4.33, and 4.34 expose the corresponding matrix identities as construction theorems. These declarations prove the identities from the Q-layer, the sigma-range realization, and the positive-Gram spectral construction, with the rectangular-SVD calculation retained as an auxiliary comparison. Thus the local $X/\widehat{X}/P$ algebra is formalized; the remaining chapter-level work is to compose this construction with the preceding orthonormalization step from the paper hypotheses.

Lemma 4.36 (Squared difference)**.**

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

Proof. The paper diagonalizes the positive operator $X\widehat{X}^\dagger$ using the SVD notation. The formal proof avoids storing that SVD explicitly: set $Y = X\widehat{X}^\dagger$. The primitive identities imply that Y is hermitian, nonnegative, and satisfies $Y^2 = XX^\dagger$. Hence $(X - \widehat{X})(X - \widehat{X})^\dagger = (Y - I)^2$, and positivity of Y gives $(Y - I)^2 \leq (Y^2 - I)^2 = (XX^\dagger - I)^2$. □

Lemma 4.37 (Projectivity of P)**.** *$P = \{P_a\}$ forms a projective sub-measurement.*

Proof. For outcomes a, b,

$$\begin{aligned}
P_a \cdot P_b &= (\widehat{X}^\dagger \cdot T_a \cdot \widehat{X})(\widehat{X}^\dagger \cdot T_b \cdot \widehat{X}) \\
&= \widehat{X}^\dagger \cdot T_a \cdot I_{m\times m} \cdot T_b \cdot \widehat{X} \\
&= \widehat{X}^\dagger \cdot T_a \cdot \widehat{X} \cdot \mathbf{1}[[]\, a = b] \\
&= P_a \cdot \mathbf{1}[[]\, a = b],
\end{aligned}$$

where the second line uses 4.33, the third uses that T is projective, and the last uses 4.32. Summing over a shows $\sum_a P_a \leq I$, so P is a projective sub-measurement. The formalization also records the sharper identity

$$\sum_a P_a = \widehat{X}^\dagger \widehat{X},$$

and the resulting total-domination estimates used later in the monotone-total self-improvement route. ☐

Lemma 4.38 (P is close to Q)**.**

$$Q_a \otimes I \approx_{30\zeta^{1/4}} P_a \otimes I.$$

Proof. The formal proof follows this argument directly from the primitive X, $\widehat{X}$, and P identities; the $30\zeta^{1/4}$ estimate is derived here rather than stored as part of the reduced datum. The unitary-group rectangular-SVD lemma combines the same estimate with the positive-square identification of the middle factor from 4.34; the square unitary factors are again carried as elements of the relevant unitary groups. The positive-Gram sigma-space lemma also records the coisometry $XX^\dagger = I$ for the original sigma embedding when the projective Q-family is subnormalized; this is the construction-level hypothesis later used to obtain the QXP-internal fresh-outcome comparison $Q_{\text{none}} \leq P_{\text{none}}$. An additional source-to-Q comparison is still needed for source-facing residual domination. The quantity to bound is

$$\sum_a \langle\psi|(Q_a - P_a)^2 \otimes I|\psi\rangle = \langle\psi|Q \otimes I|\psi\rangle + \langle\psi|P \otimes I|\psi\rangle - \sum_a \langle\psi|Q_a P_a \otimes I|\psi\rangle - \sum_a \langle\psi|P_a Q_a \otimes I|\psi\rangle, \tag{4.11}$$

using the projectivity of Q and P. The first term is at most $1 + 2\sqrt{\zeta}$ by 4.20, and the second is at most 1 because P is a sub-measurement.

For the last two terms,

$$\sum_a \langle\psi|Q_a P_a \otimes I|\psi\rangle + \sum_a \langle\psi|P_a Q_a \otimes I|\psi\rangle = 2 \cdot \Re\Big(\sum_a \langle\psi|Q_a P_a \otimes I|\psi\rangle\Big). \tag{4.12}$$

By 4.29,

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

The main estimate is

$$\sum_a \langle\psi|(X_a^\dagger \cdot X \cdot P_a) \otimes I|\psi\rangle \approx_{2\zeta^{1/4}} \sum_a \langle\psi|(X_a^\dagger \cdot \widehat{X} \cdot P_a) \otimes I|\psi\rangle. \tag{4.13}$$

This follows by Cauchy–Schwarz:

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

is bounded by the square root of

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

times the square root of $\sum_a \langle\psi|(P_a)^2 \otimes I|\psi\rangle \leq 1$. The first factor is at most $4\sqrt{\zeta}$ by 4.36, 4.31, and 4.23, so the total error is at most $2\zeta^{1/4}$.

The right-hand side of 4.13 is real, because

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

where the first equality uses 4.28, 4.32, and 4.33 and that T is a measurement, and the second uses 4.34. Hence 4.22 and 4.13 imply

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

Substituting this into 4.12 and then into 4.11 gives

$$\sum_a \langle\psi|(Q_a - P_a)^2 \otimes I|\psi\rangle \leq (1 + 2\sqrt{\zeta}) + 1 - 2(1 - 14\zeta^{1/4}) = 2\sqrt{\zeta} + 28\zeta^{1/4} \leq 30\zeta^{1/4}.$$

□

Finally, 4.20 gives

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

and 4.38 gives

$$Q_a \otimes I \approx_{30\zeta^{1/4}} P_a \otimes I.$$

The triangle inequality for $\approx_\delta$ yields

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

□

Chapter 5

Expansion in the hypercube graph

5.1 The graph and its spectrum

Definition 5.1 (Rerandomizing one coordinate)**.** For $u \in \mathbb{F}_q^m$, an index $i \in \{1, \dots, m\}$, and $x \in \mathbb{F}_q$, let

$$\mathrm{rerand}_i(u, x) = u + xe_i.$$

This is the move obtained by rerandomizing the i-th coordinate of u by a uniform additive shift. In Lean, the same edge distribution is represented as the push-forward of the uniform distribution on triples (u, i, x) by the map that replaces the i-th coordinate by x; since x is uniform, this is the same distribution as the additive-shift presentation. The corresponding weight-sum identity identifies this push-forward distribution with the explicit edge coefficient used in the matrix proof.

Definition 5.2 (Hypercube graph)**.** The hypercube graph $C = (V, E)$ has vertex set $V = \mathbb{F}_q^m$, and an edge between u and v whenever they differ in at most one coordinate. A random edge $(u, v) \sim C$ is sampled by drawing $u \sim \mathbb{F}_q^m$, $i \in \{1, \dots, m\}$, and $x \in \mathbb{F}_q$ uniformly and then setting $v = \mathrm{rerand}_i(u, x)$.

Definition 5.3 (Adjacency matrix and Laplacian)**.** Let $M = q^m$. The normalized adjacency matrix of C is

$$K = \mathbb{E}_{(u,v)\sim C} |u\rangle\langle v|,$$

and the Laplacian is

$$L = \frac{1}{M} I - K.$$

Lemma 5.4 (Laplacian rewrite)**.** *The Laplacian can be written as*

$$L = \frac{1}{2}\mathbb{E}_{(u,v)\sim C}(|u\rangle - |v\rangle)(\langle u| - \langle v|).$$

Proof. The two vertex marginals of a random edge are uniform on $\mathbb{F}_q^m$. Expanding the right-hand side yields $(1/M)I - K$. ☐

Lemma 5.5 (Average of a nontrivial additive character)**.** *For $\beta \in \mathbb{F}_q$,*

$$\mathbb{E}_{x\sim\mathbb{F}_q}\omega^{\mathrm{tr}(\beta x)} = \begin{cases} 1 & \text{if } \beta = 0, \\ 0 & \text{if } \beta \neq 0. \end{cases}$$

Proof. If $\beta = 0$ the character is constant. If $\beta \neq 0$, the map $x \mapsto \beta x$ permutes $\mathbb{F}_q$, so the average is the average of a nontrivial additive character, which vanishes. □

Lemma 5.6 (Orthogonality of finite-field characters)**.** *For $\alpha, \beta \in \mathbb{F}_q^m$,*

$$\frac{1}{q^m}\sum_{u\in\mathbb{F}_q^m}\omega^{\mathrm{tr}(u\cdot(\beta-\alpha))} = \begin{cases} 1 & \text{if } \alpha = \beta, \\ 0 & \text{if } \alpha \neq \beta. \end{cases}$$

Proof. This is Lemma 3.3 applied to $v = \beta - \alpha$. □

Lemma 5.7 (Eigenvectors of the hypercube graph)**.** *For each $\alpha \in \mathbb{F}_q^m$, define*

$$|\varphi_\alpha\rangle = \frac{1}{M^{1/2}} \cdot \sum_{u\in\mathbb{F}_q^m}\omega^{\mathrm{tr}(u\cdot\alpha)}|u\rangle.$$

Then the following two statements hold.

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 of α.*

Proof. First we prove 1. For $\alpha, \beta \in \mathbb{F}_q^m$,

$$\langle\varphi_\alpha \mid \varphi_\beta\rangle = \frac{1}{M}\sum_{u\in\mathbb{F}_q^m}\omega^{\mathrm{tr}(u\cdot(\beta-\alpha))} = \begin{cases} 1 & \text{if } \alpha = \beta, \\ 0 & \text{if } \alpha \neq \beta, \end{cases}$$

by Lemma 5.6. Thus the $|\varphi_\alpha\rangle$ form an orthonormal basis of $\mathbb{C}^V$.

Next we prove 2. For $\alpha \in \mathbb{F}_q^m$,

$$\begin{aligned} K \cdot |\varphi_\alpha\rangle &= \left(\mathbb{E}_{(u,v)\sim C}|u\rangle\langle v|\right)\cdot\left(\frac{1}{M^{1/2}}\cdot\sum_{u\in\mathbb{F}_q^m}\omega^{\mathrm{tr}(u\cdot\alpha)}|u\rangle\right) \\ &= \frac{1}{M^{1/2}}\cdot\mathbb{E}_{(u,v)\sim C}\omega^{\mathrm{tr}(v\cdot\alpha)}|u\rangle. \end{aligned} \tag{5.1}$$

By the edge-sampling rule from Definition 5.2, we may write $v = u + xe_i$, where i is uniformly random in $\{1, \dots, m\}$ and x is uniformly random in $\mathbb{F}_q$. Therefore

$$(5.1) = \frac{1}{M^{1/2}}\cdot\mathbb{E}_{u,i,x}\omega^{\mathrm{tr}((u+xe_i)\cdot\alpha)}|u\rangle = \left(\mathbb{E}_{i,x}\omega^{\mathrm{tr}(x\alpha_i)}\right)\cdot\frac{1}{M^{1/2}}\cdot\mathbb{E}_u\omega^{\mathrm{tr}(u\cdot\alpha)}|u\rangle = \frac{1}{M}\left(\mathbb{E}_{i,x}\omega^{\mathrm{tr}(x\alpha_i)}\right)\cdot|\varphi_\alpha\rangle.$$

Hence $|\varphi_\alpha\rangle$ is an eigenvector of K with eigenvalue

$$\frac{1}{M}\cdot\mathbb{E}_i\left[\mathbb{E}_x\omega^{\mathrm{tr}(x\alpha_i)}\right] = \frac{1}{M}\cdot\mathbb{E}_i\left[\mathbf{1}[[]\,\alpha_i = 0]\right] = \frac{1}{M}\cdot\frac{m-|\alpha|}{m},$$

by Lemma 5.5. □

Lemma 5.8 (Spectral gap of the Laplacian)**.** *If $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_{q^m}$ are the eigenvalues of L, then*

$$\lambda_1 = 0, \qquad \lambda_2 = \frac{1}{mq^m}.$$

Proof. Lemma 5.7 identifies the two largest eigenvalues of K, hence the two smallest eigenvalues of $L = (1/q^m)I - K$. □

5.2 Local and global variance

Definition 5.9 (Local and global variance)**.** Let $|\psi\rangle \in \mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}$ and let $0 \le A^u \le I$ be an operator for each $u \in \mathbb{F}_q^m$. The local variance is

$$\mathbf{Var}_{\mathrm{local}}(A, \psi) = \frac{1}{2}\mathbb{E}_{(u,v)\sim C}\langle\psi|(A^u - A^v)^2 \otimes I|\psi\rangle,$$

and the global variance is

$$\mathbf{Var}_{\mathrm{global}}(A, \psi) = \frac{1}{2}\mathbb{E}_{u,v\sim\mathbb{F}_q^m}\langle\psi|(A^u - A^v)^2 \otimes I|\psi\rangle.$$

Lemma 5.10 (Local variance as a Laplacian form)**.** *Define*

$$A_{\mathrm{comb}} = \sum_{u\in\mathbb{F}_q^m} |u\rangle \otimes A^u \otimes I.$$

Then

$$\mathrm{Tr}(A_{\mathrm{comb}}^\dagger(L \otimes |\psi\rangle\langle\psi|)A_{\mathrm{comb}}) = \mathbf{Var}_{\mathrm{local}}(A, \psi).$$

Proof. For $u, v \in \mathbb{F}_q^m$,

$$\begin{aligned}(((\langle u| - \langle v|) \otimes \langle\psi|) \cdot A_{\mathrm{comb}} &= (((\langle u| - \langle v|) \otimes \langle\psi|) \cdot \sum_{w\in\mathbb{F}_q^m} |w\rangle \otimes A^w \otimes I \\ &= \langle\psi| \cdot ((A^u - A^v) \otimes I).\end{aligned} \tag{5.2}$$

Therefore

$$A_{\mathrm{comb}}^\dagger(L \otimes |\psi\rangle\langle\psi|)A_{\mathrm{comb}} = \frac{1}{2}\mathbb{E}_{(u,v)\sim C}A_{\mathrm{comb}}^\dagger((|u\rangle - |v\rangle)(\langle u| - \langle v|) \otimes |\psi\rangle\langle\psi|)A_{\mathrm{comb}}$$

by Lemma 5.4, and hence

$$A_{\mathrm{comb}}^\dagger(L \otimes |\psi\rangle\langle\psi|)A_{\mathrm{comb}} = \frac{1}{2}\mathbb{E}_{(u,v)\sim C}((A^u - A^v) \otimes I)|\psi\rangle\langle\psi|((A^u - A^v) \otimes I)$$

by (5.2). Taking the trace gives

$$\mathrm{Tr}(A_{\mathrm{comb}}^\dagger(L \otimes |\psi\rangle\langle\psi|)A_{\mathrm{comb}}) = \frac{1}{2}\mathbb{E}_{(u,v)\sim C}\langle\psi|(A^u - A^v)^2 \otimes I|\psi\rangle = \mathbf{Var}_{\mathrm{local}}(A, \psi).$$

□

Lemma 5.11 (Global variance as the orthogonal Fourier mass)**.** *Write*

$$A_{\mathrm{comb}} = |\varphi_0\rangle \otimes A_0 + |\varphi_\perp\rangle \otimes A_\perp,$$

where $|\varphi_\perp\rangle$ is orthogonal to $|\varphi_0\rangle$. Then

$$\frac{1}{q^m}\mathrm{Tr}((\langle\varphi_\perp| \otimes A_\perp)(I \otimes |\psi\rangle\langle\psi|)(|\varphi_\perp\rangle \otimes A_\perp)) = \mathbf{Var}_{\mathrm{global}}(A, \psi).$$

Proof. We first compute

$$A_0 = (\langle \varphi_0| \otimes I) \cdot A_{\text{comb}} = \left(\frac{1}{M^{1/2}} \cdot \sum_{u \in \mathbb{F}_q^m} \langle u|\right) \otimes I \cdot \left(\sum_{u \in \mathbb{F}_q^m} |u\rangle \otimes A^u \otimes I\right) = \frac{1}{M^{1/2}} \cdot \sum_{u \in \mathbb{F}_q^m} A^u \otimes I.$$

Writing $A_{\text{avg}} = \mathbb{E}_u A^u$, it follows that

$$|\varphi_0\rangle \otimes A_0 = \sum_{u \in \mathbb{F}_q^m} |u\rangle \otimes A_{\text{avg}} \otimes I, \qquad |\varphi_\perp\rangle \otimes A_\perp = \sum_{u \in \mathbb{F}_q^m} |u\rangle \otimes (A^u - A_{\text{avg}}) \otimes I.$$

Hence

$$\frac{1}{M} \cdot \text{Tr}((\langle \varphi_\perp| \otimes A_\perp)(I \otimes |\psi\rangle\langle\psi|)(|\varphi_\perp\rangle \otimes A_\perp)) = \mathbb{E}_{u \sim \mathbb{F}_q^m} \langle \psi|(A^u - A_{\text{avg}})^2 \otimes I|\psi\rangle. \tag{5.3}$$

Moreover,

$$\mathbb{E}_{u \sim \mathbb{F}_q^m} (A^u - A_{\text{avg}})^2 = \mathbb{E}_{u \sim \mathbb{F}_q^m} ((A^u)^2 - (A_{\text{avg}})^2) = \frac{1}{2} \mathbb{E}_{u,v \sim \mathbb{F}_q^m} (A^u - A^v)^2.$$

Substituting this identity into (5.3) yields

$$\frac{1}{M} \cdot \text{Tr}((\langle \varphi_\perp| \otimes A_\perp)(I \otimes |\psi\rangle\langle\psi|)(|\varphi_\perp\rangle \otimes A_\perp)) = \frac{1}{2} \mathbb{E}_{u,v \sim \mathbb{F}_q^m} \langle \psi|(A^u - A^v)^2 \otimes I|\psi\rangle = \mathbf{Var}_{\text{global}}(A, \psi).$$

□

Lemma 5.12 (Local-to-global inequality)**.** *For every family* $\{A^u\}$*,*

$$\mathbf{Var}_{\text{global}}(A, \psi) \le m \, \mathbf{Var}_{\text{local}}(A, \psi).$$

Proof.

$$\begin{aligned} A_{\text{comb}}^\dagger (L \otimes |\psi\rangle\langle\psi|) A_{\text{comb}} &= ((\langle \varphi_0| \otimes A_0 + \langle \varphi_\perp| \otimes A_\perp)(L \otimes |\psi\rangle\langle\psi|)(|\varphi_0\rangle \otimes A_0 + |\varphi_\perp\rangle \otimes A_\perp) \\ &= ((\langle \varphi_\perp| \otimes A_\perp)(L \otimes |\psi\rangle\langle\psi|)(|\varphi_\perp\rangle \otimes A_\perp) \\ &= \langle \varphi_\perp|L|\varphi_\perp\rangle \cdot A_\perp |\psi\rangle\langle\psi| A_\perp. \end{aligned} \tag{5.4}$$

The second step uses that $|\varphi_0\rangle$ is a 0-eigenvector of L. Since $|\varphi_\perp\rangle$ is orthogonal to $|\varphi_0\rangle$, Lemma 5.8 gives

$$\langle \varphi_\perp|L|\varphi_\perp\rangle \ge \frac{1}{mM} \cdot \langle \varphi_\perp \mid \varphi_\perp\rangle.$$

Therefore

$$\mathbf{Var}_{\text{local}}(A, \psi) = \text{Tr}(A_{\text{comb}}^\dagger (L \otimes |\psi\rangle\langle\psi|) A_{\text{comb}}) = \langle \varphi_\perp|L|\varphi_\perp\rangle \cdot \text{Tr}(A_\perp |\psi\rangle\langle\psi| A_\perp)$$

by Lemma 5.10 and (5.4), so

$$\mathbf{Var}_{\text{local}}(A, \psi) \ge \frac{1}{mM} \cdot \text{Tr}((\langle \varphi_\perp| \otimes A_\perp)(I \otimes |\psi\rangle\langle\psi|)(|\varphi_\perp\rangle \otimes A_\perp)) = \frac{1}{m} \mathbf{Var}_{\text{global}}(A, \psi),$$

where the last step is Lemma 5.11. Rearranging gives the claim. □

Chapter 6

Global variance of the points measurements

Throughout this chapter, (ψ, A, B, L) is an $(\varepsilon, \delta, \gamma)$-good symmetric strategy for the (m, q, d) low individual degree test.

Lemma 6.1 (Replacing a line answer by a fixed polynomial)**.** *Let $G \in \mathrm{PolySub}(m, q, d)$. On the axis-parallel line-test distribution,*

$$B^{\ell}_{[f(u)=g(u)]} \otimes (G_g)^{1/2} \approx_{md/q} B^{\ell}_{g|_\ell} \otimes (G_g)^{1/2}.$$

Proof. Expand the $\approx_{md/q}$ norm on the weighted vector $(I \otimes (G_g)^{1/2})|\psi\rangle$. Because B^ℓ is a projective measurement, only those line polynomials f with $f \neq g|_\ell$ and $f(u) = g(u)$ contribute. Lemma 3.9 is the univariate Schwartz–Zippel wrapper used in Lean for this line-restriction event; it supplies the paper's md/q loss, and the remaining operator sum is bounded by the fact that G is a submeasurement. $\square$

Remark 6.2 (Conditional variants for the collision-residual estimate)*.* The paper proof of Lemma 6.1 bounds the collision event directly by the Schwartz–Zippel estimate. The following conditional variants instead assume this collision estimate, or one of the reindexing identities that implies it, as an explicit hypothesis. These conditional statements make explicit the intermediate estimates still needed for this route; they are not additional hypotheses in the paper statement of Lemma 6.1.

Lemma 6.3 (Local variance of the points measurements)**.** *Let $G \in \mathrm{PolySub}(m, q, d)$. On the hypercube edge distribution $(u, v) \sim C$,*

$$A^u_{g(u)} \otimes (G_g)^{1/2} \approx_{24\cdot\left(\varepsilon+\delta+\frac{md}{q}\right)} A^v_{g(v)} \otimes (G_g)^{1/2}. \tag{6.1}$$

Proof. Fix g and keep the weight $(G_g)^{1/2}$ on the second tensor factor throughout. On the vector $(I \otimes (G_g)^{1/2})|\psi\rangle$ one has the six-step chain

$$A^u_{g(u)}\otimes(G_g)^{1/2} \approx_{2\delta} I\otimes(G_g)^{1/2}A^u_{g(u)} \approx_{2\varepsilon} B^{\ell}_{[f(u)=g(u)]}\otimes(G_g)^{1/2} \approx_{md/q} B^{\ell}_{g|_\ell}\otimes(G_g)^{1/2} \approx_{md/q} B^{\ell}_{[f(v)=g(v)]}\otimes(G_g)^{1/2} \approx_{2\varepsilon} I\otimes(G$$

The first, second, fifth, and sixth comparisons use Lemma 3.21 together with Theorem 3.25 and the good-strategy hypotheses, and the middle pair is Lemma 6.1. Lemma 3.27 sums these six weighted transports. $\square$

Lemma 6.4 (Equivalent local-variance bound)**.** *The local-variance bound for the point measurements is*

$$\sum_{g\in\mathcal{P}(m,q,d)} \mathbb{E}_{(u,v)\sim C}\langle\psi|(A^u_{g(u)} - A^v_{g(v)})^2 \otimes G_g|\psi\rangle \le 24\left(\varepsilon+\delta+\frac{md}{q}\right). \tag{6.2}$$

Proof. This is the squared-norm expansion of Lemma 6.3 after summing over the polynomial outcome g. The Lean proof expands the weighted operator difference, telescopes the six transports in the local variance proof, sums the resulting bounds over g, and absorbs the transport-chain error into $24\left(\varepsilon+\delta+\frac{md}{q}\right)$. □

Remark 6.5 (Polynomial-sum formalization for 6.2)*.* The three referenced declarations and Lemma 6.4 together provide the full polynomial-sum (cardinality-free) chain proving the equation. The chain assembly lemma `localVarianceDeviation_sum_le_localVarianceOfPointsError` telescopes the six operator differences via `ev_sum_conjTranspose_mul_sum_le`, sums over g, and applies the six individual `_sum` bounds (including the reverse `generalizeB` sum) to avoid the per-polynomial cardinality blow-up.

Lemma 6.6 (Global variance of the points measurements)**.** *Let* $G \in \mathrm{PolySub}(m,q,d)$*. On the uniform distribution over independent* $u,v\in\mathbb{F}_q^m$*,*

$$A^u_{g(u)} \otimes (G_g)^{1/2} \approx_{24m\left(\varepsilon+\delta+\frac{md}{q}\right)} A^v_{g(v)} \otimes (G_g)^{1/2}. \tag{6.3}$$

Proof. We want to bound

$$\mathbb{E}_{u,v\sim\mathbb{F}_q^m} \sum_{g\in\mathcal{P}(m,q,d)} \|(A^u_{g(u)} - A^v_{g(v)}) \otimes (G_g)^{1/2}|\psi\rangle\|^2 = \mathbb{E}_{u,v\sim\mathbb{F}_q^m} \sum_{g\in\mathcal{P}(m,q,d)} \langle\psi|(A^u_{g(u)} - A^v_{g(v)})^2 \otimes G_g|\psi\rangle. \tag{6.4}$$

For each $g\in\mathcal{P}(m,q,d)$, define

$$A(g)^u := A^u_{g(u)}, \qquad |\psi_g\rangle := (I\otimes(G_g)^{1/2})|\psi\rangle.$$

Then

$$(6.4) = \sum_{g\in\mathcal{P}(m,q,d)} \mathbb{E}_{u,v\sim\mathbb{F}_q^m}\langle\psi_g|(A(g)^u - A(g)^v)^2\otimes I|\psi_g\rangle = \sum_{g\in\mathcal{P}(m,q,d)} 2\,\mathbf{Var}_{\mathrm{global}}(A(g),\psi_g).$$

Applying Lemma 5.12 gives

$$(6.4) \le \sum_{g\in\mathcal{P}(m,q,d)} 2m\,\mathbf{Var}_{\mathrm{local}}(A(g),\psi_g) = m\cdot \sum_{g\in\mathcal{P}(m,q,d)} \mathbb{E}_{(u,v)\sim C}\langle\psi|(A^u_{g(u)} - A^v_{g(v)})^2\otimes G_g|\psi\rangle.$$

The last quantity is at most $m\cdot 24\left(\varepsilon+\delta+\frac{md}{q}\right)$ by Lemma 6.4, which is exactly 6.3. □

Chapter 7

Self-improvement

Throughout this chapter, (ψ, A, B, L) is an $(\varepsilon, \delta, \gamma)$-good symmetric strategy for the (m, q, d) low individual degree test.

7.1 Non-projective Output

A large part of this chapter is devoted to proving Lemma 7.1, which is a slightly weaker form of the projective self-improvement theorem proved in Section 7.2. The key difference is that the output family H is only required to be a submeasurement, rather than a projective submeasurement. Once this non-projective statement is available, Theorem 4.4 upgrades H to a projective family without losing control of the four quantitative conclusions.

Compared with the projective theorem, the helper lemma records strong self-consistency using the explicit inequality from Definition 3.34, keeps the boundedness conclusion in terms of an auxiliary operator Z, and produces a much smaller error parameter.

Lemma 7.1 (Self-improvement with non-projective output)**.** *Let $G \in \mathrm{PolyMeas}(m, q, d)$ be a measurement with the following property:*

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

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

Let

$$\zeta = 100m \cdot \left(\varepsilon^{1/2} + \delta^{1/2} + (d/q)^{1/2}\right).$$

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

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

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

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

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

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

$$\langle\psi|Z\otimes I|\psi\rangle - \mathbb{E}_u \sum_a \langle\psi|A_a^u \otimes H_{[h(u)=a]}|\psi\rangle \le \zeta$$

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

$$Z \ge \left(\mathbb{E}_u A^u_{h(u)}\right).$$

Proof. Let $T = \{T_g\}$ and Z be the optimal solutions to the SDPs (7.23) and (7.24) given by Lemma 7.8. Then T is a measurement, and

$$\forall g \in \mathcal{P}(m,q,d): \qquad Z \ge \left(\mathbb{E}_u A^u_{g(u)}\right), \tag{7.1}$$

$$T_g \cdot Z = T_g \cdot \left(\mathbb{E}_u A^u_{g(u)}\right). \tag{7.2}$$

For each $u \in \mathbb{F}_q^m$, define $H^u = \{H_h^u\}_{h\in\mathcal{P}(m,q,d)}$ by

$$H_h^u := A^u_{h(u)} \cdot T_h \cdot A^u_{h(u)},$$

and let

$$H_h := \mathbb{E}_u H_h^u.$$

The pointwise family H^u is a submeasurement because

$$\sum_h H_h^u = \sum_a A_a^u \cdot \Big(\sum_{h:h(u)=a} T_h\Big)\cdot A_a^u \le \sum_a (A_a^u)^2 = I,$$

where the inequality uses that T is a measurement and the last identity uses that A is projective. Averaging over u gives

$$\sum_h H_h = \mathbb{E}_u \sum_h H_h^u \le I,$$

so H is also a submeasurement and therefore lies in $\mathrm{PolySub}(m,q,d)$.

Set

$$\zeta_{\text{variance}} = 24m\cdot\Big(\varepsilon+\delta+\frac{md}{q}\Big).$$

Lemma 7.9 below is the technical transfer from the averaged family H back to the sandwiched operators $A^u_{h(u)}T_hA^u_{h(u)}$. We use it in the completeness, consistency, strong self-consistency, and boundedness estimates that follow.

We now verify the four conclusions of Lemma 7.1.

Proof of 1. The completeness of H is

$$\begin{aligned}\sum_h \langle\psi|H_h\otimes I|\psi\rangle &= \mathbb{E}_u\sum_h\langle\psi|H_h^u\otimes I|\psi\rangle\\ &= \mathbb{E}_u\sum_h\langle\psi|(A^u_{h(u)}\cdot T_h\cdot A^u_{h(u)})\otimes I|\psi\rangle.\end{aligned}\tag{7.3}$$

Grouping the polynomials according to their value at u gives

$$\mathbb{E}_u\sum_h\langle\psi|(A^u_{h(u)}\cdot T_h\cdot A^u_{h(u)})\otimes I|\psi\rangle = \mathbb{E}_u\sum_a\langle\psi|(A^u_a\cdot T_{[h(u)=a]}\cdot A^u_a)\otimes I|\psi\rangle. \tag{7.4}$$

We first move the leftmost copy of A^u_a across the bipartition:

$$(7.4) \approx_{2\sqrt{\delta}} \mathbb{E}_u\sum_a\langle\psi|(T_{[h(u)=a]}\cdot A^u_a)\otimes A^u_a|\psi\rangle. \tag{7.5}$$

The difference is bounded by Cauchy–Schwarz as in (7.33); the first factor is at most $\sqrt{2\delta}$ by self-consistency of A, and the second is at most 1 because $T_{[h(u)=a]} \leq I$. We next remove the remaining copy of A^u_a on Bob's side:

$$(7.5) \approx_{\sqrt{\delta}} \mathbb{E}_u \sum_a \langle\psi|(T_{[h(u)=a]} \cdot A^u_a) \otimes I|\psi\rangle. \tag{7.6}$$

The first Cauchy–Schwarz factor is at most 1 because the fiber operators form a submeasurement after grouping T by the value of $h(u)$. Here the second Cauchy–Schwarz factor is

$$\mathbb{E}_u \sum_{a\neq b} \langle\psi|A^u_a \otimes A^u_b|\psi\rangle,$$

which is at most δ because A is self-consistent and projective. Reindexing by h and using (7.2),

$$(7.6) = \sum_h \langle\psi|(T_h \cdot \mathbb{E}_u A^u_{h(u)}) \otimes I|\psi\rangle = \sum_h \langle\psi|(T_h \cdot Z) \otimes I|\psi\rangle = \langle\psi|Z \otimes I|\psi\rangle.$$

Thus

$$\sum_h \langle\psi|H_h \otimes I|\psi\rangle \geq \langle\psi|Z \otimes I|\psi\rangle - 3\sqrt{\delta}. \tag{7.7}$$

Since G is a measurement, (7.1) and Lemma 3.18 give

$$\langle\psi|Z \otimes I|\psi\rangle \geq \mathbb{E}_u \sum_a \langle\psi|A^u_a \otimes G_{[g(u)=a]}|\psi\rangle \geq 1-\nu.$$

Therefore the completeness is at least $1-\nu-3\sqrt{\delta}$, which is bounded below by $(1-\nu)-\zeta$.

Proof of 2. The inconsistency of H with A is

$$\mathbb{E}_u \sum_{a\neq b} \langle\psi|A^u_a \otimes H_{[h(u)=b]}|\psi\rangle = \mathbb{E}_u \sum_{a,h:h(u)\neq a} \langle\psi|A^u_a \otimes H_h|\psi\rangle. \tag{7.8}$$

The theorem-side off-diagonal selection used for this ‘add-in-u‘ application is , and the corresponding left/right scalar quantities are and . The selected scalar chain for this specialization is recorded by . Apply Lemma 7.9 with $\mathcal{O} = \mathbb{F}_q$, $M = A$, and $S_u = \{(a,h) : h(u) \neq a\}$. Then

$$(7.8) \approx_{4\sqrt{\zeta_{\text{variance}}}} \mathbb{E}_u \sum_{a,h:h(u)\neq a} \langle\psi|(A^u_{h(u)} \cdot A^u_a \cdot A^u_{h(u)}) \otimes T_h|\psi\rangle,$$

and the right-hand side is 0 because A is projective. Hence

$$\mathbb{E}_u \sum_{a\neq b} \langle\psi|A^u_a \otimes H_{[h(u)=b]}|\psi\rangle \leq 4\sqrt{\zeta_{\text{variance}}}. \tag{7.9}$$

Since

$$4\sqrt{\zeta_{\text{variance}}} = 4\sqrt{24m \cdot \left(\varepsilon+\delta+\frac{md}{q}\right)} \leq 20m \cdot \left(\varepsilon^{1/2}+\delta^{1/2}+(d/q)^{1/2}\right) \leq \zeta,$$

this proves 2.

Proof of 3. From the definition of H^u_h and the projectivity of A,

$$H^u_h = A^u_{h(u)} \cdot T_h \cdot A^u_{h(u)} = A^u_{h(u)} \cdot H^u_h \cdot A^u_{h(u)}, \tag{7.10}$$

$$A^u_{h(u)} \cdot H^u_{h'} \cdot A^u_{h(u)} = H^u_{h'} \cdot A^u_{h(u)} = (A^u_{h(u)} \cdot T_{h'} \cdot A^u_{h(u)}) \cdot \mathbf{1}[h(u)=h'(u)]\,. \tag{7.11}$$

The strong self-consistency expression is

$$\sum_{h\in\mathcal{P}(m,q,d)} \langle\psi|H_h\otimes H_h|\psi\rangle = \mathbb{E}_u \sum_{h\in\mathcal{P}(m,q,d)} \langle\psi|H^u_h\otimes H_h|\psi\rangle. \tag{7.12}$$

Apply Lemma 7.9 with $\mathcal{O}=\mathcal{P}(m,q,d)$, $M=H$, and $S_u=\{(h,h): h\in\mathcal{P}(m,q,d)\}$. This gives

$$(7.12) \approx_{4\sqrt{\zeta_{\text{variance}}}} \mathbb{E}_u \sum_{h\in\mathcal{P}(m,q,d)} \langle\psi|(A^u_{h(u)}\cdot H^u_h\cdot A^u_{h(u)})\otimes T_h|\psi\rangle. \tag{7.13}$$

We next enlarge the sum from $h'=h$ to all pairs h,h':

$$(7.13) \approx_{2\sqrt{\zeta_{\text{variance}}}+\frac{md}{q}} \mathbb{E}_u \sum_{h,h'\in\mathcal{P}(m,q,d)} \langle\psi|(A^u_{h(u)}\cdot H^u_{h'}\cdot A^u_{h(u)})\otimes T_h|\psi\rangle. \tag{7.14}$$

The added off-diagonal contribution is

$$\begin{aligned}(7.14)-(7.13) &= \mathbb{E}_u \sum_{h\neq h'} \langle\psi|(A^u_{h(u)}\cdot H^u_{h'}\cdot A^u_{h(u)})\otimes T_h|\psi\rangle \\ &= \mathbb{E}_u \sum_{h\neq h'} \langle\psi|(A^u_{h(u)}\cdot T_{h'}\cdot A^u_{h(u)})\otimes T_h|\psi\rangle\cdot\mathbf{1}[h(u)=h'(u)]\,,\end{aligned} \tag{7.15}$$

where the second equality is (7.11). To estimate (7.15), first replace the left copy of $A^u_{h(u)}$ by $A^v_{h(v)}$:

$$(7.15) \approx_{\sqrt{\zeta_{\text{variance}}}} \mathbb{E}_{u,v} \sum_{h\neq h'} \langle\psi|(A^v_{h(v)}\cdot T_{h'}\cdot A^u_{h(u)})\otimes T_h|\psi\rangle\cdot\mathbf{1}[h(u)=h'(u)]\,. \tag{7.16}$$

The corresponding Cauchy–Schwarz bound is

$$\begin{aligned}&\left|\mathbb{E}_{u,v} \sum_{h\neq h'} \langle\psi|((A^u_{h(u)}-A^v_{h(v)})\cdot T_{h'}\cdot A^u_{h(u)})\otimes T_h|\psi\rangle\cdot\mathbf{1}[h(u)=h'(u)]\right| \\ &\qquad\leq \sqrt{\mathbb{E}_{u,v} \sum_{h\neq h'} \langle\psi|((A^u_{h(u)}-A^v_{h(v)})\cdot T_{h'}\cdot(A^u_{h(u)}-A^v_{h(v)}))\otimes T_h|\psi\rangle} \\ &\qquad\qquad\cdot\sqrt{\mathbb{E}_{u,v} \sum_{h\neq h'} \langle\psi|(A^u_{h(u)}\cdot T_{h'}\cdot A^u_{h(u)})\otimes T_h|\psi\rangle\cdot\mathbf{1}[h(u)=h'(u)]}.\end{aligned} \tag{7.17}$$

The first factor is bounded by $\sqrt{\zeta_{\text{variance}}}$ via Lemma 6.6, and the second by 1 because T and H^u are submeasurements.

Repeating the same argument for the remaining copy of A yields

$$(7.16) \approx_{\sqrt{\zeta_{\text{variance}}}} \mathbb{E}_{u,v} \sum_{h\neq h'} \langle\psi|(A^v_{h(v)}\cdot T_{h'}\cdot A^v_{h(v)})\otimes T_h|\psi\rangle\cdot\mathbf{1}[h(u)=h'(u)]\,. \tag{7.18}$$

The same intermediate estimate gives

$$\mathbb{E}_{u,v} \sum_{h\neq h'} \langle\psi|(A^v_{h(v)}\cdot T_{h'}\cdot A^v_{h(v)})\otimes T_h|\psi\rangle \leq 1, \tag{7.19}$$

so Schwartz–Zippel bounds the indicator average by md/q and proves (7.14).

Using (7.11), the right-hand side of (7.14) is

$$\mathbb{E}_u \sum_{h,h' \in \mathcal{P}(m,q,d)} \langle \psi | (H^u_{h'} \cdot A^u_{h(u)}) \otimes T_h | \psi \rangle. \tag{7.20}$$

Replace the remaining copy of $A^u_{h(u)}$ by $A^v_{h(v)}$:

$$(7.20) \approx_{\sqrt{\zeta_{\text{variance}}}} \mathbb{E}_{u,v} \sum_{h,h' \in \mathcal{P}(m,q,d)} \langle \psi | (H^u_{h'} \cdot A^v_{h(v)}) \otimes T_h | \psi \rangle. \tag{7.21}$$

Here the first Cauchy–Schwarz factor is at most 1 because T and H^u are submeasurements, and the second is again bounded by Lemma 6.6. Finally move this last copy of $A^v_{h(v)}$ to Bob's side:

$$(7.21) \approx_{\sqrt{2\delta}} \mathbb{E}_{u,v} \sum_{h,h' \in \mathcal{P}(m,q,d)} \langle \psi | H^u_{h'} \otimes (T_h \cdot A^v_{h(v)}) | \psi \rangle. \tag{7.22}$$

Substituting (7.2), then (7.1), and finally the already proved bound (7.9), one obtains

$$(7.22) \geq \sum_h \langle \psi | H_h \otimes I | \psi \rangle - 4\sqrt{\zeta_{\text{variance}}}.$$

Combining (7.13), (7.14), (7.21), and (7.22) gives

$$\sum_h \langle \psi | H_h \otimes H_h | \psi \rangle \geq \sum_h \langle \psi | H_h \otimes I | \psi \rangle - 11\sqrt{\zeta_{\text{variance}}} - \sqrt{2\delta} - \frac{md}{q}.$$

The Lean formalization records this step as named internal bounds, starting with the structure and the exact residual-side expansion and then uses the proved lemma together with to recover the final helper-stage self-consistency conclusion. The theorem-level derivation is therefore not moved into an additional hypothesis. The paper's arithmetic shows that the right-hand error is at most

$$57m \cdot \left(\varepsilon^{1/2} + \delta^{1/2} + (d/q)^{1/2}\right) \leq \zeta,$$

which proves 3.

Proof of 4. The helper-stage boundedness quantity is

$$\langle \psi | Z \otimes I | \psi \rangle - \mathbb{E}_u \sum_a \langle \psi | A^u_a \otimes H_{[h(u)=a]} | \psi \rangle.$$

The reindexing of the sum by h is the algebraic identity

$$\sum_a A^u_a \otimes H_{[h(u)=a]} = \sum_h A^u_{h(u)} \otimes H_h,$$

whose averaged scalar form reads

$$\mathbb{E}_u \sum_a \langle \psi |\, A^u_a \otimes H_{[h(u)=a]} \,| \psi \rangle = \mathbb{E}_u \sum_h \langle \psi |\, A^u_{h(u)} \otimes H_h \,| \psi \rangle.$$

Using $\sum_a A^u_a = I$, the pointwise difference between the right-placed total $I \otimes H.\text{total} = \sum_h I \otimes H_h$ and the helper-agreement operator at u reindexes as the off-diagonal sum

$$\sum_h I \otimes H_h - \sum_a A^u_a \otimes H_{[h(u)=a]} = \sum_h \sum_{a \neq h(u)} A^u_a \otimes H_h,$$

whose averaged scalar form reads

$$\sum_h \langle\psi|\, I \otimes H_h \,|\psi\rangle - \mathbb{E}_u \sum_a \langle\psi|\, A_a^u \otimes H_{[h(u)=a]} \,|\psi\rangle = \mathbb{E}_u \sum_h \sum_{a \neq h(u)} \langle\psi|\, A_a^u \otimes H_h \,|\psi\rangle.$$

Combined with (7.9) this gives

$$\langle\psi|Z \otimes I|\psi\rangle - \mathbb{E}_u \sum_a \langle\psi|A_a^u \otimes H_{[h(u)=a]}|\psi\rangle \leq \langle\psi|Z \otimes I|\psi\rangle - \sum_h \langle\psi|I \otimes H_h|\psi\rangle + 4\sqrt{\zeta_{\text{variance}}}.$$

By (7.7), the right-hand side is at most $3\sqrt{\delta}+4\sqrt{\zeta_{\text{variance}}}$. Substituting the definition of ζ_{variance} and collecting terms gives

$$23m \cdot \left(\varepsilon^{1/2} + \delta^{1/2} + (d/q)^{1/2}\right) \leq \zeta.$$

Together with (7.1), this proves 4. □

Remark 7.2 (Slackness-carrying self-improvement helper). The formalization records a strengthened helper theorem whose output carries the complementary-slackness equations used in the completeness chain. The previous top-level residual-domination assembly theorems have been removed: the full self-improvement conclusion is now stated only by the paper theorem below. Its proof follows the paper's expectation-level total-difference transport route; issue #1642 records why the sharper generic operator-total strengthening is not supplied by the present route.

7.1.1 A Semidefinite Program

Lemma 7.3 (Uniform SDP feasible witness)**.** *The uniform primal family $T_g = (2|\mathcal{P}(m,q,d)|)^{-1}I$ has total mass $(1/2)I$, and the dual operator $Z = 2I$ is positive semidefinite and dominates the identity.*

Proof. This is the elementary Slater-type feasible witness recorded in the basic self-improvement definitions. □

Lemma 7.4 (Explicit matrix feasible bounds)**.** *In every concrete matrix realization, the uniform family $T_g = (2|\mathcal{P}(m,q,d)|)^{-1}I$ has total mass $(1/2)I$. The dual operator $Z = 2I$ dominates the identity and satisfies $Z - A_g \geq I$, hence $Z \geq A_g$, for every polynomial g. In the canonical block SDP, the corresponding block matrix is feasible, its slack block is $(1/2)I$, and the canonical dual slack for the same dual witness also dominates the block identity. More generally, dual feasibility implies positivity of Z, since the averaged point operators A_g are positive semidefinite.*

Proof. The formal proof is the matrix specialization of the uniform SDP witness recorded in the basic definitions. The inequality $A_g \leq I$ follows by averaging the point-measurement effects and using the submeasurement bound at each point. □

Lemma 7.5 (Matrix slackness output)**.** *From a matrix-level strong-duality witness with complementary slackness of the shape asserted by Lemma 7.8, one obtains a complete measurement $\{T_g\}$, a feasible dual operator Z, equality of the primal and dual objectives, and the equations $T_g Z = T_g A_g$ for all g. Lean isolates the saturated canonical optimal-pair output required from strong duality: a feasible canonical primal matrix, a dual-feasible operator with equal objective value, canonical complementary slackness, and vanishing of the extra canonical slack block. This datum gives the matrix-level slackness-carrying SDP statement, whose witness is already a complete matrix measurement, and is then transported to the abstract Section 9 statement without* *assuming $I \leq Z$. The displayed measurement witness is read directly from that abstract statement. The canonical saturation datum is now the only matrix-level input used to pass from strong duality to the abstract SDP statement. The paper-facing self-improvement helper consumes the resulting abstract SDP statement.*

Proof. The identity $\sum_g T_g = I$ turns the optimal primal submeasurement into a measurement; Lean performs this completion before storing the matrix and abstract slackness statements. The zero-defect form $T_g(Z - A_g) = 0$ is equivalent, by distributivity, to the displayed equation $T_g Z = T_g A_g$. □

Remark 7.6 (Removed residual-dominating construction theorems). Earlier Lean versions exposed top-level conditional construction theorems that combined the matrix slackness output with residual-dominating orthonormalization and QXP-repair inputs. These declarations have been removed. The matrix slackness material is retained only up to the slackness-carrying helper output, and the full projective-output conclusion is represented by the paper theorem `MIPStarRE.LDT.SelfImprovement.selfImprovement`.

Remark 7.7 (Removed reduced SDP interface). Earlier Lean versions also contained a reduced theorem `sdp` proving only the measurement-total and dual-feasibility fragment used before strong duality was supplied, as well as dominance-carrying declarations with the auxiliary condition $I \leq Z$. These declarations have been removed, and they are not advertised as formalizations of Lemma 7.8. The active Lean route for Lemma 7.8 is the slackness-carrying statement `sdp_statement_with_slackness`.

Lemma 7.8 (Dual SDP and complementary slackness)**.** *Let*

$$A_g = \mathbb{E}_{u \sim \mathbb{F}_q^m} A^u_{g(u)}.$$

Then the primal is

$$\begin{aligned} \sup \quad & \sum_g \mathrm{Tr}(T_g \cdot A_g) \\ s.t. \quad & T_g \geq 0 \qquad \forall g \in \mathcal{P}(m,q,d), \\ & \sum_g T_g \leq I, \end{aligned} \tag{7.23}$$

and the dual is

$$\inf \quad \mathrm{Tr}(Z) \tag{7.24}$$

$$s.t. \quad Z \geq A_g. \tag{7.25}$$

These semidefinite programs are dual to each other. Moreover there is an optimal pair of solutions $\{T_g\}$ to (7.23) and Z to (7.24) such that $\sum_g T_g = I$ and

$$T_g Z = T_g A_g, \qquad \forall g \in \mathcal{P}(m,q,d). \tag{7.26}$$

Proof. Rewrite (7.23) in canonical block form. Let r be the dimension of the space on which A acts, let $M = |\mathcal{P}(m,q,d)|$, and fix an ordering $g_1, \dots, g_M$ of the polynomials. Consider

$$\begin{aligned} \sup \quad & \mathrm{Tr}(C^\dagger X) \\ \text{s.t.} \quad & \mathrm{Tr}(D_{ij}^\dagger X) = b_{ij} \qquad \forall i,j \in \{1,\dots,r\}, \\ & X \geq 0, \end{aligned} \tag{7.27}$$

where

$$C = \sum_{i=1}^{M} |i\rangle\langle i| \otimes A_{g_i},$$

and, for $i, j \in \{1, \dots, r\}$,

$$D_{ij} = \sum_{k=1}^{M+1} |k\rangle\langle k| \otimes |i\rangle\langle j|, \qquad b_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Writing $X = \sum_{i,j=1}^{M+1} |i\rangle\langle j| \otimes X_{ij}$, the constraints say precisely that $\sum_{i=1}^{M+1} X_{ii} = I$, hence $\sum_{i=1}^{M} X_{ii} \leq I$, and the objective is

$$\mathrm{Tr}(C^\dagger X) = \sum_{i=1}^{M} \mathrm{Tr}(A_{g_i} \cdot X_{ii}).$$

Thus (7.27) is equivalent to (7.23) under the identification $T_{g_i} = X_{ii}$.

The canonical dual is

$$\inf \quad \sum_{i} z_{ij} b_{ij} \tag{7.28}$$

$$\text{s.t.} \quad \sum_{i,j} z_{ij} D_{ij} \geq C. \tag{7.29}$$

Since

$$\sum_{i,j=1}^{r} z_{ij} D_{ij} = \sum_{k=1}^{M+1} |k\rangle\langle k| \otimes \left(\sum_{i,j=1}^{r} z_{ij} |i\rangle\langle j| \right) =: \sum_{k=1}^{M+1} |k\rangle\langle k| \otimes Z,$$

the constraint (7.29) is exactly $Z \geq A_{g_i}$ for every i, and the objective is $\mathrm{Tr}(Z)$. In Lean, the latter statement is recorded as the trace-pairing identity

$$\Re\, \mathrm{Tr}(D(Z)X) = \Re\, \mathrm{Tr}(Z)$$

for every feasible canonical primal matrix X, where $D(Z)$ denotes the block-diagonal canonical dual operator. Consequently the canonical primal-dual gap is the real trace pairing of X with the canonical dual slack $D(Z) - C$. Since the real trace pairing of positive semidefinite operators is nonnegative, Lean also records the resulting weak-duality inequality for any feasible canonical primal and dual pair. Hence (7.28) is equivalent to (7.24).

Slater's condition holds because

$$T_g = \frac{1}{2M} \cdot I \qquad \forall g \in \mathcal{P}(m, q, d), \qquad Z = 2I$$

are strictly feasible for the primal and dual respectively. Therefore strong duality holds. For an optimal pair $(X, (z_{ij}))$ in canonical form, complementary slackness gives

$$X\Big(\sum_{i,j} z_{ij} D_{ij} - C \Big) \; = \; 0. \tag{7.30}$$

We may assume that X is block diagonal. Translating (7.30) back to the variables $\{T_g\}$ and Z yields $T_g(Z - A_g) = 0$ for every g and $SZ = 0$, where $S = I - \sum_g T_g$ is the slack block. The paper then treats this as forcing $S = 0$. The Lean development records this saturation as part of the canonical optimal-pair datum for the block SDP. The canonical calculation still uses the auxiliary expression `matrixSdpComplementarySlacknessDefect` to read off the polynomial blocks of (7.30). The matrix optimal-witness interface, however, stores the displayed equation $T_g Z = T_g A_g$ directly; `MatrixSdpOptimalWitness.primalMeasurement` expresses $\sum_g T_g = I$ as a complete measurement. The statement `MatrixSdpStatementWithSlackness` records this matrix-level strong-duality output and the abstract theorem `sdp_statement_with_slackness` specializes it to the paper's self-improvement setting. The explicit Slater witnesses used above are isolated in Lemma 7.4 and separately recorded by `matrixSdpStrictPrimalSubmeasurement`, `matrixSdpStrictDualWitness`, and `matrixSdpFeasibleBounds_canonical`; the dual feasibility of $2I$ follows from `matrixAveragedPointOperator_le_one`. The block-diagonal reduction is also recorded for arbitrary feasible canonical matrices: extracting the diagonal blocks preserves the objective and transfers canonical complementary slackness to the extracted paper primal submeasurement. □

Lemma 7.9 (Averaging over the evaluation point)**.** *Let $T = \{T_h\}$ be an optimal solution of the primal SDP, let*

$$H_h = \mathbb{E}_{u\sim\mathbb{F}_q^m} A^u_{h(u)} T_h A^u_{h(u)},$$

and set

$$\zeta_{\text{variance}} = 24m \cdot \left(\varepsilon + \delta + \frac{md}{q}\right).$$

Suppose $M = \{M^u_o\}$ is a submeasurement with outcomes in some set $\mathcal{O}$. For each $u \in \mathbb{F}_q^m$, let S_u be a subset of $\mathcal{O} \times \mathcal{P}(m, q, d)$. Then

$$\mathbb{E}_u \sum_{(o,h)\in S_u} \langle\psi| M^u_o \otimes H_h |\psi\rangle \approx_{4\sqrt{\zeta_{\text{variance}}}} \mathbb{E}_u \sum_{(o,h)\in S_u} \langle\psi| (A^u_{h(u)} \cdot M^u_o \cdot A^u_{h(u)}) \otimes T_h |\psi\rangle.$$

The reduced variance-bound specialization keeps only the variance-bound consequence used later in the helper theorem; it is not itself the full formal counterpart of Lemma 7.9.

Proof. We begin by expanding

$$\begin{aligned}\mathbb{E}_u \sum_{(o,h)\in S_u} \langle\psi| M^u_o \otimes H_h |\psi\rangle &= \mathbb{E}_{u,v} \sum_{(o,h)\in S_u} \langle\psi| M^u_o \otimes H^v_h |\psi\rangle \\ &= \mathbb{E}_{u,v} \sum_{(o,h)\in S_u} \langle\psi| M^u_o \otimes (A^v_{h(v)} \cdot T_h \cdot A^v_{h(v)}) |\psi\rangle. \qquad (7.31)\end{aligned}$$

We claim that

$$(7.31) \approx_{\sqrt{2\delta}} \mathbb{E}_{u,v} \sum_{(o,h)\in S_u} \langle\psi| (A^v_{h(v)} \cdot M^u_o) \otimes (T_h \cdot A^v_{h(v)}) |\psi\rangle. \qquad (7.32)$$

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

$$\begin{aligned}&\left|\mathbb{E}_{u,v} \sum_{(o,h)\in S_u} \langle\psi| (A^v_{h(v)} \otimes I - I \otimes A^v_{h(v)}) \cdot (M^u_o \otimes (T_h \cdot A^v_{h(v)})) |\psi\rangle\right| \\ &\quad\le \sqrt{\mathbb{E}_{u,v} \sum_{(o,h)\in S_u} \langle\psi| (A^v_{h(v)} \otimes I - I \otimes A^v_{h(v)}) \cdot (M^u_o \otimes T_h) \cdot (A^v_{h(v)} \otimes I - I \otimes A^v_{h(v)}) |\psi\rangle} \\ &\qquad\cdot \sqrt{\mathbb{E}_{u,v} \sum_{(o,h)\in S_u} \langle\psi| M^u_o \otimes (A^v_{h(v)} \cdot T_h \cdot A^v_{h(v)}) |\psi\rangle}. \qquad (7.33)\end{aligned}$$

The first factor is controlled by summing over the value $a = h(v)$:

$$\mathbb{E}_{u,v} \sum_{a \in \mathbb{F}_q} \langle \psi | (A^v_a \otimes I - I \otimes A^v_a) \cdot \Big(\sum_{(o,h) \in S_u,\, h(v) = a} M^u_o \otimes T_h \Big) \cdot (A^v_a \otimes I - I \otimes A^v_a) | \psi \rangle$$
$$\leq \mathbb{E}_{u,v} \sum_{a \in \mathbb{F}_q} \langle \psi | (A^v_a \otimes I - I \otimes A^v_a)^2 | \psi \rangle,$$

where the inequality uses that M^u and T are submeasurements. By Lemma 3.21 and the self-consistency of A, this is at most 2δ. The second factor is

$$\mathbb{E}_{u,v} \sum_{(o,h) \in S_u} \langle \psi | M^u_o \otimes H^v_h | \psi \rangle,$$

which is at most 1 because M^u and H^v are submeasurements.

Next, we claim that

$$(7.32) \approx_{\sqrt{2\delta}} \mathbb{E}_{u,v} \sum_{(o,h) \in S_u} \langle \psi | (A^v_{h(v)} \cdot M^u_o \cdot A^v_{h(v)}) \otimes T_h | \psi \rangle. \tag{7.34}$$

Again, we bound the magnitude of the difference:

$$\Big| \mathbb{E}_{u,v} \sum_{(o,h) \in S_u} \langle \psi | ((A^v_{h(v)} \cdot M^u_o) \otimes T_h) \cdot (A^v_{h(v)} \otimes I - I \otimes A^v_{h(v)}) | \psi \rangle \Big|$$
$$\leq \sqrt{\mathbb{E}_{u,v} \sum_{(o,h) \in S_u} \langle \psi | (A^v_{h(v)} \cdot M^u_o \cdot A^v_{h(v)}) \otimes T_h | \psi \rangle} \tag{7.35}$$
$$\cdot \sqrt{\mathbb{E}_{u,v} \sum_{(o,h) \in S_u} \langle \psi | (A^v_{h(v)} \otimes I - I \otimes A^v_{h(v)}) \cdot (M^u_o \otimes T_h) \cdot (A^v_{h(v)} \otimes I - I \otimes A^v_{h(v)}) | \psi \rangle}.$$

The first factor is bounded by

$$\mathbb{E}_{u,v} \sum_{h} \langle \psi | \Big(A^v_{h(v)} \cdot \sum_{o : (o,h) \in S_u} M^u_o \cdot A^v_{h(v)} \Big) \otimes T_h | \psi \rangle$$
$$\leq \mathbb{E}_{u,v} \sum_{h} \langle \psi | (A^v_{h(v)})^2 \otimes T_h | \psi \rangle \leq \mathbb{E}_{u,v} \sum_{h} \langle \psi | I \otimes T_h | \psi \rangle \leq 1,$$

using first that M is a submeasurement, then that $A^v_{h(v)} \leq I$, and finally that T is a submeasurement ($\sum_g T_g \leq I$). The second factor is exactly the same expression as in (7.33), hence at most $\sqrt{2\delta}$.

Having moved both copies of A to the left-hand side, we next show that

$$(7.34) \approx_{\sqrt{\zeta_{\text{variance}}}} \mathbb{E}_{u,v} \sum_{(o,h) \in S_u} \langle \psi | (A^u_{h(u)} \cdot M^u_o \cdot A^v_{h(v)}) \otimes T_h | \psi \rangle. \tag{7.36}$$

The difference is bounded by

$$\Big| \mathbb{E}_{u,v} \sum_{(o,h) \in S_u} \langle \psi | ((A^v_{h(v)} - A^u_{h(u)}) \cdot M^u_o \cdot A^v_{h(v)}) \otimes T_h | \psi \rangle \Big|$$
$$\leq \sqrt{\mathbb{E}_{u,v} \sum_{(o,h) \in S_u} \langle \psi | ((A^v_{h(v)} - A^u_{h(u)}) \cdot M^u_o \cdot (A^v_{h(v)} - A^u_{h(u)})) \otimes T_h | \psi \rangle}$$
$$\cdot \sqrt{\mathbb{E}_{u,v} \sum_{(o,h) \in S_u} \langle \psi | (A^v_{h(v)} \cdot M^u_o \cdot A^v_{h(v)}) \otimes T_h | \psi \rangle}. \tag{7.37}$$

The first factor can be rewritten as

$$\begin{aligned}&\mathbb{E}_{u,v}\sum_h \langle\psi|\Big((A^v_{h(v)}-A^u_{h(u)})\cdot \sum_{o:(o,h)\in S_u} M^u_o\cdot (A^v_{h(v)}-A^u_{h(u)})\Big)\otimes T_h|\psi\rangle\\ &\le \mathbb{E}_{u,v}\sum_h \langle\psi|(A^v_{h(v)}-A^u_{h(u)})^2\otimes T_h|\psi\rangle,\end{aligned}$$

because M^u is a submeasurement. Lemma 6.6 bounds this by ζ_{variance} since $T\in \mathrm{PolySub}(m,q,d)$. The second factor is exactly the first factor from (7.35), so it is at most 1.

Finally, we show that

$$(7.36)\approx_{\sqrt{\zeta_{\text{variance}}}} \mathbb{E}_{u,v}\sum_{(o,h)\in S_u}\langle\psi|(A^u_{h(u)}\cdot M^u_o\cdot A^u_{h(u)})\otimes T_h|\psi\rangle. \tag{7.38}$$

The same Cauchy–Schwarz estimate as in (7.37), with the final copy of $A^v_{h(v)}$ replaced by $A^u_{h(u)}$, gives the same bound $\sqrt{\zeta_{\text{variance}}}$. Indeed, the factor containing $A^u_{h(u)}\cdot M^u_o\cdot A^u_{h(u)}$ is bounded by 1 by the same submeasurement argument as above, while the factor containing $(A^v_{h(v)}-A^u_{h(u)})^2$ is unchanged. Summing the four transports gives an error of $2\sqrt{2\delta}+2\sqrt{\zeta_{\text{variance}}}$, and the lemma follows from $2\delta\le\zeta_{\text{variance}}$. □

7.2 Projective Output

This final step applies orthonormalization to the non-projective family produced by Lemma 7.1. The bookkeeping here is entirely about showing that completeness, consistency, self-consistency, and boundedness survive the passage from $\widehat{H}$ to a projective submeasurement H.

Remark 7.10 (Auxiliary self-improvement lemmas). Several auxiliary results are used to prove Theorem 7.11 from the helper output and the orthonormalization theorem. They record local intermediate estimates and final-field assembly lemmas. The former orthonormalization spectral-transport record has been retired; the spectral-truncation part of the construction is now represented directly by `spectralTruncationStatement_of_sourceAlmostProjective`. This result is not a locality-preserving repair theorem. The theorem below records the paper-facing conclusion. Its proof follows the expectation-level completeness argument from the paper; the obstruction to a generic operator-total strengthening is recorded in issue #1642.

Theorem 7.11 (Self-improvement)**.** *Let $G\in\mathrm{PolyMeas}(m,q,d)$ be a measurement with the following properties:*

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

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

Let

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

Then there exists a projective submeasurement $H\in\mathrm{PolySub}(m,q,d)$ 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 $u \sim \mathbb{F}_q^m$,*

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

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

$$Z \geq \left(\mathbb{E}_u A_{h(u)}^u \right).$$

Proof. The bound is trivial if one of ε, δ, or d/q is at least 1, so assume all three are at most 1. In Lean, the third assumption is recorded as $d \leq q$, which is equivalent to $d/q \leq 1$ since $q > 0$. Apply Lemma 7.1 to G and let $\widehat{H} \in \mathrm{PolySub}(m, q, d)$ be the output, with

$$\hat{\zeta} = 100m \cdot \left(\varepsilon^{1/2} + \delta^{1/2} + (d/q)^{1/2} \right).$$

By 3 of Lemma 7.1,

$$\sum_h \langle \psi | \widehat{H}_h \otimes \widehat{H}_h | \psi \rangle \geq \langle \psi | \widehat{H} \otimes I | \psi \rangle - \hat{\zeta}.$$

Therefore Theorem 4.4 supplies a projective submeasurement $H \in \mathrm{PolySub}(m, q, d)$ such that

$$\widehat{H}_h \otimes I \approx_{\hat{\zeta}_{\mathrm{ortho}}} H_h \otimes I, \tag{7.39}$$

where

$$\hat{\zeta}_{\mathrm{ortho}} = 100 \hat{\zeta}^{1/4}.$$

In addition, Lemma 3.39 implies that

$$\widehat{H}_{[h(u)=a]} \otimes I \approx_{\hat{\zeta}_{\mathrm{dataprocess}}} H_{[h(u)=a]} \otimes I, \tag{7.40}$$

where

$$\hat{\zeta}_{\mathrm{dataprocess}} = 8\hat{\zeta} + 8\sqrt{\hat{\zeta}_{\mathrm{ortho}}}.$$

The four conclusions are transferred one by one.

Proof of 1. Lemma 3.38 implies

$$\langle \psi | H \otimes I | \psi \rangle \geq \langle \psi | \widehat{H} \otimes I | \psi \rangle - \hat{\zeta} - 2\sqrt{\hat{\zeta}_{\mathrm{ortho}}} \geq (1 - \nu) - 2\hat{\zeta} - 2\sqrt{\hat{\zeta}_{\mathrm{ortho}}},$$

where the second inequality is 1 of Lemma 7.1.

Proof of 2. Lemma 3.20 applied to 2 of Lemma 7.1 and the post-processed comparison (7.40) gives

$$A_a^u \otimes I \simeq_{\hat{\zeta} + \sqrt{\hat{\zeta}_{\mathrm{dataprocess}}}} I \otimes H_{[h(u)=a]}.$$

In Lean the corresponding transport is stated for submeasurements with the total-overlap displacement recorded explicitly; this is the term which vanishes in the measurement-valued form of Lemma 3.20. The monotone-total route records a sufficient structural replacement: if the orthonormalization repair preserves the completed residual outcome, then the projective total

is bounded by the helper total as an operator, hence also has no larger right-register expectation in the strategy state. Under this structural hypothesis the alphabet-size displacement term is not introduced. The QXP-layer interface used by the formalization records both the total domination of the canonical projective family and the sharper residual comparison for the fresh option-completion outcome. At present the formal QXP algebra only isolates the construction-level statement $Q_{\text{none}} \leq P_{\text{none}}$, obtained when the repair preserves the fresh row block $\widehat{X}_{\text{none}} = X_{\text{none}}$. The former generic "RestrictSome" monotone-total route would have required one to turn this into the source-facing inequality $\widehat{A}_{\perp} \leq \widehat{P}_{\perp}$, by adding a comparison between the completed source residual and the "Q"-layer fresh outcome. The present formal development no longer uses that generic route, and the comparison is false without additional hypotheses; see `docs/reports/issue-1642-restrictsome-residual-domination-obstruction.md`. Si
$A^u_{\text{tot}} = I$ and postprocessing does not change the total operator, Lean also records the equivalent form in which this displacement is the single scalar difference

$$\left|\langle \psi, I \otimes H_{\text{tot}}\psi\rangle - \langle \psi, I \otimes \widehat{H}_{\text{tot}}\psi\rangle\right|$$

between the right-register totals of H and $\widehat{H}$. The outcomewise data-processing bound controls this scalar by applying the vector triangle inequality to the sum over $a \in \mathbb{F}_q$; the formal estimate gives the additional term

$$\sqrt{q\,\hat{\zeta}_{\text{dataprocess}}}.$$

Thus the total-overlap term is no longer an independent hypothesis, although the numerical absorption of this alphabet-size loss is recorded separately. There is also a sharper formal route: if the right-register total of the projective family has no larger expectation than that of $\widehat{H}$, then the total-overlap term can only improve the consistency defect. Under this monotonicity hypothesis the transported error is again $\hat{\zeta} + \sqrt{\hat{\zeta}_{\text{dataprocess}}}$, and the standard final-stage numerical estimate absorbs it into ζ.

Proof of 3. Lemma 3.36 implies

$$\widehat{H}_h \otimes I \approx_{2\hat{\zeta}} I \otimes \widehat{H}_h.$$

Combining this with (7.39) on both sides and summing the three transports with Lemma 3.27 yields

$$H_h \otimes I \approx_{6\hat{\zeta}+6\hat{\zeta}_{\text{ortho}}} I \otimes H_h.$$

Proof of 4. The boundedness quantity is

$$\begin{aligned}
&\langle \psi | Z \otimes (I - H) | \psi \rangle \\
&= \langle \psi | Z \otimes I | \psi \rangle - \sum_h \langle \psi | Z \otimes H_h | \psi \rangle \\
&\leq \langle \psi | Z \otimes I | \psi \rangle - \sum_h \langle \psi | \mathbb{E}_u A^u_{h(u)} \otimes H_h | \psi \rangle \\
&= \langle \psi | Z \otimes I | \psi \rangle - \mathbb{E}_u \sum_h \langle \psi | A^u_{h(u)} \otimes H_h | \psi \rangle \\
&= \langle \psi | Z \otimes I | \psi \rangle - \mathbb{E}_u \sum_a \langle \psi | A^u_a \otimes H_{[h(u)=a]} | \psi \rangle. \qquad (7.41)
\end{aligned}$$

Proposition 3.24 upgrades (7.40) to the scalar estimate

$$(7.41) \leq \langle \psi | Z \otimes I | \psi \rangle - \mathbb{E}_u \sum_a \langle \psi | A^u_a \otimes \widehat{H}_{[h(u)=a]} | \psi \rangle + \sqrt{\hat{\zeta}_{\text{dataprocess}}},$$

and now 4 of Lemma 7.1 bounds this by $\hat{\zeta} + \sqrt{\hat{\zeta}_{\text{dataprocess}}}$. The formal boundedness field is the corresponding operator-monotonicity step: once the SDP dual witness satisfies $I \le Z$, the total mass of any left-placed submeasurement, and in particular of the final projective submeasurement, is bounded by $Z \otimes I$ with zero boundedness error.

Thus the four conclusions hold with errors

$$2\hat{\zeta} + 2\sqrt{\hat{\zeta}_{\text{ortho}}}, \qquad \hat{\zeta} + \sqrt{\hat{\zeta}_{\text{dataprocess}}}, \qquad 6\hat{\zeta} + 6\hat{\zeta}_{\text{ortho}}, \qquad \hat{\zeta} + \sqrt{\hat{\zeta}_{\text{dataprocess}}},$$

respectively. The paper's exponent bookkeeping then shows the following bounds. The formal bookkeeping uses the three-term power sum $\varepsilon^p + \delta^p + (d/q)^p$, its monotonicity on the unit interval, and the corresponding square-root estimate. The displayed Lean theorems record these arithmetic identities together with the completeness, self-closeness, and projective-residual absorptions $2\hat{\zeta} + 2\sqrt{\hat{\zeta}_{\text{ortho}}} \le \zeta$, $6\hat{\zeta} + 6\hat{\zeta}_{\text{ortho}} \le \zeta$ and $\hat{\zeta} + \sqrt{\hat{\zeta}_{\text{dataprocess}}} \le \zeta$. The two threshold lemmas involving $|\mathbb{F}_q|$ record the scalar part of the alphabet-size obstruction in the current final-fields point-consistency transport: the data-processing threshold contains $8\hat{\zeta}$, but the transported estimate still carries the factor $\sqrt{|\mathbb{F}_q|}$.

$$\hat{\zeta}_{\text{ortho}} \le 400m \cdot \left(\varepsilon^{1/8} + \delta^{1/8} + (d/q)^{1/8}\right),$$

$$\hat{\zeta}_{\text{dataprocess}} \le 960m \cdot \left(\varepsilon^{1/16} + \delta^{1/16} + (d/q)^{1/16}\right),$$

and hence

$$6\hat{\zeta} + 6\hat{\zeta}_{\text{ortho}} \le \zeta, \qquad \hat{\zeta} + \sqrt{\hat{\zeta}_{\text{dataprocess}}} \le \zeta.$$

Since $2\hat{\zeta} + 2\sqrt{\hat{\zeta}_{\text{ortho}}} \le \hat{\zeta}_{\text{dataprocess}} \le \zeta$, all four bounds are at most ζ. $\square$

Chapter 8

Commutativity

8.1 Commutativity of the point measurements

Definition 8.1 (Shared diagonal-line sample)**.** Sample a uniformly random ordered point pair (u, v) and a uniformly random parameter $t \in \mathbb{F}_q$. Package this data as the diagonal line whose direction is $v - u$ and whose parametrization visits u at t and v at $t + 1$.

Theorem 8.2 (Commutativity of the point measurements)**.** *Let (ψ, A, B, L) be an $(\varepsilon, \delta, \gamma)$-good symmetric strategy for the (m, q, d) low individual degree test. On average over independent and uniformly random $u, v \sim \mathbb{F}_q^m$,*

$$(A_a^u A_b^v) \otimes I \approx_{32\gamma m} (A_b^v A_a^u) \otimes I.$$

Proof. By Definition 2.13, the strategy passes the m-restricted diagonal lines test with probability $1 - \gamma m$. Hence

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

on average over uniformly random $u \sim \mathbb{F}_q^m$ and a uniformly random line ℓ in $\mathbb{F}_q^m$ containing u. By 3.21,

$$A_a^u \otimes I \approx_{2\gamma m} I \otimes L_{[f(u)=a]}^{\ell}. \tag{8.1}$$

Let u and v be independent uniformly random points in $\mathbb{F}_q^m$, and let ℓ be a uniformly random line containing both points. The marginal distribution on (u, ℓ) and on (v, ℓ) is the same as above, so

$$\begin{aligned}
(A_a^u A_b^v) \otimes I &\approx_{2\gamma m} A_a^u \otimes L_{[f(v)=b]}^{\ell} && \text{(by 8.1)} \\
&\approx_{2\gamma m} I \otimes (L_{[f(v)=b]}^{\ell} L_{[f(u)=a]}^{\ell}) && \text{(by 8.1)} \\
&= I \otimes (L_{[f(u)=a]}^{\ell} L_{[f(v)=b]}^{\ell}) && \text{(because L is projective)} \\
&\approx_{2\gamma m} A_b^v \otimes L_{[f(u)=a]}^{\ell} && \text{(by 8.1)} \\
&\approx_{2\gamma m} (A_b^v A_a^u) \otimes I. && \text{(by 8.1)}
\end{aligned}$$

The theorem follows from 3.27. □

8.2 Commutativity of G after evaluation

Lemma 8.3 (Commutativity of G after evaluation)**.** *Let (ψ, A, B, L) be an $(\varepsilon, \delta, \gamma)$-good symmetric strategy for the $(m+1, q, d)$ low individual degree test. Let $\{G^x\} \in \mathrm{PolySub}(m, q, d)$ 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 $(u, x) \sim \mathbb{F}_q^{m+1}$,*
$$A_a^{u,x} \otimes I \simeq_\zeta I \otimes G^x_{[g(u)=a]}.$$
2. *(Strong self-consistency) On average over $x \sim \mathbb{F}_q$,*
$$G^x_g \otimes I \approx_\zeta I \otimes G^x_g.$$
3. *(Boundedness) There exists a positive-semidefinite matrix Z^x for each $x \in \mathbb{F}_q$ such that*
$$\mathbb{E}_x \langle \psi | (I - G^x) \otimes Z^x | \psi \rangle \leq \zeta$$
and for each $x \in \mathbb{F}_q$ and $g \in \mathcal{P}(m, q, d)$,
$$Z^x \geq \left(\mathbb{E}_u A^{u,x}_{g(u)} \right).$$

Let
$$\nu = 48m(\gamma^{1/2} + \zeta^{1/2}).$$
Then on average over independent and uniformly random $(u, x), (v, y) \sim \mathbb{F}_q^{m+1}$,
$$G^x_{[g(u)=a]} G^y_{[h(v)=b]} \otimes I \approx_\nu G^y_{[h(v)=b]} G^x_{[g(u)=a]} \otimes I.$$

Proof. Write $G_a^{u,x} = G^x_{[g(u)=a]}$. Expanding the commutator square gives

$$\begin{aligned} &\mathbb{E}_{u,v,x,y} \sum_{a,b} \langle \psi | (G_a^{u,x} G_b^{v,y} - G_b^{v,y} G_a^{u,x})^\dagger (G_a^{u,x} G_b^{v,y} - G_b^{v,y} G_a^{u,x}) \otimes I | \psi \rangle \\ &= 2\mathbb{E}_{u,v,x,y} \sum_{a,b} \Big(\langle \psi | G_b^{v,y} G_a^{u,x} G_b^{v,y} \otimes I | \psi \rangle - \langle \psi | G_a^{u,x} G_b^{v,y} G_a^{u,x} G_b^{v,y} \otimes I | \psi \rangle \Big), \end{aligned} \tag{8.2}$$

using the projectivity of G. To compare the two quartic terms, first note that

$$G^{u,x} = \sum_a G_a^{u,x} = \sum_a G^x_{[g(u)=a]} = G^x. \tag{8.3}$$

The identity (8.3) is exactly Proposition 3.13 for the evaluation map $g \mapsto g(u)$. By 1 and 3.31,

$$\begin{aligned} G_a^{u,x} \otimes I &\approx_{4\zeta} G^{u,x} \otimes A_a^{u,x} \\ &= G^x \otimes A_a^{u,x}, \end{aligned} \tag{8.4}$$

where the last equality is 8.3. Applying 3.23 once to the second term in 8.2 gives

$$\begin{aligned} &\mathbb{E}_{u,v,x,y} \sum_{a,b} \langle \psi | G_a^{u,x} G_b^{v,y} G_a^{u,x} G_b^{v,y} \otimes I | \psi \rangle \\ &\approx_{2\sqrt{\zeta}} \mathbb{E}_{u,v,x,y} \sum_{a,b} \langle \psi | G_a^{u,x} G_b^{v,y} G_a^{u,x} G^y \otimes A_b^{v,y} | \psi \rangle. \end{aligned} \tag{8.5}$$

The first boundedness-driven stability step is

$$(8.5) \approx_{\sqrt{\zeta}} \mathbb{E}_{u,v,x,y} \sum_{a,b} \langle\psi|G_a^{u,x} G_b^{v,y} G_a^{u,x} \otimes A_b^{v,y}|\psi\rangle, \tag{8.6}$$

which is 8.4. Applying 8.4 once more and then 8.2, we obtain

$$\begin{aligned}(8.6) &\approx_{2\sqrt{\zeta}} \mathbb{E}_{u,v,x,y} \sum_{a,b} \langle\psi|G_a^{u,x} G_b^{v,y} G^x \otimes A_a^{u,x} A_b^{v,y}|\psi\rangle \\ &\approx_{6\sqrt{\gamma(m+1)}} \mathbb{E}_{u,v,x,y} \sum_{a,b} \langle\psi|G_a^{u,x} G_b^{v,y} G^x \otimes A_b^{v,y} A_a^{u,x}|\psi\rangle. \end{aligned} \tag{8.7}$$

The second boundedness-driven stability step is

$$(8.7) \approx_{\sqrt{\zeta}+6\sqrt{\gamma(m+1)}} \mathbb{E}_{u,v,x,y} \sum_{a,b} \langle\psi|G_a^{u,x} G_b^{v,y} \otimes A_a^{u,x} A_b^{v,y}|\psi\rangle, \tag{8.8}$$

which is 8.5. Using 8.4 twice more and 3.23,

$$\begin{aligned}(8.8) &\approx_{2\sqrt{\zeta}} \mathbb{E}_{u,v,x,y} \sum_{a,b} \langle\psi|G_a^{u,x} G_b^{v,y} \otimes A_b^{v,y}|\psi\rangle \\ &\approx_{2\sqrt{\zeta}} \mathbb{E}_{u,v,x,y} \sum_{a,b} \langle\psi|G_a^{u,x} G_b^{v,y} G_b^{v,y} \otimes I|\psi\rangle. \end{aligned} \tag{8.9}$$

Next, 2 and the projectivity of G imply via 3.36 and 3.37 that

$$G_a^{u,x} \otimes I \approx_{\zeta} I \otimes G_a^{u,x}. \tag{8.10}$$

Applying 8.10 twice with 3.23 turns 8.9 into

$$\mathbb{E}_{u,v,x,y} \sum_{a,b} \langle\psi|G_b^{v,y} G_a^{u,x} G_b^{v,y} \otimes I|\psi\rangle,$$

which is the first quartic term from 8.2. Collecting the ten approximation steps appearing between 8.2 and 8.10 yields the bound

$$48m(\gamma^{1/2} + \zeta^{1/2}),$$

so the commutator norm is at most ν. $\square$

Lemma 8.4. *This claim is made under the same standing hypotheses used in the commutativity-of-G theorem above. The linked `SliceBoundednessInput` declarations are precisely the two displayed parts of item 3: the averaged $(I - G^y) \otimes Z^y$ residual bound and the domination $Z^y \geq \mathbb{E}_u A_{g(u)}^{u,y}$.*

$$\begin{aligned}&\mathbb{E}_{u,v,x,y} \sum_{a,b} \langle\psi|G_a^{u,x} G_b^{v,y} G_a^{u,x} G^y \otimes A_b^{v,y}|\psi\rangle \\ &\approx_{\sqrt{\zeta}} \mathbb{E}_{u,v,x,y} \sum_{a,b} \langle\psi|G_a^{u,x} G_b^{v,y} G_a^{u,x} \otimes A_b^{v,y}|\psi\rangle. \end{aligned} \tag{8.11}$$

Proof. For each $y \in \mathbb{F}_q$ and $g \in \mathcal{P}(m, q, d)$, define

$$R_g^y = \mathbb{E}_{u,x} \sum_a G_a^{u,x} G_g^y G_a^{u,x}.$$

Since G is a sub-measurement, $R^y = \{R^y_g\}$ is also a sub-measurement. The difference between the two sides of the lemma is

$$\begin{aligned} &\mathbb{E}_{u,v,x,y} \sum_{a,b} \langle\psi| G^{u,x}_a G^{v,y}_b G^{u,x}_a (I - G^y) \otimes A^{v,y}_b |\psi\rangle \\ &= \mathbb{E}_{v,y} \sum_{g} \langle\psi| R^y_g (I - G^y) \otimes A^{v,y}_{g(v)} |\psi\rangle. \end{aligned} \tag{8.12}$$

Apply Cauchy–Schwarz to 8.12. The first factor is at most 1 because R^y is a sub-measurement. For the second factor, average over v and use 3 to replace $\mathbb{E}_v A^{v,y}_{g(v)}$ by Z^y; the resulting quantity is at most

$$\mathbb{E}_y \langle\psi|(I - G^y) \otimes Z^y|\psi\rangle \le \zeta.$$

Hence the difference has magnitude at most $\sqrt{\zeta}$. □

Lemma 8.5. *This claim is made under the same standing hypotheses used in the commutativity-of-G theorem above, in particular the boundedness item 3. The linked `SliceBoundednessInput` declarations expose those boundedness hypotheses in Lean; they are not auxiliary assumptions added to the paper statement.*

$$\begin{aligned} &\mathbb{E}_{u,v,x,y} \sum_{a,b} \langle\psi| G^{u,x}_a G^{v,y}_b G^x \otimes A^{u,x}_a A^{v,y}_b |\psi\rangle \\ \approx_{\sqrt{\zeta}+6\sqrt{\gamma(m+1)}} &\mathbb{E}_{u,v,x,y} \sum_{a,b} \langle\psi| G^{u,x}_a G^{v,y}_b \otimes A^{u,x}_a A^{v,y}_b |\psi\rangle. \end{aligned}$$

Proof. The difference between the two sides is

$$\begin{aligned} &\mathbb{E}_{u,v,x,y} \sum_{a,b} \langle\psi| G^{u,x}_a G^{v,y}_b (I - G^x) \otimes A^{u,x}_a A^{v,y}_b |\psi\rangle \\ \approx_{6\sqrt{\gamma(m+1)}} &\mathbb{E}_{u,v,x,y} \sum_{a,b} \langle\psi| G^{u,x}_a G^{v,y}_b (I - G^x) \otimes A^{v,y}_b A^{u,x}_a |\psi\rangle, \end{aligned} \tag{8.13}$$

where 8.2 commutes the point measurements on the right register. Expanding $G^{u,x}_a = \sum_{g:g(u)=a} G^x_g$, we rewrite

$$(8.13) = \mathbb{E}_{u,v,x,y} \sum_{g,b} \langle\psi| G^x_g G^{v,y}_b (I - G^x) \otimes A^{v,y}_b A^{u,x}_{g(u)} |\psi\rangle. \tag{8.14}$$

Apply Cauchy–Schwarz to 8.14. The first factor is at most 1 because both G and A are sub-measurements. In the second factor, average over u and use 3 to replace $\mathbb{E}_u A^{u,x}_{g(u)}$ by Z^x; then sum over g and b using the sub-measurement property. This leaves

$$\mathbb{E}_x \langle\psi|(I - G^x) \otimes Z^x|\psi\rangle \le \zeta.$$

Hence 8.14 has magnitude at most $\sqrt{\zeta}$, and the total loss is $\sqrt{\zeta} + 6\sqrt{\gamma(m+1)}$. □

8.3 Commutativity of G

Lemma 8.6 (Normalization condition for a sandwiched family)**.** *Let $P = \{P_a\}$ be a sub-measurement and $Q = \{Q_b\}$ be a projective sub-measurement. Define $C_{a,b} = Q_b P_a Q_b$. Then*

$$\sum_a \left(\sum_b C_{a,b}\right)^\dagger \left(\sum_b C_{a,b}\right) = \sum_a \left(\sum_b C_{a,b}\right) \left(\sum_b C_{a,b}\right)^\dagger \le I.$$

Proof. Expand the square, use the orthogonality $Q_b Q_{b'} = 0$ for $b \neq b'$, and then sum over the sub-measurement P. $\square$

Theorem 8.7 (Commutativity of G)**.** *Let (ψ, A, B, L) be an $(\varepsilon, \delta, \gamma)$-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. *(Consistency with A) On average over $(u, x) \sim \mathbb{F}_q^{m+1}$,*
$$A_a^{u,x} \otimes I \simeq_\zeta I \otimes G^x_{[g(u)=a]}.$$
2. *(Strong self-consistency) On average over $x \sim \mathbb{F}_q$,*
$$G_g^x \otimes I \approx_\zeta I \otimes G_g^x.$$
3. *(Boundedness) There exists a positive-semidefinite matrix Z^x for each $x \in \mathbb{F}_q$ such that*
$$\mathbb{E}_x \langle \psi | (I - G^x) \otimes Z^x | \psi \rangle \leq \zeta$$
and for each $x \in \mathbb{F}_q$ and $g \in \mathcal{P}(m, q, d)$,
$$Z^x \geq \left(\mathbb{E}_u A^{u,x}_{g(u)} \right).$$

Let
$$\nu = 30m \left(\gamma^{1/4} + \zeta^{1/4} + (d/q)^{1/4} \right).$$
Then on average over independent and uniformly random $(u, x), (v, y) \sim \mathbb{F}_q^{m+1}$,
$$G_g^x G_h^y \otimes I \approx_\nu G_h^y G_g^x \otimes I.$$

Proof. If at least one of γ, ζ, or d/q is at least 1, then $\nu \geq 30$ and the estimate is trivial from 3.26. We therefore assume $\gamma, \zeta, d/q \leq 1$.

Expanding the commutator square gives

$$\begin{aligned}
&\mathbb{E}_{x,y} \sum_{g,h} \langle \psi | (G_g^x G_h^y - G_h^y G_g^x)^\dagger (G_g^x G_h^y - G_h^y G_g^x) \otimes I | \psi \rangle \\
&= 2\mathbb{E}_{x,y} \sum_{g,h} \langle \psi | G_g^x G_h^y G_g^x \otimes I | \psi \rangle - 2\mathbb{E}_{x,y} \sum_{g,h} \langle \psi | G_g^x G_h^y G_g^x G_h^y \otimes I | \psi \rangle. \qquad (8.15)
\end{aligned}$$

The first term is close to $\langle \psi | G \otimes G | \psi \rangle$, where $G = \mathbb{E}_x G^x$, by 3.32 and 2. For the second term, 3.23 and 8.6 yield

$$\mathbb{E}_{x,y} \sum_{g,h} \langle \psi | G_g^x G_h^y G_g^x G_h^y \otimes I | \psi \rangle \approx_{\sqrt{\zeta}} \mathbb{E}_{x,y} \sum_{g,h} \langle \psi | G_h^y G_g^x G_h^y \otimes G_g^x | \psi \rangle. \qquad (8.16)$$

The first Schwartz–Zippel reduction evaluates the left factor at a random point:

$$\begin{aligned}
(8.16) &= \mathbb{E}_{x,y} \sum_{g,h} \langle \psi | G_h^y G_g^x G_h^y \otimes G_g^x | \psi \rangle \\
&\approx_{\frac{dm}{q}} \mathbb{E}_{u,x,y} \sum_{a,h} \langle \psi | G_h^y G^x_{[g(u)=a]} G_h^y \otimes G^x_{[g(u)=a]} | \psi \rangle. \qquad (8.17)
\end{aligned}$$

Indeed, the difference is

$$\mathbb{E}_{u,x,y} \sum_{g \neq g', h} \mathbf{1}[[]\, g(u) = g'(u)] \langle \psi | G^y_h G^x_g G^y_h \otimes G^x_{g'} | \psi \rangle, \tag{8.18}$$

and 3.7 bounds this by dm/q.

Next, apply 3.23 twice:

$$\begin{aligned}
(8.17) &= \mathbb{E}_{u,x,y} \sum_{a,h} \langle \psi | G^y_h G^x_{[g(u)=a]} G^y_h \otimes G^x_{[g(u)=a]} | \psi \rangle \\
&\approx_{\sqrt{\zeta}} \mathbb{E}_{u,x,y} \sum_{a,h} \langle \psi | G^x_{[g(u)=a]} G^y_h G^x_{[g(u)=a]} G^y_h \otimes I | \psi \rangle \\
&\approx_{\sqrt{\zeta}} \mathbb{E}_{u,x,y} \sum_{a,h} \langle \psi | G^x_{[g(u)=a]} G^y_h G^x_{[g(u)=a]} \otimes G^y_h | \psi \rangle. \qquad (8.19)
\end{aligned}$$

A second Schwartz–Zippel reduction evaluates the right factor at an independent point:

$$\begin{aligned}
(8.19) &= \mathbb{E}_{u,x,y} \sum_{a,h} \langle \psi | G^x_{[g(u)=a]} G^y_h G^x_{[g(u)=a]} \otimes G^y_h | \psi \rangle \\
&\approx_{\frac{dm}{q}} \mathbb{E}_{u,v,x,y} \sum_{a,b} \langle \psi | G^x_{[g(u)=a]} G^y_{[h(v)=b]} G^x_{[g(u)=a]} \otimes G^y_{[h(v)=b]} | \psi \rangle. \qquad (8.20)
\end{aligned}$$

The difference is

$$\mathbb{E}_{u,v,x,y} \sum_{a, h \neq h'} \mathbf{1}[[]\, h(v) = h'(v)] \langle \psi | G^x_{[g(u)=a]} G^y_h G^x_{[g(u)=a]} \otimes G^y_{h'} | \psi \rangle, \tag{8.21}$$

and the same Schwartz–Zippel estimate again gives the loss dm/q.

Now set

$$\nu_{\text{evaluation}} = 48m(\gamma^{1/2} + \zeta^{1/2}),$$

the error from 8.3. Apply 8.3 to commute the evaluated families in 8.20; after one more use of 3.23 with 2, the resulting quantity is $\langle \psi | G \otimes G | \psi \rangle$. Thus the second term in 8.15 is also close to $\langle \psi | G \otimes G | \psi \rangle$.

Combining the first-term estimate, 8.16, the two Schwartz–Zippel losses, the two intermediate $\sqrt{\zeta}$ losses, and the evaluated commutation error $\nu_{\text{evaluation}}^{1/2}$ gives

$$30m \left(\gamma^{1/4} + \zeta^{1/4} + (d/q)^{1/4} \right),$$

where the constant $30m$ accounts for the $10m$ Schwartz–Zippel evaluation losses plus $20m$ from the $\approx$-chain in the first and second terms. This proves 8.7. $\square$

Chapter 9

Pasting

Theorem 9.1 (Pasting)**.** *Let* (ψ, A, B, L) *be an* $(\varepsilon, \delta, \gamma)$*-good symmetric strategy for the* $(m+1, q, d)$ *low individual degree test. Let* $\{G^x\}_{x \in \mathbb{F}_q}$ *be projective submeasurements in* $\mathrm{PolySub}(m, q, d)$ *with the following properties:*

1. *(Completeness): If* $G = \mathbb{E}_x \sum_g G_g^x$*, then*
$$\langle \psi | G \otimes I | \psi \rangle \geq 1 - \kappa.$$
2. *(Consistency with A): On average over* $(u, x) \sim \mathbb{F}_q^{m+1}$*,*
$$A_a^{u,x} \otimes I \simeq_\zeta I \otimes G^x_{[g(u)=a]}.$$
3. *(Strong self-consistency): On average over* $x \sim \mathbb{F}_q$*,*
$$G_g^x \otimes I \approx_\zeta I \otimes G_g^x.$$
4. *(Boundedness): There exists a positive-semidefinite matrix* Z^x *for each* $x \in \mathbb{F}_q$ *such that*
$$\mathbb{E}_x \langle \psi | (I - G^x) \otimes Z^x | \psi \rangle \leq \zeta$$
and for each $x \in \mathbb{F}_q$ *and* $g \in \mathcal{P}(m, q, d)$*,*
$$Z^x \geq \mathbb{E}_u A^{u,x}_{g(u)}.$$

Let $k \geq 400md$ *be an integer, and set*

$$\nu = 100k^2 m \left(\varepsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32} \right),$$

$$\sigma = \kappa \left(1 + \frac{1}{100m} \right) + 2\nu + e^{-k/(80000m^2)}.$$

Then there exists a pasted measurement $H \in \mathrm{PolyMeas}(m+1, q, d)$ *with the following property.*

1. *(Consistency with A): On average over* $u \sim \mathbb{F}_q^{m+1}$*,*
$$A_a^u \otimes I \simeq_\sigma I \otimes H_{[h(u)=a]}.$$

Proof. The argument below is the paper proof, and the unrestricted Lean theorem now formalizes it directly. Let H be the submeasurement from Lemma 9.6, and fix an arbitrary polynomial $h^* \in \mathcal{P}(m+1,q,d)$. Define a measurement $H_{\mathrm{meas}} \in \mathrm{PolyMeas}(m+1,q,d)$ by

$$(H_{\mathrm{meas}})_h = \begin{cases} H_{h^*} + (I - H) & \text{if } h = h^*, \\ H_h & \text{otherwise.} \end{cases}$$

Then $\sum_h (H_{\mathrm{meas}})_h = H + (I - H) = I$. Moreover,

$$\begin{aligned}
&\mathbb{E}_u \sum_{a \neq b} \langle \psi | A^u_a \otimes (H_{\mathrm{meas}})_{[h(u)=b]} | \psi \rangle \\
&= \mathbb{E}_u \sum_a \sum_{h: h(u) \neq a} \langle \psi | A^u_a \otimes (H_{\mathrm{meas}})_h | \psi \rangle \\
&= \mathbb{E}_u \sum_a \sum_{h: h(u) \neq a} \langle \psi | A^u_a \otimes H_h | \psi \rangle + \mathbb{E}_u \sum_{a: h^*(u) \neq a} \langle \psi | A^u_a \otimes (I - H) | \psi \rangle \\
&\leq \mathbb{E}_u \sum_a \sum_{h: h(u) \neq a} \langle \psi | A^u_a \otimes H_h | \psi \rangle + \mathbb{E}_u \langle \psi | I \otimes (I - H) | \psi \rangle \\
&\leq \nu + \left(\kappa \left(1 + \frac{1}{100m} \right) + \nu + e^{-k/(80000m^2)} \right) && \text{(by 1 and 2)} \\
&= \sigma.
\end{aligned}$$

Thus $A^u_a \otimes I \simeq_\sigma I \otimes (H_{\mathrm{meas}})_{[h(u)=a]}$. □

Remark 9.2 (Restricted Lean form of the pasting theorem)*.* The Lean declaration linked here is the restricted nontrivial-regime form of Theorem 9.1. In addition to the hypotheses displayed in the source theorem, it assumes

$$\gamma \leq 1, \qquad \zeta \leq 1, \qquad d \leq q, \qquad 0 < d, \qquad 1 \leq k.$$

The paper theorem is stated in `references/ldt-paper/ld-pasting.tex`, lines 12–50. Lines 52–55 explain that the proof may restrict to $\varepsilon, \delta, \gamma, \zeta, d/q \leq 1$, since the complementary cases are trivial. The unrestricted source statement is linked from Theorem 9.1. All complementary branches, including the degree-zero case, are now proved in Lean.

The bound is trivial when at least one of ε, δ, γ, ζ, or d/q is at least 1, since then $\nu \geq 1$. In the remainder of the chapter we therefore assume

$$\varepsilon, \delta, \gamma, \zeta, d/q \leq 1.$$

Definition 9.3 (Nontrivial-regime pasting context)**.** The Lean context linked here is the nontrivial-regime context used after the reduction explained above. It consists of the data and hypotheses from the pasting theorem at the beginning of this chapter: an $(\varepsilon, \delta, \gamma)$-good symmetric strategy (ψ, A, B, L) for the $(m+1, q, d)$ low individual degree test, a family of projective submeasurements $\{G^x\}_{x \in \mathbb{F}_q}$ in $\mathrm{PolySub}(m,q,d)$ satisfying 1–4, an integer $k \geq 400md$, and the quantities ν and σ defined there. It also records the nontrivial-regime assumptions

$$\gamma \leq 1, \qquad \zeta \leq 1, \qquad d \leq q, \qquad 0 < d, \qquad 1 \leq k.$$

These extra fields are not assumptions of the unrestricted source theorem; they belong to the restricted proof stage. The unrestricted paper-facing Lean theorem remains `MIPStarRE.LDT.Pasting.ldPasting`.

9.0.1 From Measurements to Submeasurements

In the remainder of the chapter we only use the axis-parallel lines test in the last coordinate direction. We therefore write B^u_f for the line measurement on the line $\{(u,x) \mid x \in \mathbb{F}_q\}$ and first record the consistency relations that follow formally from the nontrivial-regime pasting context of Definition 9.3.

Conditioning the low individual degree test on the last coordinate direction gives, on average over $(u,x) \sim \mathbb{F}_q^{m+1}$,

$$A^{u,x}_a \otimes I \simeq_{(m+1)\varepsilon} I \otimes B^u_{[f(x)=a]}.$$

Hence 3.21 and $(m+1) \le 2m$ imply

$$A^{u,x}_a \otimes I \approx_{4m\varepsilon} I \otimes B^u_{[f(x)=a]}.$$

Using also the $(\varepsilon,\delta,\gamma)$-goodness assumption and 3.21,

$$I \otimes A^{u,x}_a \approx_{2\delta} A^{u,x}_a \otimes I \approx_{4m\varepsilon} I \otimes B^u_{[f(x)=a]}.$$

Therefore 3.27 yields

$$I \otimes A^{u,x}_a \approx_{8m\varepsilon+4\delta} I \otimes B^u_{[f(x)=a]}. \tag{9.1}$$

Applying 3.20 to 2 and 9.1 gives

$$G^x_{[g(u)=a]} \otimes I \simeq_{\nu_1} I \otimes B^u_{[f(x)=a]}, \tag{9.2}$$

where

$$\nu_1 = \zeta + \sqrt{8m\varepsilon + 4\delta}. \tag{9.3}$$

Finally, 8.7 gives

$$G^x_g G^y_h \otimes I \approx_{\nu_{\mathrm{com}}} G^y_h G^x_g \otimes I, \tag{9.4}$$

where

$$\nu_{\mathrm{com}} = 30m\left(\gamma^{1/4} + \zeta^{1/4} + (d/q)^{1/4}\right).$$

Lemma 9.4 (Slice measurements are consistent with line answers)**.** *In the nontrivial-regime pasting context of Definition 9.3, set*

$$\nu_1 = \zeta + \sqrt{8m\varepsilon + 4\delta}.$$

Then, on average over $(u,x) \sim \mathbb{F}_q^{m+1}$,

$$G^x_{[g(u)=a]} \otimes I \simeq_{\nu_1} I \otimes B^u_{[f(x)=a]}.$$

Proof. This is exactly 9.2, with 9.3. □

Lemma 9.5 (Lean auxiliary form of the vertical-line SDD estimate)**.** *The formalization records the point-to-vertical-line state-dependent-distance estimate with the vertical-line side written as the submeasurement family* `liftedVerticalLineAnswerFamily`*. The error term is the same $8m\varepsilon + 4\delta$ as in the complete measurement-valued construction.*

Proof. This is not a separate paper assertion. It is the representation of the same estimate obtained by replacing the complete vertical-line measurement family with its underlying submeasurement family, which is the form needed in the degree-zero branch. □

Lemma 9.6 (Pasted submeasurement)**.** *In the nontrivial-regime pasting context of Definition 9.3, there exists a submeasurement $H \in \mathrm{PolySub}(m+1, q, d)$ with the following properties.*

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

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

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

$$\langle \psi | H \otimes I | \psi \rangle \geq 1 - \kappa \left(1 + \frac{1}{100m}\right) - \nu - e^{-k/(80000m^2)}.$$

Proof. Let H be the pasted submeasurement constructed in Definition 9.13. Lemma 9.31 implies

$$H_{[h(u,x)=a]} \otimes I \simeq_{\nu_6} I \otimes B^u_{[f(x)=a]}. \tag{9.5}$$

Applying 3.20 to 9.5 and 9.1 gives

$$H_{[h(u,x)=a]} \otimes I \simeq_{\nu_6 + \sqrt{8m\varepsilon + 4\delta}} I \otimes A_a^{u,x}.$$

Since $\sqrt{8m\varepsilon + 4\delta} \leq 3m(\varepsilon^{1/32} + \delta^{1/32})$ for $\varepsilon, \delta \leq 1$, the total error is at most

$$\sqrt{8m\varepsilon + 4\delta} + \nu_6 \leq 47k^2 m \left(\varepsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right) \leq \nu,$$

which proves 1. The completeness claim is exactly Corollary 9.44. □

9.0.2 The Pasted Submeasurement

Definition 9.7 (Distinct tuples)**.** For $k \geq 1$, let $\mathsf{Distinct}_k \subseteq \mathbb{F}_q^k$ be the set of tuples $(x_1, \ldots, x_k)$ with pairwise distinct coordinates.

Proposition 9.8 (Distinct tuples versus independent tuples)**.** *If $x = (x_1, \ldots, x_k)$ is uniformly random in $\mathbb{F}_q^k$ and $y = (y_1, \ldots, y_k)$ is uniformly random in $\mathsf{Distinct}_k$, then*

$$d_{\mathrm{TV}}(x, y) \leq \frac{k^2}{q}.$$

Proof. The total variation distance is exactly the probability mass of the collision event for a uniformly random tuple, because the distinct-tuple distribution is the uniform distribution conditioned on landing in $\mathsf{Distinct}_k$. A uniformly random tuple lies outside $\mathsf{Distinct}_k$ only if two coordinates collide, so a union bound gives

$$\Pr[x \notin \mathsf{Distinct}_k] \leq \sum_{i=2}^{k} \frac{i-1}{q} = \frac{k(k-1)}{2q} \leq \frac{k^2}{q},$$

and this bounds the total variation distance. □

We first record the natural construction obtained by sampling $d+1$ slices and interpolating. It motivates the eventual argument but is not the construction used in the formal theorem.

The First Construction

The first construction samples $d+1$ slice outcomes $g_1, \dots, g_{d+1} \in \mathcal{P}(m,q,d)$, interpolates them to a global polynomial $h \in \mathcal{P}(m+1,q,d)$, and then averages over distinct interpolation points. Its consistency is governed by the commutation of the slice measurements, but its completeness is subtler. The heuristic comparison is

$$\langle\psi|H \otimes I|\psi\rangle \approx \langle\psi|G^{d+1} \otimes I|\psi\rangle.$$

Thus one would like to show

$$\langle\psi|G^{d+1} \otimes I|\psi\rangle \approx \langle\psi|G \otimes I|\psi\rangle. \tag{9.6}$$

The naive estimate would instead suggest

$$\langle\psi|G^{d+1} \otimes I|\psi\rangle \approx 1-(d+1)\kappa, \tag{9.7}$$

which is too weak for the induction.

The key input is that one can first show

$$\langle\psi|G^{d+2} \otimes I|\psi\rangle \approx_\Delta \langle\psi|G^{d+1} \otimes I|\psi\rangle, \tag{9.8}$$

for a small error $\Delta = \mathrm{poly}(m) \cdot \mathrm{poly}(\varepsilon, \delta, \gamma, \zeta, d/q)$. Writing the eigendecomposition $G = \sum_i \lambda_i |v_i\rangle\langle v_i|$ and the induced spectral distribution $\mu(i) = \langle\psi|(|v_i\rangle\langle v_i| \otimes I)|\psi\rangle$, this is equivalent to

$$\Delta \geq \langle\psi|G^{d+1} \otimes I|\psi\rangle - \langle\psi|G^{d+2} \otimes I|\psi\rangle = \mathbb{E}_{i\sim\mu}[\lambda_i^{d+1}(1-\lambda_i)]. \tag{9.9}$$

The remaining comparison between G and G^{d+1} is then reduced to a scalar inequality.

Lemma 9.9 (A scalar inequality for the first construction)**.** *For every real number $0 \leq \lambda \leq 1$,*

$$\lambda(1-\lambda^d) \leq 2(\lambda^{d+1}(1-\lambda))^{1/(d+1)}.$$

Proof. Compare $1-\lambda^d$ with $(1-\lambda)(1+\lambda+\cdots+\lambda^{d-1})$ and use that $d^{1/(d+1)} \leq 2$. ☐

Corollary 9.10 (First-construction completeness comparison)**.** *Lemma 9.9 converts (9.9) into the estimate*

$$\mathbb{E}_{i\sim\mu}[\lambda_i(1-\lambda_i^d)] \leq 2 \cdot \mathbb{E}_{i\sim\mu}\left[(\lambda_i^{d+1}(1-\lambda_i))^{1/(d+1)}\right] \leq 2 \cdot \Delta^{1/(d+1)},$$

where the middle step uses concavity of $a \mapsto a^{1/(d+1)}$. Equivalently,

$$\langle\psi|G \otimes I|\psi\rangle - \langle\psi|G^{d+1} \otimes I|\psi\rangle \leq 2 \cdot \Delta^{1/(d+1)}.$$

For constant d, this is exactly the heuristic completeness comparison for the first construction recorded in ‘references/ldt-paper/ld-pasting.tex‘.

Proof. Apply Lemma 9.9 pointwise to the spectral distribution of G, and then use the concavity of $a \mapsto a^{1/(d+1)}$. The displayed operator expression is the same quantity written in the eigenbasis of G. ☐

The Second Construction

Definition 9.11 (The incomplete part and the completed measurement)**.** For each $x \in \mathbb{F}_q$, write $G^x = \sum_g G_g^x$ and define the incomplete outcome $G_\perp^x = I - G^x$. The completed measurement $\widehat{G}^x$ has outcome set $\mathcal{P}(m, q, d) \cup \{\perp\}$ and is given by

$$\widehat{G}_g^x = G_g^x \quad (g \in \mathcal{P}(m, q, d)), \qquad \widehat{G}_\perp^x = G_\perp^x.$$

Definition 9.12 (Types)**.** A type is a vector $\tau \in \{0, 1\}^k$. Its weight $|\tau|$ records how many coordinates lie in $\mathcal{P}(m, q, d)$ rather than equal $\perp$.

Definition 9.13 (The pasted measurement)**.** Fix $k \geq d + 1$. For $x_1, \dots, x_k \in \mathbb{F}_q$ and outcomes $g_1, \dots, g_k \in \mathcal{P}(m, q, d) \cup \{\perp\}$, define

$$\widehat{H}_{g_1,\dots,g_k}^{x_1,\dots,x_k} = \widehat{G}_{g_1}^{x_1} \widehat{G}_{g_2}^{x_2} \cdots \widehat{G}_{g_k}^{x_k} \cdots \widehat{G}_{g_2}^{x_2} \widehat{G}_{g_1}^{x_1}.$$

If $(x_1, \dots, x_k) \in \mathsf{Distinct}_k$ and $h \in \mathcal{P}(m + 1, q, d)$, define

$$H_h^{x_1,\dots,x_k} = \sum_{\tau : |\tau| \geq d+1} \widehat{H}_{h_\tau}^{x_1,\dots,x_k},$$

where the tuple h_τ has i-th entry $h|_{x_i}$ when $\tau_i = 1$ and $\perp$ when $\tau_i = 0$. The pasted submeasurement is then

$$H_h = \mathbb{E}_{(x_1,\dots,x_k) \sim \mathsf{Distinct}_k} H_h^{x_1,\dots,x_k}.$$

Definition 9.14 (Completed-outcome reindexing)**.** The completed outcome tuples admit the reindexings obtained by separating the first coordinate, moving a third slice to the front, and identifying the one-coordinate split with the original slice question and outcome spaces.

Lemma 9.15 (Vertical restriction identities)**.** *Restricting an interpolated polynomial to a vertical line and then postprocessing at a point agrees with evaluating the corresponding slice. The same identities commute with averaging over indexed submeasurements.*

Proof. The compatibility calculation for the restriction and evaluation maps used in the construction of H expands the vertical restriction, the interpolation support condition, postprocessing, and averaging over indexed submeasurements. $\square$

9.0.3 Strong Self-Consistency and Commutation of $\widehat{G}$

Lemma 9.16 (Strong self-consistency of the complete part)**.** *In the nontrivial-regime pasting context of Definition 9.3,*

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

Proof. Because G is projective, 3.36 identifies the assumed strong self-consistency of the outcome-level family with

$$\mathbb{E}_x \sum_g \langle \psi | G_g^x \otimes G_g^x | \psi \rangle \geq \mathbb{E}_x \sum_g \langle \psi | G_g^x \otimes I | \psi \rangle - \frac{\zeta}{2}. \tag{9.10}$$

Hence

$$\begin{aligned}
&\mathbb{E}_x \| (G^x \otimes I - I \otimes G^x) |\psi\rangle \|^2 \\
&= 2\mathbb{E}_x \langle \psi | G^x \otimes I | \psi \rangle - 2\mathbb{E}_x \langle \psi | G^x \otimes G^x | \psi \rangle \\
&= 2\mathbb{E}_x \sum_g \langle \psi | G_g^x \otimes I | \psi \rangle - 2\mathbb{E}_x \sum_{g,h} \langle \psi | G_g^x \otimes G_h^x | \psi \rangle \\
&\leq 2\mathbb{E}_x \sum_g \langle \psi | G_g^x \otimes I | \psi \rangle - 2\mathbb{E}_x \sum_g \langle \psi | G_g^x \otimes G_g^x | \psi \rangle \leq \zeta
\end{aligned}$$

by 9.10. ☐

Corollary 9.17 (Strong self-consistency of the incomplete part)**.** *In the nontrivial-regime pasting context of Definition 9.3,*

$$G_{\perp}^{x} \otimes I \approx_{\zeta} I \otimes G_{\perp}^{x}.$$

Proof. Since $G_{\perp}^{x} = I - G^{x}$,

$$G_{\perp}^{x} \otimes I - I \otimes G_{\perp}^{x} = I \otimes G^{x} - G^{x} \otimes I,$$

so the claim is immediate from Lemma 9.16. ☐

Commutativity of $G_{\perp}$

Lemma 9.18 (Switcheroo contraction estimates)**.** *After one commutation of a G-factor through an auxiliary projective family, the two fourth-term contractions are bounded by the same self-consistency and commutation errors used in the switcheroo argument.*

Proof. After expanding the fourth switcheroo term, the inserted complete part forms a positive contraction. The left and right Cauchy–Schwarz side conditions are therefore the corresponding sums of $X_g X_g^{*}$ and $X_g^{*} X_g$, where $X_g = \sum_o G_{\perp}^{x} M_o^{y} G_g^{x} M_o^{y}$. Orthogonality of the projective slice family and the submeasurement inequalities for M and $G_{\perp}$ bound both sums by the identity. These are precisely the two contraction witnesses used in the fourth-term switcheroo chain. ☐

Lemma 9.19 (Commutativity with G_g^{x} implies commutativity with G^{x})**.** *Let $M = \{M_o^{x}\}$ be a projective submeasurement with outcomes in a set $\mathcal{O}$. Suppose that*

$$M_o^{x} \otimes I \approx_{\omega} I \otimes M_o^{x}, \tag{9.11}$$

and

$$G_g^{x} M_o^{y} \otimes I \approx_{\chi} M_o^{y} G_g^{x} \otimes I \tag{9.12}$$

on average over independent uniformly random $x, y \sim \mathbb{F}_q$. 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. Let $M = \mathbb{E}_y \sum_o M_o^{y}$. The error to bound is

$$\begin{aligned}
&\mathbb{E}_{x,y} \sum_o \|(G^{x} M_o^{y} - M_o^{y} G^{x}) \otimes I|\psi\rangle\|^2 \\
&= \mathbb{E}_{x,y} \sum_o \langle\psi| M_o^{y} (G^{x})^2 M_o^{y} \otimes I|\psi\rangle + \mathbb{E}_{x,y} \sum_o \langle\psi| G^{x} (M_o^{y})^2 G^{x} \otimes I|\psi\rangle \\
&\qquad - \mathbb{E}_{x,y} \sum_o \langle\psi| G^{x} M_o^{y} G^{x} M_o^{y} \otimes I|\psi\rangle - \mathbb{E}_{x,y} \sum_o \langle\psi| M_o^{y} G^{x} M_o^{y} G^{x} \otimes I|\psi\rangle. \qquad (9.13)
\end{aligned}$$

We compare all four terms with $\langle\psi|G \otimes M|\psi\rangle$.

For the first term in 9.13,

$$\begin{aligned}
\mathbb{E}_{x,y} \sum_o \langle\psi| M_o^{y} (G^{x})^2 M_o^{y} \otimes I|\psi\rangle &= \mathbb{E}_{x,y} \sum_o \langle\psi| M_o^{y} G^{x} M_o^{y} \otimes I|\psi\rangle \\
&= \mathbb{E}_{y} \sum_o \langle\psi| M_o^{y} G M_o^{y} \otimes I|\psi\rangle \\
&\approx_{2\sqrt{\omega}} \mathbb{E}_{y} \sum_o \langle\psi| G \otimes M_o^{y}|\psi\rangle \qquad \text{(by 3.32 and 9.11)} \\
&= \langle\psi|G \otimes M|\psi\rangle.
\end{aligned}$$

For the second term,

$$\begin{aligned}
\mathbb{E}_{x,y}\sum_o \langle\psi|G^x(M^y_o)^2G^x\otimes I|\psi\rangle &= \mathbb{E}_{x,y}\sum_o \langle\psi|G^xM^y_oG^x\otimes I|\psi\rangle \\
&= \mathbb{E}_x\langle\psi|G^xMG^x\otimes I|\psi\rangle \\
&\approx_{2\sqrt{\zeta}} \mathbb{E}_x\langle\psi|M\otimes G^x|\psi\rangle && \text{(by 3.32 and 9.16)} \\
&= \langle\psi|M\otimes G|\psi\rangle.
\end{aligned}$$

For the third term, first

$$\begin{aligned}
\mathbb{E}_{x,y}\sum_o \langle\psi|G^xM^y_oG^xM^y_o\otimes I|\psi\rangle &= \mathbb{E}_{x,y}\sum_{o,g} \langle\psi|G^xM^y_oG^x_gG^x_gM^y_o\otimes I|\psi\rangle \\
&\approx_{\sqrt{\chi}} \mathbb{E}_{x,y}\sum_{o,g} \langle\psi|G^xM^y_oG^x_gM^y_oG^x_g\otimes I|\psi\rangle. && (9.14)
\end{aligned}$$

This is the Cauchy–Schwarz estimate coming from 9.12. Next,

$$(9.14) \approx_{\sqrt{\zeta}} \mathbb{E}_{x,y}\sum_{o,g} \langle\psi|G^xM^y_oG^x_gM^y_o\otimes G^x_g|\psi\rangle. \tag{9.15}$$

Here the error is bounded using 3. Then

$$(9.15) \approx_{\sqrt{\zeta}} \mathbb{E}_{x,y}\sum_{o,g} \langle\psi|G^x_gG^xM^y_oG^x_gM^y_o\otimes I|\psi\rangle, \tag{9.16}$$

again by strong self-consistency of G^x_g. Since G is projective,

$$\begin{aligned}
(9.16) &= \mathbb{E}_{x,y}\sum_{o,g} \langle\psi|G^x_gM^y_oG^x_gM^y_o\otimes I|\psi\rangle \\
&\approx_{\sqrt{\chi}} \mathbb{E}_{x,y}\sum_{o,g} \langle\psi|M^y_oG^x_gG^x_gM^y_o\otimes I|\psi\rangle. && (9.17)
\end{aligned}$$

Finally,

$$\begin{aligned}
(9.17) &= \mathbb{E}_{x,y}\sum_{o,g} \langle\psi|M^y_oG^x_gM^y_o\otimes I|\psi\rangle \\
&= \mathbb{E}_y\sum_o \langle\psi|M^y_oGM^y_o\otimes I|\psi\rangle \\
&\approx_{2\sqrt{\omega}} \mathbb{E}_y\sum_o \langle\psi|G\otimes M^y_o|\psi\rangle && \text{(by 3.32 and 9.11)} \\
&= \langle\psi|G\otimes M|\psi\rangle.
\end{aligned}$$

Therefore

$$\mathbb{E}_{x,y}\sum_o \langle\psi|G^xM^y_oG^xM^y_o\otimes I|\psi\rangle \approx_{2\sqrt{\zeta}+2\sqrt{\omega}+2\sqrt{\chi}} \langle\psi|G\otimes M|\psi\rangle. \tag{9.18}$$

The fourth term in 9.13 is the Hermitian conjugate of the third term, so it satisfies the same bound. Summing the four errors gives

$$2\sqrt{\omega}+2\sqrt{\zeta}+2(2\sqrt{\zeta}+2\sqrt{\omega}+2\sqrt{\chi}) = 6\sqrt{\zeta}+6\sqrt{\omega}+4\sqrt{\chi}.$$

□

Remark 9.20. The Lean theorem corresponding to Lemma 9.19 takes one extra symmetry hypothesis, `hfix : swapDensity bi.density = bi.density`. This is the `density_swap` component of the paper's permutation-invariance assumption on the symmetric two-prover strategy; the stronger `swap_ev` field of `PermInvState` is not assumed by this theorem. The hypothesis is used to identify $\langle\psi|G \otimes M|\psi\rangle$ with $\langle\psi|M \otimes G|\psi\rangle$ at the center of the triangle-inequality chain.

Corollary 9.21 (Commutativity of the complete part)**.** *In the nontrivial-regime pasting context of Definition 9.3, the following commutation relations hold:*

$$G_g^x G^y \otimes I \approx_{\nu_2} G^y G_g^x \otimes I, \qquad G^x G^y \otimes I \approx_{\nu_2} G^y G^x \otimes I,$$

where

$$\nu_2 = 36m\left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16}\right).$$

Proof. By 9.4,

$$G_g^x G_h^y \otimes I \approx_{\nu_{\text{com}}} G_h^y G_g^x \otimes I. \tag{9.19}$$

Applying Lemma 9.19 to 9.19, with $\{M_o^y\}$ equal to the outcome family $\{G_h^y\}$, gives

$$G_g^x G^y \otimes I \approx_{\theta_1} G^y G_g^x \otimes I, \tag{9.20}$$

where $\theta_1 = 12\sqrt{\zeta} + 4\sqrt{\nu_{\text{com}}}$. Applying Lemma 9.19 again to 9.20, now with the one-outcome family $M_o^y = G^y$, gives

$$G^x G^y \otimes I \approx_{\theta_2} G^y G^x \otimes I,$$

where $\theta_2 = 12\sqrt{\zeta} + 4\sqrt{\theta_1}$.

Since $\nu_{\text{com}} = 30m(\gamma^{1/4} + \zeta^{1/4} + (d/q)^{1/4})$, these errors satisfy

$$\theta_1 \le 36m\left(\gamma^{1/8} + \zeta^{1/8} + (d/q)^{1/8}\right) \le \nu_2,$$
$$\theta_2 \le 36m\left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16}\right) = \nu_2.$$

□

Corollary 9.22 (Commutativity of the incomplete part)**.** *In the nontrivial-regime pasting context of Definition 9.3,*

$$G_g^x G_\perp^y \otimes I \approx_{\nu_2} G_\perp^y G_g^x \otimes I, \qquad G_\perp^x G_\perp^y \otimes I \approx_{\nu_2} G_\perp^y G_\perp^x \otimes I.$$

Proof. The identities

$$G_g^x G_\perp^y - G_\perp^y G_g^x = G^y G_g^x - G_g^x G^y$$

and

$$G_\perp^x G_\perp^y - G_\perp^y G_\perp^x = G^x G^y - G^y G^x$$

reduce both claims to Corollary 9.21. □

Putting Everything Together

Corollary 9.23 (Strong self-consistency and commutation of $\widehat{G}$)**.** *The completed measurements $\widehat{G}^x$ satisfy*

$$\widehat{G}_g^x \otimes I \approx_{2\zeta} I \otimes \widehat{G}_g^x, \tag{9.21}$$
$$\widehat{G}_g^x \widehat{G}_h^y \otimes I \approx_{\nu_3} \widehat{G}_h^y \widehat{G}_g^x \otimes I, \tag{9.22}$$

where

$$\nu_3 = 138m\left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16}\right).$$

Proof. For 9.21, split the outcome set into the genuine outcomes and $\perp$:

$$\mathbb{E}_x \sum_g \|(\widehat{G}_g^x \otimes I - I \otimes \widehat{G}_g^x)|\psi\rangle\|^2 \leq \zeta + \zeta = 2\zeta$$

by 3 and 9.17. For 9.22, the same decomposition yields four cases: polynomial–polynomial, $\perp$–$\perp$, polynomial–$\perp$, and $\perp$–polynomial. These contribute ν_{com} and three copies of ν_2, so

$$\nu_{\mathrm{com}} + 3\nu_2 \leq 138m\left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16}\right) = \nu_3.$$

□

Lemma 9.24 (Completed part bounded by a slice)**.** *For every $x \in \mathbb{F}_q$, the squared-distance defect of the one-outcome complete part is bounded by that of the original slice submeasurement:*

$$\left\|(G^x \otimes I - I \otimes G^x)|\psi\rangle\right\|^2 \;\leq\; \sum_g \left\|(G_g^x \otimes I - I \otimes G_g^x)|\psi\rangle\right\|^2,$$

where $G^x = \sum_g G_g^x$ is the total operator of the slice submeasurement.

Proof. This is the monotonicity estimate for the completed part of $\widehat{G}$: the completed contribution is one summand of the corresponding slice squared-distance defect, and the remaining summands are nonnegative. □

9.0.4 Sandwiching Lemmas

Lemma 9.25 (Completed sandwich normalization)**.**

1. ***Head–tail split identities.*** *For every $k \geq 0$, point tuple $x = (x_1, \dots, x_{k+1})$, and outcome tuple $g = (g_1, \dots, g_{k+1})$,*
$$\widehat{G}_{g_1}^{x_1} \cdots \widehat{G}_{g_{k+1}}^{x_{k+1}} = \widehat{G}_{g_1}^{x_1}(\widehat{G}_{g_2}^{x_2} \cdots \widehat{G}_{g_{k+1}}^{x_{k+1}}),$$
and the right half-sandwich analogously rotates the head to the tail.

2. ***Sum-of-adjoint-products bounds.*** *For every $r \geq 0$ and $x = (x_1, \dots, x_r) \in \mathbb{F}_q^r$,*
$$\sum_{g_1,\dots,g_r} (\widehat{G}_{g_1}^{x_1} \cdots \widehat{G}_{g_r}^{x_r})^\dagger (\widehat{G}_{g_1}^{x_1} \cdots \widehat{G}_{g_r}^{x_r}) \;\leq\; I,$$
and the same bound holds for the reverse-order product family.

3. ***Generic tensor contraction.*** *If families $\{A_a\}_a$ and $\{B_b\}_b$ satisfy $\sum_a A_a^\dagger A_a \leq I$ and $\sum_b B_b^\dagger B_b \leq I$, then*
$$\sum_{a,b} (A_a \otimes B_b)^\dagger (A_a \otimes B_b) \leq I.$$

Proof. The split identities follow from the completed-outcome equivalences. The normalization inequalities are the corresponding submeasurement bounds for the left and right half-products after reindexing the sums. □

Lemma 9.26 (One-step completed sandwich commutation)**.** *Given the completed-measurement self-consistency and commutation estimates from Corollary 9.23, the following one-step estimates hold.*

1. ***Self-consistency step.*** *For every $r \geq 0$,*

$$\mathbb{E}_{x,y,z}\mathbb{E}_{x_1,\dots,x_r}\sum_{g}\left\|(\widehat{G}^z_g \otimes I - I \otimes \widehat{G}^z_g)|\psi\rangle\right\|^2 \;\leq\; 2\zeta.$$

2. ***Commutation step.*** *For every $r \geq 0$,*

$$\mathbb{E}_{x,y}\mathbb{E}_{x_1,\dots,x_r}\sum_{g,h}\sum_{g_1,\dots,g_r}\left\|(\widehat{G}^x_g\widehat{G}^y_h \otimes \widehat{G}^{x_r}_{g_r}\cdots\widehat{G}^{x_1}_{g_1} - \widehat{G}^y_h\widehat{G}^x_g \otimes \widehat{G}^{x_r}_{g_r}\cdots\widehat{G}^{x_1}_{g_1})|\psi\rangle\right\|^2 \;\leq\; \nu_3.$$

3. ***Two-factor base case.*** *For $k=2$, the commutation chain reduces to the same tensor-placement convention with the reverse tail product equal to the identity, hence to the pairwise estimate ν_3.*

4. ***Split equivalences.*** *The half-sandwich families are equivalent, under the head–tail reindexing of Lemma 9.25, to the chain-families used in the move estimates above, with ordered products on the first tensor factor and reverse products on the second tensor factor where these appear in the commutation step.*

These are the elementary building blocks assembled by the recursive chain construction in the following lemma.

Proof. For the one-step commutation move, rewrite the relevant head–tail split, apply the self-consistency and commutation estimates for $\widehat{G}$, and assemble the resulting comparison by the triangle inequality for $\approx_\delta$. □

Lemma 9.27 (Commuting past multiple completed measurements)**.** *For every $k \geq 2$,*

$$\widehat{G}^{x_1}_{g_1}\widehat{G}^{x_2}_{g_2}\cdots\widehat{G}^{x_k}_{g_k} \otimes I \approx_{\nu_4} \widehat{G}^{x_2}_{g_2}\cdots\widehat{G}^{x_k}_{g_k}\widehat{G}^{x_1}_{g_1} \otimes I,$$

where

$$\nu_4 = 426k^2m\left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16}\right).$$

Proof. Repeatedly move the leftmost factor across the remaining product, using 9.21 and 9.22 at each step and summing the resulting errors with 3.27. The paper's calculation gives

$$3k(4k\zeta + k\nu_3) \leq 426k^2m\left(\gamma^{1/16} + \zeta^{1/16} + (d/q)^{1/16}\right).$$

□

Lemma 9.28 (Line-interpolation mismatch estimates)**.** *If the tuple of completed slices does not interpolate to the sampled vertical-line value, then some active coordinate witnesses a line-value mismatch. Concretely, for every direction u and injective tuple $\mathbf{x} = (x_1, \dots, x_k)$, let $\widehat{H}^{\mathbf{x}}|_u$ be the vertical-line restriction of the pasted interpolation family. Then*

$$\mathrm{Defect}_\psi(\widehat{H}^{\mathbf{x}}|_u,\, B^u) \;\leq\; \mathrm{Bad}(\psi; u, \mathbf{x}) \;\leq\; \sum_{i=1}^{k}\mathrm{Defect}_\psi(\widehat{H}^{(i)}(u,\mathbf{x}),\, B^{(i)}(u,\mathbf{x})),$$

where $\mathrm{Defect}_\psi(A,B)$ *denotes the bipartite consistency defect* $\max(0,\, \langle\psi|A_{\mathrm{tot}}\otimes B_{\mathrm{tot}}|\psi\rangle - \sum_a\langle\psi|A_a\otimes B_a|\psi\rangle)$,

$$\mathrm{Bad}(\psi; u, \mathbf{x}) = \sum_f\langle\psi|\Big(\sum_{\substack{g_1,\dots,g_k \\ \exists i:\, g_i\neq\perp,\ g_i(u)\neq f(x_i)}}\widehat{H}^{\mathbf{x}}_{g_1,\dots,g_k}\Big) \otimes B^u_f|\psi\rangle,$$

and $\widehat{H}^{(i)}, B^{(i)}$ are the one-point line comparison families (Lemma 9.30) restricted to coordinate i, and $\mathrm{Bad}(\psi; u, \mathbf{x}) \leq 1$.

Proof. The proof first identifies a failed vertical-line interpolation with a particular active coordinate whose slice value disagrees with the line answer. The operator sum over all such failures is then bounded by the sum of the corresponding one-coordinate line defects, using the vertical restriction identities above. □

Lemma 9.29 (Averaging the line-interpolation defect)**.** *Replacing distinct tuples by independent tuples costs the total-variation error k^2/q from Proposition 9.8. After averaging over u and $\mathbf{x} \sim \mathsf{Distinct}_k$, the expected bad mass and the expected pasted interpolation defect are bounded by the one-point sandwich error:*

$$\mathbb{E}_u \, \mathbb{E}_{\mathbf{x}\sim\mathsf{Distinct}_k} \; \mathrm{Bad}(\psi; u, \mathbf{x}) \;\leq\; k \cdot \nu_5 \;+\; \frac{k^2}{q},$$

and consequently

$$\mathbb{E}_u \, \mathbb{E}_{\mathbf{x}\sim\mathsf{Distinct}_k} \; \mathrm{Defect}_\psi(\widehat{H}^{\mathbf{x}}|_u,\, B^u) \;\leq\; k \cdot \nu_5 \;+\; \frac{k^2}{q},$$

where $\nu_5 = 43km(\varepsilon^{1/32} + \cdots + (d/q)^{1/32})$ (Lemma 9.30). Using $1/q \leq (d/q)^{1/32}$ and $m \geq 1$, this is further bounded by $44k^2m(\varepsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32})$ as used in Lemma 9.31.

Proof. The proof compares the distinct-tuple average with the independent-tuple average, paying the total-variation cost from Proposition 9.8. The independent average separates into the one-coordinate line defects from the previous lemma, and the scalar inequality $1/q \leq (d/q)^{1/32}$ absorbs the sampling error into the stated pasting parameter. □

Lemma 9.30 (Consistency of the sandwiched measurement with one line value)**.** *For any $1 \leq i \leq k$,*

$$\mathbb{E}_u \mathbb{E}_{x_1,\ldots,x_k} \sum_{g_1,\ldots,g_k \,:\, g_i \neq \perp} \; \sum_{a \neq g_i(u)} \langle \psi | \widehat{H}^{x_1,\ldots,x_k}_{g_1,\ldots,g_k} \otimes B^u_{[f(x_i)=a]} | \psi \rangle \leq \nu_5,$$

where

$$\nu_5 = 43km \left(\varepsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32} \right).$$

Proof. Summing out the coordinates to the right of i gives

$$\begin{aligned} &\mathbb{E}_u \mathbb{E}_{x_1,\ldots,x_k} \sum_{g_1,\ldots,g_k \,:\, g_i \neq \perp} \; \sum_{a \neq g_i(u)} \langle \psi | \widehat{H}^{x_1,\ldots,x_k}_{g_1,\ldots,g_k} \otimes B^u_{[f(x_i)=a]} | \psi \rangle \\ &= \mathbb{E}_u \mathbb{E}_{x_1,\ldots,x_i} \sum_{g_1,\ldots,g_i \,:\, g_i \neq \perp} \; \sum_{a \neq g_i(u)} \langle \psi | \widehat{H}^{x_1,\ldots,x_i}_{g_1,\ldots,g_i} \otimes B^u_{[f(x_i)=a]} | \psi \rangle. \end{aligned} \tag{9.23}$$

Write $\widehat{G}^{x_{<i}}_{g_{<i}} = \widehat{G}^{x_1}_{g_1} \cdots \widehat{G}^{x_{i-1}}_{g_{i-1}}$. Then

$$\begin{aligned} (9.23) &= \mathbb{E}_u \mathbb{E}_{x_1,\ldots,x_i} \sum_{g_1,\ldots,g_i \,:\, g_i \neq \perp} \langle \psi | \widehat{G}^{x_{<i}}_{g_{<i}} \widehat{G}^{x_i}_{g_i} \widehat{G}^{x_i}_{g_i} (\widehat{G}^{x_{<i}}_{g_{<i}})^\dagger \otimes (I - B^u_{[f(x_i)=g_i(u)]}) | \psi \rangle \\ &\approx_{\sqrt{\nu_4}} \mathbb{E}_u \mathbb{E}_{x_1,\ldots,x_i} \sum_{g_1,\ldots,g_i \,:\, g_i \neq \perp} \langle \psi | \widehat{G}^{x_i}_{g_i} \widehat{G}^{x_{<i}}_{g_{<i}} \widehat{G}^{x_i}_{g_i} (\widehat{G}^{x_{<i}}_{g_{<i}})^\dagger \otimes (I - B^u_{[f(x_i)=g_i(u)]}) | \psi \rangle. \end{aligned} \tag{9.24}$$

The Cauchy–Schwarz argument uses

$$\mathbb{E}_u \mathbb{E}_{x_1,\ldots,x_i} \sum_{g_1,\ldots,g_i} \langle \psi | ((\widehat{G}^{x_{<i}}_{g_{<i}} \widehat{G}^{x_i}_{g_i} - \widehat{G}^{x_i}_{g_i} \widehat{G}^{x_{<i}}_{g_{<i}})(\widehat{G}^{x_i}_{g_i} (\widehat{G}^{x_{<i}}_{g_{<i}})^\dagger - (\widehat{G}^{x_{<i}}_{g_{<i}})^\dagger \widehat{G}^{x_i}_{g_i})) \otimes I | \psi \rangle, \tag{9.25}$$

which is at most ν_4 by Lemma 9.27. A second Cauchy–Schwarz step gives

$$(9.24) \approx_{\sqrt{\nu_4}} \mathbb{E}_u \mathbb{E}_{x_1,\dots,x_i} \sum_{g_1,\dots,g_i\,:\,g_i\neq\perp} \langle\psi|\widehat{G}^{x_i}_{g_i} \widehat{G}^{x_{<i}}_{g_{<i}} (\widehat{G}^{x_{<i}}_{g_{<i}})^\dagger \widehat{G}^{x_i}_{g_i} \otimes (I - B^u_{[f(x_i)=g_i(u)]})|\psi\rangle. \tag{9.26}$$

Summing over $g_1,\dots,g_{i-1}$ collapses the middle factor to I, so 9.26 becomes

$$\mathbb{E}_u \mathbb{E}_{x_1,\dots,x_i} \sum_{g_i\,:\,g_i\neq\perp} \langle\psi|\widehat{G}^{x_i}_{g_i} \otimes (I - B^u_{[f(x_i)=g_i(u)]})|\psi\rangle \le \nu_1$$

by 9.2. Thus the total error is

$$\nu_1 + 2\sqrt{\nu_4} \le 43km\left(\varepsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right).$$

□

9.0.5 Consistency of H with A

Lemma 9.31 (Consistency of the pasted measurement with the line measurement)**.** *The pasted submeasurement satisfies*

$$H_{[h|_u=f]} \otimes I \simeq_{\nu_6} I \otimes B^u_f,$$

where

$$\nu_6 = 44k^2m\left(\varepsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right).$$

Proof. The inconsistency with B^u expands as

$$\begin{aligned}
&\mathbb{E}_u \sum_{f\neq f'} \langle\psi|H_{[h|_u=f']} \otimes B^u_f|\psi\rangle \\
&= \mathbb{E}_u \sum_h \sum_{f\neq h|_u} \langle\psi|H_h \otimes B^u_f|\psi\rangle \\
&= \mathbb{E}_u \mathbb{E}_{x_1,\dots,x_k} \sum_h \sum_{f\neq h|_u} \langle\psi|H^{x_1,\dots,x_k}_h \otimes B^u_f|\psi\rangle \\
&= \mathbb{E}_u \mathbb{E}_{x_1,\dots,x_k} \sum_h \sum_{w:|w|\ge d+1} \sum_{(g_1,\dots,g_k)=h_w} \sum_{f\neq h|_u} \langle\psi|\widehat{H}^{x_1,\dots,x_k}_{g_1,\dots,g_k} \otimes B^u_f|\psi\rangle.
\end{aligned} \tag{9.27}$$

If $|w| \ge d+1$ and $f \neq h|_u$, then some active coordinate i must satisfy $g_i \neq \perp$ and $g_i(u) \neq f(x_i)$. Hence

$$(9.27) \le \mathbb{E}_u \mathbb{E}_{x_1,\dots,x_k} \sum_{g_1,\dots,g_k} \sum_{\substack{f:\exists i,\ g_i\neq\perp,\\ g_i(u)\neq f(x_i)}} \langle\psi|\widehat{H}^{x_1,\dots,x_k}_{g_1,\dots,g_k} \otimes B^u_f|\psi\rangle. \tag{9.28}$$

Passing from independent tuples to distinct tuples costs at most k^2/q by Lemma 9.8. A union bound over the offending coordinate then gives

$$(9.28) \le \sum_{i=1}^k \mathbb{E}_u \mathbb{E}_{y_1,\dots,y_k} \sum_{g_1,\dots,g_k\,:\,g_i\neq\perp} \sum_{a\neq g_i(u)} \langle\psi|\widehat{H}^{y_1,\dots,y_k}_{g_1,\dots,g_k} \otimes B^u_{[f(y_i)=a]}|\psi\rangle \le k\nu_5.$$

Therefore

$$\frac{k^2}{q} + k\nu_5 \le 44k^2m\left(\varepsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right) = \nu_6.$$

□

Corollary 9.32 (Consistency of H with A)**.** *On average over $(u,x) \sim \mathbb{F}_q^{m+1}$,*

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

Proof. Lemma 9.31 implies

$$H_{[h(u,x)=a]} \otimes I \simeq_{\nu_6} I \otimes B^u_{[f(x)=a]}.$$

Applying 3.20 to this relation and 9.1 gives

$$H_{[h(u,x)=a]} \otimes I \simeq_{\nu_6+\sqrt{8m\varepsilon+4\delta}} I \otimes A_a^{u,x}.$$

Since

$$\sqrt{8m\varepsilon+4\delta} + \nu_6 \le 47k^2 m \left(\varepsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right) \le \nu,$$

the claim follows. □

Lemma 9.33 (Point consistency from vertical-line consistency)**.** *Let H be any polynomial-valued submeasurement. Suppose that the strategy is good, that $\gamma, \zeta \ge 0$, that $k \ge 1$, and that the restriction of H to every vertical line is consistent with the vertical-line measurement with error ν_6. Then the pointwise evaluations of H are consistent with Alice's point measurement with error ν.*

Proof. The proof transports the vertical-line consistency estimate to point evaluations and combines it with the good-strategy point-to-line comparison. The displayed numerical assumptions are used only to absorb the intermediate errors into the final parameter ν. □

Corollary 9.34 (Consistency of the completed H with A)**.** *Let H_{meas} be the completion of H at the fallback outcome. On average over $(u,x) \sim \mathbb{F}_q^{m+1}$,*

$$(H_{\text{meas}})_{[h(u,x)=a]} \otimes I \simeq_{\sigma} I \otimes A_a^{u,x}.$$

Proof. Corollary 9.32 gives

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

Corollary 9.44 bounds the missing mass of H by

$$\kappa \cdot \left(1 + \frac{1}{100m}\right) + \nu + e^{-k/(80000m^2)}.$$

Adding this completion mass to the ν-bound gives

$$\nu + \kappa \cdot \left(1 + \frac{1}{100m}\right) + \nu + e^{-k/(80000m^2)} = \sigma.$$

□

Lemma 9.35 (Completion from submeasurement consistency)**.** *Let H be any polynomial-valued submeasurement. Suppose that the pointwise evaluations of H are consistent with Alice's point measurement with error ν, and that H has the lower mass bound used in Corollary 9.44. Then completing H at the fallback polynomial preserves point consistency with the enlarged error appearing in the induction statement.*

Proof. The completed measurement differs from the original submeasurement only at the fallback polynomial. The additional term is bounded by the missing-mass estimate and then added to the given point-consistency error. □

9.0.6 Completeness of H

Definition 9.36 (Outcomes of a fixed type)**.** For a type $\tau \in \{0,1\}^k$, let $\mathsf{Outcomes}_\tau$ be the set of tuples $(g_1, \dots, g_k)$ such that $g_i \in \mathcal{P}(m,q,d)$ when $\tau_i = 1$ and $g_i = \perp$ when $\tau_i = 0$. If $x_1, \dots, x_k \in \mathbb{F}_q$, let $\mathsf{Global}_\tau(x)$ be the subset of $\mathsf{Outcomes}_\tau$ arising from restrictions of a single polynomial in $\mathcal{P}(m+1,q,d)$, and let $\overline{\mathsf{Global}_\tau(x)} = \mathsf{Outcomes}_\tau \setminus \mathsf{Global}_\tau(x)$.

Lemma 9.37 (Summing over all outcomes)**.** *If $x_1, \dots, x_k$ are sampled independently and uniformly from $\mathbb{F}_q$, then*

$$\langle \psi | H \otimes I | \psi \rangle \approx_{\nu_7} \mathbb{E}_{x_1,\dots,x_k} \sum_{\tau : |\tau| \geq d+1} \sum_{(g_1,\dots,g_k) \in \mathsf{Outcomes}_\tau} \langle \psi | \widehat{H}^{x_1,\dots,x_k}_{g_1,\dots,g_k} \otimes I | \psi \rangle,$$

where

$$\nu_7 = 46k^2 m \left(\varepsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32} \right).$$

Proof. Let $(y_1, \dots, y_k) \sim \mathsf{Distinct}_k$. By definition,

$$\begin{aligned}
\langle \psi | H \otimes I | \psi \rangle &= \sum_h \langle \psi | H_h \otimes I | \psi \rangle \\
&= \mathbb{E}_{y_1,\dots,y_k} \sum_h \langle \psi | H_h^{y_1,\dots,y_k} \otimes I | \psi \rangle \\
&= \mathbb{E}_{y_1,\dots,y_k} \sum_h \sum_{\tau : |\tau| \geq d+1} \sum_{(g_1,\dots,g_k) = h_\tau} \langle \psi | \widehat{H}^{y_1,\dots,y_k}_{g_1,\dots,g_k} \otimes I | \psi \rangle \\
&= \mathbb{E}_{y_1,\dots,y_k} \sum_{\tau : |\tau| \geq d+1} \sum_{(g_1,\dots,g_k) \in \mathsf{Global}_\tau(y)} \langle \psi | \widehat{H}^{y_1,\dots,y_k}_{g_1,\dots,g_k} \otimes I | \psi \rangle. \qquad (9.29)
\end{aligned}$$

We now remove the restriction to globally consistent tuples:

$$(9.29) \approx_{\frac{k^2}{q} + k\nu_5 + \frac{md}{q}} \mathbb{E}_{y_1,\dots,y_k} \sum_{\tau : |\tau| \geq d+1} \sum_{(g_1,\dots,g_k) \in \mathsf{Outcomes}_\tau} \langle \psi | \widehat{H}^{y_1,\dots,y_k}_{g_1,\dots,g_k} \otimes I | \psi \rangle. \qquad (9.30)$$

Since the right-hand side is at least the left-hand side, it suffices to bound the difference:

$$\begin{aligned}
(9.30) - (9.29) &= \mathbb{E}_{y_1,\dots,y_k} \sum_{\tau : |\tau| \geq d+1} \sum_{(g_1,\dots,g_k) \in \overline{\mathsf{Global}_\tau(y)}} \langle \psi | \widehat{H}^{y_1,\dots,y_k}_{g_1,\dots,g_k} \otimes I | \psi \rangle \\
&= \mathbb{E}_u \mathbb{E}_{y_1,\dots,y_k} \sum_{\tau : |\tau| \geq d+1} \sum_f \sum_{(g_1,\dots,g_k) \in \overline{\mathsf{Global}_\tau(y)}} \langle \psi | \widehat{H}^{y_1,\dots,y_k}_{g_1,\dots,g_k} \otimes B^u_f | \psi \rangle. \qquad (9.31)
\end{aligned}$$

We next insert an indicator recording consistency along the sampled line:

$$(9.31) \approx_{\frac{k^2}{q} + k\nu_5} \mathbb{E}_u \mathbb{E}_{y_1,\dots,y_k} \sum_{\tau : |\tau| \geq d+1} \sum_f \sum_{(g_1,\dots,g_k) \in \overline{\mathsf{Global}_\tau(y)}} \langle \psi | \widehat{H}^{y_1,\dots,y_k}_{g_1,\dots,g_k} \otimes B^u_f | \psi \rangle \cdot \mathbf{1}[[]\, \forall i \in \tau,\ f(y_i) = g_i(u)]. \qquad (9.32)$$

The discarded part is

$$\mathbb{E}_u \mathbb{E}_{y_1,\dots,y_k} \sum_{\tau : |\tau| \geq d+1} \sum_f \sum_{(g_1,\dots,g_k) \in \overline{\mathsf{Global}_\tau(y)}} \langle \psi | \widehat{H}^{y_1,\dots,y_k}_{g_1,\dots,g_k} \otimes B^u_f | \psi \rangle \cdot \mathbf{1}[[]\, \exists i \in \tau,\ f(y_i) \neq g_i(u)], \qquad (9.33)$$

and after switching to independent tuples and applying Lemma 9.30 coordinatewise, this contributes at most $k^2/q + k\nu_5$.

Let $\mathsf{Consistent}_\tau(g,y,u)$ denote the event that there exists a degree-d polynomial f with $f(y_i) = g_i(u)$ for all $i \in \tau$. Then

$$\mathbf{1}[[]\, \forall i \in \tau,\ f(y_i) = g_i(u)] \le \mathbf{1}[[]\, \mathsf{Consistent}_\tau(g,y,u)],$$

so 9.32 is bounded by

$$\begin{aligned}
&\mathbb{E}_u \mathbb{E}_{y_1,\dots,y_k} \sum_{\tau:\,|\tau|\ge d+1} \sum_{(g_1,\dots,g_k)\in \mathsf{Global}_\tau(y)} \langle \psi | \widehat{H}^{y_1,\dots,y_k}_{g_1,\dots,g_k} \otimes I | \psi \rangle \cdot \mathbf{1}[[]\, \mathsf{Consistent}_\tau(g,y,u)] \\
&= \mathbb{E}_{y_1,\dots,y_k} \sum_{\tau:\,|\tau|\ge d+1} \sum_{(g_1,\dots,g_k)\in \mathsf{Global}_\tau(y)} \langle \psi | \widehat{H}^{y_1,\dots,y_k}_{g_1,\dots,g_k} \otimes I | \psi \rangle \cdot \Pr_u[\mathsf{Consistent}_\tau(g,y,u)]. \qquad (9.34)
\end{aligned}$$

For a fixed globally inconsistent tuple, choose $d+1$ active coordinates and let h^* be the unique interpolant through them. Since the tuple is not globally consistent, some active coordinate i^* satisfies $g_{i^*} \ne h^*|_{y_{i^*}}$, and then Schwartz–Zippel gives

$$\Pr_u[\mathsf{Consistent}_\tau(g,y,u)] \le \Pr_u[g_{i^*}(u) = h^*(u, y_{i^*})] \le \frac{md}{q}.$$

Therefore 9.34 is at most md/q, which proves 9.30. A final application of Lemma 9.8 replaces distinct tuples by independent tuples, giving a total error of

$$2\frac{k^2}{q} + \frac{md}{q} + k\nu_5 \le 46k^2 m \left(\varepsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right) = \nu_7.$$

□

Definition 9.38 (Truncated type sums)**.** For $1 \le \ell \le k+1$ and a tail type $\tau_{\ge \ell} \in \{0,1\}^{k-\ell+1}$, define

$$S_{\tau_{\ge\ell}} = \sum_{\tau_{<\ell}\,:\,|\tau|\ge d+1} G^{|\tau_{<\ell}|}(I-G)^{(\ell-1)-|\tau_{<\ell}|}.$$

This records the contribution of all prefixes that can still be completed to total weight at least $d+1$ once the tail $\tau_{\ge\ell}$ is fixed.

Lemma 9.39 (Recurrence for truncated type sums)**.** *Each $S_{\tau_{\ge\ell}}$ is Hermitian, positive semidefinite, and bounded by I. Moreover, if $\tau_{>\ell}$ is a tail type of length $k-\ell$, then*

$$S_{\tau_{>\ell}} = \sum_{\tau_\ell \in \{0,1\}} S_{\tau_{\ge\ell}} G^{\tau_\ell}(I-G)^{1-\tau_\ell}.$$

Proof. Every summand is a polynomial in the commuting positive operators G and $I-G$, so $S_{\tau_{\ge\ell}}$ is Hermitian and positive semidefinite. Summing over all prefixes gives

$$S_{\tau_{\ge\ell}} \le (G + (I-G))^{\ell-1} = I.$$

Splitting the next type bit according to whether $\tau_\ell = 1$ or $\tau_\ell = 0$ gives the stated recurrence. □

Lemma 9.40 (From the pasted sum to a polynomial in G)**.** *Let $|\psi_{\mathrm{bi}}\rangle$ be a bipartite state on $\iota \otimes \iota$. If $x_1,\dots,x_k$ are sampled independently and uniformly from $\mathbb{F}_q$, then*

$$\mathbb{E}_{x_1,\dots,x_k} \sum_{\tau:\,|\tau|\ge d+1} \sum_{(g_1,\dots,g_k)\in \mathsf{Outcomes}_\tau} \langle \psi_{\mathrm{bi}} | \widehat{H}^{x_1,\dots,x_k}_{g_1,\dots,g_k} \otimes I | \psi_{\mathrm{bi}} \rangle \approx_{\nu_8} \sum_{r=d+1}^{k} \binom{k}{r} \langle \psi_{\mathrm{bi}} | G^r (I-G)^{k-r} \otimes I | \psi_{\mathrm{bi}} \rangle,$$

where

$$\nu_8 = 46k^2 m \left(\gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right).$$

Proof. For a type $\tau \in \{0,1\}^k$, write $\tau_{<\ell}$, $\tau_{>\ell}$, $\tau_{\le\ell}$, and $\tau_{\ge\ell}$ for the obvious truncations, and similarly for tuples g. Also write

$$\widehat{G}^{x_{\ge\ell}}_{g_{\ge\ell}} = \widehat{G}^{x_\ell}_{g_\ell} \cdots \widehat{G}^{x_k}_{g_k}.$$

Then

$$\widehat{G}^{x_{\ge\ell}}_{g_{\ge\ell}} = \widehat{G}^{x_\ell}_{g_\ell} \widehat{G}^{x_{>\ell}}_{g_{>\ell}}, \tag{9.35}$$

and

$$\widehat{H}^{x_{\ge\ell}}_{g_{\ge\ell}} = \widehat{G}^{x_{\ge\ell}}_{g_{\ge\ell}} (\widehat{G}^{x_{\ge\ell}}_{g_{\ge\ell}})^\dagger. \tag{9.36}$$

Write $\mathsf{O}_\tau = \mathsf{Outcomes}_\tau$. The main iterative step is that for each $1 \le \ell \le k$,

$$\begin{aligned} &\mathbb{E}_{x_{\ge\ell}} \sum_{\tau : |\tau| \ge d+1} \sum_{g_{\ge\ell} \in \mathsf{O}_{\tau_{\ge\ell}}} \langle\psi| \widehat{H}^{x_{\ge\ell}}_{g_{\ge\ell}} \otimes \left(G^{|\tau_{<\ell}|}(I-G)^{(\ell-1)-|\tau_{<\ell}|}\right) |\psi\rangle \\ \approx_{2\sqrt{2\zeta}+2\sqrt{\nu_4}}\ &\mathbb{E}_{x_{>\ell}} \sum_{\tau : |\tau| \ge d+1} \sum_{g_{>\ell} \in \mathsf{O}_{\tau_{>\ell}}} \langle\psi| \widehat{H}^{x_{>\ell}}_{g_{>\ell}} \otimes \left(G^{|\tau_{\le\ell}|}(I-G)^{\ell-|\tau_{\le\ell}|}\right) |\psi\rangle. \end{aligned} \tag{9.37}$$

Iterating 9.37 for $\ell = 1, \dots, k$ produces the Bernoulli polynomial

$$\sum_{r=d+1}^{k} \binom{k}{r} \langle\psi| G^r (I-G)^{k-r} \otimes I |\psi\rangle,$$

and the accumulated error is

$$k(2\sqrt{2\zeta} + 2\sqrt{\nu_4}) \le 46k^2 m \left(\gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32}\right) = \nu_8.$$

This corrects the displayed arithmetic in the paper: the iterated adjacent estimate contributes k copies of the $2\sqrt{\nu_4}$ term, and the definition of ν_4 already contains a factor k^2.

To prove 9.37, define for each $1 \le \ell \le k+1$ and each tail type $\tau_{\ge\ell}$,

$$S_{\tau_{\ge\ell}} = \sum_{\tau_{<\ell} : |\tau| \ge d+1} G^{|\tau_{<\ell}|}(I-G)^{(\ell-1)-|\tau_{<\ell}|}. \tag{9.38}$$

Then 9.37 may be rewritten as

$$\begin{aligned} &\mathbb{E}_{x_{\ge\ell}} \sum_{\tau_{\ge\ell}} \sum_{g_{\ge\ell} \in \mathsf{O}_{\tau_{\ge\ell}}} \langle\psi| \widehat{H}^{x_{\ge\ell}}_{g_{\ge\ell}} \otimes S_{\tau_{\ge\ell}} |\psi\rangle \\ \approx_{2\sqrt{2\zeta}+2\sqrt{\nu_4}}\ &\mathbb{E}_{x_{>\ell}} \sum_{\tau_{>\ell}} \sum_{g_{>\ell} \in \mathsf{O}_{\tau_{>\ell}}} \langle\psi| \widehat{H}^{x_{>\ell}}_{g_{>\ell}} \otimes S_{\tau_{>\ell}} |\psi\rangle. \end{aligned} \tag{9.39}$$

Remark 9.41 (Lean proof-stage notation). In Lean, the type-restricted averaged submeasurement, the per-tail operator/mass, the aggregated stage quantity, and the final Bernoulli-tail scalar from this proof are recorded by the declarations above. Their indexing is shifted to the repository's 0-based convention: Lean stage ℓ is the paper's stage $\ell + 1$.

The operators $S_{\tau_{\ge\ell}}$ are Hermitian and positive semidefinite, and

$$\begin{aligned} S_{\tau_{\ge\ell}} &= \sum_{\tau_{<\ell} : |\tau| \ge d+1} G^{|\tau_{<\ell}|}(I-G)^{(\ell-1)-|\tau_{<\ell}|} \\ &\le \sum_{\tau_{<\ell}} G^{|\tau_{<\ell}|}(I-G)^{(\ell-1)-|\tau_{<\ell}|} \\ &= (G + (I-G))^{\ell-1} = I. \end{aligned} \tag{9.40}$$

For any $\tau_\ell \in \{0,1\}$,

$$\mathbb{E}_{x_\ell} \sum_{g_\ell \in \mathsf{O}_{\tau_\ell}} \widehat{G}^{x_\ell}_{g_\ell} = G^{\tau_\ell}(I-G)^{1-\tau_\ell}. \tag{9.41}$$

Therefore, for every $\tau_{>\ell}$,

$$\begin{aligned} \sum_{\tau_\ell} S_{\tau_{\geq \ell}} \cdot \Big(\mathbb{E}_{x_\ell} \sum_{g_\ell \in \mathsf{O}_{\tau_\ell}} \widehat{G}^{x_\ell}_{g_\ell} \Big) &= \sum_{\tau_\ell} S_{\tau_{\geq \ell}} G^{\tau_\ell}(I-G)^{1-\tau_\ell} \\ &= S_{\tau_{>\ell}}. \end{aligned} \tag{9.42}$$

Moreover,

$$S_{\tau_{\geq \ell}} \cdot \Big(\mathbb{E}_{x_\ell} \sum_{g_\ell \in \mathsf{O}_{\tau_\ell}} \widehat{G}^{x_\ell}_{g_\ell} \Big) \cdot S_{\tau_{\geq \ell}} \leq \mathbb{E}_{x_\ell} \sum_{g_\ell \in \mathsf{O}_{\tau_\ell}} \widehat{G}^{x_\ell}_{g_\ell}, \tag{9.43}$$

because $S_{\tau_{\geq \ell}}$ commutes with G and $(I-G)$ and is bounded by I.

We now prove 9.39. Writing $\widehat{H}$ as a sandwich of $\widehat{G}$ operators and moving the rightmost $\widehat{G}^{x_\ell}_{g_\ell}$ to the second tensor factor gives

$$\begin{aligned} &\mathbb{E}_{x_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | \widehat{H}^{x_{\geq \ell}}_{g_{\geq \ell}} \otimes S_{\tau_{\geq \ell}} | \psi \rangle \\ &= \mathbb{E}_{x_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | (\widehat{G}^{x_{\geq \ell}}_{g_{\geq \ell}} (\widehat{G}^{x_{>\ell}}_{g_{>\ell}})^\dagger \widehat{G}^{x_\ell}_{g_\ell}) \otimes S_{\tau_{\geq \ell}} | \psi \rangle \\ &\approx_{\sqrt{2\zeta}} \mathbb{E}_{x_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | (\widehat{G}^{x_{\geq \ell}}_{g_{\geq \ell}} (\widehat{G}^{x_{>\ell}}_{g_{>\ell}})^\dagger) \otimes (S_{\tau_{\geq \ell}} \widehat{G}^{x_\ell}_{g_\ell}) | \psi \rangle. \end{aligned} \tag{9.44}$$

The corresponding Cauchy–Schwarz bound is

$$\begin{aligned} &\Big| \mathbb{E}_{x_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | (\widehat{G}^{x_{\geq \ell}}_{g_{\geq \ell}} \otimes S_{\tau_{\geq \ell}}) ((\widehat{G}^{x_{>\ell}}_{g_{>\ell}})^\dagger \otimes I)(\widehat{G}^{x_\ell}_{g_\ell} \otimes I - I \otimes \widehat{G}^{x_\ell}_{g_\ell}) | \psi \rangle \Big| \\ &\leq \sqrt{\mathbb{E}_{x_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | \widehat{H}^{x_{\geq \ell}}_{g_{\geq \ell}} \otimes (S_{\tau_{\geq \ell}})^2 | \psi \rangle} \cdot \sqrt{\mathbb{E}_{x_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | (\widehat{G}^{x_\ell}_{g_\ell} \otimes I - I \otimes \widehat{G}^{x_\ell}_{g_\ell})(\widehat{H}^{x_{>\ell}}_{g_{>\ell}} \otimes I)(\widehat{G}^{x_\ell}_{g_\ell} \otimes I - I \otimes } \end{aligned}$$

The first square root is at most 1 because $S_{\tau_{\geq \ell}} \leq I$ and $\widehat{H}$ is a submeasurement, and the second is at most 2ζ by 9.21.

Next, commute the leftmost $\widehat{G}^{x_\ell}_{g_\ell}$ across the left half-sandwich:

$$\begin{aligned} (9.44) &= \mathbb{E}_{x_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | (\widehat{G}^{x_\ell}_{g_\ell} \widehat{G}^{x_{>\ell}}_{g_{>\ell}} (\widehat{G}^{x_{>\ell}}_{g_{>\ell}})^\dagger) \otimes (S_{\tau_{\geq \ell}} \widehat{G}^{x_\ell}_{g_\ell}) | \psi \rangle \\ &\approx_{\sqrt{\nu_4}} \mathbb{E}_{x_{\geq \ell}} \sum_{\tau_{\geq \ell}} \sum_{(g_\ell, g_{>\ell}) \in \mathsf{O}_{\tau_{\geq \ell}}} \langle \psi | (\widehat{G}^{x_{>\ell}}_{g_{>\ell}} \widehat{G}^{x_\ell}_{g_\ell} (\widehat{G}^{x_{>\ell}}_{g_{>\ell}})^\dagger) \otimes (S_{\tau_{\geq \ell}} \widehat{G}^{x_\ell}_{g_\ell}) | \psi \rangle. \end{aligned} \tag{9.46}$$

The associated Cauchy–Schwarz estimate is

$$\left|\mathbb{E}_{x_{\geq\ell}}\sum_{\tau_{\geq\ell}}\sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}}\langle\psi|([\widehat{G}^{x_{>\ell}}_{g_{>\ell}},\widehat{G}^{x_\ell}_{g_\ell}]\otimes I)((\widehat{G}^{x_{>\ell}}_{g_{>\ell}})^\dagger\otimes(S_{\tau_{\geq\ell}}\widehat{G}^{x_\ell}_{g_\ell}))|\psi\rangle\right|$$
$$\leq\sqrt{\mathbb{E}_{x_{\geq\ell}}\sum_{\tau_{\geq\ell}}\sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}}\langle\psi|[\widehat{G}^{x_{>\ell}}_{g_{>\ell}},\widehat{G}^{x_\ell}_{g_\ell}][\widehat{G}^{x_{>\ell}}_{g_{>\ell}},\widehat{G}^{x_\ell}_{g_\ell}]^\dagger\otimes I|\psi\rangle}\cdot\sqrt{\mathbb{E}_{x_{\geq\ell}}\sum_{\tau_{\geq\ell}}\sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}}\langle\psi|\widehat{H}^{x_{>\ell}}_{g_{>\ell}}\otimes(\widehat{G}^{x_\ell}_{g_\ell}(S_{\tau_{\geq\ell}})^2\widehat{G}^{x_\ell}_{g_\ell})|\psi\rangle}. \tag{9.47}$$

The first square root is bounded by ν_4 via Lemma 9.27, and the second by 1 since $S_{\tau_{\geq\ell}}\leq I$.

Continue commuting:

$$(9.46)\approx_{\sqrt{\nu_4}}\mathbb{E}_{x_{\geq\ell}}\sum_{\tau_{\geq\ell}}\sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}}\langle\psi|(\widehat{G}^{x_{>\ell}}_{g_{>\ell}}(\widehat{G}^{x_{>\ell}}_{g_{>\ell}})^\dagger\widehat{G}^{x_\ell}_{g_\ell})\otimes(S_{\tau_{\geq\ell}}\widehat{G}^{x_\ell}_{g_\ell})|\psi\rangle. \tag{9.48}$$

This time the Cauchy–Schwarz step is

$$\left|\mathbb{E}_{x_{\geq\ell}}\sum_{\tau_{\geq\ell}}\sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}}\langle\psi|(\widehat{G}^{x_{>\ell}}_{g_{>\ell}}\otimes(S_{\tau_{\geq\ell}}\widehat{G}^{x_\ell}_{g_\ell}))([(\widehat{G}^{x_{>\ell}}_{g_{>\ell}})^\dagger,\widehat{G}^{x_\ell}_{g_\ell}]\otimes I)|\psi\rangle\right|$$
$$\leq\sqrt{\mathbb{E}_{x_{\geq\ell}}\sum_{\tau_{\geq\ell}}\sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}}\langle\psi|\widehat{H}^{x_{>\ell}}_{g_{>\ell}}\otimes(S_{\tau_{\geq\ell}}\widehat{G}^{x_\ell}_{g_\ell}S_{\tau_{\geq\ell}})|\psi\rangle}\cdot\sqrt{\mathbb{E}_{x_{\geq\ell}}\sum_{\tau_{\geq\ell}}\sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}}\langle\psi|[(\widehat{G}^{x_{>\ell}}_{g_{>\ell}})^\dagger,\widehat{G}^{x_\ell}_{g_\ell}]^\dagger[(\widehat{G}^{x_{>\ell}}_{g_{>\ell}})^\dagger,\widehat{G}^{x_\ell}_{g_\ell}]\otimes I|\psi\rangle}. \tag{9.49}$$

By 9.43, the first square root is at most 1, and the second matches the first square root in 9.47, hence is at most ν_4.

Finally move the remaining $\widehat{G}^{x_\ell}_{g_\ell}$ to the second tensor factor:

$$(9.48)\approx_{\sqrt{2\zeta}}\mathbb{E}_{x_{\geq\ell}}\sum_{\tau_{\geq\ell}}\sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}}\langle\psi|\widehat{H}^{x_{>\ell}}_{g_{>\ell}}\otimes(S_{\tau_{\geq\ell}}\widehat{G}^{x_\ell}_{g_\ell}\widehat{G}^{x_\ell}_{g_\ell})|\psi\rangle. \tag{9.50}$$

The omitted Cauchy–Schwarz bound reuses the first square root from 9.49 and the second square root from 9.45. Since $\widehat{G}$ is projective,

$$\begin{aligned}(9.50)&=\mathbb{E}_{x_{\geq\ell}}\sum_{\tau_{\geq\ell}}\sum_{(g_\ell,g_{>\ell})\in\mathsf{O}_{\tau_{\geq\ell}}}\langle\psi|\widehat{H}^{x_{>\ell}}_{g_{>\ell}}\otimes(S_{\tau_{\geq\ell}}\widehat{G}^{x_\ell}_{g_\ell})|\psi\rangle\\&=\mathbb{E}_{x_{>\ell}}\sum_{\tau_{>\ell}}\sum_{g_{>\ell}\in\mathsf{O}_{\tau_{>\ell}}}\langle\psi|\widehat{H}^{x_{>\ell}}_{g_{>\ell}}\otimes\Big(\sum_{\tau_\ell}S_{\tau_{\geq\ell}}\Big(\mathbb{E}_{x_\ell}\sum_{g_\ell\in\mathsf{O}_{\tau_\ell}}\widehat{G}^{x_\ell}_{g_\ell}\Big)\Big)|\psi\rangle\\&=\mathbb{E}_{x_{>\ell}}\sum_{\tau_{>\ell}}\sum_{g_{>\ell}\in\mathsf{O}_{\tau_{>\ell}}}\langle\psi|\widehat{H}^{x_{>\ell}}_{g_{>\ell}}\otimes S_{\tau_{>\ell}}|\psi\rangle, &&\text{(by 9.42)}\end{aligned}$$

which is exactly 9.39. □

Lemma 9.42 (Context specializations for the pasted measurement)**.** *The nontrivial-regime pasting context specializes the consistency of H with A and the conversion from the pasted sum to the polynomial in G to the parameters and error bounds fixed in Definition 9.3.*

Proof. These are direct specializations of the two preceding construction theorems to the notation and numerical hypotheses packaged in the nontrivial-regime pasting context. □

Lemma 9.43 (Matrix Chernoff bound for the Bernoulli polynomial)**.** *Let $0 < \theta < 1$ and integers $k, d > 0$ with $k \geq 2d/\theta$. Define*

$$F(X) = \sum_{r=d+1}^{k} \binom{k}{r} X^r (I - X)^{k-r}.$$

If X is Hermitian, $0 \leq X \leq I$, and $\langle \psi | X \otimes I | \psi \rangle \geq 1 - \kappa$, then

$$\langle \psi | F(X) \otimes I | \psi \rangle \geq 1 - \frac{\kappa}{1-\theta} - e^{-\theta^2 k/2}.$$

Proof. Let ρ be the reduced state of $|\psi\rangle$ on one prover's subsystem, and write the eigendecomposition $X = \sum_i \lambda_i |v_i\rangle\langle v_i|$. This defines a probability distribution $\mu(i) = \langle v_i | \rho | v_i \rangle$, with

$$\mathbb{E}_{i\sim\mu} \lambda_i = \mathrm{Tr}(X\rho) = \langle \psi | X \otimes I | \psi \rangle \geq 1 - \kappa.$$

Equivalently, $\mathbb{E}_{i\sim\mu}(1 - \lambda_i) \leq \kappa$, so Markov's inequality shows that most of the spectral mass of X lies on eigenvalues at least θ:

$$\Pr_{i\sim\mu}[\lambda_i > \theta] \geq 1 - \frac{\kappa}{1-\theta}. \tag{9.51}$$

For a scalar $p \in [d/k, 1]$, the quantity

$$F(p) = \sum_{r=d+1}^{k} \binom{k}{r} p^r (1-p)^{k-r}$$

is the probability of observing at least $d+1$ successes in k Bernoulli(p) trials. Applying the scalar Chernoff bound eigenvalue-by-eigenvalue therefore yields

$$F(p) \geq 1 - \exp\Big(-2\Big(p - \frac{d}{k}\Big)^2 k\Big). \tag{9.52}$$

Hence

$$\begin{aligned}
\langle \psi | F(X) \otimes I | \psi \rangle &= \mathrm{Tr}(F(X)\rho) \\
&= \sum_i F(\lambda_i) \langle v_i | \rho | v_i \rangle \\
&= \mathbb{E}_{i\sim\mu} F(\lambda_i) \\
&\geq \Pr_{i\sim\mu}[\lambda_i \geq \theta] \cdot F(\theta) \\
&\geq \Big(1 - \frac{\kappa}{1-\theta}\Big) \cdot F(\theta) && \text{(by 9.51)} \\
&\geq \Big(1 - \frac{\kappa}{1-\theta}\Big) \cdot \Big(1 - \exp\Big(-2\Big(\theta - \frac{d}{k}\Big)^2 k\Big)\Big). && (9.53)
\end{aligned}$$

Since $(1-b)(1-c) \geq 1 - b - c$ for $b, c \geq 0$, 9.53 implies

$$\langle \psi | F(X) \otimes I | \psi \rangle \geq 1 - \frac{\kappa}{1-\theta} - \exp\Big(-2\Big(\theta - \frac{d}{k}\Big)^2 k\Big).$$

Finally, $k \geq 2d/\theta$ implies $d/k \leq \theta/2$, hence $(\theta - d/k)^2 \geq \theta^2/4$, which yields the claimed bound. $\square$

Corollary 9.44 (Completeness of the pasted submeasurement)**.** *If $k \geq 400md$, then*

$$\langle\psi|H \otimes I|\psi\rangle \geq 1 - \kappa\left(1 + \frac{1}{100m}\right) - \nu - e^{-k/(80000m^2)}.$$

Proof. Lemma 9.37 gives

$$\langle\psi|H \otimes I|\psi\rangle \approx_{\nu_7} \mathbb{E}_{x_1,\ldots,x_k} \sum_{\tau:\,|\tau|\geq d+1} \sum_{(g_1,\ldots,g_k)\in\mathsf{Outcomes}_\tau} \langle\psi|\widehat{H}^{x_1,\ldots,x_k}_{g_1,\ldots,g_k} \otimes I|\psi\rangle.$$

Lemma 9.40 turns the right-hand side into

$$\sum_{r=d+1}^{k} \binom{k}{r} \langle\psi|G^r(I-G)^{k-r} \otimes I|\psi\rangle$$

up to error ν_8. Apply Lemma 9.43 with $\theta = 1/(200m)$. Since $k \geq 400md$, the hypothesis $k \geq 2d/\theta$ is satisfied, and the resulting lower bound is

$$1 - \frac{\kappa}{1 - 1/(200m)} - e^{-k/(80000m^2)} \geq 1 - \kappa\left(1 + \frac{1}{100m}\right) - e^{-k/(80000m^2)}.$$

The paper then absorbs the approximation losses $\nu_7 + \nu_8$ into ν. ∎

Chapter 10

The main induction step

10.1 Restricting to one slice

Definition 10.1 (Appending a line at height x)**.** For $x \in \mathbb{F}_q$, let $\mathrm{append}_x(\ell)$ be the line in $\mathbb{F}_q^{m+1}$ obtained by appending the coordinate x to every point of a line $\ell \subseteq \mathbb{F}_q^m$. If $f\colon \ell \to \mathbb{F}_q$, let $\mathrm{append}_x(f)$ be the polynomial on $\mathrm{append}_x(\ell)$ defined by $\mathrm{append}_x(f)(u,x) = f(u)$.

Definition 10.2 (x-restricted strategy)**.** Let (ψ, A, B, L) be a symmetric strategy for the $(m+1, q, d)$ low individual degree test. For each $x \in \mathbb{F}_q$, the x-restricted strategy (ψ, A^x, B^x, L^x) for the (m, q, d) test is defined by

$$(A^x)^u_a = A^{u,x}_a, \qquad (B^x)^\ell_f = B^{\mathrm{append}_x(\ell)}_{\mathrm{append}_x(f)}, \qquad (L^x)^\ell_f = L^{\mathrm{append}_x(\ell)}_{\mathrm{append}_x(f)}.$$

Lemma 10.3 (Restricted success probabilities)**.** *Let (ψ, A, B, L) be an $(\varepsilon, \delta, \gamma)$-good symmetric strategy for the $(m+1, q, d)$ test. For each $x \in \mathbb{F}_q$, let (ψ, A^x, B^x, L^x) be the restricted strategy and let $\varepsilon_x, \delta_x, \gamma_x$ be its three failure probabilities. Then*

$$\mathbb{E}_x \varepsilon_x \le \frac{m+1}{m}\varepsilon, \qquad \mathbb{E}_x \delta_x \le \delta, \qquad \mathbb{E}_x \gamma_x \le \frac{m+1}{m}\gamma.$$

Proof. The self-consistency test restricts exactly to the slice at height x. For the axis-parallel and diagonal-line tests, conditioning on the sampled line not being parallel to the last coordinate identifies the test with the corresponding restricted game. ☐

Definition 10.4 (Average restricted errors)**.** The average restricted errors are the expectations over $x \in \mathbb{F}_q$ of the axis-parallel, self-consistency, and restricted diagonal-line failure probabilities of the x-restricted strategies.

Lemma 10.5 (Weighted restricted-probability bounds)**.** *The weighted bounds from Lemma 10.3 imply the restricted-probability data used in the successor step.*

Proof. The weighted axis-parallel and diagonal-line estimates are divided by the transverse-direction weight, while the self-consistency estimate is exactly the average restricted self-consistency probability. These three estimates form the restricted-probability record used in the successor step. ☐

10.2 The inductive construction

Theorem 10.6 (Self-improvement)**.** *Let (ψ, A, B, L) be an $(\varepsilon, \delta, \gamma)$-good symmetric strategy for the (m, q, d) low individual degree test. Let $G \in \mathrm{PolyMeas}(m, q, d)$ be a measurement with the following property:*

1. *(Consistency with A): On average over $u \sim \mathbb{F}_q^m$,*
$$A_a^u \otimes I \simeq_\nu I \otimes G_{[g(u)=a]}.$$

Let
$$\zeta = 3000m\left(\varepsilon^{1/32} + \delta^{1/32} + (d/q)^{1/32}\right).$$
Then there exists a projective submeasurement $H \in \mathrm{PolySub}(m, q, d)$ with the following properties:

1. *(Completeness): If $H = \sum_h H_h$, then*
$$\langle\psi|H \otimes I|\psi\rangle \geq (1 - \nu) - \zeta.$$
2. *(Consistency with A): On average over $u \sim \mathbb{F}_q^m$,*
$$A_a^u \otimes I \simeq_\zeta I \otimes H_{[h(u)=a]}.$$
3. *(Strong self-consistency):*
$$H_h \otimes I \approx_\zeta I \otimes H_h.$$
4. *(Boundedness): There exists a positive semidefinite operator Z such that*
$$\langle\psi|Z \otimes (I - H)|\psi\rangle \leq \zeta$$
and for each $h \in \mathcal{P}(m, q, d)$,
$$Z \geq \mathbb{E}_u A_{h(u)}^u.$$

Proof. This is Theorem 7.11, with the conclusions reformulated in the notation used by the induction step. Since the output is projective, Lemma 3.36 converts the strong self-consistency statement to the usual bipartite formulation. □

Remark 10.7 (Lean realization of induction self-improvement). The Lean declaration for Theorem 10.6 has the paper hypotheses, in particular that G is a complete polynomial measurement. The labelled theorem in `references/ldt-paper/inductive_step.tex:249--286` states the input as a submeasurement, while the proved form in `references/ldt-paper/self_improvement.tex:635--671` states it as a measurement. The blueprint and Lean follow the proved measurement-valued form, which is also the form used at the point where the theorem is applied in `references/ldt-paper/inductive_step.tex` The proof applies `SelfImprovement.selfImprovement` to G, then uses `selfImprovementInInductionSectionConcl` to transport the fields. The Section 7 SDP slackness input is supplied by the formalized theorem `sdp_statement_with_slackness`; it is not an additional hypothesis of the induction-section theorem.

Theorem 10.8 (Pasting)**.** *Let (ψ, A, B, L) be an $(\varepsilon, \delta, \gamma)$-good symmetric strategy for the $(m + 1, q, d)$ low individual degree test. Let $\{G^x\}_{x \in \mathbb{F}_q}$ be projective submeasurements in* $\mathrm{PolySub}(m, q, d)$ *with the following properties:*

1. *(Completeness): If $G = \mathbb{E}_x \sum_g G^x_g$, then*

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

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

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

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

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

4. *(Boundedness): There are positive semidefinite operators Z^x such that*

$$\mathbb{E}_x \langle \psi | (I - G^x) \otimes Z^x | \psi \rangle \leq \zeta$$

and for each $x \in \mathbb{F}_q$ and $g \in \mathcal{P}(m, q, d)$,

$$Z^x \geq \mathbb{E}_u A^{u,x}_{g(u)}.$$

Let $k \geq 400md$ and set

$$\nu = 100k^2 m \left(\varepsilon^{1/32} + \delta^{1/32} + \gamma^{1/32} + \zeta^{1/32} + (d/q)^{1/32} \right)$$

and

$$\sigma = \kappa \left(1 + \frac{1}{100m} \right) + 2\nu + e^{-k/(80000m^2)}.$$

Then there exists a measurement $H \in \mathrm{PolyMeas}(m + 1, q, d)$ with the following property:

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

$$A^u_a \otimes I \simeq_\sigma I \otimes H_{[h(u)=a]}.$$

Proof. This is exactly Theorem 9.1, restated in the induction chapter because it is the pasting input used in the proof of Theorem 10.14. The Lean declaration linked above applies the unrestricted formal pasting theorem and transports its point-consistency field to the induction-section conclusion record. □

Remark 10.9 (Unrestricted Lean form of induction-section pasting). The Lean declaration linked to Theorem 10.8 is now the unrestricted source-facing restatement. It calls the formal Section 9 theorem `Pasting.ldPasting` and therefore carries exactly the source hypotheses displayed above, together with the faithful formal boundary hypotheses needed to interpret the objects in Lean.

The nontrivial-regime proof reduction remains represented in the Chapter 9 declaration `Pasting.ldPastingNontrivial`. That helper assumes

$$\gamma \leq 1, \qquad \zeta \leq 1, \qquad d \leq q, \qquad 0 < d, \qquad 1 \leq k.$$

These inequalities are not hypotheses of the paper theorem: in `references/ldt-paper/ld-pasting.tex`, lines 12–50 state the unrestricted result, while lines 52–55 use the inequalities only to reduce the proof to the nontrivial regime. The formal theorem linked from the blueprint has already incorporated the complementary cases through `Pasting.ldPasting`.

Remark 10.10 (Lean self-improvement slice transport for the successor step). The successor step records the slice-wise self-improvement output in `SelfImprovementData`, and the answer-valued route records the analogous output in `AnswerSelfImprovementData`. Their constructors call `selfImprovementInInductionSection`; they no longer carry a separate bundle of Section 7 proof-stage data as an input. The `SliceStrategyTransport` structures are narrow formal interfaces: they assume concrete per-slice symmetric strategies and measurement-transport data, while averaged point-operator compatibility is derived structurally from point-measurement transport. They then convert the remaining data into the ordinary or answer-valued self-improvement output.

Definition 10.11 (Pasting data for the successor step)**.** The successor step forms the pasting input by combining the restricted probability estimates with the slice-wise self-improvement data. The Lean construction applies the unrestricted induction-section pasting theorem directly, and the small-error constructor supplies the scalar side conditions $\gamma \leq 1$, $\zeta \leq 1$, and $d \leq q$ internally. Thus the successor proof can call the small-error pasting constructor without carrying $0 < d$, $1 \leq k$, $\gamma \leq 1$, $\zeta \leq 1$, or $d \leq q$ as separate proof inputs. The answer-valued small-error assembly then converts the answer-valued slice data to this pasting interface and applies the checked stage-data theorem.

Lemma 10.12 (Base case of main induction)**.** *Under the hypotheses of Theorem 10.14 with $m = 1$, there exists a measurement $G \in \mathrm{PolyMeas}(1, q, d)$ such that, on average over $u \sim \mathbb{F}_q$,*

$$A^u_a \otimes I \simeq_\varepsilon I \otimes G_{[g(u)=a]},$$

which is stronger than the bound required by Theorem 10.14 for $m = 1$. Lean also proves the same one-dimensional construction for the answer-valued strategy interface used by the recursive-slice route.

Proof. For $m = 1$, there is only one axis-parallel line ℓ in $\mathbb{F}_q$, so the line measurement B^ℓ is already an element of $\mathrm{PolyMeas}(1, q, d)$. Because the strategy fails the axis-parallel lines test with probability at most ε, the consistency bound follows directly from the axis-parallel consistency hypothesis. □

Lemma 10.13 (Large-error branch of main induction)**.** *If the target error parameter in Theorem 10.14 is at least 1, then the conclusion follows from a distinguished trivial polynomial measurement. Lean also proves the same large-error branch for the answer-valued strategy interface used by the recursive-slice route.*

Proof. The normalized bipartite consistency defect of two submeasurements is at most 1. Hence any polynomial measurement, in particular the measurement concentrated on one polynomial, satisfies the claimed consistency relation when the allowed error is at least 1. □

Theorem 10.14 (Main induction)**.** *This is the corrected numerical form of the theorem; the paper-gap note [con26b] records the replacement of the printed bound $k \geq md$ by $k \geq 400md$. Let (ψ, A, B, L) be an $(\varepsilon, \delta, \gamma)$-good symmetric strategy for the (m, q, d) low individual degree test. Let $k \geq 400md$ and set*

$$\nu = 1000k^2m^2 \left(\varepsilon^{1/1024} + \delta^{1/1024} + \gamma^{1/1024} + (d/q)^{1/1024}\right)$$

and

$$\sigma = m^2 \left(\nu + e^{-k/(80000m^2)}\right).$$

Then there exists a measurement $G \in \mathrm{PolyMeas}(m, q, d)$ such that, on average over $u \sim \mathbb{F}_q^m$,

$$A^u_a \otimes I \simeq_\sigma I \otimes G_{[g(u)=a]}.$$

Proof. The base case, the large-error branch, and the small-error successor construction are all formalized for the corrected large-k interface. The paper prints the weaker bound $k \geq md$; the implication needed to apply the pasting theorem is false in the interval $md \leq k < 400md$. Thus the factor 400 is recorded as a confirmed statement correction, as documented in [con26b]. □

Proposition 10.15 (Small-error successor construction for the induction theorem)**.** *This is the nontrivial successor branch in the corrected large-k induction interface. The ordinary successor theorem now reduces to the internal answer-valued induction theorem. That theorem supplies the predecessor induction argument for the restricted slices by a genuine strong induction on the dimension. The answer-valued pasting invocation is recorded in Proposition 10.18, and the answer-valued self-improvement construction is checked by Proposition 10.17. This is a construction internal to the proof, not an assumption of Theorem 10.14. The recursive slice route applies also when* $\mathrm{d} = 0$*; the successor reduction no longer needs a separate degree-zero family construction.*

Proof. The checked assembly theorem `mainInductionSuccessorNext_ofSmallErrorConstruction` reduces the ordinary branch to the answer-valued induction theorem through the answer-carrier construction. The answer-valued induction theorem supplies the predecessor argument internally; the answer-valued slice self-improvement construction and the averaged family-field assembly are checked, and the answer-valued pasting construction below supplies the final pasting input. □

Proposition 10.16 (Predecessor induction argument for successor slices)**.** *In the small-error successor branch, the proof must apply the induction hypothesis in dimension m to the restricted strategies indexed by $x \in \mathbb{F}_q$, with the same integer k. The successor large-k hypothesis implies the predecessor large-k side condition in the corrected formal interface, and the nontrivial small-error branch supplies $k \geq 1$. This is the genuine recursive use of the theorem; it is not an additional hypothesis on the successor strategy. The answer-valued restriction map for an answer-valued successor strategy is now explicit in Lean; it preserves the full function-valued diagonal answer on the slice, so it does not use the ordinary low-degree realization that is unavailable for arbitrary answer-valued diagonal measurements. The corresponding restricted-probability theorem for answer-valued successor strategies is also checked: the axis-parallel and diagonal slice averages are derived from the ambient answer-valued goodness assumptions, and the diagonal calculation uses the full function-valued answer rather than an ordinary polynomial-valued realization. Lean now also checks the recursive application itself: after these restricted probabilities are constructed, the predecessor answer-valued induction hypothesis gives the main-induction conclusion for every answer-valued restricted slice. The theorem `answerMainInduction` now supplies this predecessor hypothesis by strong induction on the dimension. Thus the predecessor induction argument is a checked internal construction, not an additional theorem hypothesis.*

Proof. The Lean proof is contained in the simultaneous answer-valued induction theorem and its restricted-slice constructors. □

Proposition 10.17 (Answer-valued slice self-improvement)**.** *The restricted slice strategy is represented in Lean by an answer-valued strategy. The slice self-improvement data are constructed directly: replace only the diagonal part by an inert ordinary covariant measurement and apply the induction-section self-improvement theorem in the form whose hypotheses are the axis-parallel and point self-consistency estimates. Since the Section 9 conclusion is independent of the diagonal-line error parameter, no low-degree support theorem for the answer-valued diagonal outcomes is needed for this step. Such a support theorem would still be necessary for the stronger route which realizes the answer-valued diagonal measurement as an ordinary polynomial-valued covariant measurement, but that realization is not used here.*

Proof. The Lean proof is the carrier construction described in the statement. The diagonal component of the carrier is a fixed covariant projective measurement; the hypotheses consumed by self-improvement are the axis-parallel and point self-consistency estimates of the answer-valued restricted slice. □

Proposition 10.18 (Answer-valued pasting construction for the successor step)**.** *After the recursive answer-valued slice measurements and the slice-wise self-improvement outputs have been averaged, Lean has a polynomial family which is complete, point-consistent with the actual answer-valued point measurement, strongly self-consistent, and bounded by the witnesses obtained from self-improvement. The boundedness input is presently typed through the point-equivalent ordinary carrier, because the existing boundedness interface for pasting is an ordinary-strategy interface.*

The final induction-section pasting invocation is proved by first deriving the answer-valued Section 10 point-commutativity estimate from the answer-valued diagonal-line verifier relation, then invoking the Section 11 scalar chain in the form which assumes this point-commutativity estimate. The proof does not replace the answer-valued strategy by the ordinary carrier for the diagonal-line test; the carrier is used only for the ordinary boundedness interface. This construction is therefore a Lean-only internal construction for the simultaneous successor proof, not a new hypothesis of Theorem 10.14 and not a replacement for Theorem 10.8.

Proof. The checked reduction `answerMainInductionSuccessorNext_ofRecursiveHypothesisAndAnswerPasting` shows that, with the predecessor answer-valued induction hypothesis in the successor context, the successor branch reduces to the answer-valued pasting theorem `answerLdPastingInInductionSectionOfSmallError`. That theorem is now proved from the answer-valued commutativity route, the degree-zero branch, and the scalar absorption estimate. □

Checked status of the corrected induction interface. In Lean, the base case is formalized through . The answer-valued analogue needed by the simultaneous recursive route uses the same one-dimensional axis-parallel-line construction and does not involve the diagonal answer interface. The printed source range $md \le k < 400md$ is excluded from the corrected theorem statement: the missing implication to the pasting size hypothesis is a confirmed statement gap, and Theorem 10.14 is stated with the corrected large-k hypothesis. The successor step of is isolated in its native $m \to m+1$ form as the corrected large-k theorem . Its nontrivial branch is named as the small-error construction . The non-base branch only decomposes an arbitrary parameter bundle with $m \neq 1$ into this successor form. Neither declaration takes restricted-probability, recursive-slice, self-improvement, or pasting records as hypotheses. The nearby formalization still names the stage data that mirror the four paper stages restrict $\to$ induct $\to$ self-improve $\to$ paste, but no theorem with these data as additional source-level hypotheses is advertised as Theorem 10.14. For the answer-valued restriction route used in the successor branch of the main formal theorem, Lean now also exposes , , which record the predecessor induction hypothesis for the restricted strategies, transport induction data in both directions between the answer-valued and ordinary slice interfaces, and then forget only the answer-valued diagonal alphabet before the existing pasting assembly. The weighted restricted-probability estimates are now public via , The Lean proof now closes the large-error branch by . The answer-valued large-error branch is checked by . The internal answer-valued induction theorem is ; it is now a genuine strong induction on the dimension. The surrounding development also contains the internal reductions , , which derive the successor conclusion from the answer-valued predecessor induction hypothesis. The answer-carrier forms construct the answer-valued self-improvement data directly from the restricted slices. The successor-bound form also derives the predecessor large-k and $k \ge 1$ side conditions from the successor large-k hypothesis together with the small-error assumption, so these elementary arithmetic conditions are no longer additional hypotheses in the successor proof. The answer-carrier route shows that the separate degree-zero family route is not needed for the successor reduction: the predecessor induction hypothesis is used for the restricted slices also when $d = 0$, while $k \geq 1$ comes from the nontrivial error branch. The obsolete slice-transport and degree-split family routes have therefore been removed from the checked interface. Thus the ordinary small-error construction theorem is now proved from the internal answer-valued induction theorem, and the answer-valued pasting theorem is also proved by the answer-valued commutativity route, the degree-zero branch, and the scalar absorption estimate. None of these constructions is a separate source hypothesis. The former answer-valued self-improvement interface mismatch is resolved for the induction step by the checked carrier construction. The older ordinary-realization constructors remain available and show that, under the stronger ordinary-realization route, the Section 9 theorem can also be applied slice by slice through the checked transport interface. Such a route still requires a low-degree support theorem for the answer-valued diagonal outcomes; otherwise the ordinary polynomial diagonal measurement cannot satisfy the required covariance after rebasing. The answer-valued ambient point-consistency averaging step is now checked by and . These lemmas prove the last-coordinate reindexing identity and the averaged consistency estimate for an ambient answer-valued strategy. The answer-valued successor scalar averaging estimates are checked by , and ; these are the ζ_x, ν_x, and σ_x Jensen and conditioning bounds for the actual `AnswerSymStrat` restricted successor profiles. The theorem now extracts the recursive slice measurements from the predecessor answer-valued induction hypothesis and proves the averaged σ_x-bound. The theorem then applies the axis-parallel/self-consistency form of self-improvement slice by slice, producing the projective $\widehat{G^x}$'s and witnesses Z^x together with the averaged ζ_x and σ_x bounds. These slice outputs are assembled at the ambient answer-valued level by , which proves averaged completeness, point consistency with the actual answer-valued point measurement, strong self-consistency, the carrier-typed slice boundedness input, and the κ- and ζ-bounds. This is not a final pasting theorem: the boundedness field uses the point-equivalent ordinary carrier only because the current boundedness interface is typed for ordinary strategies, and no claim is made that the carrier's dummy diagonal measurement satisfies the answer-valued diagonal test. The answer-valued pasting invocation is proved as . The checked reduction shows that, with the predecessor answer-valued induction hypothesis available in the successor context, the successor branch reduces to this answer-valued pasting theorem. The conditional reduction is an internal lemma used inside the strong-induction proof, not a proof-level formalization link for a source theorem. The remaining ordinary family-level averaging components are also now isolated as checked lemmas: and . For answer-valued self-improvement data attached to an ordinary ambient strategy, the corresponding pasting fields are checked directly by and . The direct ordinary-ambient answer-valued assembly now invokes the induction-section pasting theorem from these fields and the answer-valued averaged scalar estimates; the exported theorem calls this direct proof. The remaining ambient answer-valued assembly is therefore the `AnswerSymStrat` analogue needed inside the simultaneous successor proof, together with the final pasting interface, not the ordinary answer-valued slice averaging calculation. The answer-valued pasting invocation and predecessor induction argument are now supplied internally in the proof of Theorem 10.14; they are not additional hypotheses about the successor strategy. We may assume $\varepsilon, \delta, \gamma, d/q \leq 1$, since otherwise $\nu \geq 1$ and the bound is trivial.

The proof is by induction on m. The base case $m = 1$ is Lemma 10.12.

For the step from m to $m+1$, let (ψ, A, B, L) be an $(\varepsilon, \delta, \gamma)$-good symmetric strategy for the $(m+1, q, d)$ test, and let $k \geq 400(m+1)d$. For each $x \in \mathbb{F}_q$, let (ψ, A^x, B^x, L^x) be the x-restricted strategy, with failure parameters $\varepsilon_x, \delta_x, \gamma_x$. Apply the induction hypothesis to each restricted strategy with the same integer k and define

$$\nu_x = 1000k^2m^2\left(\varepsilon_x^{1/1024} + \delta_x^{1/1024} + \gamma_x^{1/1024} + (d/q)^{1/1024}\right), \qquad \sigma_x = m^2\left(\nu_x + e^{-k/(80000m^2)}\right).$$

This yields a measurement $G^x \in \mathrm{PolyMeas}(m, q, d)$ such that

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

Next define

$$\zeta_x = 3000m\left(\varepsilon_x^{1/32} + \delta_x^{1/32} + (d/q)^{1/32}\right).$$

Theorem 10.6 produces a projective submeasurement $\widehat{G^x} \in \mathrm{PolySub}(m, q, d)$ such that, for each x:

1. 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. on average over $u \sim \mathbb{F}_q^m$,
$$(A^x)^u_a \otimes I \simeq_{\zeta_x} I \otimes \widehat{G^x}_{[g(u)=a]};$$
3.
$$\widehat{G^x_g} \otimes I \approx_{\zeta_x} I \otimes \widehat{G^x_g};$$
4. there exists a positive semidefinite operator Z^x such that
$$\langle\psi|Z^x \otimes (I - \widehat{G^x})|\psi\rangle \leq \zeta_x$$
and for each $g \in \mathcal{P}(m, q, d)$,
$$Z^x \geq \mathbb{E}_u (A^x)^u_{g(u)}.$$

We now average the slice-dependent parameters. By concavity of $\alpha \mapsto \alpha^c$ for $c \leq 1$ and Lemma 10.3,

$$\begin{aligned}
\mathbb{E}_x\nu_x &\leq 1000k^2m^2\left((\mathbb{E}_x\varepsilon_x)^{1/1024} + (\mathbb{E}_x\delta_x)^{1/1024} + (\mathbb{E}_x\gamma_x)^{1/1024} + (d/q)^{1/1024}\right)\\
&\leq 1000k^2m^2\left(\left(\frac{m+1}{m}\varepsilon\right)^{1/1024} + \delta^{1/1024} + \left(\frac{m+1}{m}\gamma\right)^{1/1024} + (d/q)^{1/1024}\right)\\
&\leq 1000k^2(m+1)^2\left(\varepsilon^{1/1024} + \delta^{1/1024} + \gamma^{1/1024} + (d/q)^{1/1024}\right).
\end{aligned}$$

Call this final quantity ν. Then, defining

$$\sigma = m^2\left(\nu + e^{-k/(80000m^2)}\right),$$

we have

$$\sigma \geq \mathbb{E}_x\sigma_x.$$

Similarly,

$$\begin{aligned}
\mathbb{E}_x\zeta_x &\leq 3000m\left((\mathbb{E}_x\varepsilon_x)^{1/32} + (\mathbb{E}_x\delta_x)^{1/32} + (d/q)^{1/32}\right)\\
&\leq 3000m\left(\left(\frac{m+1}{m}\varepsilon\right)^{1/32} + \delta^{1/32} + (d/q)^{1/32}\right)\\
&\leq 3000(m+1)\left(\varepsilon^{1/32} + \delta^{1/32} + (d/q)^{1/32}\right).
\end{aligned}$$

Call this last quantity ζ. For later use,

$$\begin{aligned}\zeta &= 3000(m+1)\left(\varepsilon^{1/32}+\delta^{1/32}+(d/q)^{1/32}\right)\\ &\le 1000k^2(m+1)^2\left(\varepsilon^{1/32}+\delta^{1/32}+(d/q)^{1/32}\right)\\ &\le 1000k^2(m+1)^2\left(\varepsilon^{1/1024}+\delta^{1/1024}+\gamma^{1/1024}+(d/q)^{1/1024}\right)\\ &= \nu. \end{aligned}\tag{10.1}$$

The averaged slice family $\{\widehat{G^x}\}_{x\in\mathbb{F}_q}$ therefore satisfies the four hypotheses of Theorem 10.8 with completeness parameter $\kappa = \sigma + \zeta$. To verify the new ν bound in that theorem, observe that $(3000)^{1/32} \le 2$, so

$$\zeta^{1/32} \le 2(m+1)\left(\varepsilon^{1/1024}+\delta^{1/1024}+(d/q)^{1/1024}\right),$$

and hence

$$\begin{aligned}&100k^2m\left(\varepsilon^{1/32}+\delta^{1/32}+\gamma^{1/32}+\zeta^{1/32}+(d/q)^{1/32}\right)\\ &\le 1000k^2(m+1)^2\left(\varepsilon^{1/1024}+\delta^{1/1024}+\gamma^{1/1024}+(d/q)^{1/1024}\right)\\ &= \nu.\end{aligned}$$

Therefore Theorem 10.8 gives a pasted measurement $H \in \mathrm{PolyMeas}(m+1,q,d)$ such that, on average over $u \sim \mathbb{F}_q^{m+1}$,

$$A^u_a \otimes I \simeq_{\sigma^*} I \otimes H_{[h(u)=a]},$$

where

$$\sigma^* = (\sigma+\zeta)\left(1+\frac{1}{100m}\right) + 2\nu + e^{-k/(80000m^2)}.$$

It remains to compare σ^* with the next-stage target. The pasting parameter in Theorem 10.8 is bounded by $\nu/5$ in this regime. Using this sharper estimate and (10.1), we get

$$\begin{aligned}\sigma^* &\le \left(m^2\left(\nu+e^{-k/(80000m^2)}\right)+\nu\right)\left(1+\frac{1}{100m}\right)+\frac{2}{5}\nu+e^{-k/(80000m^2)}\\ &= \left((m^2+1)\left(1+\frac{1}{100m}\right)+\frac{2}{5}\right)\nu + \left(m^2\left(1+\frac{1}{100m}\right)+1\right)e^{-k/(80000m^2)}.\end{aligned}$$

For every $m \ge 1$,

$$(m^2+1)\left(1+\frac{1}{100m}\right)+\frac{2}{5} \le (m+1)^2, \qquad m^2\left(1+\frac{1}{100m}\right)+1 \le (m+1)^2.$$

Therefore

$$\sigma^* \le (m+1)^2\left(\nu+e^{-k/(80000m^2)}\right) \le (m+1)^2\left(\nu+e^{-k/(80000(m+1)^2)}\right),$$

which is exactly the bound required for dimension $m+1$.

Remark 10.19 (Lean status of the successor-step assembly). The Lean declaration for Theorem 10.14 has the corrected large-k theorem hypotheses and no additional stage-data hypotheses. Its native successor step `mainInductionSuccessorNext` is now proved. The weighted slice bounds feeding Lemma 10.3 remain exposed through the public restriction-package constructor above, but the recursive slice witnesses, the slice-wise self-improvement outputs, and the answer-valued pasting comparison are derived inside the successor proof. They are not represented by a separate theorem whose assumptions are added to the paper statement.

10.3 Proof of the main formal theorem

Lemma 10.20 (Heterogeneous role-register symmetrization)**.** *Let* $(\psi, A^{\mathrm{A}}, B^{\mathrm{A}}, L^{\mathrm{A}}, A^{\mathrm{B}}, B^{\mathrm{B}}, L^{\mathrm{B}})$ *be a projective strategy on two local spaces passing the* (m, q, d) *low individual degree test with probability at least* $1-\varepsilon$*. Adding a role register and placing the two local spaces in the direct sum* $H_{\mathrm{A}} \oplus H_{\mathrm{B}}$ *yields a symmetric projective strategy that is* $(3\varepsilon, 3\varepsilon, 3\varepsilon)$*-good.*

Proof. Each of the three subtests (axis-parallel, self-consistency, diagonal) occurs with probability 1/3, so the total failure probability at most ε implies each branch is at most 3ε. Block-diagonalizing the measurements over the role register preserves the axis-parallel and diagonal branch averages exactly, and the cross-register self-consistency branch of the symmetrized strategy equals the original point-agreement defect, which is itself bounded by 3ε. □

Remark 10.21 (Lean auxiliary data for role-register symmetrization)*.* The formal proof of Theorem 2.14 uses the heterogeneous role-register strategy and its goodness proof. This construction supports the final assembly and is not an additional hypothesis in Lemma 10.20.

Proposition 10.22 (Lean role-register main-induction handoff for the Section 3 proof)**.** *For a two-space projective strategy passing the low individual degree test with error* ε*, the role-register symmetrization is* $(3\varepsilon, 3\varepsilon, 3\varepsilon)$*-good. Under the corrected hypothesis* $k \geq 400md$*, Theorem 10.14 therefore supplies the Section 6 polynomial measurement used in the final theorem route.*

Proof. This is a Lean-only handoff theorem: combine Lemma 10.20 with the corrected large-k theorem . □

Removed successor restricted-recursion targets. The former Section 3 Lean-only answer-valued restricted-recursion records have been removed along with the same-space final-theorem assembly. The successor work belongs to the proof of Theorem 10.14, not to additional hypotheses or auxiliary inputs in Theorem 2.14.

Remark 10.23 (Lean Step 6 constructions for the Section 3 proof)*.* The old same-space base-case and completion-witness records have been removed. The current route starts with the heterogeneous role-register application of Theorem 10.14, extracts the two occupied role blocks, proves the Step 5 full-polynomial self-consistency relation, and then invokes the theorems for making measurements projective and for completion inside the two-space source route. These constructions are proof steps in Theorem 2.14, not hypotheses of its statement.

Lean successor-dependent Step 6 targets. The role-register Section 6 measurement is obtained by applying the corrected large-k theorem `MainInductionStep.mainInduction` to the role-register symmetrization. The final source-route theorem then carries this measurement through the two unsymmetrized point-consistency relations, the Step 5 self-consistency calculation, the step making measurements projective, completion, and the final point-evaluation transport. These targets are proof-complete for the corrected large-k interface and are not added hypotheses of Theorem 2.14.

Remark 10.24 (Projective-completion construction)*.* The projective-completion construction is now internal to the two-space source route. It is not a separate assumption in the statement of Theorem 2.14. The corrected source theorem is proof-complete: the general two-space role-register passage has been assembled, and the theorem statement records the documented $k \geq 400md$ and $k > 0$ boundary corrections.

Remark 10.25 (Lean measurement-unsymmetrization consistency)*.* Lean now isolates the Step 3 extraction from `inductive_step.tex` lines 84–108. In the same-space route, the principal-block map `extractRoleBlock` turns a role-register POVM G into the two POVMs G^{A} and G^{B} on the original prover space and proves that point-evaluation postprocessing commutes with this extraction. In the heterogeneous two-space route, the corresponding principal-block maps `extractRoleRegisterAlice` and `extractRoleRegisterBob` extract the occupied $(\mathrm{A}, \mathrm{inl})$ and $(\mathrm{B}, \mathrm{inr})$ blocks of a measurement on $\{\mathrm{A}, \mathrm{B}\} \times (\mathscr{H}_{\mathrm{A}} \oplus \mathscr{H}_{\mathrm{B}})$. Lean proves the trace-level factor-two identities for arbitrary role-register measurements in `qBipartiteConsDefect_roleRegisterProjMeas_arbitrary_eq` and its two factor-two consequences. No projectivity of arbitrary principal blocks is claimed, matching the paper's later step making measurements projective.

Definition 10.26 (Final error envelope)**.** Set

$$E = \varepsilon^{1/40000} + (d/q)^{1/40000} + e^{-k/(2560000m^2)}.$$

The remaining entries in this scalar-cascade subsection are Lean bookkeeping lemmas for the final proof of Theorem 2.14. They record explicit inequalities between the named error parameters; they are not independent source theorems from the paper.

Theorem 10.27 (Basic facts about the final error envelope)**.** *If* $0 \leq \varepsilon$, *then*

$$\nu_{\mathrm{main}} = 100000k^2m^4 \left(\varepsilon^{1/40000} + (d/q)^{1/40000} + e^{-k/(2560000m^2)}\right) = 100000k^2m^4E,$$

and $E \geq 0$.

Proof. Both identities are by definition: the first unfolds ν_{main} and the second unfolds E. Nonnegativity of E follows summand-wise from `Real.rpow_nonneg` (using $0 \leq \varepsilon$) and `Real.exp_nonneg`. □

Definition 10.28 (Cascade parameters and standing hypotheses)**.** Assume $q > 0$, $k \geq 1$, $m \geq 1$, $0 \leq \varepsilon \leq 1$, and $d \leq q$. Fix an incoming error parameter ν. Define

$$\sigma = m^2 \left(\nu + e^{-k/(80000m^2)}\right), \qquad \zeta_1 = 2\sigma + 2\sqrt{3\varepsilon + 2\sigma} + md/q,$$

$$\zeta_2 = 200\zeta_1^{1/4} + 42\zeta_1^{1/8}, \qquad \zeta_3 = 6\zeta_1 + 6\zeta_2, \qquad \zeta_4 = 2\sigma + 2\sqrt{\zeta_1 + \zeta_3/2}.$$

The coefficient 42 is the Lean cascade scalar; it widens the paper's printed coefficient 40 so that the completion theorem's extra $2\zeta_1$ term is absorbed while preserving the final error envelope.

Theorem 10.29 (Bound for the first cascade term)**.** *Under the standing hypotheses, if*

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

then

$$\sigma \leq 10000k^2m^4E.$$

Proof. Multiply the assumed bound by m^2, replace m^2 by m^4, coarsen the exponents from $1/1024$ to $1/40000$, and enlarge the exponential denominator from $80000m^2$ to $2560000m^2$. □

Theorem 10.30 (Cascade bounds for the later cascade terms)**.** *Under the standing hypotheses, assume*

$$0 \leq \nu \leq 10000k^2m^2 \left(\varepsilon^{1/1024} + (d/q)^{1/1024}\right).$$

Define

$$\sigma = m^2 \left(\nu + e^{-k/(80000m^2)}\right),$$

$$\zeta_1 = 2\sigma + 2\sqrt{3\varepsilon + 2\sigma} + md/q, \qquad \zeta_2 = 200\zeta_1^{1/4} + 42\zeta_1^{1/8},$$
$$\zeta_3 = 6\zeta_1 + 6\zeta_2, \qquad \zeta_4 = 2\sigma + 2\sqrt{\zeta_1 + \zeta_3/2}.$$

Then

$$\zeta_1, \zeta_2, \zeta_4 \leq 100000k^2m^4E, \qquad \zeta_3 \leq 2 \cdot 100000k^2m^4E.$$

Proof. The Lean declarations first derive the native bounds for $\zeta_1, \ldots, \zeta_4$ from the cascade definitions using square-root and *rpow* monotonicity, with ζ_2 using the widened coefficient 42, and then coarsen those bounds to the common envelope E. The ζ_2 proof additionally uses $k \leq k^2$, $m \leq m^4$, and the nonnegativity of the earlier cascade variables. □

Theorem 10.31 (Consolidated error cascade)**.** *Assume the standing hypotheses and*

$$0 \leq \nu \leq 10000k^2m^2 \left(\varepsilon^{1/1024} + (d/q)^{1/1024}\right).$$

Let $\sigma, \zeta_1, \zeta_2, \zeta_3$ be the cascade quantities defined above. Then

$$\sigma \leq 100000k^2m^4E, \qquad \zeta_1, \zeta_2 \leq 100000k^2m^4E, \qquad \zeta_3 \leq 2 \cdot 100000k^2m^4E,$$

and the concrete quantity

$$\zeta_4 = 2\sigma + 2\sqrt{\zeta_1 + \zeta_3/2}$$

also satisfies

$$\zeta_4 \leq 100000k^2m^4E.$$

Proof. This is the direct combination of Theorem 10.29 and Theorem 10.30. The Lean declaration uses the explicit equalities naming $\sigma, \zeta_1, \zeta_2, \zeta_3$ as the corresponding cascade expressions. □

Remark 10.32 (Lean Step 5 expansion theorem). For the Step 5 passage below, Lean now names the weighted polynomial-collision mass, proves its md/q bound by specializing the existing tensor Schwartz–Zippel helper, and proves the algebraic expansion/reindexing statement comparing the evaluated consistency defect with the full-polynomial defect plus that collision mass. The theorem for Step 5 therefore no longer takes the expansion statement as an external hypothesis.

Remark 10.33 (Lean witness structures for the main-formal assembly). The following Lean declarations are bookkeeping infrastructure for the final assembly witness: They split off the vacuous branch where the theorem error is at least 1, derive the Step 8 scalar hypotheses in the non-vacuous branch, and isolate the role-register Section 6 measurement behind a concrete role-witness record. The role-register measurement is now obtained directly from the `mainInduction` theorem by `roleRegisterSymmStrategy_sourceMainInduction`; the successor case is handled by the checked Section 6 theorem, so it is not a separate boundary hypothesis in Section 3. Lean also records that every non-base parameter choice has a predecessor via `Parameters.successorDecompositionOfNeOne`, and it retains the explicit transport helpers for the strategy and passing proof across that predecessor equality. The older branch-witness records, which assembled ordinary or answer-valued successor data from recursive slice witnesses and supplied self-improvement inputs, have been removed. The current completion witness carries the concrete role witness obtained from the Section 6 theorem rather than a separate decorative branch witness. Once the role-measurement record is obtained, Lean proves the two factor-two unsymmetrization links from its symmetrized consistency estimate. The subsequent proved lemmas use the proved Section 5 Schwartz–Zippel theorem to convert evaluated pre-projective consistency to full-polynomial consistency at ζ_1, prove the paper line-116 triangle step, and reconstruct the line-156 $\approx_{\zeta_3}$ relation obtained while making measurements projective from pre-projective consistency plus left/right completion closeness at ζ_2. Lean also derives the evaluated line-164 projective consistency. The error-parametric lemmas `pointAConsistency_ofLine169Consistency` and `pointBConsistency_ofLine169Consistency` show that a polynomial line-169 estimate with error η yields the final point error

$$2\sigma + 2\sqrt{\eta + \zeta_3/2}.$$

The active formal route specializes this with the repaired value $\eta = \zeta_1 + 10\zeta_1^{1/8}$. The existing `pointAConsistency` and `pointBConsistency` declarations are the corresponding repaired specializations. After the #869 right-register completion transport, the remaining active post-role witness asks only for Alice's line-146 completion closeness, Bob's right-register completion closeness, and the two repaired polynomial line-169 links. Lean reconstructs pre-projective full-polynomial consistency from the concrete role witness, uses those repaired line-169 links together with the evaluated line-164 projective consistency, and converts directly to the projective completion witness used by the final assembly. It no longer asks for an arbitrary role-measurement record or an arbitrary Section 6 witness as an independent field. It also no longer asks for stored factor-two role-block estimates, a stored Section 6 consistency input, or a duplicated pre-projective consistency field; those are rebuilt from the concrete role-register measurement and `hpass`. Lean then applies the projective converse of `simeq-to-approx` inside the two-space source route.

As stated in `preliminaries.tex`, `prop:triangle-sub` gives a $\delta + \sqrt{\epsilon}$ loss, so applying it directly to $G^{\mathrm{A}} \simeq_{\zeta_1} G^{\mathrm{B}}$ and $G^{\mathrm{A}} \approx_{\zeta_2} Q^{\mathrm{A}}$ yields $Q^{\mathrm{A}} \simeq_{\zeta_1+\sqrt{\zeta_2}} G^{\mathrm{B}}$, not the printed ζ_1 estimate. The generic corrected route would degrade the final point error to scale $t^{1/65536}$, outside the current `mainFormalError` envelope. The sharper local pre-completion repair from Remark 4.6 avoids that widening. Its completion step uses only the formal monotonicity helper `ProjectivizationMatchMassMonotonicity.completeAtOutcomeProj_left_matchMass_ge`, not a separate exact match-mass witness route. For the corrected boundary assumptions $k \geq 400md$ and $k > 0$, the source theorem now has a checked proof-level route.

Proof. Let

$$(\psi, A^{\mathrm{A}}, B^{\mathrm{A}}, L^{\mathrm{A}}, A^{\mathrm{B}}, B^{\mathrm{B}}, L^{\mathrm{B}})$$

be a projective strategy passing the (m, q, d) low individual degree test with probability at least $1 - \varepsilon$. Since the three subtests occur with probability $1/3$, this strategy is $(3\varepsilon, 3\varepsilon, 3\varepsilon)$-good.

Introduce a two-dimensional role register on each side and define the symmetrized state

$$|\psi_{\mathrm{sym}}\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle_{\mathrm{A}'}|1\rangle_{\mathrm{B}'}|\psi\rangle_{\mathrm{AB}} + |1\rangle_{\mathrm{A}'}|0\rangle_{\mathrm{B}'}|\psi_{\mathrm{swap}}\rangle_{\mathrm{AB}}\right).$$

Define the symmetrized point measurement by

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

and similarly for B_{sym} and L_{sym}. This yields a symmetric $(3\varepsilon, 3\varepsilon, 3\varepsilon)$-good strategy. Set

$$\nu = 10000k^2m^2\left(\varepsilon^{1/1024} + (d/q)^{1/1024}\right), \qquad \sigma = m^2\left(\nu + e^{-k/(80000m^2)}\right).$$

Theorem 10.14 applied to the symmetrized strategy produces a measurement $G = \{G_g\} \in \mathrm{PolyMeas}(m, q, d)$ such that

$$(A_{\mathrm{sym}})^u_a \otimes I \simeq_\sigma I \otimes G_{[g(u)=a]}, \qquad G_{[g(u)=a]} \otimes I \simeq_\sigma I \otimes (A_{\mathrm{sym}})^u_a. \tag{10.2}$$

Now unsymmetrize by defining

$$G_g^{\mathrm{A}} = (\langle 0| \otimes I)\, G_g\, (|0\rangle \otimes I), \qquad G_g^{\mathrm{B}} = (\langle 1| \otimes I)\, G_g\, (|1\rangle \otimes I).$$

These are POVMs on the original prover space. Since A_{sym} is a convex combination of $(|0\rangle\langle 0| \otimes I)$ and $(|1\rangle\langle 1| \otimes I)$ weighted by the test probabilities, the cross-terms $\langle 0| \cdot A_{\mathrm{sym}} \cdot |1\rangle$ vanish. Expanding the consistency of G gives

$$G^{\mathrm{A}}_{[g(u)=a]} \otimes I \simeq_{2\sigma} I \otimes A_a^{\mathrm{B},u}, \tag{10.3}$$

$$I \otimes G^{\mathrm{B}}_{[g(u)=a]} \simeq_{2\sigma} A_a^{\mathrm{A},u} \otimes I. \tag{10.4}$$

Since the original strategy is $(3\varepsilon, 3\varepsilon, 3\varepsilon)$-good,

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

Therefore Lemma 3.29 yields

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

Expanding this consistency relation and applying Lemma 3.7 to replace point evaluations by polynomial evaluations (at cost md/q) gives

$$G_g^{\mathrm{A}} \otimes I \simeq_{\zeta_1} I \otimes G_g^{\mathrm{B}}, \tag{10.5}$$

where

$$\zeta_1 = 2\sigma + 2\sqrt{3\varepsilon + 2\sigma} + \frac{md}{q}.$$

The Lean Step 5 theorem formalizes this expansion/reindexing step and combines it with the proved Schwartz–Zippel tensor bound . The Lean statement is now formulated for a bipartite state on $H_A \otimes H_B$, so this particular Schwartz–Zippel loss is no longer a same-space restriction in the source-theorem route.

Apply Lemma 4.9 to 10.5. This gives projective submeasurements $P^{\mathrm{A}}, P^{\mathrm{B}} \in \mathrm{PolySub}(m, q, d)$ with

$$G_g^{\mathrm{A}} \otimes I \approx_{100\zeta_1^{1/4}} P_g^{\mathrm{A}} \otimes I, \qquad I \otimes G_g^{\mathrm{B}} \approx_{100\zeta_1^{1/4}} I \otimes P_g^{\mathrm{B}}.$$

Completing them by Lemma 3.42 yields projective measurements $Q^{\mathrm{A}}, Q^{\mathrm{B}} \in \mathrm{PolyMeas}(m, q, d)$ such that

$$\begin{aligned} G_g^{\mathrm{A}} \otimes I &\approx_{\zeta_2} Q_g^{\mathrm{A}} \otimes I, \\ I \otimes G_g^{\mathrm{B}} &\approx_{\zeta_2} I \otimes Q_g^{\mathrm{B}}, \end{aligned} \tag{10.6}$$

where

$$\zeta_2 = 200\zeta_1^{1/4} + 42\zeta_1^{1/8}.$$

The paper prints the coefficient 40 in this display. The formal Lean cascade uses the widened coefficient 42 to absorb the literal completion error $2 \cdot (100\zeta_1^{1/4}) + 4\sqrt{100\zeta_1^{1/4}} + 2\zeta_1$: under the non-vacuous scalar regime $0 \le \zeta_1 \le 1$, the extra term $2\zeta_1$ is bounded by $2\zeta_1^{1/8}$.

By 10.5 and Lemma 3.21,

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

Hence Lemma 3.27 gives

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

where

$$\zeta_3 = 6\zeta_1 + 6\zeta_2.$$

Because both measurements are projective, Lemma 3.21 converts this to

$$Q_g^{\mathrm{A}} \otimes I \simeq_{\zeta_3/2} I \otimes Q_g^{\mathrm{B}}. \tag{10.7}$$

By data processing,

$$Q_{[g(u)=a]}^{\mathrm{A}} \otimes I \simeq_{\zeta_3/2} I \otimes Q_{[g(u)=a]}^{\mathrm{B}}. \tag{10.8}$$

Next, Lemma 3.20, applied to 10.5 and 10.6, implies

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

Data processing then gives

$$Q_{[g(u)=a]}^{\mathrm{A}} \otimes I \simeq_{\zeta_1} I \otimes G_{[g(u)=a]}^{\mathrm{B}}. \tag{10.9}$$

Applying Lemma 3.29 to 10.4, 10.9, and 10.8 yields

$$A_a^{\mathrm{A},u} \otimes I \simeq_{\zeta_4} I \otimes Q_{[g(u)=a]}^{\mathrm{B}}, \tag{10.10}$$

where

$$\zeta_4 = 2\sigma + 2\sqrt{\zeta_1 + \zeta_3/2}.$$

The same argument with the roles reversed gives

$$I \otimes A_a^{\mathrm{B},u} \simeq_{\zeta_4} Q_{[g(u)=a]}^{\mathrm{A}} \otimes I. \tag{10.11}$$

Remark 10.33 records the Lean witness structures that separate the scalar cascade and the Step 5 Schwartz–Zippel handoff from the remaining projective-stage transport work. The scalar side discharges the paper's line 71–73 coarsening and the Step 8 hypotheses in the non-vacuous branch; the Step 5 theorem now turns the evaluated estimate at line 116 into full-polynomial consistency at ζ_1. The remaining Lean witness is the completion-and-line-169 construction: supply the actual role-register measurement, Alice's line-146 completion estimate, Bob's left-register completion estimate, and the two exact polynomial-level ζ_1 links. Lean transports the Bob-side completion estimate to the line-147 right-register form, reconstructs the duplicated Step 6 pre-projective consistency field from these inputs and `hpass`, then data-processes the line-169 links to line 172 and derives the two ζ_4 point goals and the $\zeta_3/2$ self-consistency target.

With E as in Definition 10.26, the quantities $\sigma, \zeta_1, \zeta_2, \zeta_3, \zeta_4$ form the cascade of Definition 10.28. Theorem 10.29 and Theorem 10.30 give the final inflation to the scale $100000k^2m^4E$. The underlying estimates, with the formal widened ζ_2, are:

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

Then

$$\zeta_1 \leq 20204k^2m^4 \left(\varepsilon^{1/2048} + (d/q)^{1/2048} + e^{-k/(160000m^2)}\right),$$

and

$$\zeta_2 \leq 2568km \left(\varepsilon^{1/16384} + (d/q)^{1/16384} + e^{-k/(1280000m^2)}\right).$$

Consequently,

$$\zeta_3 \leq 150000k^2m^4 \left(\varepsilon^{1/16384} + (d/q)^{1/16384} + e^{-k/(1280000m^2)}\right),$$

and

$$\zeta_4 \leq 40000k^2m^4\left(\varepsilon^{1/32768} + (d/q)^{1/32768} + e^{-k/(2560000m^2)}\right).$$

Both $\zeta_3/2$ and ζ_4 are therefore bounded by

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

Hence 10.10, 10.11, and 10.7 give the three claimed conclusions, with Q^{A} and Q^{B} renamed as the theorem's G^{A} and G^{B}. $\square$

Bibliography

[con26a] The MIPStarRE contributors. The constant in the completion-route orthonormalization theorem. Paper-gap note issue-1032-orthonormalization-constant, MIPStarRE, 2026. https://sirui-lu.com/MIPStarRE/paper-gaps/issue-1032-orthonormalization-constant.pdf.

[con26b] The MIPStarRE contributors. The large-k side condition in the main induction. Paper-gap note issue-906-main-formal-k-bound, MIPStarRE, 2026. https://sirui-lu.com/MIPStarRE/paper-gaps/issue-906-main-formal-k-bound.pdf.

[con26c] The MIPStarRE contributors. The zero-sampling boundary in the main formal theorem. Paper-gap note issue-422-main-formal-zero-k-boundary, MIPStarRE, 2026. https://sirui-lu.com/MIPStarRE/paper-gaps/issue-422-main-formal-zero-k-boundary.pdf.

[JNV+20] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. Quantum soundness of the classical low individual degree test, 2020. arXiv:2009.12982.

[Sch80] Jacob T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM*, 27(4):701–717, 1980.

[Zip79] Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Proceedings of the 2nd International Symposium on Symbolic and Algebraic Manipulation*, pages 216–226, 1979.