Electronic Colloquium on Computational Complexity, Report No. 179 (2026)

On the Hardness of 4-to-1 Games with Perfect Completeness

Yumou Fei[*] Dor Minzer[†] Shuo Wang[‡]

Abstract

We prove that for all $\varepsilon > 0$, there exists a positive integer k such that given a 4-to-1 Game Ψ with alphabet size at most k, it is NP-hard to distinguish between the case that $\mathrm{val}(\Psi) = 1$ and the case that $\mathrm{val}(\Psi) \leqslant \delta$. This confirms the 4-to-1 Games Conjecture from [Khot, *CCC 2002*]. Previously, the best known result, due to [Dinur, Khot, Kindler, Minzer, Safra], established the almost-perfect completeness version (but applied to the stricter problem of 2-to-1 Games). Using results from the literature, we get the following implications:

(1) For all $k \in \mathbb{N}$, given a 3-colorable graph G, it is **NP**-hard to find a proper k-coloring.

(2) For all $\delta > 0$, given a 2-colorable 3-uniform hypergraph G, it is **NP**-hard to find in it an independent set containing at least δ fraction of the vertices.

Our proof is a three-step construction that builds on the two-step framework of [Dinur, Khot, Kindler, Minzer, Safra]. In the outer-PCP step, we use quadratic equations to gain perfect completeness. We then construct a new middle PCP that performs low-rank tests while preserving a key covering property. Finally, we construct a new inner PCP based on a tensor of the standard Grassmann encoding with its low-rank variant due to [Golowich, FOCS 2023].

The current version of the manuscript is complete mathematically, but it is **not** in the shape we wished to share in. We have chosen to do so due to rumors that surfaced around Friday, September 11th. We will work on a more complete version of the manuscript in the close future.

[*]Department of EECS, Massachusetts Institute of Technology. Supported by NSF CCF award 2310818 and NSF DMS award 2022448.

[†]Department of Mathematics, Massachusetts Institute of Technology. Supported by NSF CCF award 2227876 and NSF CAREER award 2239160.

[‡]Department of Mathematics, Massachusetts Institute of Technology. Supported by NSF award 2239160.

ISSN 1433-8092

Contents

1 Introduction

The PCP Theorem [FGL+96, AS98, ALM+98] is a fundamental result in theoretical computer science with many consequences. Most relevant to the current paper are the applications in the area of hardness of approximation, in which the PCP theorem often serves as the starting point for reductions. The combination of Fourier analysis with the long-code (or, in some cases, other error correcting codes) has proven itself to be very successful recipe, leading to many strong hardness of approximation results [Fei98, Hås01, DS05, Zuc07]. However, this paradigm appears to be incapable of establishing optimal results for many other problems, and below are a few notable examples:

(1) The Max-Cut problem: given a graph $G = (V, E)$, the goal is to partition the vertex set into two sides A, B so as to maximize the number of edges crossing between A and B. A polynomial time algorithm due to [GW95] achieves the approximation ratio $\alpha_{GW} \approx 0.878$, and the best known **NP**-hardness result [TSSW00] has ratio $\frac{16}{17} + \varepsilon$, for any $\varepsilon > 0$.

(2) The Vertex-Cover problem: given a graph $G = (V, E)$, the goal is to find the smallest set of vertices $A \subseteq V$ that touches all edges of E. Known polynomial time algorithms achieve approximation ratio 2, and the best known **NP**-hardness result [KMS25, DKK+25, DKK+21, KMS23] has ratio $\sqrt{2} - \varepsilon$, for any $\varepsilon > 0$.

(3) The 3-coloring problem: given a graph $G = (V, E)$ promised to be 3-colorable, the goal is to find a proper k-coloring of the vertices of G with k as small as possible. Polynomial time algorithms for this problem have been studied extensively over the years [Wig83, KMS98, BK97, ACC06, Chl07, KT17, KTY24, BHL26], and the state-of-the-art algorithm achieves $k = |V|^{0.2-\varepsilon}$ for a small explicit $\varepsilon > 0$. The **NP**-hardness of this problem has also been studied [KLS00, GK04, BBKO21], and the state-of-the-art **NP**-hardness result is for $k = 5$.

To make progress on these problems, Khot [Kho02] introduced the Unique-Games Conjecture and the d-to-1 Games conjectures, which we present next.

1.1 Label-Cover, Unique Games and d-to-1 Games

It is convenient to discuss PCPs in the language of the label-cover problem, defined as follows.

Definition 1.1. An instance Ψ of the label-cover problem consists of a bipartite graph $G = (L \sqcup R, E)$, alphabets Σ_L, Σ_R, and a collection of constraints $\{\Phi_e \subseteq \Sigma_L \times \Sigma_R\}_{e \in E}$. A constraint Φ_e is called a projection constraint if there is a map $\phi_e \colon \Sigma_L \to \Sigma_R$ such that

$$\Phi_e = \{(\sigma, \phi_e(\sigma)) \mid \sigma \in \Sigma_L\}.$$

A constraint Φ_e is called a d-to-1 constraint if it is a projection constraint, and the map ϕ_e is d-to-1, namely, $|\phi_e^{-1}(\sigma)| = d$ for all $\sigma \in \Sigma_R$.

Given a label-cover instance Ψ and assignments $A_L \colon L \to \Sigma_L$, $A_R \colon R \to \Sigma_R$, we define

$$\mathrm{val}_\Psi(A_L, A_R) = \frac{|\{e = (u, v) \in E \mid (A_L(u), A_R(v)) \in \Phi_e\}|}{|E|},$$

that is, the fraction of constraints that are satisfied by A_L, A_R. We define the value of Ψ as

$$\mathrm{val}(\Psi) = \max_{A_L, A_R} \mathrm{val}_\Psi(A_L, A_R).$$

For parameters $0 < s < c \leqslant 1$, an input to the promise problem $\mathsf{GapPLC}(c, s)$ is a label cover instance Ψ with projection constraints that either has $\mathrm{val}(\Psi) \geqslant c$ or $\mathrm{val}(\Psi) \leqslant s$, and the goal is to distinguish between the two cases. The basic PCP theorem [FGL+96, AS98, ALM+98] combined with the parallel repetition theorem [Raz98] gives:

Theorem 1.2. *For all $\varepsilon > 0$, there exists a positive integer r such that $\mathsf{GapPLC}_r(1, \varepsilon)$ is **NP**-hard. Here GapPLC_r is the GapPLC problem in which both alphabets have size at most r.*

Khot conjectured that versions of Theorem 1.2 continue to hold even when the constraints are d-to-1 for a small d. Towards presenting these conjectures, we make the following definition.

Definition 1.3. Let d be a positive integer. An instance Ψ of label-cover is called a d-to-1 game if all of its constraints are d-to-1. In the case that $d = 1$, Ψ is called a unique-game. We write $\mathsf{GapUG}_r(c, s)$ for the restriction of $\mathsf{GapPLC}_r(c, s)$ to unique-game instances, and $\mathsf{Gap}\text{-}d\text{-}\mathsf{to}\text{-}1_r(c, s)$ for its restriction to d-to-1 game instances.

It is easy to see that one cannot hope for Theorem 1.2 to hold for unique-games, and Khot [Kho02] conjectured that allowing for imperfect completeness remedies that:

Conjecture 1.4. *For any $\varepsilon > 0$, there exists $r \in \mathbb{N}^+$, such that $\mathsf{GapUG}_r(1 - \varepsilon, \varepsilon)$ is **NP**-hard.*

For $d \geqslant 2$, he conjectured that the assertion of Theorem 1.2 should hold as is. More precisely:

Conjecture 1.5. *For any $\varepsilon > 0$ and any integer $d \geqslant 2$, there exists a positive integer r such that $\mathsf{Gap}\text{-}d\text{-}\mathsf{to}\text{-}1_r(1, \varepsilon)$ is **NP**-hard.*

These conjectures, and especially Conjecture 1.4, have been shown to have remarkable implications, see for example [KR08, KKMO07, Rag08, KV15] and [Kho10b, Kho10a, Tre12, Kho14] for surveys, and significant effort has gone into studying them. On the algorithmic front, connections with semi-definite programming, sum-of-squares algorithms and graph expansion have led to the development of very powerful techniques and a number of algorithms for unique-games with non-trivial guarantees [CMM06, AKK+08, ABS15]. On the hardness front, a line of work [KMS25, DKK+25, DKK+21, KMS23] (with contributions from [BKS19, KMMS25]) established the imperfect completeness version of Conjecture 1.5, showing that for all $\varepsilon > 0$, there is $r \in \mathbb{N}^+$ such that $\mathsf{Gap}\text{-}2\text{-}\mathsf{to}\text{-}1_r(1 - \varepsilon, \varepsilon)$ is **NP**-hard. This implies that for all $\varepsilon > 0$ there exists $r \in \mathbb{N}^+$ such that $\mathsf{GapUG}_r(1/2, \varepsilon)$ is **NP**-hard, which to date is the best evidence towards Conjecture 1.4.

The difference between Conjecture 1.5 and its imperfect completeness version is significant in a number of ways:

(1) **Implications:** there are a number of hardness of approximations results that crucially rely on perfect completeness, most prominently graph coloring problems.

(2) **Techniques:** the approach taken by [KMS25, DKK+25, DKK+21, KMS23] inherently incurs imperfect completeness. Currently, constructions of low-soundness PCPs with perfect completeness follow composition of PCPs, for example [MR10, DH13, DHK15], which at the very least require constructions that already have perfect completeness and small soundness (with perhaps a large alphabet which is then reduced), or from parallel repetition [Raz98, Hol09, Rao11], which inherently increases the d-parameter rapidly.

1.2 Main Result

Our main result is that Conjecture 1.5 holds for $d = 4$.

Theorem 1.6. *For all $\varepsilon > 0$, there exists $r \in \mathbb{N}^+$, such that $\mathsf{Gap}\text{-}4\text{-}\mathsf{to}\text{-}1_r(1, \varepsilon)$ is **NP**-hard.*

1.2.1 Implications

We now discuss some implications of Theorem 1.6. Using known connection between d-to-1-Games and graph coloring/independent sets [DMR09, GS20], we get the following results.

Corollary 1.7. *For all $k \in \mathbb{N}$, given a graph G, it is* **NP***-hard to distinguish between the following two cases:*

(1) ***Yes case:*** *G is 3-colorable.*

(2) ***No case:*** *G is not k-colorable.*

Previously, Corollary 1.7 was known [BBKO21] for $k = 5$, or in the case the condition in the yes-case only required a proper 3-coloring of $1 - \varepsilon$ fraction of the vertices [HMS23].

Remark 1.8. Under slightly super-polynomial time reductions, our techniques give versions of Corollary 1.7 for k growing with the instant size, but at a very slow rate. Namely, assuming **NP** has no algorithms running in time $n^{t(n)}$, we get that there are no polynomial time algorithms that colors a given 3-colorable graph with $O(\log \cdots \log t(n))$ colors, where the number of logs is finite (probably at most 10, say).

In the next corollary, the condition in the completeness case is weaker, but the condition in the soundness case is stronger:

Corollary 1.9. *For all $\delta > 0$, given a graph G, it is* **NP***-hard to distinguish between the following two cases:*

(1) ***Yes case:*** *G is 8-colorable.*

(2) ***No case:*** *G has no independent set of relative size at least δ.*

Using [KS14], we get the following result for hypergraphs:

Corollary 1.10. *For all $\delta > 0$, given a 3-uniform hypergraph G, it is* **NP***-hard to distinguish between the following two cases:*

(1) ***Yes case:*** *G is 2-colorable.*

(2) ***No case:*** *G has no independent set of relative size at least δ.*

Using the results of [HMRW14], Corollaries 1.7 and 1.9 show hardness result for low-rank matrix completion. An instance in the (k, r, ε, c)-completion problem is a matrix $A \in (\mathbb{R} \cup \{\perp\})^{n \times n}$, such that there exists $M \in \mathbb{R}^{n \times n}$ satisfying:

(1) $|M(i,j)| \leqslant c$ and M has rank at most r.

(2) Letting $\Omega \subseteq [n] \times [n]$ be the set of (i,j) such that $A(i,j) \neq \perp$, we have that $|\Omega| \geqslant 0.9n^2$ and for each $(i,j) \in \Omega$ we have $M(i,j) = A(i,j)$.

The goal is to find a matrix $B \in \mathbb{R}^{n \times n}$ of rank at most k, such that $|B(i,j)| \leqslant c$ for all (i,j) and

$$\sum_{(i,j) \in \Omega} |A(i,j) - B(i,j)|^2 \leqslant \varepsilon n.$$

Applying [HMRW14, Theorem 1.8] on Corollary 1.7 gives:

Corollary 1.11. *For all $k \geqslant 3$ and $c > 0$, the $(k, 3, 0, c)$-completion problem is* **NP***-hard.*

Applying [HMRW14, Theorem 1.9] on Corollary 1.9 gives:

Corollary 1.12. *For all $k \geqslant 3$ and $c > 0$, the $(k, 8, \varepsilon, c)$-completion problem is* **NP***-hard for all $0 < \varepsilon < \frac{1}{(2cr)^2}$.*

Remark 1.13. In the accompanying supplementary material [FMW] we give a few more applications of Theorem 1.6 that were found and proved by ChatGPT 5.6.

1.3 Our Techniques

In this section we give a brief account of the components that go into the proof of Theorem 1.6. We defer the reader to Section 2 for a more thorough description.

For 2-to-1 games with almost-perfect completeness, the proof of [KMS25, DKK+25, DKK+21, KMS23] combines the following two components:

(1) An outer PCP with completeness close to 1, vanishing soundness, and a feature called the "covering property" (originally appearing in [KS13]). This game is very far from being a 2-to-1 game, and to turn it into a 2-to-1 game it is composed with an appropriate inner PCP. The goal of the covering property is to facilitate the composition step.

(2) An inner PCP with perfect completeness and a non-standard notion of soundness. The inner PCP is based on the Grassmann graph, the Grassmann encoding and the Grassmann agreement test. The Grassmann agreement test can be viewed as a local tester for the Grassmann code, but it turns out not to be sound in the traditional error correcting codes sense. Instead, it satisfies a weaker notion of soundness, which we refer to as non-standard soundness.

The outer PCP and inner PCP are then composed, and the covering property of the outer PCP is used in the soundness analysis of the composed PCP (even though the inner PCP only satisfies a non-standard notion of soundness).

The outer PCP construction of [KMS25, DKK+25, DKK+21, KMS23] is based on the hardness of approximation result of [Hås01] for solving linear equations, which inherently has imperfect completeness, and this propagates throughout their construction. To remedy that, a key difficulty is to come up with an outer PCP that has perfect completeness and simultaneously admits (some version of) the covering property.

We start our outer PCP construction based on hardness of solving a system of quadratic equations (which has perfect completeness). However, the Grassmann agreement test (which is still the only tool we have to achieve d-to-1-ness in the inner PCP) is inherently a "linear" test and cannot incorporate quadratic equations. In order for quadratic equations to be accommodated in a d-to-1 way, we add a layer of construction called the "middle PCP" between the outer and inner PCPs. In the middle PCP, we carry out a low-rank test that allows us to convert the quadratic equations into linear ones.

However, this linearization of quadratic equations destroys the covering property that is essential for composing with a Grassmann-based inner PCP. To remedy that, we have to come up with another two key ideas:

(1) We do a "lazy parallel repetition" in the middle PCP. This creates room for covering properties while preserving the soundness.

(2) The covering property we get from lazy parallel repetition is still not sufficient for composing with the Grassmann agreement test of [KMS25, DKK+25, DKK+21, KMS23]. As a workaround, We base our inner PCP on a combination of standard Grassmann graphs over matrix spaces $\mathbb{F}_2^{n\times n}$ with a sparser version of them, known as the low-rank Grassmann graph over $\mathbb{F}_2^{n\times n}$, recently defined by [Gol23].

Statement of AI use: the mathematical ideas and the overall approach presented in this paper is due to the authors. ChatGPT 5.6 was used for polishing.

2 Proof Overview

In this section we give an overview of the proof of Theorem 1.6. It will be useful to first discuss the construction of [KMS25, DKK+25, DKK+21, KMS23] which achieves imperfect completeness. We then discuss why it fails to have perfect completeness, and then in Section 2.2 we explain preliminary attempts at fixing that, building towards our construction.

2.1 The Imperfect Completeness Construction

In this section we explain the components that go into the construction of [KMS25, DKK+25, DKK+21, KMS23].

2.1.1 The Outer PCP

It will often be easier for us to describe label-cover instances in the equivalent, active language of 2-prover-1-round games, which we introduce next. In a 2-prover-1-round game, a verifier samples a challenge $e = (u, v)$, sends u to a prover named Alice, sends v to a prover named Bob. Upon receiving their challenges, they respond with labels (they are not allowed to communicate once they get their challenges), and based on their responses the verifier decides to accept or reject. The value of the game is the maximum, over all strategies of Alice and Bob, of the probability that the verifier accepts. Identifying the individual challenges the players receive as vertices of a bipartite graph, the pairs of challenges the verifier may sample as the edges of that graph gives the equivalence, and the acceptance criteria of the verifier as the constraints, one gets the equivalence. It is easy to see that the value of a game is exactly the value of the resulting label cover instance.

With this in mind, the reduction of [KMS25, DKK+25, DKK+21, KMS23] starts with the result of [Hås01] regarding the approximability of $\mathsf{3Lin}_{\mathbb{F}_2}$.

Definition 2.1. An instance of the $\mathsf{3Lin}_{\mathbb{F}_2}$ problem consists of a set of variables $\mathscr{V}$ and a set of equations $\mathscr{E}$, where each $e \in \mathscr{E}$ is a linear equation over $\mathbb{F}_2$ containing 3 distinct variables.

Håstad proved that for all $\varepsilon > 0$, $\mathsf{Gap3Lin}_{\mathbb{F}_2}(1-\varepsilon, 1/2+\varepsilon)$ is **NP**-hard, and the reduction of [KMS25, DKK+25, DKK+21, KMS23] starts off with (some regularized version of) this problem. More specifically, they use the parallel repetition theorem to prove **NP**-hardness for certain type of label-cover instances $\Psi_{q,\beta,k} = \Psi_{q,\beta,k}(\mathscr{V}, \mathscr{E})$, which we present below in the active 2-prover-1-round game view, and refer to as the β-smooth k-fold repeated game. For technical reasons, this game contains an additional feature called "advice", which we encourage the reader to ignore in first reading.

(1) **Alice's challenge:** the verifier samples equations $e_1, \ldots, e_k$ from $\mathscr{E}$ uniformly and independently, and sets E to be the set of all variables that appear in them.

(2) **Bob's challenge:** for each $i \in \{1, \ldots, k\}$ independently, with probability $1 - \beta$, the verifier takes E'_i to be the set of variables in e_i, and with probability β the verifier chooses E'_i to consist of a single uniformly chosen variable from e_i. The verifier sets $E' = E'_1 \cup \ldots \cup E'_k$.

(3) **Advice:** the verifier chooses $b^{(1)}, \ldots, b^{(q)} \in \mathbb{F}_2^{E'}$ uniformly, and then takes $a^{(1)}, \ldots, a^{(q)} \in \mathbb{F}_2^E$ where $a^{(i)}$ is obtained from $b^{(i)}$ by filling the missing coordinates with 0's. We write $B = \left(b^{(1)}, \ldots, b^{(q)}\right)$ and $A = \left(a^{(1)}, \ldots, a^{(q)}\right)$.

(4) **Acceptance criteria:** the verifier sends (E, A) to Alice, and (E', B) to Bob. They respond with vectors $x \in \mathbb{F}_2^E$ and $y \in \mathbb{F}_2^{E'}$ respectively, and the verifier accepts if and only if x satisfies all of the equations $e_1, \ldots, e_k$ and $x|_{E'} = y$.

The works [KMS25, DKK+25, DKK+21, KMS23] prove that for all $\varepsilon > 0$, it is **NP**-hard to distinguish between the case that $\mathrm{val}(\Psi_{q,\beta,k}) \geqslant 1 - \varepsilon$, and the case that $\mathrm{val}(\Psi_{q,\beta,k}) \leqslant 2^{-\Omega_q(\beta k)}$. Here and throughout, one should think of q as an absolute constant, and of β as slightly larger than $1/k$, say $\beta = k^{-3/4}$. The game $\Psi_{q,\beta,k}$ has several useful features in the context of 2-to-1 Games:

(1) **Simplicity of the constraints:** the constraints of $\Psi_{q,\beta,k}$ can all be seen to be linear equations in the responses of the provers. This is because the starting point, the $\mathsf{3Lin}_{\mathbb{F}_2}$, consisted of linear constraints, and on top of that we have the consistency constraints between the responses of Alice and Bob, which is also linear.

(2) **The covering property:** for a fixed E, the spaces $\mathbb{F}_2^{E'}$ living on Bob's side cover Alice's space $\mathbb{F}_2^E$ almost perfectly. For example, fixing E, taking E' randomly as described above conditioned on E, then $y \in \mathbb{F}_2^{E'}$ and lifting it to $x \in \mathbb{F}_2^E$ by filling in the missing coordinates with 0's, we have that the distribution of x is statistically close to the uniform distribution over $\mathbb{F}_2^E$ (this property first appeared in [KS13] in a somewhat different context).

The relevance of these properties will become clear in the next part, the inner PCP.

2.1.2 The Inner PCP

The 2-prover 1-round game $\Psi_{q,\beta,k}$ constructed in Section 2.1.1 is far from 2-to-1, because given a Bob-answer $y \in \mathbb{F}_2^{E'}$, there are many possible Alice-answers $x \in \mathbb{F}_2^E$ consistent with it. In order to make it a 2-to-1 game, in a single query we can ask the provers only for a smaller piece of information.

For each possible question (E, A) to Alice, we ask her to prepare her answer in "Grassmann encoding," and instead of requesting for a full answer $x \in \mathbb{F}_2^E$, we ask her to return only the information in one location of the Grassmann encoding of x.

The Grassmann encoding scheme is based on the Grassmann graph $\mathsf{Gr}(\mathbb{F}_2^n, \ell)$, whose vertices are the ℓ-dimensional linear subspaces of $\mathbb{F}_2^n$. Two vertices L and L' are adjacent if and only if $\dim(L \cap L') = \ell - 1$. By a slight abuse of notation, we also use $\mathsf{Gr}(\mathbb{F}_2^n, \ell)$ to denote the vertex set of this graph. The Grassmann encoding of a vector $x \in \mathbb{F}_2^n$ stores, at each location $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$, the local linear function

$$F[L] : L \to \mathbb{F}_2 \text{ defined by } z \mapsto \langle z, x \rangle .$$

Similarly, we ask Bob to prepare his answer $y \in \mathbb{F}_2^{E'}$ in Grassmann encoding, where at each location $L' \in \mathsf{Gr}(\mathbb{F}_2^{E'}, \ell - 1)$ stores the local function

$$G[L] : L' \to \mathbb{F}_2 \text{ defined by } z \mapsto \langle z, y \rangle .$$

The final 2-to-1 game is the composition of the outer PCP with the Grassmann graph gadget. In the composed PCP, upon choosing (E, A) and (E', B) as in $\Psi_{q,\beta,k}$, the verifier chooses a space L' from $\mathsf{Gr}(\mathbb{F}_2^{E'}, \ell-1)$ and a space $L \in \mathsf{Gr}(\mathbb{F}_2^{E'}, \ell)$ containing L'. The subspace $L \subseteq \mathbb{F}_2^{E'}$ will be regarded as a subspace of Alice's space $\mathbb{F}_2^{E}$ by appending 0's in the missing coordinates. The verifier then sends (E, A, L) to Alice and (E', B, L') to Bob, expecting to get linear functions $F[E, A, L]\colon L \to \mathbb{F}_2$ and $G[E', B, L']\colon L' \to \mathbb{F}_2$. The verifier accepts only if $G[E', B, L']$ agrees with the restriction $F[E, A, L]|_{L'}$. This is essentially a 2-to-1 game because for any possible choice of $G[E', B, L']$, there are at most two choices of $F[E, A, L]$ consistent with it.

Covering property and consistency: consider a question (E, A) for Alice in the outer PCP (think of ℓ as much smaller than k), and the Grassmann graph $\mathsf{Gr}(\mathbb{F}_2^{E}, \ell)$. The covering property above asserts that if we sample E' as in the outer PCP and a vertex in the graph $\mathsf{Gr}(\mathbb{F}_2^{E'}, \ell)$, lifting it to $\mathsf{Gr}(\mathbb{F}_2^{E}, \ell)$ yields a nearly uniform vertex in the latter graph. This implies that if Alice's assignment $F[E, A, \cdot]$ is has agreement ε' with an actual Grassmann codeword F, then with decent probability over E', Bob's assignment $G[E', B, \cdot]$ agrees with the projection of F to $\mathbb{F}_2^{E'}$. In short, the covering property ensures the consistency between decodings of assignments of the provers holds.

The linear equations side conditions: recall that in the outer PCP we want to ensure that Alice's response $x \in \mathbb{F}_2^{E}$ satisfies a collection of k linear equations. Note that these equations correspond to requiring that the function $z \mapsto \langle z, x\rangle$ attains specific values on k inputs. More precisely, letting the equations be $e_1, \ldots, e_k$, letting $s^{(i)} \in \mathbb{F}_2^{E}$ be the indicator vector of the variables appearing in e_i, and letting $t_i \in \mathbb{F}_2$ be the constant term of the equation e_i, we want to ensure that $\left\langle s^{(i)}, x\right\rangle = t_i$. To do so, we define the space $N_E = \mathrm{span}(s^{(1)}, \ldots, s^{(k)})$, and request Alice to give assignment $F[E, A, L + N_E]\colon (L + N_E) \to \mathbb{F}_2$ (instead of just $L \to \mathbb{F}_2$). The verifier then checks consistency between Alice and Bob similarly to the above, and additionally that $F[E, A, L + N_E]|_{N_E} = g_E$, where the linear function $g_E\colon N_E \to \mathbb{F}_2$ is defined by $g(s^{(i)}) = t_i$.

2.1.3 On the Non-Standard Notion of Soundness

One would wish that the Grassmann gadget in the inner PCP has the following type of soundness: suppose $F[L]$ is a linear map $L \to \mathbb{F}_2$ for each $L \in \mathsf{Gr}(\mathbb{F}_2^{E}, \ell)$, such that

$$\mathop{\mathbb{P}}_{L_1, L_2 \in \mathsf{Gr}(\mathbb{F}_2^{E}, \ell)}\left[F[L_1]|_{L_1\cap L_2} = F[L_2]|_{L_1\cap L_2} \;\middle|\; \dim(L_1 \cap L_2) = \ell - 1\right] \geqslant \delta \tag{2.1}$$

for some constant $\delta \in (0, 1)$, then there exists a linear function $f : \mathbb{F}_2^{E} \to \mathbb{F}_2$ such that

$$\mathop{\mathbb{P}}_{L \in \mathsf{Gr}(\mathbb{F}_2^{E}, \ell)}\left[F[L] = f|_L\right] \geqslant \delta'$$

for some $\delta' \in (0, 1)$. However, this conclusion cannot be made, as we next explain. Take $\varepsilon > 0$ to be a small constant, and take an arbitrary set $X \subseteq \mathbb{F}_2^{E} \setminus \{0\}$ of size $\varepsilon \cdot 2^{|E|}$. For each $x \in X$ choose a random linear function $f_x\colon \mathbb{F}_2^{E} \to \mathbb{F}_2$, and set a table F on ℓ-dimensional subspaces L as follows:

- If L contains a unique point $x \in X$, define $F[L] = f_x|_L$.
- Otherwise, choose $F[L]$ as a random linear function on L.

Using the inclusion-exclusion formula, one can prove that for a random edge $\{L_1, L_2\}$ in the Grassmann graph, with probability at least $\Omega(\varepsilon)$ there is $x \in X$ such that $L_1 \cap X = L_2 \cap X = \{x\}$, and F satisfies such constraints. Thus, F passes the Grassmann test Equation (2.1) with $\delta = \Omega(\varepsilon)$. However, it only agrees on $o(1)$ locations with any given codeword from the Grassmann code.

This example tells us that while there might be no linear function that agrees with F globally, it may be possible to extract a linear function by focusing on certain substructures inside the Grassmann graph. Indeed, it turns out that what can be concluded from Equation (2.1) is something of this form: there exists $Q \in \mathsf{Gr}(\mathbb{F}_2^E, q)$ and $W \in \mathsf{Gr}(\mathbb{F}_2^E, |E| - q)$, where q depends only on δ, and a linear function $f : W \to \mathbb{F}_2$ such that

$$\underset{L \in \mathsf{Gr}(\mathbb{F}_2^E, \ell)}{\mathbb{P}} \left[F[L] \equiv g|_L \,\middle|\, Q \subseteq L \subseteq W \right] \geqslant \delta'. \tag{2.2}$$

We refer to the spaces Q and W in Equation (2.2) as zoom-ins and zoom-outs, respectively, and encourage the reader to ignore W (it will irrelevant for the majority of the discussion). In words, Equation (2.2) asserts that if F passes the Grassmann test with noticeable probability, then inside non-trivial fraction of zoom-ins of constant dimension q, it does have a global linear structure.

The proof of [KMS25, DKK+25, DKK+21, KMS23] has to deal with this non-standard notion of soundness, and this is the purpose of the advice feature in the game $\Psi_{q,\beta,k}$. More precisely, given Alice and Bob's tables of assignments F and G for the composed PCP, one needs to be able to derive a strategy for the original game $\Psi_{q,\beta,k}$. However, the strategy that one can derive using Equation (2.2) heavily depends on the choice of what zoom-in Q is used. The advice feature allows the two provers in $\Psi_{q,\beta,k}$ to use (essentially) the same Q for decoding – namely the span of their advice vectors – which guarantees consistency between their answers.

2.1.4 Difficulties in Achieving Perfect Completeness

The construction for 2-to-1 games with imperfect completeness can be illustrated using Figure 1.

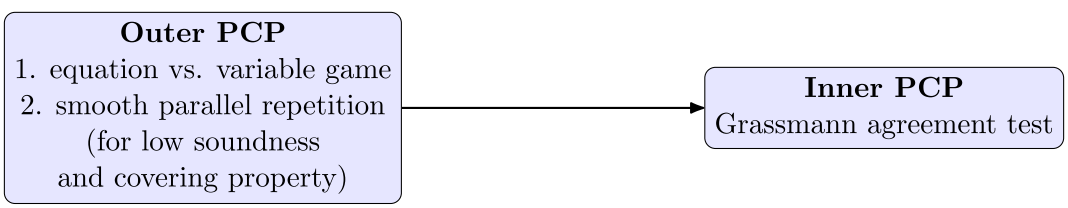

Figure 1: The imperfect-completeness construction of [KMS25, DKK+25, DKK+21, KMS23]

The above construction loses perfect completeness right from the start: it uses the $\mathsf{3Lin}_{\mathbb{F}_2}$ problem, which inherently has imperfect completeness. Moreover, in some sense this seems necessary: the Grassmann encoding presented above only allows for linear checks in the encoded string $x \in \mathbb{F}_2^n$.

Thus, to achieve perfect completeness it is necessary to come up with an encoding scheme which is more powerful than the Grassmann encoding, in the sense that it allows for non-linear checks on the encoded string $x \in \mathbb{F}_2^E$. By itself this is not a hard task; for example, the long-code would work. However, the key difficulty is that one additionally needs the encoding scheme to satisfy a covering property, and then build an outer PCP with a suitable advice feature.

2.2 Our Techniques

In this section we discuss the new ideas that go into our construction.

2.2.1 A First Attempt: the Quadratic Grassmann Encoding

We begin by outlining a natural first attempt at a reduction achieving perfect completeness. It starts with the problem of quadratic equations, $\mathsf{3QS}_{\mathbb{F}_2}$, defined as follows.

Definition 2.2. An instance of the $\mathsf{3QS}_{\mathbb{F}_2}$ problem consists of a set of variables $\mathscr{V}$ and a set of equations $\mathscr{E}$, where each $e \in \mathscr{E}$ is a quadratic equation over $\mathbb{F}_2$ containing 3 distinct variables.

Using standard results, it can be proved that there is $s < 1$ such that $\mathsf{Gap3QS}_{\mathbb{F}_2}[1, s]$ is **NP**-hard. Starting with an instance $(\mathscr{V}, \mathscr{E})$ of $\mathsf{3QS}_{\mathbb{F}_2}$, one can construct the β-smooth k-fold repeated game corresponding to it, $\Psi_{\beta,k} = \Psi_{\beta,k}(\mathscr{V}, \mathscr{E})$ in the same way as in Section 2.1.1 (we ignore the advice for now). The main question is how to compose it with a Grassmann-style encoding so as to get to a d-to-1 Game.

In the game $\Psi_{\beta,k}$, the verifier performs two types of checks: consistency between the responses of Alice and Bob, and quadratic equations in the response of Alice. The Grassmann encoding is compatible with the former tye of checks, but not with the latter type of checks, in the sense that we do not know how to incorporate checking quadratic equations on the encoded message into the Grassmann encoding itself. Instead, it is natural to consider the *quadratic Grassmann encoding*.

Let $x \in \mathbb{F}_2^E$ represent Alice's response in the game $\Psi_{\beta,k}$. In the quadratic Grassmann encoding, one considers the functions over $\mathbb{F}_2^{E\times E}$ defined by $z \mapsto \langle z, x \otimes x\rangle$, and takes its Grassmann encoding QF of x, which assigns to each ℓ-dimensional subspaces $L \subseteq \mathbb{F}_2^{E\times E}$ the linear function

$$QF[L] : L \to \mathbb{F}_2 \text{ defined by } z \mapsto \langle z, x \otimes x\rangle .$$

The main benefit of the quadratic Grassmann encoding is that it indeed allows us to test quadratic equations as side conditions, as we explain next.

The quadratic equations side conditions: Given legitimate quadratic Grassmann codeword QF, it is possible to incorporate with it checks of quadratic equations on the encoded message. Indeed, suppose that $e_1, \ldots, e_k$ are quadratic equations over E, let $s^{(1)}, \ldots, s^{(k)} \in \mathbb{F}_2^{E\times E}$ be the indicator vectors of the pairs that appear in $e_1, \ldots, e_k$ respectively, and let $t_1, \ldots, t_k$ be the constant terms. Then the equation e_i can be written as $\langle s^i, x \otimes x\rangle = t_i$ for each $i \in \{1, \ldots, k\}$. As before, this can be folded into the encoding by defining $N_E = \mathrm{span}(s^{(1)}, \ldots, s^{(k)})$, asking for the table to assign values to the sum $L + N_E$ (instead of to just L), and require that their projection onto N_E completely agrees with the linear function $g\colon N_E \to \mathbb{F}_2$ defined by $g(s^{(i)}) = t_i$. This way of incorporating quadratic side conditions will work for all of the constructions we present below, so for simplicity of presentation they will be omitted henceforth.

On the other hand, the quadratic Grassmann encoding has 2 very significant drawbacks:

(1) **Codeword testing:** Even ignoring the issue of non-standard-type soundness, the Grassmann agreement test from Section 2.1.2 is not actually a tester for the quadratic Grassmann code. Namely, if a table G that assigns each each ℓ-dimensional subspace $L \subseteq \mathbb{F}_2^{E\times E}$ a linear function $G[L]\colon L \to \mathbb{F}_2$ passes it with non-negligible probability, then at best we will be able to argue that it has significant agreement with the Grassmann encoding of an arbitrary $x' \in \mathbb{F}_2^{E\otimes E}$, i.e., not necessarily of rank 1. We need an additional feature that will guarantee that x' has rank 1 (or at the very least small rank) and is symmetric.

(2) **The covering property:** The covering property does not hold for the Grassmann encoding. Namely, fixing a question for Alice E and looking at all possible questions E' for Bob (conditioned on E), the (lifting of the) spaces $\mathbb{F}_2^{E'\times E'}$ do not cover $\mathbb{F}_2^{E\times E}$. Indeed, the number of

possible questions for Bob is roughly $\binom{3k}{3\beta k}$, and the size of Bob's space is roughly $2^{9(1-\beta)^2k^2}$, so in total they can cover at most

$$\binom{3k}{3\beta k}2^{(1-\beta)^2k^2} \leqslant 2^{O(\beta k\log(1/\beta))}2^{9k^2-9\beta k^2} = o(2^{9k^2})$$

points in $\mathbb{F}_2^{E\times E}$, which is a tiny fraction.

The first issue, namely ensuring that the encoded message has a low rank, is significantly easier to handle, and to do so we incorporate tensor tests. More precisely, instead of asking for the value of the Grassmann encoding at a subspace $L \subseteq \mathbb{F}_2^{E\times E}$, we will also take a random subspace $U \subseteq \mathbb{F}_2^E$ of dimension ℓ, and then ask for the value at the subspace $L + U \otimes U$, requiring that the projection onto $U\otimes U$ is of rank 1. This yields the following 4-to-4 Grassmann agreement + tensor test. Upon receiving a table F assigning linear functions to spaces $L + U \otimes U$ as above, do the following:

1. Sample $L_1, L_2 \subseteq \mathbb{F}_2^{E\times E}$ of dimension ℓ intersecting at dimension $\ell - 1$, and $U_1, U_2 \subseteq \mathbb{F}_2^E$ of dimension ℓ intersecting at dimension $\ell - 1$.

2. Query $F[L_1 + U_1 \otimes U_1]$ and $F[L_2 + U_2 \otimes U_2]$. Accept if they agree on

$$L_1 \cap L_2 + (U_1 \cap U_2) \otimes (U_1 \cap U_2),$$

and their respective projections onto $U_1 \otimes U_2$ and $U_1 \otimes U_2$ are of rank at most 1.

Given the (non-standard) soundness of the basic Grassmann agreement test, one can analyze the Grassmann agreement + tensor test and show that it satisfies a similar (non-standard) notion of soundness. Also, this test is 4-to-4: given a linear function on L_1, there are at most 2 linear functions on L_2 that agree with it, and given a rank-1 linear function on $U_1\otimes U_1$, there are at most 2 rank-1 linear functions on $U_2 \otimes U_2$ agreeing with it.

Ensuring symmetry is more straightforward. To do that, define for distinct points $i, j \in [k]$ define the matrix $M_{i,j} \in \mathbb{F}_2^{E\times E}$ which is 1 only at locations (i, j) and (j, i), and set $N_{\mathsf{sym}} = \text{span}(\{M_{i,j} \mid i \neq j\})$. Then, $x' \in \mathbb{F}_2^{E\times E}$ is symmetric if and only if the function $z \mapsto \langle z, x'\rangle$ vanishes on N_{sym}. Thus, instead of requesting an assignment on $L+U\otimes U$, the verifier requests an assignment on $L + U \otimes U + N_{\mathsf{sym}}$, and checks that its projection on N_{sym} vanishes.

The second issue, namely the fact that the covering property fails, is substantially more difficult to handle, and in the next section we begin the discussion about it.

2.2.2 Variations on the Quadratic Grassmann Encoding

From now, it will be convenient for us to think of points from $\mathbb{F}_2^{E\times E}$ and $\mathbb{F}_2^{E'\times E'}$ as matrices. The covering property fails for the quadratic Grassmann encoding because if, letting Bob's question be denoted by E', we have $i \in E \setminus E'$, then i-th row and column will always be 0 in liftings from $\mathbb{F}_2^{E'\times E'}$ to $\mathbb{F}_2^{E\times E}$. On the other hand, typical matrices from $\mathbb{F}_2^{E\times E}$ do not contain fully-0 rows. This raises the question of whether we can replace the structure of $E \times E$ with a sparser structure that circumvents this atypical behaviour of liftings. Moreover, requiring a response from the provers for each pair of variables in their respective variable sets is not necessary, as the quadratic equations we wish to check only require access to $\Theta(k)$ pairs (out of the $\binom{k}{2}$ many ones that exist).

This view motivates the following general recipe: can we find a graph H with vertex set E and edge set $\mathcal{E}$, such that we can restrict the quadratic Grassmann encoding to $\mathbb{F}_2^{\mathcal{E}}$ in a meaningful way? The quadratic Grassmann encoding discussed above corresponds to the choosing H to be a complete graph, but there are many other options for H. Below are a few simple choices we tried:

(1) Expander graphs: One can try to take H to be an arbirary sparse expander graph. It can be shown that if the degree of H is small then now the covering property will hold. Unfortunately, an arbitrary expander graph does not support any form of tensor tests.

(2) Cliques in a sparse high-dimensional expander: one can take a sparse high-dimensional expander $X = (X(0), \dots, X(k))$, where $X(0)$ is the vertex set, and for each i, $X(i) \subseteq \binom{X(0)}{i+1}$, and take H to be the graph underlying it. Because the degree of X is constant, the covering property holds. Also, we can perform local type of tensor tests: if $K \in X(k)$ is a $(k+1)$-face, then K is a clique in X, so we can choose a subspace $U \subseteq \mathbb{F}_2^{X(0)}$ fully supported in K, and have access to all of the entries in $U \otimes U$.

This observation initially seems very promising, but it has one serious issue. We can only afford to perform such a local tensor test for a single clique (or perhaps $O(1)$-many cliques if we are fine with increasing the value of d for which we are proving Conjecture 1.5). Unfortunately, these localized tensor tests do not actually gaurantee that the encoded function has a small rank, and Alice and Bob can use a cheating strategy in the following way. When Alice receives a space of the type $L + U \otimes U$, she extracts from it a rank-1 vector, and takes any element v in its support. Then it is highly likely that v belong to K, and because X has constant degree, the neighbourhood $N(v)$ of v has constant size. Thus, Alice can choose a linear function on $\mathbb{F}_2^{X(1)}$ which is rank-1 on $\mathbb{F}_2^{X(1)\cap(N(v)\times N(v))}$ but otherwise arbitrary, and responds according to it. Bob proceeds in the same way, where they agree beforehand on what is the rank-1 function they will use, and what the arbitrary function they will use on the rest of the space.

We remark that another attempt at implementing tensor tests via high-dimensional expanders is based on coboundary-type expansion, and we have not explored this thoroughly.

In summary, the attempts above suggest that there are meaningful variants of the quadratic Grassmann encoding that enjoy the covering property, but they comes at the cost of, at best, being able to ensure small rank locally. This type of guarantee seems insufficient, at least with an outer PCP along the lines of Section 2.1.1.

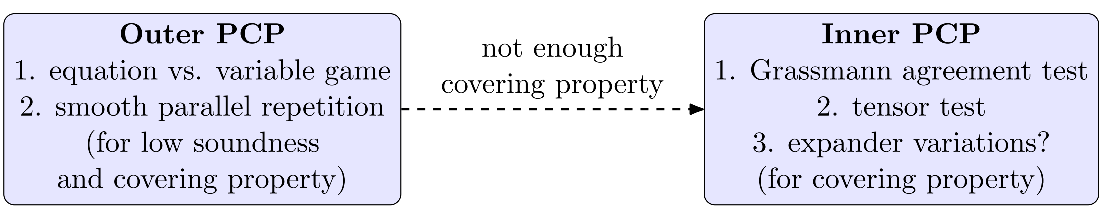

Figure 2: Our failed attempt for perfect completeness

2.2.3 A Second Attempt: Three-Step Construction

So far, the described attempts to create covering property for a quadratic-Grassmann-based inner PCP are not successful. In particular, it seems that the idea of "sparsifying" the quadratic-Grassmann encoding conflicts with tensor test. To address this conflict, we separate the tensor test from the inner PCP and name it the "middle PCP" (see Figure 3). Notice that the tensor test (proposed in Section 2.2.1) is a Grassmann structure over $\mathbb{F}_2^{E}$, instead of over $\mathbb{F}_2^{E\times E}$. Therefore, the tensor test by itself is compatible with the covering property provided by the smooth parallel repetition in the outer PCP. We only have to create more covering property for the inner PCP (which now consists only of a Grassmann agreement test) *after* incorporating tensor tests into the middle PCP. This will be provided by a "lazy parallel repetition" on the level of the middle PCP (see Figure 3).

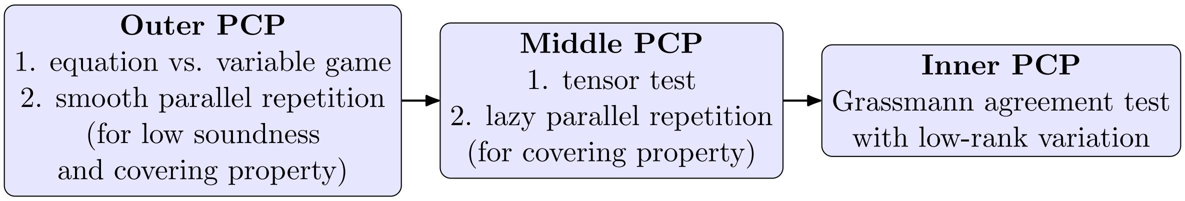

Figure 3: Our construction for perfect completeness

3 Preliminaries

In this section we set up basic tools that will be used throughout the paper.

3.1 General Notations

Linear Algebra. Given a finite dimensional $\mathbb{F}_2$-vector space V, we let V^* denote the dual space of V, which consists of all linear functionals $f : V \to \mathbb{F}_2$. For any finite-dimensional $\mathbb{F}_2$-vector spaces V_1, V_2 and any bilinear form $f : V_1 \otimes V_2 \to \mathbb{F}_2$, we define the linear map $T_f : V_1 \to V_2^*$ by letting $T_f(x)$ be the functional $y \mapsto f(x \otimes y)$.

Grassmann Graph. Given a finite dimensional $\mathbb{F}_2$-vector space V and a positive integer $\ell \leqslant \dim V$, the Grassmann graph $\mathsf{Gr}(V, \ell)$ has as vertices all ℓ-dimensional subspaces $L \subseteq V$, and two vertices are adjacent if they intersect in dimension $\ell - 1$. We denote by $\mathcal{G}(V, \ell)$ the uniform distribution over the edges of $\mathsf{Gr}(V, \ell)$.

3.2 Technical Ingredients from Prior Work

3.2.1 Approximate Coloring of Hypergraphs

Theorem 3.1. *For every fixed positive integer c, there exists a constant $\varepsilon \geqslant \exp(-\exp(O(c^3)))$ such that given a 3-uniform hypergraph H, it is* **NP***-hard to distinguish between the following cases:*

(1) Yes case: H is 2-colorable.

(2) No case: in every c-coloring of H, at least ε fraction of the edges are monochromatic.

Proof. Deferred to Appendix A. □

3.2.2 Grassmann Agreement Test

In this section we state a more precise version of the claim in Equation (2.2) regarding the soundness of the Grassmann test. It will be convenient for us to think of partial assignment, namely tables F that assign to each ℓ-dimensional subspace $L \subseteq \mathbb{F}_2^n$ either a linear function $F[L]\colon L \to \mathbb{F}_2$, or the nullity element `nil`. Given a table F the passing probability of the Grassmann test is

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[F[L_1]|_{L_1\cap L_2} = F[L_2]|_{L_1\cap L_2}\right],$$

where the event does not hold if either $F[L_1]$ or $F[L_2]$ are $\texttt{nil}$. The following result is due to [KMS23], but we use a slightly different formulation which we derive directly from their result.

Theorem 3.2. *There is an absolute constant $C > 0$ such that the following holds. Let n, ℓ, q be nonnegative integers such that $n - q \geqslant \ell \geqslant q + 1$. Suppose $F[\cdot]$ is a table where $F[L] \in L^* \cup \{\texttt{nil}\}$ for each $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$, such that*

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)} \left[F[L_1]|_{L_1\cap L_2} = F[L_2]|_{L_1\cap L_2} \right] \geqslant 2^{-q+5}. \tag{3.1}$$

Then for at least $2^{-q-\ell^2}$ fraction of subspaces $Q \in \mathsf{Gr}(\mathbb{F}_2^n, q)$, there exist a subspace $W \in \mathsf{Gr}(\mathbb{F}_2^n, n-q)$ containing Q and a linear map $f : W \to \mathbb{F}_2$ such that

$$\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)} \left[F[L] = f|_L \,\middle|\, Q \subseteq L \subseteq W \right] \geqslant 2^{-Cq^2}.$$

Proof. Deferred to Appendix D. □

We also need the following result, which asserts that if the simpler "equality test" passes with noticeable probability, then for many zoom-ins Q of constant dimension, one could fine a zoom-out W of constant codimension such that the table F has significant agreement with a constant on $Q \subseteq L \subseteq W$.

Theorem 3.3. *There is an absolute constant $C > 0$ such that the following holds. Let n, ℓ, q be nonnegative integers such that $n - q \geqslant \ell \geqslant q + 1$. Suppose Σ is a finite set and $F[\cdot]$ is a table where $F[L] \in \Sigma \cup \{\texttt{nil}\}$ for each $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$, such that*

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)} \left[F[L_1] = F[L_2] \neq \texttt{nil} \right] \geqslant 2^{-q+4}.$$

Then for at least $2^{-q-\ell^2}$ fraction of subspaces $Q \in \mathsf{Gr}(\mathbb{F}_2^n, q)$, there exist an element $\sigma \in \Sigma$ and a subspaces $W \in \mathsf{Gr}(\mathbb{F}_2^n, n-q)$ containing Q such that

$$\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)} \left[F[L] = \sigma \,\middle|\, Q \subseteq L \subseteq W \right] \geqslant 2^{-Cq^2}.$$

Proof. Deferred to Appendix D. □

3.2.3 List Decoding Bounds

When we use Theorems 3.2 and 3.3 we will often need to argue that the number of possible decodings is not too large. In this section we precisely define this notion and state the bounds we will use, all are from [DKK+25].

Fix a vector space $X := \mathbb{F}_2^n$ and let $F[\cdot]$ to be a table which assigns a linear function to every ℓ-dimensional subspace of X.

Definition 3.4 ([DKK+25])**.** Let n, ℓ, q, w be nonnegative integers such that $n - w \geqslant \ell \geqslant q + 1$. Fix parameters $\tau_0, \tau_1, \ldots, \tau_w \in (0, 1)$ and write $\vec{\tau} = (\tau_0, \ldots, \tau_w)$. Fix an $\mathbb{F}_2$-vector space V of dimension n and a subspace $Q \in \mathsf{Gr}(V, q)$. Suppose $F[\cdot]$ is a table where $F[L] \in L^* \cup \{\texttt{nil}\}$ for each $L \in \mathsf{Gr}(V, \ell)$. For a linear subspace $W \subseteq V$ with $\operatorname{codim} W \leqslant w$ and a linear functional $f : W \to \mathbb{F}_2$, we say that the pair (f, W) is *$\vec{\tau}$-occurring in $F[\cdot]$ modulo Q* if

$$\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(W,\ell)} \left[F[L] = f|_L \,\middle|\, Q \subseteq L \right] \geqslant \tau_i, \qquad \text{where } i = n - \dim W.$$

Definition 3.5. In the setting and notation of Definition 3.4, we say that (f, W) is *$\vec{\tau}$-maximal in $F[\cdot]$ modulo Q* if

(1) (f, W) is $\vec{\tau}$-occurring in $F[\cdot]$ modulo Q, but

(2) for any linear subspace $W' \subseteq \mathbb{F}_2^n$ that properly contains W and any linear functional $f' : W' \to \mathbb{F}_2$, the pair (f', W') is not $\vec{\tau}$-occurring in $F[\cdot]$ modulo Q.

The set of all $\vec{\tau}$-maximal pairs will be denoted by $\mathsf{LIST}_Q^{\vec{\tau}}(F)$.

We use the following bound on the size of maximal pairs proved by [DKK+25].

Lemma 3.6 ([DKK+25])**.** *There is an absolute constant $C \geqslant 40$ such that the following holds. Let n, ℓ, q, w be nonnegative integers such that $n - w \geqslant 2\ell \geqslant C(q + 1)$. Suppose $\vec{\tau} = (\tau_0, \dots, \tau_w)$ is a vector such that*

$$\tau_0 \geqslant 2^{-\ell/C} \qquad \textit{and} \qquad \tau_{i-1} \leqslant C^{-1}\tau_i^C \text{ for each } i \in [w].$$

Then for any table $F[\cdot]$ where $F[L] \in L^* \cup \{\mathtt{nil}\}$ for each $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$, we have

$$\left|\mathsf{LIST}_Q^{\vec{\tau}}(F)\right| \leqslant \exp(Cw^2\ell).$$

To handle large collections of zoom-outs we will need to make use of the notion of *sunflower* of zoom-outs, defined as follows.

Definition 3.7. Let V be a vector space over $\mathbb{F}_2$, and let $W_1, \dots, W_m \subseteq V$ be subspaces of co-dimension r. For $W \subseteq V$, we say that $(W, W_1, \dots, W_m)$ forms a sunflower if

1. $W_i \subseteq W$ for all i, and we denote its relative codimension by $\mathsf{codim}_W(W_i) = s$.
2. For all $i \neq j$, $\mathsf{codim}_W(W_i \cap W_j) = 2s$.

In words, a sunflower is a collection of subspaces $W_1, \dots, W_m$ living inside a subspace W that have as small as possible pairwise intersections in it. The following result, from [DKK+25], asserts that any sufficiently large collection of zoom-outs of constant codimension contains a large sunflower.

Lemma 3.8 ([DKK+25])**.** *Let $W_1, \dots, W_N \subseteq X$ be distinct vector spaces of codimension r. Let m be an integer such that $(m \cdot 2^r)^r \leqslant N$, then, there exists a subspace $W \subseteq X$ of codimension $r - s < r$ that W contains m of W_is, say, $W_1, \dots, W_m$, and for every $i \neq j \in [m]$, $\mathrm{codim}_W(W_i \cap W_j) = 2s$. Such tuple $(W, W_1, \dots, W_m)$ is called a sunflower.*

3.3 Parameter Relations

For future reference, we explicitly state the relation between the various parameters that will be used throughout the proof.

Condition 3.9. Let $k, q_0, q_1, q_2, t, r, \ell, s$ be positive integers, let $\delta, \beta, \varepsilon \in (0, 1)$, and let $\vec{\tau} = (\tau_0, \tau_1, \dots, \tau_{2q_0})$ be a vector such that for a sufficiently large constant $C > 0$, the conditions in Table 1 hold.

Parameter	Requirement	First appears in	Why require this
$\delta \in (0,1)$		inner PCP	
$q_0 \in \mathbb{N}^+$	$q_0 \geqslant C \log_2(1/\delta)$	inner PCP	inner PCP
$\tau_{2q_0} \in (0,1)$	$\tau_{2q_0} \leqslant \exp(-q_0^C)$	inner PCP	inner PCP
$\tau_{2q_0-1}, \dots, \tau_0 \in (0,1)$	$\tau_{i-1} \leqslant C^{-1}\tau_i^C,\ \forall i \in [2q_0]$	inner PCP	inner PCP
$\ell_1 \in \mathbb{N}^+$	$\ell_1 \geqslant C \log_2(1/\tau_0)$	inner PCP	inner PCP
$s \in \mathbb{N}^+$	$s = 3\ell_1$	inner PCP	inner PCP
$q_1 \in \mathbb{N}^+$	$q_1 = 2^{2\ell_1 - q_0} s$	outer PCP	inner PCP
$q_2 \in \mathbb{N}^+$	$q_2 \geqslant C\ell_1^2$	outer PCP	middle PCP
$\ell_2 \in \mathbb{N}^+$	$\ell_2 \geqslant Cq_2^2$	middle PCP	middle PCP
$r \in \mathbb{N}^+$	$r = 2q_2 + 1$	outer PCP	middle PCP
$\varepsilon \in (0,1)$	$\varepsilon \leqslant \exp(-\exp(\exp(Cr)))$	outer PCP	outer PCP
$k \in \mathbb{N}^+$	$k \geqslant \varepsilon^{-C} \exp(\exp(\ell_2))$	outer PCP	outer PCP
$\beta \in (0,1)$	$\beta = k^{-3/4}$	outer PCP	outer PCP
$t \in \mathbb{N}^+$	$t \geqslant \exp(Ck^4)$	middle PCP	inner PCP

Table 1: Parameter relations

4 The Outer PCP

4.1 Tripled Sets

In this section we set up some convenient notations for our application of the parallel repetition theorem. First, it will be convenient for us to remember the order of challenges, and we will use formalize it using the following two definitions.

Definition 4.1. A *tripled set* is a finite set E equipped with a partition into subsets of size 3. We refer to the parts of this implicit partition as the *triples* of E, and we write $\mathrm{Triples}(E)$ for collection of its triples.

Definition 4.2. Let H be a 3-uniform hypergraph with vertex set $\mathscr{V}(H)$ and edge set $\mathscr{E}(H)$. Given a positive integer k, we let $H^{\sqcup k}$ denote the k-fold vertex-disjoint union of H. More explicitly, the vertex set of $H^{\sqcup k}$ is $[k] \times \mathscr{V}(H)$, and we denote $\mathscr{V}(H,k) = [k] \times \mathscr{V}(H)$; the edge set of $H^{\sqcup k}$ is $\bigcup_{i=1}^k \mathscr{E}_i(H,k)$, where

$$\mathscr{E}_i(H,k) = \big\{\{(i,\mathsf{v}_1),(i,\mathsf{v}_2),(i,\mathsf{v}_3)\} \,\big|\, \{\mathsf{v}_1,\mathsf{v}_2,\mathsf{v}_3\} \in \mathscr{E}(H)\big\}.$$

We let $\mathscr{E}(H,k)$ be the collection of all vertex subsets $E \subseteq [k] \times \mathscr{V}(H)$ that result from picking one edge from each $\mathscr{E}_i(H,k)$. That is, it is E of the form

$$E = \{v_{i,j} \mid i \in [k],\ j \in [3]\} \qquad \text{such that} \qquad \{v_{i,1}, v_{i,2}, v_{i,3}\} \in \mathscr{E}_i(H,k) \text{ for all } i \in [k], \tag{4.1}$$

and we view it naturally as a tripled set, with $\mathrm{Triples}(E) = \big\{\{v_{i,1}, v_{i,2}, v_{i,3}\} \,\big|\, i \in [k]\big\}$.

In words, viewing the set of hyperedges in H as the challenges that a verifier samples, the edges in $H^{\sqcup k}$ are naturally viewed as the challenges that the verifier chooses in the k-fold repeated game.

Remark 4.3. For a tripled set E of the form (4.1), the k triples of E are naturally ordered. For each $i \in [k]$, we will call $\{v_{i,1}, v_{i,2}, v_{i,3}\}$ the i-th triple of E.

4.1.1 Thinning, and Thinning with Advice

Next, we set up some notations that will be helpful in formalizing the smoothness in our application of parallel repetition. The following definition will be helpful in formalizing the fact that in smooth parallel repetition, on each coordinate the two players get identical challenge with probability $1-\beta$, and otherwise Alice gets the whole challenge and Bob only gets a single variable from it.

Definition 4.4. Let E be a tripled set and let $\beta \in [0,1]$. We write $E' \sim \mathrm{Thin}_\beta(E)$ for the random subset $E' \subseteq E$ obtained as follows. Independently for each triple $T \in \mathrm{Triples}(E)$, we include all elements of T in E' with probability $1-\beta$, and include a uniformly random element of T in E' with probability β. We stress that the output E' itself is only a set, not a tripled set.

Next, we set up notations that will be helpful in handling the advice feature in smooth parallel repetition.

Definition 4.5. Let E be a tripled set. Fix nonnegative integers q_1, q_2 and a parameter $\beta \in [0,1]$. We let $\mathcal{J}_{\mathsf{out}}(E, q_1, q_2, \beta)$ denote the output distribution of the following process:

1. Sample a β-thinning $E' \sim \mathrm{Thin}_\beta(E)$.
2. Sample $q = q_1 + q_2$ vectors $a^{(1)}, \ldots, a^{(q)} \in \mathbb{F}_2^{E'}$ independently and uniformly at random. We denote $A_{\mathsf{in}} = \left(a^{(1)}, \ldots, a^{(q_1)}\right)$ and $A_{\mathsf{mid}} = \left(a^{(q_1+1)}, \ldots, a^{(q_1+q_2)}\right)$.
3. Output the tuple $(E', A_{\mathsf{in}}, A_{\mathsf{mid}})$.

Remark 4.6. For a subset E' of a finite set E, we always view $\mathbb{F}_2^{E'}$ as a linear subspace of the ambient space $\mathbb{F}_2^{E}$. In particular, the vectors $a^{(1)}, \ldots, a^{(q)}$ in Definition 4.5 are viewed also as elements of the vector space $\mathbb{F}_2^{E}$.

4.1.2 The Covering Property

We need the following two simple propositions essentially from [DKK+25]. The first one asserts that fixing E and sampling $(E', A_{\mathsf{in}}, A_{\mathsf{mid}})$ as above conditioned on the value of $A_{\mathsf{in}}, A_{\mathsf{mid}}$, the triplets in E' are independent, the for each i, the probability E' contains the entire ith triplet does not change too much.

Proposition 4.7. Let E be a tripled set. Fix nonnegative integers q_1, q_2 and a parameter $\beta \in [0,1)$. Fix arbitrary tuples

$$A_1 \in (\mathbb{F}_2^E)^{q_1} \qquad \text{and} \qquad A_2 \in (\mathbb{F}_2^E)^{q_2}.$$

Sample

$$(E', A_{\mathsf{in}}, A_{\mathsf{mid}}) \sim \mathcal{J}_{\mathsf{out}}(E, q_1, q_2, \beta)$$

conditioned on the event $\{A_{\mathsf{in}} = A_1\} \cap \{A_{\mathsf{mid}} = A_2\}$. Then the resulting marginal distribution of E' is independent over triples of E. Furthermore, each triple of E is entirely included in E' with conditional probability at least $1 - 4^{q_1+q_2}\beta$.

Proof. Deferred to Appendix C. □

The second one asserts that the distribution of $(A_{\mathsf{in}}, A_{\mathsf{mid}})$, once lifted to $\mathbb{F}_2^E$ as discussed in Remark 4.6, is close to uniform so long as $\beta = o\left(\frac{1}{\sqrt{|E|}}\right)$.

Proposition 4.8. Let E be a tripled set. Fix nonnegative integers q_1, q_2 and a parameter $\beta \in [0, 1)$. Fix an arbitrary tuple $A_1 \in (\mathbb{F}_2^E)^{q_1}$. Sample

$$(E', A_{\mathsf{in}}, A_{\mathsf{mid}}) \sim \mathcal{J}_{\mathsf{out}}(E, q_1, q_2, \beta)$$

conditioned on the event $\{A_{\mathsf{in}} = A_1\}$. Then the resulting marginal distribution of A_{mid} has total variation distance at most $4^{q_1+q_2}\beta\sqrt{|E|/3}$ to the uniform distribution over $(\mathbb{F}_2^E)^{q_2}$.

Proof. Deferred to Appendix C. □

4.2 The NAE Predicate and Satisfiability Notions

In this section we formally present the Not-All-Equal predicate that we will use, along with the notions of satisfiablity and joint NAE-satisfiability that are arguments rely on.

Definition 4.9. We define a function $\mathrm{NAE}_3 : \mathbb{F}_2^3 \to \mathbb{F}_2$ defined by

$$\mathrm{NAE}_3(x_1, x_2, x_3) = x_1 + x_2 + x_3 + x_1x_2 + x_2x_3 + x_3x_1.$$

We note that $\mathrm{NAE}_3(x_1, x_2, x_3) = 1$ if and only if x_1, x_2, x_3 are not all equal.

As discussed in the introduction, we will use the notion of jointly NAE-satisfying, which we recall below. When we get a linear function on $\mathbb{F}_2^E$ which is an encoding of an assignment to variables, we use the following definition.

Definition 4.10. Let E be a tripled set and let k, r be positive integers. Given r linear functionals $f^{(1)}, \ldots, f^{(r)} : \mathbb{F}_2^E \to \mathbb{F}_2$, let $a^{(1)}, \ldots, a^{(r)} \in \mathbb{F}_2^E$ be the unique vectors such that

$$f^{(i)}(x) = \sum_{v \in E} a_v^{(i)} x_v \qquad \text{for all } i \in [r] \text{ and } x \in \mathbb{F}_2^E.$$

If we have

$$\sum_{i=1}^{r} \mathrm{NAE}_3\big(a_{v_1}^{(i)}, a_{v_2}^{(i)}, a_{v_3}^{(i)}\big) = 1 \qquad \text{for all } \{v_1, v_2, v_3\} \in \mathrm{Triples}(E),$$

then the linear functionals $f^{(1)}, \ldots, f^{(r)}$ are said to be *jointly* NAE-*satisfying*. In the case $r = 1$, then the functional $f^{(1)}$ is simply said to be NAE-*satisfying*.

More generally, when we get a symmetric linear function on $\mathbb{F}_2^E \otimes \mathbb{F}_2^E$, which is a supposed quadratic Hadamard encoding of a message, we use the following definition.

Definition 4.11. Let E be a tripled set and let k be a positive integer. Given a symmetric bilinear map $f : \mathbb{F}_2^E \otimes \mathbb{F}_2^E \to \mathbb{F}_2$, let $X \in \mathbb{F}_2^{E \times E}$ be the unique matrix such that

$$f(x \otimes y) = \sum_{u,v \in E} X_{uv} x_u y_v \qquad \text{for all } x, y \in \mathbb{F}_2^E.$$

If we have

$$X_{v_1v_1} + X_{v_2v_2} + X_{v_3v_3} + X_{v_1v_2} + X_{v_2v_3} + X_{v_3v_1} = 1 \qquad \text{for all } \{v_1, v_2, v_3\} \in \mathrm{Triples}(E),$$

then the bilinear map f is said to be NAE-*satisfying*.

4.2.1 Linear Functionals vs. Bilinear Forms

Recall that in our reduction, we will want to switch from low-value with respect to joint NAE-satisfiability, to low-value with respect to bilinear forms of low rank. To facilitate this move, we will use the following definition and proposition.

Definition 4.12. Let V be a finite-dimensional $\mathbb{F}_2$-vector space and let r be a positive integer. Given a symmetric bilinear map $f : V \otimes V \to \mathbb{F}_2$ and r linear functionals $f^{(1)}, \ldots, f^{(r)} : V \to \mathbb{F}_2$, we say that $\left(f^{(1)}, \ldots, f^{(r)}\right)$ is a *tensor decomposition* of f if the following conditions hold:

(1) For each $i \in [r]$, we have $f^{(i)} \in \mathrm{image}(T_f)$. In other words, there exists a vector $x^{(i)} \in V$ such that $f^{(i)}(y) = f(x^{(i)} \otimes y)$ for each $y \in V$.

(2) For each $x, y \in V$, we have

$$f(x \otimes y) = \sum_{i=1}^{r} f^{(i)}(x) f^{(i)}(y). \tag{4.2}$$

The following proposition shows that if a given symmetric bilinear form that NAE-satisfies a constraint, then any tensor decomposition of it also satisfies that constraint.

Proposition 4.13. Let E be a tripled set and let r be a positive integer. Suppose $f : \mathbb{F}_2^E \otimes \mathbb{F}_2^E \to \mathbb{F}_2$ is an NAE-satisfying symmetric bilinear map. Then any tensor decomposition $\left(f^{(1)}, \ldots, f^{(r)}\right)$ of f is jointly NAE-satisfying.

Proof. let $a^{(1)}, \ldots, a^{(r)} \in \mathbb{F}_2^E$ be the unique vectors such that

$$f^{(i)}(x) = \sum_{v \in E} a_v^{(i)} x_v \qquad \text{for all } i \in [r] \text{ and } x \in \mathbb{F}_2^E.$$

Let $X \in \mathbb{F}_2^{E \times E}$ be the matrix defined by

$$X_{uv} = \sum_{i=1}^{r} a_u^{(i)} a_v^{(i)} \qquad \text{for all } u, v \in E.$$

Plugging the expansion of $f^{(i)}$ into (4.2) and changing the order of summation, we get that

$$f(x \otimes y) = \sum_{u,v \in E} X_{uv} x_u y_v \qquad \text{for all } x, y \in \mathbb{F}_2^E.$$

Since f is NAE-satisfying, for each triple $\{v_1, v_2, v_3\} \in \mathrm{Triples}(E)$ we must have

$$X_{v_1 v_1} + X_{v_2 v_2} + X_{v_3 v_3} + X_{v_1 v_2} + X_{v_2 v_3} + X_{v_3 v_1} = 1,$$

and hence

$$\sum_{i=1}^{r} \left(a_{v_1}^{(i)} a_{v_1}^{(i)} + a_{v_2}^{(i)} a_{v_2}^{(i)} + a_{v_3}^{(i)} a_{v_3}^{(i)} + a_{v_1}^{(i)} a_{v_2}^{(i)} + a_{v_2}^{(i)} a_{v_3}^{(i)} + a_{v_3}^{(i)} a_{v_1}^{(i)} \right) = 1.$$

By Definition 4.9 we thus have

$$\sum_{i=1}^{r} \mathrm{NAE}_3\left(a_{v_1}^{(i)}, a_{v_2}^{(i)}, a_{v_3}^{(i)}\right) = 1 \qquad \text{for all } \{v_1, v_2, v_3\} \in \mathrm{Triples}(E),$$

and thus $f^{(1)}, \ldots, f^{(r)}$ are jointly NAE-satisfying. $\square$

The following proposition shows that the other direction also holds. It, and the straightforward corollary following it will be used in the soundness analysis of our reduction.

Proposition 4.14. Let E be a tripled set and let r be a positive integer. Suppose $f : \mathbb{F}_2^E \otimes \mathbb{F}_2^E \to \mathbb{F}_2$ is a symmetric bilinear map and $f^{(1)}, \dots, f^{(r)} : \mathbb{F}_2^E \to \mathbb{F}_2$ are jointly NAE-satisfying linear functionals. If (4.2) holds for all $x, y \in \mathbb{F}_2^E$, then f is NAE-satisfying.

Proof. The proof follows the same lines as in Proposition 4.13, and we omit the details. □

Corollary 4.15. *Let E be a nonempty tripled set. There exists an* NAE*-satisfying symmetric bilinear map $f : \mathbb{F}_2^E \otimes \mathbb{F}_2^E \to \mathbb{F}_2^E$ of rank 1.*

Proof. There clearly exists a nonzero linear functional $f^{(1)} : \mathbb{F}_2^E \to \mathbb{F}_2$ that is NAE-satisfying (by itself). We can then let $f : \mathbb{F}_2^E \otimes \mathbb{F}_2^E \to \mathbb{F}_2$ be the rank-1 bilinear map defined by $f(x \otimes y) = f^{(1)}(x) f^{(1)}(y)$ for all $x, y \in \mathbb{F}_2^E$, and apply Proposition 4.14 to show that f is NAE-satisfying. □

In this section we construct our outer PCP, following the steps outlined in **??**. Our reduction starts with the following result:

4.3 The Equation vs. Variable Game

The goal of this section is to show the second step in our outer PCP construction, which transforms a 3-NAE instance into a 2-Prover-1-Round game, and then performs on it smooth parallel repetition with advice. The 2-Prover-1-Round game we will work with is formally defined as follows:

Definition 4.16. Given a positive integer r, nonnegative integers q_1, q_2, a parameter $\beta \in [0, 1]$ and a 3-uniform hypergraph H, let $\mathsf{Game1a}[H, k, q_1, q_2, r, \beta]$ be the 2-Prover-1-Round game where the verifier executes the following:

1. Sample a uniformly random tripled set $E \in \mathscr{E}(H, k)$.
2. Sample $(E', A_{\mathsf{in}}, A_{\mathsf{mid}}) \sim \mathcal{J}_{\mathsf{out}}(E, q_1, q_2, \beta)$. Send $(E, A_{\mathsf{in}}, A_{\mathsf{mid}})$ to Alice and $(E', A_{\mathsf{in}}, A_{\mathsf{mid}})$ to Bob.
3. Alice must return r linear functionals $f^{(1)}, \dots, f^{(r)} : \mathbb{F}_2^E \to \mathbb{F}_2$ that are jointly NAE-satisfying; Bob must return r vectors $g^{(1)}, \dots, g^{(r)} : \mathbb{F}_2^{E'} \to \mathbb{F}_2$. Accept if

$$f^{(i)}\Big|_{\mathbb{F}_2^{E'}} = g^{(i)} \qquad \text{for all } i \in [r].$$

Remark 4.17. We remark that the game $\mathsf{Game1a}[H, k, q_1, q_2, r, \beta]$ can be viewed as the k-fold parallel repetition of $\mathsf{Game1a}[H, 1, q_1, q_2, r, \beta]$, and the game $\mathsf{Game1a}[H, 1, 0, 0, r, 1]$ is the equation versus variable game corresponding to the the result of **??**.

We now analyze the completeness and the soundness of the reduction outlined in Definition 4.16. The following lemma establishes the completeness:

Lemma 4.18 (Game1a completeness)**.** *For any positive integer k, nonnegative integers q_1, q_2, parameter $\beta \in [0, 1]$ and 2-colorable 3-uniform hypergraph H, we have*

$$\mathrm{val}\Big(\mathsf{Game1a}[H, k, q_1, q_2, 1, \beta]\Big) = 1.$$

Proof. By the completeness of the reduction of **??**, we conclude a strategy for the players for the case that $k = 1$: the players simply ignore their advice vectors, and encode an assignment to their set of variables E by a linear function $f\colon \mathbb{F}_2^E \to \mathbb{F}_2$. The players can thus lift this strategy to Ψ by extending their assignment to k-tuples in the obvious way. $\square$

We now move on to the soundness. We start by analyzing the soundness of the basic game $\mathsf{Game1a}[H, 1, 0, 0, r, 1]$.

Lemma 4.19. *Fix a positive integer r and a parameter $\varepsilon \in (0, 1)$. Suppose H is a 3-uniform hypergraph such that in every 2^r-coloring of H, at least ε fraction of the edges are monochromatic. Then we have*

$$\mathrm{val}\Big(\mathsf{Game1a}[H, 1, 0, 0, r, 1]\Big) \leqslant 1 - \frac{\varepsilon}{3}.$$

Proof. Let V be the vertex set of $H^{\sqcup 1}$ (which is isomorphic to H).

We fix an arbitrary strategy for Alice and Bob in the game $\mathsf{Game1a}[H, 1, 0, 0, r, 1]$, and we will prove that the verifier rejects with probability at least $\varepsilon/3$. Since there are no advice vectors in this game, the question for Alice is simply an random edge of $H^{\sqcup 1}$; and as the smoothness parameter is set to 1, the question for Bob will be a single vertex.

For a vertex v, denote the response of Bob as $g_v^{(1)}, \ldots, g_v^{(r)} : \mathbb{F}_2^{\{v\}} \to \mathbb{F}_2$. We may view it as r assignments for the 3-NAE instance (X, C) from **??**, so from the soundness there we conclude that it jointly satisfies at most $1 - \varepsilon$ of the constraints of (X, C). Fix such an edge $E = \{v_1, v_2, v_3\}$ whose constraint is not jointly satisfied by Bob's response, and let $f^{(1)}, \ldots, f^{(r)} : \mathbb{F}_2^E \to \mathbb{F}_2$ be Alice's answer to the question E. Since $f^{(1)}, \ldots, f^{(r)}$ are jointly NAE-satisfying ,there exists a vertex $v \in \{v_1, v_2, v_3\}$ with

$$\left(f^{(1)}\Big|_{\mathbb{F}_2^{\{v\}}}, f^{(2)}\Big|_{\mathbb{F}_2^{\{v\}}}, \ldots, f^{(r)}\Big|_{\mathbb{F}_2^{\{v\}}} \right) \neq \left(g_v^{(1)}, g_v^{(2)}, \ldots, g_v^{(r)} \right). \tag{4.3}$$

Conditioned on the event that Alice receives the questions E, with probability $1/3$ the question for Bob is $\{v\}$, in which case their answers are not consistent due to Equation (4.3). Since this is true for at least ε-fraction of the edges $E \in \mathscr{E}(H, 1)$, the verifier must reject with probability at least $\varepsilon/3$. $\square$

Next, we incorporate the advice as well as constantly many repetitions, and show that the value does not increase by too much:

Lemma 4.20. *Fix a positive integers k, r, nonnegative integers q_1, q_2 and parameters $\beta, \varepsilon \in (0, 1)$ such that $k \geqslant 2^{q_1+q_2}\beta^{-1}$. Suppose H is a 3-uniform hypergraph such that in every 2^r-coloring of H, at least ε fraction of the edges are monochromatic. Then we have*

$$\mathrm{val}\Big(\mathsf{Game1a}[H, k, q_1, q_2, r, \beta]\Big) \leqslant 1 - \frac{\varepsilon}{6}.$$

Proof. We write $q = q_1 + q_2$. As in Remark 4.17, we view $\mathsf{Game1a}[H, k, q_1, q_2, r, \beta]$ as the k-fold parallel repetition of the game $\mathsf{Game1a}[H, 1, q_1, q_2, r, \beta]$. Let $(E, A_{\mathsf{in}}, A_{\mathsf{mid}})$ and $(E', A_{\mathsf{in}}, A_{\mathsf{mid}})$ be the questions for Alice and Bob in the k-fold repetition game, respectively. We write

$$A_{\mathsf{in}} = \big(a^{(1)}, \ldots, a^{(q_1)}\big) \qquad \text{and} \qquad A_{\mathsf{mid}} = \big(a^{(q_1+1)}, \ldots, a^{(q_1+q_2)}\big).$$

We say a triple $\{v_1, v_2, v_3\} \in \mathrm{Triples}(E)$ is "critical" if we have

$$a_{v_1}^{(i)} = a_{v_2}^{(i)} = a_{v_3}^{(i)} = 0 \text{ for all } i \in [q] \qquad \text{and} \qquad \big|E' \cap \{v_1, v_2, v_3\}\big| = 1.$$

Let $I \subseteq [k]$ be the collection of indices $i \in [k]$ such that the i-th triple of E is critical (see Remark 4.3). Conditioned on the event that $I \neq \emptyset$, we pick a uniformly random $i \in I$ and note that the i-th "coordinate" of the parallel repetition is a copy of the equation vs. variable game $\mathsf{Game1a}[H, 1, 0, 0, r, 1]$ and is independent of the remaining coordinates. Therefore, Lemma 4.19 implies that conditioned on $\{I \neq \emptyset\}$, the verifier rejects with probability at least $\varepsilon/3$. Since

$$\mathbb{P}[I \neq \emptyset] = 1 - \left(1 - 2^{-q}\beta\right)^k \geqslant 1 - 2^{-k \cdot 2^q \beta^{-1}} \geqslant 1 - 2^{-1} = \frac{1}{2},$$

we conclude that in the k-fold repetition game $\mathsf{Game1a}[H, k, q_1, q_2, r, \beta]$, the verifier rejects with probability at least $\frac{1}{2} \cdot \varepsilon/3 = \varepsilon/6$. ☐

Finally, we apply the parallel repetition theorem for projection games by [Rao11] to get the soundness of the reduction in Definition 4.16:

Lemma 4.21 (Game1a soundness)**.** *Assume Condition 3.9. Suppose H is a 3-uniform hypergraph such that in every 4^r-coloring of H, at least ε fraction of the edges are monochromatic. Then we have*

$$\mathrm{val}\Big(\mathsf{Game1a}[H, k, q_1, q_2, 2r, \beta]\Big) \leqslant \exp(-k^{1/5}).$$

Proof. Let $k_0 = \lceil 2^{q_1+q_2}\beta^{-1} \rceil$. By Lemma 4.20 we know that

$$\mathrm{val}\Big(\mathsf{Game1a}[H, k_0, q_1, q_2, 2r, \beta]\Big) \leqslant 1 - \frac{\varepsilon}{6}.$$

Since $\mathsf{Game1a}[H, k_0, q_1, q_2, 2r, \beta]$ is a projection game, we may apply [Rao11, Theorem 1.4] to conclude that its $\lfloor k/k_0 \rfloor$-fold parallel repetition has value at most $\exp(-c\varepsilon^2 \lfloor k/k_0 \rfloor)$, for some absolute constant $c \in (0, 1)$. It is then easy to see (via Remark 4.17) that

$$\mathrm{val}\Big(\mathsf{Game1a}[H, k, q_1, q_2, 2r, \beta]\Big) \leqslant \exp(-c\varepsilon^2 \lfloor k/k_0 \rfloor) \leqslant \exp(-k^{-1/5}),$$

where in the last transition we use the relation $k/k_0 \gg \varepsilon^{-2} k^{1/5}$ guaranteed by Condition 3.9. ☐

4.4 From Joint Satisfaction to Bilinear Forms

The goal of this section is to show the third (and last) step in our outer PCP construction, which replaces the responses of Alice and Bob by symmetric bilinear forms. More precisely, we will consider the following 2-Prover-1-Round game defined below. It is identical to $\mathsf{Game1a}[H, k, q_1, q_2, r, \beta]$, except that the provers are accepted to provide a single Bilinear form of low-rank (instead of r linear forms).

Definition 4.22. Given positive integers k, r, nonnegative integers q_1, q_2, a parameter $\beta \in [0, 1)$ and a 3-uniform hypergraph H, let $\mathsf{Game1b}[H, k, q_1, q_2, r, \beta]$ be the 2-Prover-1-Round where the verifier executes the following:

1. Sample a uniformly random tripled set $E \in \mathscr{E}(H, k)$.
2. Sample $(E', A_{\mathsf{in}}, A_{\mathsf{mid}}) \sim \mathcal{J}_{\mathsf{out}}(E, q_1, q_2, \beta)$. Send $(E, A_{\mathsf{in}}, A_{\mathsf{mid}})$ to Alice and $(E', A_{\mathsf{in}}, A_{\mathsf{mid}})$ to Bob.
3. Alice must return a symmetric bilinear form $f : \mathbb{F}_2^E \otimes \mathbb{F}_2^E \to \mathbb{F}_2$ that is NAE-satisfying and has rank at most r; Bob must return a bilinear form $g : \mathbb{F}_2^{E'} \otimes \mathbb{F}_2^{E'} \to \mathbb{F}_2$. Accept if

$$f|_{\mathbb{F}_2^{E'} \otimes \mathbb{F}_2^{E'}} = g.$$

We now analyze the completeness and the soundess of the reduction outlined in Definition 4.22. The following lemma establishes the completeness:

Lemma 4.23 (Game1b completeness)**.** *For any positive integer k, nonnegative integers q_1, q_2, parameter $\beta \in [0,1)$ and 2-colorable 3-uniform hypergraph H, we have*

$$\mathrm{val}\Big(\mathsf{Game1b}[H, k, q_1, q_2, 1, \beta]\Big) = 1.$$

Proof. Using Lemma 4.18, we get that the game $\mathsf{Game1a}[H, k, q_1, q_2, 1, \beta]$ has value 1. We denote by the linear functionals $f(E, A_{\mathsf{in}}, A_{\mathsf{mid}}) \colon \mathbb{F}_2^E \to \mathbb{F}_2$ and $g(E', A_{\mathsf{in}}, A_{\mathsf{mid}}) \colon \mathbb{F}_2^E \to \mathbb{F}_2$ strategies of Alice and Bob, respectively, that achieve value 1 in $\mathsf{Game1a}[H, k, q_1, q_2, 1, \beta]$. Write $f(E, A_{\mathsf{in}}, A_{\mathsf{mid}})(z) = \langle u, z\rangle$ for some $u \in \mathbb{F}_2^E$ and $g(E', A_{\mathsf{in}}, A_{\mathsf{mid}})(z) = \langle u', z\rangle$ for some $u' \in \mathbb{F}_2^{E'}$. The bilinear forms corresponding to $u \otimes u$ and $v \otimes v$ agree on $\mathbb{F}_2^{E'} \otimes \mathbb{F}_2^{E'}$, are rank 1, and the bilinear form of $u \otimes u$ NAE-satisfies all of E. The statement of the lemma follows. □

We now move on to the soundness. The following lemma establishes a connection between a connection between the soundness of $\mathsf{Game1a}$ and $\mathsf{Game1b}$, and more precisely it shows how to derive a strategy for the former from a strategy for the latter. Intuitively, this come from the fact that a rank r bilinear form has at most $O_r(1)$ many decompositions into rank 1 bilinear form, but some care is needed (specifically, the event that the rank of the Alice's bilinear drops in Bob's space is problematic, but thankfully it has low probability).

Lemma 4.24. *For any positive integer k, r, nonnegative integers q_1, q_2, parameter $\beta \in [0,1)$ and 3-uniform hypergraph H, we have*

$$\mathrm{val}\left(\mathsf{Game1a}[H, k, q_1, q_2, 2r, \beta]\right) \geqslant 4^{-r^2}\left(\mathrm{val}\left(\mathsf{Game1b}[H, k, q_1, q_2, r, \beta]\right) - 4^{q_1+q_2} r\beta\right).$$

Proof. We organize the proof into the following three steps.

Step 1: controlling rank drop. We denote

$$s = \mathrm{val}\left(\mathsf{Game1b}[H, k, q_1, q_2, r, \beta]\right).$$

Let $F[\cdot]$ and $G[\cdot]$ be the strategies of Alice and Bob, respectively, under which the verifier accepts with probability s in $\mathsf{Game1b}[H, k, q_1, q_2, r, \beta]$. For each $E \in \mathscr{E}(H, k)$ and $(E', A_{\mathsf{in}}, A_{\mathsf{mid}})$ in the support of $\mathcal{J}_{\mathsf{out}}(E, q_1, q_2, \beta)$, we have bilinear maps

$$F[E, A_{\mathsf{in}}, A_{\mathsf{mid}}] : \mathbb{F}_2^E \otimes \mathbb{F}_2^E \to \mathbb{F}_2 \qquad \text{and} \qquad G[E', A_{\mathsf{in}}, A_{\mathsf{mid}}] : \mathbb{F}_2^{E'} \otimes \mathbb{F}_2^{E'} \to \mathbb{F}_2.$$

A pair of questions $(E, A_{\mathsf{in}}, A_{\mathsf{mid}})$ and $(E', A_{\mathsf{in}}, A_{\mathsf{mid}})$ for Alice and Bob is said to be "good" if

$$F[E, A_{\mathsf{in}}, A_{\mathsf{mid}}]|_{\mathbb{F}_2^{E'} \otimes \mathbb{F}_2^{E'}} = G[E', A_{\mathsf{in}}, A_{\mathsf{mid}}];$$

it is said to be "excellent" if in addition to being good, it also satisfies

$$\mathrm{rank}\big(F[E, A_{\mathsf{in}}, A_{\mathsf{mid}}]\big) = \mathrm{rank}\big(G[E', A_{\mathsf{in}}, A_{\mathsf{mid}}]\big).$$

By definition, a random pair of questions is "good" with probability s. Furthermore, conditioned on each fixed Alice question $(E, A_{\mathsf{in}}, A_{\mathsf{mid}})$, since $\mathrm{rank}\big(F[E, A_{\mathsf{in}}, A_{\mathsf{mid}}]\big) \leqslant r$, we know from Corollary B.3 and Proposition 4.7 that

$$\mathop{\mathbb{P}}_{E'}\left[\mathrm{rank}\Big(F[E, A_{\mathsf{in}}, A_{\mathsf{mid}}]|_{\mathbb{F}_2^{E'} \otimes \mathbb{F}_2^{E'}}\Big) = \mathrm{rank}\Big(F[E, A_{\mathsf{in}}, A_{\mathsf{mid}}]\Big) \,\Big|\, E, A_{\mathsf{in}}, A_{\mathsf{mid}}\right] \geqslant (1 - 4^{q_1+q_2}\beta)^r$$

$$\geqslant 1 - 4^{q_1+q_2} r\beta.$$

Therefore, the probability that a random pair of questions is "good" but not "excellent" is at most $4^{q_1+q_2} r\beta$. Consequently, a random pair of questions is "excellent" with probability at least $s - 4^{q_1+q_2} r\beta$.

Step 2: deriving a strategy for Game1a. Note that in $\mathsf{Game1a}[H, k, q_1, q_2, 2r, \beta]$ the joint distribution of questions to Alice and Bob is the same as in $\mathsf{Game1b}[H, k, q_1, q_2, r, \beta]$. We next derive strategies for the provers in the former game from the tables $F[\cdot]$ and $G[\cdot]$. The strategy for Alice is as follows:

1. Receive the question $(E, A_{\mathsf{in}}, A_{\mathsf{mid}})$, and let $f = F[E, A_{\mathsf{in}}, A_{\mathsf{mid}}]$. Note that $f : \mathbb{F}_2^E \otimes \mathbb{F}_2^E \to \mathbb{F}_2$ is symmetric and NAE-satisfying.
2. Use Lemma B.4 to obtain a tensor decomposition $\left(f^{(1)}, \ldots, f^{(2r)}\right)$ of f.
3. Return the linear functionals $f^{(1)}, \ldots, f^{(2r)} : \mathbb{F}_2^E \to \mathbb{F}_2$. They must be jointly NAE-satisfying, due to Proposition 4.13.

Accordingly, the strategy for Bob is:

1. Receive the question $(E', A_{\mathsf{in}}, A_{\mathsf{mid}})$, and let $g = G[E, A_{\mathsf{in}}, A_{\mathsf{mid}}]$.
2. For each $i \in [2r]$, select a linear functional $g^{(i)}$ from the subspace $\mathrm{image}(T_g) \subseteq (\mathbb{F}_2^{E'})^*$ independently and uniformly at random.
3. Return the linear functionals $g^{(1)}, \ldots, g^{(2r)} : \mathbb{F}_2^{E'} \to \mathbb{F}_2$.

Step 3: analysis of the strategy. We show that the above strategy answers every "excellent" pair of questions consistently with probability at least 4^{-r^2}, and thereby concluding the desired inequality

$$\mathrm{val}\left(\mathsf{Game1a}[H, k, q_1, q_2, 2r, \beta]\right) \geqslant 4^{-r^2}(s - 4^{q_1+q_2} r\beta).$$

Suppose $(E, A_{\mathsf{in}}, A_{\mathsf{mid}})$ and $(E', A_{\mathsf{in}}, A_{\mathsf{mid}})$ are an "excellent" pair of questions for Alice and Bob, and let

$$f = F[E, A_{\mathsf{in}}, A_{\mathsf{mid}}] \qquad \text{and} \qquad g = G[E, A_{\mathsf{in}}, A_{\mathsf{mid}}].$$

Let $\left(f^{(1)}, \ldots, f^{(2r)}\right)$ be an arbitrary tensor decomposition of f. By Definition 4.12, we know that $f^{(i)} \in \mathrm{image}(T_f)$ for each $i \in [2r]$. Furthermore, since $\mathrm{rank}(f) = \mathrm{rank}(g)$ (due to the excellence condition), it follows from Corollary B.2 that

$$f^{(i)}\Big|_{\mathbb{F}_2^{E'}} \in \mathrm{image}(T_g) \qquad \text{for all } i \in [2r].$$

Since the cardinality of $\mathrm{image}(T_g)$ is at most $2^{\mathrm{rank}(g)} \leqslant 2^r$, if $g^{(1)}, \ldots, g^{(2r)}$ are independently and uniformly sampled from $\mathrm{image}(T_g)$, we have

$$\mathbb{P}\left[f^{(i)}\Big|_{\mathbb{F}_2^{E'}} = g^{(i)} \text{ for all } i \in [2r]\right] \geqslant (2^{-r})^{2r} = 4^{-r^2},$$

as desired. □

Using Lemma 4.24, we can now analyze the soundness of $\mathsf{Game1b}$.

Lemma 4.25 (Game1b soundness). *Assume Condition 3.9. Suppose H is a 3-uniform hypergraph such that in every 4^r-coloring of H, at least ε fraction of the edges are monochromatic. Then we have*

$$\operatorname{val}\Big(\mathsf{Game1b}[H, k, q_1, q_2, r, \beta]\Big) \leqslant k^{-1/2}.$$

Proof. Combining Lemmas 4.21 and 4.24, we have

$$\operatorname{val}\Big(\mathsf{Game1b}[H, k, q_1, q_2, r, \beta]\Big) \leqslant 4^{q_1+q_2} r\beta + 4^{r^2} \exp(-k^{1/5}) \leqslant k^{-1/2},$$

where the last transition is due to Condition 3.9. □

5 The Middle PCP

In this section we construct our middle PCP, following the steps outlined in **??**.

5.1 Setup for the Local Tensor Tests

Our first goal is to replace the global rank tests in $\mathsf{Game1b}$ with local tensor checks, and towards this end we make the following definition.

Definition 5.1. Let E be a tripled set. Fix a parameter $\beta \in [0, 1]$ and nonnegative integers q_1, q_2, ℓ such that $\ell \geqslant q_2$. We let $\mathcal{J}_{\mathsf{mid}}(E, q_1, q_2, \beta, \ell)$ denote the output distribution of the following process:

1. Sample $(E', A_{\mathsf{in}}, A_{\mathsf{mid}}) \sim \mathcal{J}_{\mathsf{out}}(E, q_1, q_2, \beta)$.
2. Sample $\ell - q_2$ vectors $a^{(1)}, \dots, a^{(\ell - q_2)} \in \mathbb{F}_2^{E'}$ independently and uniformly at random. Let $A' = \big(a^{(1)}, \dots, a^{(\ell-q_2)}\big)$, and obtain a subspace $U = \operatorname{span}(A_{\mathsf{mid}}) + \operatorname{span}(A') \subseteq \mathbb{F}_2^{E'}$.
3. Output the tuple $(E', A_{\mathsf{in}}, A_{\mathsf{mid}}, U)$.

The role of U will be that, instead of requiring that Alice's response is low-rank globally, we will only require it to be low-rank on $U \otimes U$.

Remark 5.2. Similarly to Remark 4.6, the subspace $U \subseteq \mathbb{F}_2^{E'}$ in Definition 5.1 is viewed also as a subspace of $\mathbb{F}_2^{E}$.

Remark 5.3. Consider a sample $(E', A_{\mathsf{in}}, A_{\mathsf{mid}}, U) \sim \mathcal{J}_{\mathsf{mid}}(E, q_1, q_2, \beta, \ell)$ as defined in Definition 5.1. We note that the distribution of (E', A_{in}, U) does not depend on q_2. This is because after fixing E' and A_{in}, the advice sequence A_{mid} is uniformly random over $(\mathbb{F}_2^{E'})^{q_2}$.

We will need the following two technical propositions. The first one asserts that for any fixing of E, $A_{\mathsf{in}}, A_{\mathsf{mid}}$, the distribution of U over the choice of E' and A' is statistically close to being of a random ℓ-dimensional subspace in $\mathbb{F}_2^E$ containing A_{in}.

Proposition 5.4. Let E be a tripled set. Fix a parameter $\beta \in [0, 1)$ and nonnegative integers q_1, q_2, ℓ such that $|E| \geqslant \ell \geqslant q_2$. Fix arbitrary tuples

$$A_1 \in (\mathbb{F}_2^E)^{q_1} \qquad \text{and} \qquad A_2 \in (\mathbb{F}_2^E)^{q_2} \text{ such that } \dim\operatorname{span}(A_2) = q_2.$$

Sample

$$(E', A_{\mathsf{in}}, A_{\mathsf{mid}}, U) \sim \mathcal{J}_{\mathsf{mid}}(E, q_1, q_2, \beta, \ell)$$

conditioned on the event $\{A_{\mathsf{in}} = A_1\} \cap \{A_{\mathsf{mid}} = A_2\}$. Then the resulting marginal distribution of U has total variation distance at most

$$4^{q_1+\ell}\beta\sqrt{|E|/3} + 2^{\ell-|E|}$$

to the uniform distribution over the set $\left\{U \in \mathsf{Gr}(\mathbb{F}_2^E, \ell) \,\middle|\, \operatorname{span}(A_2) \subseteq U\right\}$.

Proof. Deferred to Appendix C. □

The next proposition asserts that for any variable in E, conditioned on any fixing of A_{in} and U, the probability it is omitted from E is still small (without the conditioning, it is at most β).

Proposition 5.5. Let E be a tripled set. Fix a parameter $\beta \in [0, 1)$ and nonnegative integers q_1, ℓ. Fix an arbitrary tuple $A \in (\mathbb{F}_2^E)^{q_1}$ and an arbitrary subspace $U_0 \in \mathsf{Gr}(\mathbb{F}_2^E, \ell)$. Sample

$$(E', A_{\mathsf{in}}, \texttt{empty}, U) \sim \mathcal{J}_{\mathsf{mid}}(E, q_1, 0, \beta, \ell)$$

conditioned on the event $\{A_{\mathsf{in}} = A\} \cap \{U = U_0\}$. Then for any subset $S \subseteq E$ with $|S| \leqslant 2$, the probability that $S \subseteq E'$ is at least $1 - 4^{q_1+\ell+1}\beta$.

Proof. According to Definition 5.1, when $(E', A_{\mathsf{in}}, A_{\mathsf{mid}}, U)$ is sampled from $\mathcal{J}_{\mathsf{mid}}(E, q_1, 0, \beta, \ell)$, a vector sequence $A' \in (\mathbb{F}_2^{E'})^\ell$ is generated, and then U is taken to be $\operatorname{span}(A')$. Let

$$\mathcal{B} = \left\{\left(a^{(1)}, \ldots, a^{(\ell)}\right) \in (\mathbb{F}_2^E)^\ell \,\middle|\, \operatorname{span}\left(a^{(1)}, \ldots, a^{(\ell)}\right) = U_0\right\}$$

be the collection of bases of U_0. The event $\{U = U_0\}$ is thus equal to the disjoint union of events $\bigcup_{B\in\mathcal{B}}\{A' = B\}$. It suffices to prove that for each $B \in \mathcal{B}$, conditioned on the event $\{A_{\mathsf{in}} = A_0\} \cap \{A' = B\}$, we have $\mathbb{P}\left[S \subseteq E'\right] \geqslant 1 - 4^{q_1+\ell+1}\beta$ for any two-element subset $S \subseteq E$. It is easy to see that the distribution of E' conditioned on $\{A_{\mathsf{in}} = A_0\} \cap \{A' = B\}$ is the same as in the following process (by comparing Definitions 4.5 and 5.1):

- sample $(E', A_{\mathsf{in}}, A_{\mathsf{mid}}) \sim \mathcal{J}_{\mathsf{out}}(E, q_1, \ell, \beta)$ conditioned on $\{A_{\mathsf{in}} = A\} \cap \{A_{\mathsf{mid}} = B\}$.

The desired conclusion thus easily follows from Proposition 4.7. □

For future reference, we introduce here a notation corresponding to the game above with no advice. It will be useful for us notationally in the end to describe our reduction.

Definition 5.6. Let E be a tripled set. Fix a parameter $\beta \in [0, 1)$ and a nonnegative integer ℓ. Given any subspace $U_0 \subseteq \mathbb{F}_2^E$ with dimension at most ℓ, we let $\mathcal{J}_{\mathsf{rev}}(E, U_0, \beta, \ell)$ be the output distribution of the following process:

1. Sample $(E', \texttt{empty}, \texttt{empty}, U) \sim \mathcal{J}_{\mathsf{mid}}(E, 0, 0, \beta, \ell)$ conditioned on the event $\{U = U_0\}$.
2. Output E'.

5.2 The First Step: the Local Tensor Check Game

We now show the first step in our middle PCP construction, which replaces the global rank checks on Alice with local rank checks. More precisely, we consider the 2-Prover-1-Round game presented below. It is similar to $\mathsf{Game1b}$ but with one important distinction. Letting f be Alice's response, instead of requiring $\mathsf{rank}(f)$ is small, we require that $\mathsf{rank}(f|_{U\otimes U})$ is small, where U is (roughly) a random ℓ-dimensional space on Alice's space which is given to Alice as part of her challenge. Additionally, as part of his challenge we give Bob a random subspace U' of dimension $\ell + 1$ that contains U.

Definition 5.7. Given positive integers k, ℓ_2, nonnegative integers q_1, q_2 such that $\ell_2 \geqslant q_2$, a parameter $\beta \in [0,1)$ and a 3-uniform hypergraph H, let $\mathsf{Game2a}[H, k, \ell_2, q_1, q_2, \beta]$ be the 2-Prover-1-Round game where the verifier executes the following:

1. Sample a uniformly random tripled set $E \in \mathscr{E}(H, k)$.
2. Sample $(E', A_{\mathsf{in}}, A_{\mathsf{mid}}, U) \sim \mathcal{J}_{\mathsf{mid}}(E, q_1, q_2, \beta, \ell_2)$. Sample a random vector $x \in \mathbb{F}_2^{E'}$ and let $U' = U + \mathrm{span}(x)$.
3. Send (E, A_{in}, U) to Alice and $(E', A_{\mathsf{in}}, U')$ to Bob.
4. Alice must return an NAE-satisfying symmetric bilinear form $f : \mathbb{F}_2^E \otimes \mathbb{F}_2^E \to \mathbb{F}_2$ such that $\mathrm{rank}(f|_{U\otimes U}) \leqslant 1$; Bob must return a bilinear form $g : \mathbb{F}_2^{E'} \otimes \mathbb{F}_2^{E'} \to \mathbb{F}_2$. Accept if

$$f|_{\mathbb{F}_2^{E'}\otimes\mathbb{F}_2^{E'}} = g.$$

Remark 5.8. In Definition 5.7, the verifier does not use A_{mid} after generating it. Therefore, the distribution of $(E, A_{\mathsf{in}}, U, U')$ does not depend on q_2. Consequently, the value of the game $\mathsf{Game2a}[H, k, \ell_2, q_1, q_2, \beta]$ does not depend on q_2 either.

Remark 5.9. In Definition 5.7, it is easy to see that the distribution of $(E', A_{\mathsf{in}}, A_{\mathsf{mid}}, U')$ is identical to $\mathcal{J}_{\mathsf{mid}}(E, q_1, q_2, \beta, \ell_2 + 1)$.

We now analyze the completeness and the soundness of the reduction outlined in Definition 5.7. The completeness is straightforward and is given in the following lemma.

Lemma 5.10 (Game2a completeness)**.** *For any positive integers k, ℓ_2, nonnegative integer q_1, q_2 such that $k \geqslant \ell_2 \geqslant q_2$, parameter $\beta \in [0,1)$ and 2-colorable 3-uniform hypergraph H, we have*

$$\mathrm{val}\Big(\mathsf{Game2a}[H, k, \ell_2, q_1, q_2, \beta]\Big) = 1.$$

Proof. It is clear that the provers can ignore the subspaces U or U' given to them and pretend that they are in the game $\mathsf{Game1b}[H, k, q_1, 0, 1, \beta]$. Since there is a strategy for $\mathsf{Game1b}[H, k, q_1, 0, 1, \beta]$ with 0 rejection probability due to Lemma 4.23, it follows that the same strategy also passes $\mathsf{Game2a}[H, k, \ell_2, q_1, q_2, \beta]$ with probability 1. $\square$

The soundness analysis of the reduction from Definition 5.7, presented in Section 5.3, is much more involved. Intuitively, if we did not supply Bob with the space U', the analysis would be relatively straightforward. Indeed, in that case, while in principle Alice's response could depend on the space U she receives, in actuality such strategies do not achieve a high value, as they have to be consistent with Bob (whose response does not depend on U at all). Thus, in that case Alice's response can be thought of as not being dependent on U, and hence the tensor test on U is random and faithfully checks whether Alice's response has low-rank globally.

To make out final game 4-to-1 we must supply Bob with the space U', and the soundness analysis in that case becomes more complicated. In this case, there are genuine cheating strategies that Alice and Bob can use, and our argument below analyzes them (using the soundness of the Grassmann test) and shows that they can still be converted to good strategies for $\mathsf{Game1b}$.

5.3 Soundness Analysis

In this subsection, we fix a 3-uniform hypergraph H, positive integers k, ℓ_2, nonnegative integers q_1, q_2 and a parameter $\beta \in (0,1)$. We consider the game $\mathsf{Game2a}[H, k, \ell_2, q_1, q_2, \beta]$ as defined in Definition 5.7. The following notational convention will be used throughout Section 5.3 (but *not* in later parts of the paper).

Notation 5.11. We use $\mathcal{D}_{\mathsf{mid}}$ to denote the joint distribution of $(E, A_{\mathsf{in}}, A_{\mathsf{mid}}, U, E', U')$ in the game $\mathsf{Game2a}[H, k, \ell_2, q_1, q_2, \beta]$, where (E, A_{in}, U) is the question sent to Alice and $(E', A_{\mathsf{in}}, U')$ is the question sent to Bob. We use the symbol $*$ to denote marginalization over the corresponding component. For instance, sampling from the marginal distribution of (A_{in}, U, E') under $\mathcal{D}_{\mathsf{mid}}$ is denoted by $(*, A_{\mathsf{in}}, *, U, E', *) \sim \mathcal{D}_{\mathsf{mid}}$.

Notation 5.12. We will often condition the distribution $\mathcal{D}_{\mathsf{mid}}$ on some subset of the six components $(E, A_{\mathsf{in}}, A_{\mathsf{mid}}, U, E', U')$ and consider the induced joint distribution of the remaining components. For instance, if E and U' are fixed, then the conditional joint distribution of (U, E') is denoted by

$$\mathop{\mathbb{P}}_{(E,*,*,U,E',U')\sim\mathcal{D}_{\mathsf{mid}}}\left[\cdot \,\middle|\, E, U'\right].$$

Notation 5.13. A table $F[\cdot]$ is called an Alice-table if for each possible Alice-question (E, A_{in}, U) under $\mathcal{D}_{\mathsf{mid}}$, the entry $F[E, A_{\mathsf{in}}, U]$ is either $\mathtt{nil}$ or an NAE-satisfying symmetric bilinear form

$$F[E, A_{\mathsf{in}}, U] : \mathbb{F}_2^E \otimes \mathbb{F}_2^E \to \mathbb{F}_2 \qquad \text{such that} \qquad \operatorname{rank}\left(F[E, A_{\mathsf{in}}, U]|_{U\otimes U}\right) \leqslant 1.$$

Similarly, a table $G[\cdot]$ is called a Bob-table if for each possible Bob-question $(E', A_{\mathsf{in}}, U')$ under $\mathcal{D}_{\mathsf{mid}}$, the entry $G[E', A_{\mathsf{in}}, U']$ is either $\mathtt{nil}$ or a bilinear form

$$G[E', A_{\mathsf{in}}, U'] : \mathbb{F}_2^{E'} \otimes \mathbb{F}_2^{E'} \to \mathbb{F}_2.$$

Notation 5.14. Fix an Alice-table $F[\cdot]$ and a Bob-table $G[\cdot]$. For any tuple $(E, A_{\mathsf{in}}, *, U, E', U')$ with nonzero probability under $\mathcal{D}_{\mathsf{mid}}$, we define

$$\mathrm{Vrf}_{F,G}(E, A_{\mathsf{in}}, U, E', U') = \begin{cases} 1, & \text{if } \ F[E, A_{\mathsf{in}}, U]|_{\mathbb{F}_2^{E'}\otimes\mathbb{F}_2^{E'}} = G[E', A_{\mathsf{in}}, U'], \\ 0, & \text{otherwise.} \end{cases}$$

We will often sample $(E, A_{\mathsf{in}}, *, U, E', U')$ from $\mathcal{D}_{\mathsf{mid}}$ conditioned on some subset of its components. For instance, if E and U' are fixed, then we denote the corresponding conditional expectation of $\mathrm{Vrf}_{F,G}(E, A_{\mathsf{in}}, U, E', U')$ by

$$\mathrm{Vrf}_{F,G}(E, *, *, *, U') = \mathop{\mathbb{P}}_{(E,A_{\mathsf{in}},*,U,E',U')\sim\mathcal{D}_{\mathsf{mid}}}\left[F[E, A_{\mathsf{in}}, U]|_{\mathbb{F}_2^{E'}\otimes\mathbb{F}_2^{E'}} = G[E', A_{\mathsf{in}}, U'] \,\middle|\, E, U'\right].$$

When none of the components is fixed, we write $\mathrm{Vrf}_{F,G} = \mathrm{Vrf}_{F,G}(*, *, *, *, *)$ for short.

5.3.1 Preparing for Grassmann Decoding

We first show that if we have Alice-table and Bob-table that achieve a high value in $\mathsf{Game2a}$, then the Alice-table must pass the Grassmann equality test with noticeable probability. More precisely, we have the following lemma:

Lemma 5.15. *Assume Condition 3.9. Let $F[\cdot]$ be an Alice-table and $G[\cdot]$ be a Bob-table. For any $E \in \mathscr{E}(H, k)$ and any $A_{\mathsf{in}} \in (\mathbb{F}_2^E)^{q_1}$, we have*

$$\mathop{\mathbb{P}}_{(U_1,U_2)\sim\mathcal{G}(\mathbb{F}_2^E,\ell_2)}\left[F[E, A_{\mathsf{in}}, U_1] = F[E, A_{\mathsf{in}}, U_2] \neq \mathtt{nil}\right] \geqslant \mathrm{Vrf}_{F,G}(E, A_{\mathsf{in}}, *, *, *)^2 - 2^{-\ell_2}.$$

Proof. We organize the proof into the following three steps:

Step 1: fixing E' and U'. Fix a subset $E' \subseteq E$ with size $|E'| \geqslant \ell_2 + 1$ and a subspace $U' \in \mathsf{Gr}(\mathbb{F}_2^{E'}, \ell_2 + 1)$ such that the tuple $(E, A_{\mathsf{in}}, *, *, E', U')$ has nonzero probability under $\mathcal{D}_{\mathsf{mid}}$. By Definitions 5.1 and 5.7, it is easy to see that sampling $(E, A_{\mathsf{in}}, *, U, E', U') \sim \mathcal{D}_{\mathsf{mid}}$ conditioned on E, A_{in}, E' and U' is equivalent to sampling a uniformly random subspace $U \in \mathsf{Gr}(U', \ell_2)$. Therefore, we have

$$\begin{aligned}\mathrm{Vrf}_{F,G}(E, A_{\mathsf{in}}, *, E', U')^2 &= \mathop{\mathbb{E}}_{(E,A_{\mathsf{in}},*,U,E,U')\sim\mathcal{D}_{\mathsf{mid}}}\left[\mathrm{Vrf}_{F,G}(E, A_{\mathsf{in}}, U, E', U') \,\middle|\, E, A_{\mathsf{in}}, E', U'\right]^2 \\ &= \mathop{\mathbb{E}}_{U_1,U_2\in\mathsf{Gr}(U',\ell_2)}\left[\mathrm{Vrf}_{F,G}(E, A_{\mathsf{in}}, U_1, E', U')\,\mathrm{Vrf}_{F,G}(E, A_{\mathsf{in}}, U_2, E', U')\right].\end{aligned} \tag{5.1}$$

Expanding the right-hand side of Equation (5.1) by definition, we get

$$\begin{aligned}&\mathrm{Vrf}_{F,G}(E, A_{\mathsf{in}}, *, E', U')^2 \\ &= \mathop{\mathbb{P}}_{U_1,U_2\in\mathsf{Gr}(U',\ell_2)}\left[F[E, A_{\mathsf{in}}, U_1]|_{\mathbb{F}_2^{E'\otimes E'}} = G[E', A_{\mathsf{in}}, U'] = F[E, A_{\mathsf{in}}, U_2]|_{\mathbb{F}_2^{E'\otimes E'}}\right] \\ &\leqslant \mathop{\mathbb{P}}_{U_1,U_2\in\mathsf{Gr}(U',\ell_2)}\left[F[E, A_{\mathsf{in}}, U_1]|_{\mathbb{F}_2^{E'\otimes E'}} = F[E, A_{\mathsf{in}}, U_2]|_{\mathbb{F}_2^{E'\otimes E'}}\right].\end{aligned} \tag{5.2}$$

Step 2: fixing only U'. For any fixed $U' \in \mathsf{Gr}(\mathbb{F}_2^{E}, \ell_2 + 1)$ and $U_1, U_2 \in \mathsf{Gr}(U', \ell_2)$ such that $F[E, A_{\mathsf{in}}, U_1] \neq F[E, A_{\mathsf{in}}, U_2]$, one of the following two possibilities holds:

(1) At least one of $F[E, A_{\mathsf{in}}, U_1]$ and $F[E, A_{\mathsf{in}}, U_2]$ is `nil`.

(2) Both of $F[E, A_{\mathsf{in}}, U_1]$ and $F[E, A_{\mathsf{in}}, U_2]$ are bilinear forms $\mathbb{F}_2^E \otimes \mathbb{F}_2^E \to \mathbb{F}_2$, and their difference $f = F[E, A_{\mathsf{in}}, U_1] - F[E, A_{\mathsf{in}}, U_2]$ is nonzero. In this case, there exist two unit vectors $x, y \in \mathbb{F}_2^E$ (i.e. $|\mathrm{supp}(x)| = |\mathrm{supp}(y)| = 1$) such that $f(x \otimes y) = 1$. Therefore, for any subset $E' \subseteq E$ that contains $\mathrm{supp}(x) \cup \mathrm{supp}(y)$, we have $F[E, A_{\mathsf{in}}, U_1]|_{\mathbb{F}_2^{E'\otimes E'}} \neq F[E, A_{\mathsf{in}}, U_2]|_{\mathbb{F}_2^{E'\otimes E'}}$.

In either of the above two cases, it follows from Proposition 5.5 and Remark 5.9 that (note that U', U_1 and U_2 remain fixed in the following inequality)

$$\mathop{\mathbb{P}}_{(E,A_{\mathsf{in}},*,*,E',U')\sim\mathcal{D}_{\mathsf{mid}}}\left[F[E, A_{\mathsf{in}}, U_1]|_{\mathbb{F}_2^{E'\otimes E'}} = F[E, A_{\mathsf{in}}, U_2]|_{\mathbb{F}_2^{E'\otimes E'}} \,\middle|\, E, A_{\mathsf{in}}, U'\right] \leqslant 4^{q_1+\ell_2+2}\beta.$$

Therefore, taking expectation of the inequality Equation (5.2) over the conditional distribution of E' yields (note that U_1 and U_2 are no longer fixed in the following inequality)

$$\begin{aligned}&\mathop{\mathbb{E}}_{(E,A_{\mathsf{in}},*,*,E',U')\sim\mathcal{D}_{\mathsf{mid}}}\left[\mathrm{Vrf}_{F,G}(E, A_{\mathsf{in}}, *, E', U')^2 \,\middle|\, E, A_{\mathsf{in}}, U'\right] \\ &\leqslant \mathop{\mathbb{P}}_{U_1,U_2\in\mathsf{Gr}(U',\ell_2)}\left[F[E, A_{\mathsf{in}}, U_1] = F[E, A_{\mathsf{in}}, U_2] \neq \texttt{nil}\right] + 4^{q_1+\ell_2+2}\beta.\end{aligned} \tag{5.3}$$

Step 3: both E' and U' unfixed. By Proposition 5.4 and Remark 5.9, we know that conditioned on the fixed E and A_{in}, the distribution of U' in the sample $(E, A_{\mathsf{in}}, *, *, *, U') \sim \mathcal{D}_{\mathsf{mid}}$ is $\left(4^{q_1+\ell_2+1}\beta\sqrt{k} + 2^{\ell_2+1-3k}\right)$-close to uniform over $\mathsf{Gr}(\mathbb{F}_2^E, \ell_2 + 1)$. Therefore, we have

$$\begin{aligned}&\mathrm{Vrf}_{F,G}(E, A_{\mathsf{in}}, *, *, *)^2 \\ &= \mathop{\mathbb{E}}_{(E,A_{\mathsf{in}},*,*,E',U')\sim\mathcal{D}_{\mathsf{mid}}}\left[\mathrm{Vrf}_{F,G}(E, A_{\mathsf{in}}, *, E', U') \,\middle|\, E, A_{\mathsf{in}}\right]^2 && \text{(by definition)}\end{aligned}$$

$$\leqslant \mathop{\mathbb{E}}_{(E,A_{\mathsf{in}},*,*,E',U')\sim\mathcal{D}_{\mathsf{mid}}}\left[\mathrm{Vrf}_{F,G}(E,A_{\mathsf{in}},*,E',U')^2 \,\middle|\, E,A_{\mathsf{in}}\right] \qquad \text{(by Cauchy-Schwarz)}$$

$$\leqslant \mathop{\mathbb{E}}_{U'\in\mathsf{Gr}(\mathbb{F}_2^E,\ell_2+1)}\left[\mathop{\mathbb{E}}_{(E,A_{\mathsf{in}},*,*,E',U')\sim\mathcal{D}_{\mathsf{mid}}}\left[\mathrm{Vrf}_{F,G}(E,A_{\mathsf{in}},*,E',U')^2 \,\middle|\, E,A_{\mathsf{in}},U'\right]\right] + 4^{q_1+\ell_2+1}\beta\sqrt{k}+2^{\ell_2+1-3k}.$$

We then apply Equation (5.3) to the above and get

$$\mathrm{Vrf}_{F,G}(E,A_{\mathsf{in}},*,*,*)^2 \leqslant \mathop{\mathbb{E}}_{U'\in\mathsf{Gr}(\mathbb{F}_2^E,\ell_2+1)}\left[\mathop{\mathbb{P}}_{U_1,U_2\in\mathsf{Gr}(U',\ell_2)}\left[F[E,A_{\mathsf{in}},U_1]=F[E,A_{\mathsf{in}},U_2]\neq\mathtt{nil}\right]\right] + 4^{q_1+\ell_2+1}\beta\left(4+\sqrt{k}\right)+2^{\ell_2+1-3k}. \tag{5.4}$$

For any fixed $U'\in\mathsf{Gr}(\mathbb{F}_2^E,\ell_2+1)$, two independent samples $U_1,U_2\in\mathsf{Gr}(U',\ell_2)$ are equal to each other with probability $(2^{\ell_2+1}-1)^{-1}$. Furthermore, if $U'\in\mathsf{Gr}(\mathbb{F}_2^E,\ell_2+1)$ is sampled uniformly at random and $U_1,U_2\in\mathsf{Gr}(U',\ell_2)$ are sampled uniformly conditioned on $U_1\neq U_2$, the distribution of (U_1,U_2) is uniform over the edges of the Grassmann graph $\mathsf{Gr}(\mathbb{F}_2^E,\ell_2)$. Therefore, Equation (5.4) implies

$$\begin{aligned}\mathrm{Vrf}_{F,G}(E,A_{\mathsf{in}},*,*,*)^2 &\leqslant \mathop{\mathbb{P}}_{(U_1,U_2)\sim\mathcal{G}(\mathbb{F}_2^E,\ell_2)}\left[F[E,A_{\mathsf{in}},U_1]=F[E,A_{\mathsf{in}},U_2]\neq\mathtt{nil}\right]+\frac{1}{2^{\ell_2+1}-1}\\ &\quad+4^{q_1+\ell_2+1}\beta\left(4+\sqrt{k}\right)+2^{\ell_2+1-3k}\\ &\leqslant \mathop{\mathbb{P}}_{(U_1,U_2)\sim\mathcal{G}(\mathbb{F}_2^E,\ell_2)}\left[F[E,A_{\mathsf{in}},U_1]=F[E,A_{\mathsf{in}},U_2]\neq\mathtt{nil}\right]+2^{-\ell_2},\end{aligned}$$

where we used Condition 3.9 in the last transition. □

5.3.2 Invoking the Grassmann Decoding Theorem

Knowing that Alice's table must pass the Grassmann equality test with noticeable probability, we next apply the corresponding decoding theorem to conclude that the given Alice-table must correspond to a low-rank bilinear form.

Lemma 5.16. *Assume Condition 3.9. Let $F[\cdot]$ be an Alice-table and $G[\cdot]$ be a Bob-table. For any $E\in\mathscr{E}(H,k)$ and any $A_{\mathsf{in}}\in(\mathbb{F}_2^E)^{q_1}$ such that $\mathrm{Vrf}_{F,G}(E,A_{\mathsf{in}},*,*,*)\geqslant 2^{-q_2/3}$, the following holds: with probability at least $4^{-\ell_2^2}$ over $(E,A_{\mathsf{in}},A,*,*,*)\sim\mathcal{D}_{\mathsf{mid}}$ conditioned on E and A_{in}, there exists an* NAE*-satisfying symmetric bilinear map $f_A:\mathbb{F}_2^E\otimes\mathbb{F}_2^E\to\mathbb{F}_2$ of rank at most $2q_2+1$ satisfying*

$$\mathop{\mathbb{P}}_{(E,A_{\mathsf{in}},A,U,*,*)\sim\mathcal{D}_{\mathsf{mid}}}\left[F[E,A_{\mathsf{in}},U]=f_A \,\middle|\, E,A_{\mathsf{in}},A\right]\geqslant 2^{-\ell_2^2}.$$

Proof. We organize the proof into the following three steps.

Step 1: Grassmann decoding. Since $\mathrm{Vrf}_{F,G}(E,A_{\mathsf{in}},*,*,*)\geqslant 2^{-q_2/3}$, Lemma 5.15 implies

$$\mathop{\mathbb{P}}_{(U_1,U_2)\sim\mathcal{G}(\mathbb{F}_2^E,\ell_2)}\left[F[E,A_{\mathsf{in}},U_1]=F[E,A_{\mathsf{in}},U_2]\neq\mathtt{nil}\right]\geqslant 2^{-2q_2/3}-2^{-\ell_2}\geqslant 2^{-q_2+4},$$

where we used Condition 3.9 in the last transition. Apply Theorem 3.3 to the table $F[E,A_{\mathsf{in}},\cdot]$ yields[1] a collection of subspaces $\mathcal{Q}\subseteq\mathsf{Gr}(\mathbb{F}_2^E,q_2)$ such that:

[1]We apply Theorem 3.3 with Σ equal to the collection of NAE-satisfying symmetric bilinear maps from $\mathbb{F}_2^E\otimes\mathbb{F}_2^E$ to $\mathbb{F}_2$.

(1) $|\mathcal{Q}| \geqslant 2^{-q_2-\ell_2^2}\left|\mathsf{Gr}(\mathbb{F}_2^E, q_2)\right|$;

(2) for each $Q \in \mathcal{Q}$, there exists a subspace $W_Q \in \mathsf{Gr}(\mathbb{F}_2^E, 3k - q_2)$ and an NAE-satisfying symmetric bilinear map $h_Q : \mathbb{F}_2^E \otimes \mathbb{F}_2^E \to \mathbb{F}_2$ satisfying

$$\mathop{\mathbb{P}}_{U\in\mathsf{Gr}(\mathbb{F}_2^E,\ell_2)}\left[F[E, A_{\mathsf{in}}, U] = h_Q \,\middle|\, Q \subseteq U \subseteq W_Q\right] \geqslant 2^{-Cq_2^2}, \tag{5.5}$$

where C is an absolute constant. For any $U \in \mathsf{Gr}(\mathbb{F}_2^E, \ell_2)$ such that $F[E, A_{\mathsf{in}}, U] = h_Q$, by definition (see Notation 5.13) we must have $\mathrm{rank}(h_Q|_{U\otimes U}) \leqslant 1$. It thus follows from Equation (5.5) that

$$\mathop{\mathbb{P}}_{U\in\mathsf{Gr}(\mathbb{F}_2^E,\ell_2)}\left[\mathrm{rank}(h_Q|_{U\otimes U}) \leqslant 1 \,\middle|\, Q \subseteq U \subseteq W_Q\right] \geqslant 2^{-Cq_2^2} > 2^{4-(\ell_2-q_2)/2},$$

where we used Condition 3.9 in the last transition. By Lemma B.5, we must have $\mathrm{rank}(h_Q) \leqslant 2q_2 + 1$.

For each $Q \in \mathcal{Q}$, since W_Q has codimension q_2 with respect to $\mathbb{F}_2^E$, we have

$$\mathop{\mathbb{P}}_{U\in\mathsf{Gr}(\mathbb{F}_2^E,\ell_2)}\left[U \subseteq W_Q \,\middle|\, Q \subseteq U\right] = \begin{bmatrix} 3k - 2q_2 \\ \ell_2 - q_2 \end{bmatrix}_2 \begin{bmatrix} 3k - q_2 \\ \ell_2 - q_2 \end{bmatrix}_2^{-1} \geqslant 2^{-\ell_2(q_2+1)}. \tag{5.6}$$

Combining Equation (5.5) and Equation (5.6) yields

$$\mathop{\mathbb{P}}_{U\in\mathsf{Gr}(\mathbb{F}_2^E,\ell_2)}\left[F[E, A_{\mathsf{in}}, U] = h_Q \,\middle|\, Q \subseteq U\right] \geqslant 2^{-Cq_2^2-\ell_2(q_2+1)} \geqslant 2^{1-\ell_2^2}, \tag{5.7}$$

where we used Condition 3.9 in the last transition.

Step 2: good Q translates to good A. Now for any $A \in (\mathbb{F}_2^E)^{q_2}$, we let

$$f_A = \begin{cases} h_{\mathrm{span}(A)}, & \text{if } \dim\mathrm{span}(A) = q_2 \text{ and } \mathrm{span}(A) \in \mathcal{Q}, \\ 0, & \text{otherwise.} \end{cases}$$

If $\dim\mathrm{span}(A) = q_2$ and $\mathrm{span}(A) \in \mathcal{Q}$, then by Equation (5.7) we have

$$\mathop{\mathbb{P}}_{U\in\mathsf{Gr}(\mathbb{F}_2^E,\ell_2)}\left[F[E, A_{\mathsf{in}}, U] = f_A \,\middle|\, \mathrm{span}(A) \subseteq U\right] \geqslant 2^{1-\ell_2^2}. \tag{5.8}$$

By Proposition 5.4, we know that conditioned on the fixed E, A_{in} and $A_{\mathsf{mid}} = A$, the distribution of U in the sample $(E, A_{\mathsf{in}}, A, U, *, *) \sim \mathcal{D}_{\mathsf{mid}}$ is $\left(4^{q_1+q_2+\ell_2}\beta\sqrt{k} + 2^{\ell_2-3k}\right)$-close to uniform over the set $\left\{U \in \mathsf{Gr}(\mathbb{F}_2^E, \ell_2) \,\middle|\, \mathrm{span}(A) \subseteq U\right\}$. Therefore, the inequality Equation (5.8) implies

$$\mathop{\mathbb{P}}_{(E,A_{\mathsf{in}},A,U,*,*)\sim\mathcal{D}_{\mathsf{mid}}}\left[F[E, A_{\mathsf{in}}, U] = f_A \,\middle|\, E, A_{\mathsf{in}}, A\right] \geqslant 2^{1-\ell_2^2} - 4^{q_1+q_2+\ell_2}\beta\sqrt{k} - 2^{\ell_2-3k} \geqslant 2^{-\ell_2^2},$$

where we used Condition 3.9 in the last transition.

Step 3: abundance of good A's. It remains to show that

$$\mathop{\mathbb{P}}_{(E,A_{\mathsf{in}},A,*,*,*)\sim\mathcal{D}_{\mathsf{mid}}}\left[\dim\operatorname{span}(A)=q_2\text{ and }\operatorname{span}(A)\in\mathcal{Q}\,\middle|\,E,A_{\mathsf{in}}\right]\geqslant 4^{-\ell_2^2}.$$

By Proposition 4.8, it suffices to show

$$\mathop{\mathbb{P}}_{A\in(\mathbb{F}_2^E)^{q_2}}\left[\dim\operatorname{span}(A)=q_2\text{ and }\operatorname{span}(A)\in\mathcal{Q}\right]\geqslant 4^{-\ell_2^2}+4^{q_1+q_2}\beta\sqrt{k}.$$

But the left-hand side of the above display is clearly equal to

$$\mathop{\mathbb{P}}_{A\in(\mathbb{F}_2^E)^{q_2}}\left[\dim\operatorname{span}(A)=q_2\right]\cdot\frac{|\mathcal{Q}|}{\left|\mathsf{Gr}(\mathbb{F}_2^E,q_2)\right|}\geqslant\frac{1}{4}\cdot 2^{-q_2-\ell_2^2}\geqslant 4^{-\ell_2^2}+4^{q_1+q_2}\beta\sqrt{k},$$

where we used Condition 3.9 in the last transition. □

5.3.3 Deriving Strategy for the Outer PCP

We now use the structural result regarding Alice's strategy to complete the soundness analysis of $\mathsf{Game2a}$. Before we do so, fixing Alice-table and Bob-table, we prune Alice-table and only keep the responses on questions which have noticeable success probability.

Definition 5.17. Let $\eta\in(0,1)$ be threshold parameter. Given an Alice-table $F[\cdot]$ and a Bob-table $G[\cdot]$ for $\mathsf{Game2a}[H,k,\ell_2,q_1,q_2,\beta]$, consider the Alice-table $F'[\cdot]$ defined by

$$F'[E,A_{\mathsf{in}},U]=\begin{cases}F[E,A_{\mathsf{in}},U], & \text{if }\operatorname{Vrf}_{F,G}(E,A_{\mathsf{in}},U,*,*)\geqslant\eta,\\ \mathtt{nil}, & \text{otherwise.}\end{cases}$$

The pair of tables (F',G) is called the *η-pruning* of (F,G).

The following straightforward proposition shows that η-pruning drops the success probability of Alice and Bob by at most η.

Proposition 5.18. Let $\eta\in(0,1)$ be threshold parameter. Given an Alice-table $F[\cdot]$ and a Bob-table $G[\cdot]$ for $\mathsf{Game2a}[H,k,\ell_2,q_1,q_2,\beta]$, let (F',G) be the η-pruning of (F,G). Then

$$\operatorname{Vrf}_{F',G}\geqslant\operatorname{Vrf}_{F,G}-\eta.$$

Proof. For any $E\in\mathscr{E}(H,k)$, any $A_{\mathsf{in}}\in(\mathbb{F}_2^E)^{q_1}$ and any $U\in\mathsf{Gr}(\mathbb{F}_2^E,\ell_2)$, from Definition 5.17 it is easy to see that

$$\operatorname{Vrf}_{F',G}(E,A_{\mathsf{in}},U,*,*)\geqslant\operatorname{Vrf}_{F,G}(E,A_{\mathsf{in}},U,*,*)-\eta.$$

Taking expectation over $(E,A_{\mathsf{in}},*,U,*,*)\sim\mathcal{D}_{\mathsf{mid}}$ then yields the desired conclusion. □

The following lemma shows the soundness of $\mathsf{Game2a}$.

Lemma 5.19 (Game2a soundness)**.** *Assume Condition 3.9. Suppose that in every 4^r-coloring of H, at least ε fraction of the edges are monochromatic. Then we have*

$$\operatorname{val}\Big(\mathsf{Game2a}[H,k,\ell_2,q_1,q_2,\beta]\Big)\leqslant 2^{-q_2/4}.$$

Proof. Assume on the contrary that the value of $\mathsf{Game2a}[H,k,\ell_2,q_1,q_2,\beta]$ is larger than $2^{-q_2/4}$. Then by definition, there are an Alice-table $F[\cdot]$ and a Bob-table $G[\cdot]$ for $\mathsf{Game2a}[H,k,\ell_2,q_1,q_2,\beta]$ such that $\operatorname{Vrf}_{F,G}>2^{-q_2/4}$. Let (F',G) be the $(2^{1-q_2/3})$-pruning of (F,G). By Proposition 5.18 we have

$$\operatorname{Vrf}_{F',G}\geqslant 2^{-q_2/4}-2^{1-q_2/3}\geqslant 2^{1-q_2/3},$$

where we used $q_2\geqslant 24$ (as guaranteed by Condition 3.9).

Step 1: strategy for Game1b. For each $E \in \mathscr{E}(H,k)$ and each possible sample $(E', A_{\mathsf{in}}, A_{\mathsf{mid}})$ from $\mathcal{J}_{\mathsf{out}}(E, q_1, q_2, \beta)$, let $\mathcal{F}(E, A_{\mathsf{in}}, A_{\mathsf{mid}})$ be the collection of all NAE-satisfying symmetric bilinear maps $f : \mathbb{F}_2^E \otimes \mathbb{F}_2^E \to \mathbb{F}_2$ such that $\mathrm{rank}(f) \leqslant 2q_2 + 1$ and

$$\mathop{\mathbb{P}}_{(E,A_{\mathsf{in}},A_{\mathsf{mid}},U,*,*)\sim\mathcal{D}_{\mathsf{mid}}}\left[F'[E, A_{\mathsf{in}}, U] = f \,\middle|\, E, A_{\mathsf{in}}, A_{\mathsf{mid}}\right] \geqslant 2^{-\ell_2^2}.$$

We now derive strategies for the provers in $\mathsf{Game1b}[H, k, q_1, q_2, r, \beta]$ from the tables $F'[\cdot]$ and $G[\cdot]$. The strategy for Alice is as follows:

1. Receive the question $(E, A_{\mathsf{in}}, A_{\mathsf{mid}})$.
2. If $\mathcal{F}[E, A_{\mathsf{in}}, A_{\mathsf{mid}}]$ is empty, return an arbitrary NAE-satisfying symmetric bilinear map $f : \mathbb{F}_2^E \otimes \mathbb{F}_2^E \to \mathbb{F}_2$ of rank 1 (see Corollary 4.15).
3. If $\mathcal{F}[E, A_{\mathsf{in}}, A_{\mathsf{mid}}]$ is nonempty, return an arbitrary NAE-satisfying symmetric bilinear map $f \in \mathcal{F}[E, A_{\mathsf{in}}, A_{\mathsf{mid}}]$. By the definition of $\mathcal{F}[E, A_{\mathsf{in}}, A_{\mathsf{mid}}]$, we have $\mathrm{rank}(f) \leqslant 2q_2 + 1 = r$, where the last transition is due to Condition 3.9.

The strategy for Bob is:

1. Receive the question $(E', A_{\mathsf{in}}, A_{\mathsf{mid}})$.
2. Conditioned on E', A_{in} and A_{mid}, sample $(*, A_{\mathsf{in}}, A_{\mathsf{mid}}, *, E', U') \sim \mathcal{D}_{\mathsf{mid}}$.
3. Return the bilinear map $G[E', A_{\mathsf{in}}, U']$ (note that by definition, the table $G[\cdot]$ does not contain `nil` entries).

Step 2: analysis of the strategy. Since

$$\mathop{\mathbb{E}}_{(E,A_{\mathsf{in}},*,*,*,*)\sim\mathcal{D}_{\mathsf{mid}}}\left[\mathrm{Vrf}_{F',G}(E, A_{\mathsf{in}}, *, *, *)\right] = \mathrm{Vrf}_{F',G} \geqslant 2^{1-q_2/3},$$

we have[2]

$$\mathop{\mathbb{P}}_{(E,A_{\mathsf{in}},*,*,*,*)\sim\mathcal{D}_{\mathsf{mid}}}\left[\mathrm{Vrf}_{F',G}(E, A_{\mathsf{in}}, *, *, *) \geqslant 2^{-q_2/3}\right] \geqslant 2^{-q_2/3}. \tag{5.9}$$

For any $E \in \mathscr{E}(H,k)$ and $A_{\mathsf{in}} \in (\mathbb{F}_2^E)^{q_1}$ such that $\mathrm{Vrf}_{F',G}(E, A_{\mathsf{in}}, *, *, *, *) \geqslant 2^{-q_2/3}$, we know from Lemma 5.16 that

$$\mathop{\mathbb{P}}_{(E,A_{\mathsf{in}},A_{\mathsf{mid}},*,*,*)\sim\mathcal{D}_{\mathsf{mid}}}\left[\mathcal{F}[E, A_{\mathsf{in}}, A_{\mathsf{mid}}] \neq \emptyset \,\middle|\, E, A_{\mathsf{in}}\right] \geqslant 4^{-\ell_2^2}.$$

Combining with Equation (5.9) yields

$$\mathop{\mathbb{P}}_{(E,A_{\mathsf{in}},A_{\mathsf{mid}},*,*,*)\sim\mathcal{D}_{\mathsf{mid}}}\left[\mathcal{F}[E, A_{\mathsf{in}}, A_{\mathsf{mid}}] \neq \emptyset\right] \geqslant 2^{-2\ell_2^2 - q_2/3}. \tag{5.10}$$

Given a tuple $(E, A_{\mathsf{in}}, A_{\mathsf{mid}})$ and an element $f \in \mathcal{F}[E, A_{\mathsf{in}}, A_{\mathsf{mid}}]$, we have

$$\begin{aligned}
&\mathop{\mathbb{P}}_{(E,A_{\mathsf{in}},A_{\mathsf{mid}},*,E',U')\sim\mathcal{D}_{\mathsf{mid}}}\left[G[E', A_{\mathsf{in}}, U'] = f|_{\mathbb{F}_2^{E'}\otimes\mathbb{F}_2^{E'}} \,\middle|\, E, A_{\mathsf{in}}, A_{\mathsf{mid}}\right] \\
\geqslant &\mathop{\mathbb{P}}_{(E,A_{\mathsf{in}},A_{\mathsf{mid}},U,E',U')\sim\mathcal{D}_{\mathsf{mid}}}\left[F'[E, A_{\mathsf{in}}, U] = f \text{ and } \mathrm{Vrf}_{F',G}(E, A_{\mathsf{in}}, U, E', U') = 1 \,\middle|\, E, A_{\mathsf{in}}, A_{\mathsf{mid}}\right]
\end{aligned}$$

[2] Here we use the fact that $\mathrm{Vrf}_{F',G}(E, A_{\mathsf{in}}, *, *, *)$ always lies in $[0, 1]$.

$$= \mathop{\mathbb{E}}_{(E,A_{\mathsf{in}},A_{\mathsf{mid}},U,*,*)\sim\mathcal{D}_{\mathsf{mid}}} \left[\mathbb{1}\left[F'[E,A_{\mathsf{in}},U]=f\right] \cdot \mathrm{Vrf}_{F',G}(E,A_{\mathsf{in}},U,*,*) \,\middle|\, E,A_{\mathsf{in}},A_{\mathsf{mid}} \right]$$

$$\geqslant 2^{1-q_2/3} \cdot \mathop{\mathbb{P}}_{(E,A_{\mathsf{in}},A_{\mathsf{mid}},U,*,*)\sim\mathcal{D}_{\mathsf{mid}}} \left[F'[E,A_{\mathsf{in}},U]=f \,\middle|\, E,A_{\mathsf{in}},A_{\mathsf{mid}} \right]$$

(because (F',G) is the $(2^{1-q_2/3})$-pruning of (F,G))

$$\geqslant 2^{1-\ell_2^2-q_2/3}.$$

(since $f \in \mathcal{F}[E,A_{\mathsf{in}},A_{\mathsf{mid}}]$)

Importantly, conditioned on E, A_{in}, A_{mid} and E', the distribution of U' in a conditional sample $(E,A_{\mathsf{in}},A_{\mathsf{mid}},*,E',U')\sim\mathcal{D}_{\mathsf{mid}}$ does not depend on E. Therefore, the above display implies that under the strategies presented in Step 1 of the proof, for any tuple $(E,A_{\mathsf{in}},A_{\mathsf{mid}})$ with any element $f\in\mathcal{F}[E,A_{\mathsf{in}},A_{\mathsf{mid}}]$, the following expectation is at least $2^{1-\ell_2^2-q_2/3}$:

$$\mathop{\mathbb{E}}_{(E,A_{\mathsf{in}},A_{\mathsf{mid}},*,E',*)\sim\mathcal{D}_{\mathsf{mid}}} \left[\mathop{\mathbb{P}}_{\text{Bob}} \left[\text{Bob returns } f|_{\mathbb{F}_2^{E'}\otimes\mathbb{F}_2^{E'}} \text{ on question } (E',A_{\mathsf{in}},A_{\mathsf{mid}}) \right] \,\middle|\, E,A_{\mathsf{in}},A_{\mathsf{mid}} \right].$$

Combining with Equation (5.9), this implies that[3]

$$\mathop{\mathbb{E}}_{(E,A_{\mathsf{in}},A_{\mathsf{mid}},*,E',*)\sim\mathcal{D}_{\mathsf{mid}}} \left[\mathop{\mathbb{P}}_{\text{Bob}} \left[\text{Alice}(E,A_{\mathsf{in}},A_{\mathsf{mid}})|_{\mathbb{F}_2^{E'}\otimes\mathbb{F}_2^{E'}} = \text{Bob}(E',A_{\mathsf{in}},A_{\mathsf{mid}}) \right] \right]$$

is at least

$$2^{-2\ell_2^2-q_2/3}\cdot 2^{1-\ell_2^2-q_2/3} > k^{-1/2},$$

where we used Condition 3.9 in the last transition. Since the distribution of $(E,A_{\mathsf{in}},A_{\mathsf{mid}},E')$ in $\mathsf{Game1b}[H,k,q_1,q_2,r,\beta]$ is identical to the distribution of the sample $(E,A_{\mathsf{in}},A_{\mathsf{mid}},*,E',*)\sim\mathcal{D}_{\mathsf{mid}}$, it follows that our prescribed strategies pass $\mathsf{Game1b}[H,k,q_1,q_2,r,\beta]$ with probability

$$\mathrm{val}\Big(\mathsf{Game1b}[H,k,q_1,q_2,r,\beta]\Big) > k^{-1/2},$$

contradicting Lemma 4.25. □

5.4 The Second Step: Applying Lazy Parallel Repetition

We now show the second step in our middle PCP construction. More precisely, we consider the following 2-Prover-1-Round game, in which Alice and Bob get t-blocks of challenges, one of which is a true challenge from $\mathsf{Game2a}$, whereas on each one of the rest of the blocks they receive the same randomly sampled challenge of Alice from $\mathsf{Game2a}$. Additionally, on each block the players receive a list of randomly chosen vectors.

Definition 5.20. Given positive integers k,ℓ_2,t, nonnegative integers q_0,q_1, a parameter $\beta\in[0,1)$ and a 3-uniform hypergraph H, let $\mathsf{Game2b}[H,k,\ell_2,t,q_0,q_1,\beta]$ be a 2-Prover 1-Round game where the verifier executes the following:

1. For each $j\in[t]$, sample a uniformly random tripled set $E_j\in\mathscr{E}(H,k)$.
2. Sample a uniformly random index $c\in[t]$. Sample $(E'_c,A_c,\mathtt{empty},U_c)\sim\mathcal{J}_{\mathsf{mid}}(E_c,q_1,0,\beta,\ell_2)$. Sample a random vector $x\in\mathbb{F}_2^{E'_c}$ and let $U'_c=U_c+\mathrm{span}(x)$.
3. For each $j\in[t]\setminus\{c\}$, do the following:

[3] Here we use Alice$(E,A_{\mathsf{in}},A_{\mathsf{mid}})$ to denote Alice's answer to the question $(E,A_{\mathsf{in}},A_{\mathsf{mid}})$, and Bob$(E',A_{\mathsf{in}},A_{\mathsf{mid}})$ denotes Bob's answer to $(E',A_{\mathsf{in}},A_{\mathsf{mid}})$.

- Sample a uniformly random advice sequence $A_j = \left(a^{(j,1)}, \ldots, a^{(j,q_0)}\right) \in \left(\mathbb{F}_2^{E_j} \otimes \mathbb{F}_2^{E_j}\right)^{q_0}$, i.e. $a^{(j,i)}$ is a uniformly random vector in $\mathbb{F}_2^{E_j} \otimes \mathbb{F}_2^{E_j}$ for each $i \in [q_0]$.
- Sample $(E_j^{\text{ghost}}, \texttt{empty}, \texttt{empty}, U_j) \sim \mathcal{J}_{\mathsf{mid}}(E_j, 0, 0, \beta, \ell_2)$.
- Let $E_j' = E_j$ and $U_j' = U_j$.

4. Denote sequences

$$\begin{aligned} \mathcal{E} &= (E_1, \ldots, E_t), & \mathcal{U} &= (U_1, \ldots, U_t), & \mathcal{A} &= (A_1, \ldots, A_t), \quad \text{and} \\ \mathcal{E}' &= (E_1', \ldots, E_t'), & \mathcal{U}' &= (U_1', \ldots, U_t'). \end{aligned}$$

5. Send $(\mathcal{E}, \mathcal{A}, \mathcal{U})$ to Alice and $(\mathcal{E}', \mathcal{A}, \mathcal{U}')$ to Bob.

6. Alice must return NAE-satisfying symmetric bilinear forms

$$f_j : \mathbb{F}_2^{E_j \otimes E_j} \to \mathbb{F}_2 \text{ such that } \operatorname{rank}(f_j|_{U_j \otimes U_j}) \leqslant 1 \qquad \text{for each } j \in [t];$$

Bob must return bilinear forms

$$g_j : \mathbb{F}_2^{E_j'} \otimes \mathbb{F}_2^{E_j'} \to \mathbb{F}_2 \qquad \text{for each } j \in [t].$$

Accept if g_j is the restriction of f_j to $\mathbb{F}_2^{E_j'} \otimes \mathbb{F}_2^{E_j'}$ for each $j \in [t]$.

Ignoring the lists A_j for a moment, the game we have above is the lazy parallel repetition version of $\mathsf{Game2a}$, and the main point in it is that Alice does not know which one of the t-blocks the real challenge lies on (Bob certainly knows, because on that block his question is shorter). Lazy parallel repetition is well-known to preserve completeness and soundness (but not improve them), as well as to improve the "smoothness" of a game. We include the lists A_j to enable the advice feature in our final PCP, and we show that they do not hurt the soundness of the construction by too much (they have no effect on other features of the construction).

The following lemma establishes the completeness of the reduction in Definition 5.20.

Lemma 5.21 (Game2b completeness)**.** *For any positive integers k, ℓ_2, t such that $k \geqslant \ell_2$, nonnegative integers q_0, q_1, parameter $\beta \in [0,1)$ and 2-colorable 3-uniform hypergraph H, we have*

$$\operatorname{val}\Big(\mathsf{Game2b}[H, k, \ell_2, t, q_1, \beta]\Big) = 1.$$

Proof. We fix strategies from Lemma 5.10 that achieve value 1. Alice ignore the vectors A_j and answers each block $j \in [t]$ separately according to her strategy from $\mathsf{Game2a}$, and Bob answers according to Alice's strategy for $\mathsf{Game2a}$ on $j \neq c$, and according to Bob's strategy for $\mathsf{Game2a}$ for $j = c$. □

The following lemma establishes the soundness of the reduction in Definition 5.20.

Lemma 5.22 (Game2b soundness)**.** *Assume Condition 3.9. Suppose that in every 4^r-coloring of H, at least ε fraction of the edges are monochromatic. Then we have*

$$\operatorname{val}\Big(\mathsf{Game2b}[H, k, \ell_2, t, q_0, q_1, \beta]\Big) \leqslant 2^{-q_2/4}.$$

Proof. It suffices to show that

$$\mathrm{val}\Big(\mathsf{Game2b}[H,k,\ell_2,t,q_0,q_1,\beta]\Big) \leqslant \mathrm{val}\Big(\mathsf{Game2a}[H,k,\ell_2,q_1,0,\beta]\Big),$$

after which the conclusion follows from Lemma 5.19 (combined with Remark 5.3).

Assume Alice and Bob have a strategy for $\mathsf{Game2b}[H,k,\ell_2,t,q_0,q_1,\beta]$, and we next show how to derive a strategy for $\mathsf{Game2a}[H,k,\ell_2,q_1,0,\beta]$. In the latter game, Alice and Bob can first agree upon the following "shared data" using public randomness:

- a uniformly random critical index $c \in [t]$, and
- a uniformly random tuple (E_j, A_j, U_j) for each non-critical index $j \in [t] \setminus \{c\}$. More precisely, sample E_j from $\mathscr{E}(H,k)$ uniformly, sample A_j from $\big(\mathbb{F}_2^{E_j} \otimes \mathbb{F}_2^{E_j}\big)^{q_0}$ uniformly, and let U_j be the span of q_0 uniformly random vectors in $\mathbb{F}_2^{E_j}$.

Given a question (E,A,U) in $\mathsf{Game2a}[H,k,\ell_2,q_1,0,\beta]$, Alice can construct $(\mathcal{E},\mathcal{A},\mathcal{U})$ by planting (E,A,U) in the critical index c and using the shared data for the non-critical indices. Similarly, given a question (E',A',U'), Bob can construct $(\mathcal{E}',\mathcal{A}',\mathcal{U}')$ by planting (E',A',U') in the critical index c and using the shared data for the non-critical indices. The provers can then pretend that they are in $\mathsf{Game2b}[H,k,\ell_2,t,q_0,q_1,\beta]$, and use their strategy for $\mathsf{Game2b}[H,k,\ell_2,t,q_0,q_1,\beta]$ to answer the verifier in $\mathsf{Game2a}[H,k,\ell_2,q_1,0,\beta]$. Whenever the strategy for $\mathsf{Game2b}[H,k,\ell_2,t,q_0,q_1,\beta]$ provides consistent answers to $(\mathcal{E},\mathcal{A},\mathcal{U})$ and $(\mathcal{E}',\mathcal{A}',\mathcal{U}')$, the consistency in the critical index implies that the derived answers to the questions (E,A,U) and (E',A',U') in $\mathsf{Game2a}[H,k,\ell_2,q_1,0,\beta]$ are also consistent. Therefore, the verifier in $\mathsf{Game2a}[H,k,\ell_2,q_1,0,\beta]$ accepts with probability at least the acceptance probability in $\mathsf{Game2b}[H,k,\ell_2,t,q_0,q_1,\beta]$. $\square$

6 The Inner PCP

6.1 Tripled Packets

Definition 6.1. Let E be a tripled set. For any $v \in E$, we let $e_v \in \mathbb{F}_2^E$ denote the standard basis vector supported on v; that is, for $u \in V$ let $(e_v)_u = 1$ if and only if $u = v$.

(1) We define $\mathcal{N}_1(E) \subseteq \mathbb{F}_2^E \otimes \mathbb{F}_2^E$ be the linear subspace spanned by all vectors of the form

$$e_{v_1} \otimes e_{v_1} + e_{v_2} \otimes e_{v_2} + e_{v_3} \otimes e_{v_3} + e_{v_1} \otimes e_{v_2} + e_{v_2} \otimes e_{v_3} + e_{v_3} \otimes e_{v_1},$$

where $\{v_1, v_2, v_3\}$ ranges in $\mathrm{Triples}(E)$.

(2) We let $\mathcal{N}_2(E) \subseteq \mathbb{F}_2^E \otimes \mathbb{F}_2^E$ be the linear subspace spanned by all anti-symmetric vectors, i.e.

$$\mathcal{N}_2(E) = \mathrm{span}\left\{ x \otimes y + y \otimes x \,\middle|\, x, y \in \mathbb{F}_2^E \right\}.$$

(3) Let $\mathcal{N}(E) = \mathcal{N}_1(E) + \mathcal{N}_2(E)$ be the sum of the two subspaces.

(4) A linear map $f : \mathcal{N}(E) \to \mathbb{F}_2$ is said to be SymNAE-*satisfying* if we have:

- $f(e_{v_1} \otimes e_{v_1} + e_{v_2} \otimes e_{v_2} + e_{v_3} \otimes e_{v_3} + e_{v_1} \otimes e_{v_2} + e_{v_2} \otimes e_{v_3} + e_{v_3} \otimes e_{v_1}) = 1$ for all triples $\{v_1, v_2, v_3\} \in \mathrm{Triples}(E)$; and
- $f(x \otimes y + y \otimes x) = 0$ for all $x, y \in \mathbb{F}_2^E$.

Proposition 6.2. For any tripled set E, there is a unique linear map $f : \mathcal{N}(E) \to \mathbb{F}_2$ that is SymNAE-satisfying.

Definition 6.3. For a sequence of (not necessarily tripled) finite sets $\mathcal{E}' = (E'_1, \ldots, E'_t)$, we let $\mathcal{V}(\mathcal{E}')$ denote the $\mathbb{F}_2$-vector space

$$\mathcal{V}(\mathcal{E}') = \bigoplus_{j=1}^{t} \left(\mathbb{F}_2^{E'_j} \otimes \mathbb{F}_2^{E'_j} \right).$$

Definition 6.4. Fix a sequence of (not necessarily tripled) finite sets $\mathcal{E}' = (E'_1, \ldots, E'_t)$ and a sequence of linear subspaces $\mathcal{U} = (U_1, \ldots, U_t)$ with $U_j \subseteq \mathbb{F}_2^{E'_j}$ for each $j \in [t]$. We let $\bigoplus \mathcal{U} \subseteq \mathcal{V}(\mathcal{E}')$ be the linear subsapce defined by

$$\bigoplus \mathcal{U} = \bigoplus_{j=1}^{t} \Big(U_j \otimes U_j \Big) \subseteq \bigoplus_{j=1}^{t} \Big(\mathbb{F}_2^{E'_j} \otimes \mathbb{F}_2^{E'_j} \Big) = \mathcal{V}(\mathcal{E}').$$

Definition 6.5. A *tripled packet* is a pair $(\mathcal{E},\mathcal{U})$, where $\mathcal{E} = (E_1, \ldots, E_t)$ is a sequence of tripled sets, and $\mathcal{U} = (U_1, \ldots, U_t)$ is a sequence of linear subspaces with $U_j \subseteq \mathbb{F}_2^{E_j}$ for each $j \in [t]$. We let $\mathcal{N}(\mathcal{E},\mathcal{U}) \subseteq \mathcal{V}(\mathcal{E})$ be the linear subspace defined by

$$\mathcal{N}(\mathcal{E},\mathcal{U}) = \bigoplus_{j=1}^{t} \Big(\mathcal{N}(E_j) + U_j \otimes U_j \Big) \subseteq \bigoplus_{j=1}^{t} \Big(\mathbb{F}_2^{E_j} \otimes \mathbb{F}_2^{E_j} \Big) = \mathcal{V}(\mathcal{E}).$$

Definition 6.6. Given an tripled packet $(\mathcal{E},\mathcal{U})$ where $\mathcal{E} = (E_1, \ldots, E_t)$ and $\mathcal{U} = (U_1, \ldots, U_t)$, a linear map $f : \mathcal{N}(\mathcal{E},\mathcal{U}) \to \mathbb{F}_2$ is said to be SNR-*satisfying* if we have:

- $f|_{\mathcal{N}(E_j)}$ is SymNAE-satisfying for each $j \in [t]$.
- $\operatorname{rank}(f|_{U_j \otimes U_j}) \leqslant 1$ for each $j \in [t]$.

6.1.1 Thinning of Tripled Packets

Definition 6.7. Let $(E_1, \ldots, E_t)$ be a sequence of tripled sets. Fix an index $c \in [t]$ and a parameter $\beta \in [0, 1]$. We write $\mathcal{E}' \sim \mathrm{Thin}_{c,\beta}(\mathcal{E})$ for the random sequence $\mathcal{E}' = (E'_1, \ldots, E'_t)$ obtained by setting $E'_j = E_j$ for all $j \in [t] \setminus \{c\}$ and sampling $E'_c \sim \mathrm{Thin}_\beta(E_c)$. A sequence $\mathcal{E}'$ in the support of the distribution $\mathrm{Thin}_{c,1/2}(\mathcal{E})$ is called a *c-critical descendant* of $\mathcal{E}$.

Definition 6.8. Let $\mathcal{E} = (E_1, \ldots, E_t)$ be a sequence of tripled sets. Fix a parameter $\beta \in [0, 1]$ and a nonnegative integer ℓ. We let $\mathcal{J}_{\mathsf{in}}(\mathcal{E}, \beta, \ell)$ denote the output distribution of the following process:

1. Sample a uniformly random index $c \in [t]$. Sample $(E'_c, \texttt{empty}, \texttt{empty}, U_c) \sim \mathcal{J}_{\mathsf{mid}}(E_c, 0, 0, \beta, \ell)$. Sample a random vector $x \in \mathbb{F}_2^{E'_c}$ and let $U'_c = U_c + \operatorname{span}(x)$.
2. For each $j \in [t] \setminus \{c\}$, sample $(E_j^{\mathrm{ghost}}, \texttt{empty}, \texttt{empty}, U_j) \sim \mathcal{J}_{\mathsf{mid}}(E_j, 0, 0, \beta, \ell)$. Let $E'_j = E_j$ and $U'_j = U_j$.
3. Let $\mathcal{E}' = (E'_1, \ldots, E'_t)$, $\mathcal{U} = (U_1, \ldots, U_t)$ and $\mathcal{U}' = (U'_1, \ldots, U'_t)$.
4. Output the tuple $(c,\mathcal{U},\mathcal{E}',\mathcal{U}')$.

Proposition 6.9. In the setting and notation of Definition 6.8, the random subspaces $U_1, \ldots, U_t$ are mutually independent. Furthermore, they are jointly independent of the random index c.

6.2 The Low-Rank Grassmann Graph

6.2.1 Construction from Hadamard Code

Definition 6.10. Let ℓ be a positive integer. For each $i \in \{0, 1, \ldots, 2^\ell - 1\}$ and $b \in \{0, 1, \ldots, \ell - 1\}$, define $\mathrm{Bit}_\ell(i, b) \in \mathbb{F}_2$ to be the mod-2 remainder of the integer $\lfloor i \cdot 2^{-b} \rfloor$. Equivalently, the sequence

$$\big(\mathrm{Bit}_\ell(i, \ell - 1), \mathrm{Bit}_\ell(i, \ell - 2), \ldots, \mathrm{Bit}_\ell(i, 0)\big)$$

is the length-ℓ binary representation of i, with the most significant bit listed first. For each integer $i \in \{0, 1, \ldots, 2^\ell - 1\}$ and vector $u \in \mathbb{F}_2^\ell$, define

$$\mathrm{Had}_\ell(u, i) = \sum_{b=1}^{\ell} u_b \, \mathrm{Bit}_\ell(i, \ell - b) \in \mathbb{F}_2$$

to be the "inner product" of u with the binary representation of i (viewed as a vector in $\mathbb{F}_2^\ell$).

Definition 6.11. Let ℓ, s be positive integers, and let $q = 2s(2^\ell - 1)$. Suppose E' is a finite set, and $A \in (\mathbb{F}_2^{E'})^q$ is a sequence of vectors

$$A = \big(a^{(1)}, \ldots, a^{(q)}\big).$$

For each $u \in \mathbb{F}_2^\ell$, define a vector

$$\mathrm{Lvec}(A, \ell, u) = \sum_{i=0}^{2^\ell - 2} \left(\mathrm{Had}_\ell(u, 2^\ell - i - 1) \sum_{j=1}^{s} a^{(2si+2j-1)} \otimes a^{(2si+2j)} \right) \in \mathbb{F}_2^{E'} \otimes \mathbb{F}_2^{E'} \tag{6.1}$$

Since $u \mapsto \mathrm{Had}_\ell(u, 2^\ell - i - 1)$ is a linear map for each i, the map $u \mapsto \mathrm{Lvec}(A, \ell, u)$ from $\mathbb{F}_2^\ell$ to $\mathbb{F}_2^{E'} \otimes \mathbb{F}_2^{E'}$ is also linear. For any linear subspace $U \subseteq \mathbb{F}_2^\ell$, we denote the image of the latter linear map restricted to U by

$$\mathrm{Lspan}(A, \ell, U) = \{\mathrm{Lvec}(A, \ell, u) \mid u \in U\} \subseteq \mathbb{F}_2^{E'} \otimes \mathbb{F}_2^{E'}.$$

Remark 6.12. In the setting and notation of Definition 6.11, suppose that the odd-indexed vectors $a^{(1)}, a^{(3)}, \ldots, a^{(q-1)}$ are linearly independent, and that the even-indexed vectors $a^{(2)}, a^{(4)}, \ldots, a^{(q)}$ are also linearly independent. Then for any $u \in \mathbb{F}_2^\ell \setminus \{0\}$, since $\mathrm{Had}_\ell(u, i) = 1$ for exactly $2^{\ell-1}$ of the i's in $\{1, 2, \ldots, 2^\ell - 1\}$, it follows from Equation (6.1) that

$$\mathrm{rank}\,(\mathrm{Lvec}(A, \ell, u)) = 2^{\ell-1} s.$$

6.2.2 Mixed-Type Advice Sequences

Definition 6.13. Given a sequence of finite sets $\mathcal{E}' = (E'_1, \ldots, E'_t)$, a subset of indices $J \subseteq [t]$ and positive integers ℓ, s, we let $\mathscr{A}_J(\mathcal{E}', \ell, s)$ be the collection of advice sequences $(A_1, \ldots, A_t)$ such that:

(1) For each $j \in [t] \setminus J$, the component A_j is a sequence of ℓ vectors $A_j = (a^{(j,1)}, \ldots, a^{(j,\ell)})$, where

$$a^{(j,i)} \in \mathbb{F}_2^{E'_j} \otimes \mathbb{F}_2^{E'_j} \qquad \text{for all } i \in [\ell].$$

(2) For each $c \in J$, the component A_c is a sequence of $q = 2s(2^\ell - 1)$ vectors $A_c = (a^{(c,1)}, \dots, a^{(c,q)})$, where

$$a^{(c,i)} \in \mathbb{F}_2^{E'_c} \qquad \text{for all } i \in [q].$$

If both sequences $\left(a^{(c,1)}, a^{(c,3)} \dots, a^{(c,q-1)}\right)$ and $\left(a^{(c,2)}, a^{(c,4)}, \dots, a^{(c,q)}\right)$ are linearly independent, we say that A_c is *saturated*.

We let $\mathscr{A}^*_J(\mathcal{E}', \ell, s)$ denote the sub-collection of advices sequences $(A_1, \dots, A_t) \in \mathscr{A}_J(\mathcal{E}', \ell, s)$ such that A_c is saturated for each $c \in J$.

Definition 6.14. Fix a sequence of finite sets $\mathcal{E}' = (E'_1, \dots, E'_t)$, a proper subset of indices $J \subsetneq [t]$ and positive integers ℓ, s. Given an advice sequence $\mathcal{A} = (A_1, \dots, A_t) \in \mathscr{A}_J(\mathcal{E}', \ell, s)$, write $A_j = \left(a^{(j,1)}, \dots, a^{(j,\ell)}\right)$ for each $j \in [t] \setminus J$. For each $u \in \mathbb{F}_2^\ell$, define

$$\mathrm{Mvec}_J(\mathcal{A}, u) = \sum_{c \in J} \mathrm{Lvec}(A_c, \ell, u) + \sum_{j \in [t] \setminus J} \left(\sum_{b=1}^{\ell} u_b \cdot a^{(j,b)} \right) \in \mathcal{V}(\mathcal{E}'). \tag{6.2}$$

For any linear subspace $U \subseteq \mathbb{F}_2^\ell$, we denote

$$\mathrm{Mspan}_J(\mathcal{A}, U) = \{\mathrm{Mvec}_J(A, u) \,|\, u \in U\} \subseteq \mathcal{V}(\mathcal{E}'). \tag{6.3}$$

Notation 6.15. In the remainder of the paper, we will often use the notation $\mathrm{Mspan}_J(\mathcal{A}, U)$ (defined in Equation (6.3)) with $U = \mathbb{F}_2^\ell$. When J is clear from context, we abbreviate

$$\sigma(\mathcal{A}) := \mathrm{Mspan}_J\big(\mathcal{A}, \mathbb{F}_2^\ell\big).$$

We also write

$$\rho(\mathcal{A}) := \big(\mathrm{Mvec}_J(\mathcal{A}, e_1), \dots, \mathrm{Mvec}_J(\mathcal{A}, e_\ell)\big) \in (\mathcal{V}(\mathcal{E}'))^\ell,$$

where $e_1, \dots, e_\ell$ denote the standard basis vectors of $\mathbb{F}_2^\ell$. Thus, $\sigma(\mathcal{A})$ is the linear span of the vectors in the sequence $\rho(\mathcal{A})$, and we write $\sigma(\mathcal{A}) = \mathrm{span}(\rho(\mathcal{A}))$.

Definition 6.16. Fix a sequence of finite sets $\mathcal{E}' = (E'_1, \dots, E'_t)$, a proper subset of indices $J \subsetneq [t]$, positive integers ℓ, s and a nonnegative integer $q_0 \leqslant \ell$. Given an advice sequence $\mathcal{A} = (A_1, \dots, A_t) \in \mathscr{A}_J(\mathcal{E}', \ell, s)$, we define a *sub-advice* $\mathcal{A}_{[q_0]} = (A'_1, \dots, A'_t) \in \mathscr{A}_J(\mathcal{E}', q_0, 2^{\ell - q_0} s)$ as follows:

(1) For each $j \in [t] \setminus J$, write $A_j = \left(a^{(j,1)}, \dots, a^{(j,\ell)}\right)$ and let $A'_j = \left(a^{(j,1)}, \dots, a^{(j,q_0)}\right)$.

(2) For each $c \in J$, write $A_c = \left(a^{(c,1)}, \dots, a^{(c,q)}\right)$, where $q = 2s(2^\ell - 1)$. Denote $q_1 = 2s(2^\ell - 2^{\ell - q_0})$ and let $A'_c = \left(a^{(c,1)}, \dots, a^{(c,q_1)}\right)$.

Proposition 6.17. In the setting and notation of Definition 6.16, we have[4]

$$\mathrm{Mvec}_J(\mathcal{A}, u) = \mathrm{Mvec}_J(\mathcal{A}_{[q_0]}, u) \qquad \text{for all } u \in \mathbb{F}_2^{q_0},$$

Consequently, by Equation (6.3) we have

$$\mathrm{Mspan}_J(\mathcal{A}, \mathbb{F}_2^{q_0}) = \mathrm{Mspan}_J(\mathcal{A}_{[q_0]}, \mathbb{F}_2^{q_0}).$$

[4]Here, we identify $\mathbb{F}_2^{q_0}$ with the subspace $\mathbb{F}_2^{q_0} \times \{0\}^{\ell - q_0} \subseteq \mathbb{F}_2^\ell$; that is, a q_0-bit string is extended to an ℓ-bit string by appending $\ell - q_0$ zeros.

Proof. Denote $s_1 = 2^{\ell-q_0}s$. As in Definition 6.16, let $q = 2s(2^\ell - 1)$ and $q_1 = 2s(2^\ell - 2^{\ell-q_0})$.

Let i be an integer ranging in $\{0, 1, \ldots, 2^\ell - 2\}$. For each $u \in \mathbb{F}_2^{q_0}$ considered as an ℓ-bit string, since its last $(\ell - q_0)$ bits are all 0's, the value of

$$\mathrm{Had}_\ell(u, 2^\ell - i - 1) = \sum_{b=1}^{q_0} u_b(1 - \mathrm{Bit}_\ell(i, \ell - b))$$

does not depend on the last $(\ell - q_0)$ bits of the binary representation of i. Furthermore, if the first q_0 bits of i are all 1's, then $\mathrm{Had}_\ell(u, 2^\ell - i - 1) = 0$. Therefore, if we define $\pi(i) = 2^{\ell-q_0}\lfloor i/2^{\ell-q_0}\rfloor$ for each $i \in \{0, 1, \ldots, 2^\ell - 2^{\ell-q_0} - 1\}$, for each $c \in J$ and $u \in \mathbb{F}_2^{q_0}$ we can deduce from Equation (6.1) that

$$\begin{aligned}
\mathrm{Lvec}(A_c, \ell, u) &= \sum_{i=0}^{2^\ell-2^{\ell-q_0}-1} \left(\mathrm{Had}_\ell(u, 2^\ell - \pi(i) - 1) \sum_{j=1}^{s} a^{(c,2si+2j-1)} \otimes a^{(c,2si+2j)} \right) \\
&= \sum_{i=0}^{2^{q_0}-2} \left(\mathrm{Had}_\ell(u, 2^\ell - 2^{\ell-q_0}i - 1) \sum_{j=1}^{s_1} a^{(c,2s_1i+2j-1)} \otimes a^{(c,2s_1i+2j)} \right) \\
&= \sum_{i=0}^{2^{q_0}-2} \left(\mathrm{Had}_{q_0}(u, 2^{q_0} - i - 1) \sum_{j=1}^{s_1} a^{(c,2s_1i+2j-1)} \otimes a^{(c,2s_1i+2j)} \right) = \mathrm{Lvec}(A'_c, q_0, u).
\end{aligned}$$

It then follows from Equation (6.2) that

$$\mathrm{Mvec}_J(\mathcal{A}, u) = \sum_{c\in J} \mathrm{Lvec}(A'_c, q_0, u) + \sum_{j\in[t]\setminus J} \left(\sum_{b=1}^{q_0} u_b \cdot a^{(j,b)} \right) = \mathrm{Mvec}_J(\mathcal{A}_{[q_0]}, u),$$

for any $u \in \mathbb{F}_2^{q_0}$. $\square$

Notation 6.18. Throughout the remainder of the paper, the set $J \subseteq [t]$ of critical indices will always satisfy $|J| \leqslant 2$. When $J = \{c_1, c_2\}$ with $c_1 \neq c_2$, we suppress the braces and write c_1, c_2 in place of $\{c_1, c_2\}$ whenever J appears as a subscript. For example, we write

$$\mathscr{A}_{c_1,c_2}(\mathcal{E}', \ell, s) \qquad \text{instead of} \qquad \mathscr{A}_{\{c_1,c_2\}}(\mathcal{E}', \ell, s).$$

Similarly, when $J = \{c\}$, we write, for instance,

$$\mathrm{Mspan}_c(\mathcal{A}, U) \qquad \text{instead of} \qquad \mathrm{Mspan}_{\{c\}}(\mathcal{A}, U).$$

6.2.3 Zoom-out Probabilities

Proposition 6.19. Fix a sequence of non-empty finite sets $\mathcal{E}' = (E'_1, \ldots, E'_t)$, a subset of indices $J \subsetneq [t]$, positive integers ℓ, s and a nonnegative integer $q_0 < \ell$. Let $f : \mathcal{V}(\mathcal{E}') \to \mathbb{F}_2$ be an arbitrary linear functional that is not identically zero. For any fixed advice sequence $\mathcal{A}' \in \mathscr{A}_J(\mathcal{E}', q_0, 2^{\ell-q_0}s)$ and any vector $u \in \mathbb{F}_2^\ell \setminus \mathbb{F}_2^{q_0}$, we have

$$\frac{1}{2} - 2^{-s-1} \leqslant \mathop{\mathbb{P}}_{\mathcal{A}\in\mathscr{A}_J(\mathcal{E}',\ell,s)} \left[f\big(\mathrm{Mvec}_J(\mathcal{A}, u)\big) = 1 \,\middle|\, \mathcal{A}_{[q_0]} = \mathcal{A}' \right] \leqslant \frac{1}{2} + 2^{-s-1}.$$

Proof. As usual, we denote $\mathcal{A} = (A_1, \dots, A_t)$ and $A_j = \big(a^{(j,1)}, a^{(j,2)}, \dots\big)$ for each $j \in [t]$. Since f is not identically zero, there exists an index $j \in [t]$ such that the restriction of f to $\mathbb{F}_2^{E'_j} \otimes \mathbb{F}_2^{E'_j}$ is nonzero. Depending on whether this index lies in J, we divide into two cases:

Case 1: the restriction of f to $\mathbb{F}_2^{E'_j} \otimes \mathbb{F}_2^{E'_j}$ is nonzero for some $j \in [t] \setminus J$. In this case, if we let $b \in \{q_0+1, \dots, \ell\}$ be an index such that $u_b = 1$, even after conditioning on all vectors in $\mathcal{A}$ except for $a^{(j,b)}$, the value of $f\big(\mathrm{Mvec}_J(\mathcal{A}, u)\big)$ (see Equation (6.2)) is still an unbiased Bernoulli random variable purely due to the randomness of the vector $a^{(j,b)} \in \mathbb{F}_2^{E'_j} \otimes \mathbb{F}_2^{E'_j}$. Since $a^{(j,b)}$ is not included in the sub-advice $\mathcal{A}_{[q_0]}$, we conclude that

$$\mathop{\mathbb{P}}_{\mathcal{A} \in \mathscr{A}_J(\mathcal{E}', \ell, s)} \left[f\big(\mathrm{Mvec}_J(\mathcal{A}, u)\big) = 1 \,\middle|\, \mathcal{A}_{[q_0]} = \mathcal{A}' \right] = \frac{1}{2}.$$

Case 2: the restriction of f to $\mathbb{F}_2^{E'_c} \otimes \mathbb{F}_2^{E'_c}$ is nonzero for some $c \in J$. As in case 1, we let $b \in \{q_0+1, \dots, \ell\}$ be an index such that $u_b \neq 0$. Now let $i = 2^\ell - 2^{\ell-b} - 1$, whose ℓ-bit binary representation has a 0 in the b-th position and 1's in all other positions. Even after conditioning on all vectors in $\mathcal{A}$ except for the $2s$ vectors (none of which is included in the sub-advice $\mathcal{A}_{[q_0]}$)

$$a^{(c,2si+1)}, a^{(c,2si+2)} \dots, a^{(c,2si+2s)},$$

the value of $f\big(\mathrm{Mvec}_J(\mathcal{A}, u)\big)$ (see Equations (6.1) and (6.2)) is still the Bernoulli random variable

$$X = \sum_{j=1}^{s} f\big(a^{(c,2si+2j-1)} \otimes a^{(c,2si+2j)}\big)$$

or its negation. By independence among the vectors in $\mathcal{A}$, it is easy to see that $\mathbb{P}[X=1]$ lies in the interval $[\frac{1}{2} - 2^{-s-1}, \frac{1}{2} + 2^{-s-1}]$, and hence we have

$$\frac{1}{2} - 2^{-s-1} \leqslant \mathop{\mathbb{P}}_{\mathcal{A} \in \mathscr{A}_J(\mathcal{E}', \ell, s)} \left[f\big(\mathrm{Mvec}_J(\mathcal{A}, u)\big) = 1 \,\middle|\, \mathcal{A}_{[q_0]} = \mathcal{A}' \right] \leqslant \frac{1}{2} + 2^{-s-1}. \qquad \square$$

Proposition 6.20. Fix a sequence of non-empty finite sets $\mathcal{E}' = (E'_1, \dots, E'_t)$, a subset of indices $J \subsetneq [t]$, and positive integers ℓ, s. Let $W \subseteq \mathcal{V}(\mathcal{E}')$ be a linear subspace with codimension w. For any fixed advice sequence $\mathcal{A}' \in \mathscr{A}_J(\mathcal{E}', \ell-1, 2s)$ such that $\sigma(\mathcal{A}') \subseteq W$, we have

$$2^{-w} - 2^{-s} \leqslant \mathop{\mathbb{P}}_{\mathcal{A} \in \mathscr{A}_J(\mathcal{E}', \ell, s)} \left[\sigma(\mathcal{A}) \subseteq W \,\middle|\, \mathcal{A}_{[\ell-1]} = \mathcal{A}' \right] \leqslant 2^{-w} + 2^{-s}.$$

Proof. We fix an arbitrary vector $u \in \mathbb{F}_2^\ell \setminus \mathbb{F}_2^{\ell-1}$. Since $\sigma(\mathcal{A}') \subseteq W$ and

$$\sigma(\mathcal{A}) = \sigma(\mathcal{A}_{[\ell-1]}) + \mathrm{span}(\mathrm{Mvec}_J(\mathcal{A}, u)),$$

the desired inequality is equivalent to

$$2^{-w} - 2^{-s} \leqslant \mathop{\mathbb{P}}_{\mathcal{A} \in \mathscr{A}_J(\mathcal{E}', \ell, s)} \left[\mathrm{Mvec}_J(\mathcal{A}, u) \in W \,\middle|\, \mathcal{A}_{[\ell-1]} = \mathcal{A}' \right] \leqslant 2^{-w} + 2^{-s}. \tag{6.4}$$

Let $\mathcal{F}$ be the collection of linear functionals $f : \mathcal{V}(\mathcal{E}') \to \mathbb{F}_2$ such that $W \subseteq \ker(f)$, and denote

$$p = \mathop{\mathbb{E}}_{f \in \mathcal{F}} \left[\mathop{\mathbb{P}}_{\mathcal{A} \in \mathscr{A}(\mathcal{E}', \ell, s)} \left[f\big(\mathrm{Mvec}_J(\mathcal{A}, u)\big) = 1 \,\middle|\, \mathcal{A}_{[\ell-1]} = \mathcal{A}' \right] \right]. \tag{6.5}$$

Note that $|\mathcal{F}| = 2^w$, and for each $f \in \mathcal{F}$ that is not identically zero we may apply Proposition 6.19 to estimate the probability expression on the right-hand side of Equation (6.5), which yields

$$\frac{2^w - 1}{2^w}\left(\frac{1}{2} - 2^{-s-1}\right) \leqslant p \leqslant \frac{2^w - 1}{2^w}\left(\frac{1}{2} + 2^{-s-1}\right). \tag{6.6}$$

On the other hand, by exchanging the expectation and the probability on the right-hand side of Equation (6.5) (i.e. double counting), we have

$$p = \frac{1}{2}\left(1 - \mathop{\mathbb{P}}_{\mathcal{A}\in\mathscr{A}_J(\mathcal{E}',\ell,s)}\left[\mathrm{Mvec}_J(\mathcal{A}, u) \in W \,\middle|\, \mathcal{A}_{[\ell-1]} = \mathcal{A}'\right]\right). \tag{6.7}$$

Comparing Equations (6.6) and (6.7) immediately yields the desired Equation (6.4). $\square$

Corollary 6.21. *Fix a sequence of non-empty finite sets $\mathcal{E}' = (E'_1, \ldots, E'_t)$, a subset of indices $J \subsetneq [t]$, and positive integers ℓ, s. Let $W \subseteq \mathcal{V}(\mathcal{E}')$ be a linear subspace with codimension w, and suppose that $s \geqslant \ell + w$. Then for any nonnegative integer $q_0 < \ell$ and any fixed advice sequence $\mathcal{A}' \in \mathscr{A}_J(\mathcal{E}', q_0, 2^{\ell-q_0}s)$ such that $\sigma(\mathcal{A}') \subseteq W$, we have*

$$2^{-w(\ell-q_0)-1} \leqslant \mathop{\mathbb{P}}_{\mathcal{A}\in\mathscr{A}_J(\mathcal{E}',\ell,s)}\left[\sigma(\mathcal{A}) \subseteq W \,\middle|\, \mathcal{A}_{[q_0]} = \mathcal{A}'\right] \leqslant 2^{-w(\ell-q_0)+1}.$$

Proof. The conclusion follows easily by induction on $\ell - q_0$, using Proposition 6.20 as induction steps. The condition $s \geqslant \ell + w$ guarantees that

$$(2^{-w} - 2^{-s})^{\ell-q_0} \geqslant 2^{-w(\ell-q_0)}\left(1 - (\ell - q_0)2^{-s+w}\right) \geqslant 2^{-w(\ell-q_0)}(1 - 2^{\ell-1-s+w}) \geqslant 2^{-w(\ell-q_0)-1},$$

and similarly for the upper bound. $\square$

6.2.4 Mixed-Type Grassmann Graph

Definition 6.22. Fix a sequence of finite sets $\mathcal{E}' = (E'_1, \ldots, E'_t)$, a nonempty subset of indices $J \subsetneq [t]$ and positive integers ℓ, s. Define $\mathsf{MGr}_J(\mathcal{E}', \ell, s)$ to be the following collection of linear subspaces:

$$\mathsf{MGr}_J(\mathcal{E}', \ell, s) = \left\{\sigma(\mathcal{A}) \,\middle|\, \mathcal{A} \in \mathscr{A}^*_J(\mathcal{E}', \ell, s)\right\} \subseteq \mathsf{Gr}(\mathcal{V}(\mathcal{E}'), \ell).$$

Proposition 6.23. Fix a sequence of non-empty finite sets $\mathcal{E}' = (E'_1, \ldots, E'_t)$, a nonempty subset of indices $J \subsetneq [t]$, positive integers ℓ, s and a nonnegative integer $q_0 < \ell$. Assume that $|E'_c| \geqslant 2^{\ell+2}s$ for each $c \in J$. For any fixed advice sequence $\mathcal{A}' \in \mathscr{A}^*_J(\mathcal{E}', q_0, 2^{\ell-q_0}s)$, if we sample $\mathcal{A} \in \mathscr{A}_J(\mathcal{E}', \ell, s)$ conditioned on $\sigma(\mathcal{A}_{[q_0]}) = \sigma(\mathcal{A}')$, the resulting distribution of $\sigma(\mathcal{A})$ has total variation at most

$$\sum_{c\in J} 2^{-|E'_c|/2}$$

to the uniform distribution over the set $\left\{L \in \mathsf{MGr}_J(\mathcal{E}', \ell, s) \,\middle|\, \sigma(\mathcal{A}') \subseteq L\right\}$.

Proof. It is not hard to see that if we sample $\mathcal{A} \in \mathscr{A}^*_J(\mathcal{E}', \ell, s)$ conditioned on $\sigma(\mathcal{A}_{[q_0]}) = \sigma(\mathcal{A}')$, the resulting distribution of $\sigma(\mathcal{A})$ is exactly the uniform distribution over $\left\{L \in \mathsf{MGr}_J(\mathcal{E}', \ell, s) \,\middle|\, \sigma(\mathcal{A}') \subseteq L\right\}$. On the other hand, writing $q = 2s(2^\ell - 1)$ and $q_1 = 2s(2^\ell - 2^{\ell-q_0})$, for any $\mathcal{A}' \in \mathscr{A}^*_J(\mathcal{E}', q_0, 2^{\ell-q_0}s)$ we have

$$\mathop{\mathbb{P}}_{\mathcal{A}\in\mathscr{A}_J(\mathcal{E}',\ell,s)}\left[\mathcal{A} \in \mathscr{A}^*_J(\mathcal{E}', \ell, s) \,\middle|\, \mathcal{A}_{[q_0]} = \mathcal{A}'\right]$$

$$= \mathbb{P}\left[\left(a^{(c,1)},\ldots,a^{(c,q)}\right) \text{ is saturated for each } c\in J \,\middle|\, \text{fixed saturated sequences } \left(a^{(c,1)},\ldots,a^{(c,q_1)}\right)\right]$$

$$\text{(where } a^{(c,i)} \text{ (for } c\in J \text{ and } i\in[q]) \text{ are sampled independently and uniformly from } \mathbb{F}_2^{E'_c})$$

$$\geqslant \prod_{c\in J}\left(1-2^{(2^\ell-1)s-|E'_c|}\right)^2 \geqslant 1-\sum_{c\in J}2^{-|E'_c|/2}. \qquad \square$$

Proposition 6.24. Fix a sequence of non-empty finite sets $\mathcal{E}'=(E'_1,\ldots,E'_t)$, a subset of indices $J\subsetneq[t]$ such that $1\leqslant|J|\leqslant 2$, positive integers ℓ,s and nonnegative integers q_0,w. Assume

$$s\geqslant \ell+w, \qquad q_0<\ell, \qquad \text{and} \qquad |E'_c|\geqslant 2^{\ell+2}s \text{ for each } c\in J.$$

Then for any fixed subspaces

$$Q\in \mathsf{MGr}_J(\mathcal{E}',q_0,2^{\ell-q_0}s) \qquad \text{and} \qquad W\in \mathsf{Gr}\big(\mathcal{V}(\mathcal{E}'),\dim\mathcal{V}(\mathcal{E}')-w\big)$$

such that $Q\subseteq W$, we have

$$2^{-w(\ell-q_0)-2}\leqslant \mathop{\mathbb{P}}_{L\in\mathsf{MGr}_J(\mathcal{E}',\ell,s)}[L\subseteq W\mid Q\subseteq L]\leqslant 2^{-w(\ell-q_0)+2}.$$

Proof. Pick $\mathcal{A}'$ uniformly at random from the set

$$\left\{\mathcal{A}'\in\mathscr{A}_J(\mathcal{E}',q_0,2^{\ell-q_0}s)\,\middle|\,\sigma(\mathcal{A}')=Q\right\}.$$

It follows from Proposition 6.23 that

$$\left|\mathop{\mathbb{P}}_{L\in\mathsf{MGr}_J(\mathcal{E}',\ell,s)}[L\subseteq W\mid Q\subseteq L]-\mathop{\mathbb{E}}_{\mathcal{A}'}\left[\mathop{\mathbb{P}}_{\mathcal{A}\in\mathscr{A}_J(\mathcal{E}',\ell,s)}\left[\sigma(\mathcal{A})\subseteq W\,\middle|\,\mathcal{A}_{[q_0]}=\mathcal{A}'\right]\right]\right|\leqslant\sum_{c\in J}2^{-|E'_c|/2}\leqslant 2^{-w\ell-2}.$$

Combining the above with Corollary 6.21 yields the desired conclusion. $\square$

6.2.5 Extension of Agreement

Lemma 6.25. *Fix a sequence of non-empty finite sets $\mathcal{E}'=(E'_1,\ldots,E'_t)$, a critical index $c\in[t]$, positive integers ℓ,s, nonnegative integers q_0,w and a parameter $\delta\in(0,1)$. Assume*

$$s\geqslant 2\ell \qquad \textit{and} \qquad \ell>2q_0+2w.$$

Let $W\subseteq\mathcal{V}(\mathcal{E}')$ be a linear subspace with codimension w, and $Q\subseteq W$ a linear subspace such that some $\sigma(\mathcal{A})=Q$ for some $\mathcal{A}\in\mathscr{A}_c(\mathcal{E}',q_0,2^{\ell+1-q_0}s)$. Suppose $F[\cdot]$ is a table where $F[L]\in L^*\cup\{\mathtt{nil}\}$ for each linear subspace $L\subseteq\mathcal{V}(\mathcal{E}')$, and there exists a linear function $f:W\to\mathbb{F}_2$ such that

$$\mathop{\mathbb{P}}_{\mathcal{A}\in\mathscr{A}_c(\mathcal{E}',\ell+1,s)}\left[F[\sigma(\mathcal{A})]|_{\sigma(\mathcal{A}_{[\ell]})}=f|_{\sigma(\mathcal{A}_{[\ell]})}\,\middle|\,\sigma(\mathcal{A}_{[q_0]})=Q \text{ and } \sigma(\mathcal{A}_{[\ell]})\subseteq W\right]\geqslant\delta. \tag{6.8}$$

Then there exists a subspace $W'\subseteq\mathcal{V}(\mathcal{E}')$ containing W and a linear function $g:W'\to\mathbb{F}_2$ such that

$$\mathop{\mathbb{P}}_{\mathcal{A}\in\mathscr{A}_c(\mathcal{E}',\ell+1,s)}\left[F[\sigma(\mathcal{A})]=g|_{\sigma(\mathcal{A})}\,\middle|\,\sigma(\mathcal{A}_{[q_0]})=Q \text{ and } \sigma(\mathcal{A})\subseteq W'\right]\geqslant\frac{\delta}{4}. \tag{6.9}$$

Proof. We let η be the maximum possible value of the left-hand side of Equation (6.9) when W' ranges in the set of linear subspaces of codimension at most w and g ranges in $(W')^*$. The goal is to show that $4\eta\geqslant\delta$. Throughout the proof, we let $L,L'\subseteq\mathcal{V}(\mathcal{E}')$ be random subspaces sampled by the following process:

1. Sample $\mathcal{A} \in \mathscr{A}_c(\mathcal{E}', \ell+1, s)$ conditioned on $\sigma(\mathcal{A}_{[q_0]}) = Q$.
2. Let $L = \sigma(\mathcal{A}_{[\ell]})$ and $L' = \sigma(\mathcal{A})$.

The most important fact about the joint distribution of (L, L') that we are going to use is the following: conditioned on L', for any linear subspace such that $Q \subseteq L_0 \subseteq L'$ and $\dim L_0 = \dim L - 1$, the probability that $L = L_0$ is exactly $(2^{\ell+1-q_0} - 1)^{-1}$.

Let $p = \mathbb{P}[L' \subseteq W]$ and denote

$$\begin{aligned}
\gamma_0 &= \mathbb{P}\left[F[L'] = f|_{L'} \,\middle|\, L' \subseteq W \right], \\
\gamma_1 &= \mathbb{P}\left[F[L']\big|_Q = f|_Q \text{ but } F[L'] \neq f|_{L'} \,\middle|\, L' \subseteq W \right].
\end{aligned}$$

By the definitions of γ_0 and η, we immediately have $\eta \geqslant \gamma_0$. We denote by $\mathcal{W}$ the collection of linear subspaces $W' \subseteq \mathcal{V}(\mathcal{E})$ such that $\dim W' = 1 + \dim W$ and $W \subseteq W'$. For each $W' \in \mathcal{W}$, we let

$$\begin{aligned}
p_{W'} &= \mathbb{P}\left[L' \subseteq W' \text{ but } L' \not\subseteq W \right], \text{ and} \\
\gamma_{W'} &= \begin{cases} \mathbb{P}\left[F[L']|_{L' \cap W} = f|_{L' \cap W} \,\middle|\, L' \subseteq W' \text{ but } L' \not\subseteq W \right], & \text{if } p_{W'} > 0 \\ 0, & \text{if } p_{W'} = 0. \end{cases}
\end{aligned}$$

For each $W' \in \mathcal{W}$, there are exactly two linear extensions of f to W'. If we choose $g : W' \to \mathbb{F}_2$ to be a uniformly random one of the two, then it is easy to see that

$$\mathop{\mathbb{E}}_g \left[\mathbb{P}\left[F[L'] = g|_{L'} \text{ and } L' \subseteq W' \right] \right] = \gamma_0 p + \frac{1}{2} \gamma_{W'} p_{W'},$$

and hence

$$\eta \geqslant \mathop{\mathbb{E}}_g \left[\mathbb{P}\left[F[L'] = g|_{L'} \,\middle|\, L' \subseteq W' \right] \right] = \frac{\gamma_0 p + \gamma_{W'} p_{W'}/2}{p + p_{W'}}. \tag{6.10}$$

It is not hard to see that

$$\mathbb{P}[L \subseteq W] = p + \frac{1}{2^{\ell+1-q_0} - 1} \sum_{W \in \mathcal{W}} p_{W'}, \tag{6.11}$$

and furthermore,

$$\begin{aligned}
&\mathbb{P}\left[F[L']\big|_L = f|_L \text{ and } L \subseteq W \right] \\
&= \gamma_0 p + \frac{1}{2^{\ell+1-q_0} - 1} \left(\gamma_1 p + \sum_{W' \in \mathcal{W}} \gamma_{W'} p_{W'} \right) \\
&\leqslant 2\gamma_0 p + \frac{2}{2^{\ell+1-q_0} - 1} \sum_{W' \in \mathcal{W}} \left(\gamma_0 p + \frac{1}{2} \gamma_{W'} p_{W'} \right) && \text{(using } \gamma_1 \leqslant 1\text{)} \\
&\leqslant 2\eta \cdot p + 2\eta \cdot \frac{1}{2^{\ell+1-q_0} - 1} \sum_{W' \in \mathcal{W}} (p + p_{W'}) && \text{(using Equation (6.10) and } \gamma_0 \leqslant \eta\text{)} \\
&\leqslant 4\eta \cdot \mathbb{P}[L \subseteq W]. && \text{(using Equation (6.11) and } |\mathcal{W}| = 2^w - 1 \leqslant 2^{\ell+1-q_0} - 1\text{)}
\end{aligned}$$

This gives the desired inequality

$$4\eta \geqslant \mathbb{P}\left[F[L']\big|_L = f|_L \,\middle|\, L \subseteq W \right] \geqslant \delta,$$

where we used Equation (6.8) in the last transition. $\square$

6.3 The 4-to-1 Game

Definition 6.26. Let H be a 3-uniform hypergraph. Given positive integers k and t, we let $\mathscr{E}(H,k,t)$ denote the collection of all sequences $\mathcal{E} = (E_1,\dots,E_t)$ such that $E_j \in \mathscr{E}(H,k)$ for all $j \in [t]$.

Definition 6.27. Given positive integers k, ℓ_2, t, s_1, q_0, a parameter $\beta \in [0,1)$ and a 3-uniform hypergraph H, let $\mathsf{Game2c}[H,k,\ell_2,t,s_1,q_0,\beta]$ be a 2-prover 1-round game where the verifier executes the following:

1. Sample a uniformly random sequence $\mathcal{E} = (E_1,\dots,E_t) \in \mathscr{E}(H,k,t)$.
2. Sample $(c,\mathcal{U},\mathcal{E}',\mathcal{U}') \sim \mathcal{J}_{\mathsf{in}}(\mathcal{E},\beta,\ell_2)$.
3. Sample a uniformly random sequence $\mathcal{A} = (A_1,\dots,A_t) \in \mathscr{A}_c(\mathcal{E}',q_0,s)$.
4. Send $(\mathcal{E},\mathcal{A},\mathcal{U})$ to Alice and $(\mathcal{E}',\mathcal{A},\mathcal{U}')$ to Bob.
5. Alice must return a linear map $f : \mathcal{V}(\mathcal{E}) \to \mathbb{F}_2$ such that $f|_{\mathcal{N}(\mathcal{E},\mathcal{U})}$ is SNR-satisfying; Bob must return a linear map $g : \mathcal{V}(\mathcal{E}') \to \mathbb{F}_2$. Accept if $f|_{\mathcal{V}(\mathcal{E}')} = g$.

Proposition 6.28. For any positive integers k, ℓ_2, t, s_1, q_0, parameter $\beta \in [0,1)$ and 3-uniform hypergraph H, we have

$$\mathrm{val}\Big(\mathsf{Game2b}[H,k,\ell_2,t,q_0,2^{q_0}s_1,\beta]\Big) = \mathrm{val}\Big(\mathsf{Game2c}[H,k,\ell_2,t,s_1,q_0,\beta]\Big).$$

Proof. It is easy to see the equivalence between $\mathsf{Game2b}$ and $\mathsf{Game2c}$ by comparing Definitions 5.20 and 6.27. $\square$

Definition 6.29. Given positive integers k, ℓ_1, ℓ_2, t, s, a parameter $\beta \in [0,1)$ and a 3-uniform hypergraph H, let $\mathsf{Game3}[H,k,\ell_1,\ell_2,t,s,\beta]$ be a 2-prover 1-round game where the verifier executes the following:

1. Sample a uniformly random sequence $\mathcal{E} = (E_1,\dots,E_t) \in \mathscr{E}(H,k,t)$.
2. Sample $(c,\mathcal{U},\mathcal{E}',\mathcal{U}') \sim \mathcal{J}_{\mathsf{in}}(\mathcal{E},\beta,\ell_2)$.
3. Sample a uniformly random sequence $\mathcal{A} = (A_1,\dots,A_t) \in \mathscr{A}_c(\mathcal{E}',2\ell_1,s)$. Let
$$L = \sigma(\mathcal{A}_{[\ell_1]}) \subseteq \mathcal{V}(\mathcal{E}') \qquad \text{and} \qquad L' = \sigma(\mathcal{A}_{[\ell_1+1]}) \subseteq \mathcal{V}(\mathcal{E}').$$
4. Send $(\mathcal{E},\mathcal{U},L+\mathcal{N}(\mathcal{E},\mathcal{U}))$ to Alice and $(\mathcal{E}',\mathcal{U}',L'+\bigoplus\mathcal{U}')$ to Bob.
5. Alice must return a linear map $f : (L+\mathcal{N}(\mathcal{E},\mathcal{U})) \to \mathbb{F}_2$ such that $f|_{\mathcal{N}(\mathcal{E},\mathcal{U})}$ is SNR-satisfying; Bob must return a linear map $g : (L'+\bigoplus\mathcal{U}') \to \mathbb{F}_2$. Accept if $g|_{(L+\bigoplus\mathcal{U})} = f|_{(L+\bigoplus\mathcal{U})}$.

The game defined in Definition 6.29 is clearly an "at most 4-to-1" game, in the sense that for any response of Alice there are at most 4 responses for Bob which would make the verifier accept. Indeed, this is because $L \subseteq L'$ is of one less dimension, and $\mathcal{U}$, $\mathcal{U}'$ are identical on all coordinates, except for the critical coordinate $i \in [t]$ where $\mathcal{U}_i \subseteq \mathcal{U}'_i$ has one less dimension. Thus, the space of Bob has at most two more dimensions than the space of Alice (but it may be the case it only has one more dimension). It is not difficult to turn this "at most 4-to-1" game to an actual 4-to-1 game as per Definition 1.1, and we will include that in the full version of the paper.

The completeness and soundness of Game3 are formulated by the following two lemmas.

Lemma 6.30 (Game3 completeness)**.** *For any positive integers k, ℓ_1, ℓ_2, t, s, a parameter $\beta \in [0, 1)$ and a 2-colorable 3-uniform hypergraph H, we have*

$$\mathrm{val}\Big(\mathsf{Game3}[H, k, \ell_1, \ell_2, t, s, \beta]\Big) = 1.$$

Proof. Let $h : \mathscr{V}(H) \to \mathbb{F}_2$ be a proper 2-coloring of H. It can naturally be lifted to a map (recall from Definition 4.2)

$$\widetilde{h} : \mathscr{V}(H, k) \to \mathbb{F}_2 \text{ defined by } \widetilde{h}(i, \mathsf{v}) = h(\mathsf{v}) \text{ for any } (i, \mathsf{v}) \in \mathscr{V}(H, k).$$

The map $\widetilde{h}$ naturally extends to a linear map

$$g : \mathbb{F}_2^{\mathscr{V}(H,k)} \to \mathbb{F}_2,$$

which induces a bilinear map

$$\widetilde{g} : \mathbb{F}_2^{\mathscr{V}(H,k)} \otimes \mathbb{F}_2^{\mathscr{V}(H,k)} \to \mathbb{F}_2 \text{ defined by } \widetilde{g}(x \otimes y) = g(x)g(y) \text{ for each } x, y \in \mathbb{F}_2^{\mathscr{V}(H)}.$$

It is clear that $\widetilde{g}$ is symmetric, NAE-satisfying and of rank 1. Finally, we take

$$f : \bigoplus_{i=1}^{t} \Big(\mathbb{F}_2^{\mathscr{V}(H,k)} \otimes \mathbb{F}_2^{\mathscr{V}(H,k)}\Big) \to \mathbb{F}_2$$

to be the linear map defined by

$$f(z_1, \dots, z_t) = g(z_1) + \cdots + g(z_t), \text{ for any } z_1, \dots, z_t \in \mathbb{F}_2^{\mathscr{V}(H,k)} \otimes \mathbb{F}_2^{\mathscr{V}(H,k)}.$$

In $\mathsf{Game3}[H, k, \ell_1, \ell_2, t, s, \beta]$, the provers can make the verifier accept with probability 1 using the following strategy:

(1) Alice returns $f|_{L+\mathcal{N}(\mathcal{E},\mathcal{U})}$ on any received question $(L + \mathcal{N}(\mathcal{E}, \mathcal{U}))$.

(2) Bob returns $f|_{L'+\bigoplus\mathcal{U}'}$ on any received question $(L' + \bigoplus\mathcal{U}')$. ☐

Lemma 6.31 (Game3 soundness)**.** *Assume Condition 3.9. Suppose that in every 4^r-coloring of H, at least ε fraction of the edges are monochromatic. Then we have*

$$\mathrm{val}\Big(\mathsf{Game3}[H, k, \ell_1, \ell_2, t, s, \beta]\Big) \leqslant \delta.$$

The proof of this lemma is the most technical part of this section. We first present several important technical lemmas in Section 6.4 and complete the proof of the soundness lemma in Section 6.5.

6.4 Preparatory Lemmas for Soundness Analysis

In this subsection, we fix a 3-uniform hypergraph H, positive integers k, ℓ_1, ℓ_2, t, s, a nonnegative integer q_0 and a parameter $\beta \in (0, 1)$. We consider the game $\mathsf{Game3}[H, k, \ell_1, \ell_2, t, s, \beta]$ as defined in Definition 6.29. The following notational convention will be used throughout Section 6.4.

The first two notational definitions below are parallels of Notations 5.11 and 5.12 in Section 5.3.

Notation 6.32. We use $\mathcal{D}_{\mathsf{in}}$ to denote the joint distribution of $(\mathcal{E}, \mathcal{U}, \mathcal{E}', \mathcal{U}', \mathcal{A})$ sampled by the verifier in $\mathsf{Game3}[H, k, \ell_1, \ell_2, t, s, \beta]$. We also need a variant of $\mathcal{D}_{\mathsf{in}}$ where the five components are expanded into seven components: denote $\mathcal{D}_{\mathsf{in}}^+$ to be the joint distribution of $(\mathcal{E}, \mathcal{U}, L, Q, \mathcal{E}', \mathcal{U}', L')$, where $(\mathcal{E}, \mathcal{U}, \mathcal{E}', \mathcal{U}', \mathcal{A})$ is sampled from $\mathcal{D}_{\mathsf{in}}$ and

$$Q := \sigma(\mathcal{A}_{[q_0]}), \qquad L := \sigma(\mathcal{A}_{[\ell_1]}), \qquad L' = \sigma(\mathcal{A}_{[\ell_1+1]}).$$

We use the symbol $*$ to denote marginalization over the corresponding component. For instance, sampling from the marginal distribution of $(\mathcal{U}, \mathcal{E}')$ under $\mathcal{D}_{\mathsf{in}}$ is denoted by $(*, \mathcal{U}, \mathcal{E}', *, *) \sim \mathcal{D}_{\mathsf{in}}$.

Notation 6.33. We will often condition the distribution $\mathcal{D}_{\mathsf{in}}$ on some subset of the five components $(\mathcal{E}, \mathcal{U}, \mathcal{E}', \mathcal{U}', \mathcal{A})$ and consider the induced joint distribution of the remaining components. For instance, if $\mathcal{E}$ and $\mathcal{U}'$ are fixed, then the conditional joint distribution of $(\mathcal{U}, \mathcal{E}')$ is denoted by

$$\underset{(\mathcal{E}, \mathcal{U}, \mathcal{E}', \mathcal{U}', *) \sim \mathcal{D}_{\mathsf{in}}}{\mathbb{P}} \left[\cdot \,\middle|\, \mathcal{E}, \mathcal{U}' \right].$$

Similar conditioning notation will be used for the seven-component joint distribution $\mathcal{D}_{\mathsf{in}}^+$.

The notions of *Alice-table* and *Bob-table* to be defined in Notation 6.35 are slightly more nuanced than their counterparts in Notation 5.13. In particular, the entries of an Alice-table do not exactly correspond to the possible questions Alice may receive.

Notation 6.34. For any possible sample $(\mathcal{E}, \mathcal{U}, *, *, *) \sim \mathcal{D}_{\mathsf{in}}$, the pair $(\mathcal{E}, \mathcal{U})$ is called an *Alice-packet*. Similarly, a pair $(\mathcal{E}', \mathcal{U}')$ is called a *Bob-packet* if $(*, *, \mathcal{E}', \mathcal{U}', *)$ has nonzero probability under $\mathcal{D}_{\mathsf{in}}$. For any possible sample $(\mathcal{E}, \mathcal{U}, L, *, *, *, *) \sim \mathcal{D}_{\mathsf{in}}^+$, the tuple $(\mathcal{E}, \mathcal{U}, L)$ is called an *unfolded-Alice-question*. Similarly, a tuple $(\mathcal{E}', \mathcal{U}', L')$ is called an *unfolded-Bob-question* if $(*, *, *, *, \mathcal{E}', \mathcal{U}', L')$ has nonzero probability under $\mathcal{D}_{\mathsf{in}}^+$.

Notation 6.35. A table $F[\cdot]$ is called an Alice-table if for each Alice-packet $(\mathcal{E}, \mathcal{U})$ and each linear subspace $L \subseteq \mathcal{V}(\mathcal{E})$, the entry $F[\mathcal{E}, \mathcal{U}, L]$ is either `nil` or a linear functional $F[\mathcal{E}, \mathcal{U}, L] : L \to \mathbb{F}_2$. Similarly, a table $G[\cdot]$ is called a Bob-table if for each Bob-packet $(\mathcal{E}', \mathcal{U}')$ and each linear subspace $L' \subseteq \mathcal{V}(\mathcal{E}')$, the entry $G[\mathcal{E}', \mathcal{U}', L']$ is either `nil` or a linear functional $G[\mathcal{E}', \mathcal{U}', L'] : L' \to \mathbb{F}_2$.

Notation 6.36. Fix an Alice-table $F[\cdot]$ and a Bob-table $G[\cdot]$. For any tuple $(\mathcal{E}, \mathcal{U}, L, *, \mathcal{E}', \mathcal{U}', L')$ with nonzero probability under $\mathcal{D}_{\mathsf{in}}^+$, we define

$$\mathrm{Vrf}_{F,G}(\mathcal{E}, \mathcal{U}, L, \mathcal{E}', \mathcal{U}', L') = \begin{cases} 1, & \text{if } \left. G[\mathcal{E}', \mathcal{U}', L'] \right|_L = F[\mathcal{E}, \mathcal{U}, L], \\ 0, & \text{otherwise.} \end{cases}$$

We will often sample $(\mathcal{E}, \mathcal{U}, L, *, \mathcal{E}', \mathcal{U}', L')$ from $\mathcal{D}_{\mathsf{in}}^+$ conditioned on some subset of its components. For instance, if $\mathcal{E}$ and $\mathcal{U}'$ are fixed, then the conditional expectation of $\mathrm{Vrf}_{F,G}(\mathcal{E}, \mathcal{U}, L, \mathcal{E}', \mathcal{U}', L')$ is denoted by

$$\mathrm{Vrf}_{F,G}(\mathcal{E}, *, *, *, \mathcal{U}', *) = \underset{(\mathcal{E}, \mathcal{U}, L, *, \mathcal{E}', \mathcal{U}', L') \sim \mathcal{D}_{\mathsf{in}}^+}{\mathbb{P}} \left[\left. G[\mathcal{E}', \mathcal{U}', L'] \right|_L = F[\mathcal{E}, \mathcal{U}, L] \,\middle|\, \mathcal{E}, \mathcal{U}' \right].$$

6.4.1 Covering Property of Mixed-Type Subspaces

Definition 6.37. For a fixed Alice-packet $(\mathcal{E}, \mathcal{U})$ and a sequence of vectors

$$B = (a^{(1)}, \ldots, a^{(q_0)}) \in (\mathcal{V}(\mathcal{E}))^{q_0},$$

let $\mathrm{TVD}_{\mathcal{E},\mathcal{U}}(B)$ denote the total variation distance between

(1) the distribution of $\rho(\mathcal{A})$ when $(\mathcal{E},\mathcal{U},*,*,\mathcal{A}) \sim \mathcal{D}_{\mathsf{in}}$ is sampled conditioned on the fixed Alice-packet $(\mathcal{E},\mathcal{U})$ and the event $\{\rho(\mathcal{A}_{[q_0]}) = B\}$ (if this event happens with zero probability, let $\mathrm{TVD}_{\mathcal{E},\mathcal{U}}(B) = 1$); and

(2) the uniform distribution over the set

$$\left\{ \left(a^{(1)}, \ldots, a^{(q_0)}, b^{(q_0+1)}, \ldots, b^{(2\ell_1)}\right) \,\middle|\, b^{(q_0+1)}, \ldots, b^{(2\ell_1)} \in \mathcal{V}(\mathcal{E}) \right\} \subseteq (\mathcal{V}(\mathcal{E}))^{2\ell_1}.$$

Lemma 6.38. *Assume Condition 3.9. For any fixed Alice-packet $(\mathcal{E},\mathcal{U})$, we have*

$$\mathop{\mathbb{E}}_{(\mathcal{E},\mathcal{U},*,*,\mathcal{A})\sim\mathcal{D}_{\mathsf{in}}} \left[\mathrm{TVD}_{\mathcal{E},\mathcal{U}}\big(\rho(\mathcal{A}_{[q_0]})\big) \,\middle|\, \mathcal{E},\mathcal{U} \right] \leqslant t^{-1/3}.$$

Proof. Towards the goal of applying Proposition C.1, we set up some notation. Write

$$\mathcal{E} = (E_1, \ldots, E_t), \qquad \mathcal{U} = (U_1, \ldots, U_t) \qquad \text{and} \qquad q_1 = (2^{2\ell_1} - 1)s.$$

For each $j \in [t]$, define sets $\mathscr{A}_j, \mathscr{B}_j$ and a map $\pi_j : \mathscr{A}_j \to \mathscr{B}_j$ by

$$\mathscr{A}_j = \left\{ \left(a^{(1)}, \ldots, a^{(2\ell_1)}\right) \,\middle|\, a^{(i)} \in \mathbb{F}_2^{E_j} \otimes \mathbb{F}_2^{E_j} \text{ for each } i \in [2\ell_1] \right\} = \left(\mathbb{F}_2^{E_j} \otimes \mathbb{F}_2^{E_j}\right)^{2\ell_1},$$
$$\mathscr{B}_j = \left(\mathbb{F}_2^{E_j} \otimes \mathbb{F}_2^{E_j}\right)^{q_1}, \qquad \text{and} \qquad \pi_j : \left(a^{(1)}, \ldots, a^{(2\ell_1)}\right) \mapsto \left(a^{(1)}, \ldots, a^{(q_1)}\right).$$

For each $j \in [t]$, let μ_j be output distribution of the following process:

1. Sample $E'_j \sim \mathcal{J}_{\mathsf{rev}}(E_j, U_j, \beta, \ell_2)$.

2. Pick vectors $a^{(j,1)}, \ldots, a^{(j,q_1)}$ from $\mathbb{F}_2^{E'_j}$ independently and uniformly at random. Let $A = \left(a^{(j,1)}, \ldots, a^{(j,q_1)}\right)$.

3. Output the sequence

$$\big(\mathrm{Lvec}(A, 2\ell_1, e_1), \ldots, \mathrm{Lvec}(A, 2\ell_1, e_{2\ell_1})\big) \in \mathscr{A}_j,$$

where $e_1, \ldots, e_{2\ell_1} \in \mathbb{F}_2^{2\ell_1}$ are the standard basis vectors.

We can then define a distribution μ over $\prod_{j=1}^{t} \mathscr{A}_j$ as in the statement of Proposition C.1; it is easy to see (using Proposition 6.9) that μ is identical to the distribution of $\rho(\mathcal{A})$ when $(\mathcal{E},\mathcal{U},*,*,\mathcal{A}) \sim \mathcal{D}_{\mathsf{in}}$ is sampled conditioned on $(\mathcal{E},\mathcal{U})$.

Now we can apply Proposition C.1 to $\mathscr{A}_j, \mathscr{B}_j$, π_j and μ_j as defined above. The conclusion of Proposition C.1 says that

$$\mathop{\mathbb{E}}_{(\mathcal{E},\mathcal{U},*,*,\mathcal{A})\sim\mathcal{D}_{\mathsf{in}}} \left[\mathrm{TVD}_{\mathcal{E},\mathcal{U}}\big(\rho(\mathcal{A}_{[q_0]})\big) \,\middle|\, \mathcal{E},\mathcal{U} \right] \leqslant \sqrt{\frac{2^{18k^2 t \ell_1}}{4t}} \leqslant t^{-1/3},$$

where we used Condition 3.9 in the last transition. □

Definition 6.39. For a fixed Alice-packet $(\mathcal{E},\mathcal{U})$, a linear subspace $Q \in \mathsf{Gr}(\mathcal{V}(\mathcal{E}), q_0)$ and a linear subspace $W \subseteq \mathcal{V}(\mathcal{E})$ containing Q, let $\mathrm{TVD}_{\mathcal{E},\mathcal{U},W}(Q)$ be the total variation distance between

(1) the distribution of $\sigma(\mathcal{A}_{[\ell_1]})$ when $(\mathcal{E},\mathcal{U},*,*,\mathcal{A}) \sim \mathcal{D}_{\mathsf{in}}$ is sampled conditioned on the fixed Alice-packet $(\mathcal{E},\mathcal{U})$ and the event $\{\sigma(\mathcal{A}_{[q_0]}) = Q\} \cap \{\sigma(\mathcal{A}_{[\ell_1]}) \subseteq W\}$ (if this event happens with zero probability, let $\mathrm{TVD}_{\mathcal{E},\mathcal{U},W}(Q) = 1$); and

(2) the uniform distribution over the set $\{L \in \mathsf{Gr}(\mathcal{V}(\mathcal{E}), \ell_1) \,|\, Q \subseteq L \subseteq W\}$.

If $W = \mathcal{V}(\mathcal{E})$, we omit the subscript W and write $\mathrm{TVD}_{\mathcal{E},\mathcal{U}}(Q) = \mathrm{TVD}_{\mathcal{E},\mathcal{U},\mathcal{V}(\mathcal{E})}(Q)$.

Lemma 6.40. *Assume Condition 3.9. Fix an Alice-packet $(\mathcal{E},\mathcal{U})$, a linear subspace $Q \in \mathsf{Gr}(\mathcal{V}(\mathcal{E}), q_0)$ and a linear subspace $W \subseteq \mathcal{V}(\mathcal{E})$ containing Q. Letting $w = \operatorname{codim} W$, we always have*

$$\mathrm{TVD}_{\mathcal{E},\mathcal{U},W}(Q) \leqslant 2^{\ell_1(w+1)} \cdot \mathrm{TVD}_{\mathcal{E},\mathcal{U}}(Q).$$

Proof. This is simply due to

$$\mathop{\mathbb{P}}_{L \in \mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)} \left[L \subseteq W \,\middle|\, Q \subseteq L\right] \geqslant 2^{-\ell_1(w+1)}. \qquad \square$$

Lemma 6.41. *Fix an Alice-packet $(\mathcal{E},\mathcal{U})$ and a linear subspace $Q \in \mathsf{Gr}(\mathcal{V}(\mathcal{E}), q_0)$. We have*

$$\mathrm{TVD}_{\mathcal{E},\mathcal{U}}(Q) \leqslant \mathop{\mathbb{E}}_{(\mathcal{E},\mathcal{U},*,*,\mathcal{A})\sim\mathcal{D}_{\mathsf{in}}} \left[\mathrm{TVD}_{\mathcal{E},\mathcal{U}}\big(\rho(\mathcal{A}_{[q_0]})\big) \,\middle|\, \mathcal{E},\mathcal{U} \text{ and } \{\sigma(\mathcal{A}_{[q_0]}) = Q\}\right] + 2^{\ell_1 - 9k^2 t}.$$

Proof. Use the following fact to pass from the uniform distribution over $(\mathcal{V}(\mathcal{E}))^{\ell_1}$ to the uniform distribution over $\mathsf{Gr}(\mathcal{V}(\mathcal{E}), \ell_1)$: for any sequence of vectors $B = (a^{(1)}, \dots, a^{(q_0)}) \in (\mathcal{V}(\mathcal{E}))^{q_0}$ such that $\mathrm{span}(B) = Q$, if $b^{(q_0+1)}, \dots, b^{(\ell_1)} \in \mathcal{V}(\mathcal{E})$ are sampled independently and uniformly at random, the probability that $a^{(1)}, \dots, a^{(q_0)}, b^{(q_0+1)}, \dots, b^{(\ell_1)}$ are not linearly independent is most $2^{\ell_1 - 9k^2 t}$. $\square$

Lemma 6.42. *Assume Condition 3.9. For any fixed Alice-packet $(\mathcal{E},\mathcal{U})$, we have*

$$\mathop{\mathbb{E}}_{(\mathcal{E},\mathcal{U},*,*,\mathcal{A})\sim\mathcal{D}_{\mathsf{in}}} \left[\mathrm{TVD}_{\mathcal{E},\mathcal{U}}\big(\sigma(\mathcal{A}_{[q_0]})\big) \,\middle|\, \mathcal{E},\mathcal{U}\right] \leqslant t^{-1/4}.$$

Proof. Using Lemma 6.41 followed by Lemma 6.38, we have

$$\begin{aligned}\mathop{\mathbb{E}}_{(\mathcal{E},\mathcal{U},*,*,\mathcal{A})\sim\mathcal{D}_{\mathsf{in}}} \left[\mathrm{TVD}_{\mathcal{E},\mathcal{U}}\big(\sigma(\mathcal{A}_{[q_0]})\big) \,\middle|\, \mathcal{E},\mathcal{U}\right] &\leqslant \mathop{\mathbb{E}}_{(\mathcal{E},\mathcal{U},*,*,\mathcal{A})\sim\mathcal{D}_{\mathsf{in}}} \left[\mathrm{TVD}_{\mathcal{E},\mathcal{U}}\big(\rho(\mathcal{A}_{[q_0]})\big) \,\middle|\, \mathcal{E},\mathcal{U}\right] + 2^{\ell_1 - 9k^2 t} \\ &\leqslant t^{-1/3} + 2^{\ell_1 - 9k^2 t} \leqslant t^{-1/4},\end{aligned}$$

where we used Condition 3.9 in the last transition. $\square$

Definition 6.43. For a fixed Alice-packet $(\mathcal{E},\mathcal{U})$, a linear subspace $Q \in \mathsf{Gr}(\mathcal{V}(\mathcal{E}), q_0)$ is said to be *smooth* with respect to $(\mathcal{E},\mathcal{U})$ if $\mathrm{TVD}_{\mathcal{E},\mathcal{U}}(Q) \leqslant t^{-1/8}$.

Corollary 6.44. *Assume Condition 3.9. For any fixed Alice-packet $(\mathcal{E},\mathcal{U})$, we have*

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,Q,*,*,*)\sim\mathcal{D}^+_{\mathsf{in}}} \left[Q \text{ is smooth with respect to } (\mathcal{E},\mathcal{U}) \,\middle|\, \mathcal{E},\mathcal{U}\right] \geqslant 1 - t^{-1/8}.$$

Proposition 6.45. Assume Condition 3.9. Fix an positive integer $\ell \leqslant 2\ell_1$. For any fixed Alice-packet $(\mathcal{E},\mathcal{U})$, if we sample $(\mathcal{E},\mathcal{U},*,*,\mathcal{A}) \sim \mathcal{D}_{\mathsf{in}}$ conditioned on $(\mathcal{E},\mathcal{U})$, the resulting distribution of $\sigma(\mathcal{A}_{[\ell]})$ has total variation distance at most $t^{-1/4}$ to the uniform distribution over $\mathsf{Gr}(\mathcal{V}(\mathcal{E}), \ell)$.

Proof. The proof is entirely analogous to the (combined) proofs of Lemmas 6.38 and 6.41, and is hence omitted. The main difference is that when one invokes Proposition C.1, this time one can take $\mathscr{B}_j$ to be a singleton set for each $j \in [t]$. $\square$

6.4.2 Alice's Existence of Decoding

Lemma 6.46. *Assume Condition 3.9. Let $F[\cdot]$ be an Alice-table and $G[\cdot]$ be a Bob-table. For any Alice-packet $(\mathcal{E},\mathcal{U})$, we have*

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathcal{V}(\mathcal{E}),\ell_1)}\left[F[\mathcal{E},\mathcal{U},L_1]|_{L_1\cap L_2}=F[\mathcal{E},\mathcal{U},L_2]|_{L_1\cap L_2}\right]\geqslant \mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},*,*,*,*)^2-2^{-\ell_1}.$$

Proof. We organize the proof into the following two steps.

Step 1: fixing $(\mathcal{E}',\mathcal{U}',L)$. Fix a Bob-packet $(\mathcal{E}',\mathcal{U}')$ and a subspace $L'\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}'),\ell_1+1)$ such that

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},\mathcal{E}',\mathcal{U}',\mathcal{A})\sim\mathcal{D}_{\mathsf{in}}}\left[\sigma(\mathcal{A}_{[\ell_1+1]})=L'\,\middle|\,\mathcal{E},\mathcal{U},\mathcal{E}',\mathcal{U}'\right]>0.$$

By Definition 6.29, it is clear that if we sample $(\mathcal{E},\mathcal{U},\mathcal{E}',\mathcal{U}',\mathcal{A})\sim\mathcal{D}_{\mathsf{in}}$ conditioned on $\mathcal{E},\mathcal{U},\mathcal{E}',\mathcal{U}'$ and the event $\{\sigma(\mathcal{A}_{[\ell_1+1]})=L'\}$, the resulting distribution of $\sigma(\mathcal{A}_{[\ell_1]})$ is uniform over the set $\mathsf{Gr}(L',\ell_1)$. Therefore, we have

$$\begin{aligned}
&\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},*,\mathcal{E}',\mathcal{U}',L')^2\\
&=\mathop{\mathbb{E}}_{(\mathcal{E},\mathcal{U},\mathcal{E}',\mathcal{U}',\mathcal{A})\sim\mathcal{D}_{\mathsf{in}}}\left[\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},*,\mathcal{E}',\mathcal{U}',L')\,\middle|\,\mathcal{E},\mathcal{U},\mathcal{E}',\mathcal{U}'\text{ and }\{\sigma(\mathcal{A}_{[\ell_1+1]})=L'\}\right]^2\\
&=\mathop{\mathbb{E}}_{L_1,L_2\in\mathsf{Gr}(L',\ell_1)}\left[\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},L_1,\mathcal{E}',\mathcal{U}',L')\,\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},L_2,\mathcal{E}',\mathcal{U}',L')\right].
\end{aligned}\tag{6.12}$$

Expanding the right-hand side of Equation (6.12), we get

$$\begin{aligned}
&\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},*,\mathcal{E}',\mathcal{U}',L')^2\\
&=\mathop{\mathbb{P}}_{L_1,L_2\in\mathsf{Gr}(L',\ell_1)}\left[F[\mathcal{E},\mathcal{U},L_1]=G[\mathcal{E}',\mathcal{U},L']\big|_{L_1}\text{ and }F[\mathcal{E},\mathcal{U},L_2]=G[\mathcal{E}',\mathcal{U},L']\big|_{L_2}\right]\\
&\leqslant\mathop{\mathbb{P}}_{L_1,L_2\in\mathsf{Gr}(L',\ell_1)}\left[F[\mathcal{E},\mathcal{U},L_1]|_{L_1\cap L_2}=F[\mathcal{E},\mathcal{U},L_2]|_{L_1\cap L_2}\right].
\end{aligned}\tag{6.13}$$

Step 2: $(\mathcal{E}',\mathcal{U}',L')$ unfixed. By Proposition 6.45, we know that conditioned on $(\mathcal{E},\mathcal{U})$, the distribution of $\sigma(\mathcal{A}_{[\ell_1+1]})$ in the sample $(\mathcal{E},\mathcal{U},*,*,\mathcal{A})\sim\mathcal{D}_{\mathsf{in}}$ is $t^{-1/4}$-close to uniform over $\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1+1)$. Therefore, we have

$$\begin{aligned}
&\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},*,*,*,*)^2\\
&=\mathop{\mathbb{E}}_{(\mathcal{E},\mathcal{U},\mathcal{E}',\mathcal{U}',\mathcal{A})\sim\mathcal{D}_{\mathsf{in}}}\left[\mathrm{Vrf}_{F,G}\left(\mathcal{E},\mathcal{U},*,\mathcal{E}',\mathcal{U}',\sigma(\mathcal{A}_{[\ell_1+1]})\right)\,\middle|\,\mathcal{E},\mathcal{U}\right]^2 && \text{(by definition)}\\
&\leqslant\mathop{\mathbb{E}}_{(\mathcal{E},\mathcal{U},\mathcal{E}',\mathcal{U}',\mathcal{A})\sim\mathcal{D}_{\mathsf{in}}}\left[\mathrm{Vrf}_{F,G}\left(\mathcal{E},\mathcal{U},*,\mathcal{E}',\mathcal{U}',\sigma(\mathcal{A}_{[\ell_1+1]})\right)^2\,\middle|\,\mathcal{E},\mathcal{U}\right] && \text{(by Cauchy-Schwarz)}\\
&\leqslant\mathop{\mathbb{E}}_{L'\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1+1)}\left[\mathop{\mathbb{E}}_{(\mathcal{E},\mathcal{U},\mathcal{E}',\mathcal{U}',\mathcal{A})\sim\mathcal{D}_{\mathsf{in}}}\left[\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},*,\mathcal{E}',\mathcal{U}',L')^2\,\middle|\,\mathcal{E},\mathcal{U}\text{ and }\{\sigma(\mathcal{A}_{[\ell_1+1]})=L'\}\right]\right]+t^{-1/4}.
\end{aligned}$$

We can now apply Equation (6.13) to the above and get

$$\begin{aligned}
&\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},*,*,*,*)^2\\
&\leqslant\mathop{\mathbb{E}}_{L'\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1+1)}\left[\mathop{\mathbb{P}}_{L_1,L_2\in\mathsf{Gr}(L',\ell_1)}\left[F[\mathcal{E},\mathcal{U},L_1]|_{L_1\cap L_2}=F[\mathcal{E},\mathcal{U},L_2]|_{L_1\cap L_2}\right]\right]+t^{-1/4}
\end{aligned}$$

$$\leqslant \mathop{\mathbb{E}}_{(L_1,L_2)\sim\mathcal{G}(\mathcal{V}(\mathcal{E}),\ell_1)}\left[F[\mathcal{E},\mathcal{U},L_1]|_{L_1\cap L_2} = F[\mathcal{E},\mathcal{U},L_2]|_{L_1\cap L_2}\right] + \frac{1}{2^{\ell_1+1}-1} + t^{-1/4}$$

$$\leqslant \mathop{\mathbb{E}}_{(L_1,L_2)\sim\mathcal{G}(\mathcal{V}(\mathcal{E}),\ell_1)}\left[F[\mathcal{E},\mathcal{U},L_1]|_{L_1\cap L_2} = F[\mathcal{E},\mathcal{U},L_2]|_{L_1\cap L_2}\right] + 2^{-\ell_1}, \qquad \text{(using Condition 3.9)}$$

as desired. $\square$

Definition 6.47. Let $F[\cdot]$ be an Alice table, and let $(\mathcal{E},\mathcal{U})$ be an Alice-packet. For any linear subspace $Q \subseteq \mathcal{V}(\mathcal{E})$ with $\dim Q \leqslant q_0$, we say that the tuple $(\mathcal{E},\mathcal{U},Q)$ is *answerable* with respect to $F[\cdot]$ if $\dim Q = q_0$ and the following statement holds: there exists a linear subspace $W \in \mathsf{Gr}(\mathcal{V}(\mathcal{E}), 9k^2t - q_0)$ and a linear map $f : W \to \mathbb{F}_2$ such that

$$\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)}\left[F[\mathcal{E},\mathcal{U},L] = f|_L \,\middle|\, Q \subseteq L \subseteq W\right] \geqslant \exp(-q_0^3).$$

Lemma 6.48. *Assume Condition 3.9. Let $F[\cdot]$ be an Alice-table and let $G[\cdot]$ be a Bob-table. For any Alice-packet $(\mathcal{E},\mathcal{U})$ such that $\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},*,*,*,*) \geqslant 2^{-q_0/3}$, we have*

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,Q,*,*,*)\sim\mathcal{D}^+_{\mathsf{in}}}\left[(\mathcal{E},\mathcal{U},Q) \text{ is answerable with respect to } F[\cdot] \,\middle|\, \mathcal{E},\mathcal{U}\right] \geqslant 4^{-\ell_1^2}.$$

Proof. Since $\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},*,*,*,*) \geqslant 2^{-q_0/3}$, Lemma 6.46 implies

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathcal{V}(\mathcal{E}),\ell_1)}\left[F[\mathcal{E},\mathcal{U},L_1]|_{L_1\cap L_2} = F[\mathcal{E},\mathcal{U},L_2]|_{L_1\cap L_2}\right] \geqslant 2^{-2q_2/3} - 2^{-\ell_1} \geqslant 2^{-q_2+5},$$

where we used Condition 3.9 in the last transition. Apply Theorem 3.2 to the table $F[\mathcal{E},\mathcal{U},L]$, where L ranges in $\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)$, yields a collection of subspaces $\mathcal{Q} \subseteq \mathsf{Gr}(\mathcal{V}(\mathcal{E}),q_0)$ such that:

(1) $|\mathcal{Q}| \geqslant 2^{-q_0-\ell_1^2}\,|\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)|$;

(2) for each $Q \in \mathcal{Q}$, there exists a linear subspace $W_Q \in \mathsf{Gr}(\mathcal{V}(\mathcal{E}), 9k^2t - q_0)$ and a linear map $f_Q : W_Q \to \mathbb{F}_2$ such that

$$\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)}\left[F[\mathcal{E},\mathcal{U},L] = f_Q|_L \,\middle|\, Q \subseteq L \subseteq W_Q\right] \geqslant \exp(-Cq_0^2) \geqslant \exp(-q_0^3),$$

where C is an absolute constant and the last transition is due to Condition 3.9. By Definition 6.47, this means that for each $Q \in \mathcal{Q}$, the tuple $(\mathcal{E},\mathcal{U},Q)$ is answerable with respect to $F[\cdot]$.

Therefore, it now suffices to show that

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,*,\mathcal{A})\sim\mathcal{D}_{\mathsf{in}}}\left[\sigma(\mathcal{A}_{[q_0]}) \in \mathcal{Q} \,\middle|\, \mathcal{E},\mathcal{U}\right] \geqslant 4^{-\ell_1^2}.$$

But it follows immediately from Proposition 6.45 that

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,*,\mathcal{A})\sim\mathcal{D}_{\mathsf{in}}}\left[\sigma(\mathcal{A}_{[q_0]}) \in \mathcal{Q} \,\middle|\, \mathcal{E},\mathcal{U}\right] \geqslant \mathop{\mathbb{P}}_{Q\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),q_0)}[Q \in \mathcal{Q}] - t^{-1/4} \geqslant 2^{-q_0-\ell_1^2} - t^{-1/4} \geqslant 4^{-\ell_1^2},$$

where we used Condition 3.9 in the last transition. $\square$

6.4.3 Bob's Existence of Decoding

Definition 6.49. Let $G[\cdot]$ be a Bob table, and let $(\mathcal{E}',\mathcal{U}')$ be a Bob-packet with critical index $c \in [t]$. For any linear subspace $Q \in \mathcal{V}(\mathcal{E}')$ with $\dim Q \leqslant q_0$, we say that the tuple $(\mathcal{E}',\mathcal{U}',Q)$ is *pre-answerable* with respect to $G[\cdot]$ if the following statements hold:

1. The tuple $(*,*,*,Q,\mathcal{E}',\mathcal{U}',*)$ has nonzero probability under $\mathcal{D}_{\mathsf{in}}^{+}$. In other words, $Q = \sigma(\mathcal{A})$ for some $\mathcal{A} \in \mathscr{A}_c(\mathcal{E}', q_0, 2^{2\ell_1 - q_0}s)$.

2. There exists a linear subspace $W \subseteq \mathcal{V}(\mathcal{E}')$ with $\operatorname{codim} W \leqslant q_0$ and a linear map $f : W \to \mathbb{F}_2$ such that

$$\mathop{\mathbb{P}}_{(*,*,L,Q,\mathcal{E}',\mathcal{U}',L')\sim\mathcal{D}_{\mathsf{in}}^{+}}\left[G[\mathcal{E}',\mathcal{U}',L']\big|_L = f|_L \,\middle|\, Q,\mathcal{E}',\mathcal{U}' \text{ and } \{L \subseteq W\}\right] \geqslant \exp(-q_0^4).$$

Definition 6.50. Let $\eta \in (0,1)$ be a threshold parameter. An Alice-table $F[\cdot]$ and a Bob-table $G[\cdot]$ is said to be *η-transferable* (with each other) if the following statements hold:

(1) For any unfolded-Alice-question $(\mathcal{E},\mathcal{U},L)$ such that $F[\mathcal{E},\mathcal{U},L] \neq \texttt{nil}$, we have

$$\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},L,*,*,*) \geqslant \eta.$$

(2) For any unfolded-Bob-question $(\mathcal{E}',\mathcal{U}',L')$ such that $G[\mathcal{E},\mathcal{U},L] \neq \texttt{nil}$, we have

$$\mathrm{Vrf}_{F,G}(*,*,*,\mathcal{E}',\mathcal{U}',L') \geqslant \eta.$$

Lemma 6.51. *Assume Condition 3.9. Let $F[\cdot]$ be an Alice-table, and let $G[\cdot]$ be a Bob-table that is 2^{-q_0}-transferable with $F[\cdot]$. Fix an Alice-packet $(\mathcal{E},\mathcal{U})$ and a linear subspace $Q \in \mathsf{Gr}(\mathcal{V}(\mathcal{E}), q_0)$ such that*

(1) Q is smooth with respect to $(\mathcal{E},\mathcal{U})$ (as per Definition 6.43), and

(2) $(\mathcal{E},\mathcal{U},Q)$ is answerable with respect to $F[\cdot]$ (as per Definition 6.47).

Then

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^{+}}\left[(\mathcal{E}',\mathcal{U}',Q) \text{ is pre-answerable with respect to } G[\cdot] \,\middle|\, \mathcal{E},\mathcal{U},Q\right] \geqslant 4^{-\ell_1^2}. \tag{6.14}$$

Proof. We organize the proof into the following three steps.

Step 1: transference. Since $(\mathcal{E},\mathcal{U},Q)$ is answerable with respect to $F[\cdot]$, by Definition 6.47 there exists a linear subspace $W \in \mathsf{Gr}(\mathcal{V}(\mathcal{E}), 9k^2t - q_0)$ and a linear map $f : W \to \mathbb{F}_2$ such that

$$\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)}\left[F[\mathcal{E},\mathcal{U},L] = f|_L \,\middle|\, Q \subseteq L \subseteq W\right] \geqslant \exp(-q_0^3).$$

We then have

$$\begin{aligned}
&\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},L,Q,*,*,*)\sim\mathcal{D}_{\mathsf{in}}^{+}}\left[F[\mathcal{E},\mathcal{U},L] = f|_L \,\middle|\, \mathcal{E},\mathcal{U},Q \text{ and } \{L \subseteq W\}\right] \\
\geqslant\ &\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)}\left[F[\mathcal{E},\mathcal{U},L] = f|_L \,\middle|\, Q \subseteq L \subseteq W\right] - \mathrm{TVD}_{\mathcal{E},\mathcal{U},W}(Q) \qquad \text{(by Definition 6.39)}
\end{aligned}$$

$$\geqslant \exp(-q_0^3) - 2^{\ell_1(q_0+1)}\,\mathrm{TVD}_{\mathcal{E},\mathcal{U}}(Q) \qquad \text{(by Lemma 6.40)}$$
$$\geqslant \exp(-q_0^3) - 2^{\ell_1(q_0+1)}t^{-1/8} \geqslant \exp(-2q_0^3). \qquad \text{(by Definition 6.43 and Condition 3.9)}$$

By the definition of $\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},L,\mathcal{E}',\mathcal{U}',L')$ in Notation 6.36, we have

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},L,Q,\mathcal{E}',\mathcal{U}',L')\sim\mathcal{D}_{\mathsf{in}}^+}\left[G[\mathcal{E}',\mathcal{U}',L']\big|_L = f|_L \,\middle|\, \mathcal{E},\mathcal{U},Q \text{ and } \{L\subseteq W\}\right]$$
$$\geqslant \mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},L,Q,\mathcal{E}',\mathcal{U}',L')\sim\mathcal{D}_{\mathsf{in}}^+}\left[F[\mathcal{E},\mathcal{U},L] = f|_L \text{ and } \mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},L,\mathcal{E}',\mathcal{U}',L') = 1 \,\middle|\, \mathcal{E},\mathcal{U},Q \text{ and } \{L\subseteq W\}\right]$$
$$\geqslant 2^{-q_0}\cdot \mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},L,Q,*,*,*)\sim\mathcal{D}_{\mathsf{in}}^+}\left[F[\mathcal{E},\mathcal{U},L] = f|_L \,\middle|\, \mathcal{E},\mathcal{U},Q \text{ and } \{L\subseteq W\}\right]$$
$$\text{(since } F[\cdot] \text{ and } G[\cdot] \text{ are } 2^{-q_0} \text{ transferable)}$$
$$\geqslant 2^{-q_0}\exp(-2q_0^3) \geqslant \exp(-3q_0^3).$$

Step 2: locating good Bob-packets. For any Bob-packet $(\mathcal{E}',\mathcal{U}')$ such that $(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)$ has nonzero probability under $\mathcal{D}_{\mathsf{in}}^+$, we say it is "good" if

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},L,Q,\mathcal{E}',\mathcal{U}',L')\sim\mathcal{D}_{\mathsf{in}}^+}\left[G[\mathcal{E}',\mathcal{U}',L']\big|_L = f|_L \,\middle|\, \mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}' \text{ and } \{L\subseteq W\}\right] \geqslant \exp(-3q_0^3)/2. \tag{6.15}$$

By the conclusion from Step 1 of the proof, it follows that

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},L,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^+}\left[(\mathcal{E}',\mathcal{U}') \text{ is good} \,\middle|\, \mathcal{E},\mathcal{U},Q \text{ and } \{L\subseteq W\}\right] \geqslant \exp(-3q_0^3)/2. \tag{6.16}$$

Since Q is smooth with respect to $(\mathcal{E},\mathcal{U})$, by Definitions 6.39 and 6.43 we have

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},L,Q,*,*,*)\sim\mathcal{D}_{\mathsf{in}}^+}\left[L\subseteq W \,\middle|\, \mathcal{E},\mathcal{U},Q\right] \geqslant \mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)}\left[L\subseteq W \,\middle|\, Q\subseteq L\right] - \mathrm{TVD}_{\mathcal{E},\mathcal{U}}(Q)$$
$$\geqslant 2^{-\ell_1(q_0+1)} - t^{-1/8} \geqslant 2^{-\ell_1^2},$$

where we used Condition 3.9 in the last transition. Consequently, Equation (6.16) implies

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^+}\left[(\mathcal{E}',\mathcal{U}') \text{ is good} \,\middle|\, \mathcal{E},\mathcal{U},Q\right] \geqslant 2^{-\ell_1^2}\cdot\exp(-3q_0^3)/2 \geqslant 4^{-\ell_1^2}, \tag{6.17}$$

where we used Condition 3.9 in the last transition.

Step 3: extension. By Definition 6.29, the condition Equation (6.15) is equivalent to

$$\mathop{\mathbb{P}}_{\mathcal{A}\in\mathscr{A}_c(\mathcal{E}',\ell_1+1,s')}\left[G[\mathcal{E}',\mathcal{U}',\sigma(\mathcal{A})]\big|_{\sigma(\mathcal{A}_{[\ell_1]})} = f|_{\sigma(\mathcal{A}_{[\ell_1]})} \,\middle|\, \sigma(\mathcal{A}_{[q_0]}) = Q \text{ and } \sigma(\mathcal{A}_{[\ell_1]})\subseteq W\right] \geqslant \exp(-3q_0^3)/2,$$

where c is the critical index of $\mathcal{E}'$ and $s' = 2^{\ell_1-1}s$. By Lemma 6.25, this implies

$$\mathop{\mathbb{P}}_{\mathcal{A}\in\mathscr{A}_c(\mathcal{E}',\ell_1+1,s')}\left[G[\mathcal{E}',\mathcal{U}',\sigma(\mathcal{A})] = g|_{\sigma(\mathcal{A})} \,\middle|\, \sigma(\mathcal{A}_{[q_0]}) = Q \text{ and } \sigma(\mathcal{A})\subseteq W'\right] \geqslant \exp(-3q_0^3)/8 \geqslant \exp(-q_0^4),$$

for some linear subspace $W'\subseteq\mathcal{V}(\mathcal{E}')$ containing $W\cap\mathcal{V}(\mathcal{E}')$ and some linear function $g: W'\to\mathbb{F}_2$. Since the codimension of W' in $\mathcal{V}(\mathcal{E}')$ is at most

$$\dim\mathcal{V}(\mathcal{E}') - \dim(W\cap\mathcal{V}(\mathcal{E}')) \leqslant \dim\mathcal{V}(\mathcal{E}) - \dim W = q_0,$$

this means that $(\mathcal{E}',\mathcal{U}',Q)$ is pre-answerable with respect to $G[\cdot]$ whenever $(\mathcal{E}',\mathcal{U}')$ is good. Therefore, Equation (6.17) implies the desired Equation (6.14). $\square$

Definition 6.52. Let $G[\cdot]$ be a Bob table, and let $(\mathcal{E}',\mathcal{U}')$ be a Bob-packet with critical index $c \in [t]$. For any linear subspace $Q \in \mathcal{V}(\mathcal{E}')$ with $\dim Q \leqslant q_0$, we say that the tuple $(\mathcal{E}',\mathcal{U}',Q)$ is *answerable* with respect to $G[\cdot]$ if $Q \in \mathsf{MGr}_c(\mathcal{E}', q_0, 2^{2\ell_1 - q_0}s)$ and the following statement holds: there exists a linear subspace $W \subseteq \mathcal{V}(\mathcal{E}')$ with $\operatorname{codim} W \leqslant q_0$ and a linear map $f : W \to \mathbb{F}_2$ such that

$$\mathop{\mathbb{P}}_{L' \in \mathsf{MGr}(\mathcal{E}', \ell_1+1, s')} \left[G[\mathcal{E}',\mathcal{U}',L'] = f|_{L'} \,\middle|\, Q \subseteq L' \subseteq W \right] \geqslant \exp(-q_0^5),$$

where $s' = 2^{\ell_1 - 1}s$. More specifically, we also say that $(\mathcal{E}',\mathcal{U}',Q)$ is *answerable with respect to* $G[\cdot]$ *using* (f, W).

Proposition 6.53. Assume Condition 3.9. Let $F[\cdot]$ be an Alice-table, and let $G[\cdot]$ be a Bob-table that is 2^{-q_0}-transferable with $F[\cdot]$. Suppose $(\mathcal{E}',\mathcal{U}')$ is a Bob-packet with critical index c and $Q \in \mathsf{MGr}_c(\mathcal{E}', q_0, 2^{2\ell_1 - q_0}s)$ is a subspace such that

(1) $|E'_c| \geqslant 3k/2$, and

(2) $(\mathcal{E}',\mathcal{U}',Q)$ is pre-answerable with respect to $G[\cdot]$.

Then $(\mathcal{E}',\mathcal{U}',Q)$ is answerable with respect to $G[\cdot]$.

Proof. By Proposition 6.23, the total variation distance between the following two distributions of L' is at most, say, k^{-2}:

(1) sample $L' \in \mathsf{MGr}_c(\mathcal{E}', \ell_1 + 1, 2^{\ell_1 - 1}s)$ conditioned on $Q \subseteq L'$; and

(2) sample $\mathcal{A} \in \mathscr{A}_c(\mathcal{E}', \ell_1 + 1, 2^{\ell_1 - 1}s)$ conditioned on $\sigma(\mathcal{A}_{[q_0]}) = Q$, and let $L' = \sigma(\mathcal{A})$.

We know from Proposition 6.24 that even if we further condition both distributions on $\{L' \subseteq W\}$, for any linear subspace $W \subseteq \mathcal{V}(\mathcal{E}')$ of codimension at most q_0, the resulting distributions are still k^{-1}-close in total variation distance. By definition, this allows one to convert pre-answerablility to answerability. □

Lemma 6.54. *Assume Condition 3.9. Let $F[\cdot]$ be an Alice-table, and let $G[\cdot]$ be a Bob-table that is 2^{-q_0}-transferable with $F[\cdot]$. Then for any Alice-packet $(\mathcal{E},\mathcal{U})$ such that $\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},*,*,*,*) \geqslant 2^{-q_0/3}$, we have*

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*) \sim \mathcal{D}_{\mathsf{in}}^+} \left[(\mathcal{E}',\mathcal{U}',Q) \text{ is answerable with respect to } G[\cdot] \,\middle|\, \mathcal{E},\mathcal{U} \right] \geqslant \exp(-6\ell_1^2).$$

Proof. First of all, combining Corollary 6.44 with Lemmas 6.48 and 6.51 yields

$$\begin{aligned}\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*) \sim \mathcal{D}_{\mathsf{in}}^+} \left[(\mathcal{E}',\mathcal{U}',Q) \text{ is pre-answerable with respect to } G[\cdot] \,\middle|\, \mathcal{E},\mathcal{U} \right] &\geqslant (4^{-\ell_1^2} - t^{-1/8}) \cdot 4^{-\ell_1^2} \\ &\geqslant \exp(-5\ell_1^2), \qquad (6.18)\end{aligned}$$

where we used Condition 3.9 in the last transition.

For each Bob-packet $(\mathcal{E}',\mathcal{U}')$ with critical index $c \in [t]$, we let $\|\mathcal{E}'\|_{\mathsf{min}} := |E_c|$. Using Condition 3.9, it is not hard to see (cf. Propositions 4.7 and 5.5)

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,*,\mathcal{E}',*,*) \sim \mathcal{D}_{\mathsf{in}}^+} \left[\|\mathcal{E}'\|_{\mathsf{min}} \geqslant 3k/2 \right] \geqslant 1 - k^{-1}, \qquad (6.19)$$

and (cf. Proposition 6.23)

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',*,*)\sim\mathcal{D}_{\mathsf{in}}^{+}}\left[Q\in\mathsf{MGr}(\mathcal{E}',q_0,2^{2\ell_1-q_0}s)\;\middle|\;\|\mathcal{E}'\|_{\mathsf{min}}\geqslant 3k/2\right]\geqslant 1-k^{-1}. \tag{6.20}$$

For any Bob-packet $(\mathcal{E}',\mathcal{U}')$ with $\|\mathcal{E}'\|_{\mathsf{min}}\geqslant 3k/2$ and any $Q\in\mathsf{MGr}(\mathcal{E}',q_0,2^{2\ell_1-q_0}s)$, the pre-answerability of $(\mathcal{E}',\mathcal{U}',Q)$ can be converted to answerability, due to Proposition 6.53. Therefore, combining Equations (6.18) to (6.20) yields the desired conclusion. □

6.5 Soundness Analysis

Now, we have all ingredients to analyze the soundness of Game3.

Proof of Lemma 6.31. Assume that

$$\mathrm{val}\Big(\mathsf{Game3}[H,k,\ell_1,\ell_2,t,s,\beta]\Big)>\delta,$$

the goal is to show that

$$\mathrm{val}\Big(\mathsf{Game2c}[H,k,\ell_2,t,s_1,q_0,\beta]\Big)>2^{-q_2/4},$$

which leads to a contradiction to Lemma 5.22.

Recall that the answers returned by Alice and Bob in $\mathsf{Game3}$ are "folded". We define the following folded tables $\widetilde{F},\widetilde{G}$ for Alice and Bob. Suppose that the verifier samples $(\mathcal{E},\mathcal{U},L,*,\mathcal{E}',\mathcal{U}',L')\sim\mathcal{D}_{\mathsf{in}}^{+}$. Then the question sent to Alice is that $(\mathcal{E},\mathcal{U},L+\mathcal{N}(\mathcal{E},\mathcal{U}))$ and she returns a linear function f on $L+\mathcal{N}(\mathcal{E},\mathcal{U})$. Then, we define $\widetilde{F}[\mathcal{E},\mathcal{U},L+\mathcal{N}(\mathcal{E},\mathcal{U})]:=f$. Similarly, $\widetilde{G}$ is defined by $\widetilde{G}[\mathcal{E}',\mathcal{U}',L'+\bigoplus\mathcal{U}']:=g$ if g is the linear function returned by Bob upon receiving the question $(\mathcal{E}',\mathcal{U},L'+\bigoplus\mathcal{U}')$.

Getting unfolded tables. To make use of the decoding lemmas (Lemmas 6.48 and 6.51), we need to define "unfolded" tables $F^*[\cdot],G^*[\cdot]$

$$F^*[\mathcal{E},\mathcal{U},L]:=\widetilde{F}[\mathcal{E},\mathcal{U},L+\mathcal{N}(\mathcal{E},\mathcal{U})]|_L,$$

and

$$G^*[\mathcal{E}',\mathcal{U}',L']:=\widetilde{G}\left[\mathcal{E}',\mathcal{U}',L'+\bigoplus\mathcal{U}'\right]|_{L'},$$

for every possible sample $(\mathcal{E},\mathcal{U},L,*,\mathcal{E}',\mathcal{U}',L')\sim\mathcal{D}_{\mathsf{in}}^{+}$. For other unfolded questions $(\mathcal{E},\mathcal{U},L),(\mathcal{E}',\mathcal{U}',L')$, we simply define $F^*[\mathcal{E},\mathcal{U},L]:=\texttt{nil}$ and $G^*[\mathcal{E}',\mathcal{U}',L']:=\texttt{nil}$.

Getting transferable tables via pruning. To use Lemma 6.51, we also want our table to be transferable. We use the following process to get transferable tables $F[\cdot],G[\cdot]$ from F^*,G^*.

Definition 6.55 (Pruned tables)**.** We use the following process to define our final tables $F[\cdot],G[\cdot]$ used for decoding.

1. Initialize with $F:=F^*$, $G:=G^*$.

2. If there exists an unfolded-Alice-question $(\mathcal{E},\mathcal{U},L)$ such that

$$\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},L,*,*,*) < 2^{-q_0},$$

let $F[\mathcal{E},\mathcal{U},L] := \texttt{nil}$. Similarly, if there exists an unfolded-Bob-question $(\mathcal{E}',\mathcal{U}',L')$ such that

$$\mathrm{Vrf}_{F,G}(*,*,*,\mathcal{E}',\mathcal{U}',L') < 2^{-q_0},$$

let $G[\mathcal{E},\mathcal{U},L] := \texttt{nil}$. Repeat this process until for every unfolded-Alice-question $(\mathcal{E},\mathcal{U},L)$ such that $F[\mathcal{E},\mathcal{U},L] \neq \texttt{nil}$, we have $\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},L,*,*,*) \geqslant 2^{-q_0}$ and for every unfolded-Bob-question $(\mathcal{E}',\mathcal{U}',L')$ such that $G[\mathcal{E}',\mathcal{U}',L'] \neq \texttt{nil}$, we have $\mathrm{Vrf}_{F,G}(*,*,*,\mathcal{E}',\mathcal{U}',L') \geqslant 2^{-q_0}$.

By definitions, $F[\cdot], G[\cdot]$ obtained are 2^{-q_0}-transferable.

Now, we want to argue that when we sample question $(\mathcal{E},\mathcal{U},L,*,\mathcal{E}',\mathcal{U}',L') \sim \mathcal{D}^+_{\mathsf{in}}$, the probability that $\mathrm{Vrf}_{F,G}(\mathcal{E},\mathcal{U},L,\mathcal{E}',\mathcal{U}',L') = 1$ holds is non-negligible. Due to the condition that

$$\mathrm{val}\Big(\mathsf{Game3}[H,k,\ell_1,\ell_2,t,s,\beta]\Big) > \delta,$$

and our definition F^*, G^*, we know that

$$\mathrm{Vrf}_{F^*,G^*}(*,*,*,*,*,*) > \delta.$$

It is easy to see from the definitions of $F[\cdot], G[\cdot]$, the probability of agreement would decrease by at most 2^{-q_0+1} since when each unfolded question of Alice or Bob is pruned, its conditional contribution to the agreement is at most 2^{-q_0}. As a result, we get

$$\mathrm{Vrf}_{F,G}(*,*,*,*,*,*) \geqslant \mathrm{Vrf}_{F^*,G^*}(*,*,*,*,*,*) - 2^{-q_0+1} \geqslant \delta/2, \tag{6.21}$$

under Condition 3.9.

Induced strategy for $\mathsf{Game2c}$**.** Now, we are ready to explain our strategy for

$$\mathsf{Game2c}[H,k,\ell_2,t,s_1,q_0,\beta]$$

and defer the analysis of its success probability later. Assume that the verifier sends $(\mathcal{E},\mathcal{A},\mathcal{U})$ to Alice and sends $(\mathcal{E}',\mathcal{A},\mathcal{U}')$ to Alice and we use Q to denote $\sigma\left(\mathcal{A}_{[q_0]}\right)$ here. The strategies are,

- Let $\vec{\tau} := (\tau_0,\ldots,\tau_{2q_0})$ as $\tau_0,\ldots,\tau_{2q_0}$ determined in Condition 3.9. Note that by Condition 3.9, $\vec{\tau}$ satisfies the constraints of Lemma 3.6. Alice first picks a subspace $Q' \subseteq Q$ of codimension at most 1 *uniformly at random*. If $\mathsf{LIST}^{\vec{\tau}}_{Q'}(F[\mathcal{E},\mathcal{U},\cdot]) \neq \emptyset$, Alice picks a *uniformly random* element $(f,W) \in \mathsf{LIST}^{\vec{\tau}}_{Q'}(F[\mathcal{E},\mathcal{U},\cdot])$ and *randomly* extends f to a linear function on $\mathcal{V}(\mathcal{E})$ as its answer; if $\mathsf{LIST}^{\vec{\tau}}_{Q'}(F[\mathcal{E},\mathcal{U},\cdot])$ is empty, Alice returns an arbitrary answer.

- Bob first picks a *uniformly random* integer $q^* \in \{0,\ldots,q_0\}$, and then, picks an *arbitrary* pair (f,W) where W is a subspace of $\mathcal{V}(\mathcal{E}')$ of codimension q^* and $f \in W^*$ among those satisfy

$$\mathop{\mathbb{P}}_{L' \in \mathsf{MGr}(\mathcal{E}',\ell_1+1,2^{\ell_1-1}\cdot s)}\Big[G[\mathcal{E}',\mathcal{U}',L'] = f|_{L'} \,\Big|\, Q \subseteq L' \subseteq W\Big] \geqslant \exp(-q_0^5),$$

if there exists such a pair; then, Bob also randomly extends the function to a linear function on $\mathcal{V}(\mathcal{E}')$ and returns it. If there does not exist such a pair (f,W), Bob returns an arbitrary answer.

It suffices to prove this strategy succeeds with probability larger than $2^{-q_2/4}$.

Applying decoding lemma for Bob. In this part, we prove that

$$\mathbb{P}_{(*,*,*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^{+}}\left[(\mathcal{E}',\mathcal{U}',Q) \text{ is answerable with respect to } G[\cdot]\right] \geqslant \exp(-7\ell_1^2). \tag{6.22}$$

Recall that answerable is defined in Definition 6.52.

By Markov's inequality and (6.21), we know that

$$\mathbb{P}_{(\mathcal{E},\mathcal{U},*,*,*,*,*)\sim\mathcal{D}_{\mathsf{in}}^{+}}\left[\mathrm{Vrf}(\mathcal{E},\mathcal{U},*,*,*,*) \geqslant 2^{-q_0/3}\right] \geqslant \delta/2 - 2^{-q_0/3} \geqslant \delta/4,$$

under Condition 3.9. After fixing a Alice-packet $(\mathcal{E},\mathcal{U})$ with $\mathrm{Vrf}(\mathcal{E},\mathcal{U},*,*,*,*) \geqslant 2^{-q_0/3}$, Lemma 6.54 gives

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^{+}}\left[(\mathcal{E}',\mathcal{U}',Q) \text{ is answerable with respect to } G[\cdot] \,\middle|\, \mathcal{E},\mathcal{U}\right] \geqslant \exp(-6\ell_1^2).$$

Combining the facts above, we conclude that

$$\mathbb{P}_{(*,*,*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^{+}}\left[(\mathcal{E}',\mathcal{U}',Q) \text{ is answerable with respect to } G[\cdot]\right] \geqslant \delta\cdot\exp(-6\ell_1^2) \geqslant \exp(-7\ell_1^2), \tag{6.23}$$

under Condition 3.9.

We define a partial list of pairs $\mathscr{G}[\cdot]$ as follows, for every possible sample $(\mathcal{E}',\mathcal{U}',Q)$ which is answerable with respect to $G[\cdot]$ and let (f,W) be an arbitrary pair that satisfies (1) $\mathrm{codim}_{\mathcal{V}(\mathcal{E}')}(W) \leqslant q_0$, and $f \in W^*$, (2)

$$\mathop{\mathbb{P}}_{L'\in\mathsf{MGr}(\mathcal{E}',\ell_1+1,2^{\ell_1-1}\cdot s)}\left[G[\mathcal{E}',\mathcal{U}',L'] = f|_{L'} \,\middle|\, Q \subseteq L' \subseteq W\right] \geqslant \exp(-q_0^5),$$

define $\mathscr{G}[\mathcal{E}',\mathcal{U}',Q] := (f,W)$. For other possible samples $(\mathcal{E}',\mathcal{U}',Q)$, define $\mathscr{G}[\mathcal{E}',\mathcal{U}',Q] := \texttt{nil}$. Note that $\mathscr{G}$ actually stands for any possible table that Bob might answer for answerable $(\mathcal{E}',\mathcal{U}',Q)$ in his strategy before randomly extending it to a linear function on $\mathcal{V}(\mathcal{E}')$. Conditioned on any possible list $\mathscr{G}$ of Bob, we want to show that the strategy still succeeds with probability larger than $2^{-q_2/4}$.

Good tuples and excellent tuples. In this part, we first introduce two important definitions. We say a sample $(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*) \sim \mathcal{D}_{\mathsf{in}}^{+}$ is *good* if

1. $(\mathcal{E}',\mathcal{U}',Q)$ is answerable, and let $(f,W) = \mathscr{G}[\mathcal{E}',\mathcal{U}',Q]$;
2. there exists a subspace $Q' \subseteq Q$ with $\mathrm{codim}_Q(Q') \leqslant 1$ such that

$$\mathop{\mathbb{P}}_{L\in\mathsf{MGr}(\mathcal{E}',\ell_1,2^{\ell_1}\cdot s)}\left[F[\mathcal{E},\mathcal{U},L] = f|_{L'} \,\middle|\, Q' \subseteq L \subseteq W\right] \geqslant \exp(-q_0^6).$$

We say a sample $(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*) \sim \mathcal{D}_{\mathsf{in}}^{+}$ is *excellent* if

1. $(\mathcal{E}',\mathcal{U}',Q)$ is answerable, and let $(f,W) = \mathscr{G}[\mathcal{E}',\mathcal{U}',Q]$;
2. there exists a subspace $Q' \subseteq Q$ with $\mathrm{codim}_Q(Q') \leqslant 1$ and a pair $(f',W') \in \mathsf{LIST}_{Q'}^{\vec{\tau}}(F[\mathcal{E},\mathcal{U},\cdot])$ such that $W \subseteq W'$ and $f'|_W = f$.

The first step towards the analysis is to prove that a sample is good with non-negligible probability.

Claim 6.56. $\mathbb{P}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^{+}}$ [$(\mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}')$ *is good*] $\geqslant \exp(-8\ell_1^2)$.

Proof. For every possible sample

$$(*,*,*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^{+}$$

where $(\mathcal{E}',\mathcal{U}',Q)$ is answerable, we know that

1. $\underset{L'\in\mathsf{MGr}(\mathcal{E}',\ell_1+1,2^{\ell_1-1}\cdot s)}{\mathbb{P}}\left[G[\mathcal{E}',\mathcal{U}',L']=f|_{L'}\,\middle|\,Q\subseteq L'\subseteq W\right]\geqslant\exp(-q_0^5)$, where we let $(f,W)=\mathscr{G}[\mathcal{E}',\mathcal{U}',Q]$;

2. for every L' such that $G[\mathcal{E}',\mathcal{U}',L']=f|_{L'}\neq\texttt{nil}$, it holds
$$\mathrm{Vrf}_{F,G}(*,*,*,\mathcal{E}',\mathcal{U}',L')\geqslant 2^{-q_0},$$
since tables $F[\cdot],G[\cdot]$ are 2^{-q_0}-transferable.

Combining the two facts above, we know that in the following distributions defined in two steps (1) sample $L'\in\mathsf{MGr}(\mathcal{E}',\ell_1+1,2^{\ell_1-1}\cdot s)\mid Q\subseteq L'\subseteq W$; (2) sample $(\mathcal{E},\mathcal{U},L,*,\mathcal{E}',\mathcal{U}',L')\sim\mathcal{D}_{\mathsf{in}}^{+}\mid(\mathcal{E}',\mathcal{U}',L')$, it holds

$$\underset{\substack{L'\in\mathsf{MGr}(\mathcal{E}',\ell_1+1,2^{\ell_1-1}\cdot s)\mid Q\subseteq L'\subseteq W\\(\mathcal{E},\mathcal{U},L,*,\mathcal{E}',\mathcal{U}',L')\sim\mathcal{D}_{\mathsf{in}}^{+}\mid(\mathcal{E}',\mathcal{U}',L')}}{\mathbb{E}}\left[F[\mathcal{E},\mathcal{U},L]=G[\mathcal{E}',\mathcal{U}',L']|_L=f|_L\right]\geqslant\exp(-q_0^5-q_0).$$

Observe that the marginal distribution of $(\mathcal{E},\mathcal{U})$ are the same in the distribution

$$(\mathcal{E},\mathcal{U},L,*,\mathcal{E}',\mathcal{U}',L')\sim\mathcal{D}_{\mathsf{in}}^{+}\mid(\mathcal{E}',\mathcal{U}',L'),$$

for every $L'\in\mathsf{MGr}(\mathcal{E}',\ell_1+1,2^{\ell_1-1}\cdot s)$ and for every fixed $L'\in\mathsf{MGr}(\mathcal{E}',\ell_1+1,2^{\ell_1-1}\cdot s)$, $(\mathcal{E},\mathcal{U})$ and L are independent. Then, by Markov's inequality, we get

$$\underset{\substack{L'\in\mathsf{MGr}(\mathcal{E}',\ell_1+1,2^{\ell_1-1}\cdot s)\mid Q\subseteq L'\subseteq W\\(\mathcal{E},\mathcal{U},L,*,\mathcal{E}',\mathcal{U}',L')\sim\mathcal{D}_{\mathsf{in}}^{+}\mid(\mathcal{E},\mathcal{U},\mathcal{E}',\mathcal{U}',L')}}{\mathbb{E}}\left[F[\mathcal{E},\mathcal{U},L]=f|_L\right]\geqslant\exp(-q_0^5-q_0)/2,\tag{6.24}$$

for at least $\exp(-q_0^5-q_0)/2$ fraction of Alice-packets $(\mathcal{E},\mathcal{U})$ sampled from $(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^{+}\mid(\mathcal{E}',\mathcal{U}',Q)$, here we also use the fact that the marginal of $(\mathcal{E},\mathcal{U})$ are the same for every Q once $(\mathcal{E}',\mathcal{U}')$ is fixed.

Now, we prove that for every fixed Alice-packet $(\mathcal{E},\mathcal{U})$ for which (6.24) holds, $(\mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}')$ is good. This concludes that

$$\begin{aligned}&\underset{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^{+}}{\mathbb{P}}\left[\ (\mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}')\text{ is good }\right]\\&\geqslant\underset{(*,*,*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^{+}}{\mathbb{P}}\left[(\mathcal{E}',\mathcal{U}',Q)\text{ is answerable}\right]\cdot\exp(-q_0^5-q_0)/2\\&\geqslant\exp(-7\ell_1^2)\cdot\exp(-q_0^5-q_0)/2\geqslant\exp(-8\ell_1^2).\qquad\text{(using (6.23))}\end{aligned}$$

Note that the distribution in LHS of (6.24) is the following, first uniformly sample an element $L'\in\mathsf{MGr}(\mathcal{E}',\ell_1+1,2^{\ell_1-1}\cdot s)$ conditioned on $Q\subseteq L'\subseteq W$; and then uniformly sample an ℓ_1-dimensional subspace $L\subseteq L'$. First, by the definition of MGr, we know that $L\in\mathsf{MGr}(\mathcal{E}',\ell_1,2^{\ell_1}\cdot s)$. We claim that this distribution over $\mathsf{MGr}(\mathcal{E}',\ell_1,2^{\ell_1}\cdot s)$ is $2^{-c\ell_1}$-close (for some absolute constant $c>0$) to the following one under Condition 3.9:

- let $Q' := Q$ with probability 2^{-q_0}; otherwise, uniformly sample a subspace $Q' \subseteq Q$ conditioned on $\operatorname{codim}_Q(Q') = 1$;
- uniformly sample $L \in \mathsf{MGr}(\mathcal{E}', \ell_1, 2^{\ell_1} \cdot s)$ conditioned on $Q' \subseteq L \subseteq W$.

This can be verified using Proposition 6.19.

Combining this claim and an averaging argument, we conclude that there exists

$$Q' \subseteq Q, \operatorname{codim}_Q(Q') \leqslant 1$$

such that

$$\underset{L \in \mathsf{MGr}(\mathcal{E}', \ell_1, 2^{\ell_1} \cdot s)}{\mathbb{P}} \left[F[\mathcal{E}, \mathcal{U}, L] = f|_{L'} \,\middle|\, Q' \subseteq L \subseteq W \right] \geqslant \exp(-q_0^5 - q_0)/2 - 2^{-c\ell_1} \geqslant \exp(-q_0^6),$$

under our Condition 3.9. □

In the last step of our soundness analysis, we prove that samples are excellent with non-negligible probability. Formally, we prove the following claim.

Claim 6.57. $\mathbb{P}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*) \sim \mathcal{D}_{\mathsf{in}}^{+}}$ [$(\mathcal{E}, \mathcal{U}, Q, \mathcal{E}', \mathcal{U}')$ *is excellent*] $\geqslant \exp(-30\ell_1^2)$.

Analysis of the strategy. Before proving the claim, we first show how to complete the soundness analysis using Claim 6.57. Assume that in $\mathsf{Game2c}[H, k, \ell_2, t, s_1, q_0, \beta]$, the verifier samples a tuple $(\mathcal{E}, \mathcal{U}, \mathcal{A}, \mathcal{E}', \mathcal{U}')$ and it sends $(\mathcal{E}, \mathcal{U}, \mathcal{A})$ to Alice and sends $(\mathcal{E}', \mathcal{U}', \mathcal{A})$ to Bob. Let $Q := \sigma(\mathcal{A}_{[q_0]})$. We prove the following claim on the success probability of our strategy on an excellent tuple.

Claim 6.58. *If* $(\mathcal{E}, \mathcal{U}, Q, \mathcal{E}', \mathcal{U}')$ *is excellent, the strategy wins with probability at least* $\exp(-3Cq_0^2\ell_1)$.

Proof. Note that by our assumption and strategies, upon getting the tuple $(\mathcal{E}, \mathcal{U}, Q, \mathcal{E}', \mathcal{U}')$, Bob picks $(f_B, W_B) = \mathscr{G}[\mathcal{E}', \mathcal{U}', Q]$ since $(\mathcal{E}', \mathcal{U}', Q)$ is answerable and randomly extends it to a linear function $\widetilde{f}_B$ on $\mathcal{V}(\mathcal{E}')$. Alice first randomly samples a subspace $Q' \subseteq Q$ of codimension at most 1 and then randomly picks a pair $(f_A, W_A) \in \mathsf{LIST}_{Q'}^{\vec{\tau}}(F[\mathcal{E}, \mathcal{U}, \cdot])$ (which is non-empty since $(\mathcal{E}, \mathcal{U}, Q, \mathcal{E}', \mathcal{U}')$ is excellent), also randomly extends it to a linear function $\widetilde{f}_A$ on $\mathcal{V}(\mathcal{E})$. The strategy wins when the all following events happen,

1. $W_B \subseteq W_A$, $f_A|_{W_B} = f_B$;
2. $\widetilde{f}_A$ is SNR-satisfying;
3. $\widetilde{f}_A|_{\mathcal{V}(\mathcal{E}')} = \widetilde{f}_B$.

Due to Lemma 3.6 and the fact that $(\mathcal{E}, \mathcal{U}, Q, \mathcal{E}', \mathcal{U}')$ is excellent, with probability at least $2^{-q_0} \cdot \exp(-Cq_0^2\ell_1) \geqslant \exp(-2Cq_0^2\ell_1)$, we have $W_B \subseteq W_A$ and $f_A|_{W_B} = f_B$.

Next, we prove that for every $(f_A, W_A) \in \mathsf{LIST}_{Q'}^{\vec{\tau}}(F[\mathcal{E}, \mathcal{U}, \cdot])$, there exists a good linear function $f' \in \mathcal{V}(\mathcal{E})^*$ such that $f'|_{W_A} = f_A$ and $f'|_{\mathcal{N}(\mathcal{E},\mathcal{U})}$ is SNR-satisfying, which implies that conditioned on the first event, the second event happens with probability at least 2^{-2q_0} since $\operatorname{codim}_{\mathcal{V}(\mathcal{E})}(W_A) \leqslant 2q_0$. Let $w_A = \operatorname{codim}_{\mathcal{V}(\mathcal{E})}(W_A) \leqslant 2q_0$. By the definition of $\mathsf{LIST}_{Q'}^{\vec{\tau}}(F[\mathcal{E}, \mathcal{U}, \cdot])$, we have

$$\underset{L \in \mathsf{Gr}(\mathcal{V}(\mathcal{E}), \ell_1)}{\mathbb{P}} \left[F[\mathcal{E}, \mathcal{U}, L] = f_A|_L \mid Q' \subseteq L \subseteq W_A \right] \geqslant \tau_{w_A}.$$

As a result, we get the same lower bound for the unfolded table $\widetilde{F}$,

$$\underset{L\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)}{\mathbb{P}}\left[\widetilde{F}[\mathcal{E},\mathcal{U},L+\mathcal{N}(\mathcal{E},\mathcal{U})]|_L = f_A|_L \mid Q'\subseteq L\subseteq W_A\right]\geqslant \tau_{w_A}.$$

By averaging argument, there exists a subspace $\widetilde{L}=L+\mathcal{N}(\mathcal{E},\mathcal{U})$, where $L\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)$ such that

$$\mathbb{P}_{L\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)}\left[\widetilde{F}[\mathcal{E},\mathcal{U},\widetilde{L}]|_L = f_A|_L \mid Q'\subseteq L\subseteq \widetilde{L}\cap W_A\wedge L+\mathcal{N}(\mathcal{E},\mathcal{U})=\widetilde{L}\right]\geqslant \tau_{w_A}/2.$$

Note that $\widetilde{F}[\mathcal{E},\mathcal{U},\widetilde{L}]\in\widetilde{L}^*$ is a linear function and its restriction $\widetilde{F}[\mathcal{E},\mathcal{U},\widetilde{L}]|_{\mathcal{N}(\mathcal{E},\mathcal{U})}$ is SNR-satisfying. If f_A does not admit such good extension, we then have $f_A|_{W_A\cap\widetilde{L}}\neq\widetilde{F}[\mathcal{E},\mathcal{U},\widetilde{L}]|_{W_A\cap\widetilde{L}}$, which means

$$\begin{aligned}&\underset{L\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)}{\mathbb{P}}\left[\widetilde{F}[\mathcal{E},\mathcal{U},\widetilde{L}]|_L = f_A|_L \mid Q'\subseteq L\subseteq \widetilde{L}\cap W_A\wedge L+\mathcal{N}(\mathcal{E},\mathcal{U})=\widetilde{L}\right]\\ &\leqslant \frac{\underset{L\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)}{\mathbb{P}}\left[\widetilde{F}[\mathcal{E},\mathcal{U},\widetilde{L}]|_L = f_A|_L \mid Q'\subseteq L\subseteq \widetilde{L}\cap W_A\right]}{\underset{L\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)}{\mathbb{P}}\left[L+\mathcal{N}(\mathcal{E},\mathcal{U})=\widetilde{L}\mid Q'\subseteq L\subseteq \widetilde{L}\cap W_A\right]}\leqslant 2^{-\Omega(\ell_1)}<\tau_{w_A}/2\end{aligned}$$

under Condition 3.9, leading to a contradiction.

Finally, it is easy to see that conditioned on the first two events, the last happened with probability at least 2^{-q_0} since $\mathrm{codim}_{\mathcal{V}(\mathcal{E}')}(W_B)\leqslant q_0$. Combining facts above, we conclude that the strategy wins with probability at least $\exp(-2Cq_0^2\ell_1)2^{-3q_0}\geqslant\exp(-3Cq_0^2\ell_1)$ under Condition 3.9. □

Combining Claims 6.57 and 6.58, we conclude that the strategy wins with probability at least

$$\exp(-30\ell_1^2)\cdot\exp(-3Cq_0^2\ell_1)>2^{-q_2/4},$$

under Condition 3.9, giving a contradiction to Lemma 5.22. □

6.6 Proof of Claim 6.57

In this section, we prove Claim 6.57.

Proof of Claim 6.57. We first introduce a useful definition.

Definition 6.59. Given an integer $k^*\in\{0,\dots,3k\}$ a tuple $(\mathcal{E},\mathcal{U},Q)$, we define $\mathcal{R}_{\mathcal{E},\mathcal{U},Q,k^*}$ to be the set of all Bob-packets $(\mathcal{E}',\mathcal{U}')$ that satisfies (1) $\|\mathcal{E}'\|_{\min}=k^*$, (2) the tuple $(\mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}')$ is in the support of the distribution $(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^+$. We say a tuple $(\mathcal{E},\mathcal{U},Q,k^*)$ is redundant if

$$|\mathcal{R}_{\mathcal{E},\mathcal{U},Q,k^*}|\geqslant\exp(Ck^4/2).$$

We prove that if $(\mathcal{E},\mathcal{U},Q)$ satisfies all following conditions,

1. $\mathbb{P}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^+|(\mathcal{E},\mathcal{U},Q)}\left[(\mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}')\text{ is good}\right]\geqslant\exp(-8\ell_1^2)/2$;
2. $\mathbb{P}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^+|(\mathcal{E},\mathcal{U},Q)}\left[\|\mathcal{E}'\|_{\min}\geqslant 3k/2\right]\geqslant 1-1/\sqrt{k}$;
3. for every $k^*\in[k]$, $k^*\geqslant 3k/2$, $(\mathcal{E},\mathcal{U},Q,k^*)$ is redundant.

we then have

$$\mathbb{P}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^+|(\mathcal{E},\mathcal{U},Q)}\left[(\mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}')\text{ is excellent}\right]\geqslant\exp(-20\ell_1^2).\tag{6.25}$$

We first show that a sample $(\mathcal{E},\mathcal{U},*,Q,*,*,*)\sim\mathcal{D}_{\mathsf{in}}^+$ satisfies the above three conditions with probability at least $\exp(-10\ell_1^2)$.

First, by Claim 6.56 and a Markov's inequality, we get

$$\mathbb{P}_{(\mathcal{E},\mathcal{U},*,Q,*,*,*)\sim\mathcal{D}_{\mathsf{in}}^+}\left[\mathbb{P}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^+|(\mathcal{E},\mathcal{U},Q)}\left[(\mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}')\text{ is good}\right]\geqslant\exp(-8\ell_1^2)/2\right]\tag{6.26}$$
$$\geqslant\exp(-8\ell_1^2)/2.\tag{6.27}$$

Also, as shown in the proof of Lemma 6.54, for every Alice-packet $(\mathcal{E},\mathcal{U})$, we have

$$\mathbb{P}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^+}\left[\|\mathcal{E}'\|_{\mathsf{min}}\geqslant 3k/2\right]\geqslant 1-k^{-1}.$$

By Markov's inequality, we get

$$\mathbb{P}_{(\mathcal{E},\mathcal{U},*,Q,*,*,*)\sim\mathcal{D}_{\mathsf{in}}^+}\left[\mathbb{P}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^+|(\mathcal{E},\mathcal{U},Q)}\left[\|\mathcal{E}'\|_{\mathsf{min}}\geqslant 3k/2\right]\geqslant 1-1/\sqrt{k}\right]\tag{6.28}$$
$$\geqslant 1-1/\sqrt{k}.\tag{6.29}$$

For the third condition, it is not hard to prove that for every integer $k^*\in[3k], k^*\geqslant 3k/2$

$$\mathbb{P}_{(\mathcal{E},\mathcal{U},*,Q,*,*,*)\sim\mathcal{D}_{\mathsf{in}}^+}\left[(\mathcal{E},\mathcal{U},Q,k^*)\text{ is redundant}\right]\geqslant 1-\exp(-\Omega(k^2)).$$

by a Chernoff-type argument under Condition 3.9.

Using a union bound over k^*, we conclude that the third condition is satisfied with probability at least $1-k\cdot\exp(-\Omega(k^2))\geqslant 1-\exp(-\Omega(k^2))$. Combining (6.26), (6.28), the statement above, and a union bound, we know that $(\mathcal{E},\mathcal{U},*,Q,*,*,*)\sim\mathcal{D}_{\mathsf{in}}^+$ satisfies all three conditions with probability at least

$$\exp(-8\ell_1^2)/2-1/\sqrt{k}-\exp(-\Omega(k^2))\geqslant\exp(-10\ell_1^2),$$

under Condition 3.9.

It suffices to prove that (6.25) holds for every $(\mathcal{E},\mathcal{U},Q)$ that satisfies the three conditions above. We first introduce a technical lemma used in the remaining proof. We defer the proof of this lemma to Section 6.7.

Lemma 6.60. *Assume Condition 3.9. Let $F[\cdot]$, $G[\cdot]$ be the pruned tables of Alice and Bob (see Definition 6.55). Fix an Alice-packet $(\mathcal{E},\mathcal{U})$ and a linear subspace $Q\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),q_0)$. Suppose $m\geqslant\exp(1000\cdot k^4)$ is a positive integer and $c_1,\ldots,c_m\in[t]$ are pairwise distinct indices. For each $i\in[m]$, assume $(\mathcal{E}_i',\mathcal{U}_i')$ is a c_i-critical Bob-packet such that the following hold:*

(1) For every $i\in[m]$, $(\mathcal{E},\mathcal{U},Q,\mathcal{E}_i',\mathcal{U}_i')$ is good; let (f_i,W_i) denote $\mathscr{G}[\mathcal{E}_i',\mathcal{U}_i',Q]$.

(2) Writing $\mathcal{E}_i=(E_{i,1},\ldots,E_{i,t})$ for each $i\in[m]$, we have $|E_{1,c_1}|=|E_{2,c_2}|=\cdots=|E_{m,c_m}|$.

Then there exists linear subspaces $W\subseteq\mathcal{V}(\mathcal{E})$ with $\operatorname{codim}_{\mathcal{V}(\mathcal{E})}(W)\leqslant 2q_0$ and $Q'\subseteq Q$ and a linear map $f:W\to\mathbb{F}_2$ such that

$$\mathbb{P}_{L\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)}\left[F[\mathcal{E},\mathcal{U},L]=f|_L\,\middle|\,Q'\subseteq L\subseteq W\right]\geqslant\exp(-100\cdot q_0^7).$$

Furthermore, for some $i\in[m]$ we have $W_i\subseteq W$ and $f|_{W_i}=f_i$.

Now, we complete the proof using the lemma above. By a union bound and the three conditions, we get

$$\begin{aligned}\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^{+}|(\mathcal{E},\mathcal{U},Q)}\left[(\mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}')\text{ is good}\wedge\|\mathcal{E}'\|_{\min}\geqslant 3k/2\right]&\geqslant\exp(-8\ell_1^2)/2-1/\sqrt{k}\\&\geqslant\exp(-9\ell_1^2),\end{aligned}$$

under Condition 3.9. Let $\mathscr{P}$ be the event that

$$\|\mathcal{E}'\|_{\min}:=k^*\geqslant 3k/2\wedge\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^{+}|(\mathcal{E},\mathcal{U},Q)}\left[(\mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}')\text{ is good}\mid\|\mathcal{E}'\|_{\min}=k^*\right]\geqslant\exp(-9\ell_1^2)/2$$

By Markov's inequality, we get

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^{+}|(\mathcal{E},\mathcal{U},Q)}[\mathscr{P}]\geqslant\exp(-9\ell_1^2)/2.\tag{6.30}$$

Furthermore, we claim that for every fixed integer $k^*\geqslant 3k/2$, if

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^{+}|(\mathcal{E},\mathcal{U},Q)}\left[(\mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}')\text{ is good}\mid\|\mathcal{E}'\|_{\min}=k^*\right]\geqslant\exp(-9\ell_1^2)/2,\tag{6.31}$$

we then have

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^{+}|(\mathcal{E},\mathcal{U},Q)}\left[(\mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}')\text{ is excellent}\mid\|\mathcal{E}'\|_{\min}=k^*\right]\geqslant\exp(-9\ell_1^2)/4.\tag{6.32}$$

We define $\mathcal{R}'_{\mathcal{E},\mathcal{U},Q,k^*}\subseteq\mathcal{R}_{\mathcal{E},\mathcal{U},Q,k^*}$ to be the Bob-packets $(\mathcal{E}',\mathcal{U}')$ such that $(\mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}')$ is good, and let $\mathcal{R}''_{\mathcal{E},\mathcal{U},Q,k^*}\subseteq\mathcal{R}_{\mathcal{E},\mathcal{U},Q,k^*}$ to be the Bob-packets $(\mathcal{E}',\mathcal{U}')$ such that $(\mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}')$ is excellent. Note that (6.31) implies that $|\mathcal{R}'_{\mathcal{E},\mathcal{U},Q,k^*}|\geqslant\exp(Ck^4/2-9\ell_1^2)/2\geqslant\exp(Ck^4/4)$ under Condition 3.9. Applying Lemma 6.60 iteratively, we get $|\mathcal{R}''_{\mathcal{E},\mathcal{U},Q,k^*}|\geqslant|\mathcal{R}'_{\mathcal{E},\mathcal{U},Q,k^*}|-\exp(1000k^4)\geqslant|\mathcal{R}'_{\mathcal{E},\mathcal{U},Q,k^*}|/2$ since if $|\mathcal{R}'_{\mathcal{E},\mathcal{U},Q,k^*}|\geqslant\exp(1000k^4)$ there exists a Bob-packet $(\mathcal{E}',\mathcal{U}')\in\mathcal{R}'_{\mathcal{E},\mathcal{U},Q,k^*}$ such that $(\mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}')$ is excellent. This completes the proof of (6.32) and gives the corollary that

$$\mathop{\mathbb{P}}_{(\mathcal{E},\mathcal{U},*,Q,\mathcal{E}',\mathcal{U}',*)\sim\mathcal{D}_{\mathsf{in}}^{+}|(\mathcal{E},\mathcal{U},Q)}\left[(\mathcal{E},\mathcal{U},Q,\mathcal{E}',\mathcal{U}')\text{ is excellent}\mid\mathscr{P}\right]\geqslant\exp(-9\ell_1^2)/4.\tag{6.33}$$

Combining (6.30) and (6.33), this proves (6.25) and then the whole claim.

6.7 Proof of Lemma 6.60

In this section, we prove Lemma 6.60, which completes the whole proof of Lemma 6.31.

Proof of Lemma 6.60. Since for every $i\in[m]$, $(\mathcal{E},\mathcal{U},Q,\mathcal{E}'_i,\mathcal{U}'_i)$ is good, we know that there exists a subspace $Q_i\subseteq Q$ with $\operatorname{codim}_Q(Q_i)\leqslant 1$ such that

$$\mathop{\mathbb{P}}_{L\in\mathsf{MGr}(\mathcal{E}'_i,\ell_1,2^{\ell_1}\cdot s)}\left[F[\mathcal{E},\mathcal{U},L]=f_i|_{L'}\,\middle|\,Q_i\subseteq L\subseteq W\right]\geqslant\exp(-q_0^6)\tag{6.34}$$

By an averaging argument, there exists a subspace $Q'\subseteq Q$ with $\operatorname{codim}_Q(Q')\leqslant 1$ and an integer $w\in\{0,\ldots,q_0\}$ such that

$$\left|\{i:Q_i=Q'\wedge\operatorname{codim}_{\mathcal{V}(\mathcal{E}'_i)}(W_i)=w\}\right|\geqslant 2^{-2q_0^2}\cdot m.$$

Let $m'=\exp(-q_0^7)/(10\cdot\lambda\cdot 2^{-(D-q^*)w})$, where $\lambda>\exp(-k^2)$ is a real number to be determined later. Applying Lemma 3.8 to the collection of subspaces W_i that satisfy $Q_i=Q'$ and $\operatorname{codim}_{\mathcal{V}(\mathcal{E}'_i)}(W_i)=w$, we know that there exist distinct indices $i_1,\ldots,i_{m'}\in[m]$ and a subspace $W\subseteq\mathcal{V}(\mathcal{E}'_i)$ such that

1. for every $j \in [m']$, $Q_{i_j} = Q'$, $\mathrm{codim}_{\mathcal{V}(\mathcal{E}'_{i_j})}(W_{i_j}) = w$;
2. $W_1 + \cdots + W_{i_{m'}} = W$;
3. there exists an integer $w' > 0$ such that (1) for every $j \in [m']$, $\mathrm{codim}_W(W_{i_j}) = w'$ and (2) for every distinct $j, j' \in [m']$, $\mathrm{codim}_W(W_{i_j} \cap W_{i_{j'}}) = 2w'$.

Note that we can get $\mathrm{codim}_{\mathcal{V}(\mathcal{E})}(W) \leqslant 2w \leqslant 2q_0$ as a direct consequence since (1) $m' \geqslant 2$, and (2)

$$\mathrm{codim}_{\mathcal{V}(\mathcal{E})}(W) \leqslant \mathrm{codim}_{\mathcal{V}(\mathcal{E})}(W_{i_1} + W_{i_2}) \leqslant \mathrm{codim}_{\mathcal{V}(\mathcal{E}'_{i_1})}(W_{i_1}) + \mathrm{codim}_{\mathcal{V}(\mathcal{E}'_{i_2})}(W_{i_2}) \leqslant 2w,$$

here the second inequality comes from the fact that $c_{i_1} \neq c_{i_2}$ and thus $\mathcal{V}(\mathcal{E}'_{i_1}) + \mathcal{V}(\mathcal{E}'_{i_2}) = \mathcal{V}(\mathcal{E})$.

In the remaining proofs, we show that there exists a linear function $f \in W^*$ and $j \in [m']$, such that

$$\mathop{\mathbb{P}}_{L \in \mathsf{Gr}(\mathcal{V}(\mathcal{E}), \ell_1)} \left[F[\mathcal{E}, \mathcal{U}, L] = f|_L \,\middle|\, Q' \subseteq L \subseteq W \right] \geqslant \exp(-100 q_0^7). \tag{6.35}$$

and $f_{i_j} = f|_{W_{i_j}}$. The plan is to first show that there exists a $f \in W^*$ that satisfies (6.35) and then prove the full statement.

Existence of decoding. To prove that there exists a linear function $f \in W^*$ such that

$$\mathop{\mathbb{P}}_{L \in \mathsf{Gr}(\mathcal{V}(\mathcal{E}), \ell_1)} \left[F[\mathcal{E}, \mathcal{U}, L] = f|_L \,\middle|\, Q' \subseteq L \subseteq W \right] \geqslant \exp(-100 q_0^7),$$

we consider the *Grassmann linearity test* proposed by [KMS25, DKK+25]. To be more specific, we consider the following test over W. Sample two subspaces $Q' \subseteq L_1, L_2 \subseteq W$ of dimension ℓ_1 conditioned on $\dim(L_1 \cap L_2) = d$, where we define $d = \lfloor \ell_1/10 \rfloor$ and $D = 2\ell_1 - d$, and check if $F[\mathcal{E}, \mathcal{U}, L_1]|_{L_1 \cap L_2} = F[\mathcal{E}, \mathcal{U}, L_2]|_{L_1 \cap L_2}$. We define $\mathsf{agree}(\mathcal{E}, \mathcal{U}, W, Q')$ to denote the probability that the test passes. [KMS25, DKK+25] proved the following decoding lemma for this linearity test.

Lemma 6.61 ([KMS25, DKK+25])**.** *Under Condition 3.9, if* $\mathsf{agree}(\mathcal{E}, \mathcal{U}, W, Q') \geqslant \eta$ *then there exists a linear function* $f \in W^*$ *such that*

$$\mathbb{P}_{L \in \mathsf{Gr}(\mathcal{V}(\mathcal{E}), \ell_1)} \left[F[\mathcal{E}, \mathcal{U}, L] = f|_L \mid Q' \subseteq L \subseteq W \right] \geqslant \eta^3/1000.$$

To avoid a dependency issue in the future, we consider the following process of sampling a matrix $\mathcal{M} \in (\mathcal{V}(\mathcal{E}))^D$, where each row $x \in [D], \mathcal{M}_x$ is a vector in $\mathcal{V}(\mathcal{E})$:

- for each $x \in [D] \setminus [q^*]$, independently and uniformly sample $\mathcal{M}_x$ from W;
- let $\{y_1, \ldots, y_{q^*}\}$ be an arbitrary fixed basis of Q', let $\mathcal{M}_x := y_x$ for every $x \in [q^*]$.

After getting the matrix $\mathcal{M}$, we define $L_1 := \mathrm{span}\{\mathcal{M}_x\}_{x \in [\ell_1]}, L_2 := \mathrm{span}\{\mathcal{M}_x\}_{x \in [q_0] \cup ([D] \setminus [\ell_1])}$. We use $\mathcal{P}$ to denote the joint distribution of L_1, L_2. It is easy to see that $\mathcal{P}$ is $\exp(-\Omega(k))$-close to the distribution in the definition of $\mathsf{agree}(\mathcal{E}, \mathcal{U}, W, Q')$. Thus, it suffices to prove that

$$\mathbb{P}_{(L_1, L_2) \sim \mathcal{P}} \left[F[\mathcal{E}, \mathcal{U}, L_1]|_{L_1 \cap L_2} = F[\mathcal{E}, \mathcal{U}, L_2]|_{L_1 \cap L_2} \right] \geqslant \exp(-5 q_0^7). \tag{6.36}$$

First-moment calculations. For every $j \in [m']$, let $\mathscr{X}_j$ denote the event that $L_1, L_2 \subseteq W_{i_j}$; it is easy to see that

$$\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}[\mathscr{X}_j] = 2^{-(D-q^*)w'}. \tag{6.37}$$

We let $\mathscr{Y}_j$ denote the event that $L_1 + L_2 \in \mathsf{MGr}(\mathcal{E}'_{i_j}, D, 2^d \cdot s)$. We want to calculate the probability that

$$\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}[\mathscr{Y}_j|\mathscr{X}_j].$$

To this end, we define $\lambda := \mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{Y}_j \mid L_1, L_2 \subseteq \mathcal{V}(\mathcal{E}'_{i_j})\right]$. Instead of explicitly computing the value $\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}[\mathscr{Y}_j|\mathscr{X}_j]$, we prove that

$$1/3 \leqslant \mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}[\mathscr{Y}_j|\mathscr{X}_j]/\lambda \leqslant 3. \tag{6.38}$$

Note that

$$\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}[\mathscr{Y}_j|\mathscr{X}_j] = \frac{\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{Y}_j \wedge \mathscr{X}_j \mid L_1, L_2 \subseteq \mathcal{V}(\mathcal{E}'_{i_j})\right]}{\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{X}_j \mid L_1, L_2 \subseteq \mathcal{V}(\mathcal{E}'_{i_j})\right]}.$$

It is easy to see that the denominator equals

$$\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{X}_j \mid L_1, L_2 \subseteq \mathcal{V}(\mathcal{E}'_{i_j})\right] = 2^{-(D-q^*)\cdot w}.$$

Also, we have

$$\begin{aligned}
&\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{Y}_j \wedge \mathscr{X}_j \mid L_1, L_2 \subseteq \mathcal{V}(\mathcal{E}'_{i_j})\right] \\
&= \mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{Y}_j \mid L_1, L_2 \subseteq \mathcal{V}(\mathcal{E}'_{i_j})\right] \cdot \mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{X}_j \mid L_1, L_2 \subseteq \mathcal{V}(\mathcal{E}'_{i_j}) \wedge \mathscr{Y}_j\right] \\
&= \lambda \cdot \mathbb{P}_{L\in\mathsf{MGr}(\mathcal{E}'_{i_j}, D, 2^{D-q}\cdot s)}\left[L \subseteq W_{i_j} \mid Q' \subseteq L\right].
\end{aligned}$$

By Corollary 6.21 and Proposition 6.23, we get

$$2^{-w(D-q^*)-1} - \exp(-\Omega(k)) \leqslant \mathbb{P}_{L\in\mathsf{MGr}(\mathcal{E}'_{i_j}, D, 2^{D-q}\cdot s)}\left[L \subseteq W_{i_j} \mid Q' \subseteq L\right] \leqslant 2^{-w(D-q^*)+1} + \exp(-\Omega(k)),$$

which completes the proof of (6.38) under Condition 3.9.

We define $\mathscr{Z}_j$ to be the following event that $F[\mathcal{E}, \mathcal{U}, L_1]|_{L_1\cap L_2} = F[\mathcal{E}, \mathcal{U}, L_2]|_{L_1\cap L_2}$, and we want to calculate the conditional probability

$$\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}[\mathscr{Z}_j \mid \mathscr{X}_j \wedge \mathscr{Y}_j].$$

Note that the distribution $(L_1, L_2) \sim \mathcal{P}$ conditioned on $\mathscr{X}_j \wedge \mathscr{Y}_j$ is $2^{-\Omega(\ell_1)}$-close to the following one (actually the two distributions are equal if we further impose $L_1 + L_2 = L$)

- uniformly sample a subspace $L \in \mathsf{MGr}(\mathcal{E}'_{i_j}, D, 2^{D-q^*} \cdot s)$ conditioned on $Q' \subseteq L \subseteq W_{i_j}$;
- independently sample two subspaces $L_1, L_2 \in \mathsf{MGr}(\mathcal{E}'_{i_j}, \ell_1, 2^{\ell_1} \cdot s)$ conditioned on $Q' \subseteq L_1, L_2 \subseteq L$.

Thus, we conclude that

$$\begin{aligned}
&\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{Z}_j \mid \mathscr{X}_j \wedge \mathscr{Y}_j\right] + 2^{-\Omega(\ell_1)}\\
&\geqslant \mathbb{P}_{\substack{L\in\mathsf{MGr}(\mathcal{E}'_{i_j},D,2^{D-q^*}\cdot s)\\ L_1,L_2\in\mathsf{MGr}(\mathcal{E}'_{i_j},\ell_1,2^{\ell_1}\cdot s)}}\left[F[\mathcal{E},\mathcal{U},L_1]|_{L_1\cap L_2} = F[\mathcal{E},\mathcal{U},L_2]|_{L_1\cap L_2} \mid Q'\subseteq L_1,L_2\subseteq L \wedge Q'\subseteq L\subseteq W_{i_j}\right]\\
&\geqslant \mathbb{P}_{L\in\mathsf{MGr}(\mathcal{E}'_{i_j},D,2^{D-q^*}\cdot s)}\left[\mathbb{P}_{L_1\in\mathsf{MGr}(\mathcal{E}'_{i_j},\ell_1,2^{\ell_1}\cdot s)}\left[F[\mathcal{E},\mathcal{U},L_1] = f_{i_j}|_{L_1} \mid Q'\subseteq L_1\subseteq L\right]^2 \mid Q'\subseteq L\subseteq W_{i_j}\right]\\
&\geqslant \mathbb{P}_{\substack{L\in\mathsf{MGr}(\mathcal{E}'_{i_j},D,2^{D-q^*}\cdot s)\\ L_1\in\mathsf{MGr}(\mathcal{E}'_{i_j},\ell_1,2^{\ell_1}\cdot s)}}\left[F[\mathcal{E},\mathcal{U},L_1] = f_{i_j}|_{L_1} \mid Q'\subseteq L_1\subseteq L, Q'\subseteq L\subseteq W_{i_j}\right]^2 \qquad \text{(Cauchy-Schwarz)}\\
&\geqslant \underset{L_1\in\mathsf{MGr}(\mathcal{E}'_{i_j},\ell_1,2^{\ell_1}\cdot s)}{\mathbb{P}}\left[F[\mathcal{E},\mathcal{U},L_1] = f_{i_j}|_{L_1}\,\middle|\, Q'\subseteq L\subseteq W_{i_j}\right]^2 - \exp(-\Omega(\ell_1))\\
&\geqslant \exp(-2q_0^6) - \exp(-\Omega(\ell_1)).
\end{aligned}$$

Here, we use the fact that the following two distributions are $\exp(-\Omega(\ell_1))$-close in TV distance,

1. uniformly sample a subspace $L\in\mathsf{MGr}(\mathcal{E}'_{i_j},D,2^{D-q^*}\cdot s)$ conditioned on $Q'\subseteq L\subseteq W_{i_j}$; sample a subspace $L_1\in\mathsf{MGr}(\mathcal{E}'_{i_j},\ell_1,2^{\ell_1}\cdot s)$ conditioned on $Q'\subseteq L_1\subseteq L$;

2. uniformly sample a subspace $L_1\in\mathsf{MGr}(\mathcal{E}'_{i_j},\ell_1,2^{\ell_1}\cdot s)$ conditioned on $Q'\subseteq L_1\subseteq W_{i_j}$.

This can be deduced from Lemma 6.20.

Under Condition 3.9, we conclude that

$$\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{Z}_j \mid \mathscr{X}_j \wedge \mathscr{Y}_j\right] \geqslant \exp(-q_0^7). \tag{6.39}$$

Second-moment calculations. Towards proving (6.36), we show that for every two distinct indices $j,j'\in[m']$, we have

$$2^{-2(D-q^*)w'-2}\lambda \leqslant \mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{X}_j\wedge\mathscr{X}_{j'}\wedge\mathscr{Y}_j\wedge\mathscr{Y}_{j'}\right] \leqslant 2^{-2(D-q^*)w'+2}\lambda \tag{6.40}$$

Again, it is easy to see that $\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{X}_j\wedge\mathscr{X}_{j'}\right] = 2^{-2(D-q^*)w'}$. It suffices to calculate the probability

$$\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{Y}_j\wedge\mathscr{Y}_{j'} \mid \mathscr{X}_j\wedge\mathscr{X}_{j'}\right]$$

The goal is to prove that

$$\lambda^2/4 \leqslant \mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{Y}_j\wedge\mathscr{Y}_{j'} \mid \mathscr{X}_j\wedge\mathscr{X}_{j'}\right] \leqslant 4\lambda^2. \tag{6.41}$$

We use a similar argument as the proof of (6.38). We first define auxiliary tripled sets $\widetilde{\mathcal{E}} = (\widetilde{E}_1,\ldots,\widetilde{E}_1)$ as follows

- for every $c\in[t]\setminus\{c_{i_j},c_{i_{j'}}\}$, let $\widetilde{E}_j = E_{j,c}$;

- let $\widetilde{E}_{c_{i_j}} = E_{i_j,c_{i_j}}$ and $\widetilde{E}_{c_{i_{j'}}} = E_{i_{j'},c_{i_{j'}}}$.

Then, we have

$$\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{Y}_j \wedge \mathscr{Y}_{j'} \mid \mathscr{X}_j \wedge \mathscr{X}_{j'}\right] = \frac{\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{Y}_j \wedge \mathscr{Y}_{j'} \wedge \mathscr{X}_j \wedge \mathscr{X}_{j'} \mid L_1, L_2 \subseteq \mathcal{V}(\widetilde{\mathcal{E}})\right]}{\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{X}_j \wedge \mathscr{X}_{j'} \mid L_1, L_2 \subseteq \mathcal{V}(\widetilde{\mathcal{E}})\right]}.$$

First, it is easy to see that $W_{i_j} \cap W_{i_{j'}} \subseteq \mathcal{V}(\widetilde{\mathcal{E}})$, and let $\widetilde{w} = \operatorname{codim}_{\mathcal{V}(\widetilde{\mathcal{E}})}(W_{i_j} \cap W_{i_{j'}}) \leqslant 2w$. Then, we have

$$\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{X}_j \wedge \mathscr{X}_{j'} \mid L_1, L_2 \subseteq \mathcal{V}(\widetilde{\mathcal{E}})\right] = 2^{-(D-q^*)\widetilde{w}}.$$

Also, we have

$$\begin{aligned}
&\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{Y}_j \wedge \mathscr{Y}_{j'} \wedge \mathscr{X}_j \wedge \mathscr{X}_{j'} \mid L_1, L_2 \subseteq \mathcal{V}(\widetilde{\mathcal{E}})\right] \\
&= \underset{(L_1,L_2)\sim\mathcal{P}}{\mathbb{P}}\left[\mathscr{Y}_j \wedge \mathscr{Y}_{j'} \mid L_1, L_2 \subseteq \mathcal{V}(\widetilde{\mathcal{E}})\right] \cdot \underset{(L_1,L_2)\sim\mathcal{P}}{\mathbb{P}}\left[\mathscr{Y}_j \wedge \mathscr{Y}_{j'} \mid \mathscr{X}_j \wedge \mathscr{X}_{j'}\right] \\
&= \lambda^2 \cdot \underset{(L_1,L_2)\sim\mathcal{P}}{\mathbb{P}}\left[\mathscr{Y}_j \wedge \mathscr{Y}_{j'} \mid \mathscr{X}_j \wedge \mathscr{X}_{j'}\right].
\end{aligned}$$

Here, the second equality comes from the fact that the projections of L_1, L_2 on coordinates $c_{i_j}, c_{i_{j'}}$ are independent under the distribution $\mathcal{P}$. Applying Corollary 6.21 and Proposition 6.23, we get

$$2^{-(D-q^*)\widetilde{w}-2} \leqslant \underset{(L_1,L_2)\sim\mathcal{P}}{\mathbb{P}}\left[\mathscr{Y}_j \wedge \mathscr{Y}_{j'} \mid \mathscr{X}_j \wedge \mathscr{X}_{j'}\right] \leqslant 2^{-(D-q^*)\widetilde{w}+2},$$

under Condition 3.9. This also concludes the proof of (6.40).

We complete the proof of (6.36) by

$$\begin{aligned}
&\mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[F[\mathcal{E},\mathcal{U},L_1]|_{L_1\cap L_2} = F[\mathcal{E},\mathcal{U},L_2]|_{L_1\cap L_2}\right] && (6.42)\\
&\geqslant \sum_{j\in[m']}\left(\underset{(L_1,L_2)\sim\mathcal{P}}{\mathbb{P}}\left[\mathscr{X}_j \wedge \mathscr{Y}_j \wedge \mathscr{Z}_j\right] - \sum_{j,j'\in[m'],j\neq j'}\underset{(L_1,L_2)\sim\mathcal{P}}{\mathbb{P}}\left[\mathscr{Y}_j \wedge \mathscr{Y}_{j'} \wedge \mathscr{X}_j \wedge \mathscr{X}_{j'}\right]\right) && (6.43)\\
&\geqslant m' \cdot 2^{-(D-q^*)w'} \cdot \lambda \cdot \exp(-q_0^7)/3 - (2m' \cdot 2^{-(D-q^*)w'}\lambda)^2 && (6.44)\\
&\geqslant \exp(-5q_0^7). && \text{(plugging in } m' = \exp(-q_0^7)/(10\cdot\lambda\cdot 2^{-(D-q^*)w}))
\end{aligned}$$

Applying Lemma 6.61 and use the fact that

$$\mathsf{agree}(\mathcal{E},\mathcal{U},W,Q') \geqslant \mathbb{P}_{(L_1,L_2)\sim\mathcal{P}}\left[F[\mathcal{E},\mathcal{U},L_1]|_{L_1\cap L_2} = F[\mathcal{E},\mathcal{U},L_2]|_{L_1\cap L_2}\right] - \exp(-\Omega(k)) \geqslant \exp(-6q_0^7),$$

we conclude that there exists $f \in W^*$ such that

$$\underset{L\in\mathsf{Gr}(\mathcal{V}(\mathcal{E}),\ell_1)}{\mathbb{P}}\left[F[\mathcal{E},\mathcal{U},L] = f|_L \,\middle|\, Q' \subseteq L \subseteq W\right] \geqslant \exp(-100\cdot q_0^7).$$

Existence of a consistent decoding. In the first part of the proof, we have shown the existence of a linear function $f \in W^*$ satisfying (6.35). In the second part, we further show that there exists $f \in W^*$ and index $j \in [m']$ such that f satisfies (6.35) and $f|_{W_{i_j}} = f_{i_j}$ at the same time. In the remaining proofs, we use another input from [DKK+25]. (6.35).

Claim 6.62 (Corollary of Lemma 3.15, [DKK+25]). *Assume Condition 3.9. Let $\{f^{\mathfrak{i}} \in W^*\}_{\mathfrak{i}\in[p]}$ to be the list of all linear functions that satisfy (6.35). We then have (1) $p \leqslant \exp(105 \cdot q_0^7)$; and (2)*

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}}\left[F[\mathcal{E},\mathcal{U},L_1]|_{L_1\cap L_2} = F[\mathcal{E},\mathcal{U},L_2]|_{L_1\cap L_2} \wedge F[\mathcal{E},\mathcal{U},L_1] \notin \{f^{\mathfrak{i}}|_{L_1}\}_{\mathfrak{i}\in[p]} \wedge F[\mathcal{E},\mathcal{U},L_2] \notin \{f^{\mathfrak{i}}|_{L_2}\}_{\mathfrak{i}\in[p]}\right]$$

is at most $\exp(-20 \cdot q_0^7)$.

For every $j \in [m']$, we define an event $\mathscr{W}_j$ by

$$\mathscr{W}_j := \wedge_{j'\in[m'],j'\neq j}\neg\left(\mathscr{X}_{j'} \wedge \mathscr{Y}_{j'}\right).$$

Then, we have

$$\begin{aligned}
&\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}}\left[\left(\vee_{j\in[m']}\mathscr{X}_j \wedge \mathscr{Y}_j \wedge \mathscr{Z}_j \wedge \mathscr{W}_j\right) \wedge \left(F[\mathcal{E},\mathcal{U},L_1] \in \{(f^{\mathfrak{i}})|_{L_1}\}_{\mathfrak{i}\in[p]} \vee F[\mathcal{E},\mathcal{U},L_2] \in \{(f^{\mathfrak{i}})|_{L_2}\}_{\mathfrak{i}\in[p]}\right)\right]\\
&\geqslant \mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}}\left[\left(\vee_{j\in[m']}\mathscr{X}_j \wedge \mathscr{Y}_j \wedge \mathscr{Z}_j \wedge \mathscr{W}_j\right)\right]\\
&- \mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}}\left[\left(\vee_{j\in[m']}\mathscr{X}_j \wedge \mathscr{Y}_j \wedge \mathscr{Z}_j \wedge \mathscr{W}_j\right) \wedge \left(F[\mathcal{E},\mathcal{U},L_1] \notin \{(f^{\mathfrak{i}})|_{L_1}\}_{\mathfrak{i}\in[p]} \wedge F[\mathcal{E},\mathcal{U},L_2] \notin \{(f^{\mathfrak{i}})|_{L_2}\}_{\mathfrak{i}\in[p]}\right)\right]
\end{aligned}$$

As shown in the calculation of (6.42), we know that

$$\begin{aligned}
&\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}}\left[\left(\vee_{j\in[m']}\mathscr{X}_j \wedge \mathscr{Y}_j \wedge \mathscr{Z}_j \wedge \mathscr{W}_j\right)\right]\\
&\geqslant \sum_{j\in[m']}\left(\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{X}_j \wedge \mathscr{Y}_j \wedge \mathscr{Z}_j\right] - \sum_{j,j'\in[m'],j\neq j'}\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}}\left[\mathscr{Y}_j \wedge \mathscr{Y}_{j'} \wedge \mathscr{X}_j \wedge \mathscr{X}_{j'}\right]\right)\\
&\geqslant \exp(-5q_0^7).
\end{aligned}$$

Also, notice that $\left(\vee_{j\in[m']}\mathscr{X}_j \wedge \mathscr{Y}_j \wedge \mathscr{Z}_j \wedge \mathscr{W}_j\right)$ implies $F[\mathcal{E},\mathcal{U},L_1]|_{L_1\cap L_2} = F[\mathcal{E},\mathcal{U},L_2]|_{L_1\cap L_2}$ by definitions and thus

$$\begin{aligned}
&\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}}\left[\left(\vee_{j\in[m']}\mathscr{X}_j \wedge \mathscr{Y}_j \wedge \mathscr{Z}_j \wedge \mathscr{W}_j\right) \wedge \left(F[\mathcal{E},\mathcal{U},L_1] \notin \{(f^{\mathfrak{i}})|_{L_1}\}_{\mathfrak{i}\in[p]} \wedge F[\mathcal{E},\mathcal{U},L_2] \notin \{(f^{\mathfrak{i}})|_{L_2}\}_{\mathfrak{i}\in[p]}\right)\right]\\
&\leqslant \mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}}\left[F[\mathcal{E},\mathcal{U},L_1]|_{L_1\cap L_2} = F[\mathcal{E},\mathcal{U},L_2]|_{L_1\cap L_2} \wedge F[\mathcal{E},\mathcal{U},L_1] \notin \{(f^{\mathfrak{i}})|_{L_1}\}_{\mathfrak{i}\in[p]} \wedge F[\mathcal{E},\mathcal{U},L_2] \notin \{(f^{\mathfrak{i}})|_{L_2}\}_{\mathfrak{i}\in[p]}\right]\\
&\leqslant \exp(-20q_0^7). \qquad \text{(by Claim 6.62)}
\end{aligned}$$

Combining the two estimates, we conclude that

$$\begin{aligned}
&\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}}\left[\left(\vee_{j\in[m']}\mathscr{X}_j \wedge \mathscr{Y}_j \wedge \mathscr{Z}_j \wedge \mathscr{W}_j\right) \wedge \left(F[\mathcal{E},\mathcal{U},L_1] \in \{(f^{\mathfrak{i}})|_{L_1}\}_{\mathfrak{i}\in[p]} \vee F[\mathcal{E},\mathcal{U},L_2] \in \{(f^{\mathfrak{i}})|_{L_2}\}_{\mathfrak{i}\in[p]}\right)\right]\\
&\geqslant \exp(-5q_0^7) - \exp(-20q_0^7) \geqslant \exp(-6q_0^7),
\end{aligned}$$

under Condition 3.9.

By the symmetry of L_1, L_2 and the union bound, we get

$$\begin{aligned}
&\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}}\left[\left(\vee_{j\in[m']}\mathscr{X}_j \wedge \mathscr{Y}_j \wedge \mathscr{Z}_j \wedge \mathscr{W}_j\right) \wedge \left(F[\mathcal{E},\mathcal{U},L_1] \in \{(f^{\mathfrak{i}})|_{L_1}\}_{\mathfrak{i}\in[p]} \vee F[\mathcal{E},\mathcal{U},L_2] \in \{(f^{\mathfrak{i}})|_{L_2}\}_{\mathfrak{i}\in[p]}\right)\right]\\
&\leqslant 2\sum_{j\in[m']}\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}}\left[\left(\mathscr{X}_j \wedge \mathscr{Y}_j \wedge \mathscr{Z}_j \wedge \mathscr{W}_j\right) \wedge \left(F[\mathcal{E},\mathcal{U},L_1] \in \{(f^{\mathfrak{i}})|_{L_1}\}_{\mathfrak{i}\in[p]}\right)\right]
\end{aligned}$$

$$\leqslant 2 \sum_{j\in[m']} \mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}} \left[(\mathscr{X}_j \wedge \mathscr{Y}_j \wedge \mathscr{Z}_j \wedge \mathscr{W}_j) \wedge \left(f_{i_j}|_{L_1} \in \{(f^{\mathfrak{i}})|_{L_1}\}_{\mathfrak{i}\in[p]}\right)\right].$$

Here, the last inequality comes from the fact that $(\mathscr{X}_j \wedge \mathscr{Y}_j \wedge \mathscr{Z}_j \wedge \mathscr{W}_j)$ implies that $(f_{i_j})|_{L_1} = F[\mathcal{E},\mathcal{U},L_1]$. By an averaging argument, there exists $j \in [m']$, such that

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}} \left[(\mathscr{Z}_j \wedge \mathscr{W}_j) \wedge \left(f_{i_j}|_{L_1} \in \{(f^{\mathfrak{i}})|_{L_1}\}_{\mathfrak{i}\in[p]}\right) \mid \mathscr{X}_j \wedge \mathscr{Y}_j\right] \geqslant \frac{\exp(-6q_0^7)}{2m' \cdot \mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}}[\mathscr{X}_j \wedge \mathscr{Y}_j]}.$$

By our choice of m', the denominator satisfies $2m' \cdot \mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}}[\mathscr{X}_j \wedge \mathscr{Y}_j] \leqslant 1$, and we get

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}} \left[f_{i_j}|_{L_1} \in \{(f^{\mathfrak{i}})|_{L_1}\}_{\mathfrak{i}\in[p]} \mid \mathscr{X}_j \wedge \mathscr{Y}_j\right] \geqslant \exp(-6q_0^7).$$

By an averaging argument again, we conclude that there exists $\mathfrak{i} \in [p]$ such that

$$\begin{aligned}\mathop{\mathbb{P}}_{L\in\mathsf{MGr}(\mathcal{E}'_{i_j},\ell_1,2^{\ell_1}\cdot s)} \left[f_{i_j}|_{L_1} = (f^{\mathfrak{i}})|_{L_1} \mid Q' \subseteq L \subseteq W_{i_j}\right] &= \mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{P}} \left[f_{i_j}|_{L_1} = (f^{\mathfrak{i}})|_{L_1} \mid \mathscr{X}_j \wedge \mathscr{Y}_j\right] \\ &\geqslant \exp(-6q_0^7)/p \geqslant \exp(-200q_0^7).\end{aligned}$$

It suffices to prove that $f_{i_j} = f^{\mathfrak{i}}|_{W_{i_j}}$. Assume that $f_{i_j} \neq f^{\mathfrak{i}}|_{W_{i_j}}$. Let $W'_{i_j} \subseteq W_{i_j}$ such that for every $y \in W'_{i_j}$, it holds $f_{i_j}(x) = f^{\mathfrak{i}}(x)$. It is easy to see that $\mathrm{codim}_{\mathcal{V}(\mathcal{E})}(W'_{i_j}) = w+1$. By Corollary 6.21, we have (1)

$$\mathop{\mathbb{P}}_{L\in\mathsf{MGr}(\mathcal{E}'_{i_j},\ell_1,2^{\ell_1}\cdot s)} \left[L \subseteq W_{i_j} \mid Q' \subseteq L\right] \geqslant 2^{-(\ell_1-q^*)w-1},$$

and (2)

$$\begin{aligned}\mathop{\mathbb{P}}_{L\in\mathsf{MGr}(\mathcal{E}'_{i_j},\ell_1,2^{\ell_1}\cdot s)} \left[L \subseteq W_{i_j} \wedge f_{i_j}|_{L_1} = (f^{\mathfrak{i}})|_{L_1} \mid Q' \subseteq L\right] &= \mathop{\mathbb{P}}_{L\in\mathsf{MGr}(\mathcal{E}'_{i_j},\ell_1,2^{\ell_1}\cdot s)} \left[L \subseteq W'_{i_j} \mid Q' \subseteq L\right] \\ &\leqslant 2^{-(\ell_1-q^*)(w+1)+1}.\end{aligned}$$

Combining two estimates, we conclude that, if $f_{i_j} \neq f^{\mathfrak{i}}|_{W_{i_j}}$, we have

$$\mathop{\mathbb{P}}_{L\in\mathsf{MGr}(\mathcal{E}'_{i_j},\ell_1,2^{\ell_1}\cdot s)} \left[f_{i_j}|_{L_1} = (f^{\mathfrak{i}})|_{L_1} \mid Q' \subseteq L \subseteq W_{i_j}\right] \leqslant 2^{-(\ell_1-q^*)+2}$$

which leads to a contradiction under Condition 3.9. □

7 Acknowledgments

We thank Mark Braverman, Irit Dinur, Guy Kindler, Subhash Khot, Ran Raz, Muli Safra for several discussions and collaborations that led to this work.

References

[ABS15] Sanjeev Arora, Boaz Barak, and David Steurer. Subexponential algorithms for Unique Games and related problems. *Journal of the ACM*, 62(5):42:1–42:25, 2015.

[ACC06] Sanjeev Arora, Eden Chlamtac, and Moses Charikar. New approximation guarantee for chromatic number. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 215–224. ACM, 2006.

[AKK+08] Sanjeev Arora, Subhash A. Khot, Alexandra Kolla, David Steurer, Madhur Tulsiani, and Nisheeth K. Vishnoi. Unique Games on expanding constraint graphs are easy. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, pages 21–28. ACM, 2008.

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, 1998.

[AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: a new characterization of NP. *Journal of the ACM*, 45(1):70–122, 1998.

[BBKO21] Libor Barto, Jakub Bulín, Andrei Krokhin, and Jakub Opršal. Algebraic approach to promise constraint satisfaction. *Journal of the ACM*, 68(4):28:1–28:66, 2021.

[BHL26] Nikhil Bansal, Neng Huang, and Euiwoong Lee. Improved SDP-based algorithm for coloring 3-colorable graphs, 2026.

[BK97] Avrim Blum and David Karger. An $\widetilde{O}(n^{3/14})$-coloring algorithm for 3-colorable graphs. *Information Processing Letters*, 61(1):49–53, 1997.

[BKS19] Boaz Barak, Pravesh K. Kothari, and David Steurer. Small-set expansion in Shortcode graph and the 2-to-2 Conjecture. In *10th Innovations in Theoretical Computer Science Conference (ITCS 2019)*, volume 124 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 9:1–9:12. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019.

[Chl07] Eden Chlamtac. Approximation algorithms using hierarchies of semidefinite programming relaxations. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, pages 691–701. IEEE Computer Society, 2007.

[CMM06] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for Unique Games. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 205–214. ACM, 2006.

[DGKR05] Irit Dinur, Venkatesan Guruswami, Subhash Khot, and Oded Regev. A new multilayered PCP and the hardness of hypergraph vertex cover. *SIAM Journal on Computing*, 34(5):1129–1146, 2005.

[DH13] Irit Dinur and Prahladh Harsha. Composition of low-error 2-query PCPs using decodable PCPs. *SIAM Journal on Computing*, 42(6):2452–2486, 2013.

[DHK15] Irit Dinur, Prahladh Harsha, and Guy Kindler. Polynomially low error PCPs with polyloglog n queries via modular composition. In *Proceedings of the 47th Annual ACM Symposium on Theory of Computing*, pages 267–276. ACM, 2015.

[DKK+21] Irit Dinur, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. On non-optimally expanding sets in Grassmann graphs. *Israel Journal of Mathematics*, 243(1):377–420, 2021.

[DKK+25] Irit Dinur, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. Towards a proof of the 2-to-1 Games Conjecture. *Theory of Computing*, 21(11):1–50, 2025.

[DMR09] Irit Dinur, Elchanan Mossel, and Oded Regev. Conditional hardness for approximate coloring. *SIAM Journal on Computing*, 39(3):843–873, 2009.

[DRS05] Irit Dinur, Oded Regev, and Clifford Smyth. The hardness of 3-uniform hypergraph coloring. *Combinatorica*, 25(5):519–535, 2005.

[DS05] Irit Dinur and Samuel Safra. On the hardness of approximating minimum vertex cover. *Annals of Mathematics*, 162(1):439–485, 2005.

[EKL23] David Ellis, Guy Kindler, and Noam Lifshitz. An analogue of Bonami's lemma for functions on spaces of linear maps, and 2-2 games. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 656–660. ACM, 2023.

[Fei98] Uriel Feige. A threshold of $\ln n$ for approximating set cover. *Journal of the ACM*, 45(4):634–652, 1998.

[FGL+96] Uriel Feige, Shafi Goldwasser, László Lovász, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *Journal of the ACM*, 43(2):268–292, 1996.

[FMW] Yumou Fei, Dor Minzer, and Shuo Wang. Supplementary material for "on the hardness of 4-to-1 games with perfect completeness". Supplementary material.

[GK04] Venkatesan Guruswami and Sanjeev Khanna. On the hardness of 4-coloring a 3-colorable graph. *SIAM Journal on Discrete Mathematics*, 18(1):30–40, 2004.

[Gol23] Louis Golowich. From Grassmannian to simplicial high-dimensional expanders. In *64th IEEE Annual Symposium on Foundations of Computer Science (FOCS 2023)*, pages 1639–1648. IEEE Computer Society, 2023.

[GS20] Venkatesan Guruswami and Sai Sandeep. d-to-1 hardness of coloring 3-colorable graphs with $O(1)$ colors. In *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*, volume 168 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 62:1–62:12. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020.

[GW95] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995.

[Hås01] Johan Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, 2001.

[HMRW14] Moritz Hardt, Raghu Meka, Prasad Raghavendra, and Benjamin Weitz. Computational limits for matrix completion. In *Proceedings of the 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 703–725. PMLR, 2014.

[HMS23] Yahli Hecht, Dor Minzer, and Muli Safra. NP-hardness of almost coloring almost 3-colorable graphs. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2023)*, volume 275 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 51:1–51:12. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023.

[Hol09] Thomas Holenstein. Parallel repetition: simplification and the no-signaling case. *Theory of Computing*, 5(8):141–172, 2009.

[Kho02] Subhash Khot. On the power of Unique 2-prover 1-round games. In *Proceedings of the 17th Annual IEEE Conference on Computational Complexity*, page 25. IEEE Computer Society, 2002.

[Kho10a] Subhash Khot. Inapproximability of NP-complete problems, discrete Fourier analysis, and geometry. In *Proceedings of the International Congress of Mathematicians*, volume 4, pages 2676–2697. Hindustan Book Agency, 2010.

[Kho10b] Subhash Khot. On the Unique Games Conjecture (invited survey). In *Proceedings of the 25th Annual IEEE Conference on Computational Complexity*, pages 99–121. IEEE Computer Society, 2010.

[Kho14] Subhash Khot. Hardness of approximation. In *Proceedings of the International Congress of Mathematicians (ICM 2014)*, volume 1, pages 711–728. Kyung Moon Sa, 2014.

[KKMO07] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM Journal on Computing*, 37(1):319–357, 2007.

[KLS00] Sanjeev Khanna, Nathan Linial, and Shmuel Safra. On the hardness of approximating the chromatic number. *Combinatorica*, 20(3):393–415, 2000.

[KMMS25] Subhash Khot, Dor Minzer, Dana Moshkovitz, and Muli Safra. Small-set expansion in the Johnson graph. *Theory of Computing*, 21(2):1–43, 2025.

[KMS98] David R. Karger, Rajeev Motwani, and Madhu Sudan. Approximate graph coloring by semidefinite programming. *Journal of the ACM*, 45(2):246–265, 1998.

[KMS23] Subhash Khot, Dor Minzer, and Muli Safra. Pseudorandom sets in Grassmann graph have near-perfect expansion. *Annals of Mathematics*, 198(1):1–92, 2023.

[KMS25] Subhash Khot, Dor Minzer, and Muli Safra. On independent sets, 2-to-2 games and Grassmann graphs. *Theory of Computing*, 21(10):1–55, 2025.

[KR08] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within $2-\varepsilon$. *Journal of Computer and System Sciences*, 74(3):335–349, 2008.

[KS13] Subhash Khot and Muli Safra. A two-prover one-round game with strong soundness. *Theory of Computing*, 9(28):863–887, 2013.

[KS14] Subhash Khot and Rishi Saket. Hardness of finding independent sets in 2-colorable and almost 2-colorable hypergraphs. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1607–1625. Society for Industrial and Applied Mathematics, 2014.

[KT17] Ken-ichi Kawarabayashi and Mikkel Thorup. Coloring 3-colorable graphs with less than $n^{1/5}$ colors. *Journal of the ACM*, 64(1):4:1–4:23, 2017.

[KTY24] Ken-ichi Kawarabayashi, Mikkel Thorup, and Hirotaka Yoneda. Better coloring of 3-colorable graphs. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 331–339. ACM, 2024.

[KV15] Subhash A. Khot and Nisheeth K. Vishnoi. The Unique Games Conjecture, integrality gap for cut problems and embeddability of negative-type metrics into ℓ_1. *Journal of the ACM*, 62(1):8:1–8:39, 2015.

[MR10] Dana Moshkovitz and Ran Raz. Two-query PCP with subconstant error. *Journal of the ACM*, 57(5):29:1–29:29, 2010.

[MZ26] Dor Minzer and Kai Zhe Zheng. Near optimal alphabet-soundness tradeoff PCPs. *Journal of the ACM*, 73(4), 2026.

[Rag08] Prasad Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, pages 245–254. ACM, 2008.

[Rao11] Anup Rao. Parallel repetition in projection games and a concentration bound. *SIAM Journal on Computing*, 40(6):1871–1891, 2011.

[Raz98] Ran Raz. A parallel repetition theorem. *SIAM Journal on Computing*, 27(3):763–803, 1998.

[Tre12] Luca Trevisan. On Khot's Unique Games Conjecture. *Bulletin of the American Mathematical Society*, 49(1):91–111, 2012.

[TSSW00] Luca Trevisan, Gregory B. Sorkin, Madhu Sudan, and David P. Williamson. Gadgets, approximation, and linear programming. *SIAM Journal on Computing*, 29(6):2074–2097, 2000.

[Wig83] Avi Wigderson. Improving the performance guarantee for approximate graph coloring. *Journal of the ACM*, 30(4):729–735, 1983.

[Zuc07] David Zuckerman. Linear degree extractors and the inapproximability of Max Clique and chromatic number. *Theory of Computing*, 3(6):103–128, 2007.

A Approximate Coloring of Hypergraphs

In this section we prove Theorem 3.1, restated below.

Theorem 3.1. *For every fixed positive integer c, there exists a constant $\varepsilon \geqslant \exp(-\exp(O(c^3)))$ such that given a 3-uniform hypergraph H, it is* **NP***-hard to distinguish between the following cases:*

(1) Yes case: H is 2-colorable.

(2) No case: in every c-coloring of H, at least ε fraction of the edges are monochromatic.

Definition A.1. Let d_0 and d_1 be positive integers. A (d_0, d_1)*-biregular projection label cover instance* $\Psi = (V_0, V_1, E, \Sigma_0, \Sigma_1, (\pi_e)_{e\in E})$ consists of the following data:

(1) two vertex sets V_0 and V_1;

(2) a set E of edges between V_0 and V_1 such that in the bipartite graph $(V_0 \sqcup V_1, E)$, every vertex in V_0 has degree d_0 and every vertex in V_1 has degree d_1;

(3) two finite alphabets Σ_0, Σ_1; and

(4) a collection of constraints $(\pi_e)_{e\in E}$, where each π_e is a map from Σ_0 to Σ_1.

The value of Ψ, denoted by $\mathrm{val}(\Psi)$, is the maximum value of

$$\frac{1}{|E|} \sum_{(v_0,v_1)\in E} \mathbb{1}\left[\pi_{(v_0,v_1)}\big(A_0(v_0)\big) = A_1(v_1)\right]$$

achievable by assignment maps $A_0 : V_0 \to \Sigma_0$ and $A_1 : V_1 \to \Sigma_1$.

Theorem A.2. *There exists an absolute constant $\gamma \in (0, 1)$ such that for every fixed positive integer u, given a $(3^u, 5^u)$-biregular projection label cover instance $\Psi = (V_0, V_1, E, \Sigma_0, \Sigma_1, (\pi_e)_{e\in E})$ such that $|\Sigma_0| = 7^u$ and $|\Sigma_1| = 2^u$, it is* **NP***-hard to distinguish between the case* $\mathrm{val}(\Psi) = 1$ *and the case* $\mathrm{val}(\Psi) \leqslant 2^{-\gamma u}$.

Proof of Theorem 3.1. We proceed by a reduction from Theorem A.2. Let $\gamma \in (0, 1)$ be as in Theorem A.2, and let $u = \lceil 40\gamma^{-2}c\rceil$. We fix a $(3^u, 5^u)$-biregular projection label cover instance $\Psi = (V_0, V_1, E, \Sigma_0, \Sigma_1, (\pi_e)_{e\in E})$ with $|\Sigma_0| = 7^u$ and $|\Sigma_1| = 2^u$; we will show how to map it to a 3-uniform hypergraph H while linking the value of Ψ to the colorability of H.

The proof is divided into 8 steps. We first construct the hypergraph H in Steps 1 and 2. The completeness of the reduction is proved in Step 3, while Steps 4 to 8 will be devoted to proving the soundness of the reduction. Except for Step 6, the proof is almost identical the one in [DRS05], and we claim no originality here.[5] Throughout the proof, we let $\ell = 2c^2$.

Step 1: multi-layered label cover. Following the ideas of [DGKR05, DRS05], the first step is to perform a "multi-layered" parallel repetition on Ψ. For each $i \in [\ell]$, let $\mathcal{V}_i$ and R_i be the cartesian products

$$\mathcal{V}_i = V_1^{i-1} \times V_0^{\ell-i} \qquad \text{and} \qquad R_i = \Sigma_1^{i-1} \times \Sigma_0^{\ell-i}.$$

The $(\ell-1)$-tuples $\mathbf{v} = (v_1, \dots, v_{\ell-1}) \in \mathcal{V}_i$ will be the "blown-up vertices" in the i-th layer, and R_i can be viewed as their alphabets. For each $i, j \in [\ell]$ such that $i < j$, let $\mathcal{E}_{i,j}$ be the collection of pairs

$$(\mathbf{v}, \mathbf{v}') = \big((v_1, \dots, v_{\ell-1}), (v'_1, \dots, v'_{\ell-1})\big) \in \mathcal{V}_i \times \mathcal{V}_j$$

such that:

(1) $v_s = v'_s$ for all $s \in \{1, 2, \dots, i-1\} \cup \{j, j+1, \dots, \ell-1\}$; and

(2) $(v_s, v'_s) \in E$ for all $s \in \{i, i+1, \dots, j-1\}$.

[5] The arguments in Steps 1, 5 and 8 are due to [DGKR05].

Given such an pair $(\mathbf{v},\mathbf{v}') \in \mathcal{E}_{i,j}$, we define a projection map $\pi_{(\mathbf{v},\mathbf{v}')} : R_i \to R_j$ by

$$\pi_{(\mathbf{v},\mathbf{v}')}(\sigma_1,\ldots,\sigma_{\ell-1}) = \Big(\sigma_1,\ldots,\sigma_{i-1},\pi_{(v_i,v_i')}(\sigma_i),\ldots,\pi_{(v_{j-1},v_{j-1}')}(\sigma_{j-1}),\sigma_j,\ldots,\sigma_{\ell-1}\Big),$$

for all $\sigma_1,\ldots,\sigma_{i-1} \in \Sigma_1$ and $\sigma_i,\ldots,\sigma_{\ell-1} \in \Sigma_0$.

For each $i,j \in [\ell]$ such that $i<j$, we have

$$|\mathcal{E}_{i,j}| = |\mathcal{V}_i| \cdot (3^u)^{j-i} = (3^u)^{j-i} \cdot |V_1|^{i-1}|V_0|^{\ell-i} = 3^{u(j-i)} \cdot (5/3)^{u(i-1)} \cdot |V_0|^{\ell-1}, \tag{A.1}$$

where we used the $(3^u,5^u)$-biregularity of Ψ and its consequence $|V_1|/|V_0| = 5^u/3^u$.

Step 2: construction of H. Following [DRS05], we construct the hypergraph H by composing the multi-layered label cover from Step 1 with an inner gadget based on the Kneser graph.

For each $i \in [\ell]$, we let $r_i = \lceil (|R_i| - c)/2 \rceil$. The vertex set of H is the disjoint union of cartesian products $\bigsqcup_{i=1}^{\ell} \left(\mathcal{V}_i \times \binom{R_i}{r_i}\right)$, and the hyperedges of H are the collection of unordered triples $\big\{(\mathbf{v},S_1),(\mathbf{v},S_2),(\mathbf{v}',S_3)\big\}$ such that:

(1) $\mathbf{v} \in \mathcal{V}_i$ and $\mathbf{v}' \in \mathcal{V}_j$ for some $i,j \in [\ell]$ such that $i<j$;

(2) $(\mathbf{v},\mathbf{v}') \in \mathcal{E}_{i,j}$;

(3) $S_1,S_2 \in \binom{R_i}{r_i}$ and $S_3 \in \binom{R_j}{r_j}$; and

(4) $S_1 \cap S_2 = \emptyset$ and $\pi_{(\mathbf{v},\mathbf{v}')}(R_i \setminus (S_1 \cup S_2)) \subseteq S_3$.

Using Equation (A.1), we can upper-bound the number of edges in H by

$$\sum_{i,j\in[\ell],\, i<j} |\mathcal{E}_{i,j}| \cdot 2^{2|R_i|} \cdot 2^{|R_j|} \leqslant 8^{7^{u(\ell-1)}} \ell^2 \cdot \max_{i,j\in[\ell],\, i<j} |\mathcal{E}_{i,j}| \leqslant 3^{u(\ell-1)} 8^{7^{u(\ell-1)}} \ell^2 \cdot |V_0|^{\ell-1}. \tag{A.2}$$

It is clear that H can be constructed from Ψ by a polynomial-time Turing machine. It suffices to prove the following statements:

(1) Completeness: if $\mathrm{val}(\Psi) = 1$, then H is 2-colorable.

(2) Soundness: if under some c-coloring of H, less than $\exp(-\exp(\Omega(c^3)))$ fraction of edges of H are monochromatic, then $\mathrm{val}(\Psi) > 2^{-\gamma u}$.

The first statement will be proved in Step 3, and the second will be proved in Steps 4 to 8.

Step 3: completeness. In this step, we assume $\mathrm{val}(\Psi) = 1$ and show that H is 2-colorable.

Suppose $A_0 : V_0 \to \Sigma_0$ and $A_1 : V_1 \to \Sigma_1$ are assignments that together satisfy every constraint of Ψ. Then for any $i \in [\ell]$, we can define a lifted assignment map $\widetilde{A}_i : \mathcal{V}_i \to R_i$ by letting

$$\widetilde{A}_i(\mathbf{v}) = \big(A_1(v_1),\ldots,A_1(v_{i-1}),A_0(v_i),\ldots,A_0(v_{\ell-1})\big), \qquad \text{for all } \mathbf{v} = (v_1,\ldots,v_{\ell-1}) \in \mathcal{V}_i.$$

For any vertex $(\mathbf{v},S) \in \mathcal{V}_i \times \binom{R_i}{r_i}$, we color it red if $\widetilde{A}_i(\mathbf{v}) \in S$, and otherwise color it blue.

Now, for any hyperedge $\big\{(\mathbf{v},S_1),(\mathbf{v},S_2),(\mathbf{v}',S_3)\big\}$ included in H (see Step 2), we have:

(1) The vertices $(\mathbf{v},S_1)$ and $(\mathbf{v},S_2)$ cannot be both red, because $S_1 \cap S_2 = \emptyset$.

(2) Let $i, j \in [\ell]$ be the indices such that $\mathbf{v} \in \mathcal{V}_i$ and $\mathbf{v}' \in \mathcal{V}_j$. If $(\mathbf{v}, S_1)$ and $(\mathbf{v}, S_2)$ are both colored blue, then $\widetilde{A}_i(\mathbf{v}) \in R_i \setminus (S_1 \cup S_2)$ and hence

$$\pi_{(\mathbf{v},\mathbf{v}')}\big(\widetilde{A}_i(\mathbf{v})\big) \in S_3, \qquad \text{due to} \qquad \pi_{(\mathbf{v},\mathbf{v}')}(R_i \setminus (S_1 \cup S_2)) \subseteq S_3.$$

By the definitions of $\pi_{(\mathbf{v},\mathbf{v}')}$ and $\widetilde{A}_i, \widetilde{A}_j$, we have $\pi_{(\mathbf{v},\mathbf{v}')}\big(\widetilde{A}_i(\mathbf{v})\big) = \widetilde{A}_j(\mathbf{v}')$. Therefore, if $(\mathbf{v}, S_1)$ and $(\mathbf{v}, S_2)$ are both colored blue, then $(\mathbf{v}', S_3)$ must be colored red.

In conclusion, the hyperedge $\big\{(\mathbf{v}, S_1), (\mathbf{v}, S_2), (\mathbf{v}', S_3)\big\}$ cannot be monochromatic under the proposed coloring. This means that H is 2-colorable as long as $\mathrm{val}(\Psi) = 1$.

Step 4: finding the leading color. In the rest of the proof, we fix a c-coloring of H, denoted by the map

$$f : \bigsqcup_{i=1}^{\ell} \left(\mathcal{V}_i \times \tbinom{R_i}{r_i}\right) \to [c].$$

We suppose that at most ε fraction of edges of H are monochromatic under f, for some $\varepsilon \in (0,1)$ to be specified later.

For each $i \in [\ell]$ and each $\mathbf{v} \in \mathcal{V}_i$, we apply [DRS05, Lemma 2.2] to the vertices in the set $\left\{(\mathbf{v}, S) \,\middle|\, S \in \binom{R_i}{r_i}\right\}$ to find a color $c_{\mathbf{v}} \in [c]$ such that:

(1) The fraction of sets $S \in \binom{R_i}{r_i}$ such that $f(\mathbf{v}, S) = c_{\mathbf{v}}$ is at least $|R_i|^{-c-1}$.

(2) There exist disjoint sets $S_1, S_2 \in \binom{R_i}{r_i}$ such that $f(\mathbf{v}, S_1) = f(\mathbf{v}, S_2) = c_{\mathbf{v}}$. The set $R_i \setminus (S_1 \cup S_2)$ will be denoted by $B(\mathbf{v})$ and used later in the proof.

For each $i \in [\ell]$, let $c^{(i)} \in [c]$ be a color that maximizes the number of tuples $\mathbf{v} \in \mathcal{V}_i$ with $c_{\mathbf{v}} = c^{(i)}$. Since $\ell = 2c^2$, by the pigeonhole principle there exists a color $c^* \in [c]$ such that $c^{(i)} = c^*$ for at least $2c$ distinct indices $i \in [\ell]$. Pick exactly $2c$ indices among them, and let I be the set of these $2c$ indices.

Step 5: locating two critical layers. Observe that the bipartite graph $(\mathcal{V}_i \sqcup \mathcal{V}_{i+1}, \mathcal{E}_{i,i+1})$ is bi-regular for all $i \in [\ell - 1]$, due to the bi-regularity of Ψ. Consider the following random walk:

1. Start at a uniformly random element $\mathbf{v}^{(1)} \in \mathcal{V}_1$.
2. For each $i \in [\ell - 1]$, let $\mathbf{v}^{(i+1)}$ be a uniformly random element of $\mathcal{V}_{i+1}$ conditioned on $(\mathbf{v}^{(i)}, \mathbf{v}^{(i+1)}) \in \mathcal{E}_{i,i+1}$.

Due to the bi-regularity of the bipartite graphs $(\mathcal{V}_i \sqcup \mathcal{V}_{i+1}, \mathcal{E}_{i,i+1})$, this random walk produces a random sequence $(\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(\ell)}) \in \mathcal{V}_1 \times \cdots \times \mathcal{V}_\ell$ that is marginally uniform over each $\mathcal{V}_i$. Furthermore, by the definition of the edge sets $\mathcal{E}_{i,j}$, it is easy to see that for any $i, j \in [\ell]$ such that $i < j$, the distribution of the pair $(\mathbf{v}^{(i)}, \mathbf{v}^{(j)})$ is uniform over $\mathcal{E}_{i,j}$.

For each $i \in [\ell]$, let Γ_i be the event that $c_{\mathbf{v}^{(i)}} = c^{(i)}$. By the marginal uniformity and the choice of the $c^{(i)}$'s we have $\mathbb{P}[\Gamma_i] \geqslant c^{-1}$. By the inclusion-exclusion principle, we have

$$\begin{aligned} 1 \geqslant \mathbb{P}\left[\bigcup\nolimits_{i\in I} \Gamma_i\right] &\geqslant \sum_{i \in I} \mathbb{P}[\Gamma_i] - \sum_{i,j \in I,\, i<j} \mathbb{P}[\Gamma_i \cap \Gamma_j] \\ &\geqslant |I| \cdot c^{-1} - \binom{|I|}{2} \cdot \max_{i,j \in I,\, i<j} \mathbb{P}[\Gamma_i \cap \Gamma_j]. \end{aligned}$$

Therefore, there exist $i, j \in I$ with $i < j$ such that

$$\mathbb{P}\left[\Gamma_i \cap \Gamma_j\right] \geqslant \binom{|I|}{2}^{-1} \cdot (c^{-1}|I| - 1) = \binom{2c}{2}^{-1} \geqslant \frac{1}{2c^2}.$$

By the uniformity of $(\mathbf{v}^{(i)}, \mathbf{v}^{(j)})$ over $\mathcal{E}_{i,j}$, this means that

$$\mathop{\mathbb{P}}_{(\mathbf{v},\mathbf{v}')\in\mathcal{E}_{i,j}}\left[c_{\mathbf{v}} = c^{(i)} \text{ and } c_{\mathbf{v}'} = c^{(j)}\right] \geqslant \frac{1}{2c^2}.$$

Since $i, j \in I$, by the definition of I, the above is equivalent to

$$\mathop{\mathbb{P}}_{(\mathbf{v},\mathbf{v}')\in\mathcal{E}_{i,j}}\left[c_{\mathbf{v}} = c_{\mathbf{v}'} = c^*\right] \geqslant \frac{1}{2c^2}. \tag{A.3}$$

In the remainder of the proof, we fix the indices i and j identified above (satisfying Equation (A.3)).

Step 6: disrecarding monochromatic edges. We let $\mathcal{E}^*$ be the collection of all pairs $(\mathbf{v}, \mathbf{v}') \in \mathcal{E}_{i,j}$ such that:

(1) $c_{\mathbf{v}} = c_{\mathbf{v}'} = c^*$; and

(2) no hyperedge of H of the form $\{(\mathbf{v}, S_1), (\mathbf{v}, S_2), (\mathbf{v}', S_3)\}$ is colored monochromatically by f.

By Equation (A.2), the total number of monochromatic edges of H under the coloring f is at most

$$\varepsilon \cdot 3^{u(\ell-1)} 8^{7^{u(\ell-1)}} \ell^2 \cdot |V_0|^{\ell-1} \leqslant \varepsilon \cdot 3^{u(\ell-1)} 8^{7^{u(\ell-1)}} \ell^2 \cdot |\mathcal{E}_{i,j}| \leqslant \frac{1}{4c^2} |\mathcal{E}_{i,j}|,$$

where we chose a sufficiently small $\varepsilon = \exp(-\exp(O(c^3)))$ that justifies the last transition. It then follows from Equation (A.3) and the definition of $\mathcal{E}^*$ that

$$|\mathcal{E}^*| \geqslant \frac{1}{2c^2} |\mathcal{E}_{i,j}| - \frac{1}{4c^2} |\mathcal{E}_{i,j}| = \frac{1}{4c^2} |\mathcal{E}_{i,j}|. \tag{A.4}$$

Step 7: strategy for the multi-layered label cover. The goal of this step is to decode from f a "good" assignment map for the two critical layers (identified in Step 5) of the ℓ-layered label cover instance (constructed in Step 1).

For each $\mathbf{v}' \in \mathcal{V}_j$, let $N^*(\mathbf{v}')$ be the collection of tuples $\mathbf{v} \in \mathcal{V}_i$ such that $(\mathbf{v}, \mathbf{v}') \in \mathcal{E}^*$.

In this paragraph, we define a deterministic map $\widetilde{A}_j : \mathcal{V}_j \to R_j$ using the coloring f. Fix a tuple $\mathbf{v}' \in \mathcal{V}_j$. If $c_{\mathbf{v}'} \neq c^*$, we assign $\widetilde{A}_j(\mathbf{v}')$ arbitrarily. Otherwise, for each $\mathbf{v} \in N^*(\mathbf{v}')$, we denote

$$B(\mathbf{v}, \mathbf{v}') := \pi_{(\mathbf{v},\mathbf{v}')}(B(\mathbf{v})).$$

Recall from Step 4 that, the set $B(\mathbf{v}) \subseteq R_i$ can be decomposed as the disjoint union of two sets $S_1, S_2 \in \binom{R_i}{r_i}$ such that $f(\mathbf{v}, S_1) = f(\mathbf{v}, S_2) = c^*$. For each $S \in \binom{R_j}{r_j}$ such that $B(\mathbf{v}, \mathbf{v}') \subseteq S$, by the construction in Step 2, the hyperedge

$$\{(\mathbf{v}, S_1), (\mathbf{v}, S_2), (\mathbf{v}', S_3)\}$$

is included in H. By the definition of $\mathcal{E}^*$, this hyperedge cannot be monochromatic under the coloring f, and therefore $f(\mathbf{v}', S) \neq c^*$. However, recall also from Step 4 that the fraction of $S \in \binom{R_j}{r_j}$ such that $f(\mathbf{v}', S) = c^*$ is at least $|R_j|^{-c-1}$. Therefore, at least an $|R_j|^{-c-1}$ fraction of

$S \in \binom{R_j}{r_j}$ contain none of the sets $B(\mathbf{v}, \mathbf{v}')$, where $\mathbf{v}$ ranges in $N^*(\mathbf{v}')$. We can then apply Lemma A.3 to find[6] an element of R_j that is contained in $B(\mathbf{v}, \mathbf{v}')$ for at least

$$\eta = \frac{6^{-c}}{c(c+1)\ln|R_j|}$$

fraction of the tuples $\mathbf{v} \in N^*(\mathbf{v}')$. Define $\widetilde{A}_j(\mathbf{v}')$ to be such an element of R_j.

We now define a random assignment map $\widetilde{\mathcal{A}}_i : \mathcal{V}_i \to R_i$ by picking each $\widetilde{\mathcal{A}}_i(\mathbf{v})$ independently and uniformly at random from $B(\mathbf{v})$. By the definition of $B(\mathbf{v})$ (in Step 4), we have $|B(\mathbf{v})| \leqslant |R_i| - 2r_i \leqslant c$. Therefore, we have

$$\begin{aligned}
&\mathbb{E}_{\widetilde{\mathcal{A}}_i}\left[\mathop{\mathbb{P}}_{(\mathbf{v},\mathbf{v}')\in\mathcal{E}^*}\left[\pi_{(\mathbf{v},\mathbf{v}')}\big(\widetilde{\mathcal{A}}_i(\mathbf{v})\big) = \widetilde{A}_j(\mathbf{v}')\right]\right] \\
&= \mathop{\mathbb{E}}_{(\mathbf{v},\mathbf{v}')\in\mathcal{E}^*}\left[\mathbb{P}_{\widetilde{\mathcal{A}}_i}\left[\pi_{(\mathbf{v},\mathbf{v}')}\big(\widetilde{\mathcal{A}}_i(\mathbf{v})\big) = \widetilde{A}_j(\mathbf{v}')\right]\right] \\
&\geqslant \mathop{\mathbb{E}}_{(\mathbf{v},\mathbf{v}')\in\mathcal{E}^*}\left[\eta \cdot \mathbb{P}_{\widetilde{\mathcal{A}}_i}\left[\pi_{(\mathbf{v},\mathbf{v}')}\big(\widetilde{\mathcal{A}}_i(\mathbf{v})\big) = \widetilde{A}_j(\mathbf{v}') \,\middle|\, \widetilde{A}_j(\mathbf{v}') \in B(\mathbf{v},\mathbf{v}')\right]\right] && \text{(by the choice of } \widetilde{A}_j(\mathbf{v}')\text{)} \\
&\geqslant \eta \cdot \mathop{\mathbb{E}}_{(\mathbf{v},\mathbf{v}')\in\mathcal{E}^*}\left[\frac{1}{|B(\mathbf{v})|}\right] \geqslant \frac{\eta}{c}. && \text{(by the definition of } B(\mathbf{v},\mathbf{v}')\text{)}
\end{aligned}$$

Therefore, there exists an deterministic instantiation $\widetilde{A}_i : \mathcal{V}_i \to R_i$ of $\widetilde{\mathcal{A}}_i$ such that

$$\mathop{\mathbb{P}}_{(\mathbf{v},\mathbf{v}')\in\mathcal{E}^*}\left[\pi_{(\mathbf{v},\mathbf{v}')}\big(\widetilde{A}_i(\mathbf{v})\big) = \widetilde{A}_j(\mathbf{v}')\right] \geqslant \frac{\eta}{c}.$$

By Equation (A.4), it follows that

$$\mathop{\mathbb{P}}_{(\mathbf{v},\mathbf{v}')\in\mathcal{E}_{i,j}}\left[\pi_{(\mathbf{v},\mathbf{v}')}\big(\widetilde{A}_i(\mathbf{v})\big) = \widetilde{A}_j(\mathbf{v}')\right] \geqslant \frac{\eta}{4c^3}. \tag{A.5}$$

In the rest of the proof, we fix $\widetilde{A}_i$ to be an instantiation of $\widetilde{\mathcal{A}}_i$ that satisfies Equation (A.5).

Step 8: strategy for the original label cover. We define a random pair of (correlated) embedding maps $\varphi_0 : V_0 \hookrightarrow \mathcal{V}_i$ and $\varphi_1 : V_1 \hookrightarrow \mathcal{V}_j$ as follows:

1. For each $s \in \{1, 2, \ldots, i-1\}$, let $v_s = v'_s$ be a uniformly random element of V_1.
2. For each $s \in \{i+1, i+2, \ldots, j-1\}$, pick a pair $(v_s, v'_s) \in E$ uniformly at random.
3. For each $s \in \{j, j+1, \ldots, \ell-1\}$, let $v_s = v'_s$ be a uniformly random element of V_0.
4. For each $w \in V_0$, let $\varphi_0(w) = (v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_{\ell-1}) \in \mathcal{V}_i$.
5. For each $w' \in V_1$, let $\varphi_1(w') = (v'_1, \ldots, v'_{i-1}, w', v'_{i+1}, \ldots, v'_{\ell-1}) \in \mathcal{V}_j$.

If we first sample the random maps φ_0, φ_1 and then independently sample a pair $(w, w') \in E$, it is clear that $(\varphi_0(w), \varphi_1(w'))$ is uniformly distributed over $\mathcal{E}_{i,j}$. Therefore, it follows from Equation (A.5) that

$$\mathop{\mathbb{E}}_{(\varphi_0,\varphi_1)}\left[\mathop{\mathbb{P}}_{(w,w')\in E}\left[\pi_{(\varphi_0(w),\varphi_1(w'))}\Big(\widetilde{A}_i\big(\varphi_0(w)\big)\Big) = \widetilde{A}_j\big(\varphi_1(w')\big)\right]\right] \geqslant \frac{\eta}{4c^3}. \tag{A.6}$$

We now define a random pair of (correlated) assignment maps $\mathcal{A}_0 : V_0 \to \Sigma_0$ and $\mathcal{A}_1 : V_1 \to \Sigma_1$ as follows:

[6] Note that for any $\mathbf{v} \in N^*(\mathbf{v}')$, we have $|B(\mathbf{v},\mathbf{v}')| = \big|\pi_{(\mathbf{v},\mathbf{v}')}(B(\mathbf{v}))\big| \leqslant |B(\mathbf{v})| \leqslant |R_i| - 2r_i \leqslant c$.

1. Sample the pair of embedding maps φ_0, φ_1 as described earlier.
2. For each $w \in V_0$, let $\mathcal{A}_0(w)$ be the i-th coordinate of $\widetilde{A}_i(\varphi_0(w))$.
3. For each $w' \in V_1$, let $\mathcal{A}_1(w')$ be the i-th coordinate of $\widetilde{A}_j(\varphi_1(w))$.

By the definition of $\pi_{(\varphi_0(w),\varphi_1(w'))}$ in Step 1, it follows from Equation (A.6) that

$$\mathop{\mathbb{E}}_{(\mathcal{A}_0,\mathcal{A}_1)}\left[\mathop{\mathbb{P}}_{(w,w')\in E}\left[\pi_{(w,w')}\big(\mathcal{A}_0(w)\big) = \mathcal{A}_1(w')\right]\right] \geqslant \frac{\eta}{4c^3}.$$

Therefore, there exists a deterministic instantiation (A_0, A_1) of the random pair of maps $(\mathcal{A}_0, \mathcal{A}_1)$ that satisfies at least $\eta/(4c^3)$ fraction of the constraints in Ψ. Thus we have

$$\mathrm{val}(\Psi) \geqslant \frac{\eta}{4c^3} \geqslant \frac{6^{-c}}{4c^4(c+1)\ln|R_j|} \geqslant \frac{6^{-c}}{4c^4(c+1)(\ell-1)\ln(7^u)} \geqslant \frac{1}{2^{20c}\cdot u} > 2^{-\gamma u},$$

where we used $u \geqslant 40\gamma^{-2}c$ in the last transition. □

Lemma A.3. *Fix a positive integer c. Let R be a nonempty finite set with $|R| \geqslant 6c$, and let r be an integer such that $|R|/3 \leqslant r \leqslant 2|R|/3$. Suppose $B_1, \ldots, B_m \subseteq R$ are a sequence of subsets such that $|B_i| \leqslant c$ for each $i \in [m]$. If at least an $|R|^{-c-1}$ fraction of r-element subsets $S \in \binom{R}{r}$ contain none of the sets $B_1, \ldots, B_m$, then there exists an element $a \in R$ contained in at least*

$$\frac{6^{-c}}{c(c+1)\ln|R|}$$

fraction of the sets $B_1, \ldots, B_m$.

Proof. Let

$$\eta := \max_{a\in R} \frac{\big|\{i \in [m] \mid a \in B_i\}\big|}{m}.$$

Without loss of generality, suppose $B_1, \ldots, B_q$ are a maximum pairwise disjoint subfamily among the sets $B_1, \ldots, B_m$, for some integer $q \in [m]$. By maximality, every B_i contains some element of

$$U := B_1 \cup \cdots \cup B_q.$$

Therefore

$$m \leqslant \sum_{a\in U} \big|\{i \in [m] \mid a \in B_i\}\big| \leqslant |U| \cdot \eta m \leqslant cq \cdot \eta m,$$

and hence $\eta \geqslant 1/(cq)$.

Next we upper-bound q using the assumptions in the lemma. Let S be chosen uniformly at random from $\binom{R}{r}$. For each $j \in [q]$, we have

$$\mathop{\mathbb{P}}_{S\in\binom{R}{r}}[B_j \subseteq S] = \binom{|R|-|B_j|}{r-|B_j|}\Big/\binom{|R|}{r} = \prod_{t=0}^{|B_j|-1} \frac{r-t}{|R|-t} \geqslant \left(\frac{1}{6}\right)^{|B_j|} \geqslant 6^{-c},$$

where we used the assumptions $r \geqslant |R|/3$ and $|B_j| \leqslant c \leqslant |R|/6$ in the third transition.

Because the sets $B_1, \ldots, B_q$ are pairwise disjoint, the events $\{B_j \subseteq S\}$, for $j \in [q]$, are negatively correlated under the uniform distribution over $\binom{R}{r}$. Hence

$$\mathop{\mathbb{P}}_{S\in\binom{R}{r}}\big[B_j \not\subseteq S \text{ for all } j \in [q]\big] \leqslant \prod_{j=1}^{q} \mathop{\mathbb{P}}_{S\in\binom{R}{r}}[B_j \not\subseteq S] \leqslant (1-6^{-c})^q.$$

Therefore, by the assumption in the lemma, we have

$$|R|^{-c-1} \leqslant \underset{S \in \binom{R}{r}}{\mathbb{P}} \big[B_i \nsubseteq S \text{ for all } i \in [m]\big] \leqslant \underset{S \in \binom{R}{r}}{\mathbb{P}} \big[B_j \nsubseteq S \text{ for all } j \in [q]\big] \leqslant (1-6^{-c})^q.$$

Taking logarithms gives

$$q \leqslant \frac{(c+1)\ln|R|}{-\ln(1-6^{-c})} \leqslant 6^c(c+1)\ln|R|.$$

Combining this with $\eta \geqslant 1/(cq)$, we obtain the desired inequality

$$\eta \geqslant \frac{6^{-c}}{c(c+1)\ln|R|}.$$

□

B Basic Linear Algebra

Proposition B.1. Let V be a finite dimensional $\mathbb{F}_2$-vector space and let $U \subseteq V$ be a linear subspace. For any bilinear form $f : V \otimes V \to \mathbb{F}_2$, we have

$$\operatorname{rank}(f) - 2\operatorname{codim} U \leqslant \operatorname{rank}(f|_{U\otimes U}) \leqslant \operatorname{rank}(f).$$

Proof. Let $\iota : U \hookrightarrow V$ be the embedding map, and let $g = f|_{U\otimes U}$. Then we have the commutative diagram

$$\begin{tikzcd} U \arrow[r, "\iota"] \arrow[rrr, bend right, "T_g"'] & V \arrow[r, "T_f"] & V^* \arrow[r, "\iota^\dagger"] & U^* \end{tikzcd}$$

Both composing with ι on the right and composing with $\iota^\dagger$ on the left reduce the rank by at most $\operatorname{codim} U$, so we have

$$\operatorname{rank}(T_f) - 2\operatorname{codim} U \leqslant \operatorname{rank}(T_g) \leqslant \operatorname{rank}(T_f),$$

as desired. □

Corollary B.2. *Let V be a finite dimensional $\mathbb{F}_2$-vector space and let $U \subseteq V$ be a linear subspace. Suppose $f : V \otimes V \to \mathbb{F}_2$ is a bilinear form such that $\operatorname{rank}(f|_{U\otimes U}) = \operatorname{rank}(f)$. Then for any $v \in V$, there exists $u \in U$ such that*

$$f(v \otimes x) = f(u \otimes x) \qquad \text{for all } x \in V.$$

In other words, letting $g = f|_{U\otimes U}$, we have $h|_U \in \operatorname{image}(T_g)$ for each $h \in \operatorname{image}(T_f)$.

Proof. Let $\iota : U \hookrightarrow V$ be the embedding map. Since $\operatorname{rank}(f|_{U\otimes U}) = \operatorname{rank}(f)$, it follows from the proof of Proposition B.1 that

$$\operatorname{image}(T_f \circ \iota) = \operatorname{image}(T_f),$$

which is exactly the desired conclusion. □

Corollary B.3. *For any positive integers n, r and any symmetric bilinear form $f : \mathbb{F}_2^n \otimes \mathbb{F}_2^n \to \mathbb{F}_2$ with rank r, there exists an r-element subset $S \subseteq [n]$ such that the restriction of f to $\mathbb{F}_2^S \otimes \mathbb{F}_2^S$ has rank r.*

Proof. Since $\dim \ker(T_f) = n - r$, there exists an r-element subset $S \subseteq [n]$ such that

$$\mathbb{F}_2^S + \ker(T_f) = \mathbb{F}_2^n. \tag{B.1}$$

Let $\iota : \mathbb{F}_2^S \to \mathbb{F}_2^n$ be the embedding map. To show that the restriction of f to $\mathbb{F}_2^S \otimes \mathbb{F}_2^S$ has rank r, it suffices to show that

$$\operatorname{rank}(\iota^\dagger \circ T_f \circ \iota) = \operatorname{rank}(T_f).$$

The symmetry of f implies

$$(T_f \circ \iota)^\dagger = \iota^\dagger \circ T_f,$$

so it suffices to show that $\operatorname{rank}(T_f \circ \iota) = \operatorname{rank}(T_f)$. But this is immediate due to Equation (B.1). □

Lemma B.4. *Let V be a finite dimensional $\mathbb{F}_2$-vector space, and let r be a positive integer. For any symmetric bilinear form $f : V \otimes V \to \mathbb{F}_2$ with rank r, there exist a nonnegative integer $k \leqslant 3r/2$ and k vectors $v_1, \ldots, v_k \in V$ such that*

$$f(x \otimes y) = \sum_{i=1}^{k} f(x \otimes v_i) f(y \otimes v_i) \qquad \textit{for all } x, y \in V.$$

In other words, any symmetric bilinear form of rank r has a tensor decomposition (see Definition 4.12) into at most $3r/2$ linear functionals.

Proof. We proceed by induction on r. The conclusion clearly holds for $r = 0$. In the rest of the proof, assume $r \geqslant 1$ and the conclusion is true for all smaller values of r. Consider the following two cases.

Case 1: there exists a vector $v \in V$ such that $f(v \otimes v) = 1$. Define a subspace

$$U = \{u \in V \mid f(u \otimes v) = 0\} \subseteq V.$$

It is clear that $\ker(T_f) \subseteq U$, and since $v \notin U$ we also have $\dim U \leqslant \dim V - 1$. Therefore,

$$\operatorname{rank}(f|_{U \otimes U}) \leqslant \dim U - \dim \ker(T_f) \leqslant \dim V - \dim \ker(T_f) - 1 = \operatorname{rank}(f) - 1.$$

Define a linear map $\pi : V \to V$ by letting $\pi(x) = x - f(x \otimes v)v$ for all $x \in V$. It is clear that

$$f(\pi(x) \otimes v) = f(x \otimes v) - f(x \otimes v)^2 = 0 \qquad \text{for all } x \in V,$$

and hence the image of π is contained in U. Furthermore, for each $x, y \in V$ we have

$$\begin{aligned} f\big(\pi(x) \otimes \pi(y)\big) &= f(x \otimes y) - 2f(x \otimes v)f(y \otimes v) + f(x \otimes v)f(y \otimes v)f(v \otimes v) \\ &= f(x \otimes y) - f(x \otimes v)f(y \otimes v). \end{aligned} \tag{B.2}$$

Applying the induction hypothesis to the bilinear form $f\big(\pi(\cdot) \otimes \pi(\cdot)\big)$ (which has rank at most $r - 1$ because $\operatorname{image}(\pi) \subseteq U$ and $\operatorname{rank}(f|_{U \otimes U}) \leqslant r - 1$), we get k vectors $v_1, \ldots, v_k \in V$ with $k \leqslant 3(r-1)/2$ such that

$$f\big(\pi(x) \otimes \pi(y)\big) = \sum_{i=1}^{k} f(x \otimes v_i) f(y \otimes v_i) \qquad \text{for all } x, y \in V. \tag{B.3}$$

Combining Equations (B.2) and (B.3), we get

$$f(x \otimes y) = f(x \otimes v)f(y \otimes v) + \sum_{i=1}^{k} f(x \otimes v_i) f(y \otimes v_i) \qquad \text{for all } x, y \in V,$$

as desired (noting that $k+1 \leqslant 3r/2$).

Case 2: $f(v\otimes v)=0$ holds for all $v \in V$. Since $\mathrm{rank}(f)\geqslant 1$, there exist $v,w\in V$ such that $f(v\otimes w)=1$. It follows from the case assumption that v and w are linearly independent. Define a subspace

$$U=\{u\in V \mid f(u\otimes v)=f(u\otimes w)=0\}\subseteq V.$$

The rest of the argument is entirely analogous to Case 1. Since $U\cap \mathrm{span}\{v,w\}=\{0\}$ we have $\dim U\leqslant \dim V-2$. Due to $\ker(T_f)\subseteq U$, we then have $\mathrm{rank}(f|_{U\otimes U})\leqslant r-2$. Define a linear map $\pi: V\to V$ by letting

$$\pi(x)=x-f(x\otimes v)w-f(x\otimes w)v \qquad \text{for all } x\in V.$$

Similarly to Case 1, it is easy to verify that

$$f(\pi(x)\otimes v)=f(\pi(x)\otimes w)=0 \qquad \text{for all } x\in V,$$

and hence the image of π is contained in U. Furthermore, for each $x,y\in V$ one gets

$$\begin{aligned}&f\big(\pi(x)\otimes\pi(y)\big)\\ &=f(x\otimes y)-f(x\otimes v)f(y\otimes w)-f(x\otimes w)f(y\otimes v)\\ &=f(x\otimes y)-f(x\otimes v)f(y\otimes v)-f(x\otimes w)f(y\otimes w)-f(x\otimes(v+w))f(y\otimes(v+w)).\end{aligned}$$

As in Case 1, we then apply the induction hypothesis to decompose $f\big(\pi(x)\otimes\pi(y)\big)$ into the sum of at most $3(r-2)/2$ terms of the form $f(x\otimes v_i)f(y\otimes v_i)$. Thus $f(x\otimes y)$ can be written as the sum of at most $3(r-2)/2+3=3r/2$ such terms, as desired. $\square$

Lemma B.5. *Let V be a finite dimensional $\mathbb{F}_2$-vector space, and let $Q,W\subseteq V$ be linear subspaces such that $Q\subseteq W$. Denote $q=\dim Q$. Let ℓ be a positive integer such that $q+2\leqslant \ell\leqslant \dim W$. Then for any bilinear form $f: V\otimes V\to\mathbb{F}_2$ such that* $\mathrm{rank}(f)\geqslant 2+2\,\mathrm{codim}\,W$, *we have*

$$\mathop{\mathbb{P}}_{U\in\mathsf{Gr}(W,\ell)}\big[\mathrm{rank}(f|_{U\otimes U})\leqslant 1 \,\big|\, Q\subseteq U\big]\leqslant 2^{4-(\ell-q)/2},$$

Proof. Let $g=f|_{W\otimes W}$. By Proposition B.1, the assumption that $\mathrm{rank}(f)\geqslant 2+2\,\mathrm{codim}\,W$ implies $\mathrm{rank}(g)\geqslant 2$. It suffices to work with the restricted map g in the rest of the proof.

Let $\ell'=\lfloor(\ell-q)/2\rfloor\geqslant 1$. To pick a uniformly random $U\in\mathsf{Gr}(W,\ell)$ that contains Q, we first pick uniformly random subspaces $U_1,U_2\in\mathsf{Gr}(W,\ell')$ independently, and then pick a uniformly random $U\in\mathsf{Gr}(W,\ell)$ conditioned on $Q,U_1,U_2\subseteq U$. Since we always have

$$\mathrm{rank}(f|_{U\otimes U})=\mathrm{rank}(g|_{U\otimes U})\geqslant \mathrm{rank}(g|_{U_1\otimes U_2}),$$

it suffices to show that

$$\mathop{\mathbb{P}}_{U_1,U_2\in\mathsf{Gr}(W,\ell')}\big[\mathrm{rank}(g|_{U_1\otimes U_2})\leqslant 1\big]\leqslant 2^{3-\ell'}.$$

As in the proof of Proposition B.1, we let $\iota_{U_1}: U_1\hookrightarrow W$ and $\iota_{U_2}: U_2\hookrightarrow W$ be the two canonically defined embedding maps. Then we have

$$\mathrm{rank}(g|_{U_1\otimes U_2})=\mathrm{rank}\left(\iota_{U_2}^{\dagger}\circ T_g\circ\iota_{U_1}\right).$$

Note that

$$\mathrm{rank}(T_g\circ\iota_{U_1})=\dim(U_1+\ker(T_g))-\dim\ker(T_g).$$

Therefore, denoting $w = \dim W$ and $r = \operatorname{rank}(g) = w - \dim\ker(T_g)$, we can calculate

$$\begin{aligned}
&\mathop{\mathbb{P}}_{U_1 \in \mathsf{Gr}(W,\ell')}\big[\operatorname{rank}(T_g \circ \iota_{U_1}) \leqslant 1\big] \\
&= \mathop{\mathbb{P}}_{U_1 \in \mathsf{Gr}(W,\ell')}\big[\dim(U_1 + \ker(T_g)) - \dim\ker(T_g) \leqslant 1\big] \\
&\leqslant \sum_{K \in \mathsf{Gr}(W, w-r+1),\ K \supseteq \ker(T_g)} \mathop{\mathbb{P}}_{U_1 \in \mathsf{Gr}(W,\ell')}\big[U_1 \subseteq K\big] \\
&\leqslant (2^r - 1) \cdot \begin{bmatrix} w-r+1 \\ \ell' \end{bmatrix}_2 \begin{bmatrix} w \\ \ell' \end{bmatrix}_2^{-1} = (2^r - 1) \cdot \prod_{i=0}^{\ell'-1} \frac{2^{w-r-i+1} - 1}{2^{w-i} - 1} \\
&\leqslant 2^r \cdot (2^{-r+1})^{\ell'} \leqslant 2^{-\ell'+2}. \qquad \text{(using } r \geqslant 2 \text{ and } \ell' \geqslant 1\text{)}
\end{aligned}$$

Similarly, since

$$\operatorname{rank}\left(\iota_{U_2}^{\dagger} \circ T_g \circ \iota_{U_1}\right) = \dim\left(U_2 + \ker\left(\iota_{U_1}^{\dagger} \circ T_g^{\dagger}\right)\right) - \dim\ker\left(\iota_{U_1}^{\dagger} \circ T_g^{\dagger}\right),$$

an analogous calculation shows

$$\mathop{\mathbb{P}}_{U_2 \in \mathsf{Gr}(W,\ell')}\left[\operatorname{rank}\left(\iota_{U_2}^{\dagger} \circ T_g \circ \iota_{U_1}\right) \leqslant 1\right] \leqslant 2^{-\ell'+2} \qquad \text{whenever} \qquad \dim\ker\left(\iota_{U_1}^{\dagger} \circ T_g^{\dagger}\right) \leqslant \dim W - 2.$$

But $\dim\ker\left(\iota_{U_1}^{\dagger} \circ T_g^{\dagger}\right) \leqslant \dim W - 2$ happens exactly when $\operatorname{rank}(T_g \circ \iota_{U_1}) \geqslant 2$, the probability of which has already been shown to be at least $1 - 2^{-\ell'+2}$. Therefore, we conclude that

$$\mathop{\mathbb{P}}_{U_1, U_2 \in \mathsf{Gr}(W,\ell')}\left[\operatorname{rank}\left(\iota_{U_2}^{\dagger} \circ T_g \circ \iota_{U_1}\right) \leqslant 1\right] \leqslant 2 \cdot 2^{-\ell'+2} = 2^{-\ell'+3},$$

as desired. □

C Covering Properties

C.1 Covering in the Outer PCP

Proposition 4.7. Let E be a tripled set. Fix nonnegative integers q_1, q_2 and a parameter $\beta \in [0,1)$. Fix arbitrary tuples

$$A_1 \in (\mathbb{F}_2^E)^{q_1} \qquad \text{and} \qquad A_2 \in (\mathbb{F}_2^E)^{q_2}.$$

Sample

$$(E', A_{\mathsf{in}}, A_{\mathsf{mid}}) \sim \mathcal{J}_{\mathsf{out}}(E, q_1, q_2, \beta)$$

conditioned on the event $\{A_{\mathsf{in}} = A_1\} \cap \{A_{\mathsf{mid}} = A_2\}$. Then the resulting marginal distribution of E' is independent over triples of E. Furthermore, each triple of E is entirely included in E' with conditional probability at least $1 - 4^{q_1+q_2}\beta$.

Proof. We write $q = q_1 + q_2$. Without loss of generality, we assume

$$E = [3k] \qquad \text{and} \qquad \text{Triples}(E) = \big\{\{3i-2, 3i-1, 3i\} \,\big|\, i \in [k]\big\}$$

for some positive integer k.

Let $P \subseteq E$ be the union of the supports of all vectors appearing in the fixed advice sequences A_1 and A_2. For each $i \in [k]$, let

$$P_i = \big\{j \in \{1,2,3\} \,\big|\, 3i - 3 + j \in P\big\}.$$

Consider a sample

$$(E', A_{\mathsf{in}}, A_{\mathsf{mid}}) \sim \mathcal{J}_{\mathsf{out}}([3k], q_1, q_2, \beta)$$

conditioned on the event $\{A_{\mathsf{in}} = A_1\} \cap \{A_{\mathsf{mid}} = A_2\}$. Under this conditioning, the marginal distribution of E' is independent over the triples of E: for every $i \in [k]$,

(1) if $|P_i| \geqslant 2$, then $3i-2, 3i-1$ and $3i$ are all included in E' with probability 1;

(2) if P_i is a singleton set $\{j\}$, then $3i-2, 3i-1$ and $3i$ are all included in E' with probability $1-\eta_1$, and with probability η_1 only $3i-3+j$ is included in E', where

$$\eta_1 = \frac{(\beta/3)(1-2^{-q})}{(\beta/3)(1-2^{-q}) + (1-\beta)(1-2^{-q})2^{-2q}} = \frac{2^{2q}\beta}{2^{2q}\beta + 3(1-\beta)} \leqslant 2^{2q}\beta; \tag{C.1}$$

(3) if $P_i = \emptyset$, then $3i-2, 3i-1$ and $3i$ are all included in E' with probability $1-\eta_0$, and with probability η_0 only (a uniformly random) one of them is included in E', where

$$\eta_0 = \frac{\beta \cdot 2^{-q}}{\beta \cdot 2^{-q} + (1-\beta)2^{-3q}} = \frac{2^{2q}\beta}{2^{2q}\beta + (1-\beta)} \leqslant 2^{2q}\beta. \tag{C.2}$$

In each of the above three cases, the triple $\{3i-2, 3i-1, 3i\}$ is included entirely in E' with probability at least $1 - 2^{2q}\beta = 1 - 4^{q_1+q_2}\beta$. $\square$

Proposition 4.8. Let E be a tripled set. Fix nonnegative integers q_1, q_2 and a parameter $\beta \in [0, 1)$. Fix an arbitrary tuple $A_1 \in (\mathbb{F}_2^E)^{q_1}$. Sample

$$(E', A_{\mathsf{in}}, A_{\mathsf{mid}}) \sim \mathcal{J}_{\mathsf{out}}(E, q_1, q_2, \beta)$$

conditioned on the event $\{A_{\mathsf{in}} = A_1\}$. Then the resulting marginal distribution of A_{mid} has total variation distance at most $4^{q_1+q_2}\beta\sqrt{|E|/3}$ to the uniform distribution over $(\mathbb{F}_2^E)^{q_2}$.

Proof. As in the proof of Proposition 4.7, we assume

$$E = [3k] \qquad \text{and} \qquad \mathrm{Triples}(E) = \big\{\{3i-2, 3i-1, 3i\} \,\big|\, i \in [k]\big\}$$

for some positive integer k. Let $P \subseteq E$ be the union of the supports of all vectors appearing in the fixed advice sequences A_1. For each $i \in [k]$, let $P_i = \big\{j \in \{1,2,3\} \,\big|\, 3i-3+j \in P\big\}$. We organize the proof into the following two steps.

Step 1: description of the distribution. Similarly to Equations (C.1) and (C.2), we let

$$\eta_1 = \frac{2^{2q_1}\beta}{2^{2q_1}\beta + 3(1-\beta)} \leqslant 2^{2q_1}\beta \qquad \text{and} \qquad \eta_0 = \frac{2^{2q_1}\beta}{2^{2q_1}\beta + (1-\beta)} \leqslant 2^{2q_1}\beta.$$

Let $\mathcal{D}_0$ be the uniform distribution over

$$(\mathbb{F}_2^{q_2})^3 = \left\{\big(x^{(1)}, x^{(2)}, x^{(3)}\big) \,\middle|\, x^{(1)}, x^{(2)}, x^{(3)} \text{ are vectors in } \mathbb{F}_2^{q_2}\right\}.$$

For each $j \in [3]$, let $\mathcal{D}_j$ be the uniform distribution over the subset

$$\left\{\big(x^{(1)}, x^{(2)}, x^{(3)}\big) \in (\mathbb{F}_2^{q_2})^3 \,\middle|\, x^{(i)} = 0 \text{ for all } i \in [3] \setminus \{j\}\right\} \subseteq (\mathbb{F}_2^{q_2})^3.$$

For each $i \in [k]$, we define a distribution $\mathcal{D}^{(i)}$ over $(\mathbb{F}_2^{\ell'})^3$ as follows (according to the conditional distribution of $E' \cap \{3i-2, 3i-1, 3i\}$ described in the proof of Proposition 4.7):

(1) if $|P_i| \geqslant 2$, let $\mathcal{D}^{(i)} = \mathcal{D}_0$;

(2) if P_i is a singleton set $\{j\}$, let $\mathcal{D}^{(i)}$ be the mixture $(1-\eta_1)\mathcal{D}_0 + \eta_1\mathcal{D}_j$;

(3) if $P_i = \emptyset$, let $\mathcal{D}^{(i)}$ be the mixture $(1-\eta_0)\mathcal{D}_0 + \frac{1}{3}\eta_0\mathcal{D}_1 + \frac{1}{3}\eta_0\mathcal{D}_2 + \frac{1}{3}\eta_0\mathcal{D}_3$.

It is straightforward to verify that in each of the above three cases, for any tuple $X \in (\mathbb{F}_2^{q_2})^3$ we have

$$\left|\sqrt{\mathcal{D}^{(i)}(X)/\mathcal{D}_0(X)} - 1\right| \leqslant \left|\mathcal{D}^{(i)}(X)/\mathcal{D}_0(X) - 1\right| \leqslant 2^{2q_2}\max\{\eta_0, \eta_1\} \leqslant 2^{2q_2+2q_1}\beta. \tag{C.3}$$

Now consider a sample

$$(E', A_{\mathsf{in}}, A_{\mathsf{mid}}) \sim \mathcal{J}_{\mathsf{out}}([3k], q_1, q_2, \beta)$$

conditioned on the event $\{A_{\mathsf{in}} = A_1\}$. The conditional distribution of E' is as described in the proof of Proposition 4.7 and does not depend on q_2. It is then easy to see from the proof of Proposition 4.7 that the conditional distribution of A_{mid} is the output distribution of the following process:

1. For each $i \in [k]$, sample a tuple of three vectors $\left(x^{(3i-2)}, x^{(3i-1)}, x^{(3i)}\right) \sim \mathcal{D}^{(i)}$.
2. Form the $3k$-by-q_2 matrix $X \in \mathbb{F}_2^{3k\times q_2}$ whose i-th row is $x^{(i)}$ for each $i \in [3k]$. Let $\mathcal{M}_P$ denote the resulting distribution of X.
3. Output the q_2 columns of X.

Step 2: tensorization of Hellinger affinity. It now suffices to show that $\mathcal{M}_P$ is close to uniform over $\mathbb{F}_2^{3k\times q_2}$ (the space of $3k$-by-q_2 matrices).

For each $i \in [k]$, by Equation (C.3) we have that the squared Hellinger distance between $\mathcal{D}^{(i)}$ and $\mathcal{D}_0$ is

$$\frac{1}{2}\cdot \mathop{\mathbb{E}}_{X\sim\mathcal{D}_0}\left[\left(\sqrt{\mathcal{D}^{(i)}(X)/\mathcal{D}_0(X)} - 1\right)^2\right] \leqslant 2^{4q_1+4q_2-1}\beta^2,$$

and therefore the Hellinger affinity between $\mathcal{D}^{(i)}$ and $\mathcal{D}_0$ is

$$\sum_{X\in(\mathbb{F}_2^{q_2})^3}\sqrt{\mathcal{D}^{(i)}(X)\mathcal{D}_0(X)} = 1 - \frac{1}{2}\sum_{X\in(\mathbb{F}_2^{q_2})^3}\left(\sqrt{\mathcal{D}^{(i)}(X)} - \sqrt{\mathcal{D}_0(X)}\right)^2 \geqslant 1 - 2^{4q_1+4q_2-1}\beta^2.$$

The Cartesian products

$$\prod_{i=1}^{k}\mathcal{D}^{(i)} = \mathcal{M}_P \qquad \text{and} \qquad \prod_{i=1}^{k}\mathcal{D}_0 = \mathcal{D}_0^{\otimes k}$$

define two distributions over $\mathbb{F}_2^{3k\times q_2}$. By the tensorization of Hellinger affinity over product distributions, we have

$$\begin{aligned}\sum_{X\in\mathbb{F}_2^{3k\times q_2}}\sqrt{\mathcal{M}_P(X)\mathcal{D}_0^{\otimes k}(X)} &= \prod_{i=1}^{k}\sum_{X\in(\mathbb{F}_2^{q_2})^3}\sqrt{\mathcal{D}^{(i)}(X)\mathcal{D}_0(X)}\\ &\geqslant (1-2^{4q_1+4q_2-1}\beta^2)^k \geqslant 1 - 2^{4q_1+4q_2-1}\beta^2 k.\end{aligned}$$

Consequently, the squared Hellinger distance between $\mathcal{M}_P$ and $\mathcal{D}_0^{\otimes k}$ is

$$\frac{1}{2}\sum_{X\in\mathbb{F}_2^{3k\times q_2}}\left(\sqrt{\mathcal{M}_P(X)}-\sqrt{\mathcal{D}_0^{\otimes k}(X)}\right)^2 \leqslant 2^{4q_1+4q_2-1}\beta^2 k,$$

and hence by Cauchy-Schwarz,[7] the total variation distance between $\mathcal{M}_P$ and $\mathcal{D}_0^{\otimes k}$ is

$$\frac{1}{2}\sum_{X\in\mathbb{F}_2^{3k\times q_2}}\left|\mathcal{M}_P(X)-\mathcal{D}_0^{\otimes k}(X)\right| \leqslant \sqrt{2}\cdot\sqrt{2^{4q_1+4q_2-1}\beta^2 k} = 2^{2q_1+2q_2}\beta\sqrt{k} \leqslant 2^{2q_1+2q_2}\beta\sqrt{k}. \qquad \square$$

C.2 Covering in the Middle PCP

Proposition 5.4. Let E be a tripled set. Fix a parameter $\beta\in[0,1)$ and nonnegative integers q_1, q_2, ℓ such that $|E|\geqslant\ell\geqslant q_2$. Fix arbitrary tuples

$$A_1\in(\mathbb{F}_2^E)^{q_1} \qquad\text{and}\qquad A_2\in(\mathbb{F}_2^E)^{q_2} \text{ such that } \dim\operatorname{span}(A_2)=q_2.$$

Sample

$$(E', A_{\mathsf{in}}, A_{\mathsf{mid}}, U)\sim\mathcal{J}_{\mathsf{mid}}(E, q_1, q_2, \beta, \ell)$$

conditioned on the event $\{A_{\mathsf{in}}=A_1\}\cap\{A_{\mathsf{mid}}=A_2\}$. Then the resulting marginal distribution of U has total variation distance at most

$$4^{q_1+\ell}\beta\sqrt{|E|/3}+2^{\ell-|E|}$$

to the uniform distribution over the set $\left\{U\in\mathsf{Gr}(\mathbb{F}_2^E,\ell)\;\middle|\;\operatorname{span}(A_2)\subseteq U\right\}$.

Proof. We write $q=q_1+q_2$ and $\ell'=\ell-q_2$.

According to Definition 5.1, when $(E', A_{\mathsf{in}}, A_{\mathsf{mid}}, U)$ is sampled from $\mathcal{J}_{\mathsf{mid}}(E, q_1, q_2, \beta, \ell)$, a vector sequence $A'\in(\mathbb{F}_2^{E'})^{\ell'}$ is generated, and then U is taken to be the sum $\operatorname{span}(A_{\mathsf{mid}})+\operatorname{span}(A')$. Let

$$A_{\mathsf{in}}\uplus A_{\mathsf{mid}}\in(\mathbb{F}_2^{E'})^{q_1+q_2}$$

denote the concatenation of the two vector sequences A_{in} and A_{mid}. Note that before any conditioning, the distribution of $(E', A_{\mathsf{in}}\uplus A_{\mathsf{mid}}, A')$ is identical to $\mathcal{J}_{\mathsf{out}}(E, q, \ell', \beta)$. Therefore, it follows from Proposition 4.8 that after conditioning on $\{A_{\mathsf{in}}=A_1\}\cap\{A_{\mathsf{mid}}=A_2\}$,

$$\text{the distribution of } A' \text{ is } \left(4^{q+\ell'}\beta\sqrt{|E|/3}\right)\text{-close to uniform over } (\mathbb{F}_2^E)^{\ell'}. \tag{C.4}$$

Since $\dim\operatorname{span}(A_2)=q_2$, it is easy to see that

$$\mathop{\mathbb{P}}_{A\in(\mathbb{F}_2^E)^{\ell'}}\left[\dim\big(\operatorname{span}(A_2)+\operatorname{span}(A)\big)=\ell\right]\geqslant 1-2^{\ell-|E|}. \tag{C.5}$$

Furthermore, if $A\in(\mathbb{F}_2^E)^{\ell'}$ is a uniformly random vector sequence, then the distribution of $\big(\operatorname{span}(A_2)+\operatorname{span}(A)\big)$ conditioned on the event

$$\left\{\dim\big(\operatorname{span}(A_2)+\operatorname{span}(A)\big)=\ell\right\}$$

is uniform over the set $\left\{U\in\mathsf{Gr}(\mathbb{F}_2^E,\ell)\;\middle|\;\operatorname{span}(A_2)\subseteq U\right\}$. Therefore, it follows from Equations (C.4) and (C.5) that conditioned on $\{A_{\mathsf{in}}=A_1\}\cap\{A_{\mathsf{mid}}=A_2\}$, the distribution of $U=\big(\operatorname{span}(A_{\mathsf{mid}})+\operatorname{span}(A)\big)$ has total variation distance at most

$$4^{q+\ell'}\beta\sqrt{|E|/3}+2^{\ell-|E|}=4^{q_1+\ell}\beta\sqrt{|E|/3}+2^{\ell-|E|}$$

to the uniform distribution over $\left\{U\in\mathsf{Gr}(\mathbb{F}_2^E,\ell)\;\middle|\;\operatorname{span}(A_2)\subseteq U\right\}$. $\square$

[7] For any two probability vectors $p,q\in\mathbb{R}^n$, we have $\|p-q\|_1\leqslant\|\sqrt{p}-\sqrt{q}\|_2\cdot\|\sqrt{p}+\sqrt{q}\|_2\leqslant 2\|\sqrt{p}-\sqrt{q}\|_2$.

C.3 Covering in the Inner PCP

Proposition C.1. Fix positive integer k, t and finite sets $\mathscr{A}_1, \ldots, \mathscr{A}_t$ with $|\mathscr{A}_j| = k$ for each $j \in [t]$. Suppose that for each $j \in [t]$ there is a probability distribution μ_j over $\mathscr{A}_j$ and a surjective map $\pi_j : \mathscr{A}_j \to \mathscr{B}_j$ onto some finite set $\mathscr{B}_j$. Let $\pi : \prod_{j=1}^t \mathscr{A}_j \to \prod_{j=1}^t \mathscr{B}_j$ be the coordinate-wise map defined by

$$\pi(a_1, \ldots, a_t) = (\pi_1(a_1), \ldots, \pi_t(a_t)) \qquad \text{for all } (a_1, \ldots, a_t) \in \textstyle\prod_{j=1}^t \mathscr{A}_j.$$

Define a probability distribution μ on $\prod_{j=1}^t \mathscr{A}_j$ to be the output distribution of the following process:

1. Sample a uniformly random index $c \in [t]$.
2. Sample $a_c \sim \mu_c$. For each $j \in [t] \setminus \{c\}$, sample a_j uniformly at random from $\mathscr{A}_j$.
3. Output $(a_1, \ldots, a_t)$.

For each $b \in \prod_{j=1}^t \mathscr{B}_j$, let $\mathrm{dist}(b)$ be the total variation distance between

(1) the distribution of $a \sim \mu$ conditioned on the event $\{\pi(a) = b\}$ (if $\mathbb{P}_{a\sim\mu}[\pi(a) = b] = 0$, let $\mathrm{dist}(b) = 1$), and

(2) the uniform distribution over the subset $\pi^{-1}(\{b\}) \subseteq \prod_{j=1}^t \mathscr{A}_j$.

Then we have

$$\mathop{\mathbb{E}}_{a\sim\mu}\left[\mathrm{dist}(\pi(a))\right] \leqslant \sqrt{\frac{k}{4t}}.$$

Proof. For each $j \in [t]$, let $f_j : \mathscr{A}_j \to [0, +\infty)$ be the probability density function of μ_j; that is, let $f_j(a) = k \cdot \mu_j(\{a\})$ for all $a \in \mathscr{A}_j$. It is then easy to see that the probability density function of μ is given by

$$f(a_1, \ldots, a_t) := \frac{1}{t}\sum_{j=1}^t f_j(a_j), \qquad \text{for all } (a_1, \ldots, a_t) \in \textstyle\prod_{j=1}^t \mathscr{A}_j.$$

For each $j \in [t]$, let $g_j : \mathscr{B}_j \to [0, +\infty)$ be defined by

$$g_j(b_j) = \begin{cases} \mathbb{E}_{a_j \in \pi_j^{-1}(\{b_j\})}\left[f_j(a_j)\right], & \text{if } \pi_j^{-1}(\{b_j\}) \neq \emptyset, \\ 0, & \text{otherwise.} \end{cases}$$

For each $b \in \prod_{j=1}^t \mathscr{B}_j$, let

$$g(b) = \frac{1}{t}\sum_{j=1}^t g_j(b_j) = \mathop{\mathbb{E}}_{a\in\pi^{-1}(\{b\})}\left[f(a)\right].$$

For any $b \in \prod_{j=1}^t \mathscr{B}_j$ that lies in the support of g, the total variation distance between the distribution of $a \sim \mu$ conditioned on $\{\pi(a) = b\}$ and the uniform distribution over $\pi^{-1}(\{b\})$ is

$$\frac{1}{2} \cdot \mathop{\mathbb{E}}_{a\in\pi^{-1}(\{b\})}\left[\left|\frac{f(a)}{g(b)} - 1\right|\right] \leqslant \frac{1}{2g(b)}\sqrt{\mathop{\mathrm{Var}}_{a\in\pi^{-1}(\{b\})}\left[f(a)\right]} \qquad \text{(by Cauchy-Schwarz)}$$

$$= \frac{1}{2t \cdot g(b)} \sqrt{\sum_{j=1}^{t} \operatorname*{Var}_{a_j \in \pi_j^{-1}(b_j)} [f_j(a_j)]}$$

(using the fact that $\pi^{-1}(\{b\})$ is a product set)

$$\leqslant \frac{1}{2t \cdot g(b)} \sqrt{\sum_{j=1}^{t} k \cdot \mathop{\mathbb{E}}_{a_j \in \pi^{-1}(b_j)} [f_j(a_j)]} \qquad (f_j \text{ is upper-bounded by } k)$$

$$\leqslant \frac{1}{2} \sqrt{\frac{k}{t \cdot g(b)}}. \qquad \text{(by the definition of } g(b))$$

Let ν be the distribution of $\pi(a)$ when a is sampled uniformly from $\prod_{j=1}^{t} \mathscr{A}_j$. We then have

$$\mathop{\mathbb{E}}_{a \sim \mu} \left[\operatorname{dist}(\pi(a))^2\right] = \mathop{\mathbb{E}}_{b \sim \nu} \left[g(b) \cdot \operatorname{dist}(\pi(a))^2\right] \leqslant \mathop{\mathbb{E}}_{b \sim \nu} \left[\frac{k}{4t}\right] = \frac{k}{4t},$$

and hence $\mathbb{E}_{a \sim \mu} [\operatorname{dist}(\pi(a))] \leqslant \sqrt{k/(4t)}$ by Cauchy-Schwarz. □

D Grassmann Decoding Theorem

D.1 The Expansion Theorem

Definition D.1. Let n, ℓ, q be nonnegative integers such that $n - q \geqslant \ell \geqslant q + 1$, and let $\varepsilon \in (0, 1)$. Suppose $\mathcal{L} \subseteq \mathsf{Gr}(\mathbb{F}_2^n, \ell)$ is a set of subspaces. We say that $\mathcal{L}$ is (q, ε)-*pseudorandom* if for any subspaces $Q \in \mathsf{Gr}(\mathbb{F}_2^n, q)$ and $W \in \mathsf{Gr}(\mathbb{F}_2^n, n - q)$ such that $Q \subseteq W$, we have

$$\mathop{\mathbb{P}}_{L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)} \left[L \in \mathcal{L} \,\middle|\, Q \subseteq L \subseteq W\right] \leqslant \varepsilon,$$

Definition D.2. Let n, ℓ, q be nonnegative integers such that $n - q \geqslant \ell \geqslant q + 1$, and let $\varepsilon \in (0, 1)$. A function $f : \mathbb{F}_2^{n \times \ell} \to \{0, 1\}$ is said to be (q, ε)-*pseudorandom* if for any matrices

$$Y_1 \in \mathbb{F}_2^{\ell \times q}, \qquad Z_1 \in \mathbb{F}_2^{n \times q}, \qquad Y_2 \in \mathbb{F}_2^{q \times n} \qquad \text{and} \qquad Z_2 \in \mathbb{F}_2^{q \times \ell},$$

such that $\left\{X \in \mathbb{F}_2^{n \times \ell} \,\middle|\, XY_1 = Z_1 \text{ and } Y_2 X = Z_2\right\}$ is nonempty, we have

$$\mathop{\mathbb{E}}_{X \in \mathbb{F}_2^{n \times \ell}} \left[f(X) \,\middle|\, XY_1 = Z_1 \text{ and } Y_2 X = Z_2\right] \leqslant \varepsilon.$$

Lemma D.3 ([MZ26, Lemma A.18])**.** *Let n, ℓ, q be nonnegative integers such that $n - q \geqslant \ell \geqslant q + 1$, and let $\varepsilon \in (0, 1)$. Suppose $\mathcal{L} \subseteq \mathsf{Gr}(\mathbb{F}_2^n, \ell)$ is (q, ε)-pseudorandom (in the sense of Definition D.1). Define a function $f : \mathbb{F}_2^{n \times \ell} \to \{0, 1\}$ by*

$$f(X) = \begin{cases} 1, & \text{if } \operatorname{rank}(X) = \ell \text{ and } \operatorname{col}(X) \in \mathcal{L}, \\ 0, & \text{otherwise.} \end{cases}$$

Then f is $(q, 4\varepsilon)$-pseudorandom (in the sense of Definition D.2).

Definition D.4. Given a function $f : \mathbb{F}_2^{n \times \ell} \to \mathbb{R}$, we let $\mathcal{T} f : \mathbb{F}_2^{n \times \ell} \to \mathbb{R}$ be the function defined by

$$\mathcal{T} f(X) = \mathop{\mathbb{E}}_{Y_1 \in \mathbb{F}_2^{n \times 1}, Y_2 \in \mathbb{F}_2^{1 \times \ell}} \left[f(X + Y_1 Y_2)\right]. \tag{D.1}$$

Lemma D.5 ([EKL23, Theorem 69])**.** *There is an absolute constant $C > 0$ such that the following holds. Let n, ℓ, q be nonnegative integers such that $n - q \geqslant \ell \geqslant q + 1$. Then any $(q, 2^{-Cq^2})$-pseudorandom function $f : \mathbb{F}_2^{n\times\ell} \to \{0, 1\}$ satisfies*

$$\langle f, \mathcal{T} f\rangle \leqslant 2^{-q+1}\mathbb{E}[f].$$

Lemma D.6. *There is an absolute constant $C > 0$ such that the following holds. Let n, ℓ, q be nonnegative integers such that $n - q \geqslant \ell \geqslant q + 1$. Then any $(q, 2^{-Cq^2})$-pseudorandom subset $\mathcal{L} \subseteq \mathsf{Gr}(\mathbb{F}_2^n, \ell)$ satisfies*

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[L_1, L_2 \in \mathcal{L}\right] \leqslant 2^{-q+2} \cdot \mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[L \in \mathcal{L}\right].$$

Proof. If $q = 0$ the conclusion obviously holds. We assume $q \geqslant 1$, and let C be the constant from Lemma D.5. Suppose $\mathcal{L} \subseteq \mathsf{Gr}(\mathbb{F}_2^n, \ell)$ is $(q, 2^{-(C+2)q^2})$-pseudorandom. Define a function $f : \mathbb{F}_2^{n\times\ell} \to \{0, 1\}$ from $\mathcal{L}$ as in Equation (D.1). By Lemma D.3 we know that f is $(q, 2^{-Cq^2})$-pseudorandom (in the sense of Definition D.2).

If $X \in \mathbb{F}_2^{n\times\ell}$, $Y_1 \in \mathbb{F}_2^{n\times 1}$ and $Y_2 \in \mathbb{F}_2^{1\times\ell}$ are independent and uniformly random, then conditioned on the event

$$\{\mathrm{rank}(X) = \ell\} \cap \{Y_1 \notin \mathrm{col}(X)\} \cap \{Y_2 \neq 0\}, \tag{D.2}$$

the pair $\big(\mathrm{col}(X), \mathrm{col}(X + Y_1Y_2)\big)$ is a uniformly random edge of the Grassmann graph $\mathsf{Gr}(\mathbb{F}_2^n, \ell)$. Note that the probability of the event Equation (D.2) is at least $1 - 2^{\ell-n} - 2^{\ell-n} - 2^{-\ell} \geqslant 1/2$. Therefore, we have

$$\mathop{\mathbb{P}}_{X\in\mathbb{F}_2^n,\, Y_1\in\mathbb{F}_2^{n\times 1},\, Y_2\in\mathbb{F}_2^{1\times\ell}}\left[f(X) = f(X + Y_1Y_2) = 1\right] \geqslant \frac{1}{2} \cdot \mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[L_1, L_2 \in \mathcal{L}\right] \tag{D.3}$$

Since the left-hand side of Equation (D.3) is exactly $\langle f, \mathcal{T} f\rangle$, using Lemma D.5 we can conclude that

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[L_1, L_2 \in \mathcal{L}\right] \leqslant 2\langle f, \mathcal{T} f\rangle \leqslant 2^{-q+2} \cdot \mathbb{E}[f] \leqslant 2^{-q+2} \cdot \mathop{\mathbb{P}}_{L\in\mathsf{Gr}(n,\ell)}[L \in \mathcal{L}]. \qquad \square$$

D.2 From Expansion to Decoding

Theorem D.7. *There is an absolute constant $C > 0$ such that the following holds. Let n, ℓ, q be nonnegative integers such that $n - q \geqslant \ell \geqslant q + 1$. Suppose $F[\cdot]$ is a table where $F[L] \in L^* \cup \{\mathtt{nil}\}$ for each $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$, such that*

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[F[L_1]|_{L_1\cap L_2} = F[L_2]|_{L_1\cap L_2}\right] \geqslant 2^{-q+4}. \tag{D.4}$$

Then there exist a linear map $f : \mathbb{F}_2^n \to \mathbb{F}_2$ and subspaces $Q \in \mathsf{Gr}(\mathbb{F}_2^n, q)$, $W \in \mathsf{Gr}(\mathbb{F}_2^n, n - q)$ such that $Q \subseteq W$ and

$$\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[F[L] = f|_L \,\middle|\, Q \subseteq L \subseteq W\right] \geqslant 2^{-Cq^2}.$$

Proof. For each linear functional $f : \mathbb{F}_2^n \to \mathbb{F}_2$, let $\mathcal{L}_f$ denote the set of subspaces $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$ such that $F[L] = f|_L$. We have

$$\mathop{\mathbb{E}}_{f\in(\mathbb{F}_2^n)^*}\left[\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[L \in \mathcal{L}_f\right]\right] = \mathop{\mathbb{E}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[\mathop{\mathbb{P}}_{f\in(\mathbb{F}_2^n)^*}\left[f|_L = F[L]\right]\right] = 2^{-\ell}. \tag{D.5}$$

Furthermore,

$$\mathop{\mathbb{E}}_{f\in(\mathbb{F}_2^n)^*}\left[\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[L_1,L_2\in\mathcal{L}_f\right]\right]=\mathop{\mathbb{E}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[\mathop{\mathbb{P}}_{f\in(\mathbb{F}_2^n)^*}\left[f|_{L_1}=F[L_1]\text{ and }\ f|_{L_2}=F[L_2]\right]\right] \tag{D.6}$$

Note that for any edge (L_1, L_2) of the Grassmann graph $\mathsf{Gr}(\mathbb{F}_2^n, \ell)$ such that

$$F[L_1]|_{L_1\cap L_2}=F[L_2]|_{L_1\cap L_2}\,,$$

we have

$$\mathop{\mathbb{P}}_{f\in(\mathbb{F}_2^n)^*}\left[f|_{L_1}=F[L_1]\text{ and }\ f|_{L_2}=F[L_2]\right]=2^{-\ell-1}.$$

Therefore, combining Equations (D.4) and (D.6) yields

$$\mathop{\mathbb{E}}_{f\in(\mathbb{F}_2^n)^*}\left[\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[L_1,L_2\in\mathcal{L}_f\right]\right]\geqslant 2^{-\ell-1}\cdot 2^{-q+4}=2^{-q-\ell+3}. \tag{D.7}$$

Combining Equations (D.5) and (D.7), we know that there exists a linear functional $f\in(\mathbb{F}_2^n)^*$ such that

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[L_1,L_2\in\mathcal{L}_f\right]>2^{-q+2}\cdot\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[L\in\mathcal{L}_f\right].$$

It then follows from Lemma D.6 that $\mathcal{L}_f$ is not $(q, 2^{-Cq^2})$-pseudorandom, which is exactly the desired conclusion. ☐

Theorem 3.2. *There is an absolute constant $C>0$ such that the following holds. Let n,ℓ,q be nonnegative integers such that $n-q\geqslant\ell\geqslant q+1$. Suppose $F[\cdot]$ is a table where $F[L]\in L^*\cup\{\texttt{nil}\}$ for each $L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)$, such that*

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[F[L_1]|_{L_1\cap L_2}=F[L_2]|_{L_1\cap L_2}\right]\geqslant 2^{-q+5}. \tag{3.1}$$

Then for at least $2^{-q-\ell^2}$ fraction of subspaces $Q\in\mathsf{Gr}(\mathbb{F}_2^n,q)$, there exist a subspace $W\in\mathsf{Gr}(\mathbb{F}_2^n,n-q)$ containing Q and a linear map $f:W\to\mathbb{F}_2$ such that

$$\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[F[L]=f|_L\;\middle|\;Q\subseteq L\subseteq W\right]\geqslant 2^{-Cq^2}.$$

Proof. We take C to be the same constant as in Theorem D.7. Let $\mathcal{Q}$ be the collection of subspaces $Q\in\mathsf{Gr}(\mathbb{F}_2^n,q)$ that satisfy the desired property, and we will prove that

$$|\mathcal{Q}|\geqslant 2^{-q-\ell^2}|\mathsf{Gr}(\mathbb{F}_2^n,q)|.$$

In the rest of the proof, we assume on the contrary that $|\mathcal{Q}|<2^{-q-\ell^2}|\mathsf{Gr}(\mathbb{F}_2^n,q)|$.

We define a table $F'[\cdot]$, where $F'[L]\in L^*\cup\{\texttt{nil}\}$ for each $L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)$, as follows:

$$F'[L]=\begin{cases}\texttt{nil}, & \text{if } Q\subseteq L \text{ for some } Q\in\mathcal{Q},\\ F[L], & \text{otherwise.}\end{cases}$$

We thus have

$$\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,q)}\left[F[L]\neq F'[L]\right]\leqslant\sum_{Q\in\mathcal{Q}}\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,q)}\left[Q\subseteq L\right]=|\mathcal{Q}|\cdot\begin{bmatrix}n\\ \ell-q\end{bmatrix}_2\begin{bmatrix}n\\ \ell\end{bmatrix}_2^{-1}$$

$$\leqslant |\mathcal{Q}| \cdot 2^{n(\ell-q)} \cdot 2^{\ell(n-\ell)} = |\mathcal{Q}| \cdot 2^{\ell^2-nq} \leqslant 2^{\ell^2}|\mathcal{Q}| \cdot \begin{bmatrix} n \\ q \end{bmatrix}_2^{-1} \leqslant 2^{-q}.$$

Therefore, it follows from the assumption Equation (3.1) that

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[F'[L_1]\big|_{L_1\cap L_2} = F'[L_2]\big|_{L_1\cap L_2}\right] \geqslant 2^{-q+5} - 2\cdot 2^{-q} \geqslant 2^{-q+4}.$$

Applying Theorem D.7 to the table F' yields subspaces $Q \in \mathsf{Gr}(\mathbb{F}_2^n, q)$, $W \in \mathsf{Gr}(\mathbb{F}_2^n, n-q)$ such that $Q \subseteq W$ and

$$\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[F'[L] = f|_L \;\middle|\; Q \subseteq L \subseteq W\right] \geqslant 2^{-Cq^2}. \tag{D.8}$$

By the definition of $\mathcal{Q}$, this implies that $Q \in \mathcal{Q}$. However, as $F'[L] = \texttt{nil}$ for any $Q \in \mathcal{Q}$, the left-hand side of Equation (D.8) must be 0, which is a contradiction. □

Lemma D.8. *There is an absolute constant $C > 0$ such that the following holds. Let n, ℓ, q be nonnegative integers such that $n - q \geqslant \ell \geqslant q + 1$. Suppose Σ is a finite set and $F[\cdot]$ is a table where $F[L] \in \Sigma \cup \{\texttt{nil}\}$ for each $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$, such that*

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[F[L_1] = F[L_2] \neq \texttt{nil}\right] \geqslant 2^{-q+3}. \tag{D.9}$$

Then there exist an element $\sigma \in \Sigma$ and subspaces $Q \in \mathsf{Gr}(\mathbb{F}_2^n, q)$, $W \in \mathsf{Gr}(\mathbb{F}_2^n, n-q)$ such that $Q \subseteq W$ and

$$\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[F[L] = \sigma \;\middle|\; Q \subseteq L \subseteq W\right] \geqslant 2^{-Cq^2}.$$

Proof. For each element $\sigma \in \Sigma$, let $\mathcal{L}_\sigma$ denote the set of subspaces $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$ such that $F[L] = \sigma$. We have

$$\mathop{\mathbb{E}}_{\sigma\in\Sigma}\left[\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[L \in \mathcal{L}_\sigma\right]\right] = \mathop{\mathbb{E}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[\mathop{\mathbb{P}}_{\sigma\in\Sigma}\left[\sigma = F[L]\right]\right] = \frac{1}{|\Sigma|}. \tag{D.10}$$

Furthermore, due to the assumption of the lemma, we have

$$\begin{aligned}\mathop{\mathbb{E}}_{\sigma\in\Sigma}\left[\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[L_1, L_2 \in \mathcal{L}_\sigma\right]\right] &= \mathop{\mathbb{E}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[\mathop{\mathbb{P}}_{\sigma\in\Sigma}\left[\sigma = F[L_1] = F[L_2]\right]\right] \\ &= \frac{1}{|\Sigma|} \cdot \mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[F[L_1] = F[L_2] \neq \texttt{nil}\right] \geqslant \frac{2^{-q+3}}{|\Sigma|}. \end{aligned} \tag{D.11}$$

Therefore, combining Equations (D.10) and (D.11), we know that there exists an element $\sigma \in \Sigma$ such that

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[L_1, L_2 \in \mathcal{L}_\sigma\right] > 2^{-q+2} \cdot \mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[L \in \mathcal{L}_\sigma\right].$$

It then follows from Lemma D.6 that $\mathcal{L}_\sigma$ is not $(q, 2^{-Cq^2})$-pseudorandom, which is exactly the desired conclusion. □

Theorem 3.3. *There is an absolute constant $C > 0$ such that the following holds. Let n, ℓ, q be nonnegative integers such that $n - q \geqslant \ell \geqslant q + 1$. Suppose Σ is a finite set and $F[\cdot]$ is a table where $F[L] \in \Sigma \cup \{\texttt{nil}\}$ for each $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$, such that*

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell)}\left[F[L_1] = F[L_2] \neq \texttt{nil}\right] \geqslant 2^{-q+4}.$$

Then for at least $2^{-q-\ell^2}$ fraction of subspaces $Q \in \mathsf{Gr}(\mathbb{F}_2^n, q)$, there exist an element $\sigma \in \Sigma$ and a subspaces $W \in \mathsf{Gr}(\mathbb{F}_2^n, n-q)$ containing Q such that

$$\mathop{\mathbb{P}}_{L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)} \left[F[L] = \sigma \,\middle|\, Q \subseteq L \subseteq W \right] \geqslant 2^{-Cq^2}.$$

Proof. One can use Lemma D.8 to prove Theorem 3.3 in an entirely analogous way to how Theorem D.7 is used to prove Theorem 3.2. □

E List Decoding Bounds

The following lemma is from [KMS25, Theorem B.1] (see also [DKK+25, Lemma 3.1]).

Lemma E.1 ([KMS25, DKK+25])**.** *There is an absolute constant $C > 0$ such that the following holds. Let n, ℓ, b be positive integers such that $n \geqslant 2\ell$ and $40 \leqslant b \leqslant \ell/4$. Suppose $F[\cdot]$ is a table where $F[L] \in L^*$ for each $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$, such that*

$$\mathop{\mathbb{P}}_{(L_1, L_2) \sim \mathcal{G}(\mathbb{F}_2^n, \ell, b)} \left[F[L_1]|_{L_1 \cap L_2} = F[L_2]|_{L_1 \cap L_2} \right] \geqslant \varepsilon,$$

for some parameter $\varepsilon \in [2^{4-b/4}, 1]$. Then there exists a linear map $f : \mathbb{F}_2^n \to \mathbb{F}_2$ such that

$$\mathop{\mathbb{P}}_{L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)} \left[F[L] = f|_L \right] \geqslant C^{-1} \varepsilon^3.$$

Lemma E.2. *There is an absolute constant $C \geqslant 40$ such that the following holds. Let n, ℓ, b be positive integers such that $n \geqslant 2\ell$ and $C \leqslant b \leqslant \ell/4$. Suppose $F[\cdot]$ is a table where $F[L] \in L^* \cup \{\mathtt{nil}\}$ for each $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$, such that*

$$\mathop{\mathbb{P}}_{(L_1, L_2) \sim \mathcal{G}(\mathbb{F}_2^n, \ell, b)} \left[F[L_1]|_{L_1 \cap L_2} = F[L_2]|_{L_1 \cap L_2} \right] \geqslant \varepsilon, \tag{E.1}$$

for some parameter $\varepsilon \in [2^{4-b/4}, 1]$. Then there exists a linear map $f : \mathbb{F}_2^n \to \mathbb{F}_2$ such that

$$\mathop{\mathbb{P}}_{L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)} \left[F[L] = f|_L \right] \geqslant C^{-1} \varepsilon^3.$$

Proof. We define a random table $F'[\cdot]$ as follows: for each $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$,

(1) if $F[L] \neq \mathtt{nil}$, let $F'[L] = F[L]$;

(2) if $F[L] = \mathtt{nil}$, pick a uniformly random linear functional $h \in L^*$ and let $F'[L] = h$.

Therefore, for each $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$ the entry $F'[L]$ always takes value in the space L^*.

For any fixed linear map $f : \mathbb{F}_2^n \to \mathbb{F}_2$, the random variables

$$Z_f(L) = \mathbb{1}\left[F'[L] = f|_L \text{ and } F[L] = \mathtt{nil} \right] \qquad \text{for } L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$$

take values in $\{0, 1\}$, have expected values at most $2^{-\ell}$, and are mutually independent. Therefore, by Chernoff bound we have

$$\mathop{\mathbb{P}}_{F'} \left[\frac{1}{|\mathsf{Gr}(\mathbb{F}_2^n, \ell)|} \sum_{L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)} Z_f(L) \geqslant 2^{1-\ell} \right] \leqslant \exp\left(-\frac{1}{3} \cdot 2^{-\ell} \cdot |\mathsf{Gr}(\mathbb{F}_2^n, \ell)| \right)$$

$$\leqslant \exp\left(-\frac{1}{3}\cdot 2^{\ell(n-\ell-1)}\right) < 2^{-n},$$

where in the last transition we used the condition that $n \geqslant 2\ell \geqslant 320$. By a union bound over all linear maps $f : \mathbb{F}_2^n \to \mathbb{F}_2$, it follows that with nonzero probability over F' we have

$$\mathop{\mathbb{E}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[Z_f(L)\right] \leqslant 2^{1-\ell}.$$

In other words, there exists a realization of the random table $F'[\cdot]$, which we denote by $F_0[\cdot]$, such that

(1) for each $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$, we have $F_0[L] \in L^*$;

(2) for each $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$, we have $F_0[L] = F[L]$ whenever $F[L] \neq \texttt{nil}$;

(3) $\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[F_0[L] = f|_L \text{ and } F[L] = \texttt{nil}\right] \leqslant 2^{1-\ell}$ for any linear map $f : \mathbb{F}_2^n \to \mathbb{F}_2$.

Since $F_0[\cdot]$ only replaces the $\texttt{nil}$ values in $F[\cdot]$, it is clear that Equation (E.1) also holds for F_0, i.e.

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell,b)}\left[F_0[L_1]|_{L_1\cap L_2} = F_0[L_2]|_{L_1\cap L_2}\right] \geqslant \varepsilon,$$

and hence we may apply Lemma E.1 to obtain a linear map $f : \mathbb{F}_2^n \to \mathbb{F}_2$ such that

$$\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[F_0[L] = f|_L\right] \geqslant C_0^{-1}\varepsilon^3,$$

for some absolute constant $C_0 > 0$. Combining this with the second and third properties of $F_0[\cdot]$ listed above, it follows that

$$\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[F[L] = f|_L\right] \geqslant C_0^{-1}\varepsilon^3 - 2^{1-\ell} \geqslant (2C_0)^{-1}\varepsilon^3,$$

where the last transition holds as long as $\varepsilon^3 \cdot 2^\ell$ is sufficiently large. $\square$

Corollary E.3. *There is an absolute constant $C \geqslant 40$ such that the following holds. Fix positive integers n, ℓ, b and a nonnegative integer q such that $n \geqslant 2\ell$ and $C(q+1) \leqslant b \leqslant \ell/4$. Let $Q \in \mathsf{Gr}(\mathbb{F}_2^n, q)$ be a subspace. Suppose $F[\cdot]$ is a table where $F[L] \in L^* \cup \{\texttt{nil}\}$ for each $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$, such that*

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell,b)}\left[F[L_1]|_{L_1\cap L_2} = F[L_2]|_{L_1\cap L_2} \,\middle|\, Q \subseteq L_1, L_2\right] \geqslant \varepsilon, \tag{E.2}$$

for some parameter $\varepsilon \in [2^{-b/8}, 1]$. Then there exists a linear map $f : \mathbb{F}_2^n \to \mathbb{F}_2$ such that

$$\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[F[L] = f|_L \,\middle|\, Q \subseteq L\right] \geqslant C^{-1}\cdot 2^{-3q}\varepsilon^3.$$

Proof. Since there are exactly 2^q linear functionals $Q \to \mathbb{F}_2$, there exists a linear functional $h \in Q^*$ such that

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(\mathbb{F}_2^n,\ell,b)}\left[F[L_1]|_{L_1\cap L_2} = F[L_2]|_{L_1\cap L_2} \text{ and } F[L_1]|_Q = h \,\middle|\, Q \subseteq L_1, L_2\right] \geqslant 2^{-q}\varepsilon, \tag{E.3}$$

Let $Q^{\mathsf{c}} \in \mathsf{Gr}(\mathbb{F}_2^n, n-q)$ be a linear subspace such that $Q \cap Q^{\mathsf{c}} = \{0\}$. Define a table $F'[\cdot]$ as follows: for each $L \in \mathsf{Gr}(Q^{\mathsf{c}}, \ell - q)$ containing Q,

(1) if $F[L+Q] \neq \texttt{nil}$ and $F[L+Q]\big|_Q = h$, let $F'[L] = F[L+Q]\big|_L$;

(2) if $F[L+Q] = \texttt{nil}$ or $F[L+Q]\big|_Q \neq h$, let $F'[L] = \texttt{nil}$.

It then follows from Equation (E.3) that

$$\mathop{\mathbb{P}}_{(L_1,L_2)\sim\mathcal{G}(Q^{\mathsf{c}},\ell-q,b-q)}\left[F'[L_1]\big|_{L_1\cap L_2} = F'[L_2]\big|_{L_1\cap L_2}\right] \geqslant 2^{-q}\varepsilon.$$

Applying Lemma E.2 to the table $F'[\cdot]$ and then translating the conclusion back to $F[\cdot]$ yields the desired result. □

Lemma E.4. *Fix positive integers n, ℓ, b, a nonnegative integer q and a parameter $\eta \in (0,1)$ such that*

$$n \geqslant 2\ell, \qquad C(q+1) \leqslant b \leqslant \ell/4, \qquad \textit{and} \qquad \eta \geqslant 2^{-b/8}.$$

Let $Q \in \mathsf{Gr}(\mathbb{F}_2^n, q)$ be a subspace. Suppose $F[\cdot]$ is a table where $F[L] \in L^* \cup \{\texttt{nil}\}$ for each $L \in \mathsf{Gr}(\mathbb{F}_2^n, \ell)$. Let $f_1, \ldots, f_m : \mathbb{F}_2^n \to \mathbb{F}_2$ be the list of all linear functions $f : \mathbb{F}_2^n \to \mathbb{F}_2$ such that

$$\mathop{\mathbb{P}}_{L\in\mathsf{Gr}(\mathbb{F}_2^n,\ell)}\left[F[L] = f|_L \;\middle|\; Q \subseteq L\right] \geqslant \eta.$$

Then we have $m \leqslant 2/\eta$ and

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il