Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 27 (2024)

Near Optimal Alphabet-Soundness Tradeoff PCPs

Dor Minzer* Kai Zhe Zheng†

Abstract

We show that for all $\varepsilon > 0$, for sufficiently large $q \in \mathbb{N}$ that is a power of 2, for all $\delta > 0$, it is NP-hard to distinguish whether a given 2-Prover-1-Round projection game with alphabet size q has value at least $1 - \delta$, or value at most $1/q^{1-\varepsilon}$. This establishes a nearly optimal alphabet-to-soundness tradeoff for 2-query PCPs with alphabet size q, improving upon a result of [Chan, J. ACM 2016]. Our result has the following implications:

1. Near optimal hardness for Quadratic Programming: it is NP-hard to approximate the value of a given Boolean Quadratic Program within factor $(\log n)^{1-o(1)}$ under quasi-polynomial time reductions. This improves upon a result of [Khot, Safra, ToC 2013] and nearly matches the performance of the best known algorithms due to [Megretski, IWOTA 2000], [Nemirovski, Roos, Terlaky, Mathematical Programming 1999] and [Charikar, Wirth, FOCS 2004] that achieve $O(\log n)$ approximation ratio.
2. Bounded degree 2-CSPs: under randomized reductions, for sufficiently large $d > 0$, it is NP-hard to approximate the value of 2-CSPs in which each variable appears in at most d constraints to within a factor of $(1 - o(1))\frac{d}{2}$, improving upon a result of [Lee, Manurangsi, ITCS 2024].
3. Improved hardness results for connectivity problems: using results of [Laekhanukit, SODA 2014] and [Manurangsi, Inf. Process. Lett., 2019], we deduce improved hardness results for the Rooted k-Connectivity Problem, the Vertex-Connectivity Survivable Network Design Problem and the Vertex-Connectivity k-Route Cut Problem.

1 Introduction

The PCP theorem is a fundamental result in theoretical computer science with many equivalent formulations [FGL+96, AS98, ALM+98]. One of the formulations asserts that there exists $\varepsilon > 0$ such that given a satisfiable 3-SAT formula ϕ, it is NP-hard to find an assignment that satisfies at least $(1 - \varepsilon)$ fraction of the constraints. The PCP theorem has a myriad of applications within theoretical computer science, and of particular interest to this paper are its applications to hardness of approximation.

The vast majority of hardness of approximation results are proved via reductions from the PCP theorem above. Oftentimes, to get a strong hardness of approximation result, one must first amplify the basic PCP theorem into a result with stronger parameters [Hås01, Hås96, Fei98, KP06] (see [Tre14] for a survey). To discuss these parameters, it is often convenient to view the PCP through the problem of 2-Prover-1-Round Games, which we define next.[1]

*Department of Mathematics, Massachusetts Institute of Technology, Cambridge, USA. Supported by NSF CCF award 2227876 and NSF CAREER award 2239160.

†Department of Mathematics, Massachusetts Institute of Technology, Cambridge, USA. Supported by the NSF GRFP DGE-2141064.

[1]Strictly speaking, the notion below is often referred to as projection 2-Prover-1-Round games. As we only deal with projection games in this paper we omit the more general definition.

ISSN 1433-8092

Definition 1.1. *An instance Ψ of 2-Prover-1-Round Games consists of a bipartite graph $G = (L \cup R, E)$, alphabets Σ_L, Σ_R, and a collection of constraints $\Phi = \{\phi_e\}_{e \in E}$ specified by maps $\phi_e \colon \Sigma_L \to \Sigma_R$.*

1. *The alphabet size of Ψ is defined to be $|\Sigma_L| + |\Sigma_R|$.*
2. *The value of Ψ is defined to be the maximum fraction of edges $e \in E$ that can be satisfied by any assignment. That is,*

$$\mathsf{val}(\Psi) = \max_{\substack{A_L \colon L \to \Sigma_L \\ A_R \colon R \to \Sigma_R}} \frac{|\{e = (u,v) \in E \mid \phi_e(A_L(u)) = A_R(v)\}|}{|E|}.$$

The combinatorial view of 2-Prover-1-Round Games has its origins in an equivalent, active view of PCP in terms of a game between a verifier and two all-powerful provers. The verifier and the two provers have access to an instance Ψ of 2-Prover-1-Round Games, and the provers may agree beforehand on a strategy; after this period they are not allowed to communicate. The verifier then picks a random edge, $e = (u, v)$, from the 2-Prover-1-Round game, sends u to the first prover and v to the second prover, and then receives a label in response from each one of them. Finally, the verifier checks that these labels satisfy the constraint ϕ_e, and if so accepts (and otherwise rejects). It is easy to see that the value of the 2-Prover-1-Round game is equal to the acceptance probability of the verifier under the best strategy of the provers. This view will be useful for us later.

The majority of hardness of approximation results are proved by combining the basic PCP theorem of [FGL$^+$96, AS98, ALM$^+$98] with Raz's parallel repetition theorem [Raz98], which together imply the following result in the language of 2-Prover-1-Round Games:

Theorem 1.2. *There exists $\gamma > 0$ such that for sufficiently large R, given a 2-Prover-1-Round game Ψ with alphabet size R, it is NP-hard to distinguish between the following two cases:*

1. *YES case:* $\mathsf{val}(\Psi) = 1$.
2. *NO case:* $\mathsf{val}(\Psi) \leqslant \frac{1}{R^\gamma}$.

While the soundness guarantee in Theorem 1.2 is often sufficient, some applications require strengthenings of it. In particular, one could require the soundness to be as small as possible with respect to the alphabet size (as opposed to just being small in absolute terms). The tradeoff between the soundness error of the PCP and the alphabet size of the PCP is the main focus of this paper.

Since a random assignment to Ψ satisfies in expectation at least $\frac{1}{R}$ fraction of the constraints, the best alphabet to soundness tradeoff one could hope for in Theorem 1.2 is with $\gamma = 1 - o(1)$. Khot and Safra [KS13] were the first to achieve a reasonable bound for γ, and their result achieves $\gamma = 1/6$ with imperfect completeness (i.e., $\mathsf{val}(\Psi) \geqslant 1 - o(1)$ instead of $\mathsf{val}(\Psi) = 1$ in the YES case). Subsequently, Chan [Cha16] obtained an improvement (using different techniques), showing that one can get $\gamma = 1/2 - o(1)$ (again with imperfect completeness).

1.1 Main Results

1.1.1 Near Optimal Alphabet vs Soundness Tradeoff

The main result of this work improves upon all prior results, and shows that one may take $\gamma = 1 - o(1)$ in Theorem 1.2, again with imperfect completeness. Formally, we show:

Theorem 1.3. *For all $\varepsilon, \tau > 0$, for sufficiently large R, given a 2-Prover-1-Round game Ψ, it is NP-hard to distinguish between the following two cases:*

1. *YES case:* $\mathsf{val}(\Psi) \geqslant 1 - \tau$.
2. *NO case:* $\mathsf{val}(\Psi) \leqslant \frac{1}{R^{1-\varepsilon}}$.

As discussed below, Theorem 1.3 has several applications to combinatorial optimization problems. These applications require additional features from the instances produced in Theorem 1.3 which are omitted from the formulation for the sake of clarity. For instance, one application requires a good tradeoff between the size of the instance and the size of the alphabet, which our construction achieves (see the discussion following Theorem 1.4). Other applications require the underlying constraint graph to be bounded-degree and biregular, which our construction also achieves (after mild modifications; see Theorem 7.2).

1.1.2 Application: NP-Hardness of Approximating Quadratic Programs

An instance of the Quadratic Programming problem consists of a quadratic form $Q(x) = \sum_{i,j=1}^{n} a_{i,j} x_i x_j$ where $a_{i,i} = 0$ for all i, and the goal is to maximize $Q(x)$ over $x \in \{-1, 1\}^n$. It was previously known that this problem admits an $O(\log n)$ approximation algorithm [Meg01, NRT99, CW04] and is NP-hard to approximate within factor $(\log n)^{1/6-o(1)}$ [ABH$^+$05, KS13] under quasi-polynomial reductions. As a first application of Theorem 1.3, we improve the hardness of approximation result for the Quadratic Programming problem:

Theorem 1.4. *It is NP-hard to approximate Quadratic Programming to within a factor of $(\log n)^{1-o(1)}$ under quasi-polynomial reductions.*

Theorem 1.4 is proved via a connection between 2-Prover-1-Round Games and Quadratic Programming due to Arora, Berger, Hazan, Kindler, and Safra [ABH$^+$05]. This connection requires a good tradeoff between the alphabet size, the soundness error of the PCP, and the size of the PCP, and the construction in Theorem 1.4 has a sufficiently good tradeoff: letting N be the size of the instance, the alphabet size is $(\log N)^{1-o(1)}$ and the soundness error is $(\log N)^{-1+o(1)}$.[2]

Relevance to the sliding scale conjecture? We do not know how to use our techniques to achieve soundness error that is smaller than inversely poly-logarithmic in the instance size. In particular, our techniques have no bearing on the sliding scale conjecture, which asserts the existence of a PCP with $O(1)$ queries, polynomial size, polynomial alphabet size, and inverse polynomial soundness [Mos19]. A key feature of our PCP reduction is the so-called "covering property", which seems to inherently limit the soundness of our construction to be at most inversely logarithmic in the instance size.

1.1.3 Application: NP-hardness of Approximating Bounded Degree 2-CSPs

A 2-Constraint Satisfaction Problem (2-CSP) instance $\Psi = (X, C, \Sigma)$ consists of a set of variables X, a set of constraints C, and an alphabet Σ. Each constraint in C has the form $P(x_i, x_j) = 1$ where $P\colon \Sigma \times \Sigma \to \{0, 1\}$ is a predicate (which may be different in distinct constraints). The degree of the instance Ψ is defined

[2] We remark that, due to the use of the long-code, Chan [Cha16] does not achieve a good enough trade-off between the alphabet size and the instance size. This is the reason his result does not yield an improvement over the inapproximability result for Quadratic Programming of Khot and Safra [KS13].

to be the maximum, over variables $x \in X$, of the number of constraints that x appears in. The goal is to find an assignment $A\colon X \to \Sigma$ that satisfies as many of the constraints as possible.

There is a simple $\frac{d+1}{2}$ approximation algorithm for the 2-CSP problem for instances with degree at most d. Lee and Manurangsi proved a nearly matching $\left(\frac{1}{2} - o(1)\right) d$ hardness of approximation result assuming the Unique-Games Conjecture [LM24]. Unconditionally, they show the problem to be NP-hard to approximate within factor $\left(\frac{1}{3} - o(1)\right) d$ under randomized reductions. Using the ideas of Lee and Manurangsi, our main result implies a nearly matching NP-hardness result for bounded degree 2-CSPs:

Theorem 1.5. *For all $\eta > 0$, for sufficiently large d, approximating the value of 2-CSPs with degree at most d within factor $\left(\frac{1}{2} - \eta\right) d$ is NP-hard under randomized reductions.*

As in [LM24], Theorem 1.5 has a further application to finding independent sets in claw free graphs. A k-claw is the $(k+1)$ vertex graph which consists of one center vertex connected to k leaves and has no other edges. A graph G is said to be k-claw-free if G does not contain a k-claw as an induced subgraph. There is a polynomial time algorithm that approximates the size of the largest independent set in a given k-claw-free graph G within factor $\frac{k}{2}$ [Ber00, TW23], and a quasi-polynomial time algorithm due to [CGM13] that achieves a better $\left(\frac{1}{3} + o(1)\right) k$. As in [LM24], using ideas from [DFRR23], Theorem 1.5 implies that for all $\varepsilon > 0$, for sufficiently large k, it is NP-hard (under randomized reductions) to approximate the size of the largest independent set in a given k-claw-free graph within factor $\left(\frac{1}{4} + \eta\right) k$. This improves upon the result of [LM24] who showed that the same result holds assuming the Unique-Games Conjecture.

1.1.4 Application: NP-hardness of Approximating Connectivity Problems

Using ideas of Laekhanukit [Lae14] and some subsequent improvements by Manurangsi [Man19], Theorem 1.3 implies improved hardness of approximation results for several graph connectivity problems. More specifically, it implies new improvements over each of the results outlined in table 1 of [Lae14], with the exception of Rooted-k-Connectivity on directed graphs where the improvement is already implied by [Man19]. Below, we briefly discuss one of these graph connectivity problems — the Rooted k-Connectivity Problem — but defer the reader to [Lae14] for a detailed discussion of the remaining graph connectivity problems considered.

In the Rooted-k-Connectivity problem there is a graph $G = (V, E)$, edge costs $c\colon E \to \mathbb{R}$, a root vertex $r \in V$ and a set of terminals $T \subseteq V \setminus \{r\}$. The goal is to find a subgraph G' of smallest cost that for each $t \in T$, has at least k vertex disjoint paths from r to t. The problem admits a trivial $|T|$-approximation algorithm (by applying minimum cost k-flow algorithm for each vertex in T), as well as an $O(k \log k)$ approximation algorithm [Nut12]. Using the ideas of [Lae14], Theorem 1.3 implies the following improved hardness of approximation results:

Theorem 1.6. *For all $\varepsilon > 0$, for sufficiently large k it is NP-hard under randomized reductions to approximate the Rooted-k-Connectivity problem on undirected graphs to within a factor of $k^{1/5-\varepsilon}$, the Vertex-Connectivity Survivable Network Design Problem with connectivity parameters at most k to within a factor of $k^{1/3-\varepsilon}$, and the Vertex-Connectivity k-Route Cut Problem to within a factor of $k^{1/3-\varepsilon}$.*

We remark that in [CCK08], a weaker form of hardness for the Vertex-Connectivity Survivable Network problem is proved. More precisely, they show an $\Omega(k^{1/3}/\log k)$ integrality gap for the set-pair relaxation of the problem. Our hardness result of $k^{1/3-\varepsilon}$ shows that, unless NP $\subseteq$ BPP, no polynomial-size relaxation can yield a better than $k^{1/3-\varepsilon}$ factor approximation algorithm.

1.2 Our Techniques

Theorem 1.3 is proved by composing an inner PCP and an outer PCP. Both of these components incorporate ideas from the proof of the 2-to-1 Games Theorem. The outer PCP is constructed using smooth parallel repetition [KS13, KMS17] while the inner PCP is based on a codeword test using the Grassmann graph [KMS17, DKK$^+$18, DKK$^+$21, KMS23].

The novelty in this paper, in terms of techniques, is twofold. First, in order to get such a strong alphabet to soundness tradeoff, we must consider a Grassmann test in a different and extreme regime of parameters. Additionally, we have to work with much lower soundness error than prior works. These differences complicate matters considerably. Second, our soundness analysis is more involved than that of the 2-to-1-Games Theorem. As is the case in [KMS17, DKK$^+$18, DKK$^+$21, KMS23], we too use global hypercontractivity, but we do so more extensively. We also require quantitatively stronger versions of global hypercontractivity over the Grassmann graph which are due to [EKL24]. Our analysis also incorporates ideas from the plane versus plane test and direct product testing [RS97, IKW12, MZ23], from classical PCP theory [KS13], as well as from error correcting codes [GRS00]. All of these tools are necessary to prove our main technical statement — Lemma 1.7 below — which is a combinatorial statement perhaps of independent interest.

We now elaborate on each one of the components separately.

1.2.1 The Inner PCP

Our inner PCP is based on the subspace vs subspace low-degree test. Below, we first give a general overview of the objective in low-degree testing. We then discuss the traditional notion of soundness as well as a non-traditional notion of soundness for low-degree tests. Finally, we explain the low-degree test used in this paper, the notion of soundness that we need from it, and the way that this notion of soundness is used.

Low-degree tests in PCPs. Low-degree tests have been a vital component in PCPs since their inception, and much attention has been devoted to improving their various parameters. The goal in low-degree testing is to encode a low-degree function $f\colon \mathbb{F}_q^n \to \mathbb{F}_q$ via a table (or a few tables) of values in a locally testable way. Traditionally, one picks a parameter $\ell \in \mathbb{N}$ (which is thought of as a constant and is most often just 2) and encodes the function f by the table T of restrictions of f to ℓ-dimensional affine subspaces of $\mathbb{F}_q^n$. For the case $\ell = 2$, the test associated with this encoding is known as the Plane vs Plane test [RS97]. The Plane vs Plane test proceeds by picking two planes P_1, P_2 intersecting on a line, and then checking that $T[P_1]$ and $T[P_2]$ agree on $P_1 \cap P_2$. It is easy to see that the test has perfect completeness, namely that a valid table of restrictions T passes the test with probability 1. In the other direction, the soundness error of the test — which is a converse type statement — is much less clear (and is crucial towards applications in PCP). In the context of the Plane vs Plane test, it is known that if a table T, that assigns to each plane a degree d function, passes the Plane vs Plane test with probability $\varepsilon \geqslant q^{-c}$ (where $c > 0$ is a small absolute constant), then there is a degree d function f such that $T[P] \equiv f|_P$ on at least $\Omega(\varepsilon)$ fraction of the planes.

Nailing down the value of the constant c for which soundness holds is an interesting open problem which is related to soundness vs alphabet size vs instance size tradeoff in PCPs [MR10, BDN17, MZ23]. Currently, the best known analysis for the Plane vs Plane test [MR10] shows that one may take $c = 1/8$. Better analysis is known for higher dimensional encoding [BDN17, MZ23], and for the 3-dimensional version of it a near optimal soundness result is known [MZ23].

Low-degree tests in this paper. In the context of the current paper, we wish to encode linear functions $f\colon \mathbb{F}_q^n \to \mathbb{F}_q$, and we do so by the subspaces encoding. Specifically, we set integer parameters $\ell_1 \geqslant \ell_2$, and encode the function f using the table T_1 of the restrictions of f to all ℓ_1-dimensional linear subspaces of $\mathbb{F}_q^n$, and the table T_2 of the restrictions of f to all ℓ_2-dimensional linear subspaces of $\mathbb{F}_q^n$. The test we consider is the natural inclusion test:

1. Sample a random ℓ_1-dimensional subspace $L_1 \subseteq \mathbb{F}_q^n$ and a random ℓ_2-dimensional subspace $L_2 \subseteq L_1$.
2. Read $T_1[L_1]$, $T_2[L_2]$ and accept if they agree on L_2.

As is often the case, the completeness of the test — namely the fact that valid tables T_1, T_2 pass the test with probability 1 — is clear. The question of most interest then is with regards to the soundness of the test. Namely, what is the smallest ε such that any two tables T_1 and T_2 that assign linear functions to subspaces and pass the test with probability ε, must necessarily "come from" a legitimate linear function f?

Traditional notion of soundness. As the alphabet vs soundness tradeoff is key to the discussion herein, we begin by remarking that the alphabet size of the above encoding is $q^{\ell_1} + q^{\ell_2} = \Theta(q^{\ell_1})$ (since there are q^{ℓ} distinct linear functions on a linear space of dimension ℓ over $\mathbb{F}_q$). Thus, ideally we would like to show that the soundness error of the above test is $q^{-(1-o(1))\ell_1}$. Alas, this is false. Indeed, it turns out that one may construct assignments that pass the test with probability at least $\Omega(\max(q^{-\ell_2}, q^{\ell_2-\ell_1}))$ that do not have significant correlation with any linear function f:

1. Taking T_1, T_2 randomly by assigning to each subspace a random linear function, one can easily see that the test passes with probability $\Theta(q^{-\ell_2})$.
2. Taking linear subspaces $W_1, \ldots, W_{100q^{\ell_1}} \subseteq \mathbb{F}_q^n$ of codimension 1 randomly, and a random linear function $f_i \colon W_i \to \mathbb{F}_q$ for each i, one may choose T_1 and T_2 as follows. For each L_1, pick a random i such that $L_1 \subseteq W_i$ (if such i exists) and assign $T_1[L_1] = f_i|_{L_1}$. For each L_2, pick a random i such that $L_2 \subseteq W_i$ (if such i exists) and assign $T_2[L_2] = f_i|_{L_2}$. Taking $L_2 \subseteq L_1$ randomly, one sees that with constant probability L_2 has $\Theta(q^{\ell_1-\ell_2})$ many possible i's, L_1 has $\Theta(1)$ many possible i's and furthermore there is at least one i that is valid for both of them. With probability $\Omega(q^{\ell_2-\ell_1})$ this common i is chosen for both L_1 and L_2, and in this case, the test on (L_1, L_2) passes. It follows that, in expectation, T_1, T_2 pass the test with probability $\Omega(q^{\ell_2-\ell_1})$.

In light of the above, it makes sense that the best possible alphabet vs soundness tradeoff we may achieve with the subspace encoding is by taking $\ell_2 = \ell_1/2$. Such a setting of the parameters would give alphabet size $R = q^{\ell_1}$ and (possibly) soundness error $\Theta(1/\sqrt{R})$. This tradeoff is not good enough for our purposes, and so we must venture beyond the traditional notion of soundness.

Non-traditional notion of soundness. The above test was first considered in the context of the 2-to-1 Games Theorem, wherein one takes $q = 2$ and $\ell_2 = \ell_1 - 1$. In this setting, the test is not sound in the traditional sense; instead, the test is shown to satisfy a non-standard notion of soundness, which nevertheless is sufficient for the purposes of constructing a PCP. More specifically, in [KMS23] it is proved that for all $\varepsilon > 0$ there is $r \in \mathbb{N}$ such that for sufficiently large ℓ and for tables T_1, T_2 as above, there are subspaces $Q \subseteq W \subseteq \mathbb{F}_q^n$ with $\mathsf{dim}(Q) + \mathsf{codim}(W) \leqslant r$ and a linear function $f \colon W \to \mathbb{F}_q$ such that

$$\Pr_{Q \subseteq L_1 \subseteq W}[T_1[L_1] \equiv f|_{L_1}] \geqslant \varepsilon'(\varepsilon) > 0.$$

We refer to the set

$$\{L \subseteq \mathbb{F}_q^n \mid \mathsf{dim}(L) = \ell_1, Q \subseteq L \subseteq W\}$$

as the zoom-in of Q and zoom-out of W. While this result is sufficient for the purposes of 2-to-1 Games, the dependency between ℓ and ε (and thus, between the soundness and the alphabet size) is not good enough for us.

Our low-degree test. It turns out that the proper setting of parameters for us is $\ell_2 = (1-\delta)\ell_1$ where $\delta > 0$ is a small constant. With these parameters, we are able to show that for $\varepsilon \geqslant q^{-(1-\delta')\ell_1}$ (where $\delta' = \delta'(\delta) > 0$ is a vanishing function of δ), if T_1, T_2 pass the test with probability at least ε, then there are subspaces $Q \subseteq W$ with $\mathsf{dim}(Q) + \mathsf{codim}(W) \leqslant r = r(\delta) \in \mathbb{N}$, and a linear function $f\colon W \to \mathbb{F}_q$ such that

$$\Pr_{Q\subseteq L_1\subseteq W}[T_1[L_1] \equiv f|_{L_1}] \geqslant \varepsilon'(\varepsilon) = \Omega(\varepsilon).$$

Working in the very small soundness regime of $\varepsilon \geqslant q^{-(1-\delta')\ell_1}$ entails many challenges, however. First, dealing with such small soundness requires us to use a strengthening of the global hypercontractivity result of [KMS23] in the form of an optimal level d inequality due to Evra, Kindler and Lifshitz [EKL24]. Second, in the context of [KMS23], ε' could be any function of ε (and indeed it ends up being a polynomial function of ε). In the context of the current paper, it is crucial that $\varepsilon' = \varepsilon^{1+o(1)}$, as opposed to, say, $\varepsilon' = \varepsilon^{1.1}$. The reason is that, as we are dealing with very small ε, the result would be trivial for $\varepsilon' = \varepsilon^{1.1}$ and not useful towards the analysis of the PCP (as then ε' would be below the threshold $q^{-\ell_1}$ which represents the agreement a random linear function f has with T_1).

1.2.2 Getting List Decoding Bounds

As is usually the case in PCP reductions, we require a list decoding version for our low-degree test. Indeed, using a standard argument we are able to show that in the setting that $\ell_2 = (1-\delta)\ell_1$ and $\varepsilon \geqslant q^{-(1-\delta')\ell_1}$, there is $r = r(\delta, \delta') \in \mathbb{N}$ such that for at least $q^{-\Theta(\ell_1)}$ fraction of subspaces $Q \subseteq \mathbb{F}_q^n$ of dimension r, there exists a subspace W with codimension at most r and $Q \subseteq W \subseteq \mathbb{F}_q^n$, as well as a linear function $f\colon W \to \mathbb{F}_q$, such that

$$\Pr_{Q\subseteq L_1\subseteq W}[T_1[L_1] \equiv f|_{L_1}] \geqslant \varepsilon'(\varepsilon) = \Omega(\varepsilon). \tag{1}$$

This list decoding version theorem alone is not enough. In our PCP construction, we compose the inner PCP with an outer PCP (that we describe below), and analyzing the composition requires decoding global linear functions (from a list decoding version theorem as above) in a coordinated manner between two non communicating parties. Oftentimes, the number of possible global functions that may be decoded is constant, in which case randomly sampling one among them works. This is not the case for us, though: if (Q, W) and (Q', W') are distinct zoom-in and zoom-out pairs for which there are linear functions $f_{Q,W}$ and $f_{Q',W'}$ satisfying (1), then the functions $f_{Q,W}$ and $f_{Q',W'}$ could be completely different. Thus, to achieve a coordinated decoding procedure, we must:

1. Facilitate a way for the two parties to agree on a zoom-in and zoom-out pair (Q, W) with noticeable probability.
2. Show that for each (Q, W) there are at most $\mathsf{poly}(1/\varepsilon)$ linear functions $f_{Q,W}$ for which

$$\Pr_{Q\subseteq L_1\subseteq W}[T_1[L_1] \equiv f_{Q,W}|_{L_1}] \geqslant \varepsilon'.$$

The second item is precisely the reason we need ε' to be $\varepsilon^{1+o(1)}$; any worse dependency, such as $\varepsilon' = \varepsilon^{1.1}$ would lead to the second item being false. We also remark that the number of functions being $\mathsf{poly}(1/\varepsilon)$

is important to us as well. There is some slack in this bound, but a weak quantitative bound such as $\exp(\exp(1/\varepsilon))$ would have been insufficient for some of our applications. Luckily, such bounds can be deduced from [GRS00] for the case of linear functions.[3]

We now discuss the first item. It turns out that agreeing on the zoom-in Q can be delegated to the outer PCP, in the sense that our outer PCP game will provide the two parties with a coordinated zoom-in Q (while still being sound). Morally speaking, this works because in our list decoding theorem, the fraction of zoom-ins Q that work is significant. Coordinating zoom-outs is more difficult, and this is where much of the novelty in our analysis lies.

1.2.3 Coordinating Zoom-outs

For the sake of simplicity and to focus on the main ideas, we ignore zoom-ins for now and assume that the list decoding statement holds with no Q. Thus, the list decoding theorem asserts that there exists a zoom-out W of constant codimension on which there is a global linear function agreeing with the table T_1. However, there could be many such zoom-outs W, say $W_1, \ldots, W_m$ and say all of them were of codimension $r = O_{\delta,\delta'}(1)$. If the number m were sufficiently large — say at least $q^{-\mathsf{poly}(\ell_1)}$ fraction of all codimension r subspaces — then we would have been able to coordinate them in the same way as we coordinate zoom-ins. If the number m were sufficiently small — say $m = q^{\mathsf{poly}(\ell_1)}$, then randomly guessing a zoom-out would work well enough. The main issue is that the number m is intermediate, say $m = q^{\sqrt{n}}$.

This issue had already appeared in [KMS17, DKK$^+$18]. Therein, this issue is resolved by showing that if there are at least $m \geqslant q^{100\ell_1^2}$ zoom-outs $W_1, \ldots, W_m$ of codimension r, and linear functions $f_1, \ldots, f_m$ on $W_1, \ldots, W_m$ respectively such that

$$\Pr_{L \subseteq W_i}[T_1[L] \equiv f_i|_L] \geqslant \varepsilon'$$

for all i, then there exists a zoom-out W of codimension *strictly less than* r and a linear function $f\colon W \to \mathbb{F}_q$ such that

$$\Pr_{L \subseteq W}[T_1[L] \equiv f|_L] \geqslant \Omega(\varepsilon'^{12}).$$

Thus, if there are too many zoom-outs of a certain codimension, then there is necessarily a zoom-out of smaller codimension that also works. In that case, the parties could go up to that codimension.

This result is not good enough for us, due to the polynomial gap between the agreement of T_1 and f_i, and the agreement of T_1 and f. Indeed, in our range of parameters, ε'^{12} will be below the trivial threshold $q^{-\ell_1}$ which is the agreement a random linear function f has with T_1, and therefore the promise on the function f above is meaningless.

We resolve this issue by showing a stronger, essentially optimal version of the above assertion still holds. Formally, we prove:

Lemma 1.7. *For all $\delta > 0$, $r \in \mathbb{N}$ there is $C > 1$ such that the following holds for $\varepsilon' \geqslant q^{-(1-\delta)\ell_1}$ and large enough ℓ_1. Suppose that T_1 is a table that assigns to each subspace L of dimension ℓ_1 a linear function, and suppose that there are at least $m \geqslant q^{C\ell_1}$ subspaces $W_1, \ldots, W_m$ of codimension r and linear functions $f_i\colon W_i \to \mathbb{F}_q$ such that for all $i = 1, \ldots, m$,*

$$\Pr_{L \subseteq W_i}[T_1[L] \equiv f_i|_L] \geqslant \varepsilon'.$$

[3] In the case of higher degree functions (even quadratic functions) some bounds are known [Gop13, BL15] but they would not have been good enough for us.

Then, there exists a zoom-out W of codimension strictly smaller than r and a linear function $f\colon W \to \mathbb{F}_q$ such that

$$\Pr_{L\subseteq W}[T_1[L] \equiv f|_L] \geqslant \Omega(\varepsilon').$$

We defer a detailed discussion about Lemma 1.7 and its proof to Section 8, but remark that our proof of Lemma 1.7 is very different from the arguments in [DKK+18] and is significantly more involved.

1.2.4 The Outer PCP

Our outer PCP game is the same as that of [KMS17, DKK+18], which is a smooth parallel repetition of the equation versus variables game of Hastad [Hås01] (or of [KP06] for the application to Quadratic Programming). As in there, we equip this game with the "advice" feature to facilitate zoom-in coordination (as discussed above). For the sake of completeness we elaborate on the construction of the outer PCP below.

We start with an instance of $\mathsf{Gap3Lin}$ that has a large gap between the soundness and completeness. Namely, we start with an instance (X, Eq) of linear equations over $\mathbb{F}_q$ in which each equation has the form $ax_{i_1} + bx_{i_2} + cx_{i_3} = d$. It is known [Hås01] that for all $\eta > 0$, it is NP-hard to distinguish between the following two cases:

1. YES case: $\mathsf{val}(X, \mathsf{Eq}) \geqslant 1 - \eta$.
2. NO case: $\mathsf{val}(X, \mathsf{Eq}) \leqslant \frac{1.1}{q}$.

The standard *Variable versus Equation game* proceeds as follows:

1. The verifier picks an equation e uniformly at random from Eq and lets $U = \{x_{i_1}, x_{i_2}, x_{i_3}\}$ be the variables in it.
2. The verifier samples $V \subseteq U$ of size 1 uniformly.

The verifier then sends U to the first prover, and V to the second prover. The verifier expects to receive from each prover an assignment to the variables they received, and accepts if their assignments are consistent and satisfy the equation e. It is easy to see that if $\mathsf{val}(X, \mathsf{Eq}) \geqslant 1-\eta$, then the provers have a strategy that makes the verifier accept with probability at least $1-\eta$, and if $\mathsf{val}(X, \mathsf{Eq}) \leqslant \frac{1.1}{q}$, then for any provers' strategy, the verifier accepts with probability at most $1 - \Omega(1)$.

Now, we modify the standard Variable versus Equation game into the *smooth Variable versus Equation game with r-advice*, which proceeds as follows. The verifier has a smoothness parameter $\beta > 0$ and picks a random equation e, say $ax_{i_1} + bx_{i_2} + cx_{i_3} = d$, from (X, Eq). Then:

1. With probability $1 - \beta$ the verifier takes $U = V = \{x_{i_1}, x_{i_2}, x_{i_3}\}$ and vectors $u_1 = v_1, \ldots, u_r = v_r \in \mathbb{F}_q^U$ sampled uniformly and independently.
2. With probability β, the verifier sets $U = \{x_{i_1}, x_{i_2}, x_{i_3}\}$, chooses a set consisting of a single variable $V \subseteq U$ uniformly at random. The verifier picks $v_1, \ldots, v_r \in \mathbb{F}_q^V$ uniformly and independently and appends to each v_i the value 0 in the coordinates of $U \setminus V$ to get $u_1, \ldots, u_r$.

After that, the verifier sends U and $u_1, \ldots, u_r$ to the first prover and V and $v_1, \ldots, v_r$ to the second prover. The verifier expects to get from them $\mathbb{F}_q$ assignments to the variables in U and in V, and accepts if and only if these assignments are consistent, and furthermore the assignment to U satisfies the equation e.

Denoting the smooth Variable versus Equation game above by Ψ, it is easy to see that if $\mathsf{val}(X, \mathsf{Eq}) \geqslant 1 - \eta$, then $\mathsf{val}(\Psi) \geqslant 1 - \eta$, and if $\mathsf{val}(X, \mathsf{Eq}) \leqslant 1.1/q$, then $\mathsf{val}(\Psi) \leqslant 1 - \Omega(q^{-r}\beta)$. Indeed, to see the last assertion, note that the probability that V is a proper subset of U and that $v_1, \ldots, v_r$ are all 0, is $q^{-r}\beta$. In that case, the game being played is the standard Variable versus Equation game, which has value at most $1 - \Omega(1)$ as discussed earlier. The gap between $1 - \eta$ and $1 - \Omega(q^{-r}\beta)$ is too weak for us, and thus we apply parallel repetition.

In the parallel repetition of the smooth Variable versus Equation game with advice, denoted by $\Psi^{\otimes k}$, the verifier picks k equations uniformly and independently $e_1, \ldots, e_k$, and picks U_i, $u_1^i, \ldots, u_r^i$ and V_i, $v_1^i, \ldots, v_r^i$ for each $i = 1, \ldots, k$ from e_i independently. Thus, the questions of the provers may be seen as $U = U_1 \cup \ldots \cup U_k$ and $V = V_1 \cup \ldots \cup V_k$ and their advice is $\vec{u}_j = (u_j^1, \ldots, u_j^k) \in \mathbb{F}_q^U$ for $j = 1, \ldots, r$ and $\vec{v}_j = (v_j^1, \ldots, v_j^k) \in \mathbb{F}_q^V$ for $j = 1, \ldots, r$ respectively. The verifier expects to get from the first prover a vector in $\mathbb{F}_q^U$ which specifies an $\mathbb{F}_q$ assignment to U, and from the second prover a vector in $\mathbb{F}_q^V$ which specifies an $\mathbb{F}_q$ assignment to V. The verifier accepts if and only if these assignments are consistent and the assignment of the first prover satisfies all of $e_1, \ldots, e_k$. It is clear that if $\mathsf{val}(X, \mathsf{Eq}) \geqslant 1 - \eta$, then $\mathsf{val}(\Psi^{\otimes m}) \geqslant 1 - k\eta$. Using the parallel repetition theorem of Rao [Rao11] we argue that if $\mathsf{val}(X, \mathsf{Eq}) \leqslant \frac{1.1}{q}$, then $\mathsf{val}(\Psi^{\otimes k}) \leqslant 2^{-\Omega(\beta q^{-r} k)}$. The game $\Psi^{\otimes k}$ is our outer PCP game.

Remark 1.8. *In the case of the Quadratic Programming application we require a hardness result in which the completeness is very close to* 1 *(see Theorem 2.6). The differences between the reduction in that case and the reduction presented above are mostly minor, and amount to picking the parameters a bit differently. There is one significant difference in the analysis, namely we require a much sharper form of the "covering property" used in [KMS17, DKK⁺18]; see Section 1.2.6 for a discussion.*

1.2.5 Composing the Outer PCP and the Inner PCP Game

To compose the outer and inner PCPs, we take the outer PCP game, only keep the questions U to the first prover and consider an induced 2-Prover-1-Round game on it. The alphabet is $\mathbb{F}_q^{3k}$, and we think of a response to a question U as an $\mathbb{F}_q$ assignment to the variables of U. There is a constraint between U and U' if there is a question V to the second prover such that $V \subseteq U \cap U'$. Denoting the assignments to U and U' by s_U and $s_{U'}$, the constraint between U and U' is that s_U satisfies all of the equations that form U, $s_{U'}$ satisfies all of the equations that form U', and s_U, $s_{U'}$ agree on $U \cap U'$.

The composition amounts to replacing each question U with a copy of our inner PCP. Namely, we identify between the question U and the space $\mathbb{F}_q^U$, and then replace U by a copy of the ℓ_2, ℓ_1 subspaces graph of $\mathbb{F}_q^U$. The answer s_U is naturally identified with the linear function $f_U(x) = \langle s_U, x \rangle$, which is then encoded by the subspaces encoding via tables of assignments $T_{1,U}$ and $T_{2,U}$.

The constraints of the composed PCP must check that: (1) side conditions: the encoded vector s_U satisfies the equations of U, and (2) consistency: s_U and $s_{U'}$ agree on $U \cap U'$.

The first set of constraints is addressed by the folding technique, which we omit from this discussion. The second set of constraints is addressed by the ℓ_1 vs ℓ_2 subspace test, except that we have to modify it so that it works across blocks U and U'. This completes the description of the composition step of the other PCP and the inner PCP, and thereby the description of our reduction.

Analyzing the Composed PCP. Traditionally, the analysis of the composed PCP relates its soundness with the soundness of the outer PCP and the soundness of the inner PCP. Such analysis, however, fails in our case. Instead, and similarly to prior works [KMS17, DKK⁺21], our analysis relates the soundness of the composed PCP with the non-standard notion of soundness of the inner PCP (as discussed above) and the soundness of a modification of the (standard) outer PCP game, as presented above. The goal of this modification is to allow the provers to correlate their choice for the zoom-in set Q, and it is achieved by the additional vectors $\{\vec{u}_j\}_{j=1,\dots,r}$ and $\{\vec{v}_j\}_{j=1,\dots,r}$ that the verifier sends to the provers in the outer PCP game.

1.2.6 The Covering Property

We end this introductory section by discussing the covering property, which is an important feature of our outer PCP construction enabling the composition step to go through. The covering property first appeared in [KS13] and later more extensively in the context of the 2-to-1 Games [KMS17, DKK$^+$18]. To discuss the covering property, let $k \in \mathbb{N}$ be thought of as large, let $\beta \in (0,1)$ be thought of as $k^{-0.99}$ and consider pairwise disjoint sets $U_1, \dots, U_k$ where each U_i has size 3 (in our context, U_i will be the set of variables in the ith equation the verifier chose). Let $U = U_1 \cup \dots \cup U_k$, and consider the following two distributions over tuples in $\mathbb{F}_q^U$:

1. Sample $x_1, \dots, x_\ell \in \mathbb{F}_q^U$ uniformly.
2. For each i independently, take $V_i = U_i$ with probability $1-\beta$ and otherwise take $V_i \subseteq U_i$ randomly of size 1, then set $V = V_1 \cup \dots \cup V_k$. Sample $x_1, \dots, x_\ell \in \mathbb{F}_q^V$ uniformly and lift them to points in $\mathbb{F}_q^U$ by appending 0's in $U \setminus V$. Output the lifted points.

In [KMS17] it is shown that the two distributions above are $q^{3\ell}\beta\sqrt{k}$ close in statistical distance, which is good enough for Theorem 1.3. This is insufficient for Theorem 1.4, though.[4] Carrying out a different analysis, one can show that the two distributions are close with better parameters and in a stronger sense: there exists a set E of ℓ tuples which has negligible measure in both distributions, such that each tuple not in E is assigned the same probability under the two distributions up to factor $(1+o(1))$. We are able to prove this statement provided that k is only slightly larger than $q^{2\ell}$, which is still insufficient.

The issue with the above two distributions is that they are actually far from each other if, say, $k = q^{1.9\ell}$. To see that, one can notice that the expected number of i's such that each one of $x_1, \dots, x_\ell$ has the form $(a,0,0) \in \mathbb{F}_q^3$ on coordinates corresponding to U_i is very different. In the first distribution, this expectation is $\Theta(q^{-2\ell}k)$ which is less than 1, whereas in the second distribution it is at least $\beta k/3 \geqslant k^{0.01}$.

To resolve this issue and to go all the way through in the Quadratic Programming application, we have to modify the distributions in the covering property so that (a) they will be close even if $k = q^{1.01\ell}$, and (b) we can still use these distributions in the composition step in our analysis of the PCP construction. Indeed, this is the route we take, and the two distributions we use are defined as follows:

1. Sample $x_1, \dots, x_\ell \in \mathbb{F}_q^U$ uniformly.
2. For each i independently, take $V_i = U_i$ with probability $1-\beta$ and otherwise take $V_i \subseteq U_i$ randomly of size 1, then set $V = V_1 \cup \dots \cup V_k$. Sample $x_1, \dots, x_\ell \in \mathbb{F}_q^V$ uniformly, and let $w_i \in \mathbb{F}_q^U$ be the vector of the coefficients of the equation e_i forming U_i. Lift the points $x_1, \dots, x_\ell$ to $x'_1, \dots, x'_\ell \in \mathbb{F}_q^U$ by appending 0's in $U \setminus V$ and take $y_j = x_j + \sum\limits_{i=1}^{k} \alpha_{i,j} w_i$ where $\alpha_{i,j}$ are independent random elements from $\mathbb{F}_q$. Output $y_1, \dots, y_\ell$.

We show that for a suitable choice of k and β, these distributions are close even in the case that $k = q^{1.01\ell}$.[5] Indeed, as a sanity check one could count the expected number of appearances of blocks of the

[4]The reason is that, letting N be the size of the instance we produce, it holds that k is roughly $\log N$ and q^ℓ is the alphabet size. To have small statistical distance, we must have $k \leqslant q^{6\ell}$, hence the soundness error could not go lower than $(\log N)^{-1/6}$.

[5]More specifically, one takes a small $c > 0$ and chooses $\beta = k^{2c/3-1}$, $k = q^{(1+c)\ell}$.

form $(0, a, 0) \in \mathbb{F}^3$ and see they are very close ($q^{-2\ell}k$ versus $(1-\beta)q^{-2\ell}k + \beta k q^{-\ell}$). In this setting of parameters, k is roughly equal to the alphabet size — which can be made to be equal $(\log N)^{1-o(1)}$ under quasi-polynomial time reductions.

Remark 1.9. *A tight covering property is crucial for obtaining the tight hardness of approximation factor in Theorem 1.4. In the reduction of [ABH+05] from 2-Prover-1-Round games to Quadratic Programming, the size of the resulting instance is exponential in the alphabet size and the soundness error remains roughly the same. In our case the alphabet size is roughly k, hence the instance size is dominated by $N = 2^{\Theta(k^{1+o(1)})}$. If our analysis required $k = q^{C\ell}$, then even showing an optimal soundness of $q^{-(1-o(1))\ell}$ for the 2-Prover-1-Round game would only yield a factor of $(\log N)^{1/C-o(1)}$ hardness for Quadratic Programming.*

1.3 Open Problems

We end this introductory section by mentioning a few open problems for future research.

1. **The Low-degree CSP Conjecture.** It would be interesting to see if our techniques have any bearing on the conjecture of [CDGT22]. This conjecture is a sort of a mixture between the d-to-1 games and sliding scale conjectures, and it focuses on the relation between the instance size and the soundness error, while allowing the alphabet to be quite large. The conjecture is motivated by improved hardness results for densest k-subgraph style problems, and below we give a rough description of it.

   The 2-to-1 (or 2-to-2) Theorem asserts that the soundness of a PCP can be arbitrarily small constant in a d-to-1 (or d-to-d) game even if d is fixed; it does not apply well in the sub-constant soundness regime, though. The parallel repetition theorem asserts that for every $d\colon \mathbb{N} \to \mathbb{N}$, one can get a $d(n)$-to-$d(n)$ PCP with soundness which is polynomial in $1/d(n)$. The Low-degree CSP conjecture asks for a PCP that has $d(n)$ growing with the instance size, whose soundness decays more rapidly than polynomial in d. Specifically, it asks for $d(n) = 2^{\log^{\delta} n}$ and soundness $2^{-\log^{1/2+\delta} n}$.

2. **Perfect completeness.** Our reduction starts with the $\mathsf{Gap3Lin}$ problem and thus inherently has imperfect completeness. It would be very interesting if one could prove a version of Theorem 1.3 with perfect completeness. As far as we know no result better than Theorem 1.2 is known to date. While we are not aware of any direct applications of it, we believe such a result may lead to techniques which are able to give strong PCPs with perfect completeness (which are quite rare).

3. **Other applications.** As shown above, it is sometimes the case that results proved assuming the Unique-Games Conjecture can be translated into NP-hardness results using Theorem 1.3. We suspect there should be other such problems.

2 Preliminaries

2.1 The Grassmann Domain

In this section we present the Grassmann domain and associated Fourier analytic tools that are required for our analysis of the inner PCP. Throughout this section we fix parameters n, ℓ with $1 \ll \ell \ll n$, and a prime power q (which we will eventually pick to be 2).

2.1.1 Basic Definitions

The Grassmann domain $\mathsf{Grass}_q(n,\ell)$ consists of the set of all ℓ-dimensional subspaces $L \subseteq \mathbb{F}_q^n$. At times we will have a vector space W over $\mathbb{F}_q$, and we may write $\mathsf{Grass}_q(W,\ell)$ to mean $\mathsf{Grass}_q(n,\ell)$, where n is the dimension of W (under some identification of W with $\mathbb{F}_q^n$). When the field size $\mathbb{F}_q$ is clear from context we drop the subscript and simply write $\mathsf{Grass}(n,\ell)$ or $\mathsf{Grass}(W,\ell)$.

The number of ℓ-dimensional subspaces of $\mathbb{F}_q^n$ is counted by the Gaussian binomial coefficient, $\begin{bmatrix} n \\ \ell \end{bmatrix}_q$. The following standard fact gives a formula for the Gaussian binomial coefficients, and we omit the proof.

Fact 2.1. *Suppose $1 \leqslant \ell \leqslant \frac{n}{2}$, then the number of vertices in $\mathsf{Grass}_q(n,\ell)$ is given by*

$$\begin{bmatrix} n \\ \ell \end{bmatrix}_q = \prod_{i=0}^{\ell-1} \frac{q^n - q^i}{q^\ell - q^i}.$$

Bipartite Inclusion Graphs We will often consider bipartite inclusion graphs between $\mathsf{Grass}_q(n,\ell)$ and $\mathsf{Grass}_q(n,\ell')$ for $\ell' < \ell$. This graph has an edge between every $L \in \mathsf{Grass}_q(n,\ell)$ and $L' \in \mathsf{Grass}_q(n,\ell')$ such that $L \supseteq L'$. We refer to the family of such graphs as *Grassmann graphs*. The edges of these graphs are the motivation for the test used in our inner PCP. While we never refer to this graph explicitly, it will be helpful to have it in mind.

Zoom-ins and Zoom-outs. For subspaces $Q \subseteq W \subseteq \mathbb{F}_q^n$ and a dimension $\ell \in \mathbb{N}$, let

$$\mathsf{Zoom}_\ell[Q,W] = \{L \in \mathsf{Grass}_q(n,\ell) \mid Q \subseteq L \subseteq W\}.$$

We refer to Q as a zoom-in and W as a zoom-out. When $W = \mathbb{F}_q^n$, $\mathsf{Zoom}_\ell[Q,W]$ is the zoom-in on Q, and when $Q = \{0\}$, $\mathsf{Zoom}_\ell[Q,W]$ is the zoom-out on W.

2.1.2 Pseudo-randomness over the Grassmann graph

One notion that will be important to us is (r,ε)-pseudo-randomness, which measures how large the fractional size of a set of subspaces $\mathcal{L}$ can be in a zoom-in/zoom-out restrictions of "size r".[6] The measure of a set of subspaces $\mathcal{L} \subseteq \mathsf{Grass}_q(n,\ell)$ is the fraction of ℓ-dimensional subspaces it contains, denoted as

$$\mu(\mathcal{L}) = \Pr_{L \in \mathsf{Grass}_q(n,\ell)}[L \in \mathcal{L}].$$

We remark that the dimension ℓ in the probability above will always be clear from context based on $\mathcal{L}$.

For subspaces $Q \subseteq W \subseteq \mathbb{F}_q^n$, the measure of $\mathcal{L}$ under zoom-in/zoom-out pair (Q,W) is defined as

$$\mu_{Q,W}(\mathcal{L}) = \Pr_{L \in \mathsf{Grass}_q(n,\ell)}[L \in \mathcal{L} \mid Q \subseteq L \subseteq W]. \tag{2}$$

We also define a similar notation for measure inside of only a zoom-in or only a zoom-out.

For a zoom-in Q, we write

$$\mu_{Q,\circ}(\mathcal{L}) = \Pr_{L \in \mathsf{Grass}_q(n,\ell)}[L \in \mathcal{L} \mid Q \subseteq L].$$

[6]The results we mention can be stated in greater generality, but we avoid it for the sake of simplicity.

and for a zoom-out W, we write

$$\mu_{\circ,W}(\mathcal{L}) = \Pr_{L\in \mathsf{Grass}_q(n,\ell)}[L \in \mathcal{L} \mid L \subseteq W].$$

Here we abuse notation and note that when we use the notation μ_Q or μ_W, it will always be clear from context which of the above two notations we are using and whether the subspace in the subscript is a zoom-in or zoom-out.

Definition 2.2. *We say that a set of subspaces $\mathcal{L} \subseteq \mathsf{Grass}_q(n,\ell)$ is (r,ε)-pseudo-random if for all $Q \subseteq W \subseteq \mathbb{F}_q^n$ satisfying $\dim(Q) + \operatorname{codim}(W) = r$, we have $\mu_{Q,W}(\mathcal{L}) \leqslant \varepsilon$.*

It turns out that a pseudo-random $\mathcal{L}$ as above has the same edge expansion (and other related edge-counts) as a random set of the same density, and we require such an assertion in our inner PCP. Below is a precise formulation.

Lemma 2.3. *Let $\mathcal{L} \subseteq \mathsf{Grass}_q(n,2\ell)$ and $\mathcal{R} \subseteq \mathsf{Grass}_q(n,2(1-\delta)\ell)$ be sets of subspaces with fractional sizes $\Pr_{L\in\mathsf{Grass}_q(n,2\ell)}[L \in \mathcal{L}] = \alpha$ and $\Pr_{R\in\mathsf{Grass}_q(n,2(1-\delta)\ell)}[R \in \mathcal{R}] = \beta$, and suppose that $\mathcal{L}$ is (r,ε) pseudo-random. Then for all $t \geqslant 4$ that are powers of 2,*

$$\Pr_{L\supseteq R}[L \in \mathcal{L}, R \in \mathcal{R}] \leqslant q^{O_{t,r}(1)}\beta^{(t-1)/t}\varepsilon^{(t-2)/t} + q^{-r\delta\ell}\sqrt{\alpha\beta}.$$

In the probability above, $L \in \mathsf{Grass}_q(n,2\ell)$ is chosen uniformly at random and $R \in \mathsf{Grass}_q(n,2(1-\delta)\ell)$ is chosen uniformly at random conditioned on being contained in L.

Proof. Deferred to Appendix A. □

We also need the following lemma, asserting that if a not-too-small set $\mathcal{L}$ is highly pseudo-random, then its density remains nearly the same on all zoom-ins of dimension 1. For a point $z \in \mathbb{F}_q^n$ we denote by $\mu_z(\cdot)$ the measure $\mu_Q(\cdot)$ as above where $Q = \operatorname{span}(z)$.

Lemma 2.4. *For all $\xi > 0$, the following holds for sufficiently large ℓ. Suppose that $\ell' \geqslant \frac{\xi}{3}\ell$, $\delta_2 = \frac{\xi}{100}$, and W is a subspace of dimension $\dim(W) \geqslant \ell'^2$. Let $\mathcal{L} \subseteq \mathsf{Grass}_q(W,\ell')$ have measure $\mu(\mathcal{L}) = \eta \geqslant q^{-2\ell}$ and set $Z = \{z \in W \mid |\mu_z(\mathcal{L}) - \eta| \leqslant \frac{\eta}{10}\}$. If $\mathcal{L}$ is $(1, q^{\delta_2\ell/100}\eta)$-pseudo-random, then*

$$|Z| \geqslant \left(1 - q^{-\frac{\ell'}{2}}\right)|W|.$$

Proof. The proof is deferred to Appendix A.4. □

2.2 Hardness of 3LIN

In this section we cite several hardness of approximation results for the problem of solving linear equations over finite fields, which are the starting point of our reduction. We begin by defining the $\mathsf{3Lin}$ and the $\mathsf{Gap3Lin}$ problem.

Definition 2.5. *For a prime power q, an instance of $\mathsf{3Lin}$ (X, Eq) consists of a set of variables X and a set of linear equations Eq over $\mathbb{F}_q$. Each equation in Eq depends on exactly three variables in X, each variable appears in at most 10 equations, and any two distinct equations in Eq share at most a single variable.*

The goal in the 3Lin problem is to find an assignment $A\colon X \to \mathbb{F}_q$ satisfying as many of the equations in E as possible. The maximum fraction of equations that can be satisfied is called the value of the instance. We remark that usually in the literature, the condition that two equations in E share at most a single variable is not included in the definition of 3Lin, as well as the bound on the number of occurrences of each variable.

For $0 < s < c \leqslant 1$, the problem $\mathsf{Gap3Lin}[c, s]$ is the promise problem wherein the input is an instance (X, Eq) of 3Lin promised to either have value at least c or at most s, and the goal is to distinguish between these two cases. The problem $\mathsf{Gap3Lin}[c, s]$ with various settings of c and s will be the starting point for our reductions.

In the proof of Theorem 1.3, we will use a classical result of Håstad [Hås01], stating that for general 3Lin instances (i.e., without the additional conditions that two equations share at most a single variable and each variable occurs in at most 10 equations), the problem $\mathsf{Gap3Lin}[1-\varepsilon, 1/q+\varepsilon]$ is NP-hard for all constant $q \in \mathbb{N}$ and $\varepsilon > 0$. This result implies the following theorem by elementary reductions (similarly to Theorem 2.7; see also [Min22, Footnote 14]):

Theorem 2.6. *There exists $s < 1$ such that for every $\eta > 0$ and prime q, $\mathsf{Gap3Lin}\,[1-\eta, s]$ is NP-hard.*

In the proof of Theorem 1.4 we will need a hardness result for 3Lin with completeness close to 1, and we will use a hardness result of Khot and Ponnuswami [KP06]. Once again, their result does not immediately guarantee the fact that any two equations share at most a single variable and that each variable occurs in at most 10 equations, however it can be achieved by an elementary reduction.

Theorem 2.7. *For any q power of 2, there is a reduction from SAT with size n to a $\mathsf{Gap3Lin}[1-\eta, 1-\varepsilon]$ instance with size N over $\mathbb{F}_q$ with $\{0,1\}$-coefficients, where*

- *Both N and the running time of the reduction are bounded by $2^{O(\log^2 n)}$;*
- *$\eta \leqslant 2^{-\Omega(\sqrt{\log N})}$;*
- *$\varepsilon \geqslant \Omega\left(\frac{1}{\log^3 N}\right)$.*

Proof. Using [KP06, Theorem 4] we get the statement above without the additional property that any two distinct equations share at most a single variable. To get this property we reduce an instance (X, Eq) to an instance (X', Eq') as follows: for each equation $e \in \mathsf{Eq}$ given as $x+y+z=b$ for $x, y, z \in X$, we produce the new variables x_e, y_e, z_e and add the equations $x + y_e + z_e = b$, $x_e + y + z_e = b$ and $x_e + y_e + z = b$. We note that each variable in (X', Eq') appears in at most 10 equations, and that if $\mathsf{val}(X, \mathsf{Eq}) \geqslant 1-\eta$, then $\mathsf{val}(X', \mathsf{Eq}') \geqslant 1-\eta$, so the completeness is clear. For the soundness analysis suppose that A is an assignment to (X', Eq') satisfying at least $1-\varepsilon$ fraction of the equations. By an averaging argument for at least $1-3\varepsilon$ fraction of $e \in \mathsf{Eq}$, we have that A satisfies all of the equations $x+y_e+z_e = b$, $x_e+y+z_e = b$ and $x_e + y_e + z = b$. By adding them up and using the fact that $\mathbb{F}_q$ has characteristic 2 we get that A also satisfies the equation $x+y+z=b$, giving that $\mathsf{val}(X, \mathsf{Eq}) \geqslant 1-3\varepsilon$ and completing the soundness analysis. □

3 The Outer PCP

In this section, we describe our outer PCP game. In short, our outer PCP is a smooth parallel repetition of a "variable versus equation" game with advice. This outer PCP was first considered in [KS13] without the advice feature, and then in [KMS17] with the advice feature.

3.1 The Outer PCP construction

Let $\varepsilon_1 < \varepsilon_2$ be parameters. Our reduction starts with an instance (X, Eq) of the $\mathsf{Gap3Lin}[1-\varepsilon_1, 1-\varepsilon_2]$ problem, and we present it gradually. We begin by presenting the basic Variable versus Equation Game, then equip it with the additional features of smoothness and advice, and finally perform parallel repetition.

3.1.1 The Variable versus Equation Game

We first convert the instance (X, Eq) into an instance of 2-Prover-1-Round Games, and it will be convenient for us to describe it in the active view with a verifier and 2 provers.

In the Variable versus Equation game, the verifier picks an equation $e \in \mathsf{Eq}$ uniformly at random, and then chooses a random variable $x \in e$. The verifier sends the question e, i.e. the three variables appearing in e, to the first prover, and sends the variable x to the second prover. The provers are expected to answer with assignments to their received variables, and the verifier accepts if and only if the two assignments agree on x and the first prover's assignment satisfies the equation e. If the verifier accepts then we also say that the provers pass. This game has the following completeness and soundness, which are both easy to see (we omit the formal proof):

1. **Completeness:** If (X, Eq) has an assignment satisfying $(1-\varepsilon)$-fraction of the equations, then the provers have a strategy that passes with probability at least $1-\varepsilon$.

2. **Soundness:** If (X, Eq) has no assignment satisfying more than $(1-\varepsilon)$-fraction of the equations, then the provers can pass with probability at most $1-\frac{\varepsilon}{3}$.

3.1.2 The Smooth Variable versus Equation Game

We next describe a smooth version of the Variable versus Equation game. In this game, the verifier has a parameter $\beta \in (0, 1]$, and it proceeds as follows:

1. The verifier chooses an equation $e \in \mathsf{Eq}$ uniformly, and lets U be the set of variables in e.

2. With probability $1-\beta$, the verifier chooses $V = U$. With probability β, the verifier chooses $V \subseteq U$ randomly of size 1.

3. The verifier sends U to the first prover, and V to the second prover.

4. The provers respond with assignments to the variables they receive, and the verifier accepts if and only if their assignments agree on V and the assignment to U satisfies the equation e.

The smooth Variable versus Equation game has the following completeness and soundness property, which are again easily seen to hold (we omit the formal proof).

1. **Completeness:** If (X, Eq) has an assignment satisfying $(1-\varepsilon)$-fraction of the equations, then the provers have a strategy that passes with probability at least $1-\varepsilon$.

2. **Soundness:** If (X, Eq) has no assignment satisfying more than $(1-\varepsilon)$-fraction of the equations, then the provers can pass with probability at most $1-\frac{\beta\varepsilon}{3}$.

3.1.3 The Smooth Variable versus Equation Game with Advice

Next, we introduce the feature of advice into the smooth Variable versus Equation Game. This "advice" acts as shared randomness which may help the provers in their strategy; we show though that it does not considerably change the soundness. The game is denoted by $G_{\beta,r}$ for $\beta \in (0,1]$ and $r \in \mathbb{N}$, and proceeds as follows:

1. The verifier chooses an equation $e \in \mathsf{Eq}$ uniformly, and lets U be the set of variables in e.
2. With probability $1-\beta$, the verifier chooses $V = U$. With probability β, the verifier chooses $V \subseteq U$ randomly of size 1.
3. The verifier picks vectors $v_1, \ldots, v_r \in \mathbb{F}_q^V$ uniformly and independently. If $U = V$ the verifier takes $u_i = v_i$ for all i, and otherwise the verifier takes the vectors $u_1, \ldots, u_r \in \mathbb{F}_q^U$ where for all $i = 1, \ldots, r$, the vector u_i agrees with v_i on the coordinate of V, and is 0 in the coordinates of $U \setminus V$.
4. The verifier sends U and $u_1, \ldots, u_r$ to the first prover, and V and $v_1, \ldots, v_r$ to the second prover.
5. The provers respond with assignments to the variables they receive, and the verifier accepts if and only if their assignments agree on V and the assignment to U satisfies the equation e.

Below we state the completeness and soundness of this game:

1. **Completeness:** If (X, Eq) has an assignment satisfying $(1-\varepsilon)$-fraction of the equations, then the provers have a strategy that passes with probability at least $1-\varepsilon$. This is easy to see.
2. **Soundness:** If (X, Eq) has no assignment satisfying more than $(1-\varepsilon)$-fraction of the equations, then the provers can pass with probability at most $1 - \frac{q^{-r}\beta\varepsilon}{3}$. Indeed, suppose that the provers can win the game with probability at least $1-\eta$. Note that with probability at least βq^{-r} it holds that $U \neq V$ and all the vectors $u_1, \ldots, u_r$ and $v_1, \ldots, v_r$ are all 0, in which case the provers play the standard Variable versus Equation game. Thus, the provers' strategy wins in the latter game with probability at least $1 - \frac{\eta}{q^{-r}\beta}$, so we must have that $\frac{\eta}{q^{-r}\beta} \geqslant \frac{\varepsilon}{3}$.

3.1.4 Parallel Repetition of the Smooth Variable versus Equation Game with Advice

Finally, our outer PCP is the k-fold parallel repetition of $G_{\beta,r}$, which we denote by $G_{\beta,r}^{\otimes k}$. Below is a full description of it:

1. The verifier chooses equations $e_1, \ldots, e_k \in \mathsf{Eq}$ uniformly and independently, and lets U_i be the set of variables in e_i.
2. For each i independently, with probability $1-\beta$, the verifier chooses $V_i = U_i$. With probability β, the verifier chooses $V_i \subseteq U_i$ randomly of size 1.
3. For each i independently, the verifier picks vectors $v_1^i, \ldots, v_r^i \in \mathbb{F}_q^{V_i}$ uniformly and independently. If $U_i = V_i$ the verifier takes $u_j^i = v_j^i$ for $j = 1, \ldots, r$, and otherwise the verifier takes the vectors $u_1^i, \ldots, u_r^i \in \mathbb{F}_q^{U_i}$ where for all $j = 1, \ldots, r$, the vector u_j^i agrees with v_j^i on the coordinate of V_i, and is 0 in the coordinates of $U_i \setminus V_i$.

4. The verifier sets $U = \bigcup_{i=1}^{k} U_i$ and $u_j = (u_j^1, \ldots, u_j^k)$ for each $j = 1, \ldots r$, and $V = \cup_{i=1}^{k} V_i$ and $v_j = (v_j^1, \ldots, v_j^k)$ for each $j = 1, \ldots, r$. The verifier sends U and $u_1, \ldots, u_r$ to the first prover, and V and $v_1, \ldots, v_r$ to the second prover.

5. The provers respond with assignments to the variables they receive, and the verifier accepts if and only if their assignments agree on V and the assignment to U satisfies the equations $e_1, \ldots, e_k$.

Next, we state the completeness and the soundness of the game $G_{\beta,r}^{\otimes k}$, and we begin with its completeness.

Claim 3.1. *If* (X, Eq) *has an assignment satisfying at least* $1 - \varepsilon$ *of the equations, then the provers can win* $G_{\beta,r}^{\otimes k}$ *with probability at least* $1 - k\varepsilon$.

Proof. Let A be an assignment that satisfies at least $(1 - \varepsilon)$-fraction of the equations in E, and consider the strategy of the provers that assigns their variables according to A. Note that whenever each one of the equations $e_1, \ldots, e_k$ the verifier chose is satisfied by A, the verifier accepts. By the union bound, the probability this happens is at least $1 - k\varepsilon$. $\square$

Next, we establish the soundness of the game $G_{\beta,r}^{\otimes k}$.

Claim 3.2. *If there is no assignment to* (X, Eq) *satisfying at least* $(1 - \varepsilon)$*-fraction of the equations, then the provers can win* $G_{\beta,r}^{\otimes k}$ *with probability at most* $2^{-\Omega(\varepsilon^2 q^{-r} \beta k)}$.

Proof. We appeal to the parallel repetition theorem for projection games of Rao [Rao11], but we have to do so carefully. That theorem states that if Ψ is a 2-Prover-1-Round game with $\mathsf{val}(\Psi) \leqslant 1 - \eta$, then $\mathsf{val}(\Psi^{\otimes k}) \leqslant 2^{-\Omega(\eta^2 k)}$. We cannot apply the theorem directly on $G_{\beta,r}$ (as the square is too costly for us). Instead, we consider the game $\Psi = G_{\beta,r}^{\lceil q^r/\beta \rceil}$ and note that it has value bounded away from 1.

Write $\mathsf{val}(\Psi) = 1 - \eta$. For each coordinate i, the probability that $U_i = V_i$ or $U_i \neq V_i$ but $v_1^i, \ldots, v_r^i$ are all 0, is equal to $1 - \beta q^{-r}$. Thus, by independence the probability that this happens for $\lceil q^r/\beta \rceil$ coordinates is at most $(1 - \beta q^{-r})^{\lceil q^r/\beta \rceil} \leqslant e^{-1+o(1)} \leqslant 0.9$. In particular, with probability at least 0.1 there exists at least one coordinate i in which $U_i \neq V_i$ and all of the advice vectors $v_1^i, \ldots, v_r^i$ and $u_1^i, \ldots, u_r^i$ are all 0. Thus, there exists a coordinate i and a fixing for the questions of the provers outside i so that the answers of the players to the ith coordinate win the standard Variable versus Equation game with probability at least $1 - 10\eta$. It follows that $1 - 10\eta \leqslant 1 - \frac{\varepsilon}{3}$, and so $\eta \geqslant \frac{\varepsilon}{30}$.

We conclude from Rao's parallel repetition theorem that

$$\mathsf{val}(G_{\beta,r}^{\otimes k}) = \mathsf{val}(\Psi^{\otimes \frac{k}{\lceil q^r/\beta \rceil}}) \leqslant 2^{-\Omega(\varepsilon^2 q^{-r} \beta k)}. \qquad \square$$

Viewing the advice as subspaces. Due to the fact that each variable appears in at most $O(1)$ equations, it can easily be seen that with probability $1 - O(k^2/n)$, all variables in $e_1, \ldots, e_k$ are distinct. In that case, note that the r vectors of advice to the second prover, $v_1, \ldots, v_r \in \mathbb{F}_q^V$, are uniform, and the second prover may consider their span Q_V. Note that, conditioned on (the likely event of) the advice vectors being linearly independent, Q_V is a uniformly random r dimensional subspace of $\mathbb{F}_q^V$. As for the first prover, the vectors $u_1, \ldots, u_r \in \mathbb{F}_q^U$ are not uniformly distributed. Nevertheless, as shown by the covering property from [KS13, KMS17] (and presented below), the distribution of $u_1, \ldots, u_r$ is close to uniform over r-tuple of vectors from $\mathbb{F}_q^U$. Thus, the first prover can also take their span, call it Q_U, and think of it as a random r-dimensional subspace of $\mathbb{F}_q^U$ (which is highly correlated to Q_V).

4 The Composed PCP Construction

In this section we describe our final PCP construction, which is a composition of the outer PCP from Section 3 with the inner PCP based on the Grassmann consistency test.

4.1 The Underlying Graph

Our reduction starts from an instance (X, Eq) of $\mathsf{Gap3Lin}$. Consider the game $G_{\beta,r}^{\otimes k}$ from Section 3, and let $\mathcal{U}$ denote the set of questions asked to the first prover. Thus $\mathcal{U}$ consists of all k-tuples of equations $U = (e_1, \dots, e_k) \in \mathsf{Eq}^k$ from the $\mathsf{Gap3Lin}$ instance (X, Eq). It will be convenient to only keep the $U = (e_1, \dots, e_k)$ that satisfy the following properties:

- The equations $e_1, \dots, e_k$ are distinct and do not share variables.
- For any $i \neq j$ and pair of variables $x \in e_i$ and $y \in e_j$, the variables x and y do not appear together in any equation in the instance (X, Eq).

The fraction of $U = (e_1, \dots, e_k)$ that do not satisfy the above is $O(k^2/n)$ which is negligible for us, and dropping them will only reduce our completeness by $o(1)$. This will not affect our analysis, and henceforth we will assume that all $U = (e_1, \dots, e_k)$ satisfy the above properties. We now describe the 2-Prover-1-Round Games instance $\Psi = (\mathcal{A}, \mathcal{B}, E, \Sigma_1, \Sigma_2, \Phi)$. All vertices in the underlying graph will correspond to subspaces of $\mathbb{F}_q^X$.

Notation: for $e \in \mathsf{Eq}$ let $v_e \in \mathbb{F}_q^X$ be the vector of coefficients of the equation e; we remark that as our instance (X, Eq) will come either from either Theorem 2.6 with $q = 2$ or Theorem 2.7, the vector v_e will be 1 on coordinates corresponding to variables in e, and 0 on all other coordinates.

4.1.1 The Vertices

For each question $U = (e_1, \dots, e_k)$, let $H_U = \mathrm{span}(v_{e_1}, \dots, v_{e_k})$, and think of it as a subspace of $\mathbb{F}_q^U$. By the first property described above, $\dim(H_U) = k$ and $\dim(\mathbb{F}_q^U) = 3k$. The vertices of Ψ are:

$$\mathcal{A} = \{L \oplus H_U \mid U \in \mathcal{U}, L \subseteq \mathbb{F}_q^U, \dim(L) = 2\ell, L \cap H_U = \{0\}\},$$
$$\mathcal{B} = \{R \mid \exists U \in \mathcal{U}, \text{ s.t. } R \subseteq \mathbb{F}_q^U, \dim(R) = 2(1-\delta)\ell\}.$$

Morally, the vertices in $\mathcal{A}$ are all 2ℓ-dimensional subspaces of some $\mathbb{F}_q^U$ for some $U \in \mathcal{U}$. For technical reasons, we require them to intersect H_U trivially (which is the case for a typical 2ℓ-dimensional space) and add to them the space H_U.[7] The vertices in $\mathcal{B}$ are all $2(1-\delta)\ell$ dimensional subspaces of $\mathbb{F}_q^U$.

4.1.2 The Alphabets

The alphabets Σ_1, Σ_2 have sizes $|\Sigma_1| = q^{2\ell}$ and $|\Sigma_2| = q^{2(1-\delta)\ell}$. For each vertex $L \oplus H_U \in \mathcal{A}$, let $\psi : H_U \to \mathbb{F}_q$ denote the linear function that satisfies the side conditions given by the equations in U. In notations, writing $e_i \in U$ as $\langle x, v_{e_i}\rangle = b_i$ for $x \in \mathbb{F}_q^U$, we set $\psi(v_{e_i}) = b_i$. We say a linear function $f\colon L \oplus H_U \to \mathbb{F}_q$ satisfies the side conditions of U if $f|_{H_U} \equiv \psi$. In this language, for a vertex $L \oplus H_U$ we identify Σ_1 with

$$\{f : L \oplus H_U \to \mathbb{F}_q \mid f \text{ is linear function satisfying the side conditions of } U\}.$$

[7] This has the effect of collapsing L and L' such that $L \oplus H_U = L' \oplus H_U$ to a single vertex.

As $L \cap H_U = \{0\}$ and $\dim(L) = 2\ell$, it is easy to see that the above set indeed has size $q^{2\ell}$. For each right vertex R, we identify Σ_2 with

$$\{f : R \to \mathbb{F}_q \mid f \text{ is linear}\}.$$

4.1.3 The Edges

To define the edges, we first need the following relation on $\mathcal{A}$. Say that $(L \oplus H_U) \sim (L' \oplus H_{U'})$ if

$$L + H_U + H_{U'} = L' + H_U + H_{U'}.$$

As all subspaces above are in $\mathbb{F}_q^X$, the equality above is well defined. We next show that this is an equivalence relation. The reflexivity and symmetry are clear, and the following lemma establishes transitivity.

Lemma 4.1. *If $L_1 + H_{U_1} + H_{U_2} = L_2 + H_{U_1} + H_{U_2}$, and $L_2 + H_{U_2} + H_{U_3} = L_3 + H_{U_2} + H_{U_3}$, then*

$$L_1 + H_{U_1} + H_{U_3} = L_3 + H_{U_1} + H_{U_3}.$$

Proof. We "add" H_{U_1} to the second equation to obtain,

$$L_2 + H_{U_1} + H_{U_2} + H_{U_3} = L_3 + H_{U_1} + H_{U_2} + H_{U_3}.$$

Next, write $H_{U_2} = A + B$, where A is the span of all vectors v_e for equations e in U_2 that are also in U_1 or U_3, while B is the span of all vectors v_e for equations $e \in U_2$ that are in neither U_1 nor U_3. It follows that $A \cap B = \{0\}$. Now note that any equation $e \in B$ has at most one variable that appears in an equation in U_1, and at most one variable that appears in an equation in U_3. Thus, each $e \in B$, has a "private variable", and as the equations in B are over disjoint sets of variables, this private variable does not appear in $U_1 \cup U_3 \cup (U_2 \setminus e)$. It follows that

$$B \cap (L_1 + L_3 + H_{U_1} + H_{U_3}) = \{0\} \subset \mathbb{F}_q^X. \tag{3}$$

Indeed, by the above discussion any nonzero vector in $B \subseteq \mathbb{F}_q^X$ is nonzero on at least one coordinate of X (corresponding to a private variable), and no vector in $\mathbb{F}_q^{U_1}$ or $\mathbb{F}_q^{U_2}$ is supported on this coordinate.

Substituting $H_{U_2} = A + B$ into the original equation yields

$$L_1 + (H_{U_1} + H_{U_3} + A) + B = L_3 + (H_{U_1} + H_{U_3} + A) + B,$$

and as $A \subset H_{U_1} + H_{U_3}$ we get $L_1 + H_{U_1} + H_{U_3} + B = L_3 + H_{U_1} + H_{U_3} + B$. Using (3) it follows that

$$L_1 + H_{U_1} + H_{U_3} = L_3 + H_{U_1} + H_{U_3}.$$

□

By Lemma 4.1 the relation $\sim$ is an equivalence relation, so we may partition $\mathcal{A}$ into equivalence classes. We denote the equivalence class of a vertex $L \oplus H_U$ by $[L \oplus H_U]$. We often refer to each equivalence class as a clique because this relation partitions $\mathcal{A}$ into cliques:

$$\mathcal{A} = \mathsf{Clique}_1 \sqcup \cdots \sqcup \mathsf{Clique}_m.$$

The actual number of cliques, m, will not be important, but it is clear that such a number exists. The edges of our graph will be between vertices $L \oplus H_U$ and R if there exists $L' \oplus H_{U'} \in [L \oplus H_U]$ such that $L' \supseteq R$. The edges will be weighted according to a sampling process that we describe in the next section, which also defines the constraints on Ψ. For future reference, the following lemma will be helpful in defining the constraints:

Lemma 4.2. *Suppose $L \oplus H_U \sim L' \oplus H_{U'}$ and let $f : L \oplus H_U \to \mathbb{F}_q$ be a linear function satisfying the side conditions on U. Then there is a unique linear function $f' : L' \oplus H_{U'} \to \mathbb{F}_q$ satisfying the side conditions on U' such that there exists a linear function $g : L + H_U + H_{U'} \to \mathbb{F}_q$ satisfying the side conditions of both U and U' such that*

$$g|_{L \oplus H_U} \equiv f \quad \text{and} \quad g|_{L' \oplus H_{U'}} \equiv f'.$$

In words, the above equations say g is a linear extension of both f and f'.

Proof. Note that there is only one way to extend f to $L + H_U + H_{U'}$ in a manner that satisfies the side conditions given by U'. Let this function be g. We take f' to be $g|_{L' \oplus H_{U'}}$. □

4.1.4 The Constraints

Suppose that T_1 is an assignment to $\mathcal{A}$ that assigns, to each vertex $L \oplus H_U$, a linear function $T_1[L \oplus H_U]$ satisfying the side conditions. Further suppose that T_2 is an assignment that assigns to each vertex $R \in \mathcal{B}$ a linear function on R. The verifier performs the following test, which also describes the constraints of Ψ:

1. Choose U uniformly at random from $\mathcal{U}$.
2. Choose $L \oplus H_U$ uniformly, where $\dim(L) = 2\ell$ and $L \cap H_U = \{0\}$, and choose $R \subseteq L$ of dimension $2(1-\delta)\ell$ uniformly.
3. Choose $L' \oplus H_{U'} \in [L \oplus H_U]$ uniformly
4. As in Lemma 4.2, extend $T_1[L' \oplus H_{U'}]$ to $L' + H_{U'} + H_U$ in the unique manner that respects the side conditions and let $\tilde{T}_1[L \oplus H_U]$ be the restriction of this extension to $L \oplus H_U$.
5. Accept if and only if $\tilde{T}_1[L \oplus H_U]|_R = T_2[R]$.

This finishes the description of our instance Ψ. It is clear that the running time and instance size is $n^{O(k)}$ and that the alphabet size is $O(q^{2\ell})$. As is often the case, showing the completeness of the reduction is relatively easy. The soundness analysis is much more complicated, and in the section below we develop some tools.

5 Tools for Soundness Analysis

In this section we will present all of the tools needed to analyze the soundness of our PCP.

5.1 The 2ℓ versus $2\ell(1-\delta)$ subspace agreement test

We begin by discussing the 2ℓ versus $2\ell(1-\delta)$ test and our decoding theorem for it. In our setting, we have a question $U \in \mathcal{U}$ for the first prover, and we consider the 2ℓ versus $2\ell(1-\delta)$ test inside the space $\mathbb{F}_q^U$. We assume that this test passes with probability at least $\varepsilon \geqslant q^{-2\ell(1-\delta')}$ (where δ' is, say, $\delta' = 1000\delta$) and we want to use this fact to devise a strategy for the first prover. Below, we first state and prove a basic decoding theorem, and then use it to deduce a quantitatively stronger version that also incorporates the side conditions.

Let T_1 be a table that assigns, to each $L \in \mathsf{Grass}(\mathbb{F}_q^U, 2\ell)$, a linear function $T_1[L] : L \to \mathbb{F}_q$, and let T_2 be a table that assigns, to each $R \in \mathsf{Grass}(\mathbb{F}_q^U, 2(1-\delta)\ell)$, a linear function $T_2[R] \colon R \to \mathbb{F}_q$. We recall that

$|U| = 3k \gg 2\ell$. In this section, we show that if tables T_1 and T_2 are ε-consistent, namely

$$\Pr_{\substack{L\in\mathsf{Grass}(\mathbb{F}_q^U,2\ell)\\ R\in\mathsf{Grass}(\mathbb{F}_q^U,2(1-\delta)\ell)}}[T_1[L]|_R \equiv T_2[R] \mid R\subseteq L] \geqslant \varepsilon.$$

for $\varepsilon \geqslant q^{-2\ell(1-1000\delta)}$, then the table T_1 must have non-trivial agreement with a linear function on some zoom-in and zoom-out combination of constant dimension. The proof uses Lemma 2.3 along with an idea from [BKS19].

Theorem 5.1. *Let $0<\delta<\frac{1}{1000}$ and let ℓ be sufficiently large relative to $1/\delta$. Suppose that tables T_1 and T_2 are ε-consistent where $\varepsilon \geqslant q^{-2(1-1000\delta)\ell}$. Then there exist subspaces $Q\subset W$ and a linear function $f: W\to\mathbb{F}_q$ such that:*

1. $\dim(Q)+\operatorname{codim}(W) = \frac{10}{\delta}$.
2. *$f|_L \equiv T_1[L]$ for ε'-fraction of 2ℓ-dimensional $L\in\mathsf{Zoom}_{2\ell}[Q,W]$, where $\varepsilon' = q^{-2(1-1000\delta^2)\ell}$.*

Proof. Consider the bipartite Grassmann inclusion graph, G, whose sides are the domains $\mathsf{Grass}(n,2\ell)$ and $\mathsf{Grass}(n,(1-\delta)2\ell)$, and denote its edge set by E. Choose a linear function $f:\mathbb{F}_q^n\to\mathbb{F}_q$ uniformly at random and define the (random) sets of vertices

$$S_{L,f} = \{L\in\mathsf{Grass}(n,2\ell)\mid f|_L\equiv T_1[L]\}\quad\text{and}\quad S_{R,f}=\{R\in\mathsf{Grass}(n,2(1-\delta)\ell)\mid f|_R\equiv T_2[R]\}.$$

Denote by $E(S_{L,f},S_{R,f})$ the set of edges with endpoints in $S_{L,f}$ and $S_{R,f}$. We lower bound the expected size of $E(S_{L,f},S_{R,f})$ over the choice of f. Note that for each edge $(L,R)\in E$ such that $T_1[L]|_R\equiv T_2[R]$, we have that $(L,R)\in E(S_{L,f},S_{R,f})$ with probability $q^{-2\ell}$. Indeed, with probability $q^{-2\ell}$ we have that $T_1[L]\equiv f|_L$, and in that case we automatically get that $T_2[R]\equiv T_1[L]|_R\equiv (f|_L)|_R\equiv f|_R$. As the number of edges (L,R) such that $T_1[L]|_R\equiv T_2[R]$ is at least $\varepsilon|E|$, we conclude that

$$\mathop{\mathbb{E}}_f[|E(S_{L,f},S_{R,f})|]\geqslant \varepsilon q^{-2\ell}|E|.$$

Note that we also have that

$$\mathop{\mathbb{E}}_f[\mu(S_{R,f})] = \mathop{\mathbb{E}}_f\left[\frac{|S_{R,f}|}{|R|}\right] = q^{-2\ell(1-\delta)}.$$

Using Linearity of Expectation, we get that

$$\mathop{\mathbb{E}}_f\left[|E(S_{L,f},S_{R,f})| - \frac{1}{2}\varepsilon q^{-2\delta\ell}\mu(S_{R,f})|E|\right]\geqslant \frac{1}{2}\varepsilon q^{-2\ell}|E|,$$

thus there exists f for which the random variable on the left hand side is at least $\frac{1}{2}\varepsilon q^{-2\ell}|E|$, and we fix f so that

$$|E(S_{L,f},S_{R,f})|\geqslant \frac{1}{2}\varepsilon q^{-2\delta\ell}\mu(S_{R,f})|E| + \frac{1}{2}\varepsilon q^{-2\ell}|E|. \tag{4}$$

We claim that $S_{L,f}$ is (r,ε')-pseudo-random for $r=\frac{10}{\delta}$ and $\varepsilon' = q^{-2\ell(1-1000\delta^2)}$. Suppose for the sake of contradiction that this is not the case, and that $S_{L,f}$ is (r,ε')-pseudo-random. Denote $\alpha=\mu(S_{L,f})$ and $\beta=\mu(S_{R,f})$. Since $S_{L,f}$ is not (r,ε')-pseudo-random, we may apply Lemma 2.3 and get:

$$\frac{1}{|E|}|E(S_{L,f},S_{R,f})| = \Pr_{L\supseteq R}[L\in S_{L,f},\in S_{R,f}]\leqslant q^{O_{t,r}(1)}\beta^{\frac{t-1}{t}}\varepsilon'^{\frac{t-2}{t}} + q^{-r\delta\ell}\sqrt{\alpha\beta}\leqslant q^{O_{t,r}(1)}\beta^{\frac{t-1}{t}}\varepsilon'^{\frac{t-2}{t}}, \tag{5}$$

for any $t \geqslant 4$ that is a power of 2. In the last inequality, we used the fact that by (4)

$$\beta|E| = |S_{R,f}|\frac{|E|}{|R|} \geqslant |E(S_{L,f}, S_{R,f})| \geqslant \frac{1}{2}\varepsilon q^{-2\ell}|E|,$$

so $\beta \geqslant \frac{1}{2}\varepsilon q^{-2\ell} \geqslant q^{-4\ell}$, and thus the second term in the middle of (5) is negligible compared to the first term there. Combining (4) and (5) gives us that

$$\frac{1}{2}\varepsilon q^{-2\delta\ell}\beta \leqslant q^{O_{t,r}(1)}\beta^{\frac{t-1}{t}}\varepsilon'^{\frac{t-2}{t}}.$$

Simplifying and using the definition of ε' along with the fact that $\varepsilon \geqslant q^{-2\ell(1-1000\delta)}$ and $\beta \geqslant q^{-4\ell}$ we get

$$\frac{1}{2}q^{-2\delta\ell} \leqslant q^{O_{t,r}(1)}q^{\frac{4\ell}{t}}q^{\left(\frac{4}{t}+2000\left(\delta^2\frac{t-2}{t}-\delta\right)\right)\ell}.$$

Investigating the last two exponents of q, we have that for $t \geqslant \frac{1}{\delta-\delta^2} \geqslant 2$,

$$\left(\frac{4}{t}+\frac{4}{t}+2000\left(\delta^2\frac{t-2}{t}-\delta\right)\right)\ell \leqslant -1992(\delta-\delta^2)\ell.$$

This implies that

$$\frac{1}{2}q^{-2\delta\ell} \leqslant q^{O_\delta(1)}q^{-1992(\delta-\delta^2)\ell},$$

and contradiction. It follows that $S_{L,f}$ is not (r,ε')-pseudo-random, and unraveling the definition of not being pseudo-random gives the conclusion of the theorem. □

5.1.1 Finding a Large Fraction of Successful Zoom-Ins

Theorem 5.1 asserts the existence of a good pair of zoom-in and zoom-out (Q, W) inside which the table T_1 has good agreement with a global linear function. As discussed in the introduction, our argument requires a quantitatively stronger version asserting that there is a good fraction of zoom-ins that work for us. Below, we state a corollary of Theorem 5.1 which achieves this. We defer its proof to Appendix B.

Theorem 5.2. *Suppose that tables T_1 and T_2 are ε-consistent for $\varepsilon \geqslant 2q^{-2\ell(1-1000\delta)}$. Then there exist positive integers r_1 and r_2 satisfying $r_1 + r_2 = r = \frac{10}{\delta}$, such that for at least $q^{-5\ell^2}$-fraction of the r_1-dimensional subspaces Q, there exists a subspace $W \supseteq Q$ of codimension r_2 and a linear function $g_{Q,W}$ such that*

$$\Pr_{L\in\mathsf{Grass}(\mathbb{F}_q^U,2\ell)}[g_{Q,W}|_L = T_1[L] \mid Q \subseteq L \subseteq W] \geqslant \frac{q^{-2\ell(1-1000\delta^2)}}{2}.$$

Proof. The proof is deferred to Section B. □

5.1.2 Incorporating Side Conditions for Zoom-Ins

Next, we require a version of Theorem 5.2 which also takes the side conditions into account.

Theorem 5.3. *Let U be a question to the first prover, let T_1 be the first prover's table and suppose that*

$$\Pr_{\substack{L\in\mathsf{Grass}(\mathbb{F}_q^U,2\ell),L\cap H_U=\{0\}\\ R\in\mathsf{Grass}(\mathbb{F}_q^U,2(1-\delta)\ell)}}[T_1[L\oplus H_U]|_R = T_2[R] \mid R\subseteq L] := \varepsilon \geqslant 4q^{-2(1-1000\delta)\ell}.$$

Then there are parameters r_1 and r_2 such that $r_1+r_2\leqslant \frac{10}{\delta}$, such that for at least $q^{-6\ell^2}$ fraction of the r_1-dimensional subspaces $Q\subseteq\mathbb{F}_q^U$, there exists $W\subseteq\mathbb{F}_q^U$ of codimension r_2 containing $Q\oplus H_U$, and a global linear function $g_{Q,W}:W\to\mathbb{F}_q$ that respects the side conditions on H_U such that

$$\Pr_{L\in\mathsf{Grass}(\mathbb{F}_q^U,2\ell),L\cap H_U=\{0\}}[g_{Q,W}|_{L\oplus H_U} = T_1[L\oplus H_U] \mid Q\subseteq L\subseteq W] \geqslant \frac{q^{-2(1-1000\delta^2)\ell}}{2}.$$

Proof. For any $2k$-dimensional subspace $A\subseteq\mathbb{F}_q^U$ such that $A\cap H_U=\{0\}$ and $H_U\oplus A=\mathbb{F}_q^U$, let T_A be the table over $\mathsf{Grass}(A,2\ell)$ defined as follows. For the 2ℓ-dimensional subspaces $L\subseteq A$ such that $L\cap H_U=\{0\}$, define $T_A[L]=T_1[L\oplus H_U]|_L$. Throughout the proof, R is always used to denote a random subspace of dimension $2(1-\delta)\ell$, so we will not specify this further in the distribution of probabilities.

Let

$$p'(A) = \Pr_{L\in\mathsf{Grass}(A,2\ell),R\subseteq L}[T_A[L]|_R \equiv T_2[R]].$$

Noting that

$$\mathop{\mathbb{E}}_{A}[p'(A)] = \Pr_{L\in\mathsf{Grass}(\mathbb{F}_q^U,2\ell),L\cap H_U=\{0\},R\subseteq L}[T_1[L\oplus H_U]|_R \equiv T_2[R]] = \varepsilon,$$

we get by an averaging argument that $p'(A)\geqslant\frac{\varepsilon}{2}$ for at least $\frac{\varepsilon}{2}$-fraction of A's. Fix any such A. By Theorem 5.2 there exist positive integers r_1 and r_2 such that for at least $q^{-5\ell^2}$-fraction of r_1-dimensional zoom-ins $Q\subseteq A$, there exists a zoom-out $W'\supset Q$ of codimension r_2 and a linear function $g_{Q,W'}$ such that,

$$\Pr_{L\in\mathsf{Grass}(A,2\ell)}[g_{Q,W'}|_L \equiv T_A[L] \mid Q\subseteq L\subseteq W'] \geqslant \frac{q^{-2(1-1000\delta^2)\ell}}{2}. \tag{6}$$

Let $W=W'\oplus H_U$ and let $g_{Q,W}:W\to\mathbb{F}_q$ be the unique extension of $g_{Q,W'}$ to W satisfying the side conditions. We have

$$\begin{aligned}
&\Pr_{L\in\mathsf{Grass}(\mathbb{F}_q^U,2\ell),L\cap H_U=\{0\}}[g_{Q,W}|_{L\oplus H_U} \equiv T_1[L\oplus H_U] \mid Q\subseteq L\subseteq W]\\
&= \Pr_{L\in\mathsf{Grass}(A,2\ell)}[g_{Q,W}|_{L\oplus H_U} \equiv T_1[L\oplus H_U] \mid Q\subseteq L\subseteq W']\\
&= \Pr_{L\in\mathsf{Grass}(A,2\ell)}[g_{Q,W'}|_L \equiv T_A[L] \mid Q\subseteq L\subseteq W']\\
&\geqslant \frac{q^{-2(1-1000\delta^2)\ell}}{2}
\end{aligned}$$

In the first transition we used the fact that $L\oplus H_U$ is distributed the same in the first two expressions. In the second transition we used the fact that $g_{Q,W}$ and $T[L\oplus H_U]$ agree on H_U because both satisfy the side conditions on H_U, along with the fact that $g_{Q,W}|_A\equiv g_{Q,W'}$.

To conclude, we see that when sampling A as above and then $Q\subseteq A$ of dimension r_1, the zoom-in Q has a zoom-out $W\supseteq Q\oplus H_U$ and a function $g_{Q,W}:W\to\mathbb{F}_q$ satisfying the conditions in the theorem with probability at least $\frac{\varepsilon}{2}q^{-5\ell^2}$. The theorem requires us to bound the probability that Q satisfies this property when Q is chosen uniformly in $\mathbb{F}_q^U$ of dimension r_1, which is a slightly different distribution than what we have. Indeed, when sampling A as above and then $Q \subseteq A$ of dimension r_1, the subspace Q is uniformly random r_1-dimensional subspace in $\mathbb{F}_q^U$ conditioned on $Q \cap H_U = \{0\}$. However, $\left(1 - q^{r_1 - 2k}\right)$-fraction of r_1-dimensional subspaces $Q \subseteq \mathbb{F}_q^U$ satisfy $Q \cap H_U = \{0\}$, so overall we get that at least $\left(\frac{\varepsilon}{2} q^{-5\ell^2} - q^{r_1 - 2k}\right) \geqslant q^{-6\ell^2}$-fraction of all r_1-dimensional $Q \subseteq \mathbb{F}_q^U$ satisfy the condition of the theorem. $\square$

5.2 The Covering Property

In this section, we present the so-called "covering property", which is a feature of our PCP construction that allows us to move between the first prover's distribution over 2ℓ-dimensional subspaces of $\mathbb{F}_q^U$ and the second prover's distribution over 2ℓ-dimensional subspaces of $\mathbb{F}_q^V$. Similar covering properties are shown in [KS13, KMS17]; however, obtaining the optimal quadratic-programming hardness result in Theorem 1.4 requires a stronger analysis that goes beyond the covering properties of [KS13, KMS17]. We are able to obtain a covering property that holds with the following parameters set in the outer PCP:

$$k = q^{2(1+c)\ell} \quad , \quad \beta = q^{-2(1+2c/3)\ell}, \tag{7}$$

where $c > 0$ is a constant arbitrarily small relative to δ.

Remark 5.4. *We remark that the content of this section is only necessary if we take $k = q^{(2+c)\ell}$ as in (7). This setting is used to get a good tradeoff between the size, alphabet size and soundness of the PCP, which is crucial in the proof of Theorem 1.4. If one is only interested in the tradeoff between the alphabet size and the soundness, i.e. only wants to prove Theorem 1.3, then one can afford to take k as a much bigger function of ℓ. For example, taking $k = q^{100\ell^2}$, one could use the covering property from [KS13, KMS17] instead of the more complicated covering property below, and replace the auxiliary lemma in Section 5.2.4 by a simple averaging argument.*

5.2.1 The Basic Covering Property

We start by stating a basic form of the improved covering property, and defer its proof to Appendix C. Fix a question $U = (e_1, \ldots, e_k)$ to the first prover and recall that $H_U = \mathrm{span}(v_{e_1}, \ldots, v_{e_k})$ where v_{e_i} is the vector of coefficients of e_i. For $V \subset U$ we view $\mathbb{F}_q^V \subset \mathbb{F}_q^U$ as the subspace consisting of vectors that are zero on all indices in $U \setminus V$. The covering property we show will relate the following two distributions:
$\mathcal{D}$:

- Choose $x_1, \ldots, x_{2\ell} \in \mathbb{F}_q^U$ uniformly.
- Output the list $(x_1, \ldots, x_{2\ell})$.

$\mathcal{D}'$:

- Choose $V \subseteq U$ according to the Outer PCP.
- Choose $x'_1, \ldots, x'_{2\ell} \in \mathbb{F}_q^V$ uniformly.
- Choose $w_1, \ldots, w_{2\ell} \in H_U$ uniformly, and set $x_i = x'_i + w_i$ for $1 \leqslant i \leqslant 2\ell$.
- Output the list $(x_1, \ldots, x_{2\ell})$.

The covering property used in prior works asserts that the distribution $\mathcal{D}$ is statistically close to a variant of the distribution $\mathcal{D}'$. This closeness is not good enough for us, however, as we will want to consider events of rather small probability under $\mathcal{D}$ and still assert that their probability is roughly the same under $\mathcal{D}'$.

Our version of the covering property deviates from those of prior works in two ways. First, in these earlier works the distribution $\mathcal{D}'$ was generated without the addition of the random vectors $w_1, \dots, w_{2\ell}$ from H_U. As explained in the introduction, this distribution is not good enough for the purpose of Theorem 1.4, and we must consider the distribution $\mathcal{D}'$ above instead. Second, the notion of statistical closeness is too strict for us, and is in fact not achievable. To get around this hurdle, we instead show a "contiguity-type" statement, asserting that almost all inputs x are assigned the same probability under these two distributions up to factor $1 + o(1)$.

More precisely, setting $\eta = q^{-100\ell^{100}}$ throughout this subsection, we show:

Lemma 5.5. *There exists $E \subseteq \left(\mathbb{F}_q^U\right)^{2\ell}$ such that both $\mathcal{D}(E)$ and $\mathcal{D}'(E)$ are at most η^{40}, and for all $(x_1, \dots, x_{2\ell}) \notin E$ we have*

$$0.9 \leqslant \frac{\mathcal{D}(x_1, \dots, x_{2\ell})}{\mathcal{D}'(x_1, \dots, x_{2\ell})} \leqslant 1.1.$$

Proof. The proof is deferred to Appendix C. □

5.2.2 The Covering Property with Zoom-ins

Lemma 5.5 represents the basic form of the covering property, but for our application we require a version which incorporates zoom-ins and advice. Namely, we will actually be interested in the case where $\mathcal{D}$ and $\mathcal{D}'$ are conditioned on some r_1-dimensional zoom-in Q, for an arbitrary dimension $r_1 \leqslant \frac{10}{\delta}$. To make notation simpler, let us write $x = (x_1, \dots, x_{2\ell})$ and use $\mathrm{span}_{r_1}(x)$ to denote $\mathrm{span}(x_1, \dots, x_{r_1})$. Additionally, for any set $\mathcal{L} \subseteq \left(\mathbb{F}_q^U\right)^{2\ell}$, define

$$\mathcal{D}_Q(\mathcal{L}) := \Pr_{x \sim \mathcal{D}}[x \in \mathcal{L} \mid \mathrm{span}_{r_1}(x) = Q] = \frac{\mathcal{D}(\{x \in \mathcal{L} \mid \mathrm{span}_{r_1}(x) = Q\})}{\mathcal{D}(\{x \mid \mathrm{span}_{r_1}(x) = Q\})}.$$

Also define $\mathcal{D}'_Q$ similarly as

$$\mathcal{D}'_Q(\mathcal{L}) := \Pr_{x \sim \mathcal{D}'}[x \in \mathcal{L} \mid \mathrm{span}_{r_1}(x) = Q] = \frac{\mathcal{D}'(\{x \in \mathcal{L} \mid \mathrm{span}_{r_1}(x) = Q\})}{\mathcal{D}'(\{x \mid \mathrm{span}_{r_1}(x) = Q\})}.$$

From Lemma 5.5 we can conclude that for any $\mathcal{L} \subseteq (\mathbb{F}_q^U)^{2\ell}$ that is not too small, the measure $\mathcal{D}'_Q(\mathcal{L})$ is within at least a constant factor of $\mathcal{D}_Q(\mathcal{L})$ for nearly all Q:

Lemma 5.6. *For any $\mathcal{L} \subseteq (\mathbb{F}_q^U)^{2\ell}$, we have*

$$\Pr_Q\left[\mathcal{D}'_Q(\mathcal{L}) \geqslant 0.8 \cdot \mathcal{D}_Q(\mathcal{L}) - \eta^{20}\right] \geqslant 1 - 3\eta^{20},$$

where Q is the span of r_1 uniformly random vectors in $\mathbb{F}_q^U$.

Proof. Throughout the proof all of the expectations and probabilities over Q choose Q as in the lemma statement. Let E be the small set of tuples from Lemma 5.5, so that

$$\mathop{\mathbb{E}}_{x \sim \mathcal{D}, Q = \mathrm{span}_{r_1}(x)}[\mathcal{D}_Q(E)] \leqslant \eta^{40} \quad \text{and} \quad \mathop{\mathbb{E}}_{x \sim \mathcal{D}', Q = \mathrm{span}_{r_1}(x)}[\mathcal{D}'_Q(E)] \leqslant \eta^{40}. \tag{8}$$

Let $\mathsf{Bad}_1 = \{x \in \left(\mathbb{F}_q^U\right)^{2\ell} \mid \mathcal{D}_Q(E) \geqslant \eta^{20}, Q := \mathrm{span}_{r_1}(x)\}$ and $\mathsf{Bad}_2 = \{x \in \left(\mathbb{F}_q^U\right)^{2\ell} \mid \mathcal{D}'_Q(E) \geqslant \eta^{20}, Q := \mathrm{span}_{r_1}(x)\}$. By the union bound we have

$$\Pr_{x\sim\mathcal{D}}[x \in \mathsf{Bad}_1 \cup \mathsf{Bad}_2] \leqslant \Pr_{x\sim\mathcal{D}}[x \in \mathsf{Bad}_1] + \Pr_{x\sim\mathcal{D}}[x \in \mathsf{Bad}_2].$$

By the first expectation (8) and Markov's inequality, we have $\Pr_{x\sim\mathcal{D}}[x \in \mathsf{Bad}_1] \leqslant \eta^{20}$. To bound the second term, we have

$$\Pr_{x\sim\mathcal{D}}[x \in \mathsf{Bad}_2] = \sum_{x\in\mathsf{Bad}_2} \mathcal{D}(x) \leqslant \sum_{x\in\mathsf{Bad}_2\setminus E} \mathcal{D}(x) + \sum_{x\in E} \mathcal{D}(x) \leqslant \sum_{x\in\mathsf{Bad}_2\setminus E} 1.1\cdot\mathcal{D}'(x) + \eta^{40} \leqslant 1.1\cdot\eta^{20} + \eta^{40}.$$

In the penultimate transition we are using Lemma 5.5 to bound $\mathcal{D}(x) \leqslant 1.1 \cdot \mathcal{D}'(x)$ if $x \notin E$ as well as to bound $\mathcal{D}(E) \leqslant \eta^{40}$. In the final transition we are using the fact $\Pr_{x\sim\mathcal{D}'}[x \in \mathsf{Bad}_2] \leqslant \eta^{20}$, which follows by an application of Markov's inequality on the second expectation in (8).

Altogether, we get that with probability at least $1-3\eta^{20}$, we have $x \notin \mathsf{Bad}_1 \cup \mathsf{Bad}_2$, and $Q := \mathrm{span}_{r_1}(x)$ satisfies $\mathcal{D}_Q(E), \mathcal{D}'_Q(E) \leqslant \eta^{20}$. In this case we have,

$$\begin{aligned}
\sum_{\mathrm{span}_{r_1}(x)=Q} \mathcal{D}(x) \geqslant \sum_{\mathrm{span}_{r_1}(x)=Q, x\notin E} \mathcal{D}(x) &\geqslant \sum_{\mathrm{span}_{r_1}(x)=Q, x\notin E} 0.9 \cdot \mathcal{D}'(x) \\
&= 0.9 \cdot \left(\sum_{\mathrm{span}_{r_1}(x)=Q} \mathcal{D}'(x) - \sum_{\mathrm{span}_{r_1}(x)=Q, x\in E} \mathcal{D}'(x) \right),
\end{aligned}$$

where we used Lemma 5.5 in the second transition. Dividing both sides by $\sum_{\mathrm{span}_{r_1}(x)=Q} \mathcal{D}'(x)$ gives that

$$\frac{\sum_{\mathrm{span}_{r_1}(x)=Q} \mathcal{D}(x)}{\sum_{\mathrm{span}_{r_1}(x)=Q} \mathcal{D}'(x)} \geqslant 0.9(1 - \mathcal{D}'_Q(E)) \geqslant 0.9(1 - \eta^{20}) \geqslant 0.89. \tag{9}$$

It follows that

$$\begin{aligned}
\mathcal{D}'_Q(\mathcal{L}) &= \frac{\sum_{x\in\mathcal{L}, \mathrm{span}_{r_1}(x)=Q} \mathcal{D}'(x)}{\sum_{\mathrm{span}_{r_1}(x)=Q} \mathcal{D}'(x)} \\
&\geqslant \frac{0.89 \sum_{x\in\mathcal{L}\cap\overline{E}, \mathrm{span}_{r_1}(x)=Q} \mathcal{D}'(x)}{\sum_{\mathrm{span}_{r_1}(x)=Q} \mathcal{D}(x)} \\
&\geqslant \frac{0.9 \cdot 0.89 \cdot \sum_{x\in\mathcal{L}\cap\overline{E}, \mathrm{span}_{r_1}(x)=Q} \mathcal{D}(x)}{\sum_{\mathrm{span}_{r_1}(x)=Q} \mathcal{D}(x)} \\
&\geqslant 0.8 \cdot \mathcal{D}_Q(\mathcal{L}) - \frac{\sum_{\mathrm{span}_{r_1}(x)=Q, x\in E} \mathcal{D}(x)}{\sum_{\mathrm{span}_{r_1}(x)=Q} \mathcal{D}(x)} \\
&= 0.8 \cdot \mathcal{D}_Q(\mathcal{L}) - \mathcal{D}_Q(E) \\
&\geqslant 0.8 \cdot \mathcal{D}_Q(\mathcal{L}) - \eta^{20},
\end{aligned}$$

where we used (9) in the second transition and $\mathcal{D}_Q(E) \leqslant \eta^{20}$ in the last transition. $\square$

5.2.3 The Covering Property for the Advice

We will also need a lemma that applies to r_1-dimensional subspaces for some constant $r_1 = O(\delta^{-1})$. This is to handle the fact that the zoom-in Q is sampled uniformly from $\mathbb{F}_q^V$ after V is chosen according to the outer PCP, and then lifted to a subspace over $\mathbb{F}_q^U$, instead of uniformly from $\mathbb{F}_q^U$. For a fixed question U to the first prover, let $\mathcal{D}_{r_1}$ denote the uniform measure over $\left(\mathbb{F}_q^U\right)^{r_1}$ and let $\mathcal{D}'_{r_1}$ denote the following distribution over $\left(\mathbb{F}_q^U\right)^{r_1}$:

- Choose $V \subseteq U$ according to the Outer PCP.
- Choose $x'_1, \ldots, x'_{r_1} \in \mathbb{F}_q^V$ uniformly.
- Choose $w_1, \ldots, w_{r_1} \in H_U$ uniformly, and set $x_i = x'_i + w_i$ for $1 \leqslant i \leqslant r_1$.
- Output the list $(x_1, \ldots, x_{r_1})$.

The distributions $\mathcal{D}_{r_1}$ and $\mathcal{D}'_{r_1}$ have the following property.

Lemma 5.7. *For $\mathcal{Q} \subseteq \left(\mathbb{F}_q^U\right)^{r_1}$ such that $\mathcal{D}_{r_1}(\mathcal{Q}) \geqslant q^{-10\ell^{10}}$, we have*

$$\mathcal{D}'_{r_1}(\mathcal{Q}) \geqslant 0.8 \cdot \mathcal{D}_{r_1}(\mathcal{Q}).$$

Proof. Take E from Lemma 5.5 and define

$$\mathcal{L} = \{(x_1, \ldots, x_{2\ell}) \in \left(\mathbb{F}_q^U\right)^{2\ell} \mid (x_1, \ldots, x_{r_1}) \in \mathcal{Q}\}.$$

Then $\mathcal{D}_{r_1}(\mathcal{Q}) = \mathcal{D}(\mathcal{L})$, $\mathcal{D}'_{r_1}(\mathcal{Q}) = \mathcal{D}'(\mathcal{L})$, and

$$\mathcal{D}'_{r_1}(\mathcal{Q}) = \mathcal{D}'(\mathcal{L}) \geqslant \sum_{x \in \mathcal{L} \setminus E} \mathcal{D}'(x) \geqslant 0.9 \cdot \sum_{x \in \mathcal{L} \setminus E} \mathcal{D}(x) \geqslant 0.9 \cdot \mathcal{D}(\mathcal{L}) - \mathcal{D}(E) \geqslant 0.8 \cdot \mathcal{D}(\mathcal{L}) = 0.8 \cdot \mathcal{D}_{r_1}(\mathcal{L}),$$

where we used Lemma 5.5 in the second inequality, and the fact that $\mathcal{D}(E) \leqslant \eta^{40}$ and $\mathcal{D}(\mathcal{L}) = \mathcal{D}_{r_1}(\mathcal{Q}) \geqslant q^{-10\ell^{10}}$ in the penultimate transition. □

5.2.4 An Auxiliary Lemma

We conclude this section with an auxiliary lemma that will be used in the analysis. Throughout this subsection, we fix a question $U = (U_1, \ldots, U_k)$ to the first prover, where U_i denotes the variables from the ith equation, say e_i, and a subspace $Q \subseteq \mathbb{F}_q^U$ of dimension r, and we consider the distribution over the second prover's question, V, conditioned on the first r advice vectors having span Q. More formally, consider the distribution $\mathcal{A}(\cdot, \cdot)$ over (V, Q') defined as follows:

- For each $i \in [k]$, choose $V_i \subseteq U_i$ according to the outer PCP, and set $V = \bigcup_{i=1}^k V_i$.
- Choose r vectors $v_1, \ldots, v_r \in \mathbb{F}_q^V$ uniformly at random.
- Output $(V, \mathrm{span}(v_1, \ldots, v_r))$.

Then the distribution we are interested in is the marginal distribution $\mathcal{A}(\cdot, Q)$. After V is chosen, we let $V_1, \ldots, V_k$ be the variables in V from equations $e_1, \ldots, e_k$ respectively. We start with a helpful intermediate result, showing that in the sampling procedure, for any Q and $i \in [k]$, the probability $V_i \neq U_i$ is roughly at most β even when conditioned on V_j already being chosen for $j \in \mathcal{I} \subseteq [k]$ (where $\mathcal{I}$ can be an arbitrary subset of indices). For a vector $v \in \mathbb{F}_q^U$, we use v^i to denote its restriction to the variables in U_i, and think of v^i as being in the 3-dimensional space $\mathbb{F}_q^{U_i}$.

Lemma 5.8. *Fix a question $U = (U_1, \ldots, U_k)$, a dimension r subspace $Q \subseteq \mathbb{F}_q^U$, a parameter $0 < \beta < q^{-2\ell}$, and let $\mathcal{A}$ be defined accordingly as above using U and β. Let $\mathcal{I} \subseteq [k]$ be an arbitrary set of indices and suppose for $j \in \mathcal{I}$, $V_j \subseteq U_j$ has already been chosen. Then, for any $i \in [k] \setminus \mathcal{I}$*

$$\Pr_{V \sim \mathcal{A}(\cdot, Q)}[V_i \neq U_i \mid \{V_j\}_{\forall j \in \mathcal{I}}] := \Pr_{V_i \subseteq U_i \; Q' \subseteq \mathbb{F}_q^V}[V_i \neq U_i \mid Q' = Q, \{V_j\}_{\forall j \in \mathcal{I}}] \leqslant \frac{\beta}{(1-\beta)q^{-3r}}.$$

In words, the above probability is the probability that variables are dropped in equation i, when sampling $V \subseteq U$, conditioned on the span of the first r advice vectors equaling Q, and conditioned on V_j being fixed for $j \in \mathcal{I}$.[8]

Proof. View each vector in $\mathbb{F}_q^U$ as indexed by the variables appearing in U. It will be helpful to briefly recall how $V \subseteq U$ and the advice subspace Q are chosen according to the conditional distribution we consider. First, for $j \in \mathcal{I}$, $V_j \subseteq U_j$ is already fixed. For each $j \in [k] \setminus \mathcal{I}$ independently we set $U_j = V_j$ with probability $1 - \beta$, or with probability β, we set V_j to be a uniformly random subset of U_j of size 1. After choosing V_j, we choose r advice vectors $v_1^j, \ldots, v_r^j \in \mathbb{F}_q^{V_i}$ uniformly at random. We then set $V = \bigcup_{j=1}^k V_j$, for each $t \in [r]$ set the advice vector $v_t \in \mathbb{F}_q^V$ to be the concatenation of $v_t^1, \ldots, v_t^k$, and set $Q' = \operatorname{span}(v_1, \ldots, v_r)$.

It is clear that there is some distribution over bases $(a_1, \ldots, a_r)$ of Q such that

$$\Pr_{V_j \subseteq U_j, \forall j \notin \mathcal{I}, \; Q' \subseteq \mathbb{F}_q^V}\left[V_i \neq U_i \mid Q' = Q\right] = \mathop{\mathbb{E}}_{a_1, \ldots, a_r}\left[\Pr_{V_j \subseteq U_j, \forall j \notin \mathcal{I}, \; Q' \subseteq \mathbb{F}_q^V}[V_i \neq U_i \mid v_t = a_t \; \forall t \in [r]]\right],$$

where the expectation on the right is over this distribution of bases $(a_1, \ldots, a_r)$ of Q. We do not need to know what this distribution is. We only use the fact that, by its existence, there is some basis $a_1, \ldots, a_r$ which maximizes the inner probability on the right, and fixing this basis $(a_1, \ldots, a_r)$ of Q, we have

$$\begin{aligned}\Pr_{V_j \subseteq U_j, \forall j \notin \mathcal{I}, \; Q' \subseteq \mathbb{F}_q^V}\left[V_i \neq U_i \mid Q' = Q\right] &\leqslant \Pr_{V_j \subseteq U_j, \forall j \notin \mathcal{I}, \; Q' \subseteq \mathbb{F}_q^V}[V_i \neq U_i \mid v_t = a_t \; \forall t \in [r]] \\ &= \frac{\Pr_{V_j \subseteq U_j, \forall j \notin \mathcal{I}, \; v_1, \ldots, v_r \in \mathbb{F}_q^V}[V_i \neq U_i \wedge (v_t = a_t \; \forall t \in [r])]}{\Pr_{V_j \subseteq U_j, \forall j \notin \mathcal{I}, \; v_1, \ldots, v_r \in \mathbb{F}_q^V}[v_t = a_t \; \forall t \in [r]]}.\end{aligned}$$

[8] We implicitly assume that the probability space we consider is nonempty, so that there exist $V_s \subseteq U_s$ for $s \notin \mathcal{I}$ such that $Q' \subseteq \mathbb{F}_q^V$ after setting $V = \bigcup_{i=1}^k V_i$, with V_j fixed as above for $j \in \mathcal{I}$.

Since $V_j \subseteq U_j$ is chosen independently, the above probability becomes

$$\frac{\Pr\limits_{V_i \subseteq U_i, v_t^i \in \mathbb{F}_q^{V_i}} \left[V_i \neq U_i \wedge (v_t^i = a_t^i \ \forall t \in [r])\right] \prod\limits_{j \in [k] \setminus \{i\}} \Pr\limits_{V_j \subseteq U_j, v_t^j \in \mathbb{F}_q^{V_j}} \left[v_t^j = a_t^j \ \forall t \in [r]\right]}{\prod\limits_{j=1}^{k} \Pr\limits_{V_j \subseteq U_j, v_t^j \in \mathbb{F}_q^{V_j}} \left[v_t^j = a_t^j \ \forall t \in [r]\right]}$$
$$= \frac{\Pr\limits_{V_i \subseteq U_i, v_t^i \in \mathbb{F}_q^{V_i}} \left[V_i \neq U_i \wedge (v_t^i = a_t^i \ \forall t \in [r])\right]}{\Pr\limits_{V_i \subseteq U_i, v_t^i \in \mathbb{F}_q^{V_i}} \left[v_t^i = a_t^i \ \forall t \in [r]\right]}$$
$$\leqslant \frac{\beta}{\Pr\limits_{V_i \subseteq U_i, v_t^i \in \mathbb{F}_q^{V_i}} \left[v_t^i = a_t^i \ \forall t \in [r]\right]}$$
$$\leqslant \frac{\beta}{(1-\beta)q^{-3r}}.$$

In the third transition, we upper bound the numerator by $\Pr[V_i \neq U_i] \leqslant \beta$, and in the last transition we lower bound the denominator by $(1-\beta)(1/|\mathbb{F}_q^{U_i}|)^r = q^{-3r}$, which is the probability that $V_i = U_i$ and $v_t^i = a_t^i$ for all $t \in [r]$. □

In the soundness analysis we will find some zoom-in and zoom-out pair (Q, W) for the larger prover, and we will be interested in projecting it into a question V of the smaller prover. In particular, we will be interested in the distribution of the codimension $W \cap \mathbb{F}_q^V$. The following consequence of Lemma 5.8 upper bounds the probability its codimension is smaller by j than the codimension of W.

Lemma 5.9. *For integer $r, s \geqslant 0$, the following holds for sufficiently large k. Let U be a fixed question to the first prover in the outer PCP consisting of $3k$ distinct variables in some set of k equations, let $Q \subseteq \mathbb{F}_q^U$ be a subspace of dimension r, and let $W \subseteq \mathbb{F}_q^U$ be a codimension s subspace that contains both H_U and Q. Let $V \subseteq U$ be a random question to the second prover chosen according to the outer PCP, conditioned on $Q \subseteq \mathbb{F}_q^V$, and let $W[V] = \mathbb{F}_q^V \cap W$. Then, for each $0 \leqslant j \leqslant s$, we have*

$$\Pr_{V \sim \mathcal{A}(\cdot, Q)}[\mathrm{codim}(W[V]) = s - j] \leqslant (2s\beta q^{3r})^j,$$

where the codimension is with respect to $\mathbb{F}_q^V$.

Proof. Let us label the variables in U as $(x_1, \ldots, x_{3k})$ so that the entries of vectors in $\mathbb{F}_q^U$ are indexed by variables in $(x_1, \ldots, x_{3k})$. Recall that the equations making up U are labeled as $e_1, \ldots, e_k$. We will choose V gradually, where at each step, we choose a new equation from $\{e_1, \ldots, e_k\}$ and then choose which variables to drop from it according to the sampling procedure of the outer PCP. Specifically, at step i, let $e'_1, \ldots, e'_{i-1} \in \{e_1, \ldots, e_k\}$ be the equations that have been processed already. We choose $e'_i \in \{e_1, \ldots, e_k\} \setminus \{e'_1, \ldots, e'_{i-1}\}$. Then, with probability $1-\beta$ we do not drop any variables from e'_i, and with probability β we drop a uniformly random two out of the three variables from e'_i. Recall that when a variable is dropped when obtaining V, every vector in $\mathbb{F}_q^V$ has coordinate zero at the index of the dropped variable.

Let $V_0 = U$ and let V_i be the state of V after step i in this process, i.e. after we have processed equations $e'_1, \ldots, e'_i \in \{e_1, \ldots, e_k\}$ and chosen which variables to drop from these equations. Also define $W[V_i] := W \cap \mathbb{F}_q^{V_i}$ and say that $\mathrm{codim}(W[V_i])$ is the codimension of $W[V_i]$ with respect to $\mathbb{F}_q^{V_i}$. We will first show that it is possible to adaptively choose the first $k-s$ equations so that the codimension does not drop, i.e. $\mathrm{codim}(W[V_0]) = \cdots = \mathrm{codim}(W[V_{k-s}])$. For each one of the remaining s steps we will (roughly) show that the codimension drops by 1 with probability at most β, and does not change with probability $1-\beta$.

At step i, call an equation e saturated if the restriction of $W[V_{i-1}]$ to the variables in e has dimension 3. If we choose $e'_i = e$, it is straightforward to check that $\mathrm{codim}(W[V_i]) = \mathrm{codim}(W[V_{i-1}])$ with probability 1. We argue inductively that for the first $k-s$ steps, we can always adaptively choose a saturated equation, and thus have $\mathrm{codim}(W[V_0]) = \cdots = \mathrm{codim}(W[V_{k-s}])$.

Fix an arbitrary i such that $1 \leqslant i \leqslant k-s$ and suppose $\mathrm{codim}(W[V_{i-1}]) = s$. We claim that there can be at most s unsaturated equations in $\{e_1, \ldots, e_k\} \setminus \{e'_1, \ldots, e'_{i-1}\}$. To see this, note that for each equation $e \in \{e_1, \ldots, e_k\} \setminus \{e'_1, \ldots, e'_{i-1}\}$ that is unsaturated there is a corresponding vector in the dual space of $W[V_{i-1}] \subset \mathbb{F}_q^{V_{i-1}}$ supported only on the indices corresponding to variables in e. Thus, for any set of distinct unsaturated equations, these vectors in the dual space are linearly independent. Since the dual space of $W[V_{i-1}]$ has dimension at most s, there are thus at most s unsaturated equations. Therefore, if $1 \leqslant i \leqslant k-s$, then there is always at least one saturated equation in $\{e_1, \ldots, e_k\} \setminus \{e'_1, \ldots, e'_{i-1}\}$, and we can adaptively choose e'_i to be this equation.

We now show that for the unsaturated equations, the codimension can drop by 1 with probability at most $\frac{\beta}{(1-\beta)q^{-3r}}$ and otherwise the codimension does not change. Let e_i be an equation that is unsaturated. By Lemma 5.8, with probability at least $1 - \frac{\beta}{(1-\beta)q^{-3r}}$ no variables in e_i are dropped and in this case we have $V_i = V_{i+1}$, $W[V_i] = W[V_{i+1}]$, and hence $\mathrm{codim}(W[V_i]) = \mathrm{codim}(W[V_{i+1}])$.

With the remaining at most $\frac{\beta}{(1-\beta)q^{-3r}}$ probability, two variables are dropped, and in this case we show that the codimension decreases by at most 1. Suppose a_1, a_2 are the characteristic vectors for the variables that are dropped. Recall that $H_U \subseteq W$ at the start, and let w_i be the coefficient vector of the ith equation so that $w_i \in H_U$. Then, $w_i \in W[V_i]$, so it must be the case that a_1 is not yet in the dual space of $W[V_i]$, and so, when viewing both $W[V_{i+1}]$ and $W[V_i]$ as subspaces of $\mathbb{F}_q^{V_i}$, we get that $\dim(W[V_i]) - \dim(W[V_{i+1}]) \geqslant 1$. Combining this with the fact that $\dim\left(\mathbb{F}_q^{V_i}\right) - \dim\left(\mathbb{F}_q^{V_{i+1}}\right) = 2$, where again both spaces are viewed as subspaces of $\mathbb{F}_q^{V_i}$ it follows that the codimension of $W[V_{i+1}]$ relative to $\mathbb{F}_q^{V_{i+1}}$ is at most one more than that of $W[V_i]$ relative to $\mathbb{F}_q^{V_i}$.

Altogether, this shows that for $k-s$ of the steps i, the codimension does not change, while for the remaining s steps, the codimension drops by 1 with probability at most $\frac{\beta}{(1-\beta)q^{-3r}}$ and remains the same otherwise. As a result,

$$\Pr_{V \sim \mathcal{A}(\cdot, Q)}[\mathrm{codim}(W[V]) = s-j] \leqslant \binom{s}{j}\left(\frac{\beta}{(1-\beta)q^{-3r}}\right)^j \leqslant (2s\beta q^{3r})^j,$$

where the second inequality is by union bounding over all possible combinations of j out of s steps where the codimension may drop by 1. $\square$

We finish this section by stating a corollary of Lemma 5.9 which will be used in the soundness analysis. The setting one should have in mind is that we have a zoom-in and zoom-out pair (Q, W) and an event $\mathcal{L} \subseteq (\mathbb{F}_q^U)^{2\ell}$ inside it with noticeable probability, where U is a question to the first prover. The lemma asserts that over the choice of the question V to the smaller prover, the projection of the event $\mathcal{L}$ to the zoom-in and zoom-out pair $(Q, W[V])$ still has noticeable probability:

Lemma 5.10. *The following holds for any integers $r, s > 0$, ℓ sufficiently large relative to r and s, $k > \ell$, $0 < \beta < q^{-2\ell}$, and $C > 2sq^{-2\ell}$. Let U be a fixed question to the first prover in the outer PCP consisting* *of $3k$ variables in some set of k equations, let $Q \subseteq \mathbb{F}_q^U$ be a zoom-in of dimension r, let $W \subseteq \mathbb{F}_q^U$ be a zoom-out of codimension s containing H_U and Q.*

Also let $\mathcal{L} \subseteq \left(\mathbb{F}_q^U\right)^{2\ell}$ be such that $\mathrm{span}(x) \subseteq W$ *for every $x \in \mathcal{L}$ and $\mathcal{D}'_Q(\mathcal{L}) \geqslant q^{-s(2\ell-r)} \cdot C$. With probability at least β^{s+2} over V (chosen according to the marginal of $\mathcal{A}$ defined above, with parameter β), we have*

$$\Pr_{x'_i \in W[V], w_i \in H_U}\left[x' + w \in \mathcal{L} \mid \mathrm{span}_r(x'+w) = Q\right] \geqslant \frac{C}{4(s+1)2^s q^{3rs}}.$$

In the probability above, the sampling procedure is that of $\mathcal{D}'_Q$ with V fixed and we use x' and w to denote $x' = (x'_1, \ldots, x'_{2\ell}), w = (w_1, \ldots, w_{2\ell})$ respectively.

Proof. For each $0 \leqslant j \leqslant s$, let E_j be the set of $V \subseteq U$ such that $\mathrm{codim}(W[V]) = s - j$, and let $p_j := \Pr_{V \sim \mathcal{A}(\cdot, Q)}[V \in E_j]$, where V is chosen according to the outer PCP. Then $\mathcal{D}'_Q(\mathcal{L}) \geqslant q^{-s(2\ell-r)} \cdot C$ implies that

$$\sum_{j=0}^{s} p_j \cdot \mathop{\mathbb{E}}_{V \sim E_j}\left[\Pr_{x'_i \in W[V], w_i \in H_U}\left[x' + w \in \mathcal{L} \mid \mathrm{span}_r(x'+w) = Q\right]\right] \geqslant q^{-s(2\ell-r)} \cdot C. \tag{10}$$

In the above expectation $V \sim E_j$ means choosing V according to the distribution $\mathcal{A}(\cdot, Q)$ conditioned on $V \in E_j$. For each j and $V \in E_j$, we can bound the inner term using the fact that if $x \in \mathcal{L}$, then $\mathrm{span}(x'+w) \subseteq W$, so as a result of $W \supseteq H_U$, we must have $\mathrm{span}(x') \subseteq W[V]$. Thus,

$$\begin{aligned}
&\Pr_{x'_i \in W[V], w_i \in H_U}\left[x' + w \in \mathcal{L} \mid \mathrm{span}_r(x'+w) = Q\right] \\
&\leqslant \Pr_{x'_i \in \mathbb{F}_q^V, w_i \in H_U}\left[x'_{r+1}, \ldots, x'_{2\ell} \in W[V]\right] \\
&\leqslant q^{-(s-j)(2\ell-r)},
\end{aligned}$$

where in the second inequality we use the fact that $\mathrm{codim}(W[V]) = s - j$. Call j rare if $p_j \leqslant \beta^{j+1}$. The contribution to the sum above from rare j is at most

$$\sum_{j \text{ rare}} \beta^{j+1} \cdot q^{-(s-j)(2\ell-r)} \leqslant s \cdot \beta \cdot q^{-s(2\ell-r)} \leqslant q^{-s(2\ell-r)}\frac{C}{2},$$

where we used $\beta \leqslant q^{-2\ell} \leqslant \frac{C}{2s}$. Removing this contribution from (10) it follows that there exists j that is not rare satisfying

$$p_j \cdot \mathop{\mathbb{E}}_{V \in E_j}\left[\Pr_{x'_i \in W[V], w_i \in H_U}\left[x' + w \in \mathcal{L} \mid \mathrm{span}_r(x'+w) = Q\right]\right] \geqslant q^{-s(2\ell-r)} \cdot \frac{C}{2(s+1)}. \tag{11}$$

Fix this j henceforth. Using again the fact that if $x'+w \in \mathcal{L}$, then $\mathrm{span}(x'+w) \subseteq W$ and $\mathrm{span}(x') \subseteq W[V]$, we get

$$
\begin{aligned}
&p_j \cdot q^{-(s-j)(2\ell-r)} \cdot \mathop{\mathbb{E}}_{V\in E_j}\left[\Pr_{x'_i\in W[V], w_i\in H_U}\left[x'+w\in\mathcal{L} \mid \operatorname{span}_r(x'+w)=Q\right]\right]\\
&= p_j \cdot q^{-(s-j)(2\ell-r)} \cdot \mathop{\mathbb{E}}_{V\in E_j}\left[\Pr_{\substack{x'_i\in\mathbb{F}_q^V, w_i\in H_U,\\ x_i=x'_i+w_i, i\in[2\ell]}}\left[x'+w\in\mathcal{L} \mid \operatorname{span}_r(x)=Q,\ \operatorname{span}(x')\subseteq W[V]\right]\right]\\
&= p_j \cdot \mathop{\mathbb{E}}_{V\in E_j}\left[\Pr_{x'_i\in\mathbb{F}_q^V, w_i\in H_U}\left[x'+w\in\mathcal{L} \mid \operatorname{span}_r(x'+w)=Q\right]\right]\\
&\geqslant q^{-s(2\ell-r)}\cdot\frac{C}{2(s+1)},
\end{aligned}
$$

where the last transition uses (11). Dividing the last inequality by p_j we conclude that

$$
\begin{aligned}
\mathop{\mathbb{E}}_{V\in E_j}\left[\Pr_{x'_i\in W[V], w_i\in H_U}\left[x'+w\in\mathcal{L} \mid \operatorname{span}_r(x'+w)=Q\right]\right] &\geqslant \frac{q^{-j(2\ell-r)}}{p_j}\cdot\frac{C}{2(s+1)}\\
&\geqslant \frac{q^{-j(2\ell-r)}}{(2s\beta q^{3r})^j}\cdot\frac{C}{2(s+1)}\\
&\geqslant \frac{C}{2(s+1)2^s q^{3rs}}.
\end{aligned}
$$

where in the second inequality we apply the upper bound on p_j from Lemma 5.9, and in the last transition we use $\beta \leqslant q^{-2\ell}$. Now let

$$
C' := \frac{C}{2(s+1)2^s q^{3rs}}.
$$

By an averaging argument, conditioned on $V \in E_j$ with probability at least $\frac{C'}{2}$ the probability in the expectation above is at least $\frac{C'}{2}$. Overall, as j is not rare, we get that the probability that the expectation above is at least $\frac{C'}{2}$ over the choice of V is at least $p_j\frac{C'}{2} \geqslant \beta^{s+2}$. □

5.3 The Number of Maximal Zoom-Outs is Bounded

Theorem 5.2 suggests that the two provers can agree on a zoom-in with reasonable probability using their advice. The same cannot be said for zoom-outs however, and to circumvent this issue we must develop further tools. In this section, we define the notion of *maximal* zoom-outs and show that for a fixed zoom-in Q, the number of maximal zoom-outs is bounded.

5.3.1 Generic Sets of Subspaces

To deal with large collections of zoom-outs we define a special property of zoom-outs that is called "genericness". To motivate it, let W_{amb} be an ambient space and suppose $W_1, W_2 \subseteq W_{\mathsf{amb}}$ are distinct subspaces of codimension r. Then $W_1 \cap W_2$ is a subspace whose codimension is between $2r$ and $r+1$. For a typical pair of subspaces, the intersection $W_1 \cap W_2$ has codimension $2r$ though, and in this case we say the two subspaces are generic. Genericness is useful probabilistically because if W_1, W_2 are generic, then the event that a randomly chosen 2ℓ-dimensional subspace is contained in W_1, and the event the subspace is contained in W_2, are almost independent. Below is a formal definition which also applies to a set of subspaces rather than just two:

Definition 5.11. *We say that a set $\mathcal{S} = \{W_1, \dots, W_N\}$ of codimension r subspaces of W_{amb} is t-generic with respect to W_{amb} if for any t distinct subspaces, say $W_{i_1}, \dots, W_{i_t} \in \mathcal{S}$, we have $\operatorname{codim}(\bigcap_{1 \leqslant j \leqslant t} W_{i_j}) = t \cdot r$. When the ambient space W_{amb} is clear from context we simply say that $\mathcal{S}$ is t-generic.*

Throughout this section, it will be convenient to denote by $\operatorname{codim}_V(W)$ the codimension of $W \cap W_{\mathsf{amb}}$ with respect to W_{amb}. When no subscript is used, the identity of the subspace W_{amb} should be clear from context.

We remark that any set of subspaces that is t-generic with respect to W_{amb} is also t'-generic with respect to W_{amb} for any $t' < t$. In this section we show a sunflower-type lemma, stating that any large set of codimension r subspaces inside W_{amb} contains a large set of subspaces that are t-generic with respect to W'_{amb} for some $W'_{\mathsf{amb}} \subseteq W_{\mathsf{amb}}$. Below is a formal statement.

Lemma 5.12. *Let $t, r \in \mathbb{N}$ be integers and let $\mathcal{S}$ be a set of N subspaces each with codimension r inside of W_{amb}. Then there exists a subspace $W'_{\mathsf{amb}} \subseteq W_{\mathsf{amb}}$ and a set of subspaces $\mathcal{S}' \subseteq \mathcal{S}$ such that:*

- $|\mathcal{S}'| \geqslant \frac{N^{\frac{1}{(r+1)\cdot(t-1)!}}}{q^{3r}}$.
- *Each $W_i \in \mathcal{S}'$ is contained in W'_{amb} and has codimension s with respect to W'_{amb}, where $s \leqslant r$.*
- *$\mathcal{S}'$ is t-generic with respect to W'_{amb}.*

To show Lemma 5.12 we need two auxiliary lemmas. The first, Lemma 5.13, states that for $j \geqslant 2$, any j-generic set of subspaces contains a large $(j+1)$-generic set of subspaces. The second, Lemma 5.14, states that a given collection of subspaces is either already 2-generic, or else many subspaces in the collection are contained in a common hyperplane.

Lemma 5.13. *Let $\mathcal{S} = \{W_1, \dots, W_N\}$ be a set of N-subspaces of codimension r inside of W_{amb} that is j-generic with respect to W_{amb} for $j \geqslant 2$, then there is a subset $\{W_1, \dots, W_{N'}\} \subseteq \mathcal{S}$ of size $N' \geqslant \frac{N^{1/j}}{q^r}$ that is $(j+1)$-generic with respect to W_{amb}.*

Proof. Fix any j distinct subspaces in $\mathcal{S}$, say $W_1, \dots, W_j$ and let $W = W_1 \cap \dots \cap W_j$. Since $\mathcal{S}$ is j-generic, $\operatorname{codim}(W) = j \cdot r$. We claim that there are at most $q^{j \cdot r}$ subspaces $W_{i'} \in \mathcal{S} \setminus \{W_1, \dots, W_j\}$ such that $\operatorname{codim}(W \cap W_{i'}) \leqslant (j+1)r - 1$. Call such subspaces bad, and suppose for the sake of contradiction that there are more than q^{jr} bad subspaces $W_{i'}$. Then for each bad $W_{i'}$ we have,

$$\begin{aligned}
\dim(W_{i'} + W) &= \dim(W_{i'}) + \dim(W) - \dim(W_{i'} \cap W) \\
&\leqslant (\dim(W_{\mathsf{amb}}) - r) + (\dim(W_{\mathsf{amb}}) - jr) - (\dim(W_{\mathsf{amb}}) - (j+1)r + 1) \\
&= \dim(W_{\mathsf{amb}}) - 1.
\end{aligned}$$

Therefore, for each $W_{i'}$, the space $W + W_{i'}$ is contained in a hyperplane H such that $H \supseteq W$. There are at most $q^{\operatorname{codim}(W)} - 1 = q^{j \cdot r} - 1$ hyperplanes H containing W, and by the pigeonhole principle it follows that there are two bad subspaces say $W_{i'_1}, W_{i'_2}$ that are both contained in the same hyperplane H. This is a contradiction however, as by the j-genericness of $\mathcal{S}$, we must have

$$\begin{aligned}
\dim(W_{i'_1} + W_{i'_2}) &= \dim(W_{i'_1}) + \dim(W_{i'_2}) - \dim(W_{i'_1} \cap W_{i'_2}) \\
&= 2(\dim(W_{\mathsf{amb}}) - r) - (\dim(W_{\mathsf{amb}}) - 2r) \\
&= \dim(W_{\mathsf{amb}}),
\end{aligned}$$

and hence $W_{i'_1}$ and $W_{i'_2}$ cannot both be contained in the hyperplane H.

The lemma now follows from the claim we have just shown. Construct a subset $\mathcal{S}'$ greedily as follows:

1. Initialize $\mathcal{S}'$ by picking j arbitrary subspaces from $\mathcal{S}$ and inserting them to $\mathcal{S}'$.

2. For any j subspaces in $\mathcal{S}'$, say $W_1, \ldots, W_j$, remove any $W' \in \mathcal{S}$ which is bad for them.

3. If $\mathcal{S}$ is not empty, pick some $W \in \mathcal{S}$, insert it to $\mathcal{S}'$ and iterate.

Note that trivially, the collection $\mathcal{S}'$ will be $(j+1)$-generic in the end of the process. To lower bound the size of $\mathcal{S}'$, note that when $|\mathcal{S}'| = s$, the number of elements from $\mathcal{S}$ that have been deleted is at most $s^j q^{jr}$, and hence so long as this value is less than N, we may do another iteration. Thus, we must have that $s \geqslant \left(\frac{N}{q^{jr}}\right)^{1/j} = \frac{N^{1/j}}{q^r}$ when the process terminates. □

Lemma 5.14. *Let $\{W_1, \ldots, W_N\}$ be a set of subspaces of W_{amb} of codimension r. Then for any integer $m \geqslant 1$, at least one of the following holds.*

- *There are m subspaces, say $W_1, \ldots, W_m$ that are 2-generic with respect to W_{amb}.*
- *There is a subspace $W'_{\mathsf{amb}} \subseteq W_{\mathsf{amb}}$ of codimension 1 that contains $N' \geqslant \frac{N}{mq^r}$ of these subspaces, say $W_1, \ldots, W_{N'}$.*

Proof. Note that for any $1 \leqslant i \neq j \leqslant N$, we have $\operatorname{codim}(W_i \cap W_j) \leqslant 2r$. Consider the graph with vertices $W_1, \ldots, W_N$ and W_i, W_j are adjacent if and only if $i \neq j$ and $\operatorname{codim}(W_i \cap W_j) \leqslant 2r-1$. If every vertex in this graph has degree at most $\lfloor \frac{N}{m} \rfloor$, then we are done as there is an independent set of size m and these subspaces satisfy the first condition. Suppose this is not the case. Then there is a vertex, say W_N, that has at least $\lceil \frac{N}{m} \rceil$ neighbors, say $W_1, \ldots, W_{\lceil \frac{N}{m} \rceil}$. For $1 \leqslant i \leqslant \lceil \frac{N}{m} \rceil$, we have $\operatorname{codim}(W_N \cap W_i) \leqslant 2r-1$, so

$$\dim(W_i + W_N) = \dim(W_i) + \dim(W_N) - \dim(W_i \cap W_N) \leqslant \dim(W_{\mathsf{amb}}) - 1.$$

Thus $W_i + W_N$ is always contained in a codimension 1 subspace of W_{amb} that contains W_N. Since the number of such subspaces is $q^r - 1$, one of these codimension 1 subspaces of W_{amb}, say W'_{amb}, contains at least

$$\frac{\lceil N/m \rceil}{q^r - 1} \geqslant \frac{N}{mq^r}$$

of the subspaces in the list $W_1, \ldots, W_{\lceil \frac{N}{m} \rceil}$. □

Repeatedly applying Lemma 5.14 yields the following corollary.

Corollary 5.15. *Let $\{W_1, \ldots, W_N\}$ be a set of subspaces of W_{amb} of codimension r with respect to W_{amb}. Then, there exists a subspace $W'_{\mathsf{amb}} \subseteq W_{\mathsf{amb}}$, an integer $1 \leqslant s \leqslant r$, and a subset of $m \geqslant \frac{N^{\frac{1}{r+1}}}{q^r}$ subspaces from this set, say $\{W_1, \ldots, W_m\}$, which are all contained in W'_{amb} and satisfy the following two conditions.*

- *Each W_i, $1 \leqslant i \leqslant m$, has codimension s with respect to W'_{amb}.*
- *$\{W_1, \ldots, W_m\}$ is 2-generic with respect to W'_{amb}.*

Proof. To start set $W'_{\mathsf{amb}} = W_{\mathsf{amb}}$. If the W_i's have codimension 1 in W'_{amb} then the result holds.

Otherwise, if the conclusion does not hold, then apply Lemma 5.14 with $m = \frac{N^{\frac{1}{r+1}}}{q^r}$. Either the first condition of Lemma 5.14 holds and we are done, or we can find a new subspace, W''_{amb}, of codimension 1 inside the current W'_{amb} containing at least $\frac{N}{mq^r}$ of the W_i's. Set $W'_{\mathsf{amb}} = W''_{\mathsf{amb}}$ and repeat. Note that the codimension of the W_i's with respect to W'_{amb} drops by 1 after every iteration, so we will repeat at most r times before reaching the desired conclusion. This yields a list of W_i's that satisfy the conditions of size at least

$$\frac{N}{(mq^r)^r} = N^{\frac{1}{r+1}} \geqslant m. \qquad \square$$

Combining Corollary 5.15 and Lemma 5.13, we get Lemma 5.12 as follows:

Proof of Lemma 5.12. By Corollary 5.15, there is a collection $\mathcal{S}'$ of size $|\mathcal{S}'| \geqslant \frac{N^{\frac{1}{r+1}}}{q^r}$ and $W'_{\mathsf{amb}} \subseteq W_{\mathsf{amb}}$ such that $\mathcal{S}'$ is 2-generic with respect to W'_{amb}, and each W_i has (the same) codimension $s \leqslant r$ with respect to W'_{amb}. Applying Lemma 5.13 $t-2$ times, there is a set of t-generic subspaces relative to W'_{amb}, $\mathcal{S}'' \subseteq \mathcal{S}'$, of size

$$\left|\mathcal{S}''\right| \geqslant \left(\left(\left(|\mathcal{S}'|^{\frac{1}{2}} \cdot \frac{1}{q^r}\right)^{\frac{1}{3}} \cdots\right) \cdot \frac{1}{q^r}\right)^{\frac{1}{t-1}} \cdot \frac{1}{q^r} \geqslant \frac{|\mathcal{S}'|^{\frac{1}{(t-1)!}}}{q^{2r}} \geqslant \frac{N^{\frac{1}{(r+1)\cdot(t-1)!}}}{q^{3r}}. \qquad \square$$

In addition to Lemma 5.12, we state another useful feature of generic sets of subspaces, formalized in Lemma 5.16 below. The lemma asserts that if a collection $\{W_1, \ldots, W_N\}$ is generic, and one zooms-outs from the ambient space W_{amb} into a hyperplane H, then one gets an induced collection $\{W_1 \cap H, \ldots, W_N \cap H\}$ which is almost as generic.

Lemma 5.16. *Let $\mathcal{S} = \{W_1, \ldots, W_N\}$ be a collection of subspaces of codimension r that is t-generic with respect to some space W_{amb} for an even integer t, and let H be a hyperplane in W_{amb}. Then the collection of subspaces $\mathcal{S}' = \{W_1 \cap H, \ldots, W_N \cap H\}$ can be made a $\frac{t}{2}$-generic set of subspaces with respect to H with codimension r inside of H by removing at most $\frac{t}{2}$ subspaces $W_i \cap H$ from it.*

Proof. Suppose that $\mathcal{S}'$ is not $\frac{t}{2}$-generic with respect to H with codimension r inside of H, as otherwise we are done. In this case, there must exist $\frac{t}{2}$ distinct subspaces, say $W_1 \cap H, \ldots, W_{\frac{t}{2}} \cap H \in \mathcal{S}'$ such that the codimension of $W_1 \cap \cdots \cap W_{\frac{t}{2}} \cap H$ with respect to H, is less than $tr/2$. As H has codimension 1 with respect to W_{amb} it follows that

$$\operatorname{codim}\left(W_1 \cap \cdots \cap W_{\frac{t}{2}} \cap H\right) \leqslant \frac{t}{2} \cdot r, \tag{12}$$

where now the codimension is with respect to W_{amb}. Since $\mathcal{S}$ is t-generic (and thus $\frac{t}{2}$-generic as well) with respect to W_{amb}, this implies that $W_1 \cap \cdots \cap W_{\frac{t}{2}} \subseteq H$. Indeed, this is because $W_1 \cap \cdots \cap W_{\frac{t}{2}}$ already has codimension $\frac{tr}{2}$, so by (12), further intersecting with H does not result in a smaller subspace, and hence this subspace must be inside H.

Now delete $W_1 \cap H, \ldots, W_{\frac{t}{2}} \cap H$ from $\mathcal{S}'$. We claim that the resulting set is $\frac{t}{2}$-generic with respect to H. Suppose for the sake of contradiction that it is not. Then there must be another $\frac{t}{2}$ distinct subspaces, say $W_{\frac{t}{2}+1} \cap H, \ldots, W_t \cap H \in \mathcal{S}'$ such that $W_{\frac{t}{2}+1} \cap \cdots \cap W_t \subseteq H$, implying

$$\left(W_1 \cap \cdots \cap W_{\frac{t}{2}}\right) + \left(W_{\frac{t}{2}+1} \cap \cdots \cap W_t\right) \subseteq H.$$

This is a contradiction however, as $\mathcal{S}$ is t-generic with respect to W_{amb}, so $\operatorname{codim}(W_1 \cap \cdots \cap W_t) = tr$, and

$$\begin{aligned}
&\dim\left((W_1 \cap \cdots \cap W_{\frac{t}{2}}) + (W_{\frac{t}{2}+1} \cap \cdots \cap W_t)\right) \\
&\qquad = \dim\left(W_1 \cap \cdots \cap W_{\frac{t}{2}}\right) + \dim\left(W_{\frac{t}{2}+1} \cap \cdots \cap W_t\right) - \dim\left(W_1 \cap W_2 \cap \cdots \cap W_t\right) \\
&\qquad = 2\left(\dim(W_{\mathsf{amb}}) - \frac{t}{2}r\right) - \dim(W_{\mathsf{amb}}) + tr = \dim(W_{\mathsf{amb}}) > \dim(H),
\end{aligned}$$

and contradiction. $\square$

The following result is a version of Lemma 5.16 for codimensions larger than 1:

Lemma 5.17. *Let $\mathcal{W}$ be a collection of subspaces of codimension r that is 2^K-generic with respect to W_{amb} and let B be a subspace of codimension j. Then, there is a set of subspaces $\mathcal{W}_B$ such that the following holds*

- *Each $W_B \in \mathcal{W}_B$ is equal to $W \cap B$ for some $W \in \mathcal{W}$.*
- *The set of subspaces $\mathcal{W}_B$ is 2^{K-j}-generic with respect to B and each subspace in $\mathcal{W}_B$ has codimension r.*
- $|\mathcal{W}| - 2^K \leqslant |\mathcal{W}_B| \leqslant |\mathcal{W}|$

Proof. We will obtain $\mathcal{W}_B$ by removing a small number of subspaces from $\{W \cap B \mid W \in \mathcal{W}\}$. There is a sequence of subspaces $W_{\mathsf{amb}} = B_0 \supseteq B_1 \supseteq \cdots \supseteq B_j = B$, such that B_{i+1} is a hyperplane inside of B_i. Do the following,

1. Initialize $\mathcal{W}_0 = \mathcal{W}$ and set $i = 1$.
2. Set $\mathcal{W}_i = \{W \cap B_i \mid W \in \mathcal{W}_{i-1}\}$. Then, using Lemma 5.16, remove at most 2^{K-i-1}-many subspaces to turn $\mathcal{W}_i$ into a 2^{K-i}-generic collection with respect to B_i.
3. Stop if $i = j$, otherwise, increase i by 1 and return to step 2.

We verify the three properties. The first property holds by construction. For the second property, it is clear that the output is a set of subspaces $\mathcal{W}_j \subseteq \mathcal{W}_B$ that is 2^{K-j}-generic with respect to B, and by Lemma 5.16, the codimension of each $W \cap B \in \mathcal{W}_B$ with respect to B is still r. Finally, for the third property, note that during each iteration, at most 2^{K-i-1} subspaces are removed by Lemma 5.16, so altogether we remove at most $\sum_{i=0}^{K-1} 2^i \leqslant 2^K$ subspaces. $\square$

5.3.2 The Sampling Lemma

As explained earlier, the notion of genericness is useful probabilistically, and in this section we state and prove a sampling lemma about generic collections which is necessary for our analysis. We work inside a subspace W_{amb}. Fix an arbitrary zoom-in $Q \subseteq W_{\mathsf{amb}}$ of dimension a, and let $\mathcal{S} = \{W_1, \ldots, W_m\}$ be a 2-generic collection of subspaces of W_{amb} of codimension r all containing Q. Also let $\mathcal{A}$ be a set of j-dimensional subspaces containing Q. Consider the following two probability measures over $\mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]$:

1. The distribution μ which is uniform over $\mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]$.

2. The distribution $\nu_{\mathcal{S}}$, wherein a subspace is sampled by first picking $i \in \{1, \ldots, m\}$ uniformly and then sampling a subspace from $\mathsf{Zoom}_j[Q, W_i]$ uniformly.

Throughout this section one should think of j as much larger than r and a, and $\dim(W_{\mathsf{amb}})$ as much larger than j. Concretely, we require $j \geqslant 2^{100} r^2 a^2$ and $\dim(W_{\mathsf{amb}}) \geqslant 2^{100} j$. The main content of this section is the following lemma, asserting that the measures μ and $\nu_{\mathcal{S}}$ are close in statistical distance provided that m is large. More precisely:

Lemma 5.18. *Let W_{amb} be the ambient space and let a, r, j be dimensions such that $j \geqslant 2^{100r^2a^2}$, $\dim(W_{\mathsf{amb}}) \geqslant 2^{100j}$. Let $Q \subseteq W_{\mathsf{amb}}$ be a subspace of dimension $\dim(Q) = a$, and let $\mathcal{S}$ be a 2-generic collection of subspaces of W_{amb} such that $|\mathcal{S}| = m$, each subspace in $\mathcal{S}$ has codimension r with respect W_{amb}, and each subspace in $\mathcal{S}$ contains Q. Then, for any set of j-dimensional subspaces $\mathcal{L} \subset \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]$, we have*

$$|\nu_{\mathcal{S}}(\mathcal{L}) - \mu(\mathcal{L})| \leqslant \frac{3q^{\frac{r}{2}(j-a)}}{\sqrt{m}}.$$

We now set up some notations for the proof of Lemma 5.18. For $L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]$ let

$$N_{\mathcal{S}}(L) = |\{W_i \in \mathcal{S} \mid L \subseteq W_i\}|, \tag{13}$$

and for an arbitrary pair of distinct $W_i, W_{i'} \in \mathcal{S}$ define the following quantities:

$$\begin{aligned} D &= |\{L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}] \mid L \subseteq W_i\}| \\ p_1 &= \Pr_{L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]}[L \subseteq W_i] \\ p_2 &= \Pr_{L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]}[L \subseteq W_i \cap W_{i'}] \end{aligned} \tag{14}$$

We note that all of these quantities are well defined as they do not depend on the identity of W_i and $W_{i'}$. The following bound on p_1 will be useful in this section.

Lemma 5.19. *Let p_1 be as in* (14) *and suppose W_{amb} and the dimensions a, r, j satisfy $j \geqslant 2^{100r^2a^2}$, $\dim(W_{\mathsf{amb}}) \geqslant 2^{100j}$. Then $|p_1 - q^{-r(j-a)}| \leqslant \frac{q^j}{|W_{\mathsf{amb}}|}$.*

Proof. We have,

$$\begin{aligned} p_1 &= \Pr_{L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]}[L \subseteq W_i] \\ &= \Pr_{v_1, \ldots, v_{j-a} \in V}[v_1, \ldots, v_{j-a} \in W_i \mid \dim(Q + \mathrm{span}(v_1, \ldots, v_{j-a})) = j] \end{aligned}$$

Additionally,

$$\Pr_{v_1, \ldots, v_{j-a} \in W_{\mathsf{amb}}}[v_1, \ldots, v_{j-a} \in W_i] = q^{-r(j-a)},$$

so

$$\left| p_1 - q^{-r(j-a)} \right| \leqslant \Pr_{v_1, \ldots, v_{j-a} \in V}[\dim(Q + \mathrm{span}(v_1, \ldots, v_{j-a})) < j] \leqslant \frac{q^j}{|W_{\mathsf{amb}}|}. \qquad \square$$

Using our notations and $|\mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]| \cdot p_1 = D$ we have that

$$\mu(L) = \frac{1}{|\mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]|} = \frac{p_1}{D}, \qquad \nu_{\mathcal{S}}(L) = \frac{N_{\mathcal{S}}(L)}{m} \cdot \frac{1}{D} = \frac{N_{\mathcal{S}}(L)}{mD}, \tag{15}$$

so comparing μ and $\nu_{\mathcal{S}}$ amounts to comparing p_1 and $N_{\mathcal{S}}/m$. In the following claim we analyze the expectation and variance of $N_{\mathcal{S}}(L)$ when L is chosen uniformly from $\mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]$:

Claim 5.20. *Let $r, a, j, Q, W_{\mathsf{amb}}, N_{\mathcal{S}}$ be as defined above. We have,*

$$\mathop{\mathbb{E}}_{L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]}[N_{\mathcal{S}}(L)] = p_1 m \quad \textit{and} \quad \mathrm{var}(N_{\mathcal{S}}(L)) \leqslant p_1 m,$$

where the variance is over uniform $L \in \mathsf{Zoom}_j[Q, W_{amb}]$.

Proof. By linearity of expectation

$$\mathop{\mathbb{E}}_{L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]}[N_{\mathcal{S}}(L)] = \sum_{i=1}^{m} \Pr_{L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]}[L \subseteq W_i] = p_1 m,$$

and we move on to the variance analysis. To bound $\mathbb{E}_{L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]}[N_{\mathcal{S}}(L)^2]$, write

$$\begin{aligned}
\mathop{\mathbb{E}}_{L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]}[N_{\mathcal{S}}(L)^2] &= \mathop{\mathbb{E}}_{L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]}\left[\left(\sum_{i=1}^{m} \mathbb{1}_{L \subseteq W_i}\right)^2\right] \\
&\leqslant m \cdot \Pr_{L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]}[L \subseteq W_i] + m^2 \cdot \Pr_{L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]}[L \subseteq W_i \cap W_{i'}] \\
&= p_1 m + p_2 m^2.
\end{aligned}$$

It follows that,

$$\mathrm{var}(N_{\mathcal{S}}(L)) \leqslant p_1 m + p_2 m^2 - p_1^2 m^2.$$

Finally note that by the 2-genericness p_2 and p_1^2 are nearly the same value, and it can be checked using (14) that $p_2 \leqslant p_1^2$ and so $\mathrm{var}(N_{\mathcal{S}}(L)) \leqslant p_1 m$. $\square$

Combining Chebyshev's inequality with Claim 5.20, we conclude the following lemma:

Lemma 5.21. *For any $c > 0$ it holds that*

$$\Pr_{L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]}\Big[|N_{\mathcal{S}}(L) - p_1 m| \geqslant c \cdot p_1 m\Big] \leqslant \frac{1}{c^2 p_1 m} \leqslant \frac{1.01 \cdot q^{r(j-a)}}{c^2 m},$$

where recall the chosen L has dimension j, $\dim(W_{\mathsf{amb}}) \geqslant 2^{100} q^j$, $\mathcal{S}$ is a 2-generic set of m codimension r subspaces all containing a fixed subspace of dimension a, and

$$p_1 = \Pr_{L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]}[L \subseteq W_i] \in \left[q^{-r(j-a)} - \frac{q^j}{|W_{\mathsf{amb}}|}, q^{-r(j-a)} + \frac{q^j}{|W_{\mathsf{amb}}|}\right]$$

and

$$\mathop{\mathbb{E}}_{L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]}[N_{\mathcal{S}}(L)] = p_1 m.$$

In words, p_1 the probability that a random $L \in \mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]$ is contained in some fixed codimension r, W_i subspace of W_{amb}.

Proof. The first inequality is an immediate result of Chebyshev's inequality with the bounds from Claim 5.20. To get the second inequality, we apply the lower bound on p_1 from Lemma 5.19 to get

$$p_1 \geqslant q^{-r(j-a)} - \frac{q^j}{|W_{\mathsf{amb}}|} \geqslant \frac{1}{1.01q^{r(j-a)}}.$$

□

Lastly, we use Claim 5.20 to prove Lemma 5.18.

Proof of Lemma 5.18. We have,

$$\begin{aligned}|\mu(\mathcal{L}) - \nu_{\mathcal{S}}(\mathcal{L})| = \frac{1}{mD}\left|\sum_{L\in\mathcal{L}} N_{\mathcal{S}}(L) - p_1 m\right| &\leqslant \frac{1}{mD}\sum_{L\in\mathcal{L}} |N_{\mathcal{S}}(L) - p_1 m| \\ &= \frac{|\mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]|}{mD} \cdot \mathop{\mathbb{E}}_{L\in\mathsf{Zoom}_j[Q,W_{\mathsf{amb}}]}[|N_{\mathcal{S}}(L) - p_1 m|],\end{aligned}$$

and by Cauchy-Schwartz we get that

$$|\mu(\mathcal{L}) - \nu_{\mathcal{S}}(\mathcal{L})| \leqslant \frac{|\mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]|}{mD}\sqrt{\mathrm{var}(N_{\mathcal{S}}(L))}.$$

Plugging in Claim 5.20 and using $|\mathsf{Zoom}_j[Q, W_{\mathsf{amb}}]| = D/p_1$ we get that

$$|\mu(\mathcal{L}) - \nu_{\mathcal{S}}(\mathcal{L})| \leqslant \frac{1}{\sqrt{p_1 m}} \leqslant \frac{3q^{\frac{r}{2}(j-a)}}{\sqrt{m}}.$$

In the second inequality, we apply the lower bound on p_1 from Lemma 5.19 to get

$$p_1 \geqslant q^{-r(j-a)} - \frac{q^j}{|W_{\mathsf{amb}}|} \geqslant \frac{1}{9q^{r(j-a)}}.$$

□

5.3.3 An Upper Bound on the Number of Maximal Zoom-outs

For this subsection, we work in the second prover's space, $\mathbb{F}_q^V$, and make the assumption that $|V| \gg \ell$, say $|V| \geqslant 2^{100}q^{\ell}$ to be concrete. We first establish several results in the simplified setting where there is no zoom-in. After that we show how to deduce an analogous result with a zoom-in. Throughout this section, we fix T to be a table that assigns, to each $L \in \mathsf{Grass}(\mathbb{F}_q^V, 2\ell)$, a linear function on L.

Definition 5.22. *Given a table T on $\mathsf{Grass}(\mathbb{F}_q^V, 2\ell)$ and a subspace $Q \subseteq \mathbb{F}_q^V$, we call a zoom-out, function pair, (W, g_W), where $Q \subseteq W \subseteq \mathbb{F}_q^V$ and $g_W : W \to \mathbb{F}_q$, (C, ζ)-maximal with respect to T on Q if*

$$\Pr_{L\in\mathsf{Grass}(\mathbb{F}_q^V, 2\ell)}[g_W|_L \equiv T[L] \mid Q \subseteq L \subseteq W] \geqslant C,$$

and there does not exist another zoom-out function pair, $(W', g_{W'})$ such that $\mathbb{F}_q^V \supseteq W' \supsetneq W$, $g_{W'} : W' \to \mathbb{F}_q$, $g_{W'}|_W \equiv g|_W$ and

$$\Pr_{L\in\mathsf{Grass}(\mathbb{F}_q^V, 2\ell)}[g_{W'}|_L \equiv T[L] \mid Q \subseteq L \subseteq W'] \geqslant \zeta C.$$

In the case that $Q = \{0\}$, we say that (W, g_W) is (C, ζ)-maximal with respect to T.

In the above statement, C should be thought of as small and $\zeta \in (0,1]$ should be thought of as an absolute constant. With this in mind, one should think of a zoom-out function pair (W, g_W) as being (C,ζ) maximal if the agreement between g_W and the table T on entries in $\mathsf{Zoom}_{2\ell}[Q,W]$ is not "explained" by the agreement of a larger zoom-out function pair, $(W', g_{W'})$, where $W' \supseteq W$ and $g_{W'}|_W = g_W$. By definition every zoom-out function pair (W, f), where f has non-trivial agreement with T inside W, is in some sense contained in a maximal zoom-out function pair (W', f') (possibly with slightly lower agreement). This point is encapsulated by the following lemma.

Lemma 5.23. *Let T be a table on $\mathsf{Grass}(\mathbb{F}_q^V, 2\ell)$, $Q \subseteq \mathbb{F}_q^V$, and $W \subseteq \mathbb{F}_q^V$ be a subspace of codimension r containing Q. Suppose that there exists a linear function $g_W : W \to \mathbb{F}_q$ such that*

$$\Pr_{L \in \mathsf{Grass}(\mathbb{F}_q^V, 2\ell)}[g_W|_L \equiv T[L] \mid Q \subseteq L \subseteq W] \geqslant C.$$

Then there exists a subspace $W' \supseteq W$ and a linear function $g_{W'} : W' \to \mathbb{F}_q$ such that $g_{W'}|_W \equiv g_W$ and $(g_{W'}, W')$ is $(C\zeta^{-r}, \zeta)$-maximal.

Proof. This is an immediate consequence of Definition 5.22. If (W, g_W) is (C,ζ)-maximal then we are done. Otherwise, there must exist W_1, g_{W_1} such that $W_1 \supsetneq W$, $g_{W_1}|_W \equiv g_W$ and

$$\Pr_{L \in \mathsf{Grass}(\mathbb{F}_q^V, 2\ell)}[g_{W_1}|_L \equiv T[L] \mid Q \subseteq L \subseteq W_1] \geqslant \zeta C.$$

We can repeat this argument at most r times before obtaining some $(g_{W'}, W')$ that is $(C\zeta^{-r}, \zeta)$ maximal and satisfies $W' \supseteq W$ and $g_{W'}|_W \equiv g_W$. □

Next, we will want to upper bound the number of maximal zoom-outs. The following lemma (which is a formal version of Lemma 1.7 from the introduction) is the key component in this upper bound. We set the following parameters for the remainder of the section, which can all be considered constant:

$$\xi = \delta^5, \quad \delta_2 = \frac{\xi}{100}, \quad t = \left(2^{2+10/\delta_2}\right)! \,.$$

Recall that the parameter $\delta > 0$ is an arbitrary (small) constant, which we may fix freely.

Lemma 5.24. *Fix $\delta > 0$, a finite field $\mathbb{F}_q$, and set ξ, δ_2, t as above. Then the following is true for sufficiently large ℓ. Let T be a table on $\mathsf{Grass}(\mathbb{F}_q^V, 2\ell)$ and set $r \leqslant \frac{10}{\delta}$, $C \geqslant q^{-2(1-\delta^5)\ell}$, and*

$$N \geqslant q^{100(t-1)! r^2 \ell \xi^{-1}}.$$

Suppose that $(W_1, f_1), \ldots, (W_N, f_N)$ are zoom-out, function pairs such that the W_i's are all distinct and of codimension r, and for each $i \in [N]$

$$\Pr_{L \in \mathsf{Grass}(\mathbb{F}_q^V, 2\ell)}[f_i|_L \equiv T[L] \mid L \subseteq W_i] \geqslant C.$$

Then there is a subspace $A \subseteq \mathbb{F}_q^V$, a linear function $h' : A \to \mathbb{F}_q$, and a set of subspaces $\mathcal{W}' \subseteq \{W_1, \ldots, W_N\}$ of size at least $q^{50r\ell\xi^{-1}}$ which satisfy the following,

- *Each $W_i \in \mathcal{W}'$ is strictly contained in A and has codimension $r' < r$ with respect to A.*

- $\mathcal{W}'$ *is* 2*-generic with respect to* A.

- *For any* $W_i \in \mathcal{W}'$, $h'|_{W_i} \equiv f_i$.

Proof. The proof is deferred to Section 8. $\square$

To upper bound the number of maximal zoom-outs we also need the following list decoding property.

Lemma 5.25. *Let T be a table on $\mathsf{Grass}(\mathbb{F}_q^V, 2\ell)$, let Q be an r_1-dimensional subspace, and let $W \supseteq Q$ be a subspace of codimension r_2. Suppose that 2ℓ is sufficiently large and $\dim(W) \geqslant 20\ell$. Let $f_1, \ldots, f_m$ be a list of distinct linear functions such that $f_i|_L \equiv T[L]$ for at least β-fraction of the 2ℓ-dimensional subspaces L such that $Q \subseteq L \subseteq W$, for $\beta \geqslant 2q^{-2\ell+r_1} + c$, and $c > 0$. Then, $m \leqslant \frac{4}{c^2}$ and in particular if $\beta > 4q^{-2\ell+r_1}$, then $m \leqslant \frac{16}{\beta^2}$.*

Proof. The proof is deferred to Appendix D. $\square$

Combining Lemmas 5.24 and 5.25 yields an upper bound on the number of (C, ζ) maximal zoom-out function pairs with respect to a table T on Q.

Theorem 5.26. *Fix $\delta, \mathbb{F}_q, \xi, \delta_2, t, \ell$ as in Lemma 5.24, and suppose $|V| \geqslant q^\ell$. Then, the following is true. For any table T on $\mathsf{Grass}(\mathbb{F}_q^V, 2\ell)$ such that $|V| \geqslant q^\ell$, any subspace $Q \subseteq \mathbb{F}_q^V$ of dimension $r_1 \leqslant \frac{10}{\delta}$ and any $C \geqslant q^{-2(1-\delta^3)\ell}$, the number of $(C, \frac{1}{5})$-maximal zoom-out, function pairs with respect to T on Q of codimension at most $\frac{10}{\delta}$ is at most $\frac{160}{\delta} \cdot C^{-2} \cdot q^{100(t-1)!(10/\delta)^2\ell\xi^{-1}}$.*

Proof. Suppose for the sake of contradiction that $(W_1, f_1), \ldots, (W_M, f_M)$ are

$$M > \frac{160}{\delta} C^{-2} q^{100(t-1)!(10/\delta)^2\ell\xi^{-1}}$$

distinct pairs that are $(C, \frac{1}{5})$ maximal with respect to T on Q. By Lemma 5.25, for each W_i, there are at most $16C^{-2}$ functions $f : W_i \to \mathbb{F}_q$ such that $f|_L \equiv T[L]$ for at least C-fraction of the $L \in \mathsf{Zoom}_{2\ell}[Q, W]$. Thus, there are $C^2 M/4$ distinct W_i's appearing in the pairs, and there is a codimension $r_2 \leqslant \frac{10}{\delta}$ such there are $\frac{M}{\frac{160}{\delta}C^{-2}} = N \geqslant q^{100(t-1)!(10/\delta)^2\ell\xi^{-1}}$ pairs, say, $(W_1, f_1), \ldots, (W_N, f_N)$ zoom-out function pairs that are $(C, \frac{1}{5})$ maximal with respect to T on Q such that the W_i's are all distinct and of codimension r_2 in $\mathbb{F}_q^V$.

Write $\mathbb{F}_q^V = Q \oplus A$. For each $L \subseteq \mathbb{F}_q^V$ of dimension 2ℓ containing Q, there is a unique $L' \subseteq A$ such that $L = Q \oplus L'$. Define the table T' that assigns linear functions to each $L' \in \mathsf{Grass}(A, 2\ell - \dim(Q))$ by

$$T'[L'] \equiv T[L' \oplus Q]|_{L'}. \tag{16}$$

For each $1 \leqslant i \leqslant N$, let $W_i' \subseteq A$ be the unique subspace such that $W_i = W_i' \oplus Q$. We have that $f_i|_{L'} = T'[L']$ for at least C-fraction of $L' \in \mathsf{Grass}(W_i', 2\ell - \dim(Q))$.

By Lemma 5.24 there exists a subspace W'_{amb}, a linear function $h' : W'_{\mathsf{amb}} \to \mathbb{F}_q$, and a subcollection $\mathcal{W}' \subseteq \{W_1', \ldots, W_m'\}$ of size at least $m \geqslant q^{50(10/\delta)\ell\xi^{-1}}$ such that

- Each $W_i' \in \mathcal{W}'$ has codimension r' where $1 \leqslant r' \leqslant r_2$ with respect to W'_{amb}.

- $\mathcal{W}'$ is 2-generic with respect to W'_{amb}.

- For any $W_i' \in \mathcal{W}$, $h'|_{W_i'} = f_i|_{W_i'}$.

Let $\mathcal{W} = \{W_i \mid W'_i \in \mathcal{W}'\}$ and linearly extend h' to the function $h^\star$ on $V^\star = W'_{\mathsf{amb}} \oplus Q$ so that, $h^\star|_{W_i} \equiv f_i$ for at least q^{-r_1} of the W_i in $\mathcal{W}$. Note that such an $h^\star$ exists because a random linear extension of h' satisfies this in expectation. It follows that there is a set $\mathcal{V} = \{W_i \in \mathcal{W} \mid h^\star|_{W_i} \equiv f_i\}$ of size $|\mathcal{V}| \geqslant mq^{-r_1} \geqslant mq^{-\frac{10}{\delta}}$.

Furthermore, because $\mathcal{W}'$ is 2-generic inside of W'_{amb}, $\mathcal{W}$ is 2-generic inside of $V^\star$. We will now finish the proof by applying Lemma 5.18 on $\mathcal{V}$. Specifically, let ν denote the measure over $\mathsf{Zoom}_{2\ell}[Q, V^\star]$ generated by choosing $W_i \in \mathcal{V}$ and then $L \in \mathsf{Zoom}_{2\ell}[Q, V^\star]$ conditioned on $L \subseteq W_i$, let μ denote the uniform measure over $\mathsf{Zoom}_{2\ell}[Q, V^\star]$, and let

$$\mathcal{L} = \{L \in \mathsf{Zoom}_{2\ell}[Q, V^\star] \mid h^\star|_L \equiv T[L]\}.$$

Since $h^\star|_{W_i} \equiv f_i$ for every $W_i \in \mathcal{V}$, we have

$$\nu(\mathcal{L}) = \mathop{\mathbb{E}}_{W_i \in \mathcal{V}}\left[\Pr_{L \in \mathsf{Zoom}_{2\ell}[Q, W_i]}[f_i|_L \equiv T[L]]\right] \geqslant C.$$

By Lemma 5.18 with $a = r_1, r = r_2$, and $j = 2\ell$, it follows that

$$\mu(\mathcal{L}) \geqslant C - \frac{3q^{\frac{r_2}{2}(2\ell - r_1)}}{\sqrt{m \cdot q^{-r_1}}} \geqslant \frac{C}{5}.$$

We get that the zoom-out function pair $(V^\star, h^\star)$ satisfies that $V^\star \supsetneq W_i$ and $h^\star|_{W_i} \equiv f_i$ for at least one i, and $\Pr_{L \in \mathsf{Grass}(\mathbb{F}_q^V, 2\ell)}[h^\star|_L \equiv T[L] \mid Q \subseteq L \subseteq V^\star] \geqslant \frac{C}{5}$. This contradicts the assumption that (W_i, f_i) is $(C, \frac{1}{5})$-maximal with respect to T on Q. $\square$

6 Analysis of the PCP

In this section we analyze the PCP construction Ψ from Section 4. As usual, the completeness analysis is straightforward and the soundness analysis will comprise the bulk of our effort.

6.1 Completeness

Suppose that the $\mathsf{3Lin}$ instance (X, Eq) has an assignment $\sigma : X \to \mathbb{F}_q$ that satisfies at least $1 - \varepsilon_1$ of the equations in Eq. Let $\mathcal{U}_{\mathsf{sat}} \subseteq \mathcal{U}$ be the set of all $U = (e_1, \ldots, e_k)$ where all k equations $e_1, \ldots, e_k$ are satisfied. Then, $|\mathcal{U}_{\mathsf{sat}}| \geqslant (1 - k\varepsilon_1)|\mathcal{U}|$. We identify σ with the linear function from $\mathbb{F}_q^X \to \mathbb{F}_q$, assigning the value $\sigma(i)$ to the ith elementary basis element e_i. Abusing notation, we denote this linear map by σ as well.

For each $U \in \mathcal{U}_{\mathsf{sat}}$ and vertex $L \oplus H_U$, we set $T_1[L \oplus H_U] \equiv \sigma|_{L \oplus H_U}$. Since $U \in \mathcal{U}_{\mathsf{sat}}$, these assignments satisfy the side conditions. For all other U's, set $T_1[L \oplus H_U]$ so that the side conditions of H_U are satisfied and $T_1[L \oplus H_U]|_L \equiv \sigma|_L$. Such an assignment is possible because $L \cap H_U = \{0\}$. Similarly, the table T_2 is defined as $T_2[R] \equiv \sigma|_R$.

Sampling a constraint, note that the constraint is satisfied whenever the $L' \oplus H_{U'}$ chosen in step 3 of the test satisfies that $U' \in \mathcal{U}_{\mathsf{sat}}$. As the marginal distribution of $L' \oplus H_{U'}$ is uniform,[9] the distribution of U' is uniform. It follows that the constraint is satisfied whenever $U' \in \mathcal{U}_{\mathsf{sat}}$, which happens with probability at least $1 - k\varepsilon_1$. Thus, $\mathsf{val}(\Psi) \geqslant 1 - k\varepsilon_1$.

[9]This is true because first a clique is chosen with probability that is proportional to its size and then a vertex is sampled uniformly from the clique.

6.2 Soundness

In this section we relate the soundness of the composed PCP to that of the outer PCP. More precisely, we show:

Lemma 6.1. *For all $\delta > 0$ there are $r \in \mathbb{N}$ and $\zeta(\delta) > 0$ such that the following holds for sufficiently large $\ell \in \mathbb{N}$, sufficiently small $c \in (0,1)$, and*

$$k = q^{2(1+c)\ell}, \quad \beta = q^{-2(1+2c/3)\ell},$$

as in (7).

Let $G_{\beta,r}^{\otimes k}$ be the parallel repetition of the Smooth Variable versus Equation Game with advice described in Section 3.1.4, and let Ψ be the composed PCP described in Section 4 using the sufficiently large ℓ above. If $\mathsf{val}(G_{\beta,r}^{\otimes k}) < q^{-\zeta(\delta)\cdot\ell^2}$, then $\mathsf{val}(\Psi) \leqslant 64q^{-2(1-1000\delta)\ell}$.

The rest of this section is devoted to the proof of Lemma 6.1, and it heavily relies on the tools from Section 5. We will show the contrapositive of Lemma 6.1. That is, we will show that if there are tables T_1 and T_2 that are ε-consistent for $\varepsilon > 64q^{-2\ell(1-1000\delta)}$, then this implies strategies for the two provers that have success probability at least $q^{-\zeta(\delta)\cdot O(\ell^2)}$.

6.2.1 Clique Consistency

To start, we will reduce to the case where T_1 satisfies a condition called *clique-consistency*.

Definition 6.2. *We say an assignment T to $\mathcal{A}$ is clique consistent if for every vertex $L_1 \oplus H_{U_1}$ and for every $L_2 \oplus H_{U_2}, L_3 \oplus H_{U_3} \in [L_1 \oplus H_{U_1}]$, the assignments $T[L_2 \oplus H_{U_2}]$ and $T[L_3 \oplus H_{U_3}]$ satisfy the 1-to-1 constraint between $L_2 \oplus H_{U_2}$ and $L_3 \oplus H_{U_3}$ specified in Lemma 4.2.*

The following lemma shows that if T_1 and T_2 are ε-consistent assignments to Ψ, then there is a clique-consistent assignment T_1' such that T_1' and T_2 are ε-consistent.

Lemma 6.3. *Suppose that the assignments T_1 and T_2 are ε-consistent, then there is a clique-consistent assignment T_1' such that T_1' and T_2 are ε-consistent.*

Proof. Partition $\mathcal{A}$ into cliques, $\mathcal{A} = \mathsf{Clique}_1 \sqcup \cdots \sqcup \mathsf{Clique}_m$. For each i, choose a random $L \oplus H_U \in \mathsf{Clique}_i$ uniformly, and for every $L' \oplus H_{U'} \in \mathsf{Clique}_i$ assign $T_1'[L' \oplus H_{U'}]$ in the unique way that is consistent with $T_1[L \oplus H_U]$ and the side conditions of U' as described in Lemma 4.2. It is clear that T_1' is clique consistent, and we next analyze the expected fraction of constraints that T_1' and T_2 satisfy.

Note that an alternative description of sampling a constraint in Ψ proceeds as follows. First choose a clique Clique_i with probability that is proportional to by its size, and then choose $L \oplus H_U \in \mathsf{Clique}_i$ in the first step. The rest of the sampling procedure is the same. Let $P(L \oplus H_U)$ be the probability that the test passes conditioned on $L \oplus H_U$ being chosen in the second step. It is clear that every vertex in the clique has equal probability of being chosen, therefore the probability of passing if Clique_i is chosen is

$$\frac{1}{|\mathsf{Clique}_i|} \sum_{L \oplus H_U \in \mathsf{Clique}_i} P(L \oplus H_U).$$

On the other hand, the expected fraction of constraints satisfied by T_1' and T_2 (over the randomness of choosing T_1') is

$$\sum_{L \oplus H_U \in \mathsf{Clique}_i} \frac{1}{|\mathsf{Clique}_i|} \cdot P(L \oplus H_U) = \frac{1}{|\mathsf{Clique}_i|} \sum_{L \oplus H_U \in \mathsf{Clique}_i} P(L \oplus H_U).$$

To see this, note that for any $L \oplus H_U$, $\frac{1}{|\mathsf{Clique}_i|}$ is the probability that $T_1[L \oplus H_U]$ is used to define T_1' on Clique_i. If this is the case, then the probability the test passes on T_1' within Clique_i is $P(L \oplus H_U)$.

Since this holds over every clique, it follows that the expected fraction of constraints satisfied by T_1' equals the fraction of constraints satisfied by T_1 and T_2. In particular, there is a choice of T_1' such that together with T_2 it satisfies at least ε fraction of the constraints. $\square$

Applying Lemma 6.3 we conclude that there are clique-consistent assignments to Ψ that are ε-consistent, and henceforth we assume that T_1 is clique-consistent to begin with. We remark that, in the notation of Section 4.1.4, the benefit of having a clique-consistent assignment is that the constraint that the verifier checks is equivalent to checking that $T_1[L \oplus H_U]|_R \equiv T_2[R]$. The latter check is a test which is performed within the space $\mathbb{F}_q^U$ of the first prover. We will use this fact in the next section.

6.2.2 A Strategy for the First Prover

Let $p(U)$ be the pass probability of tables T_1 and T_2 conditioned on U being the question to the first prover. That is, $p(U)$ is the probability that a constraint, as described in Section 4.1.4, is satisfied, conditioned on U being chosen in step 1 there. As we are assuming that the overall success probability is at least ε, $\mathbb{E}_U[p(U)] \geqslant \varepsilon$. By an averaging argument, $p(U) \geqslant \frac{\varepsilon}{2}$ for at least $\frac{\varepsilon}{2}$-fraction of the U's. Call such U's good and let $\mathcal{U}_{\mathsf{good}}$ be the set of good U's.

Let $U \in \mathcal{U}$ be the question to the first prover and let Q be the advice. If $U \notin \mathcal{U}_{\mathsf{good}}$, then the first prover gives up, so henceforth assume that $U \in \mathcal{U}_{\mathsf{good}}$. For such U, the test of the inner PCP passes with probability at least $\frac{\varepsilon}{2}$. More concretely, we have

$$\Pr_{\substack{L \in \mathsf{Grass}_q(\mathbb{F}_q^U, 2\ell), L \cap H_U = \{0\} \\ \dim(R) = 2(1-\delta)\ell}} [T_1[L \oplus H_U]|_R \equiv T_2[R] \mid R \subseteq L] \geqslant \frac{\varepsilon}{2}.$$

Next, the first prover chooses an integer $0 \leqslant r' \leqslant \frac{10}{\delta}$ uniformly, and takes Q to be the span of the first r'-advice vectors (so that r in Lemma 6.1 is taken to be $10/\delta$). By Theorem 5.3, there are r_1, r_2 satisfying $r_1 + r_2 \leqslant \frac{10}{\delta}$ such that for at least $q^{-6\ell^2}$ of the $Q \subseteq \mathbb{F}_q^U$ of dimension r_1, there exists $W_Q \subseteq \mathbb{F}_q^U$ containing $Q \oplus H_U$ of codimension $r_2 \leqslant \frac{10}{\delta}$ and a linear function $g_{Q,W_Q} : W_Q \to \mathbb{F}_q$ that satisfies the side conditions on H_U and

$$\Pr_{L \in \mathsf{Zoom}_{2\ell}[Q, W_Q], L \cap H_U = \{0\}} [g_{Q,W_Q}|_{L \oplus H_U} \equiv T_1[L \oplus H_U]] \geqslant \frac{q^{-2(1-1000\delta^2)\ell}}{2}. \tag{17}$$

Henceforth, set $C := \frac{q^{-2(1-1000\delta^2)\ell}}{2}$. With probability at least $\frac{1}{1+10/\delta}$, the first prover chooses $r' = r_1$, where r_1 is the parameter from Theorem 5.3. Say that a dimension r_1 subspace Q is lucky if there exist W_Q and g_{Q,W_Q} as above satisfying (17) and let $\mathcal{Q}_{\mathsf{lucky}}$ be the set of all lucky $Q \subseteq \mathbb{F}_q^U$.

Throughout the remainder of the section, for each r_1-dimensional Q that satisfies $Q \in \mathcal{Q}_{\mathsf{lucky}}$ and $Q \cap H_U = \{0\}$, we will write W_Q and g_{Q,W_Q} to be the zoom-out and function described that satisfy (17). Additionally, we will require $W_Q = W_{Q'}$ and $g_{Q,W_Q} \equiv g_{Q',W_{Q'}}$ for any two valid Q, Q' such that $Q \oplus H_U = Q' \oplus H_U$. First let us see why this is possible. It is straightforward to verify that $Q \cap H_U = \{0\}$ if and only if $Q' \cap H_U = \{0\}$. Next, note that for any $W \subseteq \mathbb{F}_q^U$ we have $W \supseteq Q \oplus H_U$ if and only if $W \supseteq Q' \oplus H_U$. Moreover, for any function $g : W \to \mathbb{F}_q$ agreeing with the side conditions on H_U, we have

$$\Pr_{L \in \mathsf{Zoom}_{2\ell}[Q, W], L \cap H_U = \{0\}} [g|_{L \oplus H_U} \equiv T_1[L \oplus H_U]] = \Pr_{L \in \mathsf{Zoom}_{2\ell}[Q', W], L \cap H_U = \{0\}} [g|_{L \oplus H_U} \equiv T_1[L \oplus H_U]], \tag{18}$$

because in each probability, the distribution over $L \oplus H_U$ is the same. It follows that $Q \in \mathcal{Q}_{\mathsf{lucky}}$ if and only if $Q' \in \mathcal{Q}_{\mathsf{lucky}}$. Then, by (18) we can enforce that the first prover's zoom-out and function satisfy $W_Q = W_{Q'}$ and $g_{Q,W_Q} \equiv g_{Q',W_{Q'}}$.

Moving on, for each Q such that $Q \in \mathcal{Q}_{\mathsf{lucky}}$ and $Q \cap H_U = \{0\}$, let us define

$$\mathcal{L}_Q = \{L \in \mathsf{Zoom}_{2\ell}[Q, W_Q] \mid g_{Q,W_Q}|_L \equiv T_1[L \oplus H_U]|_L\}.$$

For our analysis, we will only consider the case where the first prover chooses $r' = r_1$ and $Q \in \mathcal{Q}_{\mathsf{lucky}}$. If $Q \notin \mathcal{Q}_{\mathsf{lucky}}$ or $Q \cap H_U \neq \{0\}$ we simply define $\mathcal{L}_Q = \emptyset$, but again we will not worry about such Q's in the analysis.

Finally, define

$$\mathcal{L} = \{(x_1, \ldots, x_{2\ell}) \in \left(\mathbb{F}_q^U\right)^{2\ell} \mid \operatorname{span}(x_1, \ldots, x_{r_1}) \in \mathcal{Q}_{\mathsf{lucky}}, \operatorname{span}(x_1, \ldots, x_{2\ell}) \in \mathcal{L}_{\operatorname{span}(x_1,\ldots,x_{r_1})}\},$$

and let $\mathcal{Q}_{\mathsf{smooth}}$ denote the set of r_1-dimensional subspaces Q such that

$$\mathcal{D}'_Q(\mathcal{L}) \geqslant 0.8\mathcal{D}_Q(\mathcal{L}) - \eta^{20}, \tag{19}$$

where $\eta = q^{-100\ell^{100}}$ is the negligible constant appearing in Section 5.2.

If either $Q \notin \mathcal{Q}_{\mathsf{lucky}}$, $Q \cap H_U \neq \{0\}$, or $Q \notin \mathcal{Q}_{\mathsf{smooth}}$, then the first prover gives up. Otherwise the prover extends the function g_{Q,W_Q} to a linear function on the entire space $\mathbb{F}_q^U$ randomly, and we denote this extension by $g : \mathbb{F}_q^U \to \mathbb{F}_q$. The prover outputs the string $s_{Q,U}$ as their answer where $s_{Q,U} \in \mathbb{F}_q^U$ is the unique string such that $g(x) = \langle s_{Q,U}, x\rangle$ for all $x \in \mathbb{F}_q^U$. As g_{Q,W_Q}, and by extension g, respects the side conditions, it follows that $s_{Q,U}$ satisfies the k-linear equations of U.

At this point, we have fully described the first prover's strategy. With the strategy fresh in mind, it will be helpful to now compute some probabilities related to this strategy. Specifically, let us lower bound $\mathcal{D}_Q(\mathcal{L})$ for the Q such that $Q \in \mathcal{Q}_{\mathsf{lucky}}$ and $Q \cap H_U = \{0\}$. If Q is additionally smooth, then we also get a lower bound on $\mathcal{D}'_Q(\mathcal{L})$ by (19). By definition of $\mathcal{D}_Q$ and $\mathcal{L}$, if $Q \in \mathcal{Q}_{\mathsf{lucky}}$ and $Q \cap H_U = \{0\}$, we have

$$\begin{aligned}
\mathcal{D}_Q(\mathcal{L}) &= \Pr_{x=(x_1,\ldots,x_{2\ell})\in\left(\mathbb{F}_q^U\right)^{2\ell}}[x \in \mathcal{L} \mid \operatorname{span}_{r_1}(x) = Q] \\
&= \Pr_{x\in\left(\mathbb{F}_q^U\right)^{2\ell}}[g_{Q,W_Q}|_{\operatorname{span}(x)} \equiv T_1[\operatorname{span}(x) \oplus H_U]|_{\operatorname{span}(x)} \wedge \operatorname{span}(x) \subseteq W_Q \mid \operatorname{span}_{r_1}(x) = Q] \\
&= \Pr_{L\in\mathsf{Grass}_q(\mathbb{F}_q^U,2\ell), L\cap H_U=\{0\}}[g_{Q,W_Q}|_{L\oplus H_U} \equiv T_1[L \oplus H_U] \wedge L \subseteq W_Q \mid Q \subseteq L] \\
&\quad - \Pr_{x\in\left(\mathbb{F}_q^U\right)^{2\ell}}[\dim(\operatorname{span}(x)) < 2\ell \vee \operatorname{span}(x) \cap H_U \neq \{0\} \mid \operatorname{span}_{r_1}(x) = Q] \\
&\geqslant \Pr_{L\in\mathsf{Grass}_q(\mathbb{F}_q^U,2\ell), L\cap H_U=\{0\}}[g_{Q,W_Q}|_{L\oplus H_U} \equiv T_1[L \oplus H_U] \wedge L \subseteq W_Q \mid Q \subseteq L] - q^{2\ell-3k} - q^{2\ell-2k}.
\end{aligned}$$

where the second transition is because, by definition, every $L \in \mathcal{L}_Q$ is contained in W_Q. Recalling that $k \gg \ell$, one should think of the last two terms in the last line as negligible.

Continuing, for $Q \in \mathcal{Q}_{\mathsf{lucky}}$, we have

$$\begin{aligned}
\mathcal{D}_Q(\mathcal{L}) \geqslant &\Pr_{L\in\mathsf{Grass}_q(\mathbb{F}_q^U,2\ell), L\cap H_U=\{0\}}[g_{Q,W_Q}|_{L\oplus H_U} \equiv T_1[L \oplus H_U] \mid Q \subseteq L \subseteq W_Q] \\
&\cdot \Pr_{L\in\mathsf{Grass}_q(\mathbb{F}_q^U,2\ell), L\cap H_U=\{0\}}[L \subseteq W_Q \mid Q \subseteq L] - q^{2\ell-3k} - q^{2\ell-2k} \\
\geqslant\ & q^{-r_2(2\ell-r_1)} \cdot C - q^{2\ell-3k} - q^{2\ell-2k},
\end{aligned} \tag{20}$$

where in the third transition uses (17) to lower bound the first term by C.

6.2.3 A Strategy for the Second Prover

Let V be the question to the second prover. The second prover will use a table T_1' to derive their strategy. The table T_1' is obtained from T_1 as follows. For a question V to the second prover, let $U' \supseteq V$ be an arbitrary question to the first prover. For all 2ℓ-dimensional subspaces $L \subseteq \mathbb{F}_q^V$, define

$$T_1'[L] \equiv T_1[L \oplus H_{U'}]|_L.$$

If $L \cap H_{U'} \neq \{0\}$, then we arbitrarily define $T'[L]$ to be the 0 function. Note that provided $|V| \geqslant 2k$, this will only be the case for a negligible, $q^{2\ell-k}$, fraction of subspaces L.

Using clique consistency: we first argue that T_1' is well defined, and for that we make several observations. First, recall that the subspace $L \oplus H_{U'}$ is indeed a subspace of $\mathbb{F}_q^{U'}$ because we have $L \subseteq \mathbb{F}_q^V \subseteq \mathbb{F}_q^{U'}$, where $\mathbb{F}_q^V$ is viewed as the subspace of $\mathbb{F}_q^{U'}$ where the restriction to coordinates in $U' \setminus V$ is 0. Next, note that all choices of $U' \supseteq V$ lead to the same value of $T_1'[L]$. Indeed, for a fixed L, the vertices $L \oplus H_{U'}$ over all $U' \supseteq V$ are in the same clique, so since T_1 is clique consistent they all receive a consistent value with regards to the mapping in Lemma 4.2 and hence lead to the same function $T_1[L \oplus H_{U'}]|_L$. Therefore the second prover can construct the table T_1', and we can assume without loss of generality that the U' that was chosen is the same question that the first prover received, i.e. $U' = U$.

After constructing T_1', the second prover then chooses a dimension $0 \leqslant r' \leqslant \frac{10}{\delta}$ uniformly for the advice Q. That is, the set Q to be the span of the first r' advice vectors that they receive. Note that with probability at least $\frac{1}{10/\delta+1}$ the second prover also chooses $r' = r_1$ (that is, they choose the same dimension as the first prover for the advice). The second prover then uniformly chooses a zoom-out function pair $(W_{\mathsf{second}}, g_{Q,W_{\mathsf{second}}})$ that is

$$\left(\frac{C}{4 \cdot 5^{10/\delta}}, \frac{1}{5}\right)\text{-maximal}$$

with respect to T_1' on Q (the second prover gives up if such pair does not exist).

Finally, the second prover extends the function $g_{Q,W_{\mathsf{second}}}$ randomly to a linear function on $\mathbb{F}_q^V$ to arrive at their answer. The resulting function is linear and it is equal to the inner product function $y \to \langle s_{Q,V}, y\rangle$ for some unique string $s_{Q,V} \in \mathbb{F}_q^V$. The second prover outputs $s_{Q,V}$ as their answer.

6.2.4 The Success Probability of the Provers

In order to be successful, a series of events must occur. We go through each one and state the probability that each occurs. At the end this yields a lower bound on the provers' success probability. We remark that the analysis of this sections requires Lemmas 5.5 and 5.7, so recall that k and β are set according to (7) in Section 5.2 so that these lemmas hold.

First, the provers need $U \in \mathcal{U}_{\mathsf{good}}$, which occurs with probability at least $\frac{\varepsilon}{2}$. Assuming that this occurs, the provers then both need to choose $r' = r_1$ for the dimension of their zoom-in, which happens with probability at least $\frac{\delta^2}{101}$. If both provers do choose $r' = r_1$, then both prover's set Q to be the span of the first r_1 advice vectors that they receive and in particular, receive the same zoom-in Q as advice. Going forward, we will only analyze the provers' success probability assuming that this is the case.

The provers then need $Q \in \mathcal{Q}_{\mathsf{lucky}}$, $Q \cap H_U = \{0\}$, and $Q \in \mathcal{Q}_{\mathsf{smooth}}$. When analyzing the probability that these three events occur, we need to recall that the advice vectors are actually drawn uniformly from the second prover's space $\mathbb{F}_q^V$, rather than $\mathbb{F}_q^U$. To this end, let us define the following coupled distribution.

The Advice Distribution, $\mathcal{D}_{\mathsf{adv}}$:

- Choose $V \subseteq U$ according to the outer PCP.
- Choose $x_1, \ldots, x_{r_1} \in \mathbb{F}_q^V$ uniformly.
- Choose $w_1, \ldots, w_{r_1} \in H_U$ uniformly.
- Set $Q = \mathrm{span}(x_1, \ldots, x_{r_1})$ and $Q' = \mathrm{span}(x_1 + w_1, \ldots, x_{r_1} + w_{r_1})$.

Thus, the subspace Q is actually drawn according to the first marginal of $\mathcal{D}_{\mathsf{adv}}$, which we denote $\mathcal{D}^1_{\mathsf{adv}}$, while the second marginal of $\mathcal{D}_{\mathsf{adv}}$ is the distribution $\mathcal{D}'_{r_1}$ from Section 5.2.3.

The following claim will help us move from statements about Q' to statements about Q. It is helpful, because the covering property from Section 5.2.3 will allow us to bound probabilities about Q', while Claim 6.4 will allow us to translate these into statements about Q.

Claim 6.4. *Let (Q, Q') be output by $\mathcal{D}_{\mathsf{adv}}$. Then the following two statements are true.*

- *If $Q' \in \mathcal{Q}_{\mathsf{lucky}}$, $Q' \cap H_U = \{0\}$, then $Q \in \mathcal{Q}_{\mathsf{lucky}}$, $Q \cap H_U = \{0\}$.*
- *$\mathcal{D}_{Q'}(\mathcal{L}) = \mathcal{D}_Q(\mathcal{L})$ and $\mathcal{D}'_{Q'}(\mathcal{L}) = \mathcal{D}'_Q(\mathcal{L})$, and as a consequence if $Q' \in \mathcal{Q}_{\mathsf{smooth}}$ then $Q \in \mathcal{Q}_{\mathsf{smooth}}$.*

Proof. Let (Q, Q') be as described so that $Q = \mathrm{span}(x_1, \ldots, x_{r_1})$ and $Q' = \mathrm{span}(x_1 + w_1, \ldots, x_{r_1} + w_{r_1})$ for some $x_i \in \mathbb{F}_q^U$ and $w_i \in H_U$.

We start with the first item. Suppose $Q' \in \mathcal{Q}_{\mathsf{lucky}}$, $Q' \cap H_U = \{0\}$. To see that $Q \in \mathcal{Q}_{\mathsf{lucky}}$, note that $Q \oplus H_U = Q' \oplus H_U$, so as explained earlier we have $W_{Q'} = W_Q$, and $g_{Q,W_Q} \equiv g_{Q',W_{Q'}}$, so the result follows from (18). It is also easy to see that $Q \cap H_U = \{0\}$.

We move onto the second item of showing $\mathcal{D}_{Q'}(\mathcal{L}) = \mathcal{D}_Q(\mathcal{L})$ and $\mathcal{D}'_{Q'}(\mathcal{L}) = \mathcal{D}'_Q(\mathcal{L})$. We will first show $\mathcal{D}'_Q(\mathcal{L}) = \mathcal{D}'_{Q'}(\mathcal{L})$ and then describe how the other equality can be shown similarly. Let $A : Q' \to Q$ be the linear transformation that maps $x_i + w_i$ to x_i. Abusing notation, for $x = (x_1, \ldots, x_{2\ell}) \in \mathbb{F}_q^U$ satisfying $\mathrm{span}_{r_1}(x) = Q$, let us write $A(x) = (A(x_1), \ldots, A(x_{r_1}), x_{r_1+1}, \ldots, x_{2\ell})$. Note that A is a bijection between the following two sets

$$\mathcal{A}_1 = \{x \in \mathbb{F}_q^U \mid \mathrm{span}_{r_1}(x) = Q'\} \quad \text{and} \quad \mathcal{A}_2 = \{y \in \mathbb{F}_q^U \mid \mathrm{span}_{r_1}(y) = Q\}.$$

Then,

$$\mathcal{D}'_{Q'}(\mathcal{L}) = \frac{\sum_{x \in \mathcal{L} \cap \mathcal{A}_1} \mathcal{D}'(x)}{\sum_{x \in \mathcal{A}_1} \mathcal{D}'(x)} \quad \text{and} \quad \mathcal{D}'_Q(\mathcal{L}) = \frac{\sum_{y \in \mathcal{L} \cap \mathcal{A}_2} \mathcal{D}'(y)}{\sum_{x \in \mathcal{A}_2} \mathcal{D}'(y)} \tag{21}$$

Now we claim that the map A is a measure preserving (under measure $\mathcal{D}'$) bijection between $\mathcal{A}_1$ and $\mathcal{A}_2$ and $\mathcal{L} \cap \mathcal{A}_1$ and $\mathcal{L} \cap \mathcal{A}_2$.

First note that for any $x \in \mathcal{A}_1$, we have $A(x) - x \in (H_U)^{2\ell}$, so it is straightforward to verify that $\mathcal{D}'(x) = \mathcal{D}'(A(x))$. It follows that A is a bijection between $\mathcal{A}_1$ and $\mathcal{A}_2$ which preserves measure under $\mathcal{D}'$, and consequently the denominators in (21) are the same.

For the numerators we have noted that if $x \in \mathcal{L} \cap \mathcal{A}_1$ then $\mathcal{D}'(x) = \mathcal{D}'(A(x))$, and it remains to show that $x \in \mathcal{L} \cap \mathcal{A}_1$ if and only if $A(x) \in \mathcal{L} \cap \mathcal{A}_2$. To this end observe that if $x \in \mathcal{L} \cap \mathcal{A}_1$ then $\mathrm{span}_{r_1}(x) = Q'$ and

$$T_1[\mathrm{span}(x) \oplus H_U]|_{\mathrm{span}(x)} \equiv g_{Q',W_{Q'}}|_{\mathrm{span}(x)}.$$

Then, note that $\text{span}(A(x)) \oplus H_U = \text{span}(x) \oplus H_U$, $W_Q = W_{Q'}$, and $g_{Q,W_Q} \equiv g_{Q',W_{Q'}}$. Moreover, g_{Q,W_Q} agrees with the side conditions on H_U, so the above implies that $g_{Q,W_Q}|_{\text{span}(x)\oplus H_U} \equiv T_1[\text{span}(x) \oplus H_U]$, and consequently,

$$T_1[\text{span}(x) \oplus H_U]|_{\text{span}(A(x))} \equiv g_{Q',W_{Q'}}|_{\text{span}(A(x))}.$$

It follows that $A(x) \in \mathcal{L} \cap \mathcal{A}_2$ as desired. The other direction follows similarly. Since $x \in \mathcal{L} \cap \mathcal{A}_1$ if and only if $A(x) \in \mathcal{L} \cap \mathcal{A}_2$ and the map A is also a bijection from $\mathcal{L} \cap \mathcal{A}_1$ to $\mathcal{L} \cap \mathcal{A}_2$ which preserves measure under $\mathcal{D}'$. Thus, the numerators in (21) are the same and we can conclude that $\mathcal{D}'_{Q'}(\mathcal{L}) = \mathcal{D}'_Q(\mathcal{L})$.

The proof that $\mathcal{D}_{Q'}(\mathcal{L}) = \mathcal{D}_Q(\mathcal{L})$ proceeds similarly, except that x and $A(x)$ trivially have the same measure under $\mathcal{D}$, because $\mathcal{D}$ is the uniform distribution over $\mathbb{F}_q^U$.

We conclude that $\mathcal{D}_{Q'}(\mathcal{L}) = \mathcal{D}_Q(\mathcal{L})$ and $\mathcal{D}'_{Q'}(\mathcal{L}) = \mathcal{D}'_Q(\mathcal{L})$, so by the definition of $\mathcal{Q}_{\mathsf{smooth}}$ we get that if $Q' \in \mathcal{Q}_{\mathsf{smooth}}$, then $Q \in \mathcal{Q}_{\mathsf{smooth}}$. □

With Claim 6.4 in hand, we may instead lower bound the probability that $Q' \in \mathcal{Q}_{\mathsf{lucky}}$, $Q' \cap H_U = \{0\}$, and $Q' \in \mathcal{Q}_{\mathsf{smooth}}$ under $\mathcal{D}'_{r_1}$. To this end, we start with the probability that these events occur under $\mathcal{D}_{r_1}$, which is simply the uniform distribution over $\mathbb{F}_q^U$.

By Theorem 5.3, the first item occurs with probability at least $q^{-6\ell^2}$. On the other hand the probability that the second item does not occur is at most $\sum_{i=0}^{r_1} \frac{q^i q^k}{q^{3k}} \leqslant q^{r_1+1-2k}$, while the probability that the third item does not occur is at most $3\eta^{20}$ by Lemma 5.6. Altogether we get that with probability at least

$$q^{-6\ell^2} - q^{r_1+1-2k} - 3\eta^{20} \geqslant q^{-7\ell^2}$$

under $\mathcal{D}_{r_1}$, we have $Q' \in \mathcal{Q}_{\mathsf{lucky}}$, $Q' \cap H_U = \{0\}$, and $Q' \in \mathcal{Q}_{\mathsf{smooth}}$. By Lemma 5.7, we have that $Q' \in \mathcal{Q}_{\mathsf{lucky}}$, $Q' \cap H_U = \{0\}$, and $Q' \in \mathcal{Q}_{\mathsf{smooth}}$ with probability at least $q^{-8\ell^2}$ under $\mathcal{D}'_{r_1}$.

Applying Claim 6.4 now yields that with probability at least $q^{-8\ell^2}$ over $(Q, Q') \sim \mathcal{D}_{\mathsf{adv}}$, we have $Q \in \mathcal{Q}_{\mathsf{lucky}}$, $Q \cap H_U = \{0\}$, and $Q \in \mathcal{Q}_{\mathsf{smooth}}$ as well. Here, the subspace Q is finally distributed according to the correct distribution. Towards our soundness analysis, this means that with probability at least $q^{-8\ell^2}$, both provers receive an advice subspace Q which satisfies $Q \in \mathcal{Q}_{\mathsf{lucky}}$, $Q \cap H_U = \{0\}$, and $Q \in \mathcal{Q}_{\mathsf{smooth}}$, and we fix such a Q henceforth.

Recall that W_Q and g_{Q,W_Q} are the zoom-out and function associated with Q. Let $r_2 \leqslant \frac{10}{\delta}$ be the codimension of W_Q. By (20) and the fact that Q is smooth, we get

$$\mathcal{D}'_Q(\mathcal{L}) \geqslant q^{-r_2(2\ell - r_1)} \cdot \frac{C}{2}. \tag{22}$$

Note that by definition of $\mathcal{L}$ and $\mathcal{D}'_Q$, the only $x \in \mathcal{L}$ which contribute to the measure $\mathcal{D}'_Q(\mathcal{L})$ are those such that $\text{span}_{r_1}(x) = Q$ and $\text{span}(x) \subseteq W_Q$. It will be convenient to let $\mathcal{L}'_Q$ be precisely these tuples, and to have the following definition for $\mathcal{L}'_Q$ on hand:

$$\begin{aligned} \mathcal{L}'_Q &= \{x \in \mathcal{L} \mid \text{span}_{r_1}(x) = Q,\ \text{span}(x) \subseteq W_Q\} \\ &= \{x \in \left(\mathbb{F}_q^U\right)^{2\ell} \mid \text{span}_{r_1}(x) = Q,\ g_{Q,W_Q}|_{\text{span}(x)} \equiv T_1[\text{span}(x) \oplus H_U]|_{\text{span}(x)}\}. \end{aligned} \tag{23}$$

Note that in the definition of $\mathcal{L}'_Q$, it is implied that $\dim(\text{span}(x)) = 2\ell$ as otherwise the entry $T_1[\text{span}(x) \oplus H_U]$ is not defined.

By our previous discussion, we have

$$\mathcal{D}'_Q(\mathcal{L}) = \mathcal{D}'_Q(\mathcal{L}'_Q) \geqslant q^{-r_2(2\ell - r_1)} \cdot \frac{C}{2}. \tag{24}$$

Now, for a randomly chosen question to the second prover, $V \subseteq U$, conditioned on the advice being Q, we would like to bound the probability that the first prover's zoom-out function pair, (g_{Q,W_Q}, W_Q), still has good agreement in the second prover's table. If this is the case, then g_{Q,W_Q} remains a candidate function for the second prover, giving the prover's a chance at winning the game. Let $W_Q[V] = W_Q \cap \mathbb{F}_q^V$, and consider $V \subseteq U$ chosen conditioned on the advice subspace being Q. Note that this is precisely the marginal distribution of V in $\mathcal{A}(\cdot, Q)$ considered in Section 5.2.4. Moreover, $Q, \mathcal{L}'_Q$, and W_Q satisfy the setting of Lemma 5.10. Hence we can apply Lemma 5.10, which along with (24) yields that with probability at least $\beta^{r_2+2} \geqslant \beta^{10/\delta+2}$ over the second prover's question V conditioned on $Q \subseteq \mathbb{F}_q^V$, we have:

$$\Pr_{\substack{x'_i \in W_Q[V], w_i \in H_U, \\ x_i = x'_i + w_i, \forall i \in [2\ell]}} \left[x \in \mathcal{L}'_Q \mid \operatorname{span}_{r_1}(x) = Q\right] \geqslant \frac{C}{\frac{80}{\delta} \cdot 2^{10/\delta} q^{30/\delta}} := C'. \tag{25}$$

We call such V consistent, so that V is consistent with probability at least $\beta^{10/\delta+2}$. Now we will show that if V is consistent, the zoom-out $W_Q[V]$ is a candidate for the second prover, in the sense that, the g_{Q,W_Q} also agrees with the second prover's table T'_1, for a non-trivial fraction of $L \in \mathsf{Zoom}_{2\ell}[Q, W_Q[V]]$.

Fix a consistent V, so that (25) holds. Let us now bound the fraction of agreement that g_{Q,W_Q} has with the table T'_1 inside of $\mathsf{Zoom}_{2\ell}[Q, W_Q[V]]$, by relating this quantity to the probability from (25). We have:

$$\begin{aligned}
&\Pr_{L' \in \mathsf{Grass}_q(\mathbb{F}_q^V, 2\ell)} [T'_1[L'] \equiv g_{Q,W_Q}|_{L'} \mid Q \subseteq L' \subseteq W_Q[V]] \\
&\qquad \geqslant \Pr_{x'_i \in W_Q[V], L' = \operatorname{span}(x')} [T'_1[L'] \equiv g_{Q,W_Q}|_{L'} \mid Q \subseteq L'] - q^{2\ell-k} \\
&\qquad \geqslant \Pr_{x'_i \in W_Q[V], L' = \operatorname{span}(x')} [T_1[L' \oplus H_U]|_{L'} \equiv g_{Q,W_Q}|_{L'} \mid Q \subseteq L'] - q^{2\ell-2k} - q^{2\ell-k} \qquad (26) \\
&\qquad = \Pr_{x'_i \in W_Q[V], L' = \operatorname{span}(x')} [T_1[L' \oplus H_U]|_{L'} \equiv g_{Q,W_Q}|_{L'} \mid \operatorname{span}_{r_1}(x') = Q] \\
&\qquad\qquad - q^{2\ell-2k} - q^{2\ell-k}
\end{aligned}$$

We note that in the events of interest for the probabilities above, it is implied that the event of interest only occurs if $\dim(\operatorname{span}(x')) = 2\ell$ as otherwise the table entries for T'_1 and T_1 are not defined. In the first transition, we are using the fact that the x'_i are not linearly independent with probability at most $q^{2\ell-k}$. In the second transition we are using the fact that $L' \cap H_U \neq \{0\}$ with probability at most $q^{2\ell-k}$, and that $T'_1[L'] = T_1[L' \oplus H_U]|_{L'}$. In the third transition we are using the fact that the distribution over L' in the probabilities of the third and fourth lines are the same.

Continuing, the probability from the last line is equal to

$$
\begin{aligned}
&\Pr_{x_i' \in W_Q[V], L'=\mathrm{span}(x')}[T_1[L' \oplus H_U]|_{L'} \equiv g_{Q,W_Q}|_{L'} \mid \mathrm{span}_{r_1}(x') = Q] \\
&= \Pr_{\substack{x_i' \in W_Q[V], w_i \in H_U \\ x_i = x_i' + w_i, L = \mathrm{span}(x)}} [T_1[L \oplus H_U]|_L \equiv g_{Q,W_Q}|_L \mid \mathrm{span}_{r_1}(x') = Q] \\
&= \Pr_{\substack{x_i' \in W_Q[V], w_i \in H_U \\ x_i = x_i' + w_i, L = \mathrm{span}(x)}} [T_1[L \oplus H_U]|_L \equiv g_{Q,W_Q}|_L \mid \mathrm{span}_{r_1}(x') \oplus H_U = Q \oplus H_U] \\
&= \Pr_{\substack{x_i' \in W_Q[V], w_i \in H_U \\ x_i = x_i' + w_i, L = \mathrm{span}(x)}} [T_1[L \oplus H_U]|_L \equiv g_{Q,W_Q}|_L \mid \mathrm{span}_{r_1}(x) \oplus H_U = Q \oplus H_U] \\
&= \Pr_{\substack{x_i' \in W_Q[V], w_i \in H_U \\ x_i = x_i' + w_i, L = \mathrm{span}(x)}} [T_1[L \oplus H_U]|_L \equiv g_{Q,W_Q}|_L \mid \mathrm{span}_{r_1}(x) = Q] \\
&= \Pr_{\substack{x_i' \in W_Q[V], w_i \in H_U \\ x_i = x_i' + w_i, L = \mathrm{span}(x)}} [x \in \mathcal{L}_Q' \mid \mathrm{span}_{r_1}(x) = Q].
\end{aligned}
\tag{27}
$$

The first transition is because the distribution over $L' \oplus H_U$ in the first probability is the same as the distribution over $L \oplus H_U$ in the second probability, and the event is determined by $L \oplus H_U$. The reason is that g_{Q,W_Q} agrees with the side conditions on H_U, so $T_1[L' \oplus H_U]|_{L'} \equiv g_{Q,W_Q}|_{L'}$ if and only if $T_1[L' \oplus H_U]|_{L' \oplus H_U} \equiv g_{Q,W_Q}|_{L' \oplus H_U}$, and $T_1[L \oplus H_U]|_L \equiv g_{Q,W_Q}|_L$ if and only if $T_1[L \oplus H_U]|_{L \oplus H_U} \equiv g_{Q,W_Q}|_{L \oplus H_U}$. The second transition is because in both the second and third probabilities, the subspace $L \oplus H_U$ is a uniformly random $(2\ell + \dim(H_U))$-dimensional subspace of $W_Q[V] \oplus H_U$ containing $Q \oplus H_U$, and again in both cases the event is determined by $L \oplus H_U$. The third transition is because the conditioning in the two probabilities are the same. The fourth transition follows the same reasoning as the second transition, and the final transition is by definition of $\mathcal{L}_Q'$.

Combining (25), (26) and (27) we get

$$
\Pr_{L' \in \mathsf{Grass}_q(\mathbb{F}_q^V, 2\ell)}[T_1'[L'] \equiv g_{Q,W_Q}|_{L'} \mid Q \subseteq L' \subseteq W_Q[V]] \geqslant C' - q^{2\ell - 2k} - q^{2\ell - k} \geqslant \frac{C'}{8}.
$$

By Lemma 5.23, there exists some $(W_Q'[V], g_{Q,W_Q'[V]})$ that is $\left(\frac{C'}{8 \cdot 5^{r_2}}, \frac{1}{5}\right)$-maximal and satisfies $W_Q'[V] \supseteq W_Q[V]$, $g_{Q,W_Q'[V]} : W_Q'[V] \to \mathbb{F}_q$ is linear, and $g_{Q,W_Q'[V]}|_{W_Q[V]} = g_{Q,W_Q}|_{W_Q[V]}$. Furthermore, $\frac{C'}{8 \cdot 5^{r_2}} \geqslant \frac{C'}{8 \cdot 5^{10/\delta}}$, so this zoom-out function pair is also $\left(\frac{C'}{8 \cdot 5^{10/\delta}}, \frac{1}{5}\right)$-maximal and can potentially be chosen by the second prover. Applying Theorem 5.26, the number of $\left(\frac{C'}{8 \cdot 5^{10/\delta}}, \frac{1}{5}\right)$-maximal zoom-out function pairs containing Q that the second prover chooses from is at most

$$
M = \frac{2560}{\delta} \cdot 5^{20/\delta} \cdot \left(C'\right)^{-2} \cdot q^{100(t-1)!(10/\delta)^2 \ell \xi^{-1}} \leqslant q^{\zeta(\delta)\ell},
$$

where $\zeta(\delta)$ is some function depending only on δ. Thus, the second prover chooses $(W_Q'[V], g_{Q,W_Q'[V]})$ with probability at least $\frac{1}{M}$. Finally, if the second prover chooses $(W_Q'[V], g_{Q,W_Q'[V]})$, then the provers succeed if both provers extend their functions, $g_{Q,W_Q}|_{W[V]}$ and $g_{Q,W_Q'[V]}$ in the same manner. This occurs with probability at least $q^{-\operatorname{codim}(W[V])} \geqslant q^{-10/\delta}$.

Putting everything together, we get that the provers succeed with probability at least

$$\frac{\varepsilon}{2} \cdot \frac{\delta^2}{101} \cdot q^{-8\ell^2} \cdot \frac{C}{5} \cdot \frac{1}{M} \cdot q^{-10/\delta} = q^{-\zeta(\delta)\cdot \ell^2},$$

where the first term is the probability that $U \in \mathcal{U}_{\mathsf{good}}$, the second term is the probability that both provers choose the same zoom-in dimension, the third term is the probability that $Q \in \mathcal{Q}_{\mathsf{lucky}}$, $Q \cap H_U = \{0\}$, $Q \in \mathcal{Q}_{\mathsf{smooth}}$, the fourth term is the probability that V is consistent, the fifth term is the probability that the second prover chooses a function that extends g_{Q,W_Q}, and the final term is the probability that both provers extend their functions in the same manner. This proves Lemma 6.1.

7 Proofs of the Main Theorems

7.1 Proof of Theorem 1.3

Theorem 1.3 follows by applying our PCP construction from Section 4 starting with an instance of $\mathsf{Gap3Lin}$ from Theorem 2.6. Fix any τ, ε, which will parameterize our target completeness and soundness in Theorem 1.3 and let s be the value from Theorem 2.6.

We apply the construction of Section 4 with parameters as follows. Set $q = 2$, set some sufficiently small $\delta > 0$ with respect to ε, set $r = \lceil 10/\delta \rceil$, set ℓ sufficiently large relative to δ^{-1} and s, set c sufficiently small relative to δ, and set

$$k = q^{2(1+c)\ell} \quad \text{and} \quad q^{-2(1+2c/3)\ell}.$$

Apply this construction to a $\mathsf{Gap3Lin}[1-\eta, s]$ where the completeness value is set with $\eta \leqslant \tau/k$. The completeness and soundness of the final PCP are as follows.

If the original $\mathsf{3Lin}$ instance is at least $1-\eta$ satisfiable, our final PCP satisfies $\mathsf{val}(\Psi) \geqslant 1 - k\eta \geqslant 1-\tau$. This gives the desired completeness.

On the other hand, if the original instance is at most s satisfiable for some constant $s > 0$, then by Claim 3.2, the value of the outer PCP is at most

$$\mathsf{val}(G_{\beta,r}^{\otimes k}) \leqslant 2^{-\Omega\left((1-s)^2 q^{-r+\frac{2c}{3}\ell}\right)} < q^{-\zeta(\delta) O(\ell)^2},$$

since we take ℓ sufficiently large compared to δ^{-1}. By Lemma 6.1 it follows that if the original instance is at most s satisfiable, then $\mathsf{val}(\Psi) \leqslant 64 q^{-2(1-1000\delta)\ell}$. The proof is concluded as the alphabet size of Ψ is $O(q^{2\ell})$.

7.2 Proof of Theorem 1.4

To show quasi-NP-hardness for approximate Quadratic Programming, we rely on the following result due to [ABH+05], who show a reduction from 2-Prover-1-Round Games to Quadratic Programming.

Theorem 7.1. *There is a reduction from a 2-Prover-1-Round Games, Ψ with graph $G = (L \cup R, E)$ and alphabets Σ_L, Σ_R to a Quadratic Programming instance A such that:*

- *The running time of the reduction and the number of variables in A are both polynomial in $|L| + |R|$ and $2^{|\Sigma_L|}$.*
- *If $\mathsf{val}(\Psi) \geqslant 1-\eta$, then $\mathsf{OPT}(A) \geqslant 1 - \eta - \frac{1}{|L|+|R|}$.*

- *If* $\mathsf{val}(\Psi) \leqslant \varepsilon$*, then* $\mathsf{OPT}(A) \leqslant O(\varepsilon)$.

We are now ready to prove Theorem 1.4.

Proof of Theorem 1.4. Starting with a SAT instance of size n, we take the instance of Gap3Lin from Theorem 2.7 of size $N \leqslant 2^{O(\log^2 n)}$ with field size $q = 2$ as the starting point of our reduction. Take $\delta > 0$ to be a small constant, $r = \lceil 10/\delta \rceil$, $k = (\log n)^C$ for a large constant C and pick c sufficiently small relative to δ and ℓ and β correspondingly so that (7) holds.

This yields a $2^{O(k \log^2 n)}$-time reduction from SAT to a 2-Prover-1-Game on $G = (L \cup R, E)$, with alphabets Σ_L, Σ_R and the following properties:

- $|R| + |L| = O(N^k \cdot q^{3k+2\ell})$.
- $|\Sigma_R| \leqslant |\Sigma_L| = q^{2\ell} = k^{1/(1+c)}$.
- The completeness is at least $1 - k\eta$, where $\eta = 2^{-\Theta(\sqrt{\log n})}$.
- The soundness is at most $q^{-2(1-1000\delta)\ell}$.

Indeed, the first 3 properties are clear. For the soundness, as the original 3Lin instance is at most $1 - \varepsilon$ satisfiable for $\varepsilon = \Omega(1/\log^3 N)$, we get from Claim 3.2 that

$$\mathsf{val}\left(G_{\beta,r}^{\otimes k}\right) \leqslant 2^{-\Omega\left(\varepsilon^{-2} q^{-r+\frac{2c\ell}{3}}\right)} \leqslant q^{-\zeta(\delta) O(\ell)^2}.$$

Then the soundness of the composed PCP is at most $q^{-2(1-1000\delta)\ell}$ by Lemma 6.1. Applying the reduction of Theorem 7.1, we get a reduction from SAT to a Quadratic Programming instance A such that,

- The running time of the reduction and number of variables in A are both polynomial in

$$M = \text{poly}\left(2^{O(\log^2 n) q^{2(1+c)\ell}} (\log n)^{O(q^{2(1+c)\ell})} 2^{q^{2\ell}}\right).$$

- If the original SAT instance is satisfiable, then

$$\mathsf{OPT} \geqslant 1 - 2^{-\Omega\left(\sqrt{\log n}\right)}.$$

- If the original SAT instance is not satisfiable, then

$$\mathsf{OPT} \leqslant O\left(q^{-2(1-1000\delta)\ell}\right).$$

Note that

$$\log(M) = q^{2(1+c)\ell} O\left(\log^2 n\right),$$

whereas the gap between the satisfiable and unsatisfiable cases is $\Omega\left(q^{-2(1-1000\delta)\ell}\right) = \frac{1}{\log(M)^{1-O(\delta)}}$. Altogether, this shows that for all $\varepsilon > 0$ there is $C > 0$ such that unless $\mathsf{NP} \subseteq \mathsf{DTIME}\left(2^{\log(n)^C}\right)$, there is no $\log(M)^{1-\varepsilon}$-approximation algorithm for Quadratic Programming on M variables. $\square$

7.3 Proof of Theorem 1.5

In this section we prove Theorem 1.5, and for that we must first establish a version of Theorem 1.3 for biregular graphs of bounded degree. The proof of this requires minor modifications of our construction, as well as the right degree reduction technique of Moshkovitz and Raz [MR10].

7.3.1 Obtaining a Hard Instance of Bipartite Biregular 2-CSP

We first show that the 2-Prover-1-Round game from Theorem 1.3 can be transformed into a hard instance of biregular, bipartite 2-CSP with bounded degrees. This version may be useful for future applications, and is formally stated below. Call a bipartite 2-CSP (d_1, d_2)-regular if the left degrees of its underlying graph are all d_1, and the right degrees of its underlying graph are all d_2. Throughout this section, for a 2-CSP, Ψ, and an assignment to it, F, we will use $\mathsf{val}(F)$ to denote the fraction of constraints in Ψ that F satisfies.

Theorem 7.2. *For every $\varphi, \varepsilon > 0$, and sufficiently large $R \in \mathbb{N}$, there exist $d_1, d_2 \in \mathbb{N}$ such that given a bipartite (d_1, d_2)-regular 2-CSP, Ψ, with alphabet size R, it is NP-hard to distinguish the following two cases:*

- *Completeness:* $\mathsf{val}(\Psi) \geqslant 1 - \varphi$,
- *Soundness:* $\mathsf{val}(\Psi) \leqslant \frac{1}{R^{1-\varepsilon}}$.

To prove Theorem 7.2, we start with an instance Ψ from Theorem 1.3 and show how to modify it to be biregular, while mostly preserving its value. We will do so in two steps, first making Ψ left-regular, and then applying a generic expander-based transformation (similar to that of Moshkovitz and Raz [MR10]) to gain bounded-degree right-regularity.

Fix $\varphi, \varepsilon > 0$, and let Ψ be the 2-Prover-1-Round game constructed for Theorem 1.3. Recall that this requires us to choose some large enough ℓ relative to $\varphi^{-1}, \varepsilon^{-1}$, some large enough q relative to ℓ, and set $R = q^{2\ell}$. We also set $\delta = \frac{\varepsilon}{1000}$, $0 < c$ arbitrarily small relative to δ, and $k = q^{2(1+c)\ell}$, $q = 2$. Finally, we construct our 2-Prover-1-Round game from a hard instance of $\mathsf{Gap3Lin}$ with the appropriate completeness and soundness, so that it is NP-hard to distinguish between,

$$\mathsf{val}(\Psi) \geqslant 1 - \varphi \quad \text{and} \quad \mathsf{val}(\Psi) \leqslant \frac{1}{q^{2(1-1000\delta)\ell}} = \frac{1}{R^{1-\varepsilon}}.$$

It is clear that our 2-Prover-1-Round game can equivalently be viewed as an instance of bipartite 2-CSP, so let us analyze the underlying graph. Let $\mathcal{U}$ denote the set of possible questions to the first prover. Recall that the set of left vertices is,

$$\mathsf{Left} = \{L \oplus H_U \mid U \in \mathcal{U}, L \in \mathsf{Grass}_q(\mathbb{F}_q^U, 2\ell), L \cap H_U = \{0\}\},$$

while the set of right vertices is

$$\mathsf{Right} = \{R \in \mathsf{Grass}_q(\mathbb{F}_q^U, 2(1-\delta)\ell) \mid U \in \mathcal{U}\}.$$

The constraints of Ψ are on edges of the graph and are weighted. We use $w(\cdot)$ to denote the weight function, which takes as input edges of the form $(L \oplus H_U, R)$. Recall that the weighting is defined according to the process described in Section 4.1.4. That is, the weight of the constraint on edge $(L \oplus H_U, R)$ is exactly the probability that this edge is chosen according to the process in Section 4.1.4. For a fixed $L \oplus H_U$, let us define

$$w_{L\oplus H_U}(R) = \frac{|\{L' \oplus H_{U'} \in [L \oplus H_U] \mid R \subseteq L'\}|}{|[L \oplus H_U]|} \cdot \frac{1}{\begin{bmatrix} 2\ell \\ 2(1-\delta)\ell \end{bmatrix}_q},$$

which is the probability of choosing the edge $(L \oplus H_U, R)$ conditioned on first choosing $L \oplus H_U$. Since we choose $L \oplus H_U \in \mathsf{Left}$ uniformly, it follows that

$$w(L \oplus H_U, R) = \frac{w_{L\oplus H_U}(R)}{|\mathsf{Left}|}.$$

Define the neighborhood of a vertex as,

$$\mathsf{nb}(L \oplus H_U) = \{R \in \mathsf{Right} \mid w_{L\oplus H_U}(R) > 0\}.$$

Call $L \oplus H_U$ trivial if there is an equation $e \in U$ such that for every basis $x_1, \ldots, x_{2\ell} \in \mathbb{F}_q^U$ of L, we have that, for every $i \in [2\ell]$, the point x_i restricted to the variables in e is of the form (α, α, α) for some $\alpha \in \mathbb{F}_q$.

Claim 7.3. *The fraction of* $L \oplus H_U \in \mathsf{Left}$ *that are trivial is at most* $2q^{-(2-2c)\ell}$.

Proof. Fix a $U \in \mathcal{U}$. Note that it suffices to show that at most $2q^{-(2-2c)\ell}$ vertices of the form $L \oplus H_U$ are trivial, as for each $U \in \mathcal{U}$, there are an equal number of vertices $L \oplus H_U$.

Write $U = (x_1, \ldots, x_{3k})$, where the ith equation in U contains the variables $x_{3i-2}, x_{3i-1}, x_{3i}$. Call these three coordinates a block, so that each $x \in \mathbb{F}_q^{3k}$ consists of k blocks of consecutive coordinates. Let us bound the fraction of L such that $L \oplus H_U$ is trivial. For $y_1, \ldots, y_{2\ell} \in \mathbb{F}_q^{3k}$, let $s(y_1, \ldots, y_{2\ell})$ be the number of blocks where $y_1, \ldots, y_{2\ell}$ are all of the form (α, α, α) for some $\alpha \in \mathbb{F}_q$. Then

$$\Pr_L[L \oplus H_U \text{ is trivial}] \leqslant 2 \Pr_{y_1,\ldots,y_{2\ell}}[s(y_1, \ldots, y_{2\ell}) = 0],$$

where the factor of 2 accounts for the probability that either $y_1, \ldots, y_{2\ell}$ are not linearly dependent, or $\mathrm{span}(y_1, \ldots, y_{2\ell}) \cap H_U \neq \{0\}$. Note that the probability that a specific block is trivial is $q^{-4\ell}$, hence by linearity of expectation we get that

$$\mathop{\mathbb{E}}_{y_1,\ldots,y_{2\ell}}[s(y_1, \ldots, y_{2\ell})] = kq^{-4\ell} = q^{-(2-2c)\ell},$$

and therefore

$$\Pr_{x_1,\ldots,x_{2\ell}}[s(x_1, \ldots, x_{2\ell}) \geqslant 1] \leqslant q^{-(2-2c)\ell}. \qquad \square$$

Now, let Ψ' be the instance obtained from Ψ after removing all trivial $L \oplus H_U$ from Ψ. Let Left' denote the set of left vertices in Ψ' and let $w'(\cdot)$ denote the weight function over edges in Ψ', which is given by choosing $L \oplus H_U \in \mathsf{Left}'$ uniformly, and then choosing $R \in \mathsf{nb}(L \oplus H_U)$ with probability proportional to $w_{L\oplus H_U}(R)$. It follows that,

$$w'(L \oplus H_U, R) = \frac{w_{L\oplus H_U}(R)}{|\mathsf{Left}'|}. \tag{28}$$

Since so few vertices are removed when going from Left to Left', the value of Ψ and Ψ' are roughly the same.

Claim 7.4. *We have,*

$$\mathsf{val}(\Psi) - 2q^{-(2-2c)\ell} \leqslant \mathsf{val}(\Psi') \leqslant \mathsf{val}(\Psi) + 3q^{-(2-2c)\ell}.$$

Proof. Fix any assignment to Ψ' and let E' be the set of constraints that it satisfies. Note that E' is also a set of constraints in Ψ. Let $w(E')$ and $w'(E')$ be the sum of the weights of the edges in E' under w and w' respectively. For every $(L \oplus H_U, R) \in E'$, we have the upper bound

$$w'(L \oplus H_U, R) = w(L \oplus H_U, R) \cdot \frac{|\mathsf{Left}'|}{|\mathsf{Left}|} \leqslant w(L \oplus H_U, R) \cdot \frac{1}{1 - 2q^{-(2-2c)\ell}}$$

Since this holds for every edge in E', we get that $w'(E') \leqslant \frac{w(E')}{1-2q^{-(2-2c)\ell}}$, and hence

$$\mathsf{val}(\Psi') \leqslant \frac{\mathsf{val}(\Psi)}{1 - 2q^{-(2-2c)\ell}} \leqslant \mathsf{val}(\Psi) + 3q^{-(2-2c)\ell}$$

For the lower bound, fix the any assignment to Ψ and let E be the set of constraints that it satisfies in Ψ. Once again let $w(E)$ and $w'(E)$ be the sum of the weights of E under w and w' respectively. If the left endpoint of an edge in E is no longer in Left', we define its weight under w' to be 0. Then note that

$$w'(E) \geqslant \sum_{(L \oplus H_U, R) \in E, L \oplus H_U \in \mathsf{Left}'} w(L \oplus H_U, R) \geqslant w(E) - 2q^{-(2-2c)\ell}. \qquad \square$$

Claim 7.4 allows us to remove trivial vertices without affecting the value much. Towards going from a weighted 2-CSP to an unweighted 2-CSP, we will next bound the size of the neighborhoods in Ψ'.

Claim 7.5. *For each non-trivial $L \oplus H_U$ we have $|\mathsf{nb}(L \oplus H_U)| \leqslant 10^k q^{6k\ell}$.*

Proof. Let $U = (x_1, \ldots, x_{3k}) = (e_1, \ldots, e_k)$ and suppose equation e_i contains variables $(x_{3i-2}, x_{3i-1}, x_{3i})$. Since $L \oplus H_U$ is not trivial, for each i, there must be a point $v \in L \subseteq \mathbb{F}_q^U$ such that the values of v restricted to the coordinates of variables $(x_{3i-2}, x_{3i-1}, x_{3i})$ are not all equal. Without loss of generality, say that it is x_{3i} for each $1 \leqslant i \leqslant k$. It follows that in order to have

$$L \oplus H_U \oplus H_{U'} = L' \oplus H_U \oplus H_{U'},$$

U' must contain an equation with the variable x_{3i} for each $1 \leqslant i \leqslant k$. Let E_i denote this set of equations for each i. By the regularity assumptions on our $\mathsf{3Lin}$ instance, $|E_i| \leqslant 10$ and $E_i \cap E_j = \emptyset$ for $i \neq j$. It follows that U' must contain exactly one equation from each E_i, and that these form all k equations of U', so there are at most 10^k possible U' for which there can exist $L' \subseteq U'$, such that $L' \oplus H_{U'} \in [L \oplus H_U]$. The lemma follows from the observation that $|\mathsf{Grass}_q(3k, 2(1-\delta)\ell)| \leqslant q^{6k\ell}$. $\square$

Performing the same procedure as in [KR03, Lemma 3.4], we can turn Ψ' into a bipartite, left-regular 2-CSP instance, while again not affecting its value too much. Let $Q = 10^k q^{6k\ell}$ be the upper bound on neighborhood sizes in Claim 7.5.

Claim 7.6. *For any $C \in \mathbb{N}$, there is a polynomial time algorithm that takes Ψ' as input and outputs a bipartite 2-CSP Ψ'' that is left regular with degree $C \cdot Q$ such that*

$$\mathsf{val}(\Psi') - \frac{1}{C} \leqslant \mathsf{val}(\Psi'') \leqslant \mathsf{val}(\Psi') + \frac{1}{C}.$$

Proof. We create Ψ'' by doing the following for each vertex $L \oplus H_U \in \mathsf{Left}'$. Let $R_1, \ldots, R_m$ be the vertices in $\mathsf{nb}(L \oplus H_U)$. For each $2 \leqslant i \leqslant m$, add $\lfloor w_{L\oplus H_U}(R_i) \cdot C \cdot Q \rfloor$ edges from $L \oplus H_U$ to R_i. Then, add $C \cdot Q - \sum_{i=2}^{m} \lfloor w_{L\oplus H_U}(R_i) C \cdot Q \rfloor$ edges from $L \oplus H_U$ to R_1. It is clear that Ψ'' is left regular with degree $C \cdot Q$, and that for each R_i with $2 \leqslant i \leqslant m$, there are at most $w_{L\oplus H_U}(R_i) C \cdot Q$ edges between $L \oplus H_U$ and R_i, while for R_1, there are at most $\left(w_{L\oplus H_U}(R_1) + \frac{1}{C}\right) C \cdot Q$ edges between $L \oplus H_U$ and R_1.

To bound the value of Ψ'' fix a labeling for it, and note that this is also a labeling for Ψ', since the two CSPs have the same vertex set. For every left vertex, $L \oplus H_U$, if s-fraction of its neighboring constraints are satisfied in Ψ' (where this fraction is under the weighting in Ψ'), then between $(s - 1/C)$ and $(s + 1/C)$-fraction of its neighboring constraints are satisfied in Ψ''. □

Applying Claim 7.6 with $C = q^{10\ell}$, we obtain a bipartite 2-CSP, Ψ'', that is left regular with degree $q^{10\ell}Q$, that still has nearly the same completeness and soundness as our original instance Ψ. We will now create a 2-regular bipartite CSP from Ψ'', using a technique in the spirit of the right-degree reduction of Moshkovitz and Raz [MR10]. We need the expander mixing lemma, stated below.

Lemma 7.7 (Expander Mixing Lemma). *Let $G = (A \cup B, E)$ be a bipartite biregular graph with second largest eigenvalue λ and vertex degree D. Then for any sets of vertices $A_0 \subseteq A, B_0 \subseteq B$ with sizes $|A_0| = \alpha|A|$ and $|B_0| = \beta|B_0|$, we have*

$$\left| \frac{e(A_0, B_0)}{|E|} - \alpha\beta \right| \leqslant \frac{\lambda}{D}\sqrt{\alpha\beta},$$

where $e(A_0, B_0)$ is the number of edges between the sets of vertices A_0 and B_0

Lemma 7.8. *For any parameter D, there is a polynomial time algorithm that takes as input, a bipartite, left regular 2-CSP, Ψ_0, with left degree d_{left}, projection constraints, and left and right alphabets Σ_A and Σ_B respectively, and outputs a bipartite, $(D \cdot d_{\mathsf{left}}, D)$-regular 2-CSP, Ψ_0', with projection constraints and the same alphabets, such that*

$$\mathsf{val}(\Psi_0) \leqslant \mathsf{val}(\Psi_0') \leqslant \mathsf{val}(\Psi_0) + O\left(D^{-1/2}\right).$$

Proof. For every pair of integers $N \geqslant D$, there is a polynomial time algorithm which constructs a biregular, bipartite expander graph, with N vertices on each side, vertex degree D, and second eigenvalue $O(D^{1/2})$. When N is sufficiently large relative to D, the algorithm is from [Alo21], and otherwise, a brute force algorithm suffices.

The construction of Ψ_0' is as follows. Fix the parameter D from the lemma statement and let A and B be the left and right sides of Ψ_0 respectively, and let E be its set of edges. We first describe how to construct A' and B', which are the left and right sides of Ψ_0'. For each $b \in B$, denote by A_b the set of neighbors of b in Ψ_0, and by $d(b) = |A_b|$ the degree of b. Then construct a D-regular bipartite expander graph, $H_b = (A_b, B_b, E_b)$ such that $|A_b| = |B_b| = d(b)$ with second eigenvalue $O\left(D^{1/2}\right)$. The left vertices of Ψ_0' are unchanged, so $A' = A$, but the right vertices are now $B' = \bigcup_{b\in B} B_b$. The edges are $E' = \bigcup_{b\in B} E_b$, and for each $(a, b') \in E'$, if $b' \in B_b$, then the constraint on (a, b') is the same as that on (a, b) in Ψ_0. The alphabets of Ψ_0' are also unchanged. It is clear that Ψ_0' is (Dd_{left}, D)-regular, so it remains to show the bounds on its value.

For the lower bound, fix an assignment F to Ψ_0. Consider the following assignment F' to Ψ_0'. For $a' \in A' = A$ set $F'(a) = F(a')$, and for $b' \in B'$ set $F'(b') = F(b)$, where b is the vertex in B such that $b' \in B_{b'}$. It is clear that F' satisfies the same fraction of constraints as F, so $\mathsf{val}(\Psi_0') \geqslant \mathsf{val}(\Psi_0)$.

For the upper bound, let F' be an assignment to Ψ_0'. We show how to derive an assignment F with nearly the same value. For each $b \in B$, choose $b' \in B_b$ uniformly at random and set $F(b) = F'(b')$. To analyze the value of F, for each $\sigma \in \Sigma_B$, let

$$X_{b,\sigma} = \{a \in A_b \mid (a,b) \in E, \text{and } F'(a), \sigma \text{ satisfy the constraint on } (a,b) \text{ in } \Psi_0\}.$$

For each $b \in B$, let $Y_{b,\sigma} = \{b' \in B_b \mid F'(b) = \sigma\}$. Then,

$$\mathop{\mathbb{E}}_{F}[\mathsf{val}(F)] = \mathop{\mathbb{E}}_{b \in B}\left[\sum_{\sigma \in \Sigma_B} \frac{|X_{b,\sigma}||Y_{b,\sigma}|}{d(b)^2}\right],$$

where the expectation on the left is over the randomness when choosing assignment F and the expectation on the right chooses b proportional to $d(b)$.

By Lemma 7.7 on the graph H_b, we have that

$$\frac{|X_{b,\sigma}||Y_{b,\sigma}|}{d(b)^2} \geqslant \frac{e(X_{b,\sigma}, Y_{b,\sigma})}{d(b) \cdot D} - \frac{O(D^{1/2})}{D} \cdot \frac{\sqrt{|X_{b,\sigma}||Y_{b,\sigma}|}}{d(b)} = \frac{e(X_{b,\sigma}, Y_{b,\sigma})}{d(b) \cdot D} - O\left(D^{-1/2}\right) \cdot \frac{\sqrt{|X_{b,\sigma}||Y_{b,\sigma}|}}{d(b)}$$

so

$$\begin{aligned}
\mathop{\mathbb{E}}_{F}[\mathsf{val}(F)] &\geqslant \mathop{\mathbb{E}}_{b \in B}\left[\sum_{\sigma \in \Sigma_B} \frac{e(X_{b,\sigma}, Y_{b,\sigma})}{d(b) \cdot D}\right] - O\left(D^{-1/2}\right) \cdot \mathop{\mathbb{E}}_{b \in B}\left[\sum_{\sigma \in \Sigma_B} \frac{\sqrt{|X_{b,\sigma}||Y_{b,\sigma}|}}{d(b)}\right] \\
&\geqslant \mathsf{val}(F') + O\left(D^{-1/2}\right) \cdot \mathop{\mathbb{E}}_{b \in B}\left[\frac{1}{d(b)} \cdot \sqrt{\sum_{\sigma \in \Sigma_B} |X_{b,\sigma}|} \cdot \sqrt{\sum_{\sigma \in \Sigma_B} |Y_{b,\sigma}|}\right] \\
&\geqslant \mathsf{val}(F') - O\left(D^{-1/2}\right).
\end{aligned}$$

When transitioning to the second line, we are using the fact that the first term is exactly $\mathsf{val}(F')$ and we are bounding the second term using the Cauchy-Schwarz inequality. In the transition to the last line we are using the fact that $\sum_{\sigma \in \Sigma_B} |X_{b,\sigma}| = d(b)$ because the constraints in Ψ_0 are projections, and $\sum_{\sigma \in \Sigma_B} |Y_{b,\sigma}| = d(b)$. ☐

We are now ready to complete the proof of Theorem 7.2.

Proof of Theorem 7.2. Fix $\varphi, \varepsilon > 0$ as in Theorem 7.2 and take Ψ to be the 2-CSP from Theorem 1.3 with δ therein set to $\varphi/2$, ε therein set to $\varepsilon/2$, and set ℓ and therefore R sufficiently large, so that distinguishing between $\mathsf{val}(\Psi) \geqslant 1 - \varphi/2$ and $\mathsf{val}(\Psi) \leqslant R^{-(1-\varepsilon/2)}$ is NP-hard. Starting from Ψ, obtain Ψ' by deleting trivial vertices and applying the transformation in Claim 7.6 with $C = q^{10\ell}$, and then let Ψ'' be obtained from Ψ' via the transformation in Lemma 7.8 with $D = q^{10\ell}$. It is clear that Ψ'' is $(C \cdot Q \cdot D, D)$-regular, has alphabet size R, and is constructed from Ψ in polynomial time.

It remains to show that the transformations performed preserve the value. By Claim 7.4 and Claim 7.6, we have

$$\mathsf{val}(\Psi) - 2q^{-(2-2c)\ell} - q^{-10\ell} \leqslant \mathsf{val}(\Psi') \leqslant \mathsf{val}(\Psi) + 3q^{-(2-2c)\ell} + q^{-10\ell},$$

and by Lemma 7.8, we have

$$\mathsf{val}(\Psi') \leqslant \mathsf{val}(\Psi'') \leqslant \mathsf{val}(\Psi') + O\left(q^{-5\ell}\right),$$

so altogether we get that

$$\mathsf{val}(\Psi) - 2q^{-(2-2c)\ell} - q^{-10\ell} \leqslant \mathsf{val}(\Psi'') \leqslant \mathsf{val}(\Psi) + 3q^{-(2-2c)\ell} + q^{-10\ell} + O\left(q^{-5\ell}\right).$$

Since we set ℓ and, as a result, R sufficiently large, it is clear that if $\mathsf{val}(\Psi) \geqslant 1-\varphi/2$, then $\mathsf{val}(\Psi'') \geqslant 1-\varphi$, and if $\mathsf{val}(\Psi) \leqslant R^{-(1-\varepsilon/2)}$ then $\mathsf{val}(\Psi'') \leqslant R^{-(1-\varepsilon)}$. This establishes Theorem 7.2. □

7.3.2 Sparsification

In [LM24], Lee and Manurangsi show how to conclude Theorem 1.5 from Theorem 7.2 via a sparsification procedure. We summarize the steps here. Fix the $\eta > 0$ for Theorem 1.5. Set $\varphi = \varepsilon = 0.01\eta$ in Theorem 7.2 and let Ψ be the resulting hard bipartite (d_1, d_2)-regular 2-CSP, and it is NP-hard to distinguish between the cases

$$\mathsf{val}(\Psi) \geqslant 1-\varphi \quad \text{and} \quad \mathsf{val}(\Psi) \leqslant \frac{1}{R^{1-\varepsilon}},$$

where R is the sufficiently large alphabet size. It is shown in [LM24, Lemma 10], stated below, that the degrees of Ψ can be multiplied by arbitrary constants by simply copying vertices.

Lemma 7.9. *[LM24, Lemma 10] For any integers d_1, d_2, c_1, c_2, there is a polynomial time reduction from a bipartite (d_1, d_2)-biregular CSP, Ψ, to a bipartite $(c_2d_1d_2, c_1d_1d_2)$-biregular CSP Ψ', such that $\mathsf{val}(\Psi) = \mathsf{val}(\Psi')$, and such that the left and right alphabet sizes are preserved.*

It is then shown in [LM24, Theorem 11], stated below, that one can perform a subsampling procedure to Ψ' that significantly lowers the degree.

Theorem 7.10. *[LM24, Theorem 11] For any $0 < \nu_1 < \nu_2 \leqslant 1$, any positive integer C, and any sufficiently large positive integers $d_A, d_B \geqslant d_0(\varphi, \nu)$, and $R \geqslant R_0(\delta, \nu, d_A, d_B)$, the following holds: there is a randomized polynomial-time reduction from a bipartite (d_AC, d_BC)-biregular 2-CSP, Ψ', with alphabet size at most R, (d_A, d_B)-bounded degree 2-CSP, Ψ'', such that, with probability at least $2/3$ the following two items hold*

- *Completeness:* $\mathsf{val}(\Psi'') \geqslant \mathsf{val}(\Psi') - \nu_1$,
- *Soundness: If* $\mathsf{val}(\Psi') \leqslant \frac{1}{R^{\nu_2}}$, *then* $\mathsf{val}(\Psi'') \leqslant \frac{1}{\nu_2-\nu_1}\left(\frac{1}{d_A} + \frac{1}{d_B}\right)$

Putting everything together, we can prove Theorem 1.5.

Proof of Theorem 1.5. Recall the values η and d from Theorem 1.5. Start with an instance Ψ of 2-CSP from Theorem 7.2 with $\varphi = \varepsilon = 0.01\eta$ and sufficiently large alphabet size R. Then Ψ is (d_1, d_2)-biregular, with sufficiently large alphabet size R relative to $\varphi^{-1}, \varepsilon^{-1}$, and d. For such a Ψ, it is NP-hard to distinguish whether $\mathsf{val}(\Psi) = 1 - 0.01\eta$, or $\mathsf{val}(\Psi) \leqslant \frac{1}{R^{1-0.01\eta}}$.

Applying Lemma 7.9 with $c_1 = c_2 = d$ yields, in polynomial time, a (dd_1d_2, dd_1d_2)-biregular 2-CSP, Ψ', with alphabet size R and satisfying $\mathsf{val}(\Psi') = \mathsf{val}(\Psi)$. Next, applying Theorem 7.10, with $d_A = d, d_B = d, C = d_1d_2, \nu_1 = 0.01\eta, \nu_2 = 1 - \varepsilon$, we get (after randomized polynomial time reduction) a 2-CSP Ψ'' with degree at most d such that:

- If $\mathsf{val}(\Psi) \geqslant 1-\varphi$, then $\mathsf{val}(\Psi'') \geqslant 1 - \varphi - \nu_1 = 1 - 0.02\eta$.
- If $\mathsf{val}(\Psi) \leqslant \frac{1}{R^{1-\varepsilon}}$, then $\mathsf{val}(\Psi'') \leqslant \frac{1}{1-\varepsilon-\nu_1}\left(\frac{1}{d} + \frac{1}{d}\right) = \frac{1}{1-0.02\eta} \cdot \frac{2}{d}$.

Finally note that,

$$\frac{1-0.02\varepsilon}{\frac{1}{1-0.02\varepsilon}\cdot\frac{2}{d}} \geqslant d\left(\frac{1}{2}-\varepsilon\right).$$

Thus, by Theorem 7.2 and the randomized polynomial time reduction above, it follows that unless NP $\subseteq$ BPP, there is no polynomial time $d\left(\frac{1}{2}-\eta\right)$ approximation algorithm for 2-CSP with degree at most d. $\square$

7.4 Proof of Theorem 1.6

Combining our 2-Prover-1-Round Game with the reductions in [Lae14] we obtain improved hardness of approximation results for Rooted k-connectivity on undirected graphs, the vertex-connectivity survivable network design problem, and the vertex-connectivity k-route cut problem on undirected graphs. We give an overview of the reduction here. It will be more convenient to refer to 2-CSPs rather than 2-Prover-1-Round Game for the remainder of this subsection.

In the reduction of [Lae14] one uses a minimization variant for a 2-CSP Ψ, called Min-Rep. The input in Min-Rep is still a 2-CSP, but now each alphabet symbol is assigned a cost, and a labeling of Ψ is allowed to assign multiple alphabet symbols to each. The cost of a labeling is the total cost of the alphabet symbols used over all vertices and $\mathsf{MinRep}(\Psi)$ is the minimum cost labeling such that every constraint of Ψ is satisfied. Here, a constraint on the edge (u, v) is satisfied as long as one pair of labels from those assigned to u and v respectively satisfy the constraint. We refer to [Lae14, page 1630] for a more formal definition. The following lemma [Lae14] describes approximation preserving reductions from Min-Rep to various connectivity problems.

Theorem 7.11. *[Lae14, Theorem 3.1] There are polynomial-time, approximation-preserving reductions which take an instance of minimum cost label cover with degree* d *and alphabet size* R*, and output:*

- *an instance of rooted* k*-connectivity on undirected graphs with* $k = O(d^3 \cdot R + d^4)$*,*
- *an instance of vertex-connectivity survivable network design on undirected graphs with maximum requirement* $k = O(d \cdot R + d^2)$*,*
- *an instance of vertex-connectivity* k*-route cut on undirected graphs with* $k = O(d \cdot R)$

To obtain hardness of approximation results from Theorem 7.11, we must first obtain a hardness of approximation result for Min-Rep. This can be done by using the following result due to [Man19].

Lemma 7.12. *[Man19, Lemma 7] For any* $\varepsilon, \gamma > 0$*, there is a polynomial time randomized reduction that, given any* 2*-CSP with projection constraints* Ψ*, alphabet size* R*, and size* N*, outputs a* 2*-CSP with projection constraints* Ψ' *with the same alphabet and maximum degree at most*

$$d := 10^6\left(\frac{2\log(2R)}{\sqrt{\gamma}}\right)$$

such that

- ***(Completeness)*** *if* $\mathsf{val}(\Psi) \geqslant 1-\varepsilon$*, then* $\mathsf{MinRep}(\Psi') \leqslant (1+\varepsilon\cdot d)N$ *with probability* 0.9*, and,*
- ***(Soundness)*** *if* $\text{val}(\Pi) < \gamma$*, then* $\text{Min-Rep}(\Pi') > \left(\frac{0.06}{\sqrt{\gamma}}\right)N$ *with probability* 0.9*.*

Using Lemma 7.12, we get the following hardness of approximation result for Min-Rep.

Theorem 7.13. *For every $\varepsilon' > 0$ and sufficiently large $g > 0$, given a 2-CSP, Ψ with maximum degree $O(g \log g)$ and alphabet size $O(g^{2+\varepsilon'})$, it is NP-hard under randomized reduction to approximate the* $\mathsf{MinRep}(\Psi)$ *to within a factor of g.*

Proof. Fix sufficiently large $g > 0$ and take Ψ from Theorem 1.3 with $\delta = \frac{1}{g}$, ε such that $1/(1-\varepsilon) = 1 + \varepsilon'/2$, and alphabet R such that the value in the no case is $\gamma := R^{-(1-\varepsilon)} = 2^{-11} g^{-2}$. Let N denote the size of Ψ. Then applying Lemma 7.12 we get that if $\mathsf{val}(\Psi) \geqslant 1 - \delta$, then $\mathsf{MinRep}(\Psi') \leqslant 2N$, while if $\mathsf{val}(\Psi) \leqslant \gamma$, then $\mathsf{MinRep}(\Psi') > \left(\frac{0.06}{\sqrt{\gamma}}\right) N$. Therefore, it is NP-hard under randomized reduction to approximate $\mathsf{MinRep}(\Psi')$ within factor,

$$0.03 \cdot \gamma^{-1/2} \geqslant g.$$

The alphabet size of Ψ' is $R = O(g^{2/(1-\varepsilon)}) = O(g^{2+\varepsilon'})$ and the maximum degree is at most $d = O(g \log g)$. □

Finally, by combining Theorem 7.11 with Theorem 7.13, we can show Theorem 1.6.

Proof of Theorem 1.6. Fix $\varepsilon > 0$ and $g > 0$ sufficiently large, and let Ψ be the 2-CSP from Theorem 7.13 such that $\mathsf{MinRep}(\Psi)$ is hard to approximate within factor g, and Ψ has maximum degree $d = O(g \log g)$ and alphabet size $R = O(g^{2+\varepsilon})$. By Theorem 7.11, there is a polynomial-time algorithm that produces instances of the following problems which are hard to approximate within factor g:

- rooted k-connectivity on undirected graphs with $k = O(g^{5+\varepsilon} \log g)$,
- vertex-connectivity survivable network design on undirected graphs with maximum requirement $k = O(g^{3+\varepsilon} \log g)$,
- vertex-connectivity k-route cut on undirected graphs with $k = O(g^{3+\varepsilon} \log(g))$.

Expressing g in terms of k in each of the three cases gives the desired result. □

8 Bounding the Number of Successful Zoom-outs of a Fixed Codimension

The goal of this section is to prove Lemma 5.24. Let us recall some context first. Throughout the section, we work in the second prover's space, $\mathbb{F}_q^V$, where V is some question to the second prover (in the outer PCP).

Accordingly, we make the assumption that $n := |V| \gg \ell$, say $n \geqslant 2^{100} q^{\ell}$ to be concrete. Also, as $n = |V|$, we will write the ambient space as $\mathbb{F}_q^n$ from now on. We fix T to be a table that assigns, to each $L \in \mathsf{Grass}_q(n, 2\ell)$, a linear function on L. For ease of notation, we define $\mathcal{L} := \mathsf{Grass}_q(n, 2\ell)$ throughout this section. Now, let us review the set up of Lemma 5.24. Recall that we set

$$\xi := \delta^5, \quad \delta_2 := \xi/100, \quad t := \left(2^{2+10/\delta_2}\right)!\ .$$

Let $\mathcal{S} = \{W_1, \ldots, W_N\}$ be a set of codimension r-subspaces in $\mathbb{F}_q^n$ of size

$$N \geqslant q^{100(t-1)! r^2 \ell \xi^{-1}},$$

where $r \leqslant \frac{10}{\delta}$. For each W_i, let $f_i : W_i \to \mathbb{F}_q$ be a linear function such that $f_i|_L \equiv T[L]$ for at least C-fraction of the 2ℓ-subspaces $L \subseteq W_i$, where $C \geqslant q^{-2(1-\xi)\ell}$, and $\xi > 0$.

8.1 Step 1: Reducing to a Generic Set of Subspaces

Instead of working with all of $\mathcal{S}$, we will want to use only a subset of $\mathcal{S}$ which is generic. This reduction to a generic set of subspaces allows us to readily use the sampling lemmas from Section 5.3.2. Applying Lemma 5.12, with parameter t as define, we get that there exists a subspace $W'_{\mathsf{amb}} \subseteq \mathbb{F}_q^n$ and a set of

$$m_1 \geqslant \frac{N^{\frac{1}{(r+1)\cdot(t-1)!}}}{q^{3r}} \geqslant q^{75r\ell\xi^{-1}}$$

subspaces $\mathcal{W} = \{W_1, \dots, W_{m_1}\} \subseteq \mathcal{S}$, such that

- Each $W_i \in \mathcal{W}$ is contained in W'_{amb} and has codimension s with respect to W'_{amb}, where $s \leqslant r$.
- $\mathcal{W}$ is t-generic with respect to W'_{amb}.

We remark that this subspace W'_{amb} will ultimately be the one used for Lemma 5.24 (as the subspace "A" there). The remainder of the proof is devoted to finding the appropriate linear function $h' : W'_{\mathsf{amb}} \to \mathbb{F}_q$, and the set $\mathcal{W}'$ as required by Lemma 5.24. The set $\mathcal{W}'$ we ultimately find will be a subset of $\mathcal{W}$ above.

8.2 Step 2: Finding Local Agreement

For a subspace X and linear assignment to X, $\sigma \in \mathbb{F}_q^X$, let

$$\mathcal{L}_X = \mathsf{Zoom}_{2\ell}[X, W'_{\mathsf{amb}}] \quad \text{and} \quad \mathcal{L}_{X,\sigma} = \{L \in \mathcal{L}_X \mid T[L]|_X = \sigma\}.$$

Likewise, define

$$\mathcal{W}_X = \{W_i \in \mathcal{W} \mid X \subseteq W_i\} \quad \text{and} \quad \mathcal{W}_{X,\sigma} = \{W_i \in \mathcal{W}_X \mid f_i|_X = \sigma\}.$$

We will use $\mu_{X,\circ}(\cdot)$ to denote the uniform measure over $\mathcal{L}_X$. The first step of our proof is to find sets $\mathcal{W}_{X,\sigma}$ and $\mathcal{L}_{X,\sigma}$ that have strong agreement between them, in the sense of the following lemma. The approach of this first step is similar to that of [IKW12, BDN17, MZ23]. Fix $\gamma > 0$ to be a small constant, say $\gamma = 10^{-6}$.

Lemma 8.1. *There exists a $2\left(1 - \frac{\xi}{2}\right)\ell$-dimensional subspace X and a linear assignment, σ, to X, such that the following hold:*

- $\mu_{X,\circ}(\mathcal{L}_{X,\sigma}) \geqslant \frac{C}{6}$.
- $|\mathcal{W}_{X,\sigma}| \geqslant \frac{m_1}{q^{10r\ell}}$.
- *Choosing $L \in \mathcal{L}_{X,\sigma}$ uniformly, and $W_i \in \mathcal{W}_{X,\sigma}$ uniformly such that $W_i \supseteq L$, we have*

$$\Pr_{L,W_i}[f_i|_L \not\equiv T[L]] \leqslant 5\gamma.$$

Proof. Deferred to Appendix E.1. □

We next state Corollary 8.2, which is an immediate consequence of Lemma 8.1. There are two differences between the two statements though. First, in Corollary 8.2 we require the third condition to hold for every $L \in \mathcal{L}_{X,\sigma}$ (instead of for a random L as in Lemma 8.1). Second we require every $L \in \mathcal{L}_{X,\sigma}$ to be contained in roughly the same number of $W_i \in \mathcal{W}_{X,\sigma}$.

Corollary 8.2. *Taking $\mathcal{L}_{X,\sigma}$ and $\mathcal{W}_{X,\sigma}$ from Lemma 8.1, there is a subset $\mathcal{L}'_{X,\sigma} \subseteq \mathcal{L}_{X,\sigma}$ such that the following hold.*

- $\mu_{X,\circ}(\mathcal{L}'_{X,\sigma}) \geqslant \frac{C}{12}$
- $m_2 := |\mathcal{W}_{X,\sigma}| \geqslant \frac{m_1}{q^{10r\ell}}$.
- *For every $L \in \mathcal{L}'_{X,\sigma}$, choosing $W_i \in \mathcal{W}_{X,\sigma}$ uniformly such that $W_i \supseteq L$, we have*

$$\Pr_{W_i \supseteq L, W_i \in \mathcal{W}_{X,\sigma}}[f_i|_L \not\equiv T[L]] \leqslant 12\gamma.$$

- *For every $L \in \mathcal{L}'_{X,\sigma}$,*

$$0.95 \cdot |\mathcal{W}_{X,\sigma}| \cdot q^{-\xi\ell\cdot s} \leqslant N_{\mathcal{W}_{X,\sigma}}(L) \leqslant 1.05 \cdot |\mathcal{W}_{X,\sigma}| \cdot q^{-\xi\ell\cdot s}.$$

Recall that $N_{\mathcal{W}_{X,\sigma}}(L)$ is as defined in (13)*, and s in the fourth item is the codimension of each subspace in $\mathcal{W} \supseteq \mathcal{W}_{X,\sigma}$.*

Proof. Take $X, \sigma, \mathcal{L}_{X,\sigma}$, and $\mathcal{W}_{X,\sigma}$ as guaranteed by Lemma 8.1, so that $\mu_{X,\circ}(\mathcal{L}_{X,\sigma}) \geqslant \frac{C}{6}$. We will keep the same X, σ, but we remove some L's from $\mathcal{L}_{X,\sigma}$ to make the third and fourth items hold.

By Markov's inequality, at most $\frac{5}{12}$-fraction of $L \in \mathcal{L}_{X,\sigma}$ violate the third item. By Lemma 5.21 applied to $\mathcal{W}_{X,\sigma}$ with parameters $Q = X$, $j = 2\ell$, $a = 2\left(1 - \frac{\xi}{2}\right)\ell$, $r = s$, and $c = 0.05$, we have that

$$\begin{aligned}
&\Pr_{L \in \mathsf{Zoom}_{2\ell}[Q,V]}\left[\left|N_{\mathcal{W}_{X,\sigma}}(L) - q^{-\xi\ell\cdot s}m_2\right| \geqslant 0.1q^{-\xi\ell}m_2\right] \\
&\leqslant \Pr_{L \in \mathsf{Zoom}_{2\ell}[Q,V]}\left[\left|N_{\mathcal{W}_{X,\sigma}}(L) - \beta\right| \geqslant 0.05\beta\right] \\
&\leqslant \frac{405q^{\xi\ell\cdot s}}{m_2} \\
&\leqslant \frac{C}{100},
\end{aligned}$$

where in the above we use $\beta = \mathbb{E}_{L \in \mathsf{Zoom}_{2\ell}[Q,V]}[N_{\mathcal{W}_{X,\sigma}}(L)]$ for convenience, and note from Lemma 5.19 that $|\beta - q^{-\xi\ell\cdot s}m_2| \leqslant 0.01q^{-\xi\ell\cdot s}m_2$.

It follows that after removing the $L \in \mathcal{L}_{X,\sigma}$ that do not satisfy the third or fourth condition, we arrive at the desired $\mathcal{L}'_{X,\sigma}$, which still has measure at least

$$\frac{7}{12} \cdot \frac{C}{6} - \frac{C}{100} \geqslant \frac{C}{12},$$

under $\mu_{X,\circ}(\cdot)$. $\square$

For the rest of the argument, let us fix X, σ as well as $W_{X,\sigma}$ and $\mathcal{L}'_{X,\sigma}$ to be as in Corollary 8.2.

8.3 Step 3: A Global Set with Local Agreement

The next step is to further refine the set $\mathcal{L}'_{X,\sigma}$ so that the remaining subspaces "evenly cover" a subspace of $W^\star_{\mathsf{amb}} \subseteq W'_{\mathsf{amb}}$ with codimension $\dim(X) + O_{\delta_2}(1)$. To do this, we will reduce to the case where $\mathcal{L}'_{X,\sigma}$ is pseudo-random within some zoom-in A and zoom-out B such that $X \subseteq A \subseteq B$. This is done via the following argument, which can be informally described as:

1. While $\mathcal{L}'_{X,\sigma}$ is not pseudo-random, there must be some zoom-in or zoom-out on which its measure is significantly higher than its overall measure over the whole space, so consider the restriction to this zoom-in or zoom-out.

2. This increases the measure of $\mathcal{L}'_{X,\sigma}$, and we may repeat until we have a pseudo-random set (within some zoom-in, zoom-out combination).

3. By choosing the pseudo-randomness parameters suitably, we are able to perform the above process in relatively few times until the restriction of $\mathcal{L}'_{X,\sigma}$ that we arrive at is pseudo-random.

4. As a result, when restricting to the zoom-in, zoom-out combination, the resulting set of subspaces evenly covers a space that is still relatively large in the sense that it contains $q^{-O(\xi^{-1})}$-fraction of the present space (i.e. the space obtained after iteratively restricting to zoom-ins and zoom-outs as described in steps 1-3). In particular, this present space is described as $W_{\mathsf{amb},0}$ below.

We now move on to the formal statement. For a zoom-in A and zoom-out B such that $X \subseteq A \subseteq B \subseteq W'_{\mathsf{amb}}$, write $W'_{\mathsf{amb}} = A \oplus W_{\mathsf{amb},0}$ and $B = A \oplus W^\star_{\mathsf{amb}}$, where $W^\star_{\mathsf{amb}} \subseteq W_{\mathsf{amb},0}$. Also define

$$\mathcal{W}^\star_{[A,B]} = \{W^\star_i \mid \exists W_i \in \mathcal{W}_{X,\sigma} \text{ s.t } A \oplus W^\star_i = W_i \cap B\}.$$

It is clear that each $W^\star_i \in \mathcal{W}^\star_{[A,B]}$ is contained inside of some $W_i \in \mathcal{W}_{X,\sigma}$, so for each $W^\star_i$ we may associate a function $f^\star_i$ to each $W^\star_i$ as follows. Choose an arbitrary $W_i \in \mathcal{W}_{X,\sigma}$ such that $A \oplus W^\star_i = W_i \cap B$, and define $f^\star_i := f_i|_{W^\star_i}$.

Lemma 8.3. *Keeping the notation above, there is a zoom-in A and a zoom-out B inside of the space W'_{amb} such that the following holds. There exists a collection of subspaces $\mathcal{W}^\star = \{W^\star_1, \dots, W^\star_{m_3}\} \subseteq \mathcal{W}^\star_{[A,B]}$ of codimension s with respect to $W^\star_{\mathsf{amb}}$, a dimension $\ell' \geqslant \frac{\xi}{3}\ell$, and a set $\mathcal{L}^\star \subseteq \mathsf{Grass}_q(W^\star_{\mathsf{amb}}, \ell')$, such that the following items hold for the table $T_{\ell'}$, which assigns linear functions to $L \in \mathsf{Grass}_q(W^\star_{\mathsf{amb}}, \ell')$ as follows*

$$T_{\ell'}[L] :\equiv T[A \oplus L]|_L.$$

1. *$\mu(\mathcal{L}^\star) := \eta \geqslant \frac{C}{12}$, where the measure here is over $\mathsf{Grass}_q(W^\star_{\mathsf{amb}}, \ell')$.*

2. *The set $\mathcal{L}^\star$ is $(1, q^{\delta_2 \ell}\eta)$-pseudo-random.*

3. *Each $W^\star_i$ has codimension $s \leqslant r$ inside of $W^\star_{\mathsf{amb}}$, $\mathcal{W}^\star$ is 4-generic, with respect to $W^\star_{\mathsf{amb}}$, and $W^\star_{\mathsf{amb}}$ has codimension at most $10/\delta_2$ with respect to W'_{amb}.*

4. *The size of $\mathcal{W}^\star$ satisfies*
$$\frac{m_2 \cdot q^{-10s/\delta_2}}{2} \leqslant m_3 \leqslant m_2.$$

5. *For each $W_i^\star \in \mathcal{W}_i^\star$, there is a linear function $f_i^\star : W_i^\star \to \mathbb{F}_q$ such that the following holds. For every $L \in \mathcal{L}^\star$, choosing $W_i^\star \in \mathcal{W}^\star$ uniformly such that $W_i^\star \supseteq L$, we have*

$$\Pr_{W_i^\star \supseteq L, W_i^\star \in \mathcal{W}^\star}[f_i^\star|_L \not\equiv T_{\ell'}[L]] \leqslant 14\gamma.$$

6. *For every $L \in \mathcal{L}^\star$,*

$$0.8 \cdot p_1 \cdot m_3 \leqslant N_{\mathcal{W}^\star}(L) \leqslant 1.2 \cdot p_1 \cdot m_3,$$

where $N_{\mathcal{W}^\star}(L)$ is as defined in (13)*, and*

$$p_1 := \Pr_{L \in \mathsf{Grass}_q(W^\star_{\mathsf{amb}}, \ell')}[L \subseteq W],$$

for an arbitrary $W \subseteq W^\star_{\mathsf{amb}}$ of codimension s.

Proof. Deferred to Appendix E.2. □

As a consequence of pseudo-randomness, we may apply Lemma 2.4, to get that $\mathcal{L}^\star$ evenly covers $W^\star_{\mathsf{amb}}$.

Lemma 8.4. *Setting $Z := \{z \in W^\star_{\mathsf{amb}} \mid |\mu_{z,\circ}(\mathcal{L}^\star) - \eta| \leqslant \frac{\eta}{10}\}$, we have that,*

$$\frac{|Z|}{|W^\star_{\mathsf{amb}}|} \geqslant 1 - q^{-\frac{\ell'}{2}}.$$

Proof. This is immediate by the pseudo-randomness of $\mathcal{L}^\star$ and Lemma 2.4. □

8.4 Step 4: Local to Global Agreement

The following lemma establishes agreement, on average, between functions $f_i^\star, f_j^\star$. Specifically it shows that a random pair of functions agrees almost entirely in $W_i^\star \cap W_j^\star \cap Z$.

Lemma 8.5. *Let $\mathcal{W}^\star = \{W_1^\star, \ldots, W_{m_3}^\star\}$ and $f_1^\star, \ldots, f_{m_3}^\star$ be the 4-generic set of codimension $s \leqslant r \leqslant 10/\delta$ subspaces inside of W'_{amb} and associated linear functions obtained from Lemma E.3 respectively. We have*

$$\Pr_{\substack{W_i^\star, W_j^\star \in \mathcal{W}^\star \\ z \in W_i^\star \cap W_j^\star \cap Z}}[f_i^\star(z) \neq f_j^\star(z)] \leqslant 500\gamma,$$

and for every $W_i^\star, W_j^\star \in \mathcal{W}^\star$ we have $|W_i^\star \cap W_j^\star \cap Z| \geqslant 0.81 \cdot |W_i^\star \cap W_j^\star|$. Here $Z \subseteq W^\star_{\mathsf{amb}}$ is as defined in Lemma 8.4.

Proof. Deferred to Appendix E.3. □

Using Lemma 8.5 we conclude the proof of Lemma 5.24 by using ideas from the Raz-Safra analysis of the Plane versus Plane test [RS97]. Define a graph, G with vertex set $\mathcal{W}^\star$, where $W_i^\star, W_j^\star$ are adjacent if and only if $f_i^\star|_{W_i^\star \cap W_j^\star} = f_j^\star|_{W_i^\star \cap W_j^\star}$. We claim that this graph contains a large clique, and towards this end we first show that the graph is nearly transitive. For a graph $H = (V, E)$, define

$$\beta(H) = \max_{(u,w) \notin E} \Pr_{v \in V}[(v, u), (v, w) \in E].$$

A graph H is transitive if $\beta(H) = 0$. It is easy to see that transitive graphs are (edge) disjoint unions of cliques. The following lemma, proved in [RS97], asserts that if H is relatively dense and $\beta(H)$ is small, then one could remove only a small fraction of the edges and get a fully transitive graph.

Lemma 8.6. *[RS97, Lemma 2] Any graph H on n vertices can be made transitive by deleting at most $\lceil 3\sqrt{\beta(H)} \cdot n^2 \rceil$ edges.*

To use Lemma 8.6 we first show that the graph G we defined is highly transitive.

Claim 8.7. *We have $\beta(G) \leqslant \frac{1}{m_3}$.*

Proof. Fix a $W_i^\star, W_j^\star$ that are not adjacent. We claim that they can have at most 1 common neighbor. Suppose for the sake of contradiction that $W_a^\star, W_b^\star$ are distinct common neighbors. Then,

$$f_i^\star|_{W_i^\star \cap W_j^\star \cap W_a^\star} = f_j^\star|_{W_i^\star \cap W_j^\star \cap W_a^\star}, \qquad f_i^\star|_{W_i^\star \cap W_j^\star \cap W_b^\star} = f_j^\star|_{W_i^\star \cap W_j^\star \cap W_b^\star}.$$

It follows that $f_i^\star$ and $f_j^\star$ agree on $W_i^\star \cap W_j^\star \cap W_a^\star + W_i^\star \cap W_j^\star \cap W_b^\star$. However, since $\mathcal{W}^\star$ is 4-generic, we have

$$\operatorname{codim}(W_i^\star \cap W_j^\star \cap W_a^\star + W_i^\star \cap W_j^\star \cap W_b^\star) \leqslant 3s + 3s - 4s = 2s, \qquad \operatorname{codim}(W_i^\star \cap W_j^\star) = 2s,$$

and

$$W_i^\star \cap W_j^\star \cap W_a^\star + W_i^\star \cap W_j^\star \cap W_b^\star \subseteq W_i^\star \cap W_j^\star,$$

so it must be the case that $W_i^\star \cap W_j^\star \cap W_a^\star + W_i^\star \cap W_j^\star \cap W_b^\star = W_i^\star \cap W_j^\star$. This contradicts the assumption that $W_i^\star$ and $W_j^\star$ are not adjacent. Thus, any two non-adjacent vertices can have at most 1 common neighbor, and the result follows. $\square$

Claim 8.8. *The graph G contains a clique of size at least $\frac{m_3}{4}$*

Proof. Applying Markov's inequality to Lemma 8.5, we have that with probability at least $9/10$ over $W_i^\star$ and $W_j^\star$, we have both $\Pr_{z \in W_i^\star \cap W_j^\star \cap Z}[f_i^\star(z) \neq f_j^\star(z)] \leqslant 10001\gamma$ and $|W_i^\star \cap W_j^\star \cap Z| \geqslant 0.81 \cdot |W_i^\star \cap W_j^\star|$. In this case, $f_i^\star$ and $f_j^\star$ agree on at least $(1 - 10001\gamma)$-fraction of the points in $W_i^\star \cap W_j^\star \cap Z$, which is in turn at least $(1 - 10001\gamma) \cdot 0.81 > 1/q$-fraction of the points in $W_i^\star \cap W_j^\star$. As $f_i^\star$ and $f_j^\star$ are linear functions, the Schwartz-Zippel lemma implies that such $W_i^\star, W_j^\star$ are adjacent in G. Hence, G has at least $9\binom{m_3}{2}/10$ edges.

By Claim 8.7 and Lemma 8.6, we can delete $\lceil 3m_3^{3/2} \rceil$ edges to make G a union of cliques. Doing so yields a graph on m_3 vertices with at least $m_3^2/4$-edges that is a union of cliques. Let $C_1, \ldots, C_N$ be the cliques, with C_1 being the largest one. We have,

$$|C_1| \cdot m_3 \geqslant |C_1| \cdot \sum_{i=1}^{N} |C_i| \geqslant \sum_{i=1}^{N} |C_i|^2 \geqslant \frac{m_3^2}{4}.$$

It follows that $|C_1| \geqslant \frac{m_3}{4}$, and that G contains a clique of size at least $\frac{m_3}{4}$. $\square$

Let $\mathcal{C}$ be the clique guaranteed by Claim 8.8 and write $\mathcal{C} = \{W_1^\star, \ldots, W_{\frac{m_3}{4}}^\star\}$. To complete the proof of Lemma 5.24 we will find a linear h, such that for all $1 \leqslant i \leqslant \frac{m_3}{4}$, $f_i^\star|_{V^\star} \equiv h|_{W^\star_{\mathsf{amb}}}$, and then show that this h can be extended to $X \oplus A \oplus V^\star$ in a manner that is consistent with many of the *original* f_i's for $1 \leqslant i \leqslant m_3/4$. To this end, first define $g : V^\star \to \mathbb{F}_q$ as follows:

$$g(x) = \begin{cases} f_i^\star(x), & \text{if } \exists W_i^\star \in \mathcal{C}, x \in W_i^\star \\ 0, & \text{otherwise.} \end{cases} \tag{29}$$

Since $f_i^\star(x) = f_j^\star(x)$ whenever $x \in W_i^\star \cap W_j^\star$, it does not matter which i is chosen if there are multiple $W_i^\star \in \mathcal{C}$ containing x. Thus, g is well defined and $g|_{W_i^\star} = f_i^\star$ for all $1 \leqslant i \leqslant \frac{m_3}{4}$.

We next show that g is close to a linear function and that this linear function agrees with most of the functions $f_i^\star|_{W_i^\star}$ for $W_i^\star \in \mathcal{C}$. To begin, we show that g passes the standard linearity test with high probability.

Lemma 8.9. *We have,*

$$\Pr_{z_1,z_2 \in W^\star_{\mathsf{amb}}}[g(z_1+z_2) = g(z_1) + g(z_2)] \geqslant 1 - \frac{25q^{2s}}{m_3}.$$

Proof. Note that we have

$$\Pr_{z_1,z_2 \in W^\star_{\mathsf{amb}}}[g(z_1+z_2) = g(z_1) + g(z_2)] \geqslant \Pr_{z_1,z_2 \in W^\star_{\mathsf{amb}}}[\exists W_i^\star \in \mathcal{C},\ \text{s.t. },z_1, z_2 \in W_i^\star].$$

For every $z_1, z_2 \in W^\star_{\mathsf{amb}}$ linearly independent, we can let $N(z_1, z_2)$ denote the number of $W_i^\star \in \mathcal{C}$ containing $\mathrm{span}(z_1, z_2)$. The result them follows from Lemma 5.21 with $a = 0$, $j = 2$, $r = s$, $c = 0.99$, and using the fact that $|\mathcal{C}| = m_3/4$. Indeed,

$$\begin{aligned} Pr_{z_1,z_2 \in W^\star_{\mathsf{amb}}}[\exists W_i^\star \in \mathcal{C},\ \text{s.t. },z_1, z_2 \in W_i^\star] &\geqslant \Pr_{z_1,z_2 \in W^\star_{\mathsf{amb}}}[N(z_1,z_2) > 0 \mid \dim(\mathrm{span}(z_1,z_2)) = 2] \\ &\geqslant 1 - \frac{25q^{2s}}{m_3}. \end{aligned}$$

In the first transition, we use that if z_1, z_2 are linearly dependent then $g(z_1+z_2) = g(z_1) + g(z_2)$. $\square$

We now apply the linearity testing result of Blum, Luby, and Rubinfeld [BLR93] and get that g is $\frac{25q^{2s}}{m_3}$-close to a linear function, say $h : W^\star_{\mathsf{amb}} \to \mathbb{F}_q$.

Theorem 8.10. *[BLR93] Suppose $g : W^\star_{\mathsf{amb}} \to \mathbb{F}_q$ satisfies, $\Pr_{z_1,z_2}[g(z_1+z_2) = g(z_1)+g(z_2)] \geqslant 1-\rho$, for some $\rho < 2/9$. Then, there exists a linear function $h : W^\star_{\mathsf{amb}} \to \mathbb{F}_q$ such that $\Pr_{z \in W^\star_{\mathsf{amb}}}[h(z) \neq g(z)] \leqslant 2\rho$.*

We will conclude by showing that this h is the desired function which agrees with many of the original f_i's. To this end, we first show that it agrees with many of the $f_i^\star$'s that we have (which themselves are restrictions of the original f_i's), and then show that h can be extended to W'_{amb} in a manner that retains agreement with many of the f_i's.

Towards the first step, set $S = \{x \in W^\star_{\mathsf{amb}} \mid g(x) \neq h(x)\}$. We show that choosing $W_i^\star \in \mathcal{C}$ randomly, and then a point $x \in W_i^\star$, it is unlikely that $x \in S$. Define the measure $\nu_{\mathcal{C}}$ over nonzero points in $W^\star_{\mathsf{amb}}$ obtained by choosing $W_i \in \mathcal{C}$ uniformly at random and then $x \in W_i$ nonzero uniformly at random. Let μ be the uniform measure over $W^\star_{\mathsf{amb}}$, so $\mu(S) \leqslant \frac{25q^{2s}}{m_3}$. Then $\nu_{\mathcal{C}}(S)$ is precisely the probability of interest and can be upper bounded using Lemma 5.18, with parameters $a = 0, j = 1$, codimension s,

$$\nu_{\mathcal{C}}(S) \leqslant \mu(S) + \frac{25q^{\frac{s}{2}}}{\sqrt{m_3}} \leqslant \frac{26q^{\frac{s}{2}}}{\sqrt{m_3}}. \tag{30}$$

Lemma 8.11. *We have $h|_{W_i^\star} \equiv f_i^\star$ for at least half of the $W_i^\star \in \mathcal{C}$.*

Proof. By Markov's inequality and (30) with probability at least $1/2$, over $W_i^\star \in \mathcal{C}$, we have

$$\frac{|W_i^\star \cap S|}{|W_i^\star|} \leqslant \frac{52q^{\frac{s}{2}}}{\sqrt{m_3}} < 1 - \frac{1}{q},$$

and $f_i^\star$ and $h|_{W_i^\star}$ agree on more than $1/q$ of the points in $W_i^\star$. Since $f_i^\star$ and $h|_{W_i^\star}$ are both linear, by the Schwartz-Zippel Lemma that $h|_{W_i^\star} \equiv f_i^\star$, and the result follows. □

We are now ready to finish the proof of Lemma 5.24.

Proof of Lemma 5.24. Summarizing, we now have linear functions $f_i^\star : W_i^\star \to \mathbb{F}_q$ for $1 \leqslant i \leqslant \frac{m_3}{8}$ and a linear function $h : W_{\mathsf{amb}}^\star \to \mathbb{F}_q$ such that $h|_{W_i^\star} = f_i^\star$. Furthermore, for each $f_i^\star, W_i^\star$, there is a f_i, W_i from Lemma 5.24 such that $W_i \cap W_{\mathsf{amb}}^\star = W_i^\star$, $W_i \subseteq W'_{\mathsf{amb}}$, $f_i|_{W_i^\star} = f_i^\star$, and $f_i|_X = \sigma$.

We extend h in a manner so that it agrees with many of these original functions f_i. To this end, recall that $W_{\mathsf{amb}}^\star$ satisfies,

$$A \oplus W_{\mathsf{amb}}^\star = B \subseteq W'_{\mathsf{amb}}.$$

and $\dim(A) + \operatorname{codim}(B) - \dim(X) \leqslant \frac{10}{\delta_2}$. Therefore, we may choose a random linear function $h' : W'_{\mathsf{amb}} \to \mathbb{F}_q$ conditioned on $h'|_{W_{\mathsf{amb}}^\star} \equiv h$ and $h'|_X \equiv \sigma$. For any f_i, we have that

$$\Pr_{h'}[h'|_{W_i} \equiv f_i] \geqslant q^{-(\dim(A)+\operatorname{codim}(B)-\dim(X))} \geqslant q^{-\frac{10}{\delta_2}}.$$

Indeed there is a $q^{-(\dim(A)-\dim(X))}$ probability that $h|_A \equiv f_i|_A$, as we condition on $h'|_X \equiv \sigma \equiv f_i|_X$. Then, extending h' from B to W'_{amb}, there is at least a $q^{-\operatorname{codim}(B)}$ probability that h' is equal to f_i on these extra dimensions. It follows that there is a linear $h' : W'_{\mathsf{amb}} \to \mathbb{F}_q$ such that $h'|_{W_i} \equiv f_i$ for at least

$$\frac{m_3 q^{-\frac{10}{\delta_2}}}{8} \geqslant q^{50r\ell\xi^{-1}}$$

of the pairs f_i, W_i from Lemma 5.24. Take these W_i to be the set $\mathcal{W}'$ for Lemma 5.24. As $\mathcal{W}' \subseteq \mathcal{W}$, they are 2-generic with respect to W'_{amb} and have codimension $s \leqslant r$ in W'_{amb}. □

9 Acknowledgments

We sincerely thank anonymous reviewers for their careful reading of the manuscript and their detailed comments, which greatly improved the writing of the paper.

References

[ABH+05] Sanjeev Arora, Eli Berger, Elad Hazan, Guy Kindler, and Muli Safra. On non-approximability for quadratic programs. In *FOCS 2005*, pages 206–215. IEEE Computer Society, 2005.

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, 1998.

[Alo21] Noga Alon. Explicit expanders of every degree and size. *Combinatorica*, pages 1–17, 2021.

[AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, 1998.

[BDN17] Amey Bhangale, Irit Dinur, and Inbal Livni Navon. Cube vs. cube low degree test. In *ITCS 2017*, volume 67 of *LIPIcs*, pages 40:1–40:31. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017.

[Ber00] Piotr Berman. A $d/2$ approximation for maximum weight independent set in d-claw free graphs. *Nordic Journal of Computing*, 7(3):178–184, 2000.

[BKS19] Boaz Barak, Pravesh K. Kothari, and David Steurer. Small-set expansion in shortcode graph and the 2-to-2 conjecture. In *ITCS 2019*, volume 124 of *LIPIcs*, pages 9:1–9:12. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

[BL15] Abhishek Bhowmick and Shachar Lovett. The list decoding radius of Reed-Muller codes over small fields. In *STOC 2015*, pages 277–285. ACM, 2015.

[BLR93] Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. *J. Comput. Syst. Sci.*, 47(3):549–595, 1993.

[CCK08] Tanmoy Chakraborty, Julia Chuzhoy, and Sanjeev Khanna. Network design for vertex connectivity. In *STOC 2008*, pages 167–176. ACM, 2008.

[CDGT22] Julia Chuzhoy, Mina Dalirrooyfard, Vadim Grinberg, and Zihan Tan. A new conjecture on hardness of low-degree 2-CSP's with implications to hardness of densest k-subgraph and other problems. *CoRR*, abs/2211.05906, 2022.

[CGM13] Marek Cygan, Fabrizio Grandoni, and Monaldo Mastrolilli. How to sell hyperedges: The hypermatching assignment problem. In *SODA 2013*, pages 342–351. SIAM, 2013.

[Cha16] Siu On Chan. Approximation resistance from pairwise-independent subgroups. *J. ACM*, 63(3):1–32, 2016.

[CW04] Moses Charikar and Anthony Wirth. Maximizing quadratic programs: Extending Grothendieck's inequality. In *FOCS 2004*, pages 54–60. IEEE Computer Society, 2004.

[DFRR23] Pavel Dvořák, Andreas Emil Feldmann, Ashutosh Rai, and Paweł Rzążewski. Parameterized inapproximability of independent set in h-free graphs. *Algorithmica*, 85(4):902–928, 2023.

[DKK+18] Irit Dinur, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. Towards a proof of the 2-to-1 games conjecture? In *STOC 2018*, pages 376–389. ACM, 2018.

[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.

[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 *STOC 2023*, pages 656–660. ACM, 2023.

[EKL24] Shai Evra, Guy Kindler, and Noam Lifshitz. Polynomial Bogolyubov for special linear groups via tensor rank, 2024. arXiv:2404.00641.

[Fei98] Uriel Feige. A threshold of $\ln n$ for approximating set cover. *J. 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. *J. ACM*, 43(2):268–292, 1996.

[Gop13] Parikshit Gopalan. A Fourier-analytic approach to Reed-Muller decoding. *IEEE Trans. Inf. Theory*, 59(11):7747–7760, 2013.

[GRS00] Oded Goldreich, Ronitt Rubinfeld, and Madhu Sudan. Learning polynomials with queries: The highly noisy case. *SIAM J. Discret. Math.*, 13(4):535–570, 2000.

[Hås96] Johan Håstad. Clique is hard to approximate within $n^{1-\varepsilon}$. In *FOCS 1996*, pages 627–636. IEEE Computer Society, 1996.

[Hås01] Johan Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001.

[IKW12] Russell Impagliazzo, Valentine Kabanets, and Avi Wigderson. New direct-product testers and 2-query PCPs. *SIAM J. Comput.*, 41(6):1722–1768, 2012.

[KMS17] Subhash Khot, Dor Minzer, and Muli Safra. On independent sets, 2-to-2 games, and Grassmann graphs. In *STOC 2017*, pages 576–589. ACM, 2017.

[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.

[KP06] Subhash Khot and Ashok Kumar Ponnuswami. Better inapproximability results for maxclique, chromatic number and min-3lin-deletion. In *ICALP 2006*, volume 4051 of *Lecture Notes in Computer Science*, pages 226–237. Springer, 2006.

[KR03] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within $2-\varepsilon$. In *CCC 2003*, page 379. IEEE Computer Society, 2003.

[KS13] Subhash Khot and Muli Safra. A two-prover one-round game with strong soundness. *Theory Comput.*, 9:863–887, 2013.

[Lae14] Bundit Laekhanukit. Parameters of two-prover-one-round game and the hardness of connectivity problems. In *SODA 2014*, pages 1626–1643. SIAM, 2014.

[LM24] Euiwoong Lee and Pasin Manurangsi. Hardness of approximating bounded-degree max 2-CSP and independent set on k-claw-free graphs. In *ITCS 2024*, pages 71–1, 2024.

[Man19] Pasin Manurangsi. A note on degree vs gap of min-rep label cover and improved inapproximability for connectivity problems. *Inf. Process. Lett.*, 145:24–29, 2019.

[Meg01] Alexandre Megretski. Relaxations of quadratic programs in operator theory and system analysis. In *IWOTA 2000*, pages 365–392. Springer, 2001.

[Min22] Dor Minzer. *On Monotonicity Testing and the 2-to-2 Games Conjecture*. Morgan & Claypool, 2022.

[Mos19] Dana Moshkovitz. Sliding scale conjectures in PCP. *SIGACT News*, 50(3):25–33, 2019.

[MR10] Dana Moshkovitz and Ran Raz. Two-query PCP with subconstant error. *J. ACM*, 57(5):29:1–29:29, 2010.

[MZ23] Dor Minzer and Kai Zheng. Approaching the soundness barrier: A near optimal analysis of the cube versus cube test. In *SODA 2023*, pages 2761–2776. SIAM, 2023.

[NRT99] Arkadi Nemirovski, Cornelis Roos, and Tamás Terlaky. On maximization of quadratic form over intersection of ellipsoids with common center. *Mathematical programming*, 86(3):463–473, 1999.

[Nut12] Zeev Nutov. Approximating minimum-cost connectivity problems via uncrossable bifamilies. *ACM Transactions on Algorithms (TALG)*, 9(1):1–16, 2012.

[Rao11] Anup Rao. Parallel repetition in projection games and a concentration bound. *SIAM J. Comput.*, 40(6):1871–1891, 2011.

[Raz98] Ran Raz. A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998.

[RS97] Ran Raz and Shmuel Safra. A sub-constant error-probability low-degree test, and a subconstant error-probability PCP characterization of NP. In *STOC 1997*, pages 475–484, 1997.

[Spi19] Daniel Spielman. Spectral and algebraic graph theory. *Yale lecture notes, draft of December*, 4:47, 2019.

[Tre14] Luca Trevisan. Inapproximability of combinatorial optimization problems. *Paradigms of Combinatorial Optimization: Problems and New Approaches*, pages 381–434, 2014.

[TW23] Theophile Thiery and Justin Ward. An improved approximation for maximum weighted k-set packing. In *SODA 2023*, pages 1138–1162. SIAM, 2023.

A Proofs of Lemmas 2.3 and 2.4

In this section we prove Lemmas 2.3 and 2.4. The proofs of these lemmas requires tools from [EKL23, EKL24] regarding Fourier analysis over the Bilinear Scheme.

A.1 Fourier Analysis over the Bilinear Scheme

The key to proving Lemma 2.3 is a level-d inequality for indicator functions on the Bilinear Scheme due to Evra, Kindler, and Lifshitz [EKL24]. In order to use the result of [EKL24], however, we first give some necessary background for Fourier analysis over the Bilinear Scheme, and describe the analogues of zoom-ins, zoom-outs, and pseudo-randomness as in [EKL23, EKL24]. After doing so, we must then find a suitable map from the Grassmann graph to the Bilinear Scheme that (1) preserves the edges of our original bipartite inclusion graph between 2ℓ-dimensional and $2(1-\delta)\ell$ subspaces, and (2) maps zoom-ins and zoom-outs in the Grassmann graph to their analogues over the Bilinear Scheme.

The Bilinear Scheme: Let $\mathbb{F}_q^{n\times 2\ell}$ be the set of $n\times 2\ell$ matrices over $\mathbb{F}_q$. One can define a graph over $\mathbb{F}_q^{n\times 2\ell}$ that is similar to the Grassmann graphs by calling $M_1, M_2 \in \mathbb{F}_q^{n\times 2\ell}$ adjacent if $\dim(\ker(M_1 - M_2)) \leqslant s$ for some $s \leqslant 2\ell$. Such graphs are often referred to as the *Bilinear Scheme*. For our purposes, we will need to work with a bipartite version of this graph between $\mathbb{F}_q^{(n-2\ell)\times 2\ell}$ and $\mathbb{F}_q^{(n-2(1-\delta)\ell)\times 2(1-\delta)\ell}$.

We equip the space $L_2(\mathbb{F}_q^{n\times 2\ell})$ with the following inner product:

$$\langle F, G\rangle = \mathop{\mathbb{E}}_{M\in\mathbb{F}_q^{n\times 2\ell}}[F(M)\overline{G(M)}],$$

where the distribution taken over M is uniform. Let ω be a primitive pth root of unity, where recall p is the characteristic of $\mathbb{F}_q$. For $s \in \mathbb{F}_q^n$ and $x \in \mathbb{F}_q^n$, let $\chi_s(x) = \omega^{\mathrm{Tr}(s\cdot x)}$ where $\mathrm{Tr}\colon \mathbb{F}_q \to \mathbb{F}_p$ is the trace map given by $\mathrm{Tr}(\gamma) = \sum_{i=0}^{\log_p(q)-1} \gamma^{p^i}$. A common fact that we will used is that the trace is additive $\mathrm{Tr}(\gamma+\gamma') = \mathrm{Tr}(\gamma)+\mathrm{Tr}(\gamma')$ for $\gamma, \gamma' \in \mathbb{F}_q$. The characters $\{\chi_S : \mathbb{F}_q^{n\times 2\ell} \to \mathbb{C} \mid S = (s_1, \dots, s_{2\ell}) \in \mathbb{F}_q^{n\times 2\ell}\}$, given by

$$\chi_S(x_1, \dots, x_{2\ell}) = \prod_{i=1}^{2\ell} \chi_{s_i}(x_i) = \omega^{\sum_{i=1}^{2\ell} \mathrm{Tr}(s_i\cdot x_i)} = \omega^{\mathrm{Tr}\left(\sum_{i=1}^{2\ell}(s_i\cdot x_i)\right)}, \tag{31}$$

form an orthonormal basis of $L_2(\mathbb{F}_q^{n\times 2\ell})$. As a result, any $F \in L_2(\mathbb{F}_q^{n\times 2\ell})$ can be expressed as,

$$F = \sum_{S\in\mathbb{F}_q^{n\times 2\ell}} \widehat{F}(S)\chi_S,$$

where $\widehat{F}(S) = \langle F, \chi_S\rangle$. The level d component of F is given by $F^{=d} = \sum_{S\colon \mathrm{rank}(S)=d} \widehat{F}(S)\chi_S$. If a function F only consists of components up to level d, i.e. $\widehat{F}(S) = 0$ for all S with $\mathrm{rank}(S) > d$, then we say F is of degree d.

We now describe the analogues of zoom-ins and zoom-outs on $\mathbb{F}_q^{n\times 2\ell}$.

Definition A.1. *A zoom-in of dimension d over $\mathbb{F}_q^{n\times 2\ell}$ is given by d-pairs of vectors $(u_1, v_1), \dots, (u_r, v_r)$ where each $u_i \in \mathbb{F}_q^{2\ell}$ and each $v_i \in \mathbb{F}_q^n$. Let $U \in \mathbb{F}_q^{2\ell\times d}$ and $V \in \mathbb{F}_q^{n\times d}$ denote the matrices whose ith columns are u_i and v_i respectively. Then the zoom-in on (U, V) is the set of $M \in \mathbb{F}_q^{n\times 2\ell}$ such that $MU = V$, or equivalently, $Mu_i = v_i$ for $1 \leqslant i \leqslant d$.*

Next, we define the analog of zoom-outs.

Definition A.2. *A zoom-out of dimension d is defined similarly, except by multiplication on the left. Given $X \in \mathbb{F}_q^{d\times n}$ and $Y \in \mathbb{F}_q^{d\times 2\ell}$, whose rows are given by x_i and y_i respectively, the zoom-out (X, Y) is the $M \in \mathbb{F}_q^{n\times 2\ell}$ such that $XM = Y$, or equivalently, $x_iM = y_i$ for $1 \leqslant i \leqslant d$.*

Let $\mathsf{Zoom}[(U, V), (X, Y)]$ denote the intersections of the zoom-in on (U, V) and the zoom-out on (X, Y). The codimension of $\mathsf{Zoom}[(U, V), (X, Y)]$ is the sum of the number of columns of U and the number of rows of X, which we will denote by $\dim(U)$ and $\mathrm{codim}(X)$. For a zoom-in and zoom-out pair and a Boolean function F, we define $F_{(U,V),(X,Y)} : \mathsf{Zoom}[(U, V), (X, Y)] \to \{0, 1\}$ to be the restriction of F which is given as

$$F_{(U,V),(X,Y)}(M) = F(M) \quad \text{for} \quad M \in \mathsf{Zoom}[(U, V), (X, Y)].$$

When $\dim(U)+\operatorname{codim}(X) = d$, we say that the restriction is of size d. We define (d,ε)-pseudo-randomness in terms of the L_2-norms of restrictions of F of size d. Here and throughout, when we consider restricted functions, the underlying measure is the uniform measure over the corresponding zoom-in and zoom-out set $\mathsf{Zoom}[(U,V),(X,Y)]$.

Definition A.3. *We say that an indicator function $F \in L_2(\mathbb{F}_q^{n\times 2\ell})$ is* (d,ε)-pseudo-random *if for all zoom-in zoom-out combinations* $\mathsf{Zoom}[(U,V),(X,Y)]$ *such that* $\dim(U)+\operatorname{codim}(X) = d$*, we have*

$$\left\|F_{(U,V),(X,Y)}\right\|_2^2 \leqslant \varepsilon.$$

We note that for Boolean functions F, $\left\|F_{(U,V),(X,Y)}\right\|_2^2 = \mathbb{E}_{M\in\mathsf{Zoom}[(U,V),(X,Y)]}[F(M)]$, and hence the definition above generalizes the definition we have for Boolean functions.

Definition A.4. *We say that an indicator function $F \in L_2(\mathbb{F}_q^{n\times 2\ell})$ is* (d,ε,t)-pseudo-random *if for all* $\mathsf{Zoom}[(U,V),(X,Y)]$ *such that* $\dim(U)+\operatorname{codim}(X) = d$*, we have,*

$$\left\|F_{(U,V),(X,Y)}\right\|_{t/(t-1)} = \left(\mathop{\mathbb{E}}_{M\in\mathsf{Zoom}[(U,V),(X,Y)]}\left[|F(M)|^{\frac{t}{t-1}}\right]\right)^{\frac{t-1}{t}} \leqslant \varepsilon.$$

The following result is a combination of two results form [EKL24]. It states that if a Boolean function F is (r,ε)-pseudo-random, then its degree d parts are $(r, C_{q,d}\varepsilon^2)$-pseudo-random for $d \leqslant r$.

Lemma A.5. *[EKL24, Theorem 5.5 + Proposition 3.6] Let $t \geqslant 4$ be a power of 2 and let $F : \mathbb{F}_q^{n\times 2\ell} \to \{0,1\}$ be a function that is (d,ε,t)-pseudo-random. Then $F^{=d}$ is $(r, q^{10dr+500d^2t}\varepsilon^2)$-pseudo-random for all $r \geqslant d$.*

Proof. This lemma does not actually appear in [EKL24], but it is easy to derive by combining Theorem 5.5 with Proposition 3.6 therein. In [EKL23, EKL24], the authors introduce an additional notion of generalized influences and having small generalized influences. We refrain from defining these notions explicitly as it is slightly cumbersome, but roughly speaking, one defines a Laplacian for each zoom-in, zoom-out combination, so that having (d,ε) small generalized influences means that upon applying these Laplacians on F, the 2-norm squared of the resulting function never exceeds ε.

With this notion in hand, if a function F is (d,ε,t)-pseudo-random, then by [EKL24, Theorem 5.5] we get that $F^{=d}$ has $(d, q^{500d^2t}\varepsilon^2)$-small generalized influences. Applying [EKL24, Proposition 3.6] then implies that $F^{=d}$ is $(r, q^{10dr}\cdot q^{500d^2t}\varepsilon^2)$-pseudo-random for any $r \geqslant d$, which is the desired result. $\square$

Lastly, we need the following global hypercontractivity result also due to [EKL24].[10]

Theorem A.6. *[EKL24, Theorem 1.13] Let $t \geqslant 4$ be a power of 2 and let $F \in L_2(\mathbb{F}_q^{n\times 2\ell})$ be a function of degree d that is (d,ε)-pseudo-random. Then,*

$$\|F\|_t^t \leqslant q^{200d^2t^2}\,\|F\|_2^2\,\varepsilon^{t/2-1}.$$

Combining Lemma A.5 and Theorem A.6, we arrive at the following result which bounds the t-norm of the level d component of pseudo-random indicator functions. This result will be the key to showing an analogue of Lemma 2.3 over the Bilinear Scheme.

[10] We remark that earlier results [EKL23] showed similar statement for 4-norms, i.e. the case that $t = 4$, and the result below follows by a form of induction on t. That is, one starts with F and concludes via applying the case $t = 4$ that the function and it F^2 is $(d, C_{q,d}\varepsilon)$-pseudo-random. Then one apply the case $t = 4$ on F^2 to conclude that F^4 is $(d, C'_{q,d}\varepsilon)$-pseudo-random and so on.

Theorem A.7. *Let $t \geqslant 4$ be a power of 2. Then if $F : \mathbb{F}_q^{n\times 2\ell} \to \{0,1\}$ is (r,ε)-pseudo-random, we have*

$$\left\|F^{=d}\right\|_t \leqslant q^{500d^2t}\varepsilon^{\frac{t-2}{t}}$$

for all $d \leqslant r$.

Proof. Suppose F is (r,ε)-pseudo-random, let $t \geqslant 4$ be a power of 2, and fix a $d \leqslant r$. Since $d \leqslant r$, we also have that F is (d,ε)-pseudorandom. Therefore for any size d restriction of F, $F_{(U,V),(X,Y)}$,we have,

$$\|F_{(U,V),(X,Y)}\|_{t/(t-1)} = \left(\|F_{(U,V),(X,Y)}\|_2^2\right)^{\frac{t-1}{t}} \leqslant \varepsilon^{\frac{t-1}{t}}.$$

Thus, F is $(d, \varepsilon^{\frac{t-1}{t}}, t)$-pseudo-random, and by Lemma A.5 it follows that $F^{=d}$ is $(d, q^{10d^2+500d^2t}\varepsilon^{\frac{2t-2}{t}})$-pseudo-random. Clearly, $F^{=d}$ is degree d, so applying Theorem A.6 we get that

$$\left\|F^{=d}\right\|_t^t \leqslant q^{200d^2t^2}\left\|F^{=d}\right\|_2^2\left(q^{10d^2+500d^2t}\varepsilon^{\frac{2t-2}{t}}\right)^{t/2-1} \leqslant q^{500d^2t^2}\varepsilon^{t-2},$$

where we also use the fact that $\left\|F^{=d}\right\|_2^2 \leqslant \|F\|_2^2 \leqslant \varepsilon$ because F is an (r,ε)-pseudo-random Boolean function. Taking the t-th root of the above inequality completes the proof. □

A.2 An Analog of Lemma 2.3 for the Bilinear Scheme

In this section, we will build up some Fourier analytic tools over the bilinear scheme towards showing Lemma 2.3. Specifically, our goal will be to transform the statement from Lemma 2.3 into a statement about indicator functions over the bilinear scheme, and then apply Theorem A.6.

To this end, we start with some definitions. Say that a function $F \in L_2(\mathbb{F}_q^{n\times 2\ell})$ is *basis invariant* if it satisfies $F(M) = F(MA)$ for any full rank $A \in \mathbb{F}_q^{2\ell\times 2\ell}$. Basis invariant indicator functions are meant to correspond to indicator functions for sets of subspaces in $\mathsf{Grass}_q(n, 2\ell)$. Indeed, by definition, if two M_1 and M_2 have the same column image, then $F(M_1) = F(M_2)$. Define the adjacency operator $\mathcal{T} : L_2\left(\mathbb{F}_q^{n\times 2\ell}\right) \to L_2\left(\mathbb{F}_q^{n\times 2(1-\delta)\ell}\right)$ as follows. For any $H \in L_2(\mathbb{F}_q^{n\times 2\ell})$, the function $\mathcal{T}H : \mathbb{F}_q^{n\times 2(1-\delta)\ell} \to \mathbb{C}$ is given by

$$\mathcal{T}F(M) = \mathop{\mathbb{E}}_{v_1,\dots,v_{2\delta\ell}}[F([M, v_1, \dots, v_{2\delta\ell}])].$$

Here, $[M, v_1, \dots, v_{2\delta\ell}]$ refers to the matrix obtained by adding the columns v_i to M on the right, for all $i \in [2\delta\ell]$. In words, the operator $\mathcal{T}$ averages over extensions of the matrix M to an $n \times 2\ell$ matrix by adding to it $2\delta\ell$ random columns. This operator is morally the adjacency operator for the bipartite Grassmann graph between $\mathsf{Grass}_q(n, 2\ell)$ and $\mathsf{Grass}_q(n, 2(1-\delta)\ell)$, and as such, it will ultimately help us transform the probability of interest in Lemma 2.3, into an inner product over the bilinear scheme.

Before jumping into Lemma 2.3 though, we must first show some lemmas about basis invariant functions and the operator $\mathcal{T}$. The first lemma we need is that the level d component of a basis invariant function is also basis invariant, and to show this, the following two identities will be useful.

Lemma A.8. *For any $S = (s_1, \dots, s_{\ell'}) \in \mathbb{F}_q^{n\times \ell'}$, any $M \in \mathbb{F}_q^{n\times 2\ell}$, and any matrix $A \in \mathbb{F}_q^{2\ell\times \ell'}$ we have,*

$$\chi_S(MA) = \chi_{SA^T}(M).$$

Proof. Letting $v_1, \dots, v_{2\ell}$ denote the columns of M and $a_{i,j}$ denote the entries of A, we have,

$$\chi_S(MA) = \omega^{\sum_{i=1}^{\ell'} \mathrm{Tr}\left(s_i \cdot \left(\sum_{j=1}^{2\ell} v_j a_{j,i}\right)\right)} = \omega^{\sum_{j=1}^{2\ell} \mathrm{Tr}\left(v_j \cdot \left(\sum_{i=1}^{\ell'} s_i a_{j,i}\right)\right)} = \chi_{SA^T}(M). \qquad \square$$

Lemma A.9. *Let $S = (s_1, \dots, s_{2\ell}) \in \mathbb{F}_q^{n \times 2\ell}$ and let $F \in L_2\left(\mathbb{F}_q^{n\times 2\ell}\right)$ be basis invariant. Then for any $A \in \mathbb{F}_q^{2\ell \times 2\ell}$ that is full rank, we have $\widehat{F}(SA) = \widehat{F}(S)$.*

Proof. For any matrix full rank $B \in \mathbb{F}_q^{2\ell\times 2\ell}$ we have

$$\begin{aligned}
\widehat{F}(S) &= \mathop{\mathbb{E}}_{M \in \mathbb{F}_q^{n\times 2\ell}} \left[\overline{\chi_S(M)} F(M)\right] \\
&= \mathop{\mathbb{E}}_{M \in \mathbb{F}_q^{n\times 2\ell}} \left[\overline{\chi_S(M)} F(MB^{-1})\right] \\
&= \mathop{\mathbb{E}}_{M \in \mathbb{F}_q^{n\times 2\ell}} \left[\overline{\chi_S(MB)} F(M)\right] \\
&= E_{M \in \mathbb{F}_q^{n\times 2\ell}}[\overline{\chi_{SB^T}(M)} F(M)] \\
&= \widehat{F}(SB^T),
\end{aligned}$$

where we use that F is basis invariant in the third transition and Lemma A.8 in the fourth transition. Setting $B = A^T$ gives the result. $\square$

Using Lemma A.9, we can show that the level d component of a basis invariant function is also basis invariant.

Lemma A.10. *If $F \in L_2\left(\mathbb{F}_q^{n\times 2\ell}\right)$ is basis invariant, then $F^{=d}$ is basis invariant as well for any d.*

Proof. Fix any $M \in \mathbb{F}_q^{n\times 2\ell}$ and $A \in \mathbb{F}_q^{2\ell\times 2\ell}$ full rank. We have

$$\begin{aligned}
F^{=d}(MA) &= \sum_{S \in \mathbb{F}_q^{n\times 2\ell}, \mathrm{rank}(S)=d} \widehat{F}(S)\chi_S(MA) \\
&= \sum_{S \in \mathbb{F}_q^{n\times 2\ell}, \mathrm{rank}(S)=d} \widehat{F}(S)\chi_{SA^T}(M) \\
&= \sum_{S \in \mathbb{F}_q^{n\times 2\ell}, \mathrm{rank}(S)=d} \widehat{F}(S(A^T)^{-1})\chi_S(M) \\
&= \sum_{S \in \mathbb{F}_q^{n\times 2\ell}, \mathrm{rank}(S)=d} \widehat{F}(S)\chi_S(M) \\
&= F^{=d}(M),
\end{aligned}$$

where we use Lemma A.8 in the second transition, and Lemma A.9 in the fourth transition. $\square$

Next, we will define another operator $\mathcal{G} : L_2\left(\mathbb{F}_q^{n\times 2(1-\delta)\ell}\right) \to L_2\left(\mathbb{F}_q^{n\times 2\ell}\right)$, which acts like the adjoint of $\mathcal{T}$ for basis invariant functions. The reason that we work with $\mathcal{G}$ and not the actual adjoint of $\mathcal{T}$ is that we want to define this operator in a specific, convenient way so that we can ultimately study what the composed operator $\mathcal{GT}$ looks like. The operator $\mathcal{G}$ is given by

$$\mathcal{G}H(M) = \mathop{\mathbb{E}}_{A \in \mathbb{F}_q^{2\ell \times 2(1-\delta)\ell}} [H(MA) \mid \operatorname{rank}(A) = 2(1-\delta)\ell].$$

The next lemma shows that $\mathcal{G}$ indeed acts like the adjoint of $\mathcal{T}$ for basis invariant functions.

Lemma A.11. *For $F \in L_2\left(\mathbb{F}_q^{n \times 2\ell}\right)$ that is basis invariant and $G \in L_2\left(\mathbb{F}_q^{n \times 2(1-\delta)\ell}\right)$, we have*

$$\langle \mathcal{T}F, G \rangle = \langle F, \mathcal{G}H \rangle.$$

Proof. Let $J \in \mathbb{F}_q^{2\ell \times 2(1-\delta)\ell}$ be the matrix whose restriction to the first $2(1-\delta)\ell$ rows is the identity matrix $I_{2(1-\delta)\ell \times 2(1-\delta)\ell}$ and whose remaining rows are all 0. We have

$$\begin{aligned}
\langle \mathcal{T}F, H \rangle &= \mathop{\mathbb{E}}_{M' \in \mathbb{F}_q^{n \times 2(1-\delta)\ell}, v_1, \ldots, v_{2\delta\ell} \in \mathbb{F}_q^n} \left[F\left([M', v_1, \ldots, v_{2\delta\ell}]\right) \cdot \overline{H(M')} \right] \\
&= \mathop{\mathbb{E}}_{M' \in \mathbb{F}_q^{n \times 2(1-\delta)\ell}, v_i \in \mathbb{F}_q^n, A \in \mathbb{F}_q^{2\ell \times 2\ell}} \left[F\left([M', v_1, \ldots, v_{2\delta\ell}]A\right) \cdot \overline{H(M')} \mid \operatorname{rank}(A) = 2\ell \right] \\
&= \mathop{\mathbb{E}}_{M \in \mathbb{F}_q^{n \times 2\ell}, A \in \mathbb{F}_q^{2\ell \times 2\ell}} [F(M) \cdot \overline{H(MA^{-1}J)} \mid \operatorname{rank}(A) = 2\ell],
\end{aligned}$$

where in the second transition we used the fact that F is basis invariant, and in the third one we made a change of variables $M = [M', v_1, \ldots, v_{2\delta\ell}]A$. Now note that $A^{-1}J$ is the matrix A^{-1} restricted to its first $2(1-\delta)\ell$ columns and hence in the final distribution, $A^{-1}J$ is a uniformly random matrix in $\mathbb{F}_q^{n \times 2(1-\delta)\ell}$ with rank $2(1-\delta)\ell$. It follows that,

$$\langle \mathcal{T}F, H \rangle = \mathop{\mathbb{E}}_{M \in \mathbb{F}_q^{n \times 2\ell}, B \in \mathbb{F}_q^{2\ell \times 2(1-\delta)\ell}} [F(M) \cdot \overline{H(MB)} \mid \operatorname{rank}(B) = 2(1-\delta)\ell] = \langle F, \mathcal{G}H \rangle. \qquad \square$$

We will want to understand the operator $\mathcal{GT}$, and towards this end we define the operator

$$\Phi F(M) = \mathop{\mathbb{E}}_{\substack{B \in \mathbb{F}_q^{n \times 2\delta\ell} \\ C \in \mathbb{F}_q^{2\delta\ell \times 2\ell} \\ \operatorname{rank}(C) = 2\delta\ell}} [F(M + BC)].$$

The operator Φ is meant to be a version of $\mathcal{GT}$ which is easier to work with and acts in the same way on basis invariant functions. The upshot of working with Φ is that it is an averaging operator with respect to a Cayley graph over $\mathbb{F}_q^{n \times 2\ell}$, so each character χ_S is an eigenvector of Φ and the eigenvalues have an explicit formula (see [Spi19] for a detailed discussion).

Lemma A.12. *If $F \in L_2\left(\mathbb{F}_q^{n \times 2\ell}\right)$ is basis invariant, then $\mathcal{GT}F = \Phi F$.*

Proof. By definitions

$$\mathcal{GT}F(M) = \mathop{\mathbb{E}}_{\substack{R' \in \mathbb{F}_q^{2\ell \times 2(1-\delta)\ell}, \\ w_1, \ldots, w_{2\delta\ell} \in \mathbb{F}_q^n}} [F\left([MR', w_1, \ldots, w_{2\delta\ell}]\right) \mid \operatorname{rank}(R') = 2(1-\delta)\ell].$$

We can also view $M' = [MR', w_1, \ldots, w_{2\delta\ell}]$ as being sampled as follows. Choose $R' \in \mathbb{F}_q^{2\ell \times 2(1-\delta)\ell}$ with linearly independent columns, extend R' to a matrix $R \in \mathbb{F}_q^{2\ell\times 2\ell}$ with linearly independent columns randomly by adding $2\delta\ell$ columns on the right, sample a random matrix $[0, \ldots, 0, w_1, \ldots, w_{2\delta\ell}] \in \mathbb{F}_q^{n\times 2\ell}$, and output,

$$M' = MR + [0, \ldots, 0, w_1, \ldots, w_{2\delta\ell}].$$

Furthermore, under this distribution, it is clear that $R \in \mathbb{F}_q^{2\ell\times 2\ell}$ is a uniformly random matrix with linearly independent columns. Therefore,

$$\begin{aligned}\mathcal{GT}F(M) &= \mathop{\mathbb{E}}_{\substack{R\in\mathbb{F}_q^{2\ell\times 2\ell},\\ w_1,\ldots,w_{2\delta\ell}\in\mathbb{F}_q^n}} [F(MR + [0, \ldots, 0, w_1, \ldots, w_{2\delta\ell}]) \mid \operatorname{rank}(R) = 2\ell]\\ &= \mathop{\mathbb{E}}_{\substack{R\in\mathbb{F}_q^{2\ell\times 2\ell},\\ w_1,\ldots,w_{2\delta\ell}\in\mathbb{F}_q^n}} [F(M + [0, \ldots, 0, w_1, \ldots, w_{2\delta\ell}]R^{-1}) \mid \operatorname{rank}(R) = 2\ell],\end{aligned}$$

where we are using the fact that F is basis invariant and R is invertible. In the last expectation, note that the distribution over $[0, \ldots, 0, w_1, \ldots, w_{2\delta\ell}]R^{-1}$ is the same as that over BC where $B \in \mathbb{F}_q^{n\times 2\delta\ell}$ is uniformly random, and $C \in \mathbb{F}_q^{2\delta\ell\times 2\ell}$ is uniformly random conditioned on having linearly independent rows. More precisely, it is equal to BC where $B = [w_1, \ldots, w_{2\delta\ell}]$, and C is the last $2\delta\ell$ rows of R^{-1}. It follows that

$$\mathcal{GT}F(M) = \mathop{\mathbb{E}}_{B\in\mathbb{F}_q^{n\times 2\delta\ell}, C\in\mathbb{F}_q^{2\delta\ell\times 2\ell}} [F(M + BC) \mid \operatorname{rank}(C) = 2\delta\ell]\,. \qquad \square$$

The following lemma gives upper bound on the eigenvalues of Φ.

Lemma A.13. *Suppose that* $\operatorname{rank}(S) = t$*. If* $t = 0$*, then* χ_S *is an eigenvector of* Φ *of eigenvalue* 1*. If* $t > 0$*,* χ_S *is an eigenvector of* Φ *of eigenvalue which is at most* $3q^{t-n} + q^{-t(2\delta\ell-1)}$ *in absolute value.*

Proof. Fix S. We argued earlier that χ_S is an eigenvector of Φ, and we denote the corresponding eigevalue by $\lambda = \Phi\chi_S(0)$. If $t = 0$ the statement is clear, so we assume that $t > 0$ henceforth.

Find $A \in \mathbb{F}_q^{2\ell\times 2\ell}$ of full rank so that $SA^T = (v_1, \ldots, v_t, 0, 0, \ldots, 0)$ where $v_1, \ldots, v_t$ are linearly independent. Thus, as the distribution of C is invariant under multiplying by A^T from the right, we get that

$$\lambda = \Phi\chi_S(0) = \mathop{\mathbb{E}}_{B,C}[\chi_S(BCA^T) \mid \operatorname{rank}(C) = 2\delta\ell] = \mathop{\mathbb{E}}_{B,C}[\chi_{SA}(BC) \mid \operatorname{rank}(C) = 2\delta\ell] = \Phi\chi_{SA}(0),$$

where we used Lemma A.8. We may therefore assume that $S = (v_1, \ldots, v_t, 0, \ldots, 0)$ for linearly independent $v_1, \ldots, v_t$. Applying symmetry again, we conclude that

$$\lambda = \mathop{\mathbb{E}}_{\substack{v_1,\ldots,v_t\\ \text{linearly independent}}} \left[\Phi\chi_{(v_1,\ldots,v_t,\vec{0})}(0)\right] = \mathop{\mathbb{E}}_{\substack{v_1,\ldots,v_t\\ \text{linearly independent}}} \left[\mathop{\mathbb{E}}_{B,C} \omega^{\sum_{i=1}^{t} \operatorname{Tr}(v_i\cdot \mathsf{col}_i(BC))}\right],$$

and interchanging the order of expectations we get that

$$\lambda = \mathop{\mathbb{E}}_{B,C}\left[\mathop{\mathbb{E}}_{\substack{v_1,\ldots,v_t\\ \text{linearly independent}}} \omega^{\sum_{i=1}^{t} \operatorname{Tr}(v_i\cdot \mathsf{col}_i(BC))}\right],$$

Denote $w_i = \mathsf{col}_i(BC)$, and inspect these vectors.

Claim A.14. *If $w_i \neq 0$ for some i, then*

$$\left| \mathop{\mathbb{E}}_{\substack{v_1,\ldots,v_t \\ \textit{linearly independent}}} \left[\omega^{\sum_{i=1}^{t} \mathrm{Tr}(v_i \cdot \mathsf{col}_i(BC))} \right] \right| \leqslant 2q^{t-n}.$$

Proof. We first claim that if $v_1, \ldots, v_t$ are chosen uniformly, then the left hand side is 0, or equivalently

$$\mathop{\mathbb{E}}_{\substack{v_1,\ldots,v_t \\ \text{uniform}}} \left[\omega^{\mathrm{Tr}\left(\sum_{i=1}^{t} v_i \cdot w_i\right)} \right] = 0.$$

To see this, it suffices to show that $\sum_{i=1}^{t} v_i \cdot w_i$ takes every value in $\mathbb{F}_q$ with equal probability, and we focus on showing this. Fix i such that $w_i \neq 0$ and suppose the jth entry, $w_{i,j}$ is nonzero. We can fix all entries of the $v_1, \ldots, v_t$ uniformly except for $v_{i,j}$, and then for each $\alpha \in \mathbb{F}_q$, there is exactly one choice of $v_{i,j}$ that will result in $\sum_{i=1}^{t} v_i \cdot w_i = \alpha$.

Thus, if we took the distribution over $v_1, \ldots, v_t$ to be uniformly and independently chosen, then the magnitude of the above expectation would be 0. Hence, we conclude that the above expectation is at most twice the probability randomly chosen $v_1, \ldots, v_t$ are not linearly independent, which is at most q^{t-n}. $\square$

By Claim A.14 we conclude that $\lambda \leqslant 2q^{t-n} + \Pr_{B,C}[w_i = 0,\ \forall i = 1, \ldots, t]$, and we next bound this probability. Recalling the definition of w_i, we have that

$$w_i = \sum_{j=1}^{2\delta\ell} C(j,i)\mathsf{col}_j(B).$$

Consider the $2\delta\ell \times t$ minor of C and call it C'. First we upper bound the probability that $\mathrm{rank}(C') = 0$. Note that the distribution of C is the same as of $A|_{2\delta\ell \times 2\ell}$ where $A \in \mathbb{F}_q^{2\ell \times 2\ell}$ is a random invertible matrix. Thus, C' has the same distribution as of $A|_{2\delta\ell \times t}$, and the probability that $C' = 0$ is at most

$$\frac{q^{2\ell - 2\delta\ell}}{q^{2\ell} - 1} \cdot \frac{q^{2\ell - 2\delta\ell}}{q^{2\ell} - q} \cdots \frac{q^{2\ell - 2\delta\ell}}{q^{2\ell} - q^{t-1}} \leqslant q^{-t(2\delta\ell - 1)}.$$

It remains to bound the probability that w_i are all 0 in the case that $\mathrm{rank}(C') \geqslant 1$. In this case, assume without loss of generality that the first column of C' is non-zero. Thus, it follows that over the randomness of B, the vector w_1 is uniformly chosen from $\mathbb{F}_q^n$, and so the probability it is the all 0 vector is at most q^{-n}. Combining, we get that $\lambda \leqslant 3q^{t-n} + q^{-t(2\delta\ell-1)}$. $\square$

Finally, using Lemma A.8 again, we note that $\mathcal{G}$ does not increase the level of a function:

$$\mathcal{G}\chi_S(M) = \mathop{\mathbb{E}}_{A \in \mathbb{F}_q^{2\ell \times 2(1-\delta)\ell}} [\chi_S(MA)] = \mathop{\mathbb{E}}_{A}[\chi_{SA^T}(M)]. \tag{32}$$

This observation has the following implication:

Lemma A.15. *Let $F \in L_2\left(\mathbb{F}_q^{n \times 2\ell}\right)$ be basis invariant and let $G \in L_2\left(\mathbb{F}_q^{n \times 2(1-\delta)\ell}\right)$. Then, $\langle \mathcal{T} F^{=d}, H \rangle = \langle \mathcal{T} F^{=d}, H^{=d} \rangle$.*

Proof. Using (32), we have,

$$\mathcal{G}H^{=j}(M) = \sum_{S\in\mathbb{F}_q^{n\times 2(1-\delta)\ell},\mathrm{rank}(S)=j} \widehat{H}(S)\cdot \mathop{\mathbb{E}}_{A}\left[\chi_{SA^T}(M)\mid \mathrm{rank}(A)=2(1-\delta)\ell\right].$$

Since $\mathrm{rank}(S) = j$, it follows that the $\mathrm{rank}(SA^T)$ is at most j, so it follows that for $j < d$, we have $\langle F^{=d}, \mathcal{G}H^{=j}\rangle = 0$. As a result,

$$\langle \mathcal{T}F^{=d}, H\rangle = \langle F^{=d}, \mathcal{G}H\rangle = \sum_{j=d}^{2(1-\delta)\ell} \langle F^{=d}, \mathcal{G}H^{=j}\rangle. \tag{33}$$

In the first equality we used Lemma A.11 and the fact that $F^{=d}$ is basis invariant by Lemma A.10. Next we have,

$$\mathcal{T}F^{=d}(M) = \sum_{S\in\mathbb{F}_q^{n\times 2\ell},\mathrm{rank}(S)=d} \widehat{F}(S)\chi_S(M') = \sum_{S\in\mathbb{F}_q^{n\times 2\ell},\mathrm{rank}(S)=d} \widehat{F}(S)\chi_{S'}(M'),$$

where both M' and S' are obtained from M by removing the last $2\delta\ell$ columns. It follows that $\mathcal{T}F^{=d}$ has level at most d, so using (33) we get

$$\langle \mathcal{T}F^{=d}, H\rangle = \sum_{j=d}^{2(1-\delta)\ell} \langle F^{=d}, \mathcal{G}H^{=j}\rangle = \sum_{j=d}^{2(1-\delta)\ell} \langle \mathcal{T}F^{=d}, H^{=j}\rangle = \langle \mathcal{T}F^{=d}, H^{=d}\rangle. \qquad \square$$

We are now ready to state and prove an analog of Lemma 2.3 for basis invariant functions on the Bilinear scheme.

Lemma A.16. *Let $F \in L_2(\mathbb{F}_q^{n\times 2\ell})$ and $H \in L_2(\mathbb{F}_q^{n\times 2(1-\delta)\ell})$ be basis invariant indicator functions with $\mathbb{E}[F] = \alpha, \mathbb{E}[H] = \beta$. If F is (r,ε) pseudo-random, then for all $t \geqslant 4$ that are powers of 2, we have*

$$\langle \mathcal{T}F, H\rangle \leqslant q^{O_{t,r}(1)}\beta^{\frac{t-1}{t}}\varepsilon^{\frac{t-2}{t}} + q^{-r\delta\ell}\sqrt{\alpha\beta}.$$

Proof. Using the degree decomposition of F and Lemma A.15, we can write

$$\langle \mathcal{T}F, H\rangle = \sum_{d=0}^{2\ell} \langle \mathcal{T}F^{=d}, H^{=d}\rangle.$$

We first bound the contribution from terms in the summation with $d > r$ using Cauchy-Schwarz. For $d > r$,

$$\begin{aligned}
|\langle \mathcal{T}F^{=d}, H^{=d}\rangle|^2 &\leqslant \left\|\mathcal{T}F^{=d}\right\|_2^2 \left\|H^{=d}\right\|_2^2 \\
&= \left\|H^{=d}\right\|_2^2 \langle F^{=d}, \mathcal{G}\mathcal{T}F^{=d}\rangle \\
&= \left\|H^{=d}\right\|_2^2 \langle F^{=d}, \Phi F^{=d}\rangle \\
&\leqslant \left(q^{-2d\delta\ell} + 3q^{d-n}\right)\left\|F^{=d}\right\|_2^2 \left\|H^{=d}\right\|_2^2 \\
&\leqslant \left(q^{-2d\delta\ell} + 3q^{d-n}\right)\alpha\beta,
\end{aligned}$$

where the third transition uses Lemmas A.10 and A.12, and the fourth transition uses Lemma A.13. Thus, the total contribution from the $d > r$ terms is

$$\sum_{d=r+1}^{2\ell} \left|\langle \mathcal{T} F^{=d}, H^{=d}\rangle\right| \leqslant \sum_{d=r+1}^{2\ell} 2q^{-d\delta\ell}\sqrt{\alpha\beta} \leqslant q^{-r\delta\ell}\sqrt{\alpha\beta}.$$

Next, we bound the contribution from $d \leqslant r$ by bounding each term separately. Fix a $d \leqslant r$. By Lemma A.15 and Hölder's inequality we have

$$\left|\langle \mathcal{T} F^{=d}, H^{=d}\rangle\right| = \left|\langle \mathcal{T} F^{=d}, H\rangle\right| \leqslant \left\|\mathcal{T} F^{=d}\right\|_t \|H\|_{t/(t-1)} \leqslant \left\|F^{=d}\right\|_t \beta^{(t-1)/t} \leqslant q^{500d^2 t}\beta^{\frac{t-1}{t}}\varepsilon^{\frac{t-2}{t}},$$

where in the last inequality we use the fact that F is (r, ε)-pseudo-random, so by Theorem A.7

$$\left\|F^{=d}\right\|_t \leqslant q^{500d^2 t}\varepsilon^{\frac{t-2}{t}}.$$

Altogether, this shows

$$\langle \mathcal{T} F, H\rangle \leqslant q^{O_{t,r}(1)}\beta^{\frac{t-1}{t}}\varepsilon^{\frac{t-2}{t}} + q^{-r\delta\ell}\sqrt{\alpha\beta}. \qquad \square$$

A.3 Reduction to the Bilinear Scheme

We are now ready to prove Lemma 2.3. Let $\mathcal{L} \subseteq \mathsf{Grass}_q(n, 2\ell)$ and $\mathcal{R} \subseteq \mathsf{Grass}_q(n, 2(1-\delta)\ell)$ be as in Lemma 2.3 and suppose that $\mathcal{L}$ is (r, ε)-pseudo-random. Define the related Boolean functions $F \in L_2\left(\mathbb{F}_q^{n\times 2\ell}\right)$, $H \in L_2(\mathbb{F}_q^{n\times 2(1-\delta)\ell})$ as follows:

$$F(M') = \begin{cases} 1 & \text{if } \mathrm{im}(M') \in \mathcal{L}, \\ 0 & \text{otherwise,} \end{cases} \qquad G(M) = \begin{cases} 1 & \text{if } \mathrm{im}(M) \in \mathcal{R}, \\ 0 & \text{otherwise.} \end{cases} \tag{34}$$

In the above, $\mathrm{im}(\cdot)$ is the usual definition of matrix image and refers to the span of the columns of the matrix. Implicit in the definitions is the fact that F evaluates to 0 if the columns of M' are not linearly independent, as in this case their span is not even a dimension 2ℓ subspace, and likewise H evaluates to 0 if the columns of M are not linearly independent. We note that F and G are both basis invariant

Towards applying Lemma A.16, we want to first show that the pseudo-randomness of $\mathcal{L}$ carries over to F. To this end we begin with the following lemma that simplifies the type of zoom-ins and zoom-out combinations we have to consider when showing F is pseudo-random.

Lemma A.17. *For any $(U, V), (X, Y)$ such that $\dim(U) + \mathrm{codim}(X) < 2\ell$ and $\mathsf{Zoom}[(U, V), (X, Y)]$ is nonempty, there are r' linearly independent columns of V, say $v_1, \ldots, v_{r'} \in \mathbb{F}_q^n$ and a subset of linearly independent rows of X, say $X'' \in \mathbb{F}_q^{s'\times n}$, such that $r' \leqslant \dim(U)$, $s' \leqslant \mathrm{codim}(X)$ and*

$$\left\|F_{(U,V),(X,Y)}\right\|_2^2 \leqslant 2 \cdot \mathop{\mathbb{E}}_{M' \in \mathbb{F}_q^{n\times(2\ell-r')}}[F\left([v_1, \ldots, v_{r'}, M']\right) \mid X''M' = 0],$$

where $[v_1, \ldots, v_{r'}, M'] \in \mathbb{F}_q^{n\times 2\ell}$ is the matrix whose first r' columns are $v_1, \ldots, v_{r'}$, and remaining columns are M'.

Proof. Let $r = \dim(U)$ and $s = \operatorname{codim}(X)$. First note that we can assume that the columns of U and V respectively are both nonzero and linearly independent. Indeed, otherwise say $u_i = 0$, then either $v_i = 0$, in which case the ith columns of U and V can be removed, or $v_i \neq 0$ and $\mathsf{Zoom}[(U,V),(X,Y)]$ is an empty set. Otherwise, if, say, $v_i = 0$, then either $u_i = 0$ and again we can ignored the ith columns, or $u_i \neq 0$ and $\mathsf{Zoom}[(U,V),(X,Y)]$ consists of matrices whose columns are not linearly independent. In this case F is identically 0 on $\mathsf{Zoom}[(U,V),(X,Y)]$ and the statement is trivially true. Similarly, if the columns of V are not linearly independent, then $\mathsf{Zoom}[(U,V),(X,Y)]$ consists of matrices whose columns are not linearly independent, and again F is identically 0 on $\mathsf{Zoom}[(U,V),(X,Y)]$. Finally, if the columns of U are not linearly independent, then either $\mathsf{Zoom}[(U,V),(X,Y)]$ is empty or there must be some i such that both u_i and v_i are linear combinations of the other columns in U and V respectively, with the same coefficients. In this case, we can remove the ith columns of U and V without changing $\mathsf{Zoom}[(U,V),(X,Y)]$.

Now suppose that the columns of U and V are nonzero and linearly independent, and let $A \in \mathbb{F}_q^{2\ell\times 2\ell}$ be a full rank matrix such that the column vectors $AU_i = e_i$ for $1 \leqslant i \leqslant r$. Let $Y' \in \mathbb{F}_q^{d\times(2\ell-r)}$ denote the last $2\ell - r$ rows of YA^{-1}. We have

$$\begin{aligned}
\left\|F_{(U,V),(X,Y)}\right\|_2^2 &= \mathop{\mathbb{E}}_{M\in\mathbb{F}_q^{n\times 2\ell}}[F(M) \mid MU = V, XM = Y] \\
&= \mathop{\mathbb{E}}_{M\in\mathbb{F}_q^{n\times 2\ell}}[F(MA^{-1}) \mid MU = V, XM = Y] \\
&= \mathop{\mathbb{E}}_{M\in\mathbb{F}_q^{n\times 2\ell}}[F(M) \mid MAU = V, XMA = Y] \\
&= \mathop{\mathbb{E}}_{M'\in\mathbb{F}_q^{n\times(2\ell-r')}}[F([v_1,\dots,v_r,M']) \mid XM' = Y'].
\end{aligned}$$

In the second equality we use the fact that F is basis invariant and in the last equality we use the fact that because $\mathsf{Zoom}[(U,V),(X,Y)]$ is nonempty, the matrix product $X[v_1,\dots,v_r]$ has the same first r rows as Y.

To complete the proof, we show how to reduce to the case that Y' is the zero matrix. First note that, using the same reasoning as we did for U and V, we can assume that the nonzero rows Y' are linearly independent and the rows of X are linearly independent. Suppose that $y'_1,\dots,y'_a \in \mathbb{F}_q^{2\ell-r}$ are the nonzero (and linearly independent) rows of Y', while the remaining rows are $y'_{a+1},\dots,y'_s = 0$. Let $Y'' \in \mathbb{F}_q^{a\times(2\ell-r)}$ be the first a rows of Y', which are nonzero, let $X' \in \mathbb{F}_q^{a\times n}$ denote the first a rows of X, and let $X'' \in \mathbb{F}_q^{(2\ell-a)\times n}$ denote rows $a+1$ through s of X. For any $Z = (z_1,\dots,z_a) \in \mathbb{F}_q^{a\times(2\ell-r)}$ with a linearly independent rows let $A_Z \in \mathbb{F}_q^{(2\ell-r)\times(2\ell-r)}$ be the full rank matrix such that $Y''A_Z = Z$. Then, for any linearly independent $z_1,\dots,z_a \in \mathbb{F}_q^{2\ell-r}$,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{M'\in\mathbb{F}_q^{n\times(2\ell-r')}}[F([v_1,\dots,v_r,M']) \mid XM' = Y'] \\
&\qquad= \mathop{\mathbb{E}}_{M'}[F([v_1,\dots,v_r,M']) \mid X'M' = Y'', X''M' = 0] \\
&\qquad= \mathop{\mathbb{E}}_{M'}[F([v_1,\dots,v_r,M'A_Z^{-1}]) \mid X'M' = Y'', X''M' = 0] \\
&\qquad= \mathop{\mathbb{E}}_{M'}[F([v_1,\dots,v_r,M']) \mid X'M' = Z, X''M' = 0].
\end{aligned}$$

In the second transition, we used the fact that F is basis invariant and multiplied its input by the matrix whose top left $r \times r$ minor is the identity, its bottom right $(2\ell - r)\times(2\ell - r)$ is A_Z^{-1}, and the rest of the entries are 0. Since the above holds for any Z with a many linearly independent rows, letting E denote the event that $X'M'$ has a linearly independent rows, it follows that

$$\mathop{\mathbb{E}}_{M'\in\mathbb{F}_q^{n\times(2\ell-r')}}[F([v_1,\ldots,v_r,M'])\mid XM'=Y']=\mathop{\mathbb{E}}_{M'}[F([v_1,\ldots,v_r,M'])\mid E\wedge X''M'=0],$$

and

$$\begin{aligned}\mathop{\mathbb{E}}_{M'\in\mathbb{F}_q^{n\times(2\ell-r')}}[F([v_1,\ldots,v_r,M'])\mid X''M'=0]\geqslant{}&\Pr_{M'\in\mathbb{F}_q^{n\times(2\ell-r')}}[E\mid X''M'=0]\\&\cdot\mathop{\mathbb{E}}_{M'}[F([v_1,\ldots,v_r,M'])\mid E\wedge X''M'=0].\end{aligned}$$

Finally since $\Pr[E\mid X''M'=0]\geqslant\frac{1}{2}$ (as it is the probability of choosing $a<2\ell-r$ linearly independent vectors in $\mathbb{F}_q^{2\ell-r}$), we have,

$$\begin{aligned}\left\|F_{(U,V),(X,Y)}\right\|_2^2&=\mathop{\mathbb{E}}_{M'\in\mathbb{F}_q^{n\times(2\ell-r')}}[F([v_1,\ldots,v_r,M'])\mid E\wedge X''M'=0]\\&\leqslant2\mathop{\mathbb{E}}_{M'}[F([v_1,\ldots,v_r,M'])\mid X''M'=0],\end{aligned}$$

and the proof is concluded. □

As an immediate consequence of Lemma A.17 we have the following result.

Lemma A.18. *If $\mathcal{L}\subseteq\mathsf{Grass}_q(n,2\ell)$ is (r,ε)-pseudo-random, then F defined from $\mathcal{L}$ as in* (34) *is $(r,2\varepsilon)$-pseudo-random.*

Proof. Fix any (U,V) and (X,Y) such that $\dim(U)+\operatorname{codim}(X)=r$. Using Lemma A.17, there are linearly independent $v_1,\ldots,v_{r'}\in\mathbb{F}_q^n$ and $X'\in\mathbb{F}_q^{s'\times n}$ with linearly independent rows such that

$$\begin{aligned}&\left\|F_{(U,V),(X,Y)}\right\|_2^2\\&\qquad\leqslant2\cdot\mathop{\mathbb{E}}_{M'\in\mathbb{F}_q^{n\times(2\ell-r')}}\left[F\left([v_1,\ldots,v_{r'},M']\right)\ \middle|\ X'M'=0\right]\\&\qquad\leqslant2\cdot\mathop{\mathbb{E}}_{M'\in\mathbb{F}_q^{n\times(2\ell-r')}}\left[F\left([v_1,\ldots,v_{r'},M']\right)\ \middle|\ X'M'=0,\dim(\operatorname{im}([v_1,\ldots,v_r,M']))=2\ell\right]\\&\qquad=2\cdot\Pr_{M'\in\mathbb{F}_q^{n\times(2\ell-r')}}[\operatorname{im}([v_1,\ldots,v_{r'},M'])\in\mathcal{L}\mid X'M'=0,\dim(\operatorname{im}([v_1,\ldots,v_r,M']))=2\ell],\end{aligned}$$

where in the second transition we are using the fact that $F(M)=0$ for all M such that $\dim(\operatorname{im}(M))<2\ell$, and in the third transition we are using the definition of F. We will bound the final term by using the pseudo-randomness of F.

Choosing $M'\in\mathbb{F}_q^{n\times(2\ell-r')}$ uniformly conditioned on $X'M'=0$, and $\dim(\operatorname{im}([v_1,\ldots,v_r,M']))=2\ell$, we claim that $\operatorname{im}([v_1,\ldots,v_r,M'])$ is a uniformly random 2ℓ-dimensional subspace in $\mathsf{Zoom}[Q,Q+H]$, where H is the codimension s subspace that is dual to the rows of X' and $Q=\operatorname{span}(v_1,\ldots,v_r)$. To see why, first note that it is clear $\operatorname{im}([v_1,\ldots,v_r,M'])\in\mathsf{Zoom}[Q,Q+H]$. Additionally, each $L\in\mathsf{Zoom}[Q,Q+H]$ has an equal number of $M'\in\mathbb{F}_q^{n\times(2\ell-r')}$ such that

$$L=\operatorname{im}([v_1,\ldots,v_r,M']),$$

and therefore has an equal chance of being selected. It follows that choosing $M' \in \mathbb{F}_q^{n\times(2\ell-r')}$ uniformly conditioned on $X'M' = 0$, and $\dim(\mathrm{im}([v_1,\dots,v_r,M']))$, $\mathrm{im}([v_1,\dots,v_r,M'])$ is a uniformly random 2ℓ-dimensional subspace in $\mathsf{Zoom}[Q,Q+H]$. As a result,

$$\begin{aligned}
&\left\|F_{(U,V),(X,Y)}\right\|_2^2\\
&\qquad\leqslant 2\cdot \Pr_{M'\in\mathbb{F}_q^{n\times(2\ell-r')}}[\mathrm{im}([v_1,\dots,v_{r'},M'])\in\mathcal{L}\mid X'M'=0,\dim(\mathrm{im}([v_1,\dots,v_r,M']))=2\ell]\\
&\qquad= 2\cdot \Pr_{L\in\mathsf{Zoom}[Q,Q+H]}[L\in\mathcal{L}]\\
&\qquad\leqslant 2\varepsilon,
\end{aligned}$$

where in the last transition we use the fact that the set of subspaces $\mathcal{L}$ is (r,ε)-pseudo-random and $\dim(Q)+\mathrm{codim}(Q+H)\leqslant r'+s\leqslant r$. □

The next lemma shows that the value of the probability of interest in Lemma 2.3 is not much larger than the value $\langle \mathcal{T}F, G\rangle$:

Lemma A.19. *We have*

$$\Pr_{L\in\mathsf{Grass}_q(n,2\ell),R\in\mathsf{Grass}_q(n,2(1-\delta)\ell),L\supseteq R}[L\in\mathcal{L},R\in\mathcal{R}]\leqslant 2\langle\mathcal{T}F,G\rangle.$$

Proof. We have,

$$\begin{aligned}
\langle\mathcal{T}F,G\rangle &= \mathop{\mathbb{E}}_{M\in\mathbb{F}_q^{n\times 2\ell},A\in\mathbb{F}_q^{2\ell\times2(1-\delta)\ell}}[F(M)\cdot G(MA)\mid \mathrm{rank}(A)=2(1-\delta)\ell]\\
&\geqslant \Pr_{M,A}[\mathrm{rank}(M)=2\ell,\mathrm{rank}(MA)=2(1-\delta)\ell\mid \mathrm{rank}(A)=2(1-\delta)\ell]\\
&\cdot \mathop{\mathbb{E}}_{M,A}[F(M)\cdot G(MA)\mid \mathrm{rank}(M)=2\ell,\mathrm{rank}(MA)=2(1-\delta)\ell,\mathrm{rank}(A)=2(1-\delta)\ell]\\
&\geqslant \frac{1}{2}\mathop{\mathbb{E}}_{M,A}[F(M)\cdot G(MA)\mid \mathrm{rank}(M)=2\ell,\mathrm{rank}(MA)=2(1-\delta)\ell,\mathrm{rank}(A)=2(1-\delta)\ell]\\
&= \frac{1}{2}\mathop{\mathbb{E}}_{M,A}[F(\mathrm{im}(M))\cdot G(\mathrm{im}(MA))\mid \mathrm{rank}(M)=2\ell,\mathrm{rank}(MA)=2(1-\delta)\ell,\mathrm{rank}(A)=2(1-\delta)\ell].
\end{aligned}$$

To finish the proof, notice that in the conditional distribution $(\mathrm{im}(M),\mathrm{im}(MA))$ in the last term, $\mathrm{im}(M)$ is a uniform $L\in\mathsf{Grass}_q(n,2\ell)$ and $\mathrm{im}(MA)$ is a uniform $R\in\mathsf{Grass}_q(n,2(1-\delta)\ell)$ such that $R\subseteq L$. Thus, the final expectation is exactly the probability

$$p:=\Pr_{L\in\mathsf{Grass}_q(n,2\ell),R\in\mathsf{Grass}_q(n,2(1-\delta)\ell),L\supseteq R}[L\in\mathcal{L},R\in\mathcal{R}].$$

Therefore,

$$\langle\mathcal{T}F,G\rangle\geqslant\frac{1}{2}\mathop{\mathbb{E}}_{L,R}[F(L)\cdot G(R)\mid L\supseteq R]=\frac{p}{2}$$ □

We now prove Lemma 2.3 by combining Lemmas A.18 and A.19.

Proof of Lemma 2.3. Take $\mathcal{L} \subseteq \mathsf{Grass}_q(n, 2\ell)$ and $\mathcal{R} \subseteq \mathsf{Grass}_q(n, 2(1-\delta)\ell)$ with fractional sizes α and β as in the lemma statement. Using these sets, define the associated functions $F \in L_2\left(\mathbb{F}_q^{n\times 2\ell}\right)$ and $G \in L_2\left(\mathbb{F}_q^{n\times 2(1-\delta)\ell}\right)$ as in (34). It is clear that

$$\|F\|_2^2 \leqslant \alpha, \quad \|G\|_2^2 \leqslant \beta.$$

Furthermore, by Lemma A.19, we have that

$$\Pr_{L\supseteq R}[L \in \mathcal{L}, R \in \mathcal{R}] \leqslant 2\langle \mathcal{T}F, G\rangle.$$

By Lemma A.18 F is $(r, 2\varepsilon)$-pseudo-random, and applying Lemma A.16 we get that

$$\Pr_{L\supseteq R}[L \in \mathcal{L}, R \in \mathcal{R}] \leqslant 2\langle \mathcal{T}F, G\rangle \leqslant q^{O_{t,r}(1)}\beta^{\frac{t-1}{t}}\varepsilon^{\frac{t-2}{t}} + q^{-r\delta\ell}\sqrt{\alpha\beta}. \qquad \square$$

A.4 Proof of Lemma 2.4

We will show that if a set of ℓ'-dimensional subspaces $\mathcal{L}^\star \subseteq \mathsf{Grass}_q(W, \ell')$ is pseudo-random, then it must "evenly cover" the space W in the sense that there are very few points $z \in W$ such that $\mu_z(\mathcal{L}^\star)$ significantly deviates from $\mu(\mathcal{L}^\star)$. We will require the following result from [EKL24].

Theorem A.20. *[EKL24, Theorem 1.14] If $F \in L_2\left(\mathbb{F}_q^{n\times \ell'}\right)$ is a Boolean function which is $(1, \varepsilon)$-global, then for all t that are powers of 2, we have*

$$\left\|F^{=1}\right\|_2^2 \leqslant q^{460t}\|F\|_2^2\,\varepsilon^{1-2/t}.$$

We are now ready to prove Lemma 2.4.

Proof of Lemma 2.4. Let $\dim(W) = n$, let $F' \in L_2\left(\mathbb{F}_q^{n\times \ell'}\right)$ be the function associated with $\mathcal{L}^\star$ given by

$$F'(x_1, \dots, x_{\ell'}) = \begin{cases} 1 & \text{if } \mathrm{span}(x_1, \dots, x_{2\ell}) \in \mathcal{L}^\star, \\ 0 & \text{otherwise.} \end{cases}$$

By Lemma A.18, F' is $(1, 2q^{c\ell'}\eta)$-pseudo-random. For any point $x \in W$, we have

$$\mu_z(\mathcal{L}^\star) = \mathop{\mathbb{E}}_{x_1,\dots,x_{\ell'-1}\in W}\left[F'(x_1, \dots, x_{\ell'-1}, z) \mid \dim(\mathrm{span}(x_1, \dots, x_{\ell'-1}, z)) = \ell'\right],$$

so it follows that

$$\left|\mu_z(\mathcal{L}^\star) - \mathop{\mathbb{E}}_{x_1,\dots,x_{\ell'-1}}\left[F'(x_1, \dots, x_{\ell'-1}, z)\right]\right| \leqslant \frac{q^{\ell'}}{q^n} \quad \text{and} \quad \left|\mu(\mathcal{L}^\star) - \left\|F'\right\|_2^2\right| \leqslant \frac{q^{\ell'}}{q^n}.$$

Thus

$$\left|\mathop{\mathbb{E}}_{x_1,\dots,x_{\ell'-1}}\left[F'(x_1, \dots, x_{\ell'-1}, z)\right] - \left\|F'\right\|_2^2\right| \geqslant |\mu_z(\mathcal{L}^\star) - \mu(\mathcal{L}^\star)| - \frac{q^{\ell'}}{q^n}. \tag{35}$$

We will now relate this quantity to the level 1 weight of F'. Note that

$$\begin{aligned}\mathop{\mathbb{E}}_{x_1,\ldots,x_{\ell'-1}}\left[F'(x_1,\ldots,x_{\ell'-1},z)\right] &= \sum_{S=(s_1,\ldots,s_{\ell'})\in W}\widehat{F'}(S)\mathop{\mathbb{E}}_{x_1,\ldots,x_{\ell'-1}}[\chi_S(x_1,\ldots,x_{\ell'-1},z)] \\ &= \sum_{S=(s_1,\ldots,s_{\ell'})\in W}\widehat{F'}(S)\chi_{s_{\ell'}}(z)\prod_{i=1}^{\ell'-1}\mathop{\mathbb{E}}_{x_i}\left[\chi_{s_i}(x_i)\right].\end{aligned}$$

Now note that $\mathbb{E}_{x_i}\left[\chi_{s_i}(x_i)\right]=0$ if s_i is not the zero vector, and $\mathbb{E}_{x_i}\left[\chi_{s_i}(x_i)\right]=1$ if s_i is the zero vector. Thus,

$$\mathop{\mathbb{E}}_{x_1,\ldots,x_{\ell'-1}}\left[F'(x_1,\ldots,x_{\ell'-1},z)\right]=\widehat{F'}(0,\ldots,0)+\sum_{a\in W,a\neq 0}\widehat{F'}(0,\ldots,0,a)\chi_a(z),$$

and using the fact that $\widehat{F'}(0,\ldots,0)=\mathbb{E}[F']=\|F'\|_2^2$,

$$\left(\mathop{\mathbb{E}}_{x_1,\ldots,x_{\ell'-1}}\left[F'(x_1,\ldots,x_{\ell'-1},z)\right]-\left\|F'\right\|_2^2\right)^2=\left(\sum_{a\in W,a\neq 0}\widehat{F'}(0,\ldots,0,a)\chi_a(z)\right)^2.$$

Therefore, we get by (35) that

$$\begin{aligned}\mathop{\mathbb{E}}_{z\in W}[(\mu_z(\mathcal{L}^\star)-\eta)^2] &\leqslant \mathop{\mathbb{E}}_{z\in W}\left[\left|\mathop{\mathbb{E}}_{x_1,\ldots,x_{\ell'-1}}\left[F'(x_1,\ldots,x_{\ell'-1},z)\right]-\left\|F'\right\|_2^2\right|^2\right]+5\frac{q^{\ell'}}{q^n} \\ &\leqslant \mathop{\mathbb{E}}_{z}\left[\left|\sum_{a\in W,a\neq 0}\widehat{F'}(0,\ldots,0,a)\chi_a(z)\right|^2\right]+5\frac{q^{\ell'}}{q^n} \\ &= \sum_{a\in W,a\neq 0}|\widehat{F'}(0,\ldots,0,a)|^2+5\frac{q^{\ell'}}{q^n}.\end{aligned}$$

We now examine the summation in the last line. Since F' is basis invariant, using Lemma A.8, we have that for all $\alpha_1,\ldots,\alpha_{\ell'}\in\mathbb{F}_q$ that are not all zero,

$$\widehat{F'}\left(\alpha_1 a,\ldots,\alpha_{\ell'}a\right)=\widehat{F'}\left(0,\ldots,0,a\right).$$

It follows that

$$\sum_{a\in W,a\neq 0}|\widehat{F'}(0,\ldots,0,a)|^2=\frac{1}{q^{\ell'}-1}\sum_{\mathrm{rank}(S)=1}|\widehat{F'}(S)|^2=\frac{\left\|F'^{=1}\right\|_2^2}{q^{\ell'}-1}.$$

Using Theorem A.20 along the fact that F' is $(1,2q^{c\ell'}\eta)$-pseudo-random, we get

$$\mathop{\mathbb{E}}_{z\in W}[(\mu_z(\mathcal{L}^\star)-\eta)^2]\leqslant\frac{\left\|F'^{=1}\right\|_2^2}{q^{\ell'}-1}+5\frac{q^{\ell'}}{q^n}\leqslant\frac{q^{461t}q^{c\ell'}\eta^{2-\frac{2}{t}}}{q^{\ell'}-1},$$

for any $t\geqslant 4$ that is a power of 2, where we also used the fact that $\|F'\|_2^2\leqslant 2\mu(\mathcal{L})\leqslant 2\eta$ by Lemma A.18. By Markov's inequality it follows that

$$\frac{|\overline{Z}|}{q^n}\cdot\frac{\eta^2}{100}\leqslant\frac{100q^{461t}q^{c\ell'}}{\eta^{\frac{2}{t}}q^{\ell'}}\eta^2\leqslant\frac{100q^{461t}}{q^{\left(1-c-\frac{2C}{t}\right)\ell'}}\eta^2\leqslant\frac{q^{-\frac{\ell'}{2}}}{100}\eta^2.$$

In the third term, we set $C := \frac{6}{\xi}$, and use the fact that $\ell' \geqslant \frac{\xi}{3}\ell$, so $\eta \geqslant q^{-2\ell} \geqslant q^{-C\ell'}$. Then for the last transition we take t to be the smallest power of 2 that satisfies $1 - c + \frac{2C}{t} \geqslant \frac{2}{3}$, and use the fact that ℓ, and hence ℓ', is sufficiently large. Dividing by η^2 finishes the proof. □

B Proof of Theorem 5.2

In order to prove Theorem 5.2 we will find the subspaces Q one at a time by using Theorem 5.1. We maintain a list $\mathcal{Q}$ of all Q's collected thus far. Each time a new subspace Q is added to $\mathcal{Q}$, we randomize the assignment $T_1[L]$ for all 2ℓ-dimensional $L \supset Q$. At a high level, the effect of this randomization is that there is only a little agreement between any linear function and the assignments on subspaces containing Q, thus these entries are essentially "deleted".

Formally, we construct the set $\mathcal{Q}$ of subspaces as follows. Initially set $\widetilde{T}_1 = T_1$, $\mathcal{Q} = \emptyset$, and $\mathcal{X} = \emptyset$. Recall that initially $\widetilde{T}_1$ and T_2 are ε-consistent for $\varepsilon \geqslant 2q^{-2\ell(1-1000\delta)}$. While $\widetilde{T}_1$ and T_2 are at least $\varepsilon/2$-consistent, do the following.

1. Let $Q \subset W$ be subspaces guaranteed by Theorem 5.1. That is, Q and W satisfy $\dim(Q)+\mathrm{codim}(W) = r$ and there exists linear $g_{Q,W} : W \to \mathbb{F}_q$ such that
$$\Pr_{L \in \mathsf{Grass}(n,2\ell)}[g_{Q,W}|_L = \widetilde{T}_1[L] \mid Q \subseteq L \subseteq W] \geqslant \varepsilon' := q^{-2(1-1000\delta^2)\ell}.$$
2. Set $\mathcal{Q} \leftarrow \mathcal{Q} \cup \{Q\}$.
3. Set $\mathcal{X} \leftarrow \mathcal{X} \cup \{L \mid Q \subseteq L \subseteq W\}$.
4. For each $L \in \mathcal{X}$ independently, choose $\widetilde{T}_1[L]$ uniformly among all linear functions on L.

We have the following claim regarding the re-assignment phase.

Claim B.1. *With probability at least $1-e^{-\Omega(q^{\ell n})}$ over the random assignment $\widetilde{T}_1$, for every pair of subspaces $Q \subset W$ such that $\dim(Q)+\mathrm{codim}(W) = r$ and $\mu_{Q,W}(\mathcal{X}) \geqslant q^{-2\ell}$, and every linear function $g_{Q,W} : W \to \mathbb{F}_q$, we have*
$$\Pr_{L \in \mathcal{X}, L \in \mathsf{Zoom}_{2\ell}[Q,W]}[g_{Q,W}|_L \equiv \widetilde{T}_1[L]] \leqslant q^{1-2\ell}$$

Proof. Fix $Q, W, g_{Q,W}$ such that $\dim(Q)+\mathrm{codim}(W) = r$ and $\mu_{Q,W}(\mathcal{X}) \geqslant q^{-2\ell}$. Let $\mathcal{A} = \mathsf{Zoom}_{2\ell}[Q,W] \cap \mathcal{X}$ and for each $L \in \mathcal{A}$, let Z_L denote the indicator variable that takes value 1 if $g_{Q,W}|_L = \widetilde{T}_1[L]$ and 0 otherwise, where $\widetilde{T}_1$ is the randomly chosen assignment over $\mathcal{X}$. Over randomly chosen $\widetilde{T}_1$, the expectation of Z_L is $q^{-2\ell}$, so by a Chernoff bound, we have
$$\Pr_{\widetilde{T}}\left[\frac{1}{|\mathcal{A}|}\sum_{L \in \mathcal{A}} Z_L \geqslant q^{1-2\ell}\right] \leqslant e^{-q^2\left(q^{-2\ell}|\mathcal{A}|\right)/6}.$$

By assumption, $|\mathcal{A}| \geqslant q^{-2\ell}\,|\{L \in \mathsf{Zoom}_{2\ell}[Q,W], \dim(L) = 2\ell\}| \geqslant q^{-2\ell}q^{(2\ell-r)(n-r-2\ell)}$. Thus, using a union bound over all $Q, W, g_{Q,W}$, the probability that there exist a bad triple is at most,
$$(r+1)q^{nr}q^n e^{-q^{-4\ell+2}q^{(2\ell-r)(n-r-2\ell)}/6} \leqslant e^{-\Omega(q^{\ell n})}.$$
□

We now analyze the process. Note that when we find a new triplet $(L, Q, g_{Q,W})$, then by Chernoff's bound with probability $1 - e^{-\Omega(q^{\ell n})}$ over the randomization step, the probability $\Pr_{L\in\mathsf{Grass}_q(n,2\ell)}[g_{Q,W}|_L = \widetilde{T}_1[L] \mid Q \subseteq L \subseteq W]$ drops from at least ε' to at most $q^{1-2\ell}$. In that case, the measure of $\mathcal{X}$ increases by at least

$$(\varepsilon' - q^{1-2\ell})q^{-rn} \geqslant q^{-O(rn)}.$$

Doing a union bound over the steps, it follows that with probability $1 - e^{-\Omega(q^{\ell n})}q^{O(rn)} = 1 - o(1)$ the process terminates within $q^{O(rn)}$ steps.

Note that it is possible that the same subspace Q is added multiple times (with different zoom-outs) in the process above, so we clarify that $\mathcal{Q}$ is considered as a set without repeats. Also note that with probability $1 - o(1)$, for each $Q \in \mathcal{Q}$, W and $g_{Q,W}$ found in the process it holds that

$$\Pr_{L\in\mathsf{Grass}(n,2\ell)}[g_{Q,W}|_L = T_1[L] \mid Q \subseteq L \subseteq W] \geqslant \frac{1}{2}\varepsilon' \tag{36}$$

(the point being is that the agreement now is compared to the original T_1 and not to $\widetilde{T}_1$). Indeed, considering the step Q, W and $g_{Q,W}$ were found, $g_{Q,W}$ had agreement at least ε' with $\widetilde{T}_1$ on $\mathsf{Zoom}_{2\ell}[Q, W]$ at that point, and by Claim B.1 with probability $1 - e^{-\Omega(q^{\ell n})}$ at most $q^{1-2\ell} \leqslant \varepsilon'/2$ of that agreement came from $L \in \mathcal{X}$. Thus, by union bound over all of the steps, with probability $1 - q^{O(rn)}e^{q^{-\Omega(\ell n)}} = 1 - o(1)$ it follows that (36) holds for every Q, W and $g_{Q,W}$ found throughout the process.

The following claim shows that at the end of the process the number of Q's found in the process is large, thereby finishing the proof of Theorem 5.2.

Claim B.2. *There exists some $0 \leqslant r_1 \leqslant r$ such that $\mathcal{Q}$ contains at least a $q^{-5\ell^2}$-fraction of all r_1-dimensional subspaces.*

Proof. At the end of the process, the consistency has dropped by at least $\varepsilon'/2$, so the probability over edges (L, R) that L was reassigned must be at least $\varepsilon'/2$. For each $0 \leqslant r_1 \leqslant r$, let N_{r_1} be the number of Q of dimension r_1 in $\mathcal{Q}$.

For each Q of dimension r_1, the fraction of 2ℓ-dimensional L's that are reassigned due to Q being added to $\mathcal{Q}$ is at most the fraction of 2ℓ-dimensional subspaces that contain Q. This is,

$$\frac{\begin{bmatrix} n \\ 2\ell - r_1 \end{bmatrix}_q}{\begin{bmatrix} n \\ 2\ell \end{bmatrix}_q} \leqslant \frac{q^{n(2\ell - r_1)}}{q^{2\ell(n-2\ell)}} = q^{4\ell^2 - r_1 n}.$$

It follows that there must be some r_1 such that $N_{r_1}q^{4\ell^2 - r_1 n} \geqslant \frac{\varepsilon}{2(r+1)}$, and rearranging gives that

$$N_{r_1} \geqslant \frac{\varepsilon}{2(r+1)q^{4\ell^2}}q^{r_1 n} \geqslant q^{-5\ell^2}\begin{bmatrix} n \\ r_1 \end{bmatrix}_q.$$

Thus there exists an r_1 such that $\mathcal{Q}$ contains at least a $q^{-5\ell^2}$-fraction of all r_1-dimensional subspaces. $\square$

C Proof of Lemma 5.5

Fix a question U to the first prover. Recall from (7) that we set

$$k = q^{2(1+c)\ell} \quad \text{and} \quad \beta = q^{-2(1+2c/3)\ell},$$

where $0 < c < 1$ is some small constant close to 0 and set $\eta = q^{-100\ell^{100}}$. Let us recall the distributions $\mathcal{D}$ and $\mathcal{D}'$, as well as Lemma 5.5.
$\mathcal{D}$:

- Choose $x_1, \ldots, x_{2\ell} \in \mathbb{F}_q^U$ uniformly.
- Output the list $(x_1, \ldots, x_{2\ell})$.

$\mathcal{D}'$:

- Choose $V \subseteq U$ according to the outer PCP.
- Choose $x'_1, \ldots, x'_{2\ell} \in \mathbb{F}_q^V$ uniformly, and lift these vectors to $\mathbb{F}_q^U$ by inserting 0's into the missing coordinates.
- Choose $w_1, \ldots, w_{2\ell} \in H_U$ uniformly, and set $x_i = x'_i + w_i$ for $1 \leqslant i \leqslant 2\ell$.
- Output the list $(x_1, \ldots, x_{2\ell})$.

Lemma 5.5. *There exists $E \subseteq \left(\mathbb{F}_q^U\right)^{2\ell}$ such that both $\mathcal{D}(E)$ and $\mathcal{D}'(E)$ are at most η^{40}, and for all $(x_1, \ldots, x_{2\ell}) \notin E$ we have*

$$0.9 \leqslant \frac{\mathcal{D}(x_1, \ldots, x_{2\ell})}{\mathcal{D}'(x_1, \ldots, x_{2\ell})} \leqslant 1.1.$$

Proof. For $x_1, \ldots, x_{2\ell} \in \mathbb{F}_q^{|U|}$, let us view $x_1, \ldots, x_{2\ell}$ as the rows of a $2\ell \times 3k$ matrix, and split the columns of this matrix into k blocks - each consisting of 3 consecutive columns. The three columns of each one of these k blocks correspond to the three variables coming from one equation of the outer PCP. Let $s(x_1, \ldots, x_{2\ell})$ be the number of blocks where exactly two of the columns are equal, and set $p = 3q^{-2\ell} - 3q^{-4\ell}$ to be the probability that a fixed block has exactly two columns equal to each other. The intention is that $s(x_1, \ldots, x_{2\ell})$ equals the number of equations where we drop variables in the outer PCP.

Let $s'(x_1, \ldots, x_{2\ell})$ be the number of blocks where all 3 columns are equal, and let $p' = q^{-4\ell}$ be the probability that a fixed block has all three columns equal. We define the sets E_1, E_2, E as follows:

$$E_1 = \left\{(x_1, \ldots, x_{2\ell}) \in \left(\mathbb{F}_q^U\right)^{2\ell} \mid |s(x_1, \ldots, x_{2\ell}) - pk| > 50\sqrt{pk\log(1/\eta)}\right\}, \tag{37}$$

$$E_2 = \left\{(x_1, \ldots, x_{2\ell}) \in \left(\mathbb{F}_q^U\right)^{2\ell} \mid s'(x_1, \ldots, x_{2\ell}) > \ell^{100}\right\}, \tag{38}$$

and $E = E_1 \cup E_2$. We first argue that the event E has small probability under both $\mathcal{D}$ and $\mathcal{D}'$. Note that

$$\mathop{\mathbb{E}}_{\mathcal{D}}[s(x_1, \ldots, x_{2\ell})] = pk \quad \text{and} \quad \mathop{\mathbb{E}}_{\mathcal{D}'}[s(x_1, \ldots, x_{2\ell})] = (1-\beta)pk + \beta(1-q^{-2\ell})k,$$

and in particular,

$$\left|\mathop{\mathbb{E}}_{\mathcal{D}'}[s(x_1, \ldots, x_{2\ell})] - pk\right| \leqslant \beta k.$$

By a Chernoff bound,

$$\mathcal{D}(E_1) = \Pr_{\mathcal{D}}\left[|s(x_1,\ldots,x_{2\ell}) - pk| > 50\sqrt{pk\log(1/\eta)}\right] \leqslant \eta^{50},$$

where recall that $\eta = q^{-100\ell^{100}}$. Also, by our setting of βk, we have $\beta k = q^{2c\ell/3}$, while $pk = \Omega(q^{2c\ell})$, so the same Chernoff bound holds for $\mathcal{D}'$:

$$\mathcal{D}'(E_1) = \Pr_{\mathcal{D}'}\left[|s(x_1,\ldots,x_{2\ell}) - pk| > 50\sqrt{pk\log(1/\eta)}\right] \leqslant \eta^{50}.$$

For the event E_2 we have,

$$\mathcal{D}(E_2) = \Pr_{D}[s'(x_1,\ldots,x_{2\ell}) > \ell^{100}] \leqslant \binom{k}{\ell^{100}} p'^{\ell^{100}} \leqslant (kp')^{\ell^{100}} \leqslant \eta^{100},$$

where in the middle term, the first factor is the number of ways to choose ℓ^{100} blocks and the second factor is the probability that all of these blocks have all three columns equal. Similarly,

$$\mathcal{D}'(E_2) = \Pr_{D'}[s'(x_1,\ldots,x_{2\ell}) > \ell^{100}] \leqslant \binom{k}{\ell^{100}} \left((1-\beta)p' + \beta\frac{1}{q^{2\ell}}\right)^{\ell^{100}} \leqslant k^{\ell^{100}}\left(2q^{-4\ell}\right)^{\ell^{100}} \leqslant \eta^{100}.$$

Putting everything together, we get that

$$\mathcal{D}(E) \leqslant \mathcal{D}(E_1) + \mathcal{D}(E_2) \leqslant \eta^{40} \quad \text{and} \quad \mathcal{D}'(E) \leqslant \mathcal{D}'(E_1) + \mathcal{D}'(E_2) \leqslant \eta^{40}. \tag{39}$$

We next show that the probability measures $\mathcal{D}$ and $\mathcal{D}'$ assign roughly the same measure to each $x \notin E$.

Fix $(x_1,\ldots,x_{2\ell}) \notin E$. It is clear that $\mathcal{D}(x_1,\ldots,x_{2\ell}) = q^{-2\ell\cdot 3k}$, where we use $|U| = 3k$. Let $s = s(x_1,\ldots,x_{2\ell})$ and $s' = s'(x_1,\ldots,x_{2\ell})$. Then,

$$\begin{aligned}
\mathcal{D}'(x_1,\ldots,x_{2\ell}) &= ((1-\beta)q^{-3\cdot 2\ell})^{k-s-s'}\left((1-\beta)q^{-3\cdot 2\ell} + \frac{\beta}{3}q^{-4\ell}\right)^{s}\left((1-\beta)q^{-3\cdot 2\ell} + \beta q^{-4\ell}\right)^{s'} \\
&= q^{-2\ell\cdot 3k}(1-\beta)^{k-s-s'}\left(1-\beta+\frac{\beta}{3}q^{2\ell}\right)^{s}(1-\beta+\beta q^{2\ell})^{s'}
\end{aligned} \tag{40}$$

In the first equality, the first term is the probability of choosing the blocks that have three distinct columns. Then, $(1-\beta)$ is the probability that no variables are dropped, and $q^{-3(2\ell)}$ is the probability of choosing those three particular x_i's in that block. The second term is the probability of choosing the blocks that have exactly two equal columns. Then, $(1-\beta)q^{-3(2\ell)}$ is the probability of having no variables dropped and choosing the three x_i's, and $\frac{\beta}{3}q^{-4\ell}$ is the probability of first having the variable dropped in the column that is not equal to the other two, and then choosing the correct values for the remaining two column values. The third term similarly corresponds to blocks that have all columns being equal.

We start by showing

$$\frac{\mathcal{D}'(x_1,\ldots,x_{2\ell})}{\mathcal{D}(x_1,\ldots,x_{2\ell})} \geqslant \frac{1}{1.1}.$$

Using (40),

$$
\begin{aligned}
\frac{\mathcal{D}'(x_1,\dots,x_{2\ell})}{\mathcal{D}(x_1,\dots,x_{2\ell})} &= (1-\beta)^{k-s-s'}\left(1-\beta+\frac{\beta}{3}q^{2\ell}\right)^{s}(1-\beta+\beta q^{2\ell})^{s'} \\
&\geqslant (1-\beta)^{k-s}\left(1-\beta+\frac{\beta}{3}q^{2\ell}\right)^{s} \\
&= (1-\beta)^{k-s}\left(1+\beta\left(\frac{q^{2\ell}}{3}-1\right)\right)^{s} \\
&\geqslant (1-\beta(k-s))\left(1+\beta s\left(\frac{q^{2\ell}}{3}-1\right)\right) \\
&\geqslant 1-\beta k+\beta s\frac{q^{2\ell}}{3}.
\end{aligned}
$$

where in the fourth transition we use the bound $(1+z)^n \geqslant 1+nz$ which holds for all $n \geqslant 1$ and $z \geqslant -1$.

Now write $s = pk - v$ and let us analyze the last line. Plugging this in and using the definition of p we get that

$$
1-\beta k+\frac{q^{2\ell}}{3}\beta s = 1-\beta k+\frac{q^{2\ell}}{3}\beta(pk-v) = 1-\beta k\left(1-\frac{q^{2\ell}}{3}p\right)-\frac{q^{2\ell}}{3}\beta v = 1-\beta k q^{-2\ell}-\frac{q^{2\ell}}{3}\beta v. \quad (41)
$$

Hence, plugging in our values for β, k and p, we get

$$
\frac{\mathcal{D}'(x_1,\dots,x_{2\ell})}{\mathcal{D}(x_1,\dots,x_{2\ell})} \geqslant 1-\beta k q^{-2\ell}-\frac{q^{2\ell}}{3}\beta v = 1-q^{-2\ell+2c\ell/3}-\frac{q^{-4c\ell/3}v}{3} \geqslant \frac{1}{1.1},
$$

In the last transition we use $v \leqslant 50\sqrt{pk\log(1/\eta)} \leqslant q^{(c+o(1))\ell}$ (where above and henceforth the $o(1)$ terms are as ℓ goes to infinity) and the fact that ℓ is sufficiently large.

For the other direction, we show

$$
\frac{\mathcal{D}'(x_1,\dots,x_{2\ell})}{\mathcal{D}(x_1,\dots,x_{2\ell})} \leqslant \frac{1}{0.9},
$$

in nearly the same fashion. First note that

$$
\left(\frac{1-\beta+\beta q^{2\ell}}{1-\beta}\right)^{s'} \leqslant \left(\frac{1-\beta+\beta q^{2\ell}}{1-\beta}\right)^{\ell^{100}} = \left(1+\frac{\beta}{1-\beta}q^{2\ell}\right)^{\ell^{100}} \leqslant 1+o(1), \quad (42)
$$

By (40), we have

$$
\begin{aligned}
\frac{\mathcal{D}'(x_1,\dots,x_{2\ell})}{\mathcal{D}(x_1,\dots,x_{2\ell})} &= (1-\beta)^{k-s-s'}\left(1-\beta+\frac{\beta}{3}q^{2\ell}\right)^{s}(1-\beta+\beta q^{2\ell})^{s'} \\
&\leqslant (1+o(1))\cdot(1-\beta)^{k-s}\left(1-\beta+\frac{\beta}{3}q^{2\ell}\right)^{s} \\
&= (1+o(1))\cdot(1-\beta)^{k-s}\left(1+\beta\left(\frac{q^{2\ell}}{3}-1\right)\right)^{s} \\
&\leqslant (1+o(1))\cdot\exp\left(-\beta(k-s)+\left(\frac{q^{2\ell}}{3}-1\right)\beta s\right) \\
&= (1+o(1))\cdot\exp\left(-\beta k+\frac{q^{2\ell}}{3}\beta s\right).
\end{aligned}
$$

where in the first transition we used (42), in the fourth transition we use the fact that $1+z \leqslant e^z$. Writing $s = pk + v$ and using (41) (but with $+v$ instead of $-v$) and $v \leqslant O\left(q^{(c/2+o(1))\ell}\right)$ we have

$$\frac{\mathcal{D}'(x_1,\ldots,x_{2\ell})}{\mathcal{D}(x_1,\ldots,x_{2\ell})} \leqslant (1+o(1))\exp\left(-\beta k q^{-2\ell} + \frac{q^{2\ell}}{3}\beta v\right) \leqslant (1+o(1))\exp\left(O(q^{-c\ell/3-o(1)})\right) \leqslant \frac{1}{0.9}. \quad \square$$

D List Decoding Bound

In this section we prove Lemma 5.25, which is a direct consequence of the generic list decoding bound of [GRS00]:

Theorem D.1. *[GRS00, Theorem 15] Let $\mathcal{C} \subseteq \Sigma^N$ be a code with alphabet size $M := |\Sigma|$, blocklength N, and relative distance $1-\gamma$. Let $\delta > 0$ and $R \in \Sigma^N$. Suppose that $C_1,\ldots,C_m \in \Sigma^N$ are distinct codewords from $\mathcal{C}$ that each differ from R on at most $(1-\delta)$-fraction of entries. If*

$$\delta > \frac{1}{M} + \sqrt{\left(\gamma - \frac{1}{M}\right)\left(1-\frac{1}{M}\right)},$$

then

$$m \leqslant \frac{1}{(\delta - 1/M)^2 - (1-1/M)(\gamma - 1/M)}.$$

Proving Lemma 5.25 is simply a matter of translating to the notation of Theorem D.1, and below are the details.

Proof of Lemma 5.25. Let $\Sigma = \mathbb{F}_q^{2\ell-r_1}$ and write Q as $Q := \mathrm{span}(z_1,\ldots,z_{r_1})$. Define a code $\mathcal{C} \subseteq \{P\colon W^{2\ell-r_1} \to \mathbb{F}_q^{2\ell-r_1}\}$ consisting of $P : W^{2\ell-r_1} \to \mathbb{F}_q^{2\ell-r_1}$ such that there exists $v \in W$ satisfying

$$P(x_1,\ldots,x_{2\ell-r_1}) = (v\cdot x_1,\ldots,v\cdot x_{2\ell-r_1})$$

for all $x_1,\ldots,x_{2\ell-r_1} \in W^{2\ell-r_1}$.

Note that for distinct $v, w \in W$ we have that $v\cdot x = w\cdot x$ for at most $1/q$-fraction of $x \in W$. Thus, the relative distance of $\mathcal{C}$ is $1 - q^{-2\ell+r_1}$. We would like the table T corresponds to a word, say P', and $f_1,\ldots,f_m$ correspond to m codewords in $\mathcal{C}$, say $C_1,\ldots,C_m$. A slight issue is that T is only defined over 2ℓ-dimensional subspaces of $L \in \mathsf{Zoom}_{2\ell}[Q,W]$, while P has an entry for every $(2\ell-r_1)$-tuple of points in W. To resolve this, note that nearly every $(2\ell-r_1)$-tuple of points combined with $z_1,\ldots,z_{r_1}$ span an $L \in \mathsf{Zoom}_{2\ell}[Q,W]$. Thus, define P as follows. If $(z_1,\ldots,z_{r_1},x_1,\ldots,x_{2\ell-r_1})$ are linearly independent, then let L be the span of $(z_1,\ldots,z_{r_1},x_1,\ldots,x_{2\ell-r_1})$ and define

$$P'(x_1,\ldots,x_{2\ell-r_1}) = (T[L](x_1),\ldots,T[L](x_{2\ell-r_1}))\,.$$

Otherwise, define $P'(x_1,\ldots,x_{2\ell-r_1})$ arbitrarily. Note that the fraction of tuples $(x_1,\ldots,x_{2\ell-r_1})$ such that $(z_1,\ldots,z_{r_1},x_1,\ldots,x_{2\ell-r_1})$ are not linearly independent is at most,

$$\sum_{i=r_1+1}^{2\ell}\frac{q^{i-1}}{q^n} \leqslant q^{2\ell-n},$$

so nearly all of the entries in P' correspond to table entries in T. For the functions $f_1, \dots, f_m$ we define C_i corresponding to f_i by

$$C_i(x_1, \dots, x_{2\ell-r_1}) = (f_i(x_1), \dots, f_i(x_{2\ell-r_1})).$$

As each f_i agrees with T on at least β-fraction of the entries, we have that P' and C_i agree on at least β-fraction of the entries $(x_1, \dots, x_{2\ell-r_1})$ such that $(z_1, \dots, z_{r_1}, x_1, \dots, x_{2\ell-r_1})$ are linearly independent, so

$$\delta(P, C_i) \leqslant 1 - \beta \cdot (1 - q^{2\ell-n}) \leqslant 1 - \frac{\beta}{2}$$

for each $1 \leqslant i \leqslant m$. Finally, note that the alphabet size of $\mathcal{C}$ is $\left|\mathbb{F}_q^{2\ell-r_1}\right| = q^{2\ell-r_1}$. To bound m, we can apply Theorem D.1 with $\delta = \frac{\beta}{2} \geqslant q^{-2\ell+r_1} + \frac{c}{2}$, $M = q^{2\ell-r_1}$, and $\gamma = q^{-2\ell+r_1}$. We first note that the condition of Theorem D.1 is indeed satisfied:

$$\delta \geqslant q^{-2\ell+r_1} + \frac{c}{2} > q^{-2\ell+r_1} + 0.$$

Thus Theorem D.1 implies that $m \leqslant \frac{4}{c^2}$. $\square$

E Missing Proofs from Section 8

This section contains the missing proofs from Section 8, and we begin by recalling some notation. We recall the notation $\mu_{X,W}$, $\mu_{X,\circ}$, and $\mu_{\circ,W}$ from (2). Throughout this section, for a subspace L and a set of constant codimension subspaces $\mathcal{W}$, we use the notation $N_{\mathcal{W}}(L) = |\{W \in \mathcal{W} \mid L \subseteq W\}|$ from (13).

Let us also recall the setup in Section 8. We have the following items:

- $\delta > 0$ is a fixed small constant, ℓ is taken sufficiently large relative to $1/\delta$, r is a codimension parameter and is at most $10/\delta$, and

$$\xi := \delta^5, \quad \delta_2 := \xi/100, \quad t := \left(2^{2+10/\delta_2}\right)! \,.$$

- W'_{amb} is the ambient space and has dimension at least $n - 10/\delta$. Also, $n \geqslant 2^{100}q^{\ell}$ and ℓ is thought of as going to infinity relative to $1/\delta$. Broadly, we just want that n along with the dimension of W'_{amb} is "much larger" than ℓ.
- $\mathcal{W} = \{W_1, \dots, W_{m_1}\}$ is a set of t-generic subspaces of codimension s with respect to W'_{amb}, where $s \leqslant r$, and $m_1 \geqslant q^{75r\ell\xi^{-1}}$.
- T is a table defined over $\mathsf{Grass}(\mathbb{F}_q^n, 2\ell) \supseteq \mathsf{Grass}(W'_{\mathsf{amb}}, 2\ell)$. For this section, we mainly work in W'_{amb} and thus focus on T over $\mathsf{Grass}(W'_{\mathsf{amb}}, 2\ell)$.
- For each $i \in [m_1]$, we have a linear function $f_i : W_i \to \mathbb{F}_q$ such that

$$\Pr_{L \in \mathsf{Grass}(W, 2\ell)}[f_i|_L \equiv T[L]] \geqslant C \geqslant q^{-2(1-\xi)\ell}.$$

We are now ready to move onto the missing proofs for Section 8.

E.1 Proof of Lemma 8.1

This subsection is devoted to the proof of Lemma 8.1 which is restated below for convenience. We recall the following definitions for a subspace X (of dimension less than 2ℓ) and a linear assignment σ to X:

$$\mathcal{L}_X = \mathsf{Zoom}_{2\ell}[X, W'_{\mathsf{amb}}] \quad \text{and} \quad \mathcal{L}_{X,\sigma} = \{L \in \mathcal{L}_X \mid T[L]|_X \equiv \sigma\}.$$

Also,

$$\mathcal{W}_X = \{W_i \in \mathcal{W} \mid X \subseteq W_i\} \quad \text{and} \quad \mathcal{W}_{X,\sigma} = \{W_i \in \mathcal{W}_X \mid f_i|_X \equiv \sigma\}.$$

Finally, for the remainder of this subsection, we remind the reader that $\gamma = 10^{-6}$ is a small constant and $C \geqslant q^{-2(1-\xi)\ell}$, and $\xi > 0$ is the small constant fixed at the start of Section 8 (and recalled at the start of Appendix E).

Lemma 8.1. *There exists a $2\left(1 - \frac{\xi}{2}\right)\ell$-dimensional subspace X and a linear assignment, σ, to X, such that the following hold:*

- $\mu_{X,\circ}(\mathcal{L}_{X,\sigma}) \geqslant \frac{C}{6}$.
- $|\mathcal{W}_{X,\sigma}| \geqslant \frac{m_1}{q^{10r\ell}}$.
- *Choosing $L \in \mathcal{L}_{X,\sigma}$ uniformly, and $W_i \in \mathcal{W}_{X,\sigma}$ uniformly such that $W_i \supseteq L$, we have*
$$\Pr_{L,W_i}[f_i|_L \not\equiv T[L]] \leqslant 5\gamma.$$

For each $2\left(1 - \frac{\xi}{2}\right)\ell$-dimensional subspace X and linear assignment, σ, to X, let us define

$$p_{X,\sigma} := \Pr_{\substack{W_i \in \mathcal{W}_{X,\sigma} \\ L \in \mathsf{Zoom}_{2\ell}[X, W_i]}}[L \in \mathcal{L}_{X,\sigma} \wedge f_i|_L \neq T[L]], \qquad q_{X,\sigma} := \Pr_{\substack{W_i \in \mathcal{W}_{X,\sigma} \\ L \in \mathsf{Zoom}_{2\ell}[X, W_i]}}[L \in \mathcal{L}_{X,\sigma}],$$

where in both probabilities X and σ are fixed, and $W_i \in \mathcal{W}_{X,\sigma}$ is chosen uniformly and $L \in \mathsf{Zoom}_{2\ell}[X, W_i]$ is chosen uniformly. The intention behind these values is that for a fixed (X, σ), the quantity $p_{X,\sigma}$ should reflect how much disagreement there is between the table T and the functions f_i for $W_i \in \mathcal{W}_{X,\sigma}$, on subspaces $L \in \mathcal{L}_{X,\sigma}$, while $q_{X,\sigma}$ should reflect the size of $\mathcal{L}_{X,\sigma}$. Note that if $L \in \mathcal{L}_{X,\sigma}$ and $W_i \in \mathcal{W}_{X,\sigma}$, then by definition we already have $f_i|_X \equiv T[L]|_X \equiv \sigma$. Therefore we would expect that in fact f_i and $T[L]$ also agree on L (since it is only larger than X by $\xi\ell$ dimensions), meaning that we expect $p_{X,\sigma}$ to typically be small. Also, for each W_i, there are at least a C-fraction of $L \in \mathsf{Grass}(W_i, 2\ell)$ for which $f_i|_L \equiv T[L]$, so we would expect $q_{X,\sigma}$ to be $\Omega(C)$ for a non-trivial fraction of (X, σ). In the following claim, we formalize this intuition and show that there indeed exists an (X, σ) for which $p_{X,\sigma}$ is small, $q_{X,\sigma}$ is large, and additionally the set $\mathcal{W}_{X,\sigma}$ is large.

Remark E.1. *The idea of looking for such (X, σ) was first introduced in [IKW12] where they call these (X, σ)-excellent and was used again in [BDN17, MZ23] to analyze lower dimensional subspace versus subspace tests, which is similar in spirit to what we are ultimately trying to show in Lemma 5.24.*

Claim E.2. *There exists (X, σ) where X is a $2\left(1 - \frac{\xi}{2}\right)$-dimensional subspace and σ is a linear function on X such that:*

- $m_2 := |\mathcal{W}_{X,\sigma}| \geqslant \frac{m_1}{q^{10r\ell}}$.

- $q_{X,\sigma} \geqslant \frac{C}{2}$.
- $p_{X,\sigma} \leqslant \gamma \cdot q_{X,\sigma}$.

Proof. Consider the following process which outputs W_i, L, X, σ such that W_i is uniform in $\mathcal{W}$, $L \in \mathsf{Grass}(W_i, 2\ell)$ is uniform, $X \in \mathsf{Grass}\left(L, 2\left(1-\frac{\xi}{2}\right)\ell\right)$ is uniform, and σ is the assignment of f_i to X, i.e $\sigma :\equiv f_i|_X$.

1. Choose (X,σ) with probability proportional to $|\mathcal{W}_{X,\sigma}|$.
2. Choose $W_i \in \mathcal{W}_{X,\sigma}$ uniformly.
3. Choose a 2ℓ-dimensional subspace L uniformly conditioned on $X \subseteq L \subseteq W_i$.

Notice that the marginal distribution over (W_i, L) above is equivalent to that of choosing $W_i \in \mathcal{W}$ uniformly and $L \subset W_i$ uniformly. Furthermore, in the distribution above, conditioned on an L, $X \subseteq L$ is also uniform. Moreover, $f_i|_L \equiv T[L]$ only if $L \in \mathcal{L}_{X,\sigma}$, as f_i and $T[L]$ must agree on $X \subseteq L$ in order to agree on L. Therefore,

$$\mathop{\mathbb{E}}_{X,\sigma}[q_{X,\sigma}] \geqslant \Pr_{W_i \in \mathcal{W}, L \subseteq W_i}[f_i|_L \equiv T[L]] \geqslant C \geqslant \frac{1}{q^{2(1-\xi)\ell}}. \tag{43}$$

On the other hand,

$$\mathop{\mathbb{E}}_{X,\sigma}[p_{X,\sigma}] \leqslant \Pr_{X \subseteq L}[f_i|_X \equiv T[L]|_X \wedge f_i|_L \neq T[L]] \leqslant \frac{1}{q^{2(1-\xi/2)\ell}}. \tag{44}$$

Here the distribution of (X,σ) is proportional to the sizes $|\mathcal{W}_{X,\sigma}|$ and the second inequality is by the Schwartz-Zippel lemma. Indeed, by the Schwartz-Zippel lemma, if $f_i|_L$ and $T[L]$ are distinct then they agree on at most $1/q$-fraction of points z in L. Therefore, the middle term is bounded by the probability that $2(1-\xi/2)\ell$ uniformly random, linearly independent points are all chosen in this $1/q$-fraction.

Then,

$$\mathop{\mathbb{E}}_{X,\sigma}[q_{X,\sigma} - \gamma p_{X,\sigma}] \geqslant C - \frac{\gamma}{q^{2(1-\xi/2)\ell}} \geqslant 0.99C$$

By Markov's inequality, we get that for at least $(0.49C)$-fraction of (X,σ) we have, $q_{X,\sigma} - \gamma p_{X,\sigma} \geqslant \frac{C}{2}$, and hence for at least $(0.49C)$-fraction of (X,σ) we have both $q_{X,\sigma} \geqslant C/2$ and $p_{X,\sigma} \leqslant \gamma q_{X,\sigma}$.

Next we wish to argue that for most of these (X,σ)'s, $|\mathcal{W}_{X,\sigma}|$ is large. First note that the total number of (X,σ)'s is $\begin{bmatrix} n \\ 2(1-\xi/2)\ell \end{bmatrix}_q q^{2\left(1-\frac{\xi}{2}\right)\ell}$. For a fixed (X,σ), the probability that it is chosen is precisely,

$$\frac{|\mathcal{W}_{X,\sigma}|}{m_1} \cdot \frac{1}{\begin{bmatrix} n-s \\ 2(1-\xi/2)\ell \end{bmatrix}_q}. \tag{45}$$

Therefore,

$$\Pr_{X,\sigma}\left[|\mathcal{W}_{X,\sigma}| \leqslant \frac{m_1}{q^{10r\ell}}\right] \leqslant \frac{1}{q^{10r\ell}} \cdot \frac{\begin{bmatrix} n \\ 2(1-\xi/2)\ell \end{bmatrix}_q q^{2\ell}}{\begin{bmatrix} n-r \\ 2(1-\xi/2)\ell \end{bmatrix}_q} \leqslant \frac{1}{q^{10r\ell}} \cdot q^{4r\ell} \cdot q^{2\ell} \leqslant \frac{1}{q^{5r\ell}}. \tag{46}$$

In the first transition we union bound over the probability of choosing (X,σ) for all (X,σ) such that $|\mathcal{W}_{X,\sigma}| \leqslant \frac{m_1}{q^{10r\ell}}$. Using (45), this probability is at most,

$$\frac{1}{q^{10r\ell}} \cdot \frac{1}{\begin{bmatrix} n-s \\ 2(1-\xi/2)\ell \end{bmatrix}_q} \leqslant \frac{1}{q^{10r\ell}} \cdot \frac{1}{\begin{bmatrix} n-r \\ 2(1-\xi/2)\ell \end{bmatrix}_q}$$

for each (X,σ) (recall that $s \leqslant r$), and we have to union bound over at most $\begin{bmatrix} n \\ 2(1-\xi/2)\ell \end{bmatrix}_q q^{2\left(1-\frac{\xi}{2}\right)\ell}$ many (X,σ)'s. Altogether, it follows that with probability at least

$$0.49C - \frac{1}{q^{5r\ell}} > 0,$$

over (X,σ), we have, $q_{X,\sigma} \geqslant \tau$, $p_{X,\sigma} \leqslant \gamma\tau$, and $|\mathcal{W}_{X,\sigma}| \geqslant \frac{m_1}{q^{10r\ell}}$, which establishes the claim. $\square$

Taking the (X,σ) given by Claim E.2 almost works for Lemma 8.1. However, notice that while the probability of interest for the third item there looks similar to $p_{X,\sigma}$, it has a different distribution over L and W_i. Indeed, in Lemma 8.1 one first chooses $L \in \mathcal{L}_{X,\sigma}$ and then $W_i \in \mathcal{W}_{X,\sigma}$ containing L, whereas for $p_{X,\sigma}$, we are first choosing $W_i \in \mathcal{W}_{X,\sigma}$ and then $L \subseteq W_i$ (and in particular we do not condition on L being in the set $\mathcal{L}_{X,\sigma}$). Intuitively, we expect the following three quantities to be roughly equal

$$\Pr_{L\in\mathcal{L}_{X,\sigma}, W_i\in\mathcal{W}_{X,\sigma}}[f_i|_L \neq T[L] \mid W_i \supseteq L] \approx \Pr_{W_i\in\mathcal{W}_{X,\sigma}, L\subseteq W_i}[f_i|_L \neq T[L] \mid L \in \mathcal{L}_{X,\sigma}] \approx \frac{p_{X,\sigma}}{q_{X,\sigma}}.$$

Once we establish the above, the third item of Claim E.2 yields $\frac{p_{X,\sigma}}{q_{X,\sigma}} \leqslant \gamma$, which completes the proof of Lemma 8.1. Thus, the bulk of the transition from Claim E.2 to Lemma 8.1 is in converting from the distribution used in $p_{X,\sigma}$ to that required by the third item of Lemma 8.1 without losing too much. The rest of the argument is devoted to this goal.

Proof of Lemma 8.1. Fix (X,σ) from Claim E.2 and define $\mathcal{W}_{X,\sigma}$ and $\mathcal{L}_{X,\sigma}$ accordingly. Let $m_2 := |\mathcal{W}_{X,\sigma}| \geqslant \frac{m_1}{q^{10r\ell}}$. In order to lower bound $\mu_{X,\circ}(\mathcal{L}_{X,\sigma})$ we use Lemma 5.18 on the collection of subspaces $\mathcal{W}_{X,\sigma}$ with parameters $j = 2\ell$, $a = 2\left(1-\frac{\xi}{2}\right)\ell$. Indeed, consider the measure $\nu_{\mathcal{W}_{X,\sigma}}$ over $\mathsf{Zoom}_{2\ell}[X, W'_{\mathsf{amb}}]$, which, recall, is obtained by choosing $W_i \in \mathcal{W}_{X,\sigma}$ uniformly and then a uniform subspace in $\mathsf{Zoom}_{2\ell}[X, W_i]$, and note that this measure is precisely the measure corresponding to $q_{X,\sigma}$. Thus, applying Lemma 5.18 we get

$$\mu_{X,\circ}(\mathcal{L}_{X,\sigma}) \geqslant \nu_{\mathcal{W}_{X,\sigma}}(\mathcal{L}_{X,\sigma}) - \frac{3q^{\frac{s}{2}\xi\ell}}{\sqrt{m_2}} = q_{X,\sigma} - \frac{3q^{\frac{s}{2}\xi\ell}}{\sqrt{m_2}} \geqslant \frac{C}{6}.$$

Define $\tau := q_{X,\sigma}$ so that $\tau \geqslant C/2$ and $p_{X,\sigma} < \gamma \cdot \tau$. Note that the first two conditions of Lemma 8.1 are already satisfied, so it remains to check the third condition.

To this end it will be helpful to have in mind the bipartite graph with parts $\mathcal{W}_{X,\sigma}$ and $\mathcal{L}_{X,\sigma}$ and edges (W_i, L) if $L \subseteq W_i$. For each $L' \in \mathcal{L}_X$ and $W_i \in \mathcal{W}_{X,\sigma}$ define the following degree-like quantities:

- $N_{\mathcal{W}_{X,\sigma}}(L') := |\{W_i \in \mathcal{W}_{X,\sigma} \mid W_i \supseteq L'\}|$,
- $e_{L'} := |\{W_i \in \mathcal{W}_{X,\sigma} \mid W_i \supseteq L', f_i|_{L'} \neq T[L']\}|$,

- $d_i := |\{L \in \mathcal{L}_{X,\sigma} \mid L \subseteq W_i\}|$,
- $e_i := |\{L \in \mathcal{L}_{X,\sigma} \mid L \subseteq W_i, f_i|_L \neq T[L]\}|$.

Note that the first quantity is the same as in (13). We also let $D = |\{L \mid X \subseteq L \subseteq W_i, \dim(L) = 2\ell\}|$, where the $W_i \in \mathcal{W}_{X,\sigma}$ is arbitrary (the value is the same regardless which we pick). Then we clearly have that $\mathbb{E}_{L\in\mathcal{L}_X}[N_{\mathcal{W}_{X,\sigma}}(L)] = \frac{m_2 D}{|\mathcal{L}_X|}$ and the probability that we are interested in can be expressed as:

$$\Pr_{L\in\mathcal{L}_{X,\sigma}, W_i\in\mathcal{W}_{X,\sigma}}[f_i|_L \neq T_1[L] \mid L \subseteq W_i] = \mathop{\mathbb{E}}_{L\in\mathcal{L}_{X,\sigma}}\left[\frac{e_L}{N_{\mathcal{W}_{X,\sigma}}(L)}\right].$$

Since $q_{X,\sigma} = \tau$ and $p_{X,\sigma} \leqslant \gamma\cdot\tau$, we have

$$q_{X,\sigma}\cdot m_2 \cdot D = \sum_{L\in\mathcal{L}_{X,\sigma}} N_{\mathcal{W}_{X,\sigma}}(L) = \sum_{W_i\in\mathcal{W}_{X,\sigma}} d_i = m_2\cdot D\cdot\tau,$$

and

$$p_{X,\sigma}\cdot m_2\cdot D = \sum_{L\in\mathcal{L}_{X,\sigma}} e_L = \sum_{W_i\in\mathcal{W}_{X,\sigma}} e_i \leqslant m_2\cdot D\cdot\gamma\cdot\tau. \tag{47}$$

By Lemma 5.21, along with the fact that $\mathbb{E}_{L\in\mathcal{L}_X}[N_{\mathcal{W}_{X,\sigma}}(L)] = \frac{m_2 D}{|\mathcal{L}_X|}$, and the very loose bound $\frac{D}{|\mathcal{L}_X|} \geqslant \frac{1}{q^{r\cdot\ell}}$, we have

$$\Pr_{L\in\mathcal{L}_X}\left[N_{\mathcal{W}_{X,\sigma}}(L) \leqslant 0.9 \mathop{\mathbb{E}}_{L\in\mathcal{L}_X}[N_{\mathcal{W}_{X,\sigma}}(L)]\right] = \Pr_{L\in\mathcal{L}_X}\left[N_{\mathcal{W}_{X,\sigma}}(L) \leqslant \frac{0.9\cdot m_2 D}{|\mathcal{L}_X|}\right] \leqslant \frac{101q^{r\cdot\xi\ell}}{m_2}. \tag{48}$$

We conclude that

$$\begin{aligned}
\mathop{\mathbb{E}}_{L\in\mathcal{L}_{X,\sigma}}\left[\frac{e_L}{N_{\mathcal{W}_{X,\sigma}}(L)}\right] &\leqslant \Pr_{L\in\mathcal{L}_{X,\sigma}}\left[N_{\mathcal{W}_{X,\sigma}}(L) \leqslant \frac{0.9\cdot m_2\cdot D}{|\mathcal{L}_X|}\right] + \mathop{\mathbb{E}}_{L\in\mathcal{L}_{X,\sigma}}\left[\frac{e_L}{0.9\cdot m_2\cdot D/|\mathcal{L}_X|}\right]\\
&\leqslant \frac{101q^{r\cdot\xi\ell}/m_2}{\Pr_{L\in\mathcal{L}_X}[L\in\mathcal{L}_{X,\sigma}]} + \mathop{\mathbb{E}}_{L\in\mathcal{L}_{X,\sigma}}\left[\frac{e_L}{0.9\cdot m_2\cdot D/|\mathcal{L}_X|}\right]\\
&\leqslant \frac{101q^{r\cdot\xi\ell}/m_2}{C/6} + \mathop{\mathbb{E}}_{L\in\mathcal{L}_{X,\sigma}}\left[\frac{e_L|\mathcal{L}_X|}{0.9\cdot m_2\cdot D}\right]\\
&\leqslant \frac{606q^{r\cdot\ell}}{m_2\cdot C} + \frac{m_2\cdot D\cdot\gamma\cdot\tau}{0.9\cdot m_2\cdot D}\cdot\frac{|\mathcal{L}_X|}{|\mathcal{L}_{X,\sigma}|}\\
&\leqslant \frac{606q^{r\cdot\ell}}{m_2\cdot C} + \frac{\gamma\tau}{0.9}\cdot\frac{3}{\tau}\\
&\leqslant 5\gamma,
\end{aligned}$$

where in the second transition we used (48) and in the fourth transition we used (47). □

E.2 Proof of Lemma 8.3

Take the $X, \sigma, \mathcal{L}'_{X,\sigma}$, and $\mathcal{W}_{X,\sigma}$ from Corollary 8.2, and recall W'_{amb} is the ambient space and $\delta_2 = \frac{\xi}{100}$. We next recall a few notations: for a zoom-in A and zoom-out B such that $X \subseteq A \subseteq B \subseteq W'_{\mathsf{amb}}$, we write $W'_{\mathsf{amb}} = A \oplus W_{\mathsf{amb},0}$ and $B = A \oplus W^\star_{\mathsf{amb}}$, where $W^\star_{\mathsf{amb}} \subseteq W_{\mathsf{amb},0}$. Now define

$$\mathcal{W}^\star_{[A,B]} := \{W^\star_i \mid \exists W_i \in \mathcal{W}_{X,\sigma} \text{ s.t } A \oplus W^\star_i = W_i \cap B\}.$$

It is clear that each $W_i^\star \in \mathcal{W}_{[A,B]}^\star$ is contained inside of some $W_i \in \mathcal{W}_{X,\sigma}$, so for each $W_i^\star$, we may define $f_i^\star \equiv f_i|_{W_i^\star}$. If there are multiple such i, we choose one arbitrarily to define $f_i^\star$. With this in mind, we restate Lemma 8.3.

Lemma 8.3. *Keeping the notation above, there is a zoom-in A and a zoom-out B inside of the space W'_{amb} such that the following holds. There exists a collection of subspaces $\mathcal{W}^\star = \{W_1^\star, \dots, W_{m_3}^\star\} \subseteq \mathcal{W}_{[A,B]}^\star$ of codimension s with respect to $W_{\mathsf{amb}}^\star$, a dimension $\ell' \geqslant \frac{\xi}{3}\ell$, and a set $\mathcal{L}^\star \subseteq \mathsf{Grass}_q(W_{\mathsf{amb}}^\star, \ell')$, such that the following items hold for the table $T_{\ell'}$, which assigns linear functions to $L \in \mathsf{Grass}_q(W_{\mathsf{amb}}^\star, \ell')$ as follows*

$$T_{\ell'}[L] :\equiv T[A \oplus L]|_L.$$

1. *$\mu(\mathcal{L}^\star) := \eta \geqslant \frac{C}{12}$, where the measure here is over $\mathsf{Grass}_q(W_{\mathsf{amb}}^\star, \ell')$.*
2. *The set $\mathcal{L}^\star$ is $(1, q^{\delta_2 \ell}\eta)$-pseudo-random.*
3. *Each $W_i^\star$ has codimension $s \leqslant r$ inside of $W_{\mathsf{amb}}^\star$, $\mathcal{W}^\star$ is 4-generic, with respect to $W_{\mathsf{amb}}^\star$, and $W_{\mathsf{amb}}^\star$ has codimension at most $10/\delta_2$ with respect to W'_{amb}.*
4. *The size of $\mathcal{W}^\star$ satisfies*
$$\frac{m_2 \cdot q^{-10s/\delta_2}}{2} \leqslant m_3 \leqslant m_2.$$
5. *For each $W_i^\star \in \mathcal{W}_i^\star$, there is a linear function $f_i^\star : W_i^\star \to \mathbb{F}_q$ such that the following holds. For every $L \in \mathcal{L}^\star$, choosing $W_i^\star \in \mathcal{W}^\star$ uniformly such that $W_i^\star \supseteq L$, we have*
$$\Pr_{W_i^\star \supseteq L, W_i^\star \in \mathcal{W}^\star}[f_i^\star|_L \not\equiv T_{\ell'}[L]] \leqslant 14\gamma.$$
6. *For every $L \in \mathcal{L}^\star$,*
$$0.8 \cdot p_1 \cdot m_3 \leqslant N_{\mathcal{W}^\star}(L) \leqslant 1.2 \cdot p_1 \cdot m_3,$$
where $N_{\mathcal{W}^\star}(L)$ is as defined in (13)*, and*
$$p_1 := \Pr_{L \in \mathsf{Grass}_q(W_{\mathsf{amb}}^\star, \ell')}[L \subseteq W],$$
for an arbitrary $W \subseteq W_{\mathsf{amb}}^\star$ of codimension s.

The following result finds the zoom-in and zoom-out pair as required for Lemma 8.3, modulo a few minor alterations.

Lemma E.3. *We can find a zoom-in A and a zoom-out B such that $X \subseteq A \subseteq B \subseteq W'_{\mathsf{amb}}$, such that the following hold.*

- $\dim(A) + \operatorname{codim}(B) \leqslant \dim(X) + \frac{10}{\delta_2}$.
- *The set of subspaces $\mathcal{L}' := \mathcal{L}'_{X,\sigma} \cap \mathsf{Zoom}_{2\ell}[A, B]$ satisfies $\eta := \mu_{A,B}(\mathcal{L}') \geqslant \frac{C}{12}$ and $\mathcal{L}'$ is $(1, q^{\delta_2 \ell}\eta)$-pseudo-random* .[11]

[11] By $(1, q^{\delta_2 \ell}\eta)$-pseudo-random in $\mathsf{Zoom}_{2\ell}[A, B]$ we mean that $\mathcal{L}'$ does not increase its fractional size to $q^{\delta_2 \ell}\eta$ when restricted to any zoom-in containing A or any zoom-out contained in B.

Proof. Set $A_0 = X$, $\mathcal{L}_0 = \mathcal{L}_{X,\sigma}$, $B_0 = V$, and $\eta_0 = \mu_{X,\circ}(\mathcal{L}_{X,\sigma}) \geqslant \frac{C}{12}$. Now do the following.

1. Set $i = 0$, and initialize $A_0, \mathcal{L}_0, B_0, \eta_0$ as above.
2. If $\mathcal{L}_i$ is $(1, q^{\delta_2 \ell}\eta_i)$-pseudo-random inside of $\mathsf{Zoom}_{2\ell}[A_i, B_i]$, then stop.
3. Otherwise, there exist $A \subseteq B$ such that, $A_i \subseteq A \subseteq B \subseteq B_i$, $\dim(A) + \mathrm{codim}(B) = \dim(A_i) + \mathrm{codim}(B_i) + 1$, and $\mu_{A,B}(\mathcal{L}_i) \geqslant q^{\delta_2 \ell}\eta_i$.
4. Set $A_{i+1} := A$, $B_{i+1} := B$, $\mathcal{L}_{i+1} := \mathcal{L}_i \cap \mathsf{Zoom}_{2\ell}[A_{i+1}, B_{i+1}]$, and let $\eta_{i+1} := \mu_{A_{i+1},B_{i+1}}(\mathcal{L}_{i+1})$.
5. Increment i by 1 and return to step 2.

Suppose this process terminates on iteration j. We claim that taking $\mathcal{L}' = \mathcal{L}_j$, $A = A_j$, and $B = B_j$, satisfies the requirements of the lemma.

First, note that by construction $\eta_{i+1} \geqslant q^{\delta_2 \ell}\eta_i$. Therefore, we perform at most $\left\lceil \frac{\log(12/C)}{\log(q^{\delta_2 \ell})} \right\rceil \leqslant \frac{10}{\delta_2}$ iterations before stopping, so $j \leqslant \frac{10}{\delta_2}$. By construction $\mathcal{L}_j$ is $(1, q^{\delta_2 \ell}\eta_j)$-pseudo-random in $\mathsf{Zoom}_{2\ell}[A_j, B_j]$ and has fractional size $\eta_j \geqslant \mu_{X,\sigma}(L_{X,\sigma}) \geqslant \frac{C}{12}$ in $\mathsf{Zoom}_{2\ell}[A_j, B_j]$. Moreover, $\dim(A_j) + \mathrm{codim}(B_j) = \dim(X) + j \leqslant \dim(X) + \frac{10}{\delta_2}$, so the conditions of the lemma are satisfied. □

Take A, B and $\mathcal{L}'$ given by Lemma E.3, as well as $\mathcal{W}_{X,\sigma}$ from Lemma 8.1. We are now ready to prove Lemma 8.3.

Proof of Lemma 8.3. With A and B set as above, we now construct the $\mathcal{L}^\star, \mathcal{W}^\star$, that satisfy Lemma 8.3. Let $W^\star_{\mathsf{amb}}$ be a subspace such that $A \oplus W^\star_{\mathsf{amb}} = B$, set $\ell' = 2\ell - \dim(A)$, and let

$$\mathcal{L}^\star = \{L^\star \in \mathsf{Grass}_q(W^\star_{\mathsf{amb}}, \ell') \mid L^\star \oplus A \in \mathcal{L}'\}.$$

For each $L^\star \in \mathcal{L}^\star$, let L' denote the corresponding subspace such that $A \oplus L^\star = L' \in \mathcal{L}'$. Note the correspondence,

$$L^\star \in \mathcal{L}^\star \longleftrightarrow L' = A \oplus L^\star \in \mathcal{L}', \tag{49}$$

is a bijection between $\mathcal{L}^\star$ and $\mathcal{L}'$ because every subspace in $\mathcal{L}'$ contains A. Abusing notations we let T denote both the original table on 2ℓ-dimensional subspaces, as well as the new table on $\mathsf{Grass}(W^\star_{\mathsf{amb}}, \ell')$, given by $T[L^\star] \equiv T[L']|_{L^\star}$ for each $L^\star \in \mathsf{Grass}(W^\star_{\mathsf{amb}}, \ell')$. It will always be clear, based on the argument in $T[\cdot]$, which assignment we are referring to.

We obtain $\mathcal{W}^\star$ in a similar way as $\mathcal{L}^\star$, however, some care will be needed to ensure that it is 4-generic. Starting with $\mathcal{W}_{X,\sigma}$, we set

$$\mathcal{W}_{X,\sigma,\mathcal{L}'} = \{W \in \mathcal{W}_{X,\sigma} \mid \exists L \in \mathcal{L}', W \supseteq L\}.$$

That is, $\mathcal{W}_{X,\sigma,\mathcal{L}'}$, consists of the subspaces in $\mathcal{W}_{X,\sigma}$ containing at least one $L \in \mathcal{L}'$. Since $A \subseteq L$ for all $L \in \mathcal{L}'$, we are guaranteed at every $W \in \mathcal{W}_{X,\sigma,\mathcal{L}'}$, contains A. Now, take the collection of subspaces $\{W \cap B \mid W \in \mathcal{W}_{X,\sigma,\mathcal{L}'}\}$. This collection may not be 4-generic, and moreover may even contain duplicate subspaces, but we take $\widetilde{\mathcal{W}^\star}$ to be its *largest* subset that is 4-generic with codimension s with respect to B (note that being 4-generic implies that every subspace is distinct). Finally, set

$$\mathcal{W}^\star = \{W^\star_i \subseteq W^\star_{\mathsf{amb}} \mid A \oplus W^\star_i \in \widetilde{\mathcal{W}^\star}\},$$

and set $m_3 := |\mathcal{W}^\star|$. For each $W_i^\star$, choose an arbitrary $W_i \in \mathcal{W}_{X,\sigma}$ such that $W_i^\star = W_i \cap B$ and set $f_i^\star :\equiv f_i|_{W_i^\star}$. Summarizing, we have the following chain of relations, which took us from $\mathcal{W}_{X,\sigma}$ to $\mathcal{W}^\star$:

$$\mathcal{W}_{X,\sigma} \xrightarrow{\supseteq L} \mathcal{W}_{X,\sigma,\mathcal{L}'} \xrightarrow{\cap B \text{ and make 4-generic}} \widetilde{\mathcal{W}^\star} \xleftrightarrow{\text{Subtract subspace } A} \mathcal{W}^\star \tag{50}$$

$$W_i \longrightarrow W_i \longrightarrow W_i \cap B \longrightarrow W_i^\star \text{ s.t } W_i^\star \oplus A = W_i \cap B \tag{51}$$

It will be helpful to refer back to the chains of relations above. The double arrow transitions are bijections, while in the single arrow transitions subspaces are being removed. The second line shows what a generic member of each set looks like, where W_i are the original subspaces in $\mathcal{W}_{X,\sigma}$. It is clear from (50) that $\mathcal{W}^\star \subseteq \mathcal{W}_{[A,B]}$. When going from $\mathcal{W}_{X,\sigma,\mathcal{L}'}$, some subspaces are removed from the collection $\{W \cap B \mid W \in W_{X,\sigma,\mathcal{L}'}\}$, and we use Lemma 5.17 to upper bound the number of removed subspaces. In particular, since $\mathcal{W}_{X,\sigma}$ is $2^{2+10/\delta_2}$ generic and since by Lemma E.3 B has codimension at most

$$10/\delta_2 + (\dim(X) - \dim(A)) \leqslant 10/\delta_2 \tag{52}$$

in W'_{amb}, we get that

$$0 \leqslant |\mathcal{W}_{X,\sigma,\mathcal{L}'}| - |\widetilde{\mathcal{W}^\star}| \leqslant 2^{2+10/\delta_2}. \tag{53}$$

We remark that this argument also establishes the desired codimension of $W^\star_{\mathsf{amb}}$ inside of W'_{amb} for property 3. We now verify that the six properties of Lemma 8.3 hold.

Property 1. The subspaces of $\mathcal{L}^\star$ are of dimension ℓ' inside $W^\star_{\mathsf{amb}}$, and

$$\ell' = 2\ell - \dim(A) \geqslant 2\ell - \dim(X) - \frac{10}{\delta_2} \geqslant \frac{\xi}{3}\ell.$$

Also, $\mu(\mathcal{L}^\star) = \eta$ is the same as the measure of $\mathcal{L}'$ inside $\mathsf{Zoom}_{2\ell}[A, B]$ due to the bijection between $\mathcal{L}^\star$ and $\mathcal{L}'$ noted in (49). Therefore $\mu(\mathcal{L}^\star) = \eta \geqslant \frac{C}{12}$ by the second part of Lemma E.3.

Property 2. Since $\mathcal{L}'$ does not increase its measure to $q^{\delta_2 \ell}\eta$ on any zoom-in containing A or zoom-out inside B, it follows that $\mathcal{L}^\star$ is $(1, q^{\delta_2 \ell}\eta)$-pseudo-random.

Property 3. By construction, $\widetilde{\mathcal{W}^\star}$ is 4-generic inside of B. Since $B = A \oplus W^\star_{\mathsf{amb}}$, and all $W_i \in \widetilde{\mathcal{W}^\star}$ contain A, it follows that $\mathcal{W}^\star$ is 4-generic inside of $W^\star_{\mathsf{amb}}$. Finally, the codimension of $W^\star_{\mathsf{amb}}$ inside of W'_{amb} was established in Equation (52).

Before showing the remaining properties, we will state a few useful inequalities. Fix an $L^\star \in \mathcal{L}^\star$ and let $L' = L^\star \oplus A$, so that $L' \in \mathcal{L}' \subseteq \mathcal{L}_{X,\sigma}$. We will use the fact that, by construction of $\mathcal{W}_{X,\sigma,\mathcal{L}'}$,

$$N_{\mathcal{W}_{X,\sigma}}(L') = N_{\mathcal{W}_{X,\sigma,\mathcal{L}'}}(L'), \qquad N_{\mathcal{W}^\star}(L^\star) = N_{\widetilde{\mathcal{W}^\star}}(L').$$

Fix $L' \subseteq B$. Note that for every $W \in \mathcal{W}_{X,\sigma}$ containing L', the subspace $W \cap B$ contains L' as well. Thus, for each $W \in \mathcal{W}_{X,\sigma}$ contributing to $N_{\mathcal{W}_{X,\sigma}}(L')$, the corresponding $W \cap B$ also contributes to $N_{\widetilde{\mathcal{W}^\star}}(L')$ unless it was removed in the second transition of (50). Hence,

$$N_{\mathcal{W}_{X,\sigma}}(L') - \left(|\mathcal{W}_{X,\sigma,\mathcal{L}'}| - |\widetilde{\mathcal{W}^\star}|\right) \leqslant N_{\widetilde{\mathcal{W}^\star}}(L') \leqslant N_{\mathcal{W}_{X,\sigma}}(L'),$$

so we have

$$N_{\mathcal{W}_{X,\sigma}}(L') - \left(|\mathcal{W}_{X,\sigma,\mathcal{L}'}| - |\widetilde{\mathcal{W}^\star}|\right) \leqslant N_{\mathcal{W}^\star}(L^\star) \leqslant N_{\mathcal{W}_{X,\sigma}}(L'). \tag{54}$$

Thus, combining this with the upper bound from (53) as well as the bounds on $N_{\mathcal{W}_{X,\sigma}(L')}$ from Corollary 8.2, we have,

$$0.95 \cdot m_2 \cdot q^{-\xi\ell\cdot s} - 2^{2+10/\delta_2} \leqslant N_{\mathcal{W}^\star}(L^\star) \leqslant 1.05 \cdot m_2 \cdot q^{-\xi\ell\cdot s}. \tag{55}$$

Since we fixed $L^\star \in \mathcal{L}^\star$ arbitrarily, note that (55) holds for every $L^\star \in \mathcal{L}^\star$.

Property 5. Fix $L^\star \in \mathcal{L}^\star$ and let $L' \in \mathcal{L}'$ be the subspace such that $L' = A \oplus L^\star$. Let $G_{\mathcal{W}_{X,\sigma}}(L') := |\{W_i \in \mathcal{W}_{X,\sigma} \mid f_i|_{L'} \not\equiv T[L']\}$, and let $G_{\mathcal{W}^\star}(L^\star) := |\{W_i^\star \in \mathcal{W}^\star \mid f_i^\star|_{L^\star} \not\equiv T[L^\star]\}$. By the third part of Corollary 8.2, we have that

$$\frac{G_{\mathcal{W}_{X,\sigma}}(L')}{N_{\mathcal{W}_{X,\sigma}}(L')} \leqslant 12\gamma. \tag{56}$$

To show property 5, we must bound $G_{\mathcal{W}^\star}(L^\star)/N_{\mathcal{W}^\star}(L^\star)$, and we will accomplish this by relating it to the above quantity. Recall that for every $W_i^\star \in \mathcal{W}^\star$ and associated function $f_i^\star$, there is a corresponding $W_i \in \mathcal{W}_{X,\sigma}$ such that $W_i^\star = W_i \cap B$, and $f_i^\star = f_i|_{W_i^\star}$. In addition, $T[L']|_{L^\star} = T[L^\star]$, so if $f_i^\star|_{L^\star} \not\equiv T[L^\star]$, then $f_i|_{L'} \not\equiv T[L']$. Hence,

$$G_{\mathcal{W}^\star}(L^\star) \leqslant G_{\mathcal{W}_{X,\sigma}}(L'),$$

and

$$\frac{G_{\mathcal{W}^\star}(L^\star)}{N_{\mathcal{W}^\star}(L^\star)} \leqslant \frac{G_{\mathcal{W}_{X,\sigma}}(L')}{N_{\mathcal{W}_{X,\sigma}}(L')} \cdot \frac{N_{\mathcal{W}_{X,\sigma}}(L')}{N_{\mathcal{W}^\star}(L^\star)} \leqslant 12\gamma \cdot \frac{N_{\mathcal{W}_{X,\sigma}}(L')}{N_{\mathcal{W}^\star}(L^\star)}, \tag{57}$$

where we use (56) in the last transition. It remains to bound the ratio on the right hand side above:

$$\frac{N_{\mathcal{W}_{X,\sigma}}(L')}{N_{\mathcal{W}^\star}(L^\star)} \leqslant 1 + \frac{N_{\mathcal{W}_{X,\sigma}}(L') - N_{\mathcal{W}^\star}(L^\star)}{N_{\mathcal{W}^\star}(L^\star)} \leqslant 1 + \frac{2^{2+10/\delta_2}}{N_{\mathcal{W}^\star}(L^\star)} \leqslant 1.1, \tag{58}$$

where in the second inequality we are using (54) along with the upper bound on $|\mathcal{W}_{X,\sigma,\mathcal{L}'}| - |\widetilde{\mathcal{W}^\star}|$ from (53). Combining (56), (57), and (58), we get

$$\frac{G_{\mathcal{W}^\star}(L^\star)}{N_{\mathcal{W}^\star}(L^\star)} \leqslant 12\gamma \cdot 1.1 \leqslant 14\gamma,$$

as desired.

Property 6. The only difference between property 6 and (55) is that property 6 is expressed in terms of m_3, while (55) is expressed in terms of m_2. From (55), we already have that all of the $N_{\mathcal{W}^\star}(L^\star)$ over $L^\star \in \mathcal{L}^\star$ are within a multiplicative factor of 1.11 of each other, so to conclude property 6, we just need to show that there is at least one $L^\star$ such that $N_{\mathcal{W}^\star}(L^\star)$ is in the interval $[0.95 \cdot p_1 \cdot m_3, 1.05 \cdot p_1 \cdot m_3]$ which follows by Lemma 5.21.

Property 4. Since there exists $L^\star$ such that the lower bound from (55) and the upper bound from property 6 hold, we have:

$$m_3 \geqslant \frac{1}{1.2 \cdot p_1} \cdot \left(0.95 \cdot m_2 \cdot q^{-\xi\ell\cdot s} - 2^{2+10/\delta_2}\right) \geqslant \frac{0.6 \cdot m_2}{q^{\xi\ell\cdot s} \cdot p_1}$$

By Lemma 5.19, we have that $p_1 \leqslant 1.01 q^{-s\ell'}$, so,

$$m_3 \geqslant \frac{0.6 \cdot m_2}{1.01 q^{\xi\ell\cdot s - s\ell'}} \geqslant \frac{0.6 \cdot m_2}{1.01 q^{s(\xi\ell - 2\ell + \dim(A))}} = \frac{0.6 \cdot m_2}{1.01 q^{s(\dim(A) - \dim(X))}} \geqslant \frac{m_2 q^{-10s/\delta_2}}{2}. \qquad \square$$

E.3 Proof of Lemma 8.5

This subsection is dedicated to proving Lemma 8.5, but before going into the proof we must first show a basic Fourier analytic fact. We will use some of the set-up from Appendix A. The proof of Lemma 8.5 is given after this necessary fact is established.

E.3.1 A Necessary Fourier Analytic Fact

Let $\mathcal{A} \subseteq \mathsf{Grass}(n, j)$ and let $\eta = \mu(\mathcal{A})$. We assume that n is sufficiently large relative to q and j and that η is not too small, say

$$n > q^j \quad \text{and} \quad \eta \geqslant 100jq^{j-n} \tag{59}$$

to be concrete. Define $F \colon \mathbb{F}_q^{n\times j} \to \{0,1\}$ by

$$F(x_1, \dots, x_j) = \begin{cases} 1, & \text{if } \mathrm{span}(x_1, \dots, x_j) \in \mathcal{A}, \\ 0, & \text{otherwise.} \end{cases}$$

In the next lemma, we use the characters χ_S for $S = (s_1, \dots, s_j) \in \mathbb{F}_q^{n\times j}$ defined as in (31), except we replace the dimension parameter 2ℓ therein by j. We will use the notation $S \subseteq W^{\perp}$ to mean that $s_i \in W^{\perp}$ for each i.

Lemma E.4. *Fix a subspace $W \subseteq \mathbb{F}_q^n$, then for any $S = (s_1, \dots, s_j) \in \mathbb{F}_q^{n\times j}$ we have,*

$$\mathop{\mathbb{E}}_{x_1,\dots,x_j \in W}[\chi_S(x_1, \dots, x_j)] = \begin{cases} 1, & \textit{if } S \subseteq W^{\perp}, \\ 0, & \textit{if } S \subsetneq W^{\perp}. \end{cases}$$

Proof. If $S \subseteq W^{\perp}$, then for any $x \in W$, we have $s_i \cdot x = 0$ for all $1 \leqslant i \leqslant j$, so the first case follows.

Now suppose $S \subsetneq W^{\perp}$, and without loss of generality say that $s_1 \notin W^{\perp}$. We can write,

$$\mathop{\mathbb{E}}_{x_1,\dots,x_j \in W}[\chi_S(x)] = \mathop{\mathbb{E}}_{x_1 \in W}\left[\omega^{\mathrm{Tr}(x_1 \cdot s_1)}\right] \mathop{\mathbb{E}}_{x_2,\dots,x_j \in W}\left[\omega^{\sum_{i=2}^{j} \mathrm{Tr}(x_i \cdot s_i)}\right].$$

We will show that $\mathbb{E}_{x_1 \in W}\left[\omega^{\mathrm{Tr}(x_1 \cdot s_1)}\right] = 0$. Notice that it is sufficient to show that $x_1 \cdot s_1$ takes each value in $\mathbb{F}_q$ with equal probability over uniformly random $x_1 \in W$. First, since $s_1 \notin W^{\perp}$, $\Pr_{x_1 \in W}[x_1 \cdot s_1 = 0] = \frac{1}{q}$. Next note for any $\alpha \neq 0$,

$$\Pr_{x_1 \in W}[x_1 \cdot s_1 = 1] = \Pr_{x_1 \in W}[(\alpha x_1) \cdot s_1 = \alpha] = \Pr_{x_1 \in W}[x_1 \cdot s_1 = \alpha].$$

Therefore, $x_1 \cdot s_1$ takes each of the $q-1$ nonzero values in $\mathbb{F}_q$ with probability $\frac{1}{q}$ over uniform $x_1 \in W$, and this concludes the proof. $\square$

We can now show the necessary fact, asserting that if zooming out to W changes the measure of $\mathcal{A}$ considerably, then we can attribute it to a Fourier character S only containing elements from $W^{\perp}$.

Lemma E.5. *Suppose the parameters j, q, n, η satisfy* (59) *and let F be defined as above. If $W \subseteq \mathbb{F}_q^n$ has codimension r and satisfies*

$$|\mu_{\circ,W}(\mathcal{A}) - \eta| \geqslant 0.01\eta,$$

then there is a nonzero $S \in \mathbb{F}_q^{n\times j}$ such that $S \subseteq W^{\perp}$ and $\left|\widehat{F}(S)\right| \geqslant \frac{\eta}{20q^{r\cdot j}}$.

Proof. Note that, $\mu_{\circ,W}(\mathcal{A}) = \Pr_{x_1,\ldots,x_j\in\mathbb{F}_q^n}[\mathrm{span}(x_1,\ldots,x_j)\in\mathcal{A} \mid \dim(\mathrm{span}(x)) = j]$, so

$$\left|\mu_{\circ,W}(\mathcal{A}) - \mathop{\mathbb{E}}_{x\subseteq W}[F(x)]\right| \leqslant j\cdot q^{j-n}, \tag{60}$$

where the term on the right hand side bounds the probability that uniformly random $x_1,\ldots,x_j$ do not satisfy the conditioning $\dim(\mathrm{span}(x)) = j$. Using the Fourier decomposition of F we have

$$\mathop{\mathbb{E}}_{x\subseteq W}[F(x)] = \widehat{F}(0) + \sum_{0\neq S\subseteq W^{\perp}}\widehat{F}(S)\mathop{\mathbb{E}}_{x\subseteq W}[\chi_S(x)] + \sum_{0\neq S\subsetneq W^{\perp}}\widehat{F}(S)\mathop{\mathbb{E}}_{x\subseteq W}[\chi_S(x)].$$

Combining this with (60) and Lemma E.4 and using the fact that $\widehat{F}(0) = \eta$, we get

$$\left|\mu_{\circ,W}(\mathcal{A}) - \eta - \sum_{0\neq S\subseteq W^{\perp}}\widehat{F}(S)\right| \leqslant j\cdot q^{j-n}.$$

By the triangle inequality we conclude that

$$|\mu_{\circ,W}(\mathcal{A}) - \eta| \leqslant \left|\sum_{0\neq S\subseteq W^{\perp}}\widehat{F}(S)\right| + j\cdot q^{j-n},$$

and finally by the assumption in the lemma statement we have,

$$\left|\sum_{0\neq S\subseteq W^{\perp}}\widehat{F}(S)\right| \geqslant 0.1\cdot\eta - j\cdot q^{j-n}.$$

Since $j\cdot q^{j-n} \leqslant 0.01\cdot\eta$ by (59) and there are at most $q^{r\cdot j}$ tuples $S = (s_1,\ldots,s_j)\subseteq W^{\perp}$, the result follows. □

E.3.2 The Proof of Lemma 8.5

We are now ready to return to Lemma 8.5. Take $\mathcal{L}^\star, \mathcal{W}^\star, W^\star_{\mathsf{amb}}$ from Lemma 8.3, and recall that

$$Z := \{z\in W^\star_{\mathsf{amb}} \mid |\mu_{z,\circ}(\mathcal{L}^\star) - \eta| \leqslant \frac{\eta}{10}\},\ \eta := \mu(\mathcal{L}^\star),\ \text{and } m_3 := |\mathcal{W}^\star|.$$

Also recall that each $W_i^\star\in\mathcal{W}^\star$ has an associated linear function $f_i^\star : W_i^\star\to\mathbb{F}_q$ and $T_{\ell'}$ is a table assigning a linear function to each subspace in $\mathsf{Grass}(W^\star_{\mathsf{amb}},\ell')$, where $\ell'\geqslant\frac{\xi}{3}\ell$ is a dimension and $\xi>0$ should be thought of as a small constant.

We also restate the six items from Lemma 8.3 which the objects above satisfy

1. $\mu(\mathcal{L}^\star) := \eta \geqslant \frac{C}{12}$, where the measure here is over $\mathsf{Grass}_q(W^\star_{\mathsf{amb}},\ell')\supseteq\mathcal{L}^\star$.
2. The set $\mathcal{L}^\star$ is $(1, q^{\delta_2\ell}\eta)$-pseudo-random.

3. Each $W_i^\star$ has codimension $s \leqslant r$ inside of $W^\star_{\mathsf{amb}}$, $\mathcal{W}^\star$ is 4-generic, with respect to $W^\star_{\mathsf{amb}}$, and $W^\star_{\mathsf{amb}}$ has codimension at most $10/\delta_2$ with respect to W'_{amb}. As a consequence, $\dim(W^\star_{\mathsf{amb}})$ is much larger than ℓ'. Concretely:

$$\dim(W^\star_{\mathsf{amb}}) \geqslant \dim(W'_{\mathsf{amb}}) - 10/\delta_2 \geqslant n - 10/\delta - 10/\delta_2 \geqslant 2^{99} q^{\ell'},$$

   where here we use $n \geqslant 2^{100} q^{\ell}$.

4. The size of $\mathcal{W}^\star$ satisfies

$$\frac{m_2 \cdot q^{-10s/\delta_2}}{2} \leqslant m_3 \leqslant m_2.$$

5. For each $W_i^\star \in \mathcal{W}_i^\star$, there is a linear function $f_i^\star : W_i^\star \to \mathbb{F}_q$ such that the following holds. For every $L \in \mathcal{L}^\star$, choosing $W_i^\star \in \mathcal{W}^\star$ uniformly such that $W_i^\star \supseteq L$, we have

$$\Pr_{W_i^\star \supseteq L, W_i^\star \in \mathcal{W}^\star}[f_i^\star|_L \not\equiv T_{\ell'}[L]] \leqslant 14\gamma.$$

6. For every $L \in \mathcal{L}^\star$,

$$0.8 \cdot p_1 \cdot m_3 \leqslant N_{\mathcal{W}^\star}(L) \leqslant 1.2 \cdot p_1 \cdot m_3,$$

   where $N_{\mathcal{W}^\star}(L)$ is as defined in (13), and

$$p_1 := \Pr_{L \in \mathsf{Grass}_q(W^\star_{\mathsf{amb}}, \ell')}[L \subseteq W],$$

   for an arbitrary $W \subseteq W^\star_{\mathsf{amb}}$ of codimension s.

With this context in mind, the remainder of the section is devoted to the proof of Lemma 8.5, restated below for convenience.

Lemma 8.5. *Let $\mathcal{W}^\star = \{W_1^\star, \ldots, W_{m_3}^\star\}$ and $f_1^\star, \ldots, f_{m_3}^\star$ be the 4-generic set of codimension $s \leqslant r \leqslant 10/\delta$ subspaces inside of W'_{amb} and associated linear functions obtained from Lemma E.3 respectively. We have*

$$\Pr_{\substack{W_i^\star, W_j^\star \in \mathcal{W}^\star \\ z \in W_i^\star \cap W_j^\star \cap Z}}[f_i^\star(z) \neq f_j^\star(z)] \leqslant 500\gamma,$$

and for every $W_i^\star, W_j^\star \in \mathcal{W}^\star$ we have $|W_i^\star \cap W_j^\star \cap Z| \geqslant 0.81 \cdot |W_i^\star \cap W_j^\star|$. Here $Z \subseteq W^\star_{\mathsf{amb}}$ is as defined in Lemma 8.4.

Let us start by setting up some notation. For an arbitrary fixed point $z \in W^\star_{\mathsf{amb}}$, let D denote the number of ℓ'-dimensional subspaces $L \subseteq W^\star_{\mathsf{amb}}$ containing z. We note that D does not depend on which point z is fixed. Also let,

$$\begin{aligned}
&\mathcal{L}_z^\star = \{L \in \mathcal{L}^\star \mid z \in L\}, \\
&\mathcal{W}_z^\star = \{W_i^\star \in \mathcal{W}^\star \mid z \in W_i^\star\}, \\
&m_z = |\mathcal{W}_z^\star|, \\
&N_{2,\mathcal{W}_z^\star}(L) = |\{(i,j) \mid W_i^\star \cap W_j^\star \supseteq L,\ W_i^\star, W_j^\star \in \mathcal{W}_z^\star\}|, \\
&N_{2,\mathcal{W}^\star}(L) = |\{(i,j) \mid W_i^\star \cap W_j^\star \supseteq L, W_i^\star, W_j^\star \in \mathcal{W}^\star\}|.
\end{aligned} \tag{61}$$

We remark that in the last two equations, we do not require i and j to be distinct. Now, for an arbitrary pair of distinct $W_i^\star$ and $W_j^\star$ which both contain some point z, define

$$p_1 = \Pr_{L \in \mathsf{Grass}(W^\star_{\mathsf{amb}}, \ell')}[L \subseteq W_i^\star] \quad \text{and} \quad p_2 = \Pr_{L \in \mathsf{Grass}(W^\star_{\mathsf{amb}}, \ell')}[L \subseteq W_i^\star \cap W_j^\star \mid z \in L]. \tag{62}$$

Note that the above quantities depend neither on the specific identity of the distinct pair of subspaces $W_i^\star, W_j^\star$, nor on the identity of the point z. Also, we point out that p_1 here is the same as defined in Property 6 of Lemma 8.3. A straightforward computation shows that

$$\frac{p_2}{p_1^2} \geqslant \frac{q^{2s}}{2}, \tag{63}$$

where recall $s = \operatorname{codim}(W_i^\star)$ in $W^\star_{\mathsf{amb}}$. We start by removing all $z \in Z$ that do not satisfy

$$1.1 \cdot q^{-s} m_3 \geqslant m_z \geqslant 0.9 \cdot q^{-s} m_3. \tag{64}$$

We will abuse notation and still call the resulting set Z. By Lemma 5.21, the fraction of z removed is at most $\frac{2q^s}{m_3}$, so combined with Lemma 8.4 we still have

$$\mu(Z) \geqslant 1 - q^{-\ell'/2} - \frac{2q^s}{m_3}. \tag{65}$$

For the remainder of the section we have that all $z \in Z$ satisfy (64).

Consider the following two distributions over triples $(z, W_i^\star, W_j^\star) \in Z \times \mathcal{W}^\star \times \mathcal{W}^\star$. The first is $\mathcal{D}_1$, generated by choosing $z \in Z$ uniformly and $W_i^\star, W_j^\star \in \mathcal{W}^\star$ uniformly conditioned on $z \in W_i^\star \cap W_j^\star$. The second is $\mathcal{D}_1'$, generated by choosing $z \in Z$ uniformly, $L \in \mathcal{L}_z^\star$ uniformly, and then $W_i^\star, W_j^\star \in \mathcal{W}^\star$ uniformly conditioned on $L \subseteq W_i^\star \cap W_j^\star$. We have

$$\begin{aligned}
\mathcal{D}_1(z, W_i^\star, W_j^\star) &:= \frac{1}{|Z|} \cdot \frac{1}{|\{i', j' \mid W_{i'}^\star, W_{j'}^\star \in \mathcal{W}^\star, z \in W_{i'}^\star \cap W_{j'}^\star\}|} = \frac{1}{|Z| \cdot m_z^2}, \\
\mathcal{D}_1'(z, W_i^\star, W_j^\star) &= \frac{1}{|Z|} \cdot \frac{|\{L \in \mathcal{L}_z^\star \mid L \subseteq W_i^\star \cap W_j^\star\}|}{|\mathcal{L}_z^\star|} \cdot \mathop{\mathbb{E}}_{L \in \mathcal{L}_z^\star, L \subseteq W_i^\star \cap W_j^\star}\left[\frac{1}{N_{2,\mathcal{W}^\star}(L)}\right].
\end{aligned} \tag{66}$$

Again, we do not require i, j or i', j' to be distinct in any of the definitions above.

We will show that, for most $(z, W_i^\star, W_j^\star)$ and up to multiplicative constants, the distribution $\mathcal{D}_1'$ assigns at least as much as the distribution $\mathcal{D}_1$. For that, we use the fact that for all $z, W_i^\star, W_j^\star$:

$$\frac{1}{|Z|} \cdot \frac{1}{1.21 m_3^2 q^{-2s}} \leqslant \mathcal{D}_1(z, W_i^\star, W_j^\star) \leqslant \frac{1}{|Z|} \cdot \frac{1}{0.81 \cdot m_3^2 \cdot q^{-2s}}, \tag{67}$$

$$\mathcal{D}_1'(z, W_i^\star, W_j^\star) \geqslant \frac{1}{|Z|} \cdot \frac{\mu_{z, W_i^\star \cap W_j^\star}(\mathcal{L}^\star) \cdot D \cdot p_2}{1.1 \cdot \eta \cdot D} \cdot \frac{1}{1.44 \cdot p_1^2 \cdot m_3^2} \geqslant \frac{1}{|Z|} \cdot \frac{\mu_{z, W_i^\star \cap W_j^\star}(\mathcal{L}^\star)}{3.2 \cdot \eta \cdot m_3^2 \cdot q^{-2s}}. \tag{68}$$

To get (67) we plugged the bounds for m_z from (64) into the definition of $\mathcal{D}_1$ from (66). To get the first inequality in (68) we use the fact that, due to the definition of Z, we have $|\mathcal{L}_z^\star| \leqslant 1.1\eta \cdot D$ for all $z \in Z$. We also use the fact that $N_{2,\mathcal{W}^\star}(L) = N_{\mathcal{W}^\star}(L)^2$, along with the upper bound on $N_{\mathcal{W}^\star}(L)$ from the sixth property of Lemma 8.3. The second inequality in (68) follows from simplifying and using $p_2/p_1^2 \geqslant q^{2s}/2$ from (63).

Now, call a triplet $(z, W_i^\star, W_j^\star)$ bad if

$$\mu_{z,W_i^\star\cap W_j^\star}(\mathcal{L}^\star) \leqslant \frac{4}{5}\eta.$$

If the triplet $(z, W_i^\star, W_j^\star)$ is *not* bad, then by the above inequalities

$$\mathcal{D}_1(z, W_i^\star, W_j^\star) \leqslant \frac{1}{|Z|}\cdot\frac{1}{0.81\cdot m_3^2 q^{-2s}} \leqslant 6\cdot\frac{1}{|Z|}\cdot\frac{4\eta/5}{3.2\cdot\eta\cdot m_3^2\cdot q^{-2s}} \leqslant 6\cdot\mathcal{D}_1'(z, W_i^\star, W_j^\star). \tag{69}$$

We will now start by showing that there are very few bad triplets.

Claim E.6. *The following holds:*

1. *For each $z \in Z$, the number of i, j such that $\mu_{z,W_i^\star\cap W_j^\star}(\mathcal{L}^\star) \leqslant \frac{4}{5}\eta$ is at most $\frac{10^6 q^{4s\ell}}{\eta^2} m_3$.*
2. *for every $W_i^\star, W_j^\star \in \mathcal{W}^\star$, we have $\mu_{\circ,W_i^\star\cap W_j^\star}(Z) \geqslant 0.9\mu(Z)$.*

Proof. Fix a point $z \in Z$, let F_z be the restriction of F to the zoom-in of z, where $F(x_1, \ldots, x_{\ell'}) = 1$ if $\mathrm{span}(x_1, \ldots, x_{\ell'}) \in \mathcal{L}^\star$ and 0 otherwise. Let $\eta' = \mu_{z,\circ}(\mathcal{L}^\star)$. Since $z \in Z$, we have $\eta' \in [0.9\eta, 1.1\eta]$. Note that if i, j satisfy$\mu_{z,W_i^\star\cap W_j^\star}(\mathcal{L}^\star) \leqslant \frac{4}{5}\eta$, then (using $\eta' \in [0.9\eta, 1.1\eta]$)

$$\left|\mu_{z,W_i^\star\cap W_j^\star}(\mathcal{L}^\star) - \eta'\right| \geqslant \frac{\eta'}{20}.$$

By Lemma E.5 it follows that there is $S = (s_1, \ldots, s_{\ell'-1})$ such that $\mathrm{span}(s_1, \ldots, s_{\ell'-1}) \subseteq \left(W_i^\star \cap W_j^\star\right)^\perp$ and

$$|\widehat{F_z}(S)| \geqslant \frac{\eta'}{400 q^{2s(\ell'-1)}}. \tag{70}$$

Since by Parseval's inequality the sum of $\left|\widehat{F_z}(S)\right|^2$ is at most $\|F_z\|_2^2 \leqslant 1$, there are at most $\frac{160000 q^{4s\ell'}}{\eta'^2}$ tuples S satisfying (70). Now consider the bipartite graph whose left side consists of these tuples $S = (s_1, \ldots, s_{\ell'-1})$, right side consists of $W_i^\star \cap W_j^\star$, and the edges are between pairs that satisfy

$$\mathrm{span}(s_1, \ldots, s_{\ell'-1}) \subseteq \left(W_i^\star \cap W_j^\star\right)^\perp.$$

It follows that the number of edges in this graph is an upper bound on the number of bad triples containing z. Since $\mathcal{W}^\star$ is 4-generic, we have

$$(W_i^\star \cap W_j^\star)^\perp \cap (W_{i'}^\star \cap W_{j'}^\star)^\perp = \{0\}$$

for all i, j, i', j' distinct. Therefore, any two neighbours of a vertex on the left must either have their i or j be equal, and hence the maximum degree of a vertex on the left side is at most $2m_3$. As a result, the graph has at most $2m_3\cdot\frac{160000 q^{4s\ell}}{\eta'^2} \leqslant \frac{10^6 q^{4s\ell}}{\eta^2} m_3$ edges, where we also use that $\eta' \geqslant 0.9\eta$. This completes the proof of the first assertion of the claim.

For the second part of the lemma, note that $\mu(\overline{Z}) \leqslant \frac{q^{-\ell'}}{2} + \frac{2q^s}{m_3}$ by (65). Therefore, for any $W_i^\star \cap W_j^\star$, we have,

$$\mu_{\circ,W_i^\star\cap W_j^\star}(\overline{Z}) \leqslant q^{2s}\cdot\left(\frac{q^{-\ell'}}{2} + \frac{2q^s}{m_3}\right).$$

It follows that,

$$\mu_{\circ,W_i^\star\cap W_j^\star}(Z) \geqslant 1 - q^{2s}\cdot\left(\frac{q^{-\ell'}}{2} + \frac{2q^s}{m_3}\right) \geqslant 0.9\mu(Z). \qquad \square$$

Lemma E.7. *Let E be any event defined with respect to $(z, W_i^\star, W_j^\star)$. Then,*

$$\mathcal{D}_1(E) \leqslant 6\mathcal{D}_1'(E) + \gamma.$$

Proof. If the triple $(z, W_i^\star, W_j^\star)$ is not bad, then $\mathcal{D}_1(z, W_i^\star, W_j^\star) \leqslant 6\mathcal{D}_1'(z, W_i^\star, W_j^\star)$. Otherwise, we can use the generic bound $\mathcal{D}_1(z, W_i^\star, W_j^\star) \leqslant \frac{1}{|Z| \cdot 0.81 \cdot m_3^2 q^{-2s}}$ from (67). By the bound on the number of bad triples per z in Claim E.6, it follows that

$$\begin{aligned}
\mathcal{D}_1(E) &\leqslant 6\mathcal{D}_1'(E) + |Z| \cdot \frac{10^6 q^{4s\ell}}{\eta^2} m_3 \frac{1}{|Z| \cdot 0.81 \cdot m_3^2 q^{-2s}} \\
&= 6\mathcal{D}_1'(E) + \frac{10^7 q^{4s\ell+2s}}{\eta^2 \cdot m_3} \\
&\leqslant 6\mathcal{D}_1'(E) + \gamma.
\end{aligned}$$

Note that in the last transition we are using the fact that m_3 is large by the fourth property in Lemma 8.3. $\square$

Now let $\mathcal{D}_2$ be the distribution obtained by choosing $W_i^\star, W_j^\star \in \mathcal{W}$ uniformly, and then choosing $z \in W_i^\star \cap W_j^\star \cap Z$ uniformly. We have

$$\mathcal{D}_2(z, W_i^\star, W_j^\star) = \frac{1}{m_3^2} \cdot \frac{1}{|W_i^\star \cap W_j^\star \cap Z|}.$$

Using essentially the same proof, we get the following lemma.

Lemma E.8. *Let E be any event defined with respect to $(z, W_i^\star, W_j^\star)$. Then $\mathcal{D}_2(E) \leqslant 2\mathcal{D}_1(E)$.*

Proof. By Claim E.6, we have $\mu_{\circ, W_i^\star \cap W_j^\star}(Z) > 0.9\mu(Z)$ for all i, j, or equivalently

$$|W_i^\star \cap W_j^\star \cap Z| \geqslant 0.9 \cdot |Z| q^{-2s}.$$

Thus, for all i, j and all z,

$$\mathcal{D}_2(z, W_i^\star, W_j^\star) \leqslant \frac{1}{0.9|Z| \cdot q^{-2s}} \cdot \frac{1}{m_3^2} \leqslant \frac{2}{1.21|Z|} \cdot \frac{1}{m_3^2 q^{-2s}} \leqslant 2D_1(z, W_i^\star, W_j^\star),$$

where we use (67) for the third transition. $\square$

We are now ready to prove Lemma 8.5.

Proof of Lemma 8.5. By the fifth property in Lemma 8.3, for every $L \in \mathcal{L}^\star$, we have

$$\Pr_{W_i^\star \supseteq L, W_i^\star \in \mathcal{W}^\star} [f_i|_L \neq T_{\ell'}[L]] \leqslant 14\gamma.$$

Let E denote the event over $(z, W_i^\star, W_j^\star)$ that $f_i(z) \neq f_j(z)$. It follows that,

$$\mathcal{D}_1'(E) \leqslant \Pr_{\substack{W_i^\star, W_j^\star \supseteq L \\ W_i^\star, W_j^\star \in \mathcal{W}^\star}} [f_i|_L \neq T_{\ell'}[L] \ \vee \ f_j|_L \neq T_{\ell'}[L]] \leqslant 2 \cdot \Pr_{W_i^\star \supseteq L, W_i^\star \in \mathcal{W}^\star} [f_i|_L \neq T_{\ell'}[L]] \leqslant 28\gamma.$$

Putting Lemmas E.7 and E.8 together,

$$\mathcal{D}_2(E) \leqslant 2(6\mathcal{D}_1'(E) + \gamma) = 338\gamma,$$

proving the first part of Lemma 8.5.

For the second part, recall from the second part of Claim E.6 that for every pair i, j, we have

$$\left|W_i^\star \cap W_j^\star \cap Z\right| \geqslant 0.9 \cdot |Z|\, q^{-2s} \geqslant 0.81 \left|W_{\mathsf{amb}}^\star\right| q^{-2s} = 0.81 \cdot \left|W_i^\star \cap W_j^\star\right|. \qquad \square$$

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il