

# The Unique Games Theorem

## Abstract

We prove the Unique Games Conjecture. For every fixed  $\varepsilon, \delta \in (0, 1/2)$ , we give a deterministic polynomial-time reduction from 3SAT to explicit unweighted simple bipartite Unique Games over a fixed binary-vector alphabet, with completeness at least  $1 - \varepsilon$  and soundness at most  $\delta$ . All constraints are translations. The proof uses noise that rarely changes a nonlinear evaluation but is detected with constant probability by families of linear functionals whose restrictions to the label space have sufficiently large rank. Fourier comparison and Grassmann expansion yield agreeing local strategies from high test acceptance; parallel repetition rules out that agreement on unsatisfiable inputs.

## Contents

|          |                                                                    |           |
|----------|--------------------------------------------------------------------|-----------|
| <b>1</b> | <b>Introduction</b>                                                | <b>2</b>  |
| 1.1      | The theorem . . . . .                                              | 2         |
| 1.2      | Prior work and the completeness obstacle . . . . .                 | 3         |
| 1.3      | The construction and proof . . . . .                               | 3         |
| <b>2</b> | <b>Conventions and proof overview</b>                              | <b>5</b>  |
| 2.1      | Weighted games . . . . .                                           | 5         |
| 2.2      | Binary linear algebra and Fourier analysis . . . . .               | 5         |
| 2.3      | Probability and finite constructions . . . . .                     | 6         |
| 2.4      | Deduction of the final gap . . . . .                               | 6         |
| <b>3</b> | <b>Nonlinear stability and linear detection</b>                    | <b>6</b>  |
| 3.1      | Deduction from a fixed output space . . . . .                      | 7         |
| 3.2      | The quadratic block . . . . .                                      | 8         |
| 3.3      | Subspaces with few vanishing restrictions . . . . .                | 10        |
| 3.4      | Concatenation and its noise . . . . .                              | 11        |
| 3.5      | Linear lifts and loss of rank . . . . .                            | 13        |
| 3.6      | Realizing the fixed output space . . . . .                         | 15        |
| <b>4</b> | <b>The outer questions and the permutation test</b>                | <b>15</b> |
| 4.1      | Weighted parity equations and distinct positions . . . . .         | 16        |
| 4.2      | Tuples, local answer copies, and homogeneous coordinates . . . . . | 17        |
| 4.3      | Table keys and folding . . . . .                                   | 18        |
| 4.4      | The matrix test and completeness . . . . .                         | 19        |
| <b>5</b> | <b>Matrix extraction and Fourier decoding</b>                      | <b>20</b> |
| 5.1      | Agreement on a matrix slice . . . . .                              | 20        |
| 5.2      | The advice experiment and the decoding statement . . . . .         | 21        |
| 5.3      | The local strategies and their agreement . . . . .                 | 22        |
| 5.4      | Comparing the two matrix perturbations . . . . .                   | 23        |
| 5.5      | Obtaining many useful fibers . . . . .                             | 25        |

|          |                                                                  |           |
|----------|------------------------------------------------------------------|-----------|
| 5.6      | Transferring the chosen slice . . . . .                          | 27        |
| 5.7      | Decoding from agreement on a slice . . . . .                     | 29        |
| <b>6</b> | <b>Coordinates with zero advice and the soundness gap</b>        | <b>30</b> |
| 6.1      | Classical projection-game repetition . . . . .                   | 30        |
| 6.2      | Singleton coordinates with zero advice slope . . . . .           | 31        |
| 6.3      | The matrix-game gap . . . . .                                    | 31        |
| 6.4      | Proof of the zero-advice bound . . . . .                         | 32        |
| <b>7</b> | <b>Completion of the reduction</b>                               | <b>34</b> |
| 7.1      | Removing weights and repeated edges . . . . .                    | 34        |
| 7.2      | The final reduction . . . . .                                    | 34        |
| 7.3      | Proof of rounding and subdivision . . . . .                      | 36        |
| <b>8</b> | <b>Approximation consequences</b>                                | <b>37</b> |
| 8.1      | Constraint satisfaction and Vertex Cover . . . . .               | 37        |
| 8.2      | Ordering constraints . . . . .                                   | 39        |
| 8.3      | Cut, deletion, and clustering consequences . . . . .             | 40        |
| 8.4      | Kernel clustering . . . . .                                      | 40        |
| <b>A</b> | <b>Finite formulations of the consequences</b>                   | <b>41</b> |
| A.1      | Strong satisfaction of left vertices in a simple graph . . . . . | 41        |
| A.2      | The finite ordering transfer . . . . .                           | 44        |
| A.3      | Proofs of the local ordering estimates . . . . .                 | 50        |
| <b>B</b> | <b>Abundance of subspaces with few vanishing restrictions</b>    | <b>53</b> |
| <b>C</b> | <b>Agreement from Grassmann expansion</b>                        | <b>54</b> |

## 1 Introduction

A permutation constraint prescribes exactly one label at one endpoint for each label at the other. The Unique Games Conjecture asserts that systems of such constraints can be hard to approximate even when almost all constraints can be satisfied. We resolve this conjecture positively, with an explicit unweighted reduction from 3SAT.

### 1.1 The theorem

A Unique Games instance has a finite vertex set, a finite alphabet  $K$ , and a nonempty list of oriented constraints  $e = (u_e, v_e, \pi_e)$ , where  $\pi_e$  is a permutation of  $K$ . A labeling  $a$  satisfies  $e$  when  $a(v_e) = \pi_e(a(u_e))$ . Its value is the fraction of satisfied constraints, and  $\text{val}(G)$  is the maximum value of a labeling. The instance is explicit and unweighted when the entire vertex and constraint lists are written out, including the table of each permutation. A simple bipartite instance has two disjoint vertex sets and at most one constraint on each pair of opposite vertices.

**Theorem 1.1 (Unique Games Theorem).** *For every fixed  $\varepsilon, \delta \in (0, 1/2)$ , a deterministic polynomial-time reduction maps each 3SAT formula  $\varphi$  to an explicit unweighted simple bipartite Unique Games instance  $G_\varphi$  over  $K = \mathbb{F}_2^s$ , for an integer  $s \geq 1$  depending only on  $\varepsilon, \delta$ , with*

$$\begin{aligned} \varphi \text{ satisfiable} &\implies \text{val}(G_\varphi) \geq 1 - \varepsilon, \\ \varphi \text{ unsatisfiable} &\implies \text{val}(G_\varphi) \leq \delta. \end{aligned}$$

Every constraint has the form  $a(v_e) = a(u_e) + c_e$  for a specified  $c_e \in K$ . The degree and constants of the running-time polynomial depend only on  $\varepsilon, \delta$ .

Thus the result resolves the Unique Games Conjecture positively in its usual edge-value formulation [12, Conjecture 2.5]. The two errors are prescribed independently, before the alphabet is chosen. Binary translation constraints are a known equivalent formulation of the conjecture [13, Corollary 13 and Section 11.1]; the reduction here produces that form directly.

Theorem 1.1 supplies the Unique Games hypothesis in earlier hardness reductions for cut, covering, constraint-satisfaction, ordering, deletion, and clustering problems [13, 16, 22, 8, 3, 4]. Section 8 gives the precise problem formulations, strict thresholds, and reduction qualifications, together with proofs of the consequences.

## 1.2 Prior work and the completeness obstacle

The closest preceding hardness results come from the Grassmann-graph approach. Khot, Minzer, and Safra gave a reduction whose soundness is conditional on a Grassmann-graph hypothesis [15, Definition 1.5 and Theorem 1.8]. In its NO case, one cannot assign lists of a fixed number of labels to more than a prescribed positive fraction of vertices so that every induced 2-to-2 edge has a satisfying pair and every induced 1-to-1 edge matches the two lists exactly.

Dinur, Khot, Kindler, Minzer, and Safra obtained ordinary edge-satisfaction soundness under a Grassmann agreement hypothesis [5]. Barak, Kothari, and Steurer related that hypothesis to expansion through the degree-two shortcode graph [1], and Khot, Minzer, and Safra proved the required expansion theorem [14]. Together these results establish hardness for almost satisfiable 2-to-1 and 2-to-2 games with arbitrarily small soundness. Here a 2-to-1 constraint has two preimages per right label; a 2-to-2 constraint matches pairs of labels and allows all four combinations within each matched pair.

Their Unique Games consequence has completeness one half [5, Corollary 1.12]. Reaching completeness arbitrarily close to one requires a further argument.<sup>1</sup>

The degree-two shortcode formulation makes the loss of completeness explicit. Its vertices are binary  $\ell$ -by- $m$  matrices, and a step adds  $al^\top$  for independent uniform  $a \in K = \mathbb{F}_2^\ell$  and  $l \in \mathbb{F}_2^m$ . For a nonzero answer  $z \in \mathbb{F}_2^m$ , the evaluation  $f_z(M) = Mz$  obeys

$$f_z(M + al^\top) - f_z(M) = a(l^\top z).$$

Allowing either  $f_z(M)$  or  $f_z(M) + a$  as the reply gives perfect completeness. Requiring equality gives acceptance  $(1 + 2^{-\ell})/2$ , close to one half. Barak, Kothari, and Steurer proposed higher-degree tests with better completeness, while identifying their composition with an outer PCP as an unresolved obstacle [1, Section 1.6].

For a function on matrices, equality acceptance averages the probability of staying within each output fiber, weighted by that fiber's mass. Fixed positive acceptance therefore gives a fiber with positive retention. Combining this observation with Grassmann expansion yields the conclusion of Barak, Kothari, and Steurer [1, Hypothesis 2.1 and Lemma 2.3]: on a positive fraction of a nonempty slice specified by boundedly many row and column equations, the function agrees with an affine evaluation. Theorem 5.1 states this conclusion for arbitrary functions under the uniform rank-one step. We must compare our noise with that step before applying it.

## 1.3 The construction and proof

Lemma 3.1 constructs a binary space  $\mathcal{V} \supseteq K$ , a nonlinear map  $C : \mathcal{V} \rightarrow K$ , and a noise law  $\mu$ . For uniform  $x$  and independent  $a \sim \mu$ , the probability that  $C(x + a) \neq C(x)$  can be made arbitrarily small.

---

<sup>1</sup>A separate companion proves hardness for exact 2-to-1 games with perfect completeness and arbitrarily small fixed soundness, over fixed alphabets [20, Theorem 1.1]. It is not an input to this proof.

Yet any family of linear functions whose restrictions to  $K$  have sufficiently large rank contains a function nonzero at  $a$  with probability at least  $1/8$ . The identity  $C(x + h) = C(x) + h$  for  $h \in K$  makes  $C$  usable as a nearly satisfying labeling of the translation test.

We build  $C$  recursively from quadratic blocks and perturb one random leaf. The nonlinear change probability decays exponentially with depth. Along the same branch, we track  $H_r = \sum_{j=1}^r 1/j$ , where  $r$  is the surviving restriction rank. A bound on its expected decrease gives a positive probability that a nonzero restriction reaches the noisy leaf. The required rank threshold is fixed before the alphabet dimension.

The structural theorem of Barak–Kothari–Steurer and Khot–Minzer–Safra applies to uniform rank-one noise. To use it for our constructed noise, Lemma 5.4 compares acceptance through the Fourier expansions of the characters of the table labels. A frequency consists of one linear functional per matrix column. Averaging the selector  $l$  makes its multiplier the probability that all these functionals vanish at the sampled noise vector. The detection bound keeps this multiplier away from one at high restriction rank, so acceptance near one forces positive squared Fourier mass at lower ranks. Under uniform noise in  $K$ , a frequency of restriction rank  $r$  has multiplier  $2^{-r}$  [14, Lemma 2.10]; the same mass therefore gives positive acceptance.

Split  $\mathcal{V} = K \oplus W$  and each table into its  $K$ - and  $W$ -valued components. A uniform- $K$  perturbation leaves the  $W$ -valued component fixed. A positive fraction of the resulting restrictions satisfy Theorem 5.1’s acceptance hypothesis and hence contain the required slices.

Starting from Håstad’s parity gap [9, Theorem 5.4 and its proof], we apply the test to tables of valid answers to tuples of parity equations. A second question is obtained by replacing a small random fraction of the equations by individual variables. We identify a table on this question with its pullback to the equation tuple, using the shared-entry construction and sparse comparison of Khot–Safra [17, Section 1.1 and Lemma 3.1]. Adding a constant to a table prescribes the same change to its answer, as in folding [2, Section 3.3]; the two provers receive compatible linear observations of the hidden table, as in the advice construction of [15, Sections 3.3–3.4].

As in earlier Grassmann reductions [5, Sections 1.4.1–1.4.2], a slice selected from one prover’s observations is unavailable to the other. The equation prover chooses a slice and a short set of valid target vectors from its own observations, then samples one of those answers. The projected prover chooses a uniform binary linear functional  $\sigma$  on the label space and applies  $y \mapsto (-1)^{\sigma(y)}$  to the table labels. On the matrices compatible with its observed row equations, it samples a frequency according to the squared Fourier coefficients of this sign function.

That frequency is a linear map from the kernel of the observed row map to the projected answer space. When  $\sigma$  is nonzero on the kernel, the prover evaluates the sampled map at a kernel vector  $h$  with  $\sigma(h) = 1$ . Agreement on the transferred slice gives a positive probability that this returned vector is the projection of one of the equation prover’s candidate answers. Their independent samples therefore agree with positive probability, without sharing the selected slice (Section 5).

On a NO parity instance, we use variable positions whose coefficients vanish in every visible advice map. After the remaining data are fixed, the ordinary equation–variable question law is unchanged on these positions. Projection-game repetition then forces the agreement probability to tend to zero (Section 6).

Dinur–Steurer’s projection-game repetition bound [6, Corollary 1.2] is independent of the answer alphabet, so the final repetition exponent can be chosen before the nonlinear construction fixes that alphabet. Figure 1 shows how the two soundness bounds enter the reduction.

Section 2 fixes the game and probability conventions and gives the deduction of the final gap. Section 3 constructs the gadget, and Section 4 defines the permutation test. Sections 5 and 6 give the two incompatible bounds on the advice experiment. Section 7 completes the parameter choices and the explicit Unique Games reduction. Section 8 then derives the approximation consequences from the completed theorem and the cited reductions. Appendix A gives the formulation with almost all left vertices strongly satisfied and the finite ordering reduction used there. Appendices B and C contain the abundance and

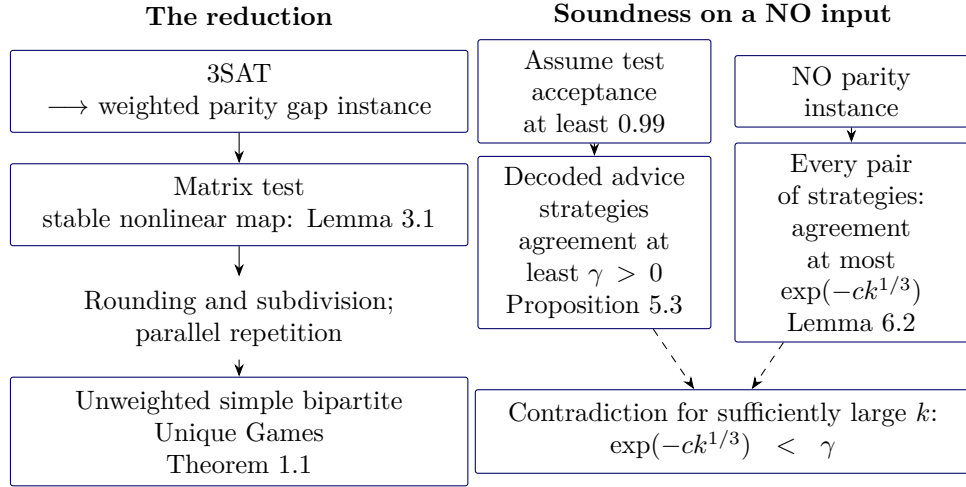


Figure 1: The solid arrows describe the reduction; the dashed arrows describe its soundness argument. The decoding lower bound follows from high test acceptance, while the upper bound uses variable positions whose coefficients are zero in all visible advice maps and follows from the NO parity gap and applies to every advice strategy. The constants  $c, \gamma > 0$  are fixed before the tuple length  $k$  grows.

Grassmann verifications used in Sections 3 and 5.

## 2 Conventions and proof overview

### 2.1 Weighted games

Intermediate instances may have loops, repeated constraint occurrences, and positive rational weights summing to one. Their value is the maximum total weight satisfied by a labeling. Occurrences retain separate IDs even if their constraints coincide. All rational numbers are encoded by binary numerators and denominators.

A bipartite projection game has a question at each vertex, finite answer sets, and a map from left answers to right answers on each constraint. Each prover's strategy assigns an answer to each of its own questions. Question-dependent answer sets of equal size can be identified with a fixed alphabet; we use this for the four satisfying triples of a parity equation. Randomness independent of the questions cannot increase value: sampling the provers' complete answer tables before the questions arrive expresses their success probability as an average over deterministic strategies. This applies to both private and shared randomness.

The parallel power  $G^{\otimes t}$  samples  $t$  independent constraint occurrences, sends the ordered question tuples, and accepts if every coordinate accepts. Answers may depend on the prover's entire tuple. Repeated question IDs retain separate coordinate positions and separate local answer copies. This is the ordinary classical product game used by the repetition theorem.

### 2.2 Binary linear algebra and Fourier analysis

Except where a larger field is specified, all vector spaces and maps are finite-dimensional over  $\mathbb{F}_2$ . Write  $\text{Hom}(E, F)$  for the space of linear maps and  $E^* = \text{Hom}(E, \mathbb{F}_2)$  for the dual. A subspace  $S \leq E^*$  restricts to its image  $S|_K \leq K^*$  on a subspace  $K \leq E$ ; its restriction rank is  $\dim(S|_K)$ .

For a real function  $f$  on a binary space  $E$ , use normalized Fourier coefficients

$$\hat{f}(g) = \mathbb{E}_{x \in E} f(x) (-1)^{g(x)}, \quad f(x) = \sum_{g \in E^*} \hat{f}(g) (-1)^{g(x)}, \quad \sum_g |\hat{f}(g)|^2 = \mathbb{E}_x |f(x)|^2.$$

These identities follow because a nonzero linear functional has equally many zero and one values, so its character has mean zero. The characters are therefore an orthonormal basis. On a space of matrices, this is the entrywise pairing with its dual matrix space; we specify tensor notation when defining the Fourier decoder.

## 2.3 Probability and finite constructions

Uniform sampling from a vector space includes zero. Samples are independent unless a dependence is stated. For finite probability laws,

$$\text{TV}(P, Q) = \frac{1}{2} \sum_x |P(x) - Q(x)|.$$

It bounds the difference of expectations of any  $[0, 1]$ -valued function. We condition only on events of positive probability.

Random experiments describe finite weighted constraint lists. The reduction enumerates those lists deterministically. Every auxiliary dimension and tuple length is fixed before the input formula varies; the final proof checks both this parameter order and the binary encoding cost. Fourier witnesses and decoders are used only to bound game value and need not be efficient.

## 2.4 Deduction of the final gap

The proof of Theorem 1.1 combines the matrix-game gap of Proposition 6.3, the rounding and subdivision bound of Lemma 7.1, and the repetition bound of Theorem 6.1.

Choose the repetition exponent  $t$  for the soundness gap  $1/800$  and error  $\delta$  before choosing the alphabet. Next choose  $p_*, \tau$ , then the gadget and advice parameters, then a sufficiently large cube  $k$  for Proposition 6.3, and finally the parity error  $\xi < 1/100$ .

Theorem 4.1 and Lemma 4.2 supply the parity gap. Proposition 6.3 produces a translation game  $H$  with YES value at least  $1 - d$ , where  $d = k\xi + p_*/2$ , and NO value at most  $99/100$ . For  $\tau \leq 1/200$ , Lemma 7.1 gives a simple bipartite unweighted game  $B$  on the same alphabet, with YES value at least  $1 - (d + \tau)/4$  and NO value at most  $1 - 1/800$ . For  $G = B^{\otimes t}$ , a coordinatewise YES labeling and a union bound give the first inequality below; Theorem 6.1 gives the second:

$$\begin{aligned} \text{YES : } \quad \text{val}(G) &\geq 1 - t(d + \tau)/4 \geq 1 - \varepsilon, \\ \text{NO : } \quad \text{val}(G) &\leq (1 - 1/10240000)^t \leq \delta. \end{aligned}$$

Section 7.2 gives rational error budgets, in the stated order, that ensure the final inequalities.

Products preserve simplicity and uniform edge weights here because  $B$  is simple. Products of translations on  $\mathbb{F}_2^\ell$  are translations on  $\mathbb{F}_2^{\ell t}$ . All parameters are fixed before the input formula varies. The full proof in Section 7 makes each parameter choice effective and verifies the polynomial encoding of every constraint. Thus Sections 3–6 need only establish the stated matrix-game gap with this parameter order; their two soundness bounds are assembled in Proposition 6.3.

## 3 Nonlinear stability and linear detection

We construct a nonlinear map that is stable under additive noise while sufficiently high-rank families of linear functions detect the same noise. The following lemma states the two probability bounds and their common equivariance requirement.

**Lemma 3.1 (Nonlinear stability and linear detection).** *For every  $p_* > 0$ , there is an integer  $r_*$  such that every sufficiently large integer  $\ell$  admits the following data: a finite-dimensional binary space  $\mathcal{V}$*

with a specified subspace  $K = \mathbb{F}_2^\ell$ , a map  $C : \mathcal{V} \rightarrow K$ , and a rational probability law  $\mu$  on  $\mathcal{V}$ . The map is equivariant under translation by this subspace:

$$C(x + k) = C(x) + k \quad (x \in \mathcal{V}, k \in K). \quad (3.1)$$

For independent uniform  $x \in \mathcal{V}$  and  $a \sim \mu$ ,

$$\Pr[C(x + a) \neq C(x)] \leq p_*. \quad (3.2)$$

Every linear subspace  $G \leq \mathcal{V}^*$  whose image under restriction to the specified  $K$  has dimension at least  $r_*$  satisfies

$$\Pr_{a \sim \mu} [\text{there exists } g \in G \text{ with } g(a) \neq 0] \geq \frac{1}{8}. \quad (3.3)$$

For rational  $p_*$ , a terminating deterministic algorithm finds  $r_*$  and constructs these finite data for every sufficiently large  $\ell$ .

### 3.1 Deduction from a fixed output space

We will construct a finite binary space  $V_0$  containing a subspace  $B$  of dimension  $b \geq 1$ , a map  $\overline{C} : V_0 \rightarrow B$  satisfying  $\overline{C}(x + b') = \overline{C}(x) + b'$  for every  $x \in V_0$  and  $b' \in B$ , and a rational noise law  $\nu$ , with

$$\Pr_{x \text{ uniform}, a \sim \nu} [\overline{C}(x + a) \neq \overline{C}(x)] \leq p_*.$$

The same noise will satisfy

$$G_0|_B = B^* \implies \Pr_{a \sim \nu} [\exists g \in G_0 : g(a) \neq 0] \geq \frac{1}{4} \quad (G_0 \leq V_0^*).$$

Here  $x$  and  $a$  are independent, and  $b$  and all the finite data depend only on  $p_*$ . For rational  $p_*$  they are obtained by a terminating deterministic construction. Sections 3.2–3.6 establish this output. We first give the deduction that allows the eventual alphabet dimension  $\ell$  to be chosen later.

*Proof of Lemma 3.1 from the fixed-space construction.* Set  $r_* = b + 2$ . A product indexed by all injections  $B \rightarrow K$  converts detection of a family filling  $B^*$  into detection of every sufficiently high-rank family on  $K$ .

Fix  $\ell \geq b + 2$  and write  $K = \mathbb{F}_2^\ell$ . Let  $\mathcal{L}$  be the finite set of all binary linear injections  $L : B \rightarrow K$ . Take a product  $P = \prod_{L \in \mathcal{L}} V_0$  and define

$$\tilde{C}((x_L)_L) = \sum_{L \in \mathcal{L}} L \overline{C}(x_L).$$

The product of shift subspaces  $T = \prod_L B \leq P$  acts on this output through the linear map

$$\Lambda : T \longrightarrow K, \quad \Lambda((b_L)_L) = \sum_L L b_L.$$

This map is onto, because every nonzero vector of  $K$  is in the image of some injection  $B \rightarrow K$ . Quotient by its kernel inside  $T$ :

$$\mathcal{V} = P / \ker \Lambda.$$

For  $h \in T$ , equivariance in each copy gives

$$\tilde{C}(x + h) = \tilde{C}(x) + \Lambda h.$$

In particular,  $\tilde{C}$  is constant on every coset of  $\ker \Lambda$ , so it descends to  $C : \mathcal{V} \rightarrow K$ . The injective map  $T / \ker \Lambda \rightarrow P / \ker \Lambda$  identifies its image with  $K$  through  $\Lambda$ . If  $q : P \rightarrow \mathcal{V}$  is the quotient map and  $\Lambda h = k$ ,

the same identity reads  $C(qx + k) = C(qx) + k$ , independently of the chosen representative  $h$ . This proves Equation (3.1).

On  $P$ , choose a uniform  $L \in \mathcal{L}$  and update only that copy by an independent sample from  $\nu$ . Let  $\mu$  be the resulting law after the quotient. On uniform product input the output sum changes exactly when the selected copy's output changes, because  $L$  is injective. Uniform input pushes to uniform input on the quotient, so the change probability is still at most  $p_*$ . This proves Equation (3.2). Every noise distribution used here is obtained from uniform finite choices by pushforward, and therefore is rational.

It remains to establish detection for an arbitrary  $G \leq \mathcal{V}^*$ . Let  $S = G|_K \leq K^*$ , of dimension  $r \geq r_*$ . Pull  $G$  back to  $P$  and then restrict to its copy of  $V_0$  indexed by  $L$ ; denote the resulting subspace of  $V_0^*$  by  $G_L$ . Its restriction to that copy's shift subspace  $B$  is exactly

$$G_L|_B = L^*S.$$

To see this, a shift  $b' \in B$  supported on copy  $L$  represents the shift  $Lb' \in K$  in the quotient, so a linear functional restricting to  $z \in S$  evaluates there as  $z(Lb')$ .

Write  $S^\perp = \{v \in K : z(v) = 0 \text{ for all } z \in S\}$ . The map  $L^* : S \rightarrow B^*$  fails to be onto if and only if there is a nonzero  $b' \in B$  with  $Lb' \in S^\perp$ . For each such fixed  $b'$ , a uniform injection  $L$  makes  $Lb'$  uniform among the nonzero vectors of  $K$ . Hence

$$\Pr_L[Lb' \in S^\perp] = \frac{2^{\ell-r} - 1}{2^\ell - 1} \leq 2^{-r}.$$

A union bound gives

$$\Pr_L[L^*S \neq B^*] \leq (2^b - 1)2^{-r} \leq 2^{b-r} \leq \frac{1}{4}.$$

On each of the remaining copies, the fixed-space detection bound gives detection probability at least  $1/4$  by  $G_L$ , and thus by  $G$  on the embedded noise. Averaging over the selected copy yields detection probability at least  $\frac{3}{4} \cdot \frac{1}{4} = \frac{3}{16} \geq \frac{1}{8}$ , proving Equation (3.3). The construction works for every  $\ell \geq b + 2$ , with  $r_* = b + 2$ , as required.

The finite product, quotient, and pushforward operations preserve effectivity once the fixed-space data have been constructed.  $\square$

It remains to construct the stated fixed-space data. We concatenate a quadratic block, tracking both its nonlinear change probability and the rank of linear restrictions along the same random branch. Lemma 3.6 and Proposition 3.8 give the two quantitative bounds. Section 3.6 chooses a common height where stability and detection both hold, and passes from the shift action to the subspace  $B$ .

### 3.2 The quadratic block

Fix an integer  $r_0 \geq 1$ , whose size will eventually depend only on  $p_*$ . Write  $H_0 = 0$  and  $H_r = \sum_{j=1}^r 1/j$ . Let  $F = \mathbb{F}_{2^d}$ ,  $q = 2^d$ , and regard  $X = B = F^3$  as binary spaces, each of dimension  $b = 3d$ . The integer  $d$  will be chosen sufficiently large in terms of  $r_0$ .

Use the trace pairing to identify both  $X^*$  and  $B^*$  with  $F^3$ : the vector  $z \in F^3$  represents the linear functional  $y \mapsto \text{Tr}_{F/\mathbb{F}_2}(z \cdot y)$ , where the dot denotes the field dot product. We abbreviate the trace by  $\text{Tr}$ . This is an identification of binary vector spaces; the corresponding real character is  $y \mapsto (-1)^{\text{Tr}(z \cdot y)}$ . The trace pairing is nondegenerate, and  $\text{Tr}(t^2) = \text{Tr}(t)$ . Every element of  $F$  has a unique square root.

Define  $Q : X \rightarrow B$  and  $D : X \times X \rightarrow B$  by

$$Q(x_1, x_2, x_3) = (x_2x_3, x_1x_3, x_1x_2), \quad D(x, a) = Q(x + a) + Q(x) + Q(a). \quad (3.4)$$

The block observable is

$$C_{\text{blk}} : X \oplus B \longrightarrow B, \quad C_{\text{blk}}(x, y) = y + Q(x).$$

It is equivariant under the vertical copy of  $B$ :  $C_{\text{blk}}(x, y + k) = C_{\text{blk}}(x, y) + k$ . For  $z \in B^*$ , the scalar observation  $z \circ C_{\text{blk}}$  changes by  $z(k)$  under this shift. We shall compare it with linear functionals whose restriction to the vertical  $B$  is also  $z$ .

For a one-dimensional  $F$ -subspace  $A = Fv$  of  $F^3$ , put

$$D_A = \{y \in B : v \cdot y = 0\}, \quad U_A = \{(a, Q(a) + w) : a \in A, w \in D_A\} \subseteq X \oplus B. \quad (3.5)$$

These definitions are independent of the chosen nonzero generator  $v$ . There are  $q^2 + q + 1$  such lines; set

$$\theta = \frac{1}{q^2 + q + 1}. \quad (3.6)$$

**Lemma 3.2 (Geometry and response of a block).** *For every one-dimensional  $F$ -subspace  $A \leq F^3$ , the set  $U_A$  in Equation (3.5) is a binary subspace of dimension  $b$ . The subspace  $D_A$  in Equation (3.5) satisfies*

$$D(x, a) \in D_A \quad (x \in X, a \in A),$$

*and the annihilator of  $D_A$  under the trace pairing is exactly  $A \leq B^*$ . The family  $(U_A)_A$  spans  $X \oplus B$ .*

*For this same  $A$ , the block map has the exact change probability*

$$\Pr[C_{\text{blk}}((x, y) + \zeta) \neq C_{\text{blk}}(x, y)] = 1 - \theta,$$

*when  $(x, y)$  is uniform in  $X \oplus B$  and  $\zeta$  is independently uniform in  $U_A \setminus \{0\}$ .*

*Proof.* The three coordinates of  $D(x, a)$  are

$$x_2a_3 + a_2x_3, \quad x_1a_3 + a_1x_3, \quad x_1a_2 + a_1x_2.$$

When  $a = tv$ , every term in  $v \cdot D(x, a)$  occurs twice, so their sum is zero. When both arguments are multiples of  $v$ , each coordinate of the polar expression is zero. Thus  $Q|_A$  is binary additive, and the parametrization of  $U_A$  by  $A \oplus D_A$  is a binary linear bijection. Its dimension is  $d + 2d = b$ .

Every  $z \in A$  annihilates  $D_A$  under the trace pairing. The annihilator has binary dimension  $3d - 2d = d$ , so it is exactly  $A$ . Finally, for each coordinate line  $A = Fe_i$ , the space  $U_A$  contains  $(Fe_i, 0)$ , because  $Q(te_i) = 0$ , as well as  $\{0\} \oplus D_A$ . These three coordinate lines and their vertical hyperplanes span  $X \oplus B$ .

The form of  $U_A$  makes the quadratic term of an update cancel. For  $(a, Q(a) + w) \in U_A$ ,

$$C_{\text{blk}}((x, y) + (a, Q(a) + w)) + C_{\text{blk}}(x, y) = w + D(x, a).$$

The right-hand side lies in  $D_A$  by the identity proved above. First take  $\zeta$  uniform in all of  $U_A$ , still independently of the uniform aggregate. Its parametrization  $(a, Q(a) + w)$  has independent uniform  $a \in A$  and  $w \in D_A$ . For every fixed  $x, a$ , the output difference  $w + D(x, a)$  is uniform in  $D_A$ , so it is nonzero with probability  $1 - q^{-2}$ . The zero perturbation has mass  $q^{-3}$  and never changes the output. Conditioning on a nonzero perturbation therefore gives

$$\frac{1 - q^{-2}}{1 - q^{-3}} = \frac{q(q + 1)}{q^2 + q + 1} = 1 - \theta.$$

□

The response calculation concerns an independent uniform aggregate and uniform nonzero update. When the update comes from a lower level of the recursion, we will verify this conditional law before applying it.

To analyze linear observations of a block, use  $(X \oplus B)^* = X^* \oplus B^*$ . An extension of  $z \in B^*$  is uniquely of the form  $(\gamma_z, z)$ , so a linear choice of extensions on a subspace  $S$  is specified by a linear map  $\gamma : S \rightarrow X^*$ . The kernel of restriction to  $U_A$  determines whether this family loses rank.

**Lemma 3.3 (Kernel of a block restriction).** *Let  $S \leq B^*$  be a binary subspace and let  $\gamma : S \rightarrow X^*$  be binary linear. Write  $g : S \rightarrow F^3$  for its trace representative, so that  $\gamma_z(x) = \text{Tr}(g(z) \cdot x)$ . For a one-dimensional  $F$ -subspace  $A \leq F^3$ , consider the restriction map*

$$\kappa_{A,\gamma} : S \longrightarrow U_A^*, \quad \kappa_{A,\gamma}(z) = (\gamma_z, z)|_{U_A}.$$

*Its nonzero kernel elements are precisely the  $z \in S \setminus \{0\}$  satisfying*

$$A = Fz \quad \text{and} \quad z \cdot g(z) = \sqrt{z_1 z_2 z_3}.$$

*Proof.* If  $z$  lies in the kernel, it annihilates every vertical vector  $(0, w)$  with  $w \in D_A$ . Lemma 3.2 therefore gives  $z \in A$ , and for  $z \neq 0$  this means  $A = Fz$ . On the remaining vectors  $(tz, Q(tz))$ ,  $t \in F$ , the functional has value

$$\text{Tr}(t g(z) \cdot z + t^2 z_1 z_2 z_3). \quad (3.7)$$

The three summands of  $z \cdot Q(tz)$  are equal, and three equals one in characteristic two. Also, for  $t, c \in F$ ,

$$\text{Tr}(t^2 c) = \text{Tr}(t \sqrt{c}).$$

Nondegeneracy of the trace pairing shows that Equation (3.7) vanishes for every  $t$  exactly when  $z \cdot g(z) = \sqrt{z_1 z_2 z_3}$ . Together with  $z \in A$ , this condition is sufficient as well: the vertical part is already annihilated.  $\square$

### 3.3 Subspaces with few vanishing restrictions

The kernel criterion suggests two sufficient controls. We shall use a family in which no field line contains two nonzero linear functionals and only a few linear functionals satisfy the alignment equation for any one choice of extensions. The uniformity over that choice will later allow a lift to depend on earlier random changes of coordinates.

**Definition 3.4.** *For each integer  $r \geq 1$ , let  $\mathfrak{S}_r(F)$  be the family of binary subspaces  $S \leq F^3$  of dimension  $r$  such that no two distinct nonzero elements of  $S$  are proportional over  $F$ , and, for every binary linear map  $g : S \rightarrow F^3$ , at most  $3r$  nonzero  $z \in S$  satisfy*

$$z \cdot g(z) = \sqrt{z_1 z_2 z_3}. \quad (3.8)$$

Fix one map  $\gamma : S \rightarrow X^*$  before choosing a uniform line  $A$ , and let  $r_A = \dim \text{im } \kappa_{A,\gamma}$ , where  $r = \dim S > 0$ . Lemma 3.3 now bounds the loss. If  $S \in \mathfrak{S}_r(F)$ , each field line contains at most one nonzero element of  $S$ , so the kernel has dimension at most one. For the fixed map  $g$  associated with  $\gamma$ , at most  $3r$  linear functionals satisfy Equation (3.8); hence at most  $3r$  lines can cause any rank loss. Consequently

$$r - r_A \leq 1, \quad \Pr_A[r_A < r] \leq 3r\theta, \quad \mathbb{E}_A[H_r - H_{r_A}] \leq 3\theta. \quad (3.9)$$

The last bound uses  $H_r - H_{r-1} = 1/r$ . For an arbitrary rank- $r$  space, a loss still requires  $A = Fz$  for some nonzero  $z \in S$ . Therefore

$$\Pr_A[r_A < r] \leq (2^r - 1)\theta \leq 2^r \theta, \quad H_r - H_{r_A} \leq H_r. \quad (3.10)$$

These estimates concern rank under restriction. Detection of the eventual noise will follow only after a positive-rank family reaches a leaf.

It remains to show that spaces satisfying Definition 3.4 are abundant.

**Lemma 3.5.** *For every fixed integer  $r \geq 1$ , if  $S$  is uniform among the binary  $r$ -dimensional subspaces of  $\mathbb{F}_{2^d}^3$ , then*

$$\Pr[S \in \mathfrak{S}_r(\mathbb{F}_{2^d})] \longrightarrow 1 \quad (d \rightarrow \infty).$$

The proof is in Appendix B. The values of  $g$  on a binary basis give  $3r$  field unknowns. Requiring Equation (3.8) at more than  $3r$  distinct nonzero binary combinations of that basis imposes a linear system on these unknowns. The appendix proves that this system is inconsistent when the basis entries are independent indeterminates. After squaring the entries, a nonzero minor of the augmented matrix becomes a polynomial that must vanish whenever the specialized finite-field system has a solution. Its bounded degree makes this event unlikely as the field grows. A union bound over the finitely many choices of binary combinations gives the required uniformity over every  $g$ .

The two conditions therefore control every possible map  $g$ . We now fix a field where they hold with high probability at every rank up to  $r_0$ ; no linear maps need to be chosen at this stage. Choose a positive  $\varepsilon_0$  with

$$H_{r_0} 2^{r_0} \varepsilon_0 \leq 1, \quad \varepsilon_0 < 1. \quad (3.11)$$

By Lemma 3.5, we may fix  $d$  sufficiently large that  $b = 3d \geq r_0$  and, for every  $1 \leq r \leq r_0$ , a uniform rank- $r$  subspace of  $B^*$  fails to belong to  $\mathfrak{S}_r(F)$  with probability at most  $\varepsilon_0$ . In particular, there exists a subspace in  $\mathfrak{S}_{r_0}(F)$ , which we fix once and for all.

### 3.4 Concatenation and its noise

We keep the output space  $B$  fixed while concatenating the quadratic blocks. The output will be a map  $C_t : P_t \rightarrow B$ , a shift action representing every translation of  $B$ , and noise  $\xi_t$  supported on one random leaf. Lemma 3.6 proves change probability  $(1 - q^{-3})(1 - \theta)^t$ . Proposition 3.8 proves detection probability at least  $1/4$  at height  $s = \lfloor H_{r_0}/(8\theta) \rfloor$ . At that height the change probability is at most  $\exp(1 - H_{r_0}/8)$ ; choosing  $r_0$  makes it at most  $p_*$ . The quotient in Section 3.6 preserves both probabilities and realizes the translations as the subspace  $B$  needed in Section 3.1.

We construct binary input spaces  $P_t$ , shift subspaces  $R_t \leq P_t$ , linear surjections  $\lambda_t : R_t \rightarrow B$ , and maps  $C_t : P_t \rightarrow B$  such that

$$C_t(u + h) = C_t(u) + \lambda_t(h) \quad (u \in P_t, h \in R_t). \quad (3.12)$$

A shift  $h \in R_t$  thus represents the output translation  $\lambda_t(h)$ ; different shifts may represent the same translation. This form of the shift action is preserved by aggregation. At the end, quotienting by the shifts of value zero will turn the output translations into an actual subspace of the input.

At height zero take  $P_0 = R_0 = B$  and  $C_0 = \lambda_0 = \text{id}_B$ . Let  $\mathcal{I}$  consist of all pairs  $(A, J)$ , where  $A$  is a line in  $F^3$  and  $J : B \rightarrow U_A$  is a binary linear isomorphism. Each line has the same number of such isomorphisms.

We include one child for every  $(A, J)$  in a single fixed parent map. Choosing a uniform noisy child will then sample a uniform line and a fresh uniform isomorphism. For  $t \geq 1$ , take  $P_t = \prod_{(A, J) \in \mathcal{I}} P_{t-1}$ . For  $u = (u_{AJ}) \in P_t$ , set

$$(x, y) = \sum_{A, J} J C_{t-1}(u_{AJ}), \quad C_t(u) = C_{\text{blk}}(x, y) = y + Q(x). \quad (3.13)$$

A tuple of child shifts changes this aggregate by the sum of their embedded output translations. To make the parent change a pure output translation, we require the  $X$ -component of this sum to vanish. Accordingly, define

$$R_t = \left\{ (h_{AJ}) \in \prod_{A, J} R_{t-1} : \sum_{A, J} J \lambda_{t-1}(h_{AJ}) \in \{0\} \oplus B \right\}, \quad (3.14)$$

and let  $\lambda_t(h) = k$  when this aggregate is  $(0, k)$ . These are linear definitions, and Equation (3.12) follows immediately from Equation (3.13). Surjectivity follows inductively from Lemma 3.2: the child output translations can be chosen arbitrarily, and their aggregate map is onto  $X \oplus B$ . Equation (3.12) and surjectivity also show that  $C_t$  sends a uniform input to a uniform element of  $B$ : for any  $k \in B$ , choose

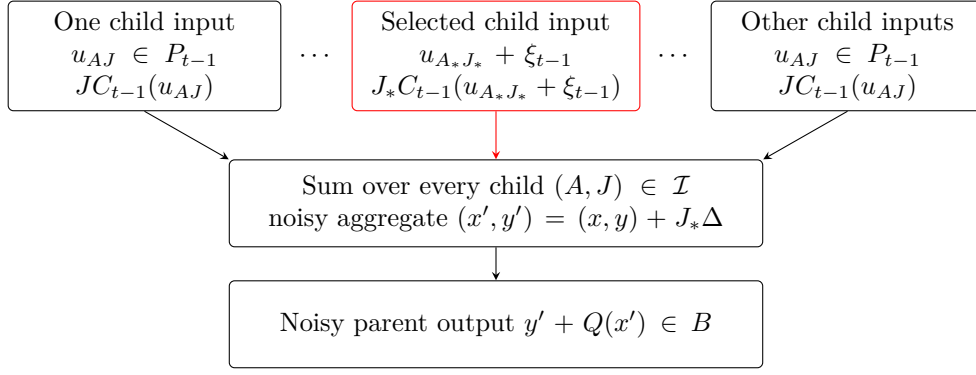


Figure 2: One level of the recursion in Section 3, shown under its additive noise. The input space is a product over all child indices  $(A, J)$ ; only one uniformly chosen child is perturbed. Each child output is embedded in  $U_A \leq X \oplus B$ , all embedded outputs are added, and the quadratic map is applied once at the parent. Here  $(x, y)$  is the initial aggregate, and  $\Delta$  is the selected child's output difference, so the noisy aggregate is  $(x', y') = (x, y) + J_*\Delta$ . The omitted children obey the same rule. The highlighted perturbation may disappear at the parent.

$h \in R_t$  with  $\lambda_t(h) = k$ . Translation by  $h$  bijects  $C_t^{-1}(y)$  with  $C_t^{-1}(y + k)$ , so all output fibers have the same size.

Let  $\xi_t$  be the following additive noise in  $P_t$ . At height zero it is uniform in  $B$ . At a positive height, choose a uniform child  $(A, J)$  and embed an independent copy of  $\xi_{t-1}$  in that child, with zero in all other children. Equivalently, choose a uniform leaf in the product tree and add a uniform symbol of  $B$  at that leaf. Each path choice can be sampled as a uniform  $A$  followed by a uniform  $J$ , independently at successive levels. All these distributions are rational. Figure 2 illustrates one level of the recursion and the selected noisy branch.

**Lemma 3.6.** *At every integer height  $t \geq 0$ , the map  $C_t : P_t \rightarrow B$  and the noise  $\xi_t$  defined above satisfy*

$$\Pr[C_t(u + \xi_t) \neq C_t(u)] = (1 - q^{-3})(1 - \theta)^t, \quad (3.15)$$

where  $u$  is uniform in  $P_t$  and independent of  $\xi_t$ .

*Proof.* Write  $p_t$  for the change probability on the left of Equation (3.15). At height zero,  $p_0 = 1 - q^{-3}$ . Consider one higher level and condition on the selected child  $(A, J)$ . Write  $\Delta$  for its output difference. All child copies have the same function and the same inner noise law, so the law of  $\Delta$  is independent of the selected index.

We need also the conditional law of the child's initial output. For any fixed child perturbation  $a$ , Equation (3.12) gives

$$C_{t-1}(u + a + h) + C_{t-1}(u + h) = C_{t-1}(u + a) + C_{t-1}(u) \quad (h \in R_{t-1}).$$

Thus the difference is invariant under these shifts, while the initial output is translated by every element of  $B$ . Fixing the perturbation and a value of the difference leaves a set of inputs preserved by every shift in  $R_{t-1}$ . Surjectivity of  $\lambda_{t-1}$  gives bijections between the initial-output fibers within this set. On uniform input its initial output therefore remains uniform conditional on the difference, even if the inner perturbation is also specified. All other initial child outputs remain independently uniform. Because the aggregate map is onto  $X \oplus B$ , the initial aggregate  $(x, y)$  is therefore uniform conditional on  $(A, J, \Delta)$ . In particular it is independent of the aggregate perturbation  $J\Delta$ .

Conditional on  $\Delta \neq 0$  and  $A$ , the isomorphism  $J$  remains uniform, because the law of  $\Delta$  is independent of the selected child index. Thus  $J\Delta$  is uniform in  $U_A \setminus \{0\}$ . The conditional uniformity just proved makes

the initial aggregate independent of this update. Lemma 3.2 therefore gives parent change probability  $1 - \theta$  conditional on a nonzero child difference. When  $\Delta = 0$ , the aggregate does not change. Consequently  $p_t = p_{t-1}(1 - \theta)$ , proving the identity.  $\square$

### 3.5 Linear lifts and loss of rank

The nonlinear change probability now decays with the height. For the fixed- $B$  gadget, our remaining target is detection whenever a family's restrictions fill  $B^*$ . We will prove this by tracking every binary-linear choice of extensions of these output linear functionals through the shift action.

For  $S \subseteq B^*$ , a lift on  $P_t$  means a binary linear map  $z \mapsto \psi_z$  from  $S$  to  $P_t^*$  satisfying

$$\psi_z|_{R_t} = z \circ \lambda_t. \quad (3.16)$$

We next describe every possible lift at a parent, without making any assumption about how it was chosen.

**Lemma 3.7.** *Let  $t \geq 1$  and let  $z \mapsto \psi_z$  be a lift  $S \rightarrow P_t^*$  in the sense of Equation (3.16). There is a binary linear map  $\gamma : S \rightarrow X^*$  such that, simultaneously for every child  $(A, J)$  and every  $z \in S$ , the restriction of  $\psi_z$  to that child's copy of  $R_{t-1}$  is*

$$J^*((\gamma_z, z)|_{U_A}) \circ \lambda_{t-1}. \quad (3.17)$$

Here  $J^* : U_A^* \rightarrow B^*$  is precomposition with  $J$ .

*Proof.* The parent lift prescribes its values on all shifts whose aggregate is vertical. We recover its remaining freedom as one functional of the horizontal aggregate.

A shift supported on one child and lying in  $\ker \lambda_{t-1}$  is a parent shift with output translation zero. Each  $\psi_z$  vanishes on it. Since  $\lambda_{t-1}$  is onto, the restriction of  $\psi_z$  to that child's shift space therefore factors through a unique  $\eta_{AJ,z} \in B^*$ .

Let  $T = \prod_{A,J} B$  be the space of child output translations, with aggregate map  $\mathcal{A} : T \rightarrow X \oplus B$ , and write  $\mathcal{A}_X, \mathcal{A}_B$  for its components. Define

$$D_z((b_{AJ})) = \sum_{A,J} \eta_{AJ,z}(b_{AJ}) - z(\mathcal{A}_B((b_{AJ}))).$$

Every tuple in  $\ker \mathcal{A}_X$  can be realized by child shifts and is then a parent shift. Equation (3.16) says exactly that  $D_z$  vanishes on this kernel. Because  $\mathcal{A}_X$  is onto, there is a unique  $\gamma_z \in X^*$  with  $D_z = \gamma_z \circ \mathcal{A}_X$ . Both  $D_z$  and this factorization depend linearly on  $z$ . On a tuple supported at  $(A, J)$ , with value  $b' \in B$ , the identity becomes

$$\eta_{AJ,z}(b') = (\gamma_z, z)(Jb').$$

Thus  $\eta_{AJ,z} = J^*((\gamma_z, z)|_{U_A})$ . Composing with  $\lambda_{t-1}$  gives (3.17), with the same  $\gamma$  for every child.  $\square$

**Proposition 3.8.** *At height*

$$s = \left\lfloor \frac{H_{r_0}}{8\theta} \right\rfloor, \quad (3.18)$$

*every binary linear lift  $z \mapsto \psi_z$  from  $B^*$  to  $P_s^*$  satisfying  $\psi_z|_{R_s} = z \circ \lambda_s$  detects the noise with probability at least  $1/4$ :*

$$\Pr[\text{there exists } z \in B^* \text{ with } \psi_z(\xi_s) \neq 0] \geq \frac{1}{4}. \quad (3.19)$$

*Proof.* We use  $H_r$  as a potential for the rank surviving along a random branch. Its expected loss is small at each level, so with positive probability a nonzero restriction reaches the leaf and detects its uniform noise symbol. Restrict the lift with domain  $B^*$  at the top to the fixed subspace in  $\mathfrak{S}_{r_0}(F)$ . We descend along a uniformly random leaf path as follows.

At a parent with current lift domain  $S$ , take the common map  $\gamma$  from Lemma 3.7 and abbreviate  $\phi_{AJ} : S \rightarrow B^*$  by  $\phi_{AJ}(z) = J^*((\gamma_z, z)|_{U_A})$ . After choosing a child, let  $S' = \text{im } \phi_{AJ}$  and choose a linear section  $\iota : S' \rightarrow S$  of  $\phi_{AJ}$ . Restricting  $\psi_{\iota(z')}$  to that child input gives a lift with domain  $S'$  on  $P_{t-1}$ : for every  $h \in R_{t-1}$ , its value is

$$\phi_{AJ}(\iota(z'))(\lambda_{t-1}h) = z'(\lambda_{t-1}h).$$

We may fix deterministic rules for all such sections. In particular they depend only on the already chosen path, and never on a future path index or on the final noise symbol. On descending the tree, every surviving functional is a restriction of a member of the original linear family.

The factorization now connects the recursion to the local restriction problem. For a fixed parent lift, its single map  $\gamma$  is common to every outgoing line  $A$ , and  $\phi_{AJ} = J^*\kappa_{A,\gamma}$ . Since  $J^*$  is an isomorphism, it does not change the kernel or rank. We may therefore apply Equations (3.9) and (3.10) when the outgoing line is chosen uniformly. The conditioning below shows that the incoming isomorphisms make membership in  $\mathfrak{S}_r(F)$  typical at each positive subsequent rank  $r$ , even though the lift itself depends on the earlier path.

Follow the path, replacing the lift at each chosen child by this image lift. Let  $S_j$  be its domain after  $j$  steps, and let  $r_j = \dim S_j$ , so initially  $r_j = r_0$  at  $j = 0$ . Continue with the zero space after extinction.

We spell out the conditioning that permits Definition 3.4 to be used at every level. Let  $\mathcal{F}_j$  be the information in all path choices before the outgoing choice  $(A_j, J_j)$  at step  $j$ . The current space, lift, and map  $\gamma_j : S_j \rightarrow X^*$  are  $\mathcal{F}_j$ -measurable; write  $\gamma_{j,z} = \gamma_j(z)$  for  $z \in S_j$ . Conditional on  $\mathcal{F}_j$  and  $A_j$ , the subspace

$$E_{j,A_j} = \text{im}[z \mapsto (\gamma_{j,z}, z)|_{U_{A_j}}] \leq U_{A_j}^*$$

is fixed. The next domain and its rank are

$$S_{j+1} = J_j^* E_{j,A_j}, \quad r_{j+1} = \dim E_{j,A_j}.$$

The rank is fixed before  $J_j$  is revealed. Binary automorphisms of  $B^*$  act transitively on subspaces of that rank, and composing  $J_j$  with any automorphism preserves its uniform law. Thus  $S_{j+1} = J_j^* E_{j,A_j}$  is a uniformly random subspace of  $B^*$  of the fixed rank. Therefore, if  $B_j$  is zero when  $r_j = 0$  and is the indicator of  $S_j \notin \mathfrak{S}_{r_j}(F)$  when  $r_j > 0$ , then

$$\mathbb{E}[B_{j+1} \mid \mathcal{F}_j, A_j] \leq \varepsilon_0. \quad (3.20)$$

The next lift may depend on  $J_j$ . This causes no restriction on the argument: the conditions of Definition 3.4 bound every binary linear map  $g$  on the selected subspace, including the map supplied by that lift.

Let  $\Delta_j = H_{r_j} - H_{r_{j+1}} \geq 0$ . Conditional on  $\mathcal{F}_j$ , the outgoing  $A_j$  is independent and uniform, and the rank-loss estimates above apply to the fixed current lift. Consequently

$$\mathbb{E}[\Delta_j \mid \mathcal{F}_j] \leq 3\theta + H_{r_0} 2^{r_0} \theta B_j. \quad (3.21)$$

For  $j = 0$   $S_0 \in \mathfrak{S}_{r_0}(F)$ , so its expected drop is at most  $3\theta$ . For  $j \geq 1$ , average Equation (3.21) over the preceding isomorphism using Equation (3.20). This gives the unconditional bound

$$\mathbb{E}\Delta_j \leq (3 + H_{r_0} 2^{r_0} \varepsilon_0) \theta \leq 4\theta.$$

No independence between an incoming orientation and its subsequent lift has been assumed.

Summing the expected drops and using Equation (3.18),

$$\mathbb{E}[H_{r_0} - H_{r_s}] \leq 4s\theta \leq \frac{H_{r_0}}{2}.$$

Extinction requires a total drop of  $H_{r_0}$ , so  $\Pr[r_s = 0] \leq 1/2$ . At a leaf,  $P_0 = R_0 = B$  and its lift consists of the linear functionals themselves. Conditional on any surviving rank  $r_s > 0$ , an independent uniform symbol in  $B$  is detected with probability  $1 - 2^{-r_s} \geq 1/2$ . Every surviving functional is a restriction of the original family along the selected leaf, so such a detection implies a detection by that family of the embedded noise  $\xi_s$ . Multiplying the two bounds proves Equation (3.19).  $\square$

### 3.6 Realizing the fixed output space

*Completion of the fixed-space construction.* Choose  $r_0$  sufficiently large that

$$\exp(1 - H_{r_0}/8) \leq p_*.$$

Then choose  $\varepsilon_0$  as in Equation (3.11) and fix  $d$ , and hence  $b = 3d$  and  $\theta$ , using Lemma 3.5. Construct the height  $s$  in Equation (3.18).

The fixed- $B$  construction is complete once its shift action is realized by an actual subspace. Put

$$V_0 = P_s / \ker \lambda_s,$$

where the kernel is taken inside  $R_s \leq P_s$ , and let  $q_0 : P_s \rightarrow V_0$  be the quotient map. Equation (3.12) makes  $C_s$  invariant under this kernel, so it descends to a map  $\overline{C} : V_0 \rightarrow B$ . The quotient  $R_s / \ker \lambda_s$  is identified with  $B$  by  $\lambda_s$ , giving an inclusion  $i_B : B \rightarrow V_0$  such that

$$q_0(h) = i_B(\lambda_s h) \quad (h \in R_s).$$

We henceforth identify this subspace with  $B$ . Its shift action satisfies

$$\overline{C}(x + b') = \overline{C}(x) + b' \quad (x \in V_0, b' \in B).$$

Let  $\nu$  be the pushforward of the noise law of  $\xi_s$ . Uniform input on  $P_s$  pushes to uniform input on  $V_0$ , so the output change probability remains  $(1 - q^{-3})(1 - \theta)^s$  by Lemma 3.6.

Every linear subspace  $G_0 \leq V_0^*$  whose restriction is all of  $B^*$  satisfies

$$\Pr_{a \sim \nu}[\text{there exists } g \in G_0 \text{ with } g(a) \neq 0] \geq \frac{1}{4}. \quad (3.22)$$

Indeed, choose a linear section  $\sigma : B^* \rightarrow G_0$  of restriction and define  $\psi_z = q_0^* \sigma(z) = \sigma(z) \circ q_0$ . For  $h \in R_s$ ,

$$\psi_z(h) = \sigma(z)(q_0 h) = z(\lambda_s h).$$

Thus this pullback is a lift with domain  $B^*$ . Moreover,  $\psi_z(\xi_s) = \sigma(z)(q_0 \xi_s)$ , so detection by the lift is detection by the chosen section on the pushed-forward noise. Proposition 3.8 gives the bound, and detection by the section implies detection by  $G_0$ .

Lemma 3.6 gives

$$\Pr_{x \text{ uniform in } V_0, a \sim \nu}[\overline{C}(x + a) \neq \overline{C}(x)] \leq (1 - \theta)^s \leq \exp(-\theta s) \leq \exp(\theta - H_{r_0}/8) \leq p_*. \quad (3.23)$$

Set  $r_* = b + 2$ . All choices so far depend only on  $p_*$ .

For completeness, all choices are effective when  $p_*$  is rational. Choose an integer  $m \geq 0$  with  $2^{-m} \leq p_*$  and then choose  $r_0$  with  $H_{r_0} \geq 8(m + 1)$ ; since  $e^{-m} \leq 2^{-m}$ , this ensures the initial bound on  $r_0$ . Choose rational  $\varepsilon_0$  satisfying Equation (3.11). For successive integers  $d$  with  $3d \geq r_0$ , enumerate the binary subspaces of  $\mathbb{F}_{2^d}^3$  and the linear maps in Definition 3.4. This tests membership in  $\mathfrak{S}_r(F)$  exactly and computes the fraction of spaces outside that family at each rank  $1, \dots, r_0$ . Lemma 3.5 ensures that eventually all these fractions are at most  $\varepsilon_0$ . Fix the first such field and a space in  $\mathfrak{S}_{r_0}(F)$ . The remaining products, isomorphisms, injections, quotients, function tables, and rational noise laws are obtained by finite enumeration and binary linear algebra. This gives the asserted terminating construction.  $\square$

## 4 The outer questions and the permutation test

We use Lemma 3.1 to construct a permutation instance from parity equations. Its vertices must give compatible answers to an equation tuple and to a question obtained by retaining individual variables in selected equations. In particular, an affine map on the projected answer space and its composition with the answer projection must receive the same answer. We enforce this by representing each map on the coordinates on which it depends, including its constant term; Section 4.3 gives the precise table representation.

## 4.1 Weighted parity equations and distinct positions

A *weighted parity instance* consists of a finite variable set  $\mathcal{X}$ , a finite nonempty set  $\mathcal{E}$  of equation occurrences, and positive rational weights  $\omega_e$  summing to one. Each variable  $v \in \mathcal{X}$  takes a value  $x_v \in \mathbb{F}_2$ . An occurrence is an ordered equation

$$x_{v_{e,1}} + x_{v_{e,2}} + x_{v_{e,3}} = b_e, \quad b_e \in \mathbb{F}_2.$$

Its value is the weighted fraction of equations satisfied, and  $\text{OPT}(\mathcal{E})$  denotes the maximum value. Different occurrences have different IDs even if their ordered equations coincide.

**Theorem 4.1 (Weighted parity gap).** *For every fixed rational  $\xi > 0$ , a deterministic polynomial-time reduction maps a Boolean satisfiability instance to a weighted parity instance  $\mathcal{E}$ , with binary-encoded rational weights, such that*

$$\begin{aligned} \text{the input is satisfiable} &\implies \text{OPT}(\mathcal{E}) \geq 1 - \xi, \\ \text{the input is unsatisfiable} &\implies \text{OPT}(\mathcal{E}) \leq \frac{1}{2} + \xi. \end{aligned}$$

*The running-time polynomial may depend on  $\xi$ .*

This is the gap consequence of Håstad's construction in [9, Theorem 5.4 and its proof]. More specifically, the proof uses a fixed dyadic verifier error  $\delta > 0$ , gives completeness  $1 - \delta$  and soundness  $(1 + \delta)/2$ , and enumerates verifier choices as equations with their rational probabilities as weights. Choosing  $\delta \leq \xi$  gives the displayed statement. In that source the proof bits are signs: write each sign as  $(-1)^{x_v}$ . An accepting equation  $s_1 s_2 s_3 = (-1)^b$  then becomes  $x_{v_1} + x_{v_2} + x_{v_3} = b$  over  $\mathbb{F}_2$ , with exactly the same satisfying assignments and weight. Repeated queried bits are allowed here and are handled by Lemma 4.2. The enumeration is a deterministic reduction; the verifier's random choices describe the resulting weighted instance.

**Lemma 4.2 (Making positions distinct).** *Every weighted parity instance  $\mathcal{E}$  can be transformed deterministically in polynomial time into a weighted parity instance  $\mathcal{E}'$  with three distinct variable IDs in every occurrence and*

$$\text{OPT}(\mathcal{E}) \leq \text{OPT}(\mathcal{E}') \leq \frac{1 + \text{OPT}(\mathcal{E})}{2} + \frac{3}{100}.$$

*The new weights are positive rationals of polynomial bit length.*

*Proof.* Replace each variable  $v$  by one hundred clones  $(v, c)$ ,  $c \in [100]$ . For each original occurrence, choose its three clones uniformly subject to choosing distinct clones whenever positions have the same underlying variable. List all the resulting occurrences and give them the corresponding conditional weights. There are at most  $100^3$  choices per original occurrence. Giving all clones their original variable's label proves the completeness assertion.

For the upper bound fix arbitrary clone labels, and let

$$t_v = \frac{1}{100} \sum_{c=1}^{100} (-1)^{x_{(v,c)}}.$$

First choose clones independently, without conditioning on distinctness. The satisfaction probability of occurrence  $e$  is

$$\frac{1 + (-1)^{b_e} t_{v_{e,1}} t_{v_{e,2}} t_{v_{e,3}}}{2}.$$

Round each bias to a majority label, resolving zero biases arbitrarily. For every equation violated by this rounded assignment, the signed product in the numerator is nonpositive, so its satisfaction probability is at most  $1/2$ . The other equations have probability at most one. The weighted mean is therefore at most  $(1 + \text{OPT}(\mathcal{E}))/2$ . The probability of a forbidden collision among the independently chosen clones is at

most  $3/100$ , by a union bound over the three pairs of positions. Conditioning on its complement changes the expectation of any  $[0, 1]$ -valued function by at most  $3/100$ . This proves the bound. All conditional probabilities have constant-size denominators apart from their original occurrence weights.  $\square$

Applying Lemma 4.2 after Theorem 4.1 with  $\xi < 1/100$  preserves completeness at least  $1 - \xi$  and gives NO optimum at most  $4/5$ .

For the rest of the construction,  $\mathcal{E}$  denotes the preprocessed instance, so every equation has three distinct variable IDs. Its *ordinary equation-versus-variable game* samples  $e$  according to  $\omega$  and  $j$  uniformly from  $[3]$ . The first prover receives the occurrence ID  $e$  and answers with a satisfying triple for that equation; the second receives the variable ID  $v_{e,j}$  and answers with a bit. Acceptance means equality at position  $j$ . Because the IDs within an equation are distinct, the question pair determines the projection constraint.

For any second-prover strategy, its bits constitute a global assignment. On an equation violated by that assignment, every valid first-prover answer disagrees in at least one position. Hence this ordinary projection game, denoted  $G_{\mathcal{E}}$ , satisfies

$$\text{val}(G_{\mathcal{E}}) \leq 1 - \frac{1 - \text{OPT}(\mathcal{E})}{3} \leq \frac{14}{15} \quad \text{if } \text{OPT}(\mathcal{E}) \leq \frac{4}{5}. \quad (4.1)$$

The answer alphabets have sizes four and two, respectively.

## 4.2 Tuples, local answer copies, and homogeneous coordinates

Fix a positive integer  $k$ . An equation question  $U = (e_1, \dots, e_k)$  is sampled from  $\omega^{\otimes k}$ , with replacement. Its answer space is the product of the  $k$  affine solution planes. Blocks always use local copies: even if the same equation or variable ID appears in several blocks, their local answers are separate coordinates.

For a parameter  $0 < \beta < 1$ , independently in every block either retain the equation, with probability  $1 - \beta$ , or retain just the variable ID at a uniformly chosen position, with probability  $\beta$ . The resulting question is denoted  $O$ , and  $\pi$  is the coordinate projection from answers on  $U$  to answers on  $O$ . The question  $O$  records an equation ID or a variable ID at every ordered block position. A singleton block accepts either bit.

Represent an equation answer by its first two bits  $(x_i, y_i)$ ; the third is  $b_{e_i} + x_i + y_i$ . A singleton is represented by its actual bit. Add a common homogeneous coordinate  $t_0$ . Thus

$$L_U = \mathbb{F}_2^{1+2k}, \quad m = 1 + 2k,$$

and  $L_O$  has one homogeneous coordinate, two coordinates for each retained equation, and one for each singleton. In either space, vectors whose first coordinate is one are exactly valid answers. Write  $e_0^\top$  for the functional reading that coordinate.

The affine projection  $\pi$  extends to a linear surjection  $\bar{\pi} : L_U \rightarrow L_O$  preserving  $t_0$ . On retained equation blocks it is the identity, and on a block projected to a singleton its coordinate is

$$\begin{cases} x_i, & \text{position 1,} \\ y_i, & \text{position 2,} \\ b_{e_i}t_0 + x_i + y_i, & \text{position 3.} \end{cases} \quad (4.2)$$

We use the blockwise projections obtained by these choices; their family is independent of  $\beta$ . Surjectivity follows by fixing  $t_0$  and choosing coordinates independently in each block. We use the same conventions for questions with some blocks omitted.

### 4.3 Table keys and folding

The identification of tables across projected questions follows Khot–Safra’s identification of entries in their virtual tables [17, Section 1.1]. We first express each table on the coordinates through which it factors. This will give the same representation to a projected table and its pullback.

Fix data  $(\mathcal{V}, K, C, \mu)$  from Lemma 3.1, where  $K = \mathbb{F}_2^\ell$ , and fix a linear complement  $\mathcal{V} = K \oplus W$ . A *table* on a question  $Q$  is an affine map from its answer space to  $\mathcal{V}$ , or equivalently its unique linear extension  $P \in \text{Hom}(L_Q, \mathcal{V})$ . Here  $Q$  may be an equation question or one of its blockwise projections, and the same definition applies to a question with omitted blocks. Write  $c$  for the coefficient of the homogeneous coordinate  $t_0$  in  $P$ .

In an equation block, inspect the rank of its slope map  $\mathbb{F}_2^2 \rightarrow \mathcal{V}$ , whose contribution we write as  $a_i x_i + d_i y_i$ . If its rank is two, retain the equation occurrence and this contribution. If its rank is zero, omit the block. In the remaining case the rank is one, so the coefficients have exactly one of the forms

$$(a_i, d_i) = (v, 0), \quad (0, v), \quad (v, v), \quad v \neq 0.$$

These forms retain only the variable at position one, two, or three, respectively, with contribution  $v z_i$  in its actual bit  $z_i$ . For the third form,

$$v(x_i + y_i) = v z_i + b_{e_i} v t_0, \quad z_i = b_{e_i} t_0 + x_i + y_i,$$

so this replacement also adds  $b_{e_i} v$  to the homogeneous intercept. In a singleton block, retain its variable ID and its original contribution exactly when its slope is nonzero; otherwise omit the block.

The retained blocks form the *support* of  $(Q, P)$ . They keep their indices in  $[k]$ , their equation/variable tags, and their IDs. A retained equation keeps its ordered first two coordinates, and a retained singleton uses its actual-bit coordinate. Thus repeated IDs in different blocks still have separate local coordinates. The support may be empty; its homogeneous coordinate remains in every case.

Let  $J_3$  be the set of equation blocks reduced to their third bit, and write  $v_i$  for the nonzero slope in such a block. The *reduced table*  $\tilde{P}$  on the support has the retained block contributions specified above and homogeneous intercept

$$\tilde{c} = c + \sum_{i \in J_3} b_{e_i} v_i. \tag{4.3}$$

To check the complete table, let  $L_{\text{supp}}$  be the homogeneous answer space of the support and let  $\rho_{Q,P} : L_Q \rightarrow L_{\text{supp}}$  retain  $t_0$  and the specified block coordinates. It is the identity on retained equations and on retained singleton blocks already present in  $Q$ , uses the corresponding actual-bit coordinate from Equation (4.2) on a reduced equation, and discards omitted blocks. Choosing the coordinates independently in each block shows that  $\rho_{Q,P}$  is surjective. The block identities and Equation (4.3) give

$$P = \tilde{P} \rho_{Q,P}.$$

Surjectivity makes  $\tilde{P}$  the unique linear map with this factorization. For this construction, we define the *key* of  $(Q, P)$  to be its support together with this entire reduced table. Thus two tables have the same key exactly when their retained coordinates agree, with the same positions, tags, and IDs, and their induced affine maps agree on those coordinates, including their intercepts.

We next check compatibility with every blockwise projection  $\bar{\pi} : L_U \rightarrow L_O$ . Let  $P' \in \text{Hom}(L_O, \mathcal{V})$ . An equation retained in  $O$  has the same contribution in  $P'$  and its pullback  $P' \bar{\pi}$ , so its reduction is identical. A singleton of slope zero disappears on both sides. A singleton of nonzero slope  $v$  pulls back to the corresponding one of the three rank-one patterns above, which reduces to the same variable ID at the same block position. For a third-position singleton, the pullback adds  $b_{e_i} v$  to the intercept, and its reduction adds  $b_{e_i} v$  once more. The two terms cancel over  $\mathbb{F}_2$ . All other intercept corrections come from retained equation blocks and are identical on the two sides. The supports and complete reduced tables therefore agree:  $(U, P' \bar{\pi})$  and  $(O, P')$  have the same key.

Define the key universe to consist of the keys of all tables on equation questions  $U \in \mathcal{E}^k$ . The compatibility just proved shows that the key of every table on a blockwise projected question already belongs to this universe, by pullback from any compatible  $U$ .

The folding construction of Bellare, Goldreich, and Sudan prescribes scalar output shifts along specified input shifts [2, Section 3.3]. We use constant  $K$ -valued tables as input shifts and translate the answer by the same element of  $K$ .

Adding  $h \in K$  to a key adds  $h$  to its reduced intercept and leaves its support and slopes unchanged. The presenting table becomes  $P + he_0^\top$ , so this action preserves the key universe. Recording the intercept makes the action free. Each orbit has a unique key whose reduced intercept has zero  $K$  component, obtained using the complement  $W$ ; this key represents one vertex of the permutation instance.

For a table  $P$  on an equation question or a blockwise projected question  $Q$ , let  $c(Q, P) \in K$  be the  $K$  component of its key's reduced intercept. Let  $v(Q, P)$  denote the orbit vertex represented by the key obtained after removing this component. A labeling  $\lambda$  of the vertices defines the restored table answer

$$F_Q(P) = \lambda(v(Q, P)) + c(Q, P).$$

Using  $\mathcal{V} = K \oplus W$ , for  $M \in \text{Hom}(L_Q, K)$  and  $T \in \text{Hom}(L_Q, W)$  we also write  $F_Q(M, T) = F_Q(P)$ , where  $P(z) = M(z) + T(z)$  for  $z \in L_Q$ . This notation identifies the underlying space of tables with the product of its two linear-map spaces. Adding a constant changes only the removed component, so

$$F_Q(P + he_0^\top) = F_Q(P) + h \quad (h \in K). \quad (4.4)$$

Tables with the same key have the same representative and removed component. The pullback compatibility therefore gives, for every permitted  $U, O, \bar{\pi}$  and every table  $P'$  on  $O$ ,

$$F_U(P' \bar{\pi}) = F_O(P'). \quad (4.5)$$

Projected questions thus use answers determined by the same vertex labeling, without an additional consistency test.

#### 4.4 The matrix test and completeness

The permutation instance, denoted  $\mathcal{G}$ , samples independently

$$U \sim \omega^{\otimes k}, \quad P \text{ uniform in } \text{Hom}(L_U, \mathcal{V}), \quad a \sim \mu, \quad l \text{ uniform in } \mathbb{F}_2^m,$$

and checks

$$F_U(P) = F_U(P + al^\top). \quad (4.6)$$

Here  $al^\top$  denotes the map  $z \mapsto a(l^\top z)$ . Each sample specifies an edge between the two vertices represented by the corresponding orbits. If their restored offsets are  $c_0, c_1$ , its constraint is the permutation

$$\lambda(v_1) = \lambda(v_0) + c_0 + c_1$$

of the fixed alphabet  $K$ . Edge occurrences have their sampling probabilities as weights. At this stage loops and repeated edges are allowed. The sparse projection experiment involving  $O$  is used to analyze these same vertex labels; it is not an additional test in  $\mathcal{G}$ .

**Lemma 4.3 (Completeness of the outer matrix test).** *Let  $\mathcal{G}$  be the matrix game in Equation (4.6), with the fixed data  $(\mathcal{V}, K, C, \mu)$  and translation equivariance of Equation (3.1). If  $\mathcal{E}$  has an assignment of value at least  $1 - \xi$  and*

$$\Pr_{x \text{ uniform in } \mathcal{V}, a \sim \mu} [C(x + a) \neq C(x)] \leq p_*$$

*for independent  $x, a$ , then*

$$\text{val}(\mathcal{G}) \geq 1 - k\xi - \frac{p_*}{2}.$$

*Proof.* Fix the near-satisfying global bit assignment. Give every reduced support a valid answer as follows. At a singleton use its global bit. At a retained equation use its first two global bits and set the third to the unique value satisfying the equation. These choices depend only on the support, including its occurrence IDs, and therefore define an answer for every key independently of its origin.

At an orbit representative, evaluate its reduced table on this answer and apply  $C$ . This defines a vertex labeling. The equivariance of  $C$  under  $K$  implies that its restored answer is evaluation of the actual reduced table followed by  $C$ , in agreement with Equation (4.4).

With probability at least  $1 - k\xi$ , every equation in  $U$  is satisfied by the global assignment. On such a tuple, the support answers used for every table are projections of the same full answer  $z_U \in L_U$ , whose first coordinate is one. Consequently the two sides of Equation (4.6) are

$$C(Pz_U) \quad \text{and} \quad C(Pz_U + a(l^\top z_U)).$$

Since  $z_U \neq 0$ , the vector  $Pz_U$  is uniform in  $\mathcal{V}$ , independently of  $a, l$ , and  $l^\top z_U$  is an independent fair bit. The test therefore fails with probability at most  $p_*/2$  on these tuples. Adding the probability of a tuple containing an unsatisfied equation proves the bound.  $\square$

## 5 Matrix extraction and Fourier decoding

We convert high acceptance of the matrix test into agreement between two strategies for the sparsely projected questions  $U, O$ . Each strategy uses only its own question and a bounded number of rows of matrix advice. The resulting agreement probability is positive independently of the tuple length  $k$  and the parity instance. The table functions enter only through their folding and sharing identities and the acceptance bound, so the proof applies to any family with these properties; the restored functions of every vertex labeling are examples. We state the precise assumptions below, after fixing the parameters and defining the advice experiment.

### 5.1 Agreement on a matrix slice

The structural input describes functions that often keep their value under a uniform rank-one perturbation. Such a function agrees with an affine evaluation on a set obtained by fixing a bounded number of linear combinations of rows and columns.

Write  $\text{Mat}_{\ell, m} = \text{Hom}(\mathbb{F}_2^m, \mathbb{F}_2^\ell)$ . We use nonempty affine slices specified by row and column equations

$$\mathcal{S} = \{M : d_i^\top M = s_i^\top \ (1 \leq i \leq n_r), \quad Mq_j = t_j \ (1 \leq j \leq n_c)\}, \quad (5.1)$$

where  $d_i \in \mathbb{F}_2^\ell$ ,  $s_i \in \mathbb{F}_2^m$ ,  $q_j \in \mathbb{F}_2^m$ , and  $t_j \in \mathbb{F}_2^\ell$ . The nonnegative integer counts  $n_r, n_c$  need not be the ranks of their respective systems.

**Theorem 5.1 (Agreement under uniform rank-one perturbations).** *For every  $\eta \in (0, 1)$ , there are  $\alpha \in (0, 1]$  and integers  $r_s \geq 1$  and  $\ell_0$ , depending only on  $\eta$ , such that the following holds. For each integer  $\ell \geq \ell_0$ , there is  $m_0 = m_0(\eta, \ell)$  such that every integer  $m \geq m_0$  and every function  $f : \text{Mat}_{\ell, m} \rightarrow \mathbb{F}_2^\ell$  with*

$$\Pr_{M, a, l} [f(M) = f(M + al^\top)] \geq \eta \quad (5.2)$$

*admit a nonempty slice  $\mathcal{S}$  of Equation (5.1), with  $n_r, n_c \leq r_s$ , and vectors  $z \in \mathbb{F}_2^m$ ,  $u \in \mathbb{F}_2^\ell$  for which*

$$\Pr_{M | \mathcal{S}} [f(M) = Mz + u] \geq \alpha. \quad (5.3)$$

*In Equation (5.2),  $M \in \text{Mat}_{\ell, m}$ ,  $a \in \mathbb{F}_2^\ell$ , and  $l \in \mathbb{F}_2^m$  are independent and uniform;  $M | \mathcal{S}$  is uniform on the specified slice. The function  $f$  is arbitrary apart from Equation (5.2).*

The complete derivation is in Appendix C. For an output value  $y$ , consider the matrices with  $f(M) = y$ . Equality acceptance is the average probability that a perturbation stays in this same set, so one such set has positive retention. Associate  $M$  with its graph subspace  $L_M = \{(x, M^\top x) : x \in \mathbb{F}_2^\ell\}$ . A nonzero rank-one change of  $M$  makes the two graph subspaces intersect in dimension  $\ell - 1$ , which is adjacency in the Grassmann graph; exactly half of the Grassmann neighbors of a graph subspace are again graph subspaces. Expansion then supplies subspaces  $A_0 \subseteq B_0$  such that, among the  $L_M$  satisfying  $A_0 \subseteq L_M \subseteq B_0$ , a positive fraction come from matrices with  $f(M) = y$ . The first containment imposes row equations on  $M$ , and the second imposes column equations. These equations define the nonempty slice in (5.1), proving (5.3) with  $z = 0$  and  $u = y$ . Allowing a general affine target remains useful in the subsequent extraction.

The constants and dimension thresholds in Theorem 5.1 can be chosen effectively, with rational  $\alpha$  when  $\eta$  is rational. This follows from the quantitative bound in [14, Theorem 2.13 and Section 2.4] and the comparison estimates in [14, Lemmas 2.7–2.8], with theorem numbers again referring to the 2018 revision.

We now fix the parameters needed below. Choose  $p_* > 0$  and the associated  $r_*$  from Lemma 3.1. Set

$$\eta = 2^{-r_*}/16, \quad 2^{-(\ell-r_s)} < \alpha/8, \quad 2^{-\ell} < \eta/4, \quad (5.4)$$

where  $\alpha, r_s$  first come from Theorem 5.1, and  $\ell$  is then chosen sufficiently large for that theorem, for Lemma 3.1, and for the displayed inequalities. Obtain the data  $(\mathcal{V}, K, C, \mu)$  at this value of  $\ell$ , and fix a linear splitting  $\mathcal{V} = K \oplus W$ . All these objects are fixed before  $k$  grows. In particular,  $\ell > r_s$ . We write a table as  $P = (M, T)$ , with  $M : L_U \rightarrow K$  and  $T : L_U \rightarrow W$ .

## 5.2 The advice experiment and the decoding statement

We first specify the local inputs for two strategies whose answer agreement will contradict the outer-game gap. This experiment is defined for every positive integer  $k$  and every  $0 < \beta < 1$ . Let  $\mathcal{Q}_k$  consist of all equation questions  $U \in \mathcal{E}^k$  and all blockwise projected questions  $O$  obtained from them by the coordinate projections of Section 4. The questions  $U, O$  and the projection  $\bar{\pi} : L_U \rightarrow L_O$  have the law in Section 4: independently in each block, an equation is retained with probability  $1 - \beta$ , and a uniformly chosen variable position is retained with probability  $\beta$ .

**Definition 5.2 (Projected advice experiment).** *Sample  $U, O, \pi$ , then an independent uniform map  $A : K \rightarrow \mathbb{F}_2^{r_s}$  and independent uniform maps*

$$T_O : L_O \rightarrow W, \quad M_O : L_O \rightarrow K.$$

Put  $S_O = AM_O$  and send the following inputs to the two provers:

$$\begin{aligned} U\text{-prover} : & (U, A, T_O \bar{\pi}, S_O \bar{\pi}), \\ O\text{-prover} : & (O, A, T_O, S_O). \end{aligned} \quad (5.5)$$

The map  $M_O$  beyond the stated advice is hidden. Each prover returns an actual answer, that is, a vector of first bit one in its homogeneous answer space. They succeed when their answers  $a_U, a_O$  satisfy  $\bar{\pi}a_U = a_O$ .

The  $U$ -message contains the maps  $(T_U, S_U) = (T_O \bar{\pi}, S_O \bar{\pi})$ , obtained by padding maps sampled on  $O$ . We will first find useful advice under a different law: an unrestricted uniform table  $(M, T)$  on  $U$ , with  $S_U = AM$ . Later we compare the full padded table  $(M_O \bar{\pi}, T_O \bar{\pi})$  with this unrestricted table after averaging over the projection choices.

**Proposition 5.3 (Matrix decoding).** *Fix the data  $(\mathcal{V}, K, C, \mu)$ , the splitting  $\mathcal{V} = K \oplus W$ , and the parameters from Lemma 3.1 and Theorem 5.1 in the order prescribed by Equation (5.4). There are  $\gamma > 0$  and an integer  $k_0 \geq 8$ , depending only on these data, such that the following holds for every integer cube  $k \geq k_0$  and every parity instance with the distinct positions stipulated in Section 4.*

Suppose a family

$$F_Q : \text{Hom}(L_Q, \mathcal{V}) \longrightarrow K \quad (Q \in \mathcal{Q}_k)$$

satisfies

$$\begin{aligned} F_Q(P + ce_0^\top) &= F_Q(P) + c, \\ F_U(P'\bar{\pi}) &= F_O(P') \end{aligned} \tag{5.6}$$

for every  $Q \in \mathcal{Q}_k$ ,  $P \in \text{Hom}(L_Q, \mathcal{V})$ ,  $c \in K$ , every blockwise projection  $\bar{\pi} : L_U \rightarrow L_O$ , and every  $P' \in \text{Hom}(L_O, \mathcal{V})$ . If its acceptance in Equation (4.6) is at least 0.99, then two local strategies in Definition 5.2, with  $\beta = k^{-2/3}$ , agree on the full answer with probability at least  $\gamma$ .

Every labeling of the vertices of  $\mathcal{G}$  supplies such a family: its restored functions satisfy Equation (5.6) by Equations (4.4) and (4.5). The family is fixed when the two response rules are defined and may be built into those rules.

### 5.3 The local strategies and their agreement

Fix a family satisfying (5.6). The equation prover will use a slice certificate determined by its visible advice, and the projected prover will sample Fourier mass on its own fiber. We specify both rules before proving the estimates that make them agree.

For an equation tuple  $U$  and maps  $T : L_U \rightarrow W$ ,  $A : K \rightarrow \mathbb{F}_2^{r_s}$  and  $S_0 : L_U \rightarrow \mathbb{F}_2^{r_s}$ , the row advice specifies the fiber  $\{M \in \text{Hom}(L_U, K) : AM = S_0\}$ . Fix the table family and tuple length. We include  $(U, T, A, S_0)$  in  $\mathfrak{R}$  exactly when there are  $0 \leq n_c \leq r_s$  column equations  $Mq_j = t_j$ , with  $q_j \in L_U$  and  $t_j \in K$ , and vectors  $z \in L_U$ ,  $u \in K$ , such that the solution set of these column equations within the fiber is nonempty and  $F_U(M, T) = Mz + u$  on at least an  $\alpha/2$  fraction of this set, under uniform measure.

The vector  $z$  need not initially be a valid answer. On the slice,  $Mw$  is fixed for every  $w \in \text{span}\{q_j\}$ , so replacing  $z$  by  $z + w$  and  $u$  by  $u + Mw$  leaves the target unchanged. These equivalent target vectors form the coset  $z + \text{span}\{q_j\}$ ; its valid vectors will supply the equation prover's answers.

Lemma 5.6, proved in Section 5.5, guarantees that every member of  $\mathfrak{R}$  has a witness with  $e_0^\top z = 1$ . Choose one by a fixed order on the finite possibilities and put  $Z = \text{span}\{q_j\}$ . Its valid target vectors are

$$\mathcal{A} = (z + Z) \cap \{v : e_0^\top v = 1\}, \quad 1 \leq |\mathcal{A}| \leq 2^{r_s}.$$

Given advice in  $\mathfrak{R}$ , the  $U$ -prover samples  $a_U$  uniformly from  $\mathcal{A}$ ; otherwise it returns any actual answer. Both the witness and this distribution depend only on its visible advice and the fixed table family. The  $O$ -prover will use neither the witness nor the answer set; their projection appears only in the agreement analysis.

The  $O$ -prover starts from its own fiber. Given  $(O, A, T_O, S_O)$ , put  $H = \ker A$  and choose, by a fixed rule, a representative  $M_0$  satisfying  $AM_0 = S_O$ . Such a representative exists on every input in the experiment. Write the fiber as  $M_0 + X$ , where  $X = \text{Hom}(L_O, H)$ . Sample uniform  $\sigma \in K^*$  and form the sign function

$$f_\sigma(N) = (-1)^{\sigma(F_O(M_0 + N, T_O))}, \quad N \in X.$$

Parseval gives  $\sum_\Phi |\hat{f}_\sigma(\Phi)|^2 = 1$ , so these squares define a probability law. Sample a Fourier frequency  $\Phi \in X^*$  with that law. A different representative  $M_0 + N_0$  translates the sign function by  $N_0 \in X$ ; its Fourier coefficient at  $\Phi$  is multiplied by  $(-1)^{\langle \Phi, N_0 \rangle}$ . Thus the sampling law is independent of this auxiliary representative. Identify  $X^* = H^* \otimes L_O$ , with the pairing defined on pure tensors by

$$\langle \varphi \otimes v, N \rangle = \varphi(Nv).$$

For  $h \in H$ , define the contraction  $\Phi(h) \in L_O$  linearly by  $(\varphi \otimes v)(h) = \varphi(h)v$ . If  $\sigma|_H \neq 0$ , choose by a fixed rule an  $h \in H$  with  $\sigma(h) = 1$  and return  $\Phi(h)$ . If the restriction is trivial, return any actual answer.

This rule always returns an actual answer at frequencies of positive mass. Indeed the folding identity in Equation (5.6) gives

$$f_\sigma(N + he_0^\top) = -f_\sigma(N) \quad \text{when } \sigma(h) = 1.$$

Changing variables in a Fourier coefficient shows that its being nonzero requires  $(-1)^{\langle \Phi, he_0^\top \rangle} = -1$ , equivalently  $e_0^\top \Phi(h) = 1$ .

Figure 3 summarizes the information and private choices of the two decoders. The transferred slice will be used only to analyze their agreement probability.

For this analysis, write  $D = (U, O, \pi, A, T_O, S_O)$  for the joint data that determine both prover messages and the projection. These data do not specify  $M_O$  beyond the constraint  $AM_O = S_O$ .

Fix  $U$ -advice in  $\mathfrak{R}$  for a moment, with its chosen witness and answer set  $\mathcal{A}$  from Lemma 5.6. For a blockwise projection put  $z' = \bar{\pi}z$  and  $Z' = \bar{\pi}Z$ . Since  $\bar{\pi}$  preserves the first coordinate,

$$\bar{\pi}\mathcal{A} = (z' + Z') \cap \{v \in L_O : e_0^\top v = 1\}. \quad (5.7)$$

Indeed, every vector in the coset on the right has a preimage in  $z + Z$ , and the preimage has the same first coordinate. We will show that, with positive probability, the vector returned by the  $O$ -prover after evaluating its sampled frequency lies in  $\bar{\pi}\mathcal{A}$ .

*Deduction of Proposition 5.3.* Assume matrix-test acceptance at least 0.99, and take  $k$  sufficiently large along cubes with  $\beta = k^{-2/3}$ . Lemma 5.8, proved in Section 5.6, gives probability at least

$$c_{\text{tr}} = \frac{5\alpha\eta^2}{4} 2^{-2\ell r_s}$$

that the chosen  $U$ -witness transfers to the nonempty slice

$$AM_O = S_O, \quad M_O \bar{\pi} q_j = t_j,$$

where  $F_O(M_O, T_O) = M_O \bar{\pi} z + u$  on at least an  $\alpha/4$  fraction. Lemma 5.9, proved in Section 5.7, applies to every such joint datum  $D$ : with probability at least  $2^{-\ell r_s} (\alpha/8)^2$ , the  $O$ -prover samples an output functional  $\sigma$  nonzero on  $\ker A$  and returns a vector  $\Phi(h) \in \bar{\pi}\mathcal{A}$ , and the independent uniform sample from  $\mathcal{A}$  agrees with it with conditional probability at least  $2^{-r_s}$ . Multiplication gives

$$\Pr[\bar{\pi} a_U = a_O] \geq c_{\text{tr}} 2^{-\ell r_s} (\alpha/8)^2 2^{-r_s} = \frac{5\alpha^3 \eta^2}{256} 2^{-(3\ell+1)r_s} =: \gamma > 0.$$

The largeness requirements for  $k$  come only from Theorem 5.1, the concentration bound in Lemma 5.5, and the total variation estimate in Lemma 5.7. They are uniform in the parity instance and table family. Witness choices and Fourier sampling need not be efficient: they define strategies in a finite game for the soundness argument. Finally, averaging over the provers' independent random tapes also gives deterministic strategies with at least the same agreement probability.  $\square$

Sections 5.4–5.6 prove a positive probability that the equation prover's chosen slice remains nonempty and retains the required agreement after projection. Section 5.7 proves the conditional Fourier bound. The extraction steps use unrestricted uniform tables for arbitrary  $k$ ; the transfer step is where  $k$  is restricted to large cubes and  $\beta = k^{-2/3}$ .

## 5.4 Comparing the two matrix perturbations

The uniform matrix perturbation has the rank-indexed Fourier spectrum computed in [14, Lemma 2.10 in the 2018 revision]. We compare that spectrum with the noise supplied by the construction of Lemma 3.1.

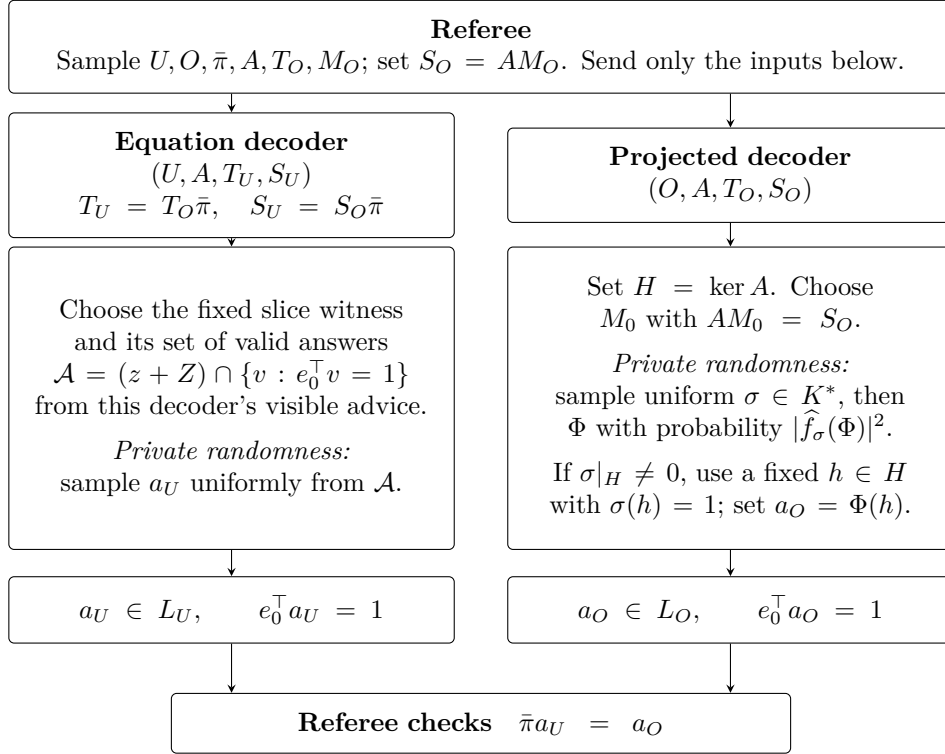


Figure 3: The two decoders use independent random tapes. Neither receives the sampled  $M_O$  or the projection  $\bar{\pi}$ ; the right decoder chooses its own representative  $M_0$  from the visible advice. Every displayed advice map includes its affine intercept, through the homogeneous coordinate. The transferred slice is used only to analyze these private rules: the right decoder does not use the left witness or answer set. If the left advice is outside  $\mathfrak{R}$ , or if  $\sigma|_H = 0$  on the right, that decoder returns an arbitrary valid answer.

**Lemma 5.4 (Spectral comparison).** *For the table family fixed above, acceptance at least 0.99 in Equation (4.6) implies*

$$\mathbb{E}_{U,T} \Pr_{M,a,l} [F_U(M,T) = F_U(M + al^\top, T)] \geq 0.9 2^{-r_*}. \quad (5.8)$$

Here  $U \sim \omega^{\otimes k}$ ; conditional on  $U$ , the maps  $T : L_U \rightarrow W$  and  $M : L_U \rightarrow K$  and the vectors  $a \in K$ ,  $l \in \mathbb{F}_2^m$  are independent and uniform. Under this law, the pairs  $(U, T)$  for which the inner probability is at least  $4\eta$  have probability at least  $10\eta$ .

*Proof.* Both averaging operators are diagonal in the Fourier-character basis. For the constructed noise, the multiplier is at most  $7/8$  at frequencies whose column functionals have restriction rank at least  $r_*$ . Acceptance near one therefore leaves little squared Fourier mass at those frequencies. For uniform noise in  $K$ , each remaining frequency has multiplier at least  $2^{-r_*}$ , so its contribution gives a positive lower bound on acceptance.

For  $\sigma \in K^*$  define the sign function  $f_{U,\sigma}(P) = (-1)^{\sigma(F_U(P))}$ . Let  $e_0, \dots, e_{m-1}$  be the standard basis of  $L_U = \mathbb{F}_2^m$ . A frequency on the matrix domain is a tuple  $g = (g_i)_{i=0}^{m-1}$  of functionals in  $\mathcal{V}^*$ ; its character at  $P$  is  $(-1)^{\sum_i g_i(Pe_i)}$ . All Fourier coefficients use uniform probability measure. Translation by  $al^\top$  multiplies this character by  $(-1)^{\sum_i l_i g_i(a)}$ . Averaging over uniform  $l$  therefore gives the eigenvalue

$$\Lambda_g = \Pr_{a \sim \mu} [g_i(a) = 0 \text{ for every } i].$$

It lies in  $[0, 1]$ . By Lemma 3.1, it is at most  $7/8$  whenever

$$\dim \text{span}\{g_i|_K : 0 \leq i < m\} \geq r_*.$$

Indeed the span of the  $g_i$  is a subspace of  $\mathcal{V}^*$  to which the lemma applies.

Expanding equality by a uniform output character, and then using Fourier orthogonality, expresses the test acceptance as

$$\mathbb{E}_{U,\sigma} \sum_g |\hat{f}_{U,\sigma}(g)|^2 \Lambda_g.$$

Parseval gives total squared mass one for each sign function. If  $B$  is the average squared mass on restricted ranks at least  $r_*$ , then  $0.99 \leq 1 - B/8$ , so  $B \leq 0.08$ . In particular, at least 0.9 of the average mass lies on ranks below  $r_*$ .

When  $a$  is instead uniform in  $K$ , the eigenvalue at  $g$  is  $2^{-\dim \text{span}\{g_i|_K\}}$ . The preceding mass thus gives the lower bound in Equation (5.8). This perturbation leaves  $T$  fixed, so the expression there is precisely the average of the equality tests under uniform rank-one perturbations at fixed  $(U, T)$ . Its lower bound is  $14.4\eta$ . If  $q$  is the probability that the inner acceptance is at least  $4\eta$ , then the average is at most  $4\eta + q$ . Thus  $q \geq 10.4\eta \geq 10\eta$ .  $\square$

## 5.5 Obtaining many useful fibers

Theorem 5.1 gives one structured slice on each qualifying  $(U, T)$  fiber. The  $U$ -prover's information about the  $K$ -valued matrix is the pair  $(A, AM)$ . We need a positive probability that these observed rows determine a useful slice. We will also identify the short set of valid answers supplied by that slice. The erasure argument below proves the probability bound, and folding then makes the answer set nonempty.

For this extraction step, sample  $U$  as before and take independent uniform maps  $T : L_U \rightarrow W$ ,  $A : K \rightarrow \mathbb{F}_2^{r_s}$ , and  $M : L_U \rightarrow K$ . The set  $\mathfrak{R}$  and its chosen witnesses were defined in Section 5.3.

**Lemma 5.5 (Advice admitting a slice with agreement).** *Fix the parameters in Equation (5.4). For all sufficiently large integers  $k$ , every table family satisfying Equation (5.6) with acceptance at least 0.99 in Equation (4.6) satisfies*

$$\Pr_{U,T,A,M} [(U, T, A, AM) \in \mathfrak{R}] \geq 10\eta^2 2^{-\ell r_s}. \quad (5.9)$$

The threshold on  $k$  depends only on the fixed parameters, uniformly over the parity instance and the table family.

**Lemma 5.6 (A finite set of valid answers).** *Fix the parameters in Equation (5.4). For every positive integer  $k$  and every family satisfying Equation (5.6), each  $(U, T, A, S_0) \in \mathfrak{R}$  has a witness to its defining condition with  $e_0^\top z = 1$ . Such a witness can be chosen by a deterministic rule depending only on this advice and the fixed family. With  $Z = \text{span}\{q_j\}$  for its chosen column directions, the set of actual answers is*

$$\mathcal{A} = (z + Z) \cap \{v \in L_U : e_0^\top v = 1\}, \quad 1 \leq |\mathcal{A}| \leq 2^{r_s}. \quad (5.10)$$

*Proof of Lemma 5.5.* We show that useful advice must cover a positive fraction of matrices. Otherwise, randomizing precisely that covered set would preserve acceptance while destroying every slice supplied by Theorem 5.1.

*Randomizing the useful fibers.* Fix  $(U, T)$  for which the acceptance under the uniform rank-one perturbation is at least  $4\eta$ , and write  $f(M) = F_U(M, T)$ . The union used below ranges over all row maps  $A$ ; this lets the argument rule out every slice in Theorem 5.1 at once. Let  $B$  be the union, in the matrix domain  $\text{Mat}_{\ell, m}$ , of all fibers whose advice  $(U, T, A, S_0)$  belongs to  $\mathfrak{R}$ , over every possible  $A, S_0$ . Suppose that its uniform measure is less than  $\eta$ . Change  $f$  at the entries of  $B$ , replacing each by an independent uniform element of  $K$ . Each matrix receives one replacement, even when several of these fibers overlap there. Both endpoints of the shortcode step are uniform matrices. Consequently every such modification changes acceptance by at most  $2\Pr[B]$ , and the modified function has acceptance at least  $2\eta$ .

*Controlling all slice targets.* We show that the random replacements can be chosen so that they contribute less than  $\alpha/4$  of the size of every slice defined by both row and column equations as matches with every target  $Mz + u$ . Put  $r = r_s$ . A nonempty slice with at most  $r$  row and  $r$  column specifications has homogeneous space

$$\text{Hom}\left(\mathbb{F}_2^m / \text{span}\{q_j\}, \bigcap_i \ker d_i^\top\right).$$

Indeed, subtract one solution  $M_*$  from another. The difference vanishes on the column span and takes values in the common kernel of the row functionals, and these are the only homogeneous conditions. If their ranks are  $c$  and  $a$ , the slice has exactly  $2^{(\ell-a)(m-c)}$  elements. Thus its size  $n_S$  is at least  $2^{(\ell-r)(m-r)}$ . Padding both lists with zero equations, each row specification contributes  $\ell + m$  bits, each column specification contributes  $m + \ell$  bits, and the target contributes  $m + \ell$  bits. This gives at most  $2^{(2r+1)(\ell+m)}$  descriptions.

For a fixed description, count only matches at randomized entries and assign a deterministic zero contribution to the other entries. These are independent variables in  $[0, 1]$ , with total mean at most  $2^{-\ell}n_S < \alpha n_S/8$ . Hoeffding's inequality [10, Theorem 2, equation (2.6)] bounds the probability of at least  $\alpha n_S/4$  such matches by  $\exp(-\alpha^2 n_S/32)$ . The union bound over all descriptions is thus at most

$$\exp\left((2r+1)(\ell+m)\log 2 - \frac{\alpha^2}{32}2^{(\ell-r)(m-r)}\right),$$

which tends to zero as  $m = 1 + 2k$  grows. Fix replacements with the asserted property.

*Contradiction and averaging.* Consider any slice of the allowed type. If its row advice belongs to  $\mathfrak{R}$ , every entry of that fiber belongs to  $B$ , and its modified agreement is less than  $\alpha/4$ . If the row advice is outside  $\mathfrak{R}$ , the original agreement with each target on the slice was less than  $\alpha/2$ . The unmodified matches form a subset of those original matches, and the randomized matches add less than  $\alpha/4$ . The modified agreement is then less than  $3\alpha/4$ . Padding shorter row systems by zeros shows that these two cases cover every slice in Theorem 5.1. That theorem now contradicts the modified function's acceptance at least  $2\eta$ . Its applicability does not require the modification to preserve folding.

It follows that  $\Pr[B] \geq \eta$ . For each  $M \in B$ , at least one matrix  $A$  satisfies  $(U, T, A, AM) \in \mathfrak{R}$ . The uniform choice of  $A$  hits that particular matrix with probability  $2^{-\ell r_s}$ . Hence, at this fixed  $(U, T)$ , the

probability of advice in  $\mathfrak{R}$  is at least  $\eta 2^{-\ell r_s}$ . Lemma 5.4 supplies at least  $10\eta$  of such pairs  $(U, T)$ , proving Equation (5.9). All counting and concentration bounds depended only on the fixed parameters and  $m$ , as required.  $\square$

*Proof of Lemma 5.6.* This assertion uses neither high test acceptance nor largeness of  $k$ . Fix any positive integer  $k$ , a family satisfying Equation (5.6), and any row advice in  $\mathfrak{R}$ . Fix a witness, write its column specifications as  $Mq_j = t_j$ , and put  $Z = \text{span}\{q_j\}$ . Nonemptiness of the slice makes the prescribed values consistent: they determine a linear map  $t : Z \rightarrow K$  with  $t(w) = Mw$  throughout the slice. Replacing  $z$  by  $z + w$  and  $u$  by  $u + t(w)$  for  $w \in Z$  therefore preserves the target. It suffices to show that the coset  $z + Z$  contains a vector of first bit one.

Otherwise  $z$  and every vector in  $Z$  have first bit zero. Put  $H = \ker A$ . For each  $h \in H$ , translation  $M \mapsto M + he_0^\top$  preserves the fiber, the column specifications, and the target. By folding it translates  $F_U(M, T)$  by  $h$ . The action is free, and in each of its  $H$ -orbits at most one matrix can satisfy the target equality. The agreement is at most

$$|H|^{-1} = 2^{-\ell + \text{rank } A} \leq 2^{-(\ell - r_s)} < \alpha/8,$$

contradicting the agreement condition defining  $\mathfrak{R}$ . Choose  $w \in Z$  for which  $z + w$  has first bit one and make the corresponding changes to  $z$  and  $u$ . They preserve the original agreement fraction. Choose a witness and such a representative using a fixed order on the finite possibilities. This choice uses only the row advice and the fixed table family. Its vector  $z$  lies in  $\mathcal{A}$ , so the set is nonempty; since  $\dim Z \leq r_s$ , its size is at most  $2^{r_s}$ .  $\square$

## 5.6 Transferring the chosen slice

Lemma 5.5 gives a positive probability of useful advice under unrestricted uniform sampling on  $U$ , and Lemma 5.6 selects a valid-answer witness for each such datum. In the actual experiment, however, tables are sampled on  $O$  and pulled back to  $U$ . Following Khot–Safrá’s distributional comparison [17, Section 3 and Lemma 3.1], with shared linear advice as in [15, Sections 3.3–3.4], we compare these two full table laws, including their affine intercepts, after averaging over projection choices. This comparison will show that the selected slice remains nonempty and retains sufficient agreement with positive probability.

From now on take  $k$  along integer cubes larger than one and set  $\beta = k^{-2/3}$  in Definition 5.2.

The choice  $\beta = k^{-2/3}$  serves two purposes:  $k\beta^2 \rightarrow 0$  makes the padded table law approach the unrestricted law in Lemma 5.7, while  $k\beta \rightarrow \infty$  makes the expected number of singleton coordinates diverge. Once the advice spaces are fixed, each singleton has a uniformly positive probability that every visible advice row has zero coefficient on its variable bit. Lemma 6.2 uses these coordinates to bound agreement.

**Lemma 5.7 (The padded matrix is asymptotically uniform).** *In the projected advice experiment, conditional on any equation tuple  $U$ , the full pulled-back table*

$$(M, T) = (M_O \bar{\pi}, T_O \bar{\pi})$$

*has total variation distance  $o(1)$  from the uniform law on  $\text{Hom}(L_U, K) \times \text{Hom}(L_U, W)$  as  $k \rightarrow \infty$  along integer cubes with  $\beta = k^{-2/3}$ . This comparison averages over the projection choices and is uniform in  $U$  and the parity instance.*

*Proof.* Combine the two components into a  $\mathcal{V}$ -valued table and put  $n = |\mathcal{V}|$ . On an unprojected equation block the two slopes are independent uniform elements of  $\mathcal{V}$ . On a singleton block the padded pair of slopes is an equal mixture of

$$(a, 0), \quad (0, a), \quad (a, a), \quad a \text{ uniform in } \mathcal{V}.$$

Call these pair laws  $P_0$  and  $Q_0$ , respectively. The likelihood ratio  $dQ_0/dP_0$  is  $n$  at  $(0, 0)$ ,  $n/3$  at the other  $3(n-1)$  points on the three displayed lines, and zero elsewhere. Thus

$$\chi^2(Q_0\|P_0) := \mathbb{E}_{P_0} \left[ \left( \frac{dQ_0}{dP_0} - 1 \right)^2 \right] = \frac{n-1}{3}.$$

The one-block mixture  $(1-\beta)P_0 + \beta Q_0$  consequently has chi-squared divergence  $\beta^2(n-1)/3$  from  $P_0$ . Blocks are sampled independently, including when their question IDs coincide. Multiplying their likelihood ratios gives divergence

$$\left( 1 + \frac{n-1}{3}\beta^2 \right)^k - 1 = o(1), \quad (5.11)$$

because  $n$  is fixed and  $k\beta^2 = k^{-1/3}$  tends to zero. Total variation is at most one half the square root of this expression.

For a projection onto the third bit of equation  $e$ , the padded intercept acquires the term  $b_e a$ , where  $a$  is that singleton's slope. The original intercept is independent and uniform in  $\mathcal{V}$ . Adding the sum of these terms leaves it uniform and independent even after all projection choices and slopes are given. The preceding slope estimate therefore proves the assertion for the entire table. Neither the slope laws nor the bound in Equation (5.11) depend on the equation IDs or their right-hand sides.  $\square$

The certifying slice does not automatically survive projection. For example, a column condition  $Mq = t$  becomes  $M_O \bar{\pi} q = t$ ; it is inconsistent if  $\bar{\pi} q = 0$  but  $t \neq 0$ . More generally, projected column directions may acquire new dependencies. The next argument uses the comparison of the full table laws to find a positive probability of samples where the chosen slice remains nonempty and retains enough agreement.

**Lemma 5.8 (Transfer of a chosen witness).** *If the matrix-test acceptance is at least 0.99, then for all sufficiently large integer cubes  $k$ , the joint data  $D$  in Definition 5.2 have the following property with probability at least*

$$\frac{5\alpha\eta^2}{4} 2^{-2\ell r_s}. \quad (5.12)$$

*The padded  $U$ -advice belongs to  $\mathfrak{R}$ , and its witness selected by Lemma 5.6 transfers to the nonempty slice*

$$AM_O = S_O, \quad M_O \bar{\pi} q_j = t_j.$$

*Writing  $z' = \bar{\pi} z$ , the uniform law on this slice satisfies*

$$\Pr[F_O(M_O, T_O) = M_O z' + u \mid AM_O = S_O, M_O \bar{\pi} q_j = t_j \text{ for all } j] \geq \alpha/4. \quad (5.13)$$

*Proof.* A positive excess of target matches over the required agreement level will survive the total variation comparison. Keeping this excess before conditioning avoids any assumption that every chosen column slice survives a projection.

*Comparison before conditioning.* First use unrestricted sampling of  $U, A, M, T$ , setting  $S_0 = AM$ . Let  $\mathcal{J}$  be the event that  $(U, T, A, S_0) \in \mathfrak{R}$  and  $M$  satisfies its chosen column specifications. Let  $\mathcal{M}$  be equality with the chosen target, declaring this event false when the advice is outside  $\mathfrak{R}$ . Within a fiber whose advice belongs to  $\mathfrak{R}$ , the witness chosen by Lemma 5.6 specifies a nonempty column slice with at most  $\ell r_s$  additional scalar linear equations. Its agreement guarantee and the probability bound in Lemma 5.5 give

$$\Pr[\mathcal{J}] \geq 10\eta^2 2^{-2\ell r_s}, \quad \Pr[\mathcal{J} \cap \mathcal{M}] \geq (\alpha/2) \Pr[\mathcal{J}].$$

It follows that

$$\Pr[\mathcal{J} \cap \mathcal{M}] - (\alpha/4) \Pr[\mathcal{J}] \geq \frac{5\alpha\eta^2}{2} 2^{-2\ell r_s}. \quad (5.14)$$

Both events in this expression depend only on  $(U, A, M, T)$ : the row value is  $AM$ , and the certificate is the fixed deterministic function of  $(U, T, A, AM)$ . Adding the independent  $A$  to Lemma 5.7 preserves its variation bound. We may therefore compare the two event probabilities with their values in the padded experiment. If the total variation distance is  $v_k = o(1)$ , the difference in Equation (5.14) changes by at most  $(1 + \alpha/4)v_k$ . For all sufficiently large  $k$  the padded difference is at least the bound in Equation (5.12).

*Locating the transferred slices.* Now condition on  $D$ . The remaining  $M_O$  is uniform on the fiber  $AM_O = S_O$ . The padded  $U$ -advice, and therefore its chosen witness and answer set, are fixed. On advice in  $\mathfrak{R}$  let  $p_D$  be the conditional probability of the projected column slice, and, when  $p_D > 0$ , let  $q_D$  be its conditional equality fraction. Put  $p_D = 0$  when the padded advice is outside  $\mathfrak{R}$ , and put  $q_D = 0$  when  $p_D = 0$ . Sharing gives  $F_U(M_O \bar{\pi}, T_O \bar{\pi}) = F_O(M_O, T_O)$ , so the padded difference is exactly

$$\mathbb{E}_D[p_D(q_D - \alpha/4)] \geq \frac{5\alpha\eta^2}{4} 2^{-2\ell r_s}.$$

The integrand is nonpositive outside the event consisting of padded advice in  $\mathfrak{R}$ ,  $p_D > 0$ , and  $q_D \geq \alpha/4$ , and it is at most one everywhere. That event thus has probability at least the bound in Equation (5.12) and has precisely the asserted properties. The law comparison applied to events depending only on  $(U, A, M, T)$ ; conditioning on  $D$  here only identifies where their positive difference occurs in the projected experiment.  $\square$

## 5.7 Decoding from agreement on a slice

Extracting an outer strategy from Fourier mass follows the approach of Håstad [9, Section 1.1] and the list decoding from Fourier coefficients in Khot–Safra [17, Section 1.1]. Here the projected prover samples a frequency with probability equal to the squared Fourier coefficient of the label character on its own advice fiber. We use agreement on the transferred slice to show that evaluating this frequency as in the decoder rule returns a vector in  $\bar{\pi}\mathcal{A}$  with positive probability.

**Lemma 5.9 (A transferred slice gives agreement).** *Fix joint data  $D$  for which the selected witness transfers to a nonempty slice with the agreement in Equation (5.13). Conditional on  $D$ , the two decoder rules above, with independent private randomness, satisfy*

$$\Pr[\bar{\pi}a_U = a_O \mid D] \geq 2^{-\ell r_s} (\alpha/8)^2 2^{-r_s}. \quad (5.15)$$

*Proof.* Agreement on the transferred slice forces Fourier mass in a coset whose contractions are projected answers. We first identify that coset, keeping the affine part of the slice in the Fourier phase.

Use the decoder’s representative  $M_0$  and write  $M_O = M_0 + N$ , with  $N \in X = \text{Hom}(L_O, H)$  and  $H = \ker A$ . Put  $Z' = \bar{\pi}Z$ , so  $\dim Z' \leq r_s$ . Each projected column condition reads

$$N(\bar{\pi}q_j) = t_j - M_0 \bar{\pi}q_j.$$

Nonemptiness ensures that these values respect all dependencies among the  $\bar{\pi}q_j$  and lie in  $H$ . They therefore define a linear map  $b : Z' \rightarrow H$ . Choose  $N_*$  satisfying  $N_*|_{Z'} = b$ . The slice is  $N_* + Y$ , where  $Y = \{N \in X : N|_{Z'} = 0\}$ . For every  $\sigma \in K^*$  define, with  $N$  uniform in  $X$ ,

$$C_\sigma = \mathbb{E}[f_\sigma(N)(-1)^{\sigma(Nz')} \mid N|_{Z'} = b].$$

The character expansion of Equation (5.13) gives

$$\mathbb{E}_\sigma[(-1)^{\sigma(M_0 z' + u)} C_\sigma] \geq \alpha/4, \quad \text{so} \quad \mathbb{E}_\sigma|C_\sigma| \geq \alpha/4.$$

Exactly a fraction  $2^{-\dim H}$  of the characters have trivial restriction to  $H$ . Since  $\dim H \geq \ell - r_s$  and  $|C_\sigma| \leq 1$ , Equation (5.4) implies

$$\mathbb{E}_\sigma[\mathbf{1}_{\{\sigma|_H \neq 0\}} |C_\sigma|] \geq \alpha/8. \quad (5.16)$$

The annihilator of  $Y$  in  $X^*$  is  $H^* \otimes Z'$ . To see both inclusions, extend a basis of  $Z'$  to one of  $L_O$ : maps in  $Y$  are zero on the first basis vectors and arbitrary  $H$ -valued maps on the remaining vectors, so annihilating them forces precisely the remaining dual coefficients to vanish. In particular, for any  $\Omega \in X^*$ ,

$$\mathbb{E}_{N \in N_* + Y} (-1)^{\langle \Omega, N \rangle} = \mathbf{1}_{\{\Omega \in H^* \otimes Z'\}} (-1)^{\langle \Omega, N_* \rangle}.$$

For  $\theta_\sigma = (\sigma|_H) \otimes z'$ , apply this identity to  $\Omega = \Phi + \theta_\sigma$  in the Fourier expansion of  $f_\sigma$ . It gives

$$C_\sigma = \sum_{\Phi \in \theta_\sigma + H^* \otimes Z'} \hat{f}_\sigma(\Phi) (-1)^{\langle \Phi + \theta_\sigma, N_* \rangle}.$$

This sum has  $2^{\dim H \dim Z'} \leq 2^{\ell r_s}$  terms. Cauchy–Schwarz, first on each sum and then in the average over  $\sigma$ , shows from Equation (5.16) that

$$\begin{aligned} \mathbb{E}_\sigma \left[ \mathbf{1}_{\{\sigma|_H \neq 0\}} \sum_{\Phi \in \theta_\sigma + H^* \otimes Z'} |\hat{f}_\sigma(\Phi)|^2 \right] &\geq 2^{-\ell r_s} \mathbb{E}_\sigma [\mathbf{1}_{\{\sigma|_H \neq 0\}} C_\sigma^2] \\ &\geq 2^{-\ell r_s} (\alpha/8)^2. \end{aligned}$$

This is a lower bound on the probability that the  $O$ -decoder samples a nontrivial restriction and a frequency in the indicated coset.

Every such frequency has the form  $\Phi = (\sigma|_H) \otimes z' + \Psi$  with  $\Psi \in H^* \otimes Z'$ . For the decoder's choice of  $h$  it satisfies

$$\Phi(h) = z' + \Psi(h) \in z' + Z'.$$

As already proved, a frequency of positive mass gives first bit one. Equation (5.7) therefore places the returned vector in  $\bar{\pi}\mathcal{A}$ . Independently, the  $U$ -decoder samples uniformly from  $\mathcal{A}$  and chooses a preimage with probability at least  $1/|\mathcal{A}| \geq 2^{-r_s}$ . Multiplying the bounds proves Equation (5.15).  $\square$

## 6 Coordinates with zero advice and the soundness gap

Proposition 5.3 converts high acceptance of the matrix test into agreement between two local strategies supplied with correlated linear advice. We now bound the value of this advice experiment on a NO parity instance. The use of advice that vanishes on selected coordinates follows [15, Sections 3.3–3.4]. The key is to identify singleton coordinates on which both advice slopes vanish. After fixing the data outside those coordinates, their question pairs have the ordinary independent equation–variable law. Each side can then reconstruct its original advice input from its own remaining questions and the fixed data. This defines strategies for an ordinary parallel repetition, whose value bounds the original agreement.

### 6.1 Classical projection-game repetition

We use the following form of the parallel repetition theorem of Dinur and Steurer. The value here is the ordinary classical value: each prover answers using only its own question. A projection constraint maps each answer on one side to its required answer on the other side.

**Theorem 6.1 (Projection-game repetition).** *For every finite bipartite projection game  $G$ , every integer  $n \geq 1$ , and every  $g \in (0, 1)$ ,*

$$\text{val}(G) \leq 1 - g \implies \text{val}(G^{\otimes n}) \leq (1 - g^2/16)^n. \quad (6.1)$$

*The constraint occurrences may have any probability distribution, including parallel occurrences. Here  $G^{\otimes n}$  is the independent product from Section 2: each prover sees its own question tuple, and all  $n$  constraints must hold. The bound is uniform over all finite question and answer sets.*

This is Corollary 1.2 of [6]. Their Section 2.1 allows nonnegative edge weights and parallel edges, and Section 2.2 uses the independent product defined in our Section 2, so the bound applies to our finite weighted question distributions.

Their projection convention lets the second answer determine the first. Identify their second prover with our equation prover and their first with our variable prover. Exchanging the question sides in every product coordinate preserves value. Either orientation works for Unique Games because each constraint is a permutation. The deterministic and randomized value conventions are those established in Section 2.

## 6.2 Singleton coordinates with zero advice slope

Let  $\mathcal{E}$  be the parity instance from Section 4, with three distinct variable IDs in each equation occurrence. Recall its ordinary equation-variable game  $G_{\mathcal{E}}$ : the first prover receives an equation and answers with a satisfying triple; the second receives a uniformly selected variable from that equation and answers with a bit. On a NO instance, Equation (4.1) gives  $\text{val}(G_{\mathcal{E}}) \leq 14/15$ .

Recall the advice experiment of Section 5, taking  $k$  along integer cubes larger than one and setting  $\beta = k^{-2/3}$ . An ordered tuple  $U$  consists of  $k$  independent equation samples. Independently in each block, with probability  $\beta$  we keep a uniformly chosen variable position, and otherwise keep the entire equation. This produces  $O$  and the answer projection  $\bar{\pi} : L_U \rightarrow L_O$  of Equation (4.2). Independently choose

$$A : K \rightarrow \mathbb{F}_2^{r_s}, \quad T_O : L_O \rightarrow W, \quad M_O : L_O \rightarrow K$$

uniformly, and put  $S_O = AM_O$ . The two inputs, respectively, are

$$(U, A, T_O \bar{\pi}, S_O \bar{\pi}) \quad \text{and} \quad (O, A, T_O, S_O). \quad (6.2)$$

The matrix  $M_O$  beyond  $S_O$  is not part of either input. Each prover must return an actual answer, equivalently a vector whose first coordinate is one in its respective answer space.

**Lemma 6.2 (Agreement bound from zero advice slopes).** *Let  $\mathcal{E}$  have three distinct variable IDs per occurrence and  $\text{OPT}(\mathcal{E}) \leq 4/5$ . For the fixed advice spaces above, every integer cube  $k > 1$  with  $\beta = k^{-2/3}$  and every pair of local randomized strategies in Definition 5.2 satisfy*

$$\Pr[\bar{\pi} a_U = a_O] \leq \left(1 - \frac{\beta 2^{-\dim W - r_s}}{3600}\right)^k \leq \exp\left(-\frac{2^{-\dim W - r_s}}{3600} k^{1/3}\right). \quad (6.3)$$

Thus the agreement tends to zero uniformly over the parity instance as  $k$  grows along cubes, with the advice spaces fixed.

## 6.3 The matrix-game gap

We can now compare the upper bound for every advice strategy with the strategies extracted from a highly accepting matrix labeling.

**Proposition 6.3 (A fixed gap with adjustable completeness).** *Fix  $p_* > 0$  and the data from Lemma 3.1 and Theorem 5.1 in the order of Equation (5.4). For every sufficiently large integer cube  $k$ , the matrix game  $\mathcal{G}$  of Section 4 satisfies*

$$\begin{aligned} \text{OPT}(\mathcal{E}) \geq 1 - \xi &\implies \text{val}(\mathcal{G}) \geq 1 - k\xi - p_*/2, \\ \text{OPT}(\mathcal{E}) \leq 4/5 &\implies \text{val}(\mathcal{G}) \leq 99/100 \end{aligned}$$

for every weighted parity instance  $\mathcal{E}$  with three distinct variable IDs in each occurrence. The threshold on  $k$  depends only on the fixed data, uniformly in  $\mathcal{E}$  and  $\xi$ .

*Proof.* Completeness is Lemma 4.3. For soundness, fix a labeling of  $\mathcal{G}$ . Its restored table functions satisfy the folding and sharing identities of Section 4. Proposition 5.3 supplies a constant  $\gamma > 0$  and, for every sufficiently large cube  $k$ , gives local advice strategies with agreement at least  $\gamma$  whenever those functions have matrix-test acceptance at least 0.99. For all sufficiently large cubes  $k$ , the upper bound in Equation (6.3) is smaller than  $\gamma$ . Lemma 6.2 contradicts the existence of those strategies whenever  $\text{OPT}(\mathcal{E}) \leq 4/5$ . All thresholds are uniform over the parity instance.  $\square$

The order matters:  $p_*$  can be made arbitrarily small before the gadget parameters are fixed, and  $\xi$  can be made arbitrarily small after  $k$  is fixed. Thus the same NO gap is compatible with any prescribed small completeness error.

## 6.4 Proof of the zero-advice bound

We keep the common affine intercept fixed while reconstructing the two local inputs on coordinates with zero slopes.

*Proof.* The zero slopes remove every dependence of the advice on the hidden equation–variable choices in those slots. We freeze the other data and use the original response rules to define an ordinary repeated-game strategy.

Fix an arbitrary pair of local randomized strategies in the advice experiment. We first condition on  $A$  and combine the two visible advice maps into

$$D_O = (T_O, S_O) : L_O \longrightarrow E_A, \quad E_A = W \oplus \text{im } A.$$

Conditional on the questions and projection choices, the intercept  $c_O$  and all block slope columns of  $D_O$  are independent uniform elements of  $E_A$ . Indeed, the columns of  $T_O$  and  $M_O$  are independent and uniform, and applying  $A$  to a uniform column gives a uniform element of  $\text{im } A$ . This coefficient law depends only on whether a block is an equation or a singleton, and not on any equation ID, variable ID, or selected position. In particular,  $c_O$  is independent of all block data.

Let  $I \subseteq [k]$  consist of the coordinates projected to a singleton whose one advice slope  $d_i \in E_A$  is zero. Conditional on  $A$ , the indicators  $\mathbf{1}_{\{i \in I\}}$  are independent, with common probability

$$p_A = \frac{\beta}{|E_A|} = \beta 2^{-\dim W - \text{rank } A}. \quad (6.4)$$

Write  $j_i \in [3]$  for the selected variable position at a singleton. More precisely, for every equation occurrence  $e$  and position  $j \in [3]$ ,

$$\Pr[e_i = e, i \in I, j_i = j \mid A] = \omega_e \frac{\beta}{3|E_A|}. \quad (6.5)$$

Thus, conditional on belonging to  $I$ , a coordinate has precisely the ordinary equation–variable distribution.

For the analysis, let  $R$  consist of  $A$ ,  $I$ , the original intercept  $c_O$ , and the following data for every  $i \notin I$ : its equation occurrence, whether the block was retained or projected, the selected position in the latter case, and its advice slope columns. No witness, decoder output, agreement event, or response randomness is included in  $R$ .

For each value  $R = r$  of positive probability, regard the set  $I$  recorded in  $r$  as fixed. When  $I$  is nonempty, we will use  $r$  as constant data when defining two strategies for  $G_{\mathcal{E}}^{\otimes |I|}$ . Each strategy will reconstruct one original advice input from its own product-game questions, apply the corresponding original response rule, and retain the answers in the slots in  $I$ .

Conditional on  $R = r$ , the pairs

$$(e_i, v_{e_i, j_i}), \quad i \in I,$$

are independent question pairs of  $G_{\mathcal{E}}$ . Indeed, conditional on  $A$ , the coordinate data are independent, and Equation (6.5) shows that conditioning on the set  $I$  preserves the ordinary law in each coordinate in  $I$ . The outside coordinate data and the independent intercept impose no further condition on these pairs.

We now verify the two reconstructions, including the affine intercept shared across the blocks. An equation with right-hand side  $b_e$  has actual bits  $x, y, b_e t_0 + x + y$ . Pulling back a singleton slope  $d_i$  thus produces the two slopes

$$(d_i, 0), \quad (0, d_i), \quad (d_i, d_i) \quad (6.6)$$

at positions one, two, and three, respectively. The intercept of  $D_U = D_O \bar{\pi}$  is

$$c_U = c_O + \sum_{\substack{i \text{ projected to a singleton} \\ j_i=3}} b_{e_i} d_i. \quad (6.7)$$

For a coordinate in  $I$ , every slope in Equation (6.6) is zero and its contribution to Equation (6.7) is zero. All remaining terms in that equation are determined by  $r$ .

The equation side receives its occurrences  $e_i$  for  $i \in I$ . Together with  $r$ , they specify all of  $U$ . Its two advice slopes in each block in  $I$  are zero, while  $r$  determines every other slope and the full intercept. These coefficients reconstruct both padded advice maps without using any selected position or variable question in  $I$ . The variable side receives the variable question for each  $i \in I$ . Together with  $r$ , these specify all of  $O$ . Its singleton slopes in  $I$  are zero, and  $r$  supplies every other slope and the original intercept. These coefficients reconstruct its two advice maps without using any equation in  $I$ . Repeated IDs across different blocks cause no additional restriction: questions were sampled independently with replacement, and answers use the local copies specified in Section 4.

These reconstructions define the promised product-game strategies: each side applies its original response rule to the reconstructed input and retains its answers in the slots  $I$ , in their original order. Their equation answers are satisfying triples because the original answers are valid. For every realization of the questions in  $I$  and response randomness, full agreement in the advice experiment implies acceptance in every slot in  $I$ . If  $I = \emptyset$ , the conditional success probability is at most one. Otherwise,

$$\Pr[\bar{\pi} a_U = a_O \mid R = r] \leq \text{val}(G_{\mathcal{E}}^{\otimes |I|}) \leq \left(1 - \frac{1}{3600}\right)^{|I|}.$$

The first inequality uses the strategies just constructed and the conditional product law. The second uses Theorem 6.1 and Equation (4.1). The original response rules are evaluated only at their original local inputs;  $r$  indexes the auxiliary product-game strategies. Thus the final bound holds for every  $r$ , including  $I = \emptyset$  when its right-hand side is interpreted as one.

Finally, conditional on  $A$ , the size of  $I$  has distribution  $\text{Binomial}(k, p_A)$ . Averaging the preceding inequality first over  $R$  conditional on  $A$ , and then over  $A$ , gives

$$\begin{aligned} \Pr[\bar{\pi}(a_U) = a_O] &\leq \mathbb{E}_A \left(1 - \frac{p_A}{3600}\right)^k \\ &\leq \left(1 - \frac{\beta 2^{-\dim W - r_s}}{3600}\right)^k \\ &\leq \exp\left(-\frac{k\beta 2^{-\dim W - r_s}}{3600}\right), \end{aligned}$$

since  $\text{rank } A \leq r_s$ . Substituting  $k\beta = k^{1/3}$  proves Equation (6.3).  $\square$

## 7 Completion of the reduction

The matrix game has a fixed soundness gap and an arbitrarily small completeness error. We now obtain the form required by Theorem 1.1: unweighted, simple bipartite games with arbitrary prescribed errors. Rounding the occurrence weights and subdividing each occurrence will preserve a fixed gap. One application of alphabet-independent repetition then completes the proof.

### 7.1 Removing weights and repeated edges

We use the following elementary construction for weighted permutation constraints. At this intermediate stage, a constraint occurrence may be a loop, and several occurrences may join the same vertices.

**Lemma 7.1 (Rounding and subdivision).** *Let  $H$  be a finite permutation-constraint game with  $h \geq 1$  positive rational occurrence weights summing to one, allowing loops and parallel occurrences. For every positive rational  $\tau$ , one can deterministically construct an unweighted simple bipartite game  $B$  on the same alphabet, with  $4\lceil h/\tau \rceil$  edges and*

$$\left| \text{val}(B) - \left( 1 - \frac{1 - \text{val}(H)}{4} \right) \right| \leq \frac{\tau}{4}. \quad (7.1)$$

*Translation constraints on a finite binary vector space remain translations. For fixed  $\tau$ , the construction is polynomial in the encoding length of the vertex list, constraint occurrences, permutation tables, and binary rational weights of  $H$ .*

The proof is given in Section 7.3. Rounding approximates the occurrence distribution in total variation. Subdivision then gives exactly four satisfied edges for a satisfied occurrence and at most three for a violated one, with disjoint interiors making these choices independent. This is the source of the factor  $1/4$  in (7.1).

### 7.2 The final reduction

*Proof of Theorem 1.1. Error bounds and repetition.* Fix  $\varepsilon, \delta \in (0, 1/2)$ . Choose positive rational  $\varepsilon_0 \leq \varepsilon$  and  $\delta_0 \leq \delta$ . We will take  $\tau \leq 1/200$  when applying the rounding and subdivision lemma. For any matrix game  $H$  satisfying the NO bound  $\text{val}(H) \leq 99/100$ , the resulting game  $B$  then satisfies

$$\text{val}(B) \leq 1 - \frac{1/100 - \tau}{4} \leq 1 - \frac{1}{800}. \quad (7.2)$$

This gap does not depend on the alphabet. We can therefore use alphabet-independent repetition to choose its exponent before the construction in Section 3 fixes that alphabet. Choose an integer  $t \geq 1$  such that

$$\left( 1 - \frac{1}{10240000} \right)^t \leq \delta_0. \quad (7.3)$$

For example,  $t \geq \lceil 10240000 \log(1/\delta_0) \rceil$  suffices. By Theorem 6.1, this exponent reduces every projection game of value at most  $1 - 1/800$  to value at most  $\delta_0$ .

Choose positive rational parameters

$$p_* \leq \frac{\varepsilon_0}{t}, \quad \tau \leq \min \left\{ \frac{1}{200}, \frac{\varepsilon_0}{t} \right\}. \quad (7.4)$$

*Noise and dimension parameters.* Choose the rank threshold  $r_*$  from Lemma 3.1, put  $\eta = 2^{-r_*}/16$ , and choose the corresponding agreement parameters  $\alpha, r_s$ , using a positive rational lower bound for  $\alpha$ . Then choose a sufficiently large  $\ell$ , as specified in Equation (5.4), and fix the corresponding data  $(\mathcal{V}, K, C, \mu)$  and a splitting  $\mathcal{V} = K \oplus W$ , with  $K = \mathbb{F}_2^\ell$ .

Once the rational budgets are fixed, these choices are effective. For rational  $p_*$ , the effective assertion of Lemma 3.1 supplies all the finite data, including a computable rational noise law, by a terminating deterministic construction. The effective bounds stated after Theorem 5.1 supply the rational choice of  $\alpha$  and the dimension thresholds, including  $m_0(\eta, \ell)$  after  $\ell$  has been fixed.

*Choosing the tuple length.* Next choose a cube  $k \geq 8$  large enough for Proposition 6.3. Such a cube can be found by checking the estimates used in its proof. Write  $m = 1 + 2k$  and  $\beta = k^{-2/3}$ ; since  $k$  is a cube,  $\beta$  is rational. Require  $m \geq m_0(\eta, \ell)$  and

$$\frac{\alpha^2}{32} 2^{(\ell-r_s)(m-r_s)} > (2r_s + 1)(\ell + m).$$

The latter is a sufficient rational condition for the union bound in Lemma 5.5 to be less than one. For witness transfer in Lemma 5.8, require

$$\frac{1}{4} \left[ \left( 1 + \frac{|\mathcal{V}| - 1}{3} \beta^2 \right)^k - 1 \right] < \frac{25\alpha^2\eta^4}{16(1 + \alpha/4)^2} 2^{-4\ell r_s},$$

using the lower bound in Equation (5.12). The left-hand side is the square of the total-variation bound. For the final contradiction, require

$$\left( 1 - \frac{\beta 2^{-\dim W - r_s}}{3600} \right)^k < \gamma,$$

with  $\gamma$  from Proposition 5.3. All these conditions are rational and hold for sufficiently large cubes  $k$ . Enumerating cubes until they hold therefore finds a suitable  $k$ ; the search cost depends only on the two prescribed errors and is incurred before the input formula is used.

Finally, with  $k$  fixed, choose a positive rational

$$\xi < \min \left\{ \frac{1}{100}, \frac{\varepsilon_0}{2kt} \right\}. \quad (7.5)$$

Every choice depends only on the two prescribed errors. In particular,  $r_*$  is fixed before  $\ell$ , the advice dimensions before  $k$ , and  $k$  before the parity error  $\xi$ .

*Gap and output structure.* Given a 3SAT formula, apply Theorem 4.1 and Lemma 4.2, and construct the matrix game  $H$  of Section 4. Proposition 6.3 gives

$$\begin{aligned} \text{YES: } \text{val}(H) &\geq 1 - d, \quad d = k\xi + p_*/2 \leq \varepsilon_0/t, \\ \text{NO: } \text{val}(H) &\leq 99/100. \end{aligned}$$

Apply Lemma 7.1 with  $\tau$  to obtain the unweighted simple bipartite game  $B$ . It satisfies

$$\begin{aligned} \text{YES: } \text{val}(B) &\geq 1 - (d + \tau)/4, \\ \text{NO: } \text{val}(B) &\leq 1 - 1/800, \end{aligned}$$

where the NO bound is Equation (7.2). The final output is  $G = B^{\otimes t}$ . Its NO value is at most  $\delta_0 \leq \delta$ , by Equations (6.1) and (7.3). On a YES input, use a labeling of  $B$  coordinatewise. A union bound gives

$$\text{val}(G) \geq 1 - \frac{t(d + \tau)}{4} \geq 1 - \varepsilon_0/2 \geq 1 - \varepsilon.$$

The output has the claimed graph and constraint form. Its two vertex sets are the  $t$ -fold products of the two sides of  $B$ . A pair of product vertices determines at most one edge in each coordinate, since  $B$  is simple, and hence at most one product edge. Independent uniform edge sampling is uniform on these product edges. The constraints of  $H$  are translations on  $K$ , by Section 4; rounding and subdivision

preserve them. The product of translations with offsets  $c_1, \dots, c_t$  is translation by  $(c_1, \dots, c_t)$  on  $K^t = \mathbb{F}_2^{\ell t}$ . This is the alphabet  $\mathbb{F}_2^s$  in the theorem, with  $s = \ell t \geq 1$  fixed by the prescribed errors.

*Polynomial encoding.* To verify the polynomial encoding bound, suppose the preprocessed parity instance has  $N_{\mathcal{E}}$  positive equation occurrences. There are  $N_{\mathcal{E}}^k$  equation tuples. The number of tables, noise outcomes, and perturbation vectors for each tuple is fixed after the parameter choices above, so the number  $h$  of listed matrix-test occurrences is a fixed constant times  $N_{\mathcal{E}}^k$ . Omit zero-probability outcomes and retain repeated occurrences separately. Each weight is a product of  $k$  parity weights, a fixed rational noise probability, and fixed uniform-sampling factors. It has polynomial binary length.

Table keys and folded representatives are computed by linear algebra in fixed dimensions, keeping the original variable and occurrence IDs. At most  $2h$  endpoints appear. Even comparing every pair of keys to identify equal vertices takes polynomial time; number the resulting distinct vertices consecutively. Largest-remainder rounding uses  $Q = \lceil h/\tau \rceil = O(h)$  copies; subdivision creates  $3Q$  new vertices and  $4Q$  edges. The final product has  $(4Q)^t$  edges and polynomially many vertices, because  $t$  is fixed. Full permutation tables on the fixed alphabet  $\mathbb{F}_2^{\ell t}$  have constant size, and endpoint indices have  $O(t \log(h+1))$  bits. These observations give a deterministic polynomial-time explicit construction throughout and complete the proof.  $\square$

### 7.3 Proof of rounding and subdivision

*Proof of Lemma 7.1.* Put  $Q = \lceil h/\tau \rceil$  and write the occurrence weights as  $w_1, \dots, w_h$ . Start with  $n_e = \lfloor Qw_e \rfloor$ , and add one to the  $Q - \sum_e \lfloor Qw_e \rfloor$  entries having largest fractional parts, breaking ties in a fixed order. Then

$$n_e \geq 0, \quad \sum_e n_e = Q, \quad |w_e - n_e/Q| \leq 1/Q.$$

The weight differences sum to zero. Consequently, for every labeling, the change in the weight of its satisfied-occurrence set is at most

$$\frac{1}{2} \sum_{e=1}^h |w_e - n_e/Q| \leq \frac{h}{2Q} \leq \tau.$$

Replace each occurrence  $e$  by  $n_e$  copies, omitting it if  $n_e = 0$ . The resulting unweighted multigraph  $H'$  has  $Q$  occurrences and  $|\text{val}(H') - \text{val}(H)| \leq \tau$ .

Orient each occurrence  $e$  of  $H'$  from  $u$  to  $v$ , writing its constraint as  $a_v = \rho_e(a_u)$ . Replace it by the path

$$u \text{ --- } p_e \text{ --- } q_e \text{ --- } r_e \text{ --- } v,$$

using three fresh interior vertices for *every occurrence*. Place all original vertices and all  $q_e$  on the left, and all  $p_e, r_e$  on the right. In the direction from  $u$  to  $v$ , put identity permutations on the first three edges and  $\rho_e$  on the last. When expressing an edge in left-to-right orientation, invert its permutation if needed. Give all  $4Q$  edges equal weight.

Fix labels at the original vertices. A satisfied occurrence extends to four satisfied edges. A violated occurrence cannot extend to four, but propagating labels along the first three edges satisfies exactly three. Since the interiors are disjoint, these extensions can be chosen independently. Maximizing over the original labels gives the exact identity

$$\text{val}(B) = 1 - \frac{1 - \text{val}(H')}{4}. \quad (7.6)$$

This proves Equation (7.1).

A loop becomes a four-cycle with three distinct new vertices. Parallel occurrences have disjoint interiors. Thus each left-right vertex pair carries at most one edge, so  $B$  is simple. An inverse binary

translation is the same translation, and an identity is translation by zero; the claimed constraint form is preserved.

Finally,  $Q \leq h/\tau + 1$  is polynomially bounded for fixed  $\tau$ . Integer division and exact comparison of rational remainders compute the  $n_e$  in polynomial time and bit complexity. Listing their  $Q$  copies and the  $3Q$  new vertices is therefore polynomial as well.  $\square$

## 8 Approximation consequences

Theorem 1.1 turns established conditional hardness results into NP-hardness results. We give the additional arguments needed for the exact instance models stated below. For a maximization problem, ratio  $r$  means returning value at least  $r$  times the optimum; for a minimization problem, factor  $c$  means returning value at most  $c$  times the optimum. All factors and tolerances below are fixed independently of input size. In each reduction using Theorem 1.1, they determine the Unique Games errors before its alphabet is chosen. If  $P \neq NP$ , the resulting hardness statements exclude deterministic polynomial-time approximations at the indicated strict thresholds.

### 8.1 Constraint satisfaction and Vertex Cover

For MAX-CUT, which maximizes the number of edges crossing a bipartition, Khot, Kindler, Mossel, and O’Donnell’s reduction gives NP-hardness of approximation at every fixed ratio in  $(\alpha_{\text{GW}}, 1)$ , where the Goemans–Williamson constant is  $\alpha_{\text{GW}} \simeq 0.87856$  [13, Theorem 1]. This matches the semidefinite approximation threshold of Goemans and Williamson [7]. The reduction uses the Majority Is Stablest theorem of Mossel, O’Donnell, and Oleszkiewicz [18]. For VERTEX COVER, which minimizes the number of vertices meeting every edge, Khot and Regev’s reduction gives NP-hardness at every fixed factor in  $(1, 2)$  [16]. To match the source convention that a right label determines the left label, use  $\pi_e^{-1}$ : for every pair of labels,  $a(v_e) = \pi_e(a(u_e))$  if and only if  $a(u_e) = \pi_e^{-1}(a(v_e))$ . Renumbering the fixed alphabet and assigning uniform rational edge weights likewise preserve each labeling’s value, so Theorem 1.1 supplies the edge-value premise of both reductions.

A constraint satisfaction problem (CSP) is specified here by a finite nonempty domain  $D$  and a finite family  $\Lambda$  of Boolean predicates on  $D$ , of bounded arity. A weighted  $\Lambda$ -instance has a finite variable set and a finite collection of predicates from  $\Lambda$  applied to tuples of variables, with nonnegative rational weights summing to one. An assignment maps the variables to  $D$ ; its value is the total satisfied weight, and OPT is the maximum of this value over all assignments. Raghavendra’s theorem characterizes the optimal approximation for every fixed language of this form, up to arbitrarily small positive error, by a basic semidefinite relaxation [22, Theorem 1.1 and Corollary 1.3]. The VERTEX COVER result uses its own reduction. Write  $\text{SDP}_R$  for the value of Raghavendra’s relaxation, as defined in [23, Section 4.5].

The fixed-instance form of Raghavendra’s theorem and the ordering theorem below use a strengthened bipartite formulation of Unique Games. In its YES case, one labeling satisfies every edge incident to at least a  $1 - \zeta$  fraction of left vertices; in its NO case, every labeling satisfies at most a  $\nu$  fraction of edges. These are uniform fractions of the left vertices and edges, respectively. Lemma A.1 gives this promise on an explicit nonempty simple graph whose common positive left degree can meet any fixed lower bound and be divisible by any fixed positive integer. It follows Khot and Regev’s normalization and sequence construction [16, Lemmas 3.3, 3.4, and 3.6], with an additional restriction that makes the graph simple. Its error parameters and degree are fixed before the alphabet is chosen, so Theorem 1.1 supplies the required premise.

**Corollary 8.1 (Raghavendra’s consequence).** *For a fixed  $\Lambda$ -instance  $J$  with  $\text{OPT}(J) < \text{SDP}_R(J)$  and a fixed tolerance*

$$0 < \eta < \frac{\text{SDP}_R(J) - \text{OPT}(J)}{2},$$

it is NP-hard to distinguish  $\Lambda$ -instances  $I$  satisfying

$$\text{OPT}(I) \geq \text{SDP}_R(J) - \eta \quad \text{from those satisfying} \quad \text{OPT}(I) \leq \text{OPT}(J) + \eta.$$

*Proof.* To match the thesis's common-arity convention [23, Definitions 7.2.1–7.2.2], identify  $D$  with a fixed  $[q]$  and add ignored trailing arguments to each predicate until all have the same positive arity. Fill the added positions with a fixed variable of  $J$ , which exists because  $\text{OPT}(J) < \text{SDP}_R(J)$ . Each applied payoff remains the same function of the original variables. In the relaxation of [23, Section 4.5], its dependence set indexes the local distribution; that set, the vector variables and the pairwise moment constraints are therefore unchanged. The same feasible solutions work in both directions with the same objective, preserving both  $\text{OPT}$  and  $\text{SDP}_R$ . On output, remove only the added trailing positions, keeping the original predicate from  $\Lambda$  and its weight.

Apply Raghavendra's fixed-instance theorem [23, Theorem 7.9] with tolerance  $\eta/2$ ; see also [22, Theorem 1.1 and Lemma 2.3]. The fixed witness and tolerance determine its strengthened Unique Games errors, so the formulation above and Theorem 1.1 supply its premise, with a fixed alphabet  $[M]$ . Its verifier instances have optimum greater than  $\text{SDP}_R(J) - \eta/2$  in the YES case and at most  $\text{OPT}(J) + \eta/2$  in the NO case.

We pass from the verifier's probability weights to the stated rational weights. For this fixed invocation, the construction in [23, Sections 7.3.1 and 7.5] gives a local probability law  $\theta$  on a finite set  $S$ , after the alphabet  $[M]$  is fixed. A record  $s \in S$  specifies an applied predicate of  $J$ , its  $a(s)$  ordered oracle-query strings in  $D^M$ , and the positions filled by their replies. This law may depend on  $J, \eta, M$ , but no varying input parameter enters it. Rational probability vectors are dense in this finite simplex, so fix  $\hat{\theta} \in \mathbb{Q}_{\geq 0}^S$  with

$$\sum_{s \in S} \hat{\theta}(s) = 1, \quad \delta := \frac{1}{2} \sum_{s \in S} |\hat{\theta}(s) - \theta(s)| < \frac{\eta}{2}.$$

This rational table is fixed finite data of the reduction; its description length may depend on  $J, \eta, M$ .

Let  $L_\Phi$  and  $V_\Phi$  be the left and right sides of the nonempty simple game supplied above, and let  $h \geq 1$  be its common left degree. List the variables  $V_\Phi \times D^M$  used in Section 7.5 of the cited thesis. Its outer verifier chooses a uniform left vertex  $u$  and, independently for each oracle query, chooses a uniform neighbor of  $u$  and relabels the query string by that edge's permutation. For each  $s \in S$ ,  $u \in L_\Phi$ , and ordered neighbor tuple in  $N(u)^{a(s)}$ , emit the  $\Lambda$ -predicate prescribed by  $s$  on the resulting query variables. Fill all positions of any syntactic variable omitted from the test with the first listed output variable; the applied predicate is independent of such variables, so this leaves its value unchanged. Give the occurrence weight

$$\frac{\hat{\theta}(s)}{|L_\Phi| h^{a(s)}}.$$

These weights are nonnegative rationals summing to one. Keeping equal occurrences separate gives  $|L_\Phi| \sum_{s \in S} h^{a(s)}$  occurrences and  $|V_\Phi| |D|^M$  variables. The local set and every  $a(s)$  are fixed, while the graph has polynomial size; hence this is a polynomial occurrence list. The explicit neighbor lists and edge permutations let us write it in polynomial time. Each weight has polynomial binary length: it can be represented with a fixed denominator multiplied by  $|L_\Phi| h^{a(s)}$ .

For an assignment  $\sigma$ , let  $g_\sigma(s) \in [0, 1]$  be its conditional acceptance probability over these outer choices. Write  $\text{val}_\theta(\sigma)$  and  $\text{val}_{\hat{\theta}}(\sigma)$  for its satisfied weight under the original and rationalized local laws, respectively. The instances use the same variable set, and

$$|\text{val}_{\hat{\theta}}(\sigma) - \text{val}_\theta(\sigma)| = \left| \sum_{s \in S} (\hat{\theta}(s) - \theta(s)) g_\sigma(s) \right| \leq \delta < \frac{\eta}{2}.$$

Taking maxima gives the same bound for their optima. The imported gap therefore yields optimum at least  $\text{SDP}_R(J) - \eta$  in the YES case and at most  $\text{OPT}(J) + \eta$  in the NO case, as required.  $\square$

## 8.2 Ordering constraints

An ordering CSP asks for a global order of its input variables; the number of variables may grow with the input. Fix an integer  $k \geq 2$  and a nonempty proper subset  $\Pi \subset S_k$ , where  $S_k$  is the set of permutations of  $\{1, \dots, k\}$ . In the finite probability-weighted model of Guruswami, Håstad, Manokaran, Raghavendra, and Charikar [8, Definition 8.2], we use an explicitly listed finite variable set  $V$  and a finite list of constraint occurrences. Their nonnegative rational weights are encoded by binary numerators and denominators and sum to one. Every occurrence is an ordered  $k$ -tuple  $(v_1, \dots, v_k)$  of distinct variables, satisfied by a linear order  $\prec$  when

$$v_{\pi(1)} \prec \dots \prec v_{\pi(k)} \quad \text{for some } \pi \in \Pi.$$

Write  $\text{OPT}_{\Pi}(I)$  for the maximum satisfied weight. A uniform random order satisfies expected weight  $\rho_{\Pi} = |\Pi|/k!$ . Since  $\text{OPT}_{\Pi}(I) \leq 1$ , it is a  $\rho_{\Pi}$ -approximation. For fixed  $k$ , conditional expectation derandomizes this guarantee. Build an order by exposing one next variable at a time. If exactly  $s$  variables of a constraint have been exposed, its conditional success probability is the number of permutations in  $\Pi$  whose first  $s$  positions are the roles of those variables in the exposed order, divided by  $(k-s)!$ . It can be computed by inspecting the fixed set  $S_k$ . For each choice of the next variable, sum these probabilities against the listed rational weights and choose a next variable whose conditional expectation is at least the average over the choices. This preserves the initial expectation  $|\Pi|/k!$ . All comparisons are exact and polynomial in the input bit length: the product of the weight denominators has bit length at most their total encoded length, and the factorial denominators have constant size for fixed  $k$ . The variable list is explicit, so the number of prefix steps and candidate choices is polynomial as well.

The size of  $V$  is part of the input, so this is not a CSP over the fixed finite domain used in the preceding corollary.

**Corollary 8.2 (Ordering-CSP consequence).** *For every fixed integer  $k \geq 2$ , nonempty proper subset  $\Pi \subset S_k$ , and  $\varepsilon \in (0, (1 - \rho_{\Pi})/2)$ , where  $\rho_{\Pi} = |\Pi|/k!$ , it is NP-hard to distinguish the finite rationally weighted  $\Pi$ -ordering instances of Section 8 satisfying*

$$\text{OPT}_{\Pi}(I) \geq 1 - \varepsilon \quad \text{from those satisfying} \quad \text{OPT}_{\Pi}(I) \leq \rho_{\Pi} + \varepsilon.$$

*Approximation at any fixed ratio  $r \in (\rho_{\Pi}, 1]$  is therefore NP-hard.*

*Proof.* Proposition A.2 constructs the stated finite instance from a simple bipartite Unique Games instance with the promise in Lemma A.1. It uses the local gap and analytic estimates of Guruswami et al. [8] and fixes the required game errors and degree conditions before the alphabet is chosen. Lemma A.1 supplies that game from Theorem 1.1; composing the reductions proves the displayed gap for the original predicate  $\Pi$ .

For a fixed  $r \in (\rho_{\Pi}, 1]$ , choose the positive error still smaller so that  $\varepsilon < (r - \rho_{\Pi})/(1 + r)$ . Then  $r(1 - \varepsilon) > \rho_{\Pi} + \varepsilon$ , so an  $r$ -approximation would distinguish the two cases.  $\square$

Together with the random-order guarantee, Corollary 8.2 gives the optimal approximation threshold  $\rho_{\Pi}$ .

For MAXIMUM ACYCLIC SUBGRAPH in the finite weighted directed-graph model, take  $k = 2$  and  $\Pi = \{12\}$  in one-line notation. A constraint  $(u, v)$  asks that the edge point forward; forward edges form an acyclic subgraph, and every acyclic subgraph has a topological order. Its threshold is therefore  $1/2$ . For BETWEENNESS, the constraint  $(a, b, c)$  asks that  $b$  lie between  $a$  and  $c$ , so  $\Pi = \{123, 321\} \subset S_3$  and the threshold is  $2/3! = 1/3$  [8, Theorem 1.1 and Section 1.2]. These finite weighted consequences have near-perfect completeness with fixed positive error.

### 8.3 Cut, deletion, and clustering consequences

For the cut problems of Chawla, Krauthgamer, Kumar, Rabani, and Sivakumar [3, Section 1], let  $G = (V, E)$  be a finite undirected graph with positive rational edge costs  $c_e$ , and let  $\mathcal{D} = \{\{s_i, t_i\} : 1 \leq i \leq k\}$  contain  $k \geq 1$  distinct demand pairs with  $s_i \neq t_i$ . For  $M \subseteq E$ , write  $c(M) = \sum_{e \in M} c_e$ , and let  $q_{\mathcal{D}}(M)$  count the demand pairs disconnected in  $G - M$ . Weighted MULTICUT minimizes  $c(M)$  subject to  $q_{\mathcal{D}}(M) = k$ . Their nonuniform SPARSEST CUT search problem minimizes  $c(M)/q_{\mathcal{D}}(M)$  over cutsets  $M$  with  $q_{\mathcal{D}}(M) > 0$ . Only the selected pairs carry unit demand; they need not be all vertex pairs.

For weighted MIN-2CNF $\equiv$  DELETION, the input consists of clauses ( $\ell \equiv \ell'$ ) with positive rational weights, where  $\ell, \ell'$  are Boolean literals, and an assignment pays the total weight of false clauses. The allowed clauses therefore express equality or disequality of variables.

In the arbitrary weighted-graph form of CORRELATION CLUSTERING, each present edge has a  $+$  or  $-$  label and a nonnegative rational weight. A partition into any number of clusters pays for  $+$  edges whose endpoints lie in different clusters and  $-$  edges whose endpoints lie in the same cluster. Missing edges carry no penalty. This is the minimum-disagreements objective of [4, Section 2.1].

**Corollary 8.3 (Cut, deletion, and clustering hardness).** *For every fixed real  $C > 1$ , approximation within factor  $C$  is NP-hard for each of the four minimization problems just defined: weighted MULTICUT, nonuniform SPARSEST CUT, weighted MIN-2CNF $\equiv$  DELETION, and minimum-disagreements CORRELATION CLUSTERING. For SPARSEST CUT, the hardness is under polynomial-time Cook reductions and concerns algorithms that output a cutset.*

*Proof.* Corollary 1.3 of [3] gives the first three conclusions under its ordinary edge-value Unique Games hypothesis. For a requested factor  $C$ , use its hardness statement at any fixed larger factor  $L$ : a  $C$ -approximation would also be an  $L$ -approximation. To supply that hypothesis, orient the bipartite constraints of Theorem 1.1 left-to-right, renumber its fixed alphabet as  $[d]$  with  $d = 2^s \geq 2$ , and give its edges equal weights summing to one. Padding the smaller vertex side with isolated questions if needed preserves both value and polynomial size. The search-version qualification for SPARSEST CUT is in Section 1.3 of the cited source.

For CORRELATION CLUSTERING, the hard MULTICUT instances in [3, Section 2.1] have distinct antipodal demand pairs inside  $d$ -cubes, while the edges of the MULTICUT graph are cube edges or join different cubes. Here  $d \geq 2$ , unchanged by Lemma 1.5 of that source, so the demand pairs are nonedges. The transformation of Demaine, Emanuel, Fiat, and Immorlica [4, Theorem 4.7 and Corollary 4.8] therefore creates no opposite-signed parallel edges. Concretely, make every original edge positive and put a negative edge of weight  $W_+ + 1$  on each demand pair, where  $W_+$  is the total original edge cost. A clustering with a negative mistake can be improved by separating one endpoint: it removes at least  $W_+ + 1$  of negative cost and adds at most  $W_+$  of positive cost. Repeating this operation leaves a clustering separating all demands, whose positive mistakes form a multicut of no greater cost. Conversely, the connected components after deleting a multicut separate all demands and have positive disagreement cost at most that multicut's cost. These two inequalities give equality of the optima and the required solution conversion. The weights remain rational with polynomial bit length, so the same fixed-factor hardness transfers.  $\square$

The separate manuscript [19, Theorem 1.1] proves NP-hardness at every fixed factor  $C > 1$  for the bipartition objective with nonnegative rational capacities and exactly unit demand between every pair of distinct vertices. This uniform-demand theorem does not depend on Theorem 1.1 and is not supplied by the nonuniform reduction above.

### 8.4 Kernel clustering

For kernel clustering with target matrix  $I_k$ , fix an integer  $k \geq 3$ , where  $I_k$  is the  $k$ -by- $k$  identity matrix. An input is an explicitly represented rational symmetric matrix  $A = (a_{pq})_{p,q=1}^N \succeq 0$  with  $A\mathbf{1} = 0$ , where

$\mathbf{1}$  is the all-ones vector. An assignment  $\sigma : \{1, \dots, N\} \rightarrow \{1, \dots, k\}$  has value

$$\text{val}_A(\sigma) = \sum_{p,q=1}^N a_{pq} \mathbf{1}_{\{\sigma(p)=\sigma(q)\}}, \quad \text{OPT}_k(A) = \max_{\sigma} \text{val}_A(\sigma).$$

The sum includes ordered pairs and the diagonal; empty clusters are allowed, and there is no normalization by cluster size. A loss factor  $\alpha \geq 1$  means returning value at least  $\text{OPT}_k(A)/\alpha$ . The Gaussian propeller companion uses Theorem 1.1 to prove NP-hardness of achieving every fixed loss factor

$$1 \leq \alpha < \alpha_k, \quad \alpha_k = \frac{8\pi}{9} \left(1 - \frac{1}{k}\right),$$

for each fixed  $k \geq 3$  [21, Theorem 6.1]. This is strict hardness below the rounding factor  $\alpha_k$  stated in that companion.

## A Finite formulations of the consequences

Some of the reductions in Section 8 use a bipartite Unique Games instance in which a nearly satisfying labeling satisfies all constraints incident to most left vertices. We first obtain a simple graph with this property and with a prescribed lower bound on its common left degree. We then use it to give the finite rational formulation of the ordering consequence. The local ordering construction follows Guruswami, Håstad, Manokaran, Raghavendra, and Charikar [8] and uses the low-influence results cited below. The composition specifies the graph, the independent neighbor choices, and the conversion to constraints on distinct variables.

### A.1 Strong satisfaction of left vertices in a simple graph

Orient the constraints of a bipartite game from its left side to its right side. A left vertex is *strongly satisfied* by a labeling when every edge incident to it is satisfied.

**Lemma A.1 (Strong satisfaction on the left).** *For every  $\zeta, \nu \in (0, 1)$  and integers  $d, H \geq 1$ , there are integers  $h, R \geq 1$  with  $h \geq H$  and  $d \mid h$ , and a deterministic polynomial-time reduction from 3SAT to nonempty explicit unweighted simple bipartite Unique Games instances  $\Phi_{\varphi} = (A, B, E)$  over  $[R]$ , such that every left vertex has degree  $h$  and*

$$\begin{aligned} \varphi \text{ satisfiable} &\implies \text{some labeling strongly satisfies at least } (1 - \zeta)|A| \text{ left vertices,} \\ \varphi \text{ unsatisfiable} &\implies \text{val}(\Phi_{\varphi}) < \nu. \end{aligned}$$

*The degree  $h$  and the two errors for Theorem 1.1 can be chosen from  $\zeta, \nu, d, H$  before that theorem supplies the alphabet size  $R$ .*

*Proof.* We amplify local completeness by making a left question an ordered sequence of incident constraints. A nearly satisfying label satisfies all constraints in most such sequences, while low edge value makes repeated demands for the same left label unlikely. We adapt the normalization and sequence construction of Khot and Regev [16, Lemmas 3.3, 3.4, and 3.6]; independent right-vertex tags then make the retained graph simple.

Choose a multiple  $h$  of  $d$  with

$$h \geq \max\{H, 2\}, \quad \frac{1}{h} < \frac{\nu}{4}.$$

For  $(T)_h = T(T-1) \cdots (T-h+1)$ , choose a fixed integer  $T \geq h$  such that

$$p_{\text{graph}} := \frac{(T)_h}{T^h} > \frac{3}{4}.$$

Such a choice exists because  $1 - (T)_h/T^h \leq \binom{h}{2}/T$ . Choose positive rationals  $\beta, \gamma, u$  and a positive integer  $L_2$  so that

$$\beta < \frac{\nu}{4}, \quad h^2\gamma < \frac{\nu}{4}, \quad (h+1)u < \frac{\zeta}{2}, \quad \frac{1}{L_2} < \min \left\{ \frac{\beta\gamma}{2}, \frac{\zeta}{2h} \right\}.$$

Finally choose positive rational  $\gamma_0, \zeta_0 < 1/2$  with

$$2\gamma_0 < \frac{\beta\gamma}{2}, \quad \sqrt{2\zeta_0} < u.$$

All these choices precede the invocation of Theorem 1.1 with completeness error  $\zeta_0$  and soundness error  $\gamma_0$ .

The construction below first normalizes the left weights and rounds them to obtain (A.1), an occurrence game of NO value below  $\beta\gamma$ . Copying its right vertices preserves this bound by (A.2). The length- $h$  sequence graph then has NO value at most  $s = \beta + 1/h + (1 - \beta)h^2\gamma < 3\nu/4$ , as proved in (A.3). Keeping only sequences with distinct tags makes the graph simple and retains a fraction  $p_{\text{graph}} > 3/4$ , so its NO value is below  $\nu$ . The same conditioning leaves the base occurrences independent in the YES case; the final strong-satisfaction bound is  $1 - (h+1)u - h/L_2 > 1 - \zeta$ . We now construct these graphs and prove each bound.

Write its output as  $\Phi_0 = (X_0, Y, E_0)$ , orienting each permutation left-to-right and renumbering the fixed alphabet as  $[R]$ . Put  $M_0 = |E_0| > 0$ , put  $w_{xy} = 1/M_0$  when  $xy \in E_0$  and  $w_{xy} = 0$  otherwise, and let  $w_x = \sum_y w_{xy}$  be the total incident weight at  $x$ . Simplicity gives one permutation at each pair of positive weight. Create

$$n_x = \lfloor 2|X_0|w_x \rfloor$$

copies of  $x$ , each with local edge weights  $w_{xy}/w_x$ . Only vertices with  $w_x > 0$  have copies. If  $X_1$  is the copy set, then

$$|X_0| \leq |X_1| = \sum_x n_x \leq 2|X_0|,$$

and the incident weight at each copy is one.

Here  $w_{\text{sat}}(x)$  denotes the sum of satisfied local weights at a copy  $x' \in X_1$  under the given labeling. For a labeling  $L$  of  $\Phi_0$ , write  $w_{\text{sat},L}(x)$  for the satisfied weight incident to  $x$  and  $w_{\text{sat},L}(\Phi_0)$  for the total satisfied weight. Each local quantity uses the labels at that left vertex and at the shared right vertices. Given any labeling of the copy game, keep the right labels and give all copies of each  $x$  a label that maximizes their common local objective. The corresponding labeling  $L$  of  $\Phi_0$  satisfies

$$\sum_{x' \in X_1} w_{\text{sat}}(x') \leq \sum_{x: n_x > 0} \frac{n_x}{w_x} w_{\text{sat},L}(x) \leq 2|X_0| w_{\text{sat},L}(\Phi_0).$$

Consequently the copy game's normalized value is at most  $2\gamma_0$  on a NO input. On a YES input, lift a labeling of  $\Phi_0$  with unsatisfied weight at most  $\zeta_0$ . The same calculation for unsatisfied weight gives mean local loss at most  $2\zeta_0 < u^2$ . Markov's inequality leaves more than a  $1 - u$  fraction of  $X_1$  with local satisfied weight at least  $1 - u$ .

To remove the weights, put  $\alpha = L_2|Y|$ . At each  $x \in X_1$ , write  $\tilde{w}_{xy}$  for its normalized local weights. Choose one neighbor  $y_0$  of positive local weight, round  $\alpha\tilde{w}_{xy}$  down for every  $y \neq y_0$ , and assign the remaining occurrences to  $y_0$ . All occurrences at a pair retain that pair's permutation. This gives an unweighted multigraph  $\Phi_{\text{occ}}$  of left degree  $\alpha$ . For any subset of the neighbors, the difference between its old weight and its new occurrence fraction has magnitude at most  $|Y|/\alpha = 1/L_2$ : sum the rounding losses over the subset if it omits  $y_0$ , and over its complement otherwise. This applies to the satisfied subset for every labeling. Hence

$$\text{val}(\Phi_{\text{occ}}) \leq 2\gamma_0 + \frac{1}{L_2} < \beta\gamma \quad \text{on a NO input.} \quad (\text{A.1})$$

On a YES input, the labeling above has satisfied occurrence fraction at least  $1 - a$ , where  $a = u + 1/L_2$ , at more than a  $1 - u$  fraction of  $X_1$ .

Index the  $\alpha$  occurrences at  $x$  by  $j$ , with right endpoint  $y_j$  and permutation  $\pi_j$ . Replace each  $y \in Y$  by  $(y, t)$ ,  $t \in [T]$ , and replace occurrence  $j$  by all  $T$  occurrences from  $x$  to  $(y_j, t)$  with permutation  $\pi_j$ . Call the resulting left- $\alpha T$ -regular multigraph  $\Phi_{\text{tag}}$ . For any labeling of  $\Phi_{\text{tag}}$ , keep its left labels and, independently for each original  $y$ , give  $y$  the label of a uniformly chosen copy  $(y, t)$ . The expected edge value in  $\Phi_{\text{occ}}$ , counting its occurrences, is exactly the given edge value in  $\Phi_{\text{tag}}$ . Thus

$$\text{val}(\Phi_{\text{tag}}) \leq \text{val}(\Phi_{\text{occ}}) < \beta\gamma \quad \text{on a NO input.} \quad (\text{A.2})$$

The tag-independent lift of the YES labeling preserves every local satisfied fraction.

Consider the full sequence graph whose left vertices are

$$(x, (j_1, t_1), \dots, (j_h, t_h)), \quad j_i \in [\alpha], \quad t_i \in [T],$$

with position  $i$  joined to  $(y_{j_i}, t_i)$  by  $\pi_{j_i}$ . This is the sequence construction of Khot and Regev; its left vertex identifier includes the entire ordered occurrence sequence. We record the soundness estimate in a form that permits the left label to depend on that entire sequence. Each  $x$  contributes exactly  $(\alpha T)^h$  full sequence vertices, all of degree  $h$ . Thus the uniform edge value is the average satisfied fraction over a uniform  $x$ , an ordered  $h$ -tuple of independent uniform occurrences at  $x$ , and a uniform position in that tuple.

Fix arbitrary right labels in  $\Phi_{\text{tag}}$ . Each occurrence at  $x$  determines the unique left label that satisfies it. Let  $p_x(\ell)$  be the fraction determining label  $\ell$ . The  $p_x(\ell)$  sum to one, and choosing the best left label at each  $x$  gives

$$\mathbb{E}_x \max_{\ell} p_x(\ell) \leq \text{val}(\Phi_{\text{tag}}) < \beta\gamma.$$

Fewer than a  $\beta$  fraction of  $x$  therefore have  $\max_{\ell} p_x(\ell) \geq \gamma$ . At every other  $x$ , the probability that two labels determined by  $h$  independent occurrences agree is at most

$$\binom{h}{2} \sum_{\ell} p_x(\ell)^2 \leq \binom{h}{2} \max_{\ell} p_x(\ell) < h^2 \gamma.$$

If no two determined labels agree, any label of the sequence vertex satisfies at most one of its  $h$  edges. On the remaining sequences its satisfied fraction is at most one. Averaging first over sequences and then over  $x$ , and using  $1/h + h^2 \gamma < 1$ , bounds the full sequence graph's value by

$$s := \beta + \frac{1}{h} + (1 - \beta)h^2 \gamma < \frac{3\nu}{4}. \quad (\text{A.3})$$

The estimate holds for every choice of the right labels and every sequence-dependent choice of left labels.

Retain only sequence vertices for which  $t_1, \dots, t_h$  are pairwise distinct. For each  $x$  and each base sequence  $(j_1, \dots, j_h)$ , exactly  $(T)_h$  tag sequences remain. Each  $x$  thus contributes  $\alpha^h (T)_h$  retained vertices, a fraction  $p_{\text{graph}}$  of its full set. The  $h$  right neighbors of each retained vertex are distinct because their tags are distinct. Different occurrence sequences remain different left vertices, so this retained graph is simple and has common left degree  $h$ .

Give each sequence vertex based at  $x$  the old label of  $x$ , and keep the tag-independent right labels. For this lifted YES labeling, satisfaction at position  $i$  depends only on  $j_i$ . Conditioning the tags to be distinct leaves the base occurrence sequence independent and uniform. If the satisfied fraction at  $x$  in  $\Phi_{\text{occ}}$  is  $f_x$ , exactly an  $f_x^h$  fraction of its retained sequence vertices are strongly satisfied. Since every  $x$  contributes equally, the strongly satisfied fraction is at least

$$(1 - u)(1 - a)^h \geq 1 - u - ha = 1 - (h + 1)u - \frac{h}{L_2} > 1 - \zeta.$$

Here  $0 < a < 1$  follows from the parameter choices.

For NO, extend any labeling of the retained graph to the full graph by giving discarded left vertices arbitrary labels and keeping all right labels. All left degrees are  $h$ , so its full edge value is at least  $p_{\text{graph}}$  times its retained edge value. By (A.3), the latter is at most

$$\frac{s}{p_{\text{graph}}} < \frac{3\nu/4}{3/4} = \nu.$$

This extension permits arbitrary dependence of the retained left labels on their tags.

All the operations are finite and deterministic. If  $N$  bounds the explicit size of  $\Phi_0$ , then  $|X_1| = O(N)$  and  $\alpha = L_2|Y| = O(N)$ . The retained graph has  $|X_1|\alpha^h(T)_h = O(N^{h+1})$  left vertices and  $h$  edges per left vertex. Its identifiers and the rational arithmetic used for rounding have polynomial bit length, and its permutation tables use the fixed alphabet inherited from  $\Phi_0$ . This proves the claimed polynomial bound and completes the reduction.  $\square$

## A.2 The finite ordering transfer

The ordering construction uses two inputs from Guruswami et al. The first is a finite instance on which every assignment of  $q$  ranks has value near the random-order value, together with probability laws on its constraint tuples whose expected payoff is near one. The second bounds the expected payoff of functions with small coordinate influences by that  $q$ -rank optimum, up to a vanishing error. Lemmas A.3 and A.5 give the precise forms used here. We use these local results to construct rationally weighted ordering constraints from a Unique Games instance, choosing the required game errors before its alphabet.

**Proposition A.2 (Finite ordering transfer).** *Fix an integer  $k \geq 2$ , a nonempty proper subset  $\Pi \subset S_k$ , and  $\delta \in (0, (1 - \rho_\Pi)/2)$ , where  $\rho_\Pi = |\Pi|/k!$ . There are integers  $q, H \geq 1$  and positive rationals  $\zeta, \nu < 1$ , depending only on  $k, \Pi, \delta$ , such that the following holds for every fixed integer  $R \geq 1$ .*

*Let  $\Phi = (A, B, E)$  be a nonempty explicit unweighted simple bipartite Unique Games instance over  $[R]$ , oriented from  $A$  to  $B$ , with common left degree  $h \geq H$  divisible by  $q$ . A deterministic transformation, polynomial in the explicit size of  $\Phi$  for fixed  $R$ , produces a finite  $\Pi$ -ordering instance  $I$  with an explicit variable list and an explicit constraint-occurrence list. Each occurrence uses distinct variables and has a nonnegative rational weight encoded by a binary numerator and positive binary denominator; the weights sum to one. The transformation satisfies*

$$\begin{aligned} \text{some labeling strongly satisfies at least } (1 - \zeta)|A| \text{ left vertices} &\implies \text{OPT}_\Pi(I) \geq 1 - \delta, \\ \text{val}(\Phi) \leq \nu &\implies \text{OPT}_\Pi(I) \leq \rho_\Pi + \delta. \end{aligned}$$

*The choices  $q, H, \zeta, \nu$  are independent of  $R$  and  $h$ .*

The output variables are  $(b, z) \in B \times [m]^R$ , with a fixed symbol set  $[m]$  chosen below. At each left vertex, coarsen the proposed order on its neighbors' output variables into  $q$  equal consecutive blocks. For a fixed query string, the fractions of neighbors landing in these blocks form a probability vector; each component has uniform mean  $1/q$ . Independent neighbor choices then express the expected coarse payoff through the single function used in Lemma A.5. Lemma A.4 bounds the probability that two selected variables share a block, where coarsening could change their order.

If a smoothed, neighbor-averaged interval indicator has a large coordinate influence, that coordinate supplies a left game label. The influence inequality gives a positive fraction of neighbors with the matching coordinate in a bounded list, from which right labels are sampled. Many such vertices would contradict low edge value. Conditioning on distinct neighbors produces legal constraint tuples, and breaking ties within the YES rank classes produces a strict order.

Choose  $\pi_0 \in \Pi$  and set

$$P_0(\sigma) = \mathbf{1}_\Pi(\pi_0 \circ \sigma), \quad \sigma \in S_k.$$

Then  $P_0(\text{id}) = 1$ , and its average is  $\rho_\Pi$ , since  $\sigma \mapsto \pi_0 \circ \sigma$  is a bijection of  $S_k$ . We use this normalization only together with the tuple relabeling below; it must preserve the payoff also when a coarse rank map

has ties. For a vector of integer ranks  $r = (r_1, \dots, r_k)$ , let  $\text{Ref}(r)$  be the permutations compatible with their weak order: they list positions in increasing rank and may order positions of equal rank arbitrarily. Define

$$\bar{P}_0(r) = \frac{1}{|\text{Ref}(r)|} \sum_{\sigma \in \text{Ref}(r)} P_0(\sigma). \quad (\text{A.4})$$

This is the uniform tie extension used in [8, Section 8.2]. It lies in  $[0, 1]$  and agrees with  $P_0$  when the ranks are distinct.

The normalization preserves the predicate after one permutation of tuple positions. Given a raw tuple  $w = (w_1, \dots, w_k)$ , define the emitted tuple  $v$  by

$$v_{\pi_0(j)} = w_j \quad (j \in [k]). \quad (\text{A.5})$$

If a strict order induces  $\sigma$  on  $w$ , it induces  $\pi_0 \circ \sigma$  on  $v$ . More generally, relabeling positions gives the bijection

$$\text{Ref}(f(v_1), \dots, f(v_k)) = \{\pi_0 \circ \sigma : \sigma \in \text{Ref}(f(w_1), \dots, f(w_k))\}$$

for every integer-valued rank map  $f$  on the tuple variables. Uniform averages over these sets therefore give equal payoffs for the original predicate  $\Pi$  and for  $P_0$ , including ties. The permutation of positions in (A.5) preserves distinctness and weights.

For a finite local instance  $J = (U, \mathcal{C}, \omega)$ , write  $t_c = (u_{c,1}, \dots, u_{c,k})$  for the tuple of occurrence  $c \in \mathcal{C}$  and  $\omega_c$  for its probability weight. Its  $q$ -ordering value is

$$\text{opt}_q(J) = \max_{g: U \rightarrow [q]} \sum_{c \in \mathcal{C}} \omega_c \bar{P}_0(g(u_{c,1}), \dots, g(u_{c,k})).$$

**Lemma A.3 (Rational local ordering data).** *For every integer  $q \geq 1$  and rational  $\eta > 0$ , there is a finite local instance  $J = (U, \mathcal{C}, \omega)$ , with  $|U| = m \geq 2$ , whose occurrence weights are rational and whose tuples each use  $k$  distinct variables, together with rational probability laws  $\mu_c$  on  $[m]^k$ , such that*

$$\text{opt}_q(J) \leq \rho_\Pi + \eta, \quad \sum_c \omega_c \mathbb{E}_{X \sim \mu_c} \bar{P}_0(X) \geq 1 - \eta. \quad (\text{A.6})$$

*Every one-position marginal of  $\mu_c$  is uniform on  $[m]$ , and a common  $\alpha \in (0, 1/2]$  satisfies  $\mu_c(x) > \alpha$  for every  $c \in \mathcal{C}$  and  $x \in [m]^k$ .*

*These specified laws admit common feasible Gram data for the relaxation in [8, Section 8.3]: a unit vector  $I$  and vectors  $b_{u,j}$ ,  $u \in U$ ,  $j \in [m]$ , satisfying*

$$\sum_j b_{u,j} = I, \quad \langle b_{u,i}, b_{u,j} \rangle = 0 \quad (i \neq j), \quad \|b_{u,j}\|^2 = 1/m,$$

*and, for every occurrence  $c$  and positions  $r, s \in [k]$ ,*

$$\langle b_{u_{c,r},i}, b_{u_{c,s},j} \rangle = \Pr_{X \sim \mu_c} [X_r = i, X_s = j] \quad (i, j \in [m]). \quad (\text{A.7})$$

*All the data depend only on  $k, \Pi, q, \eta$ .*

Its complete proof is in Section A.3.

For real scalar functions on the uniform product  $[m]^R$ , we use the coordinate influence and resampling noise conventions of [11, Definitions 2.4–2.5]:

$$\text{Inf}_\ell(F) = \mathbb{E}[\text{Var}_{z_\ell}(F \mid z_{-\ell})].$$

The noise operator  $T_\theta$ ,  $0 \leq \theta \leq 1$ , independently keeps each coordinate with probability  $\theta$  and otherwise resamples it uniformly, then averages the function. Choose a real orthonormal basis  $\chi_0 \equiv 1, \chi_1, \dots, \chi_{m-1}$

on uniform  $[m]$ . For  $\sigma \in \{0, \dots, m-1\}^R$ , put  $\chi_\sigma(z) = \prod_{j=1}^R \chi_{\sigma_j}(z_j)$ ,  $|\sigma| = |\{j : \sigma_j \neq 0\}|$ , and  $\hat{F}(\sigma) = \mathbb{E}_z F(z) \chi_\sigma(z)$ . In this normalization the formulas are

$$T_\theta F = \sum_{\sigma} \theta^{|\sigma|} \hat{F}(\sigma) \chi_\sigma, \quad \text{Inf}_\ell(F) = \sum_{\sigma: \sigma_\ell \neq 0} \hat{F}(\sigma)^2. \quad (\text{A.8})$$

In particular, further smoothing decreases every influence.

**Lemma A.4 (Uniform collision estimate).** *For every  $\kappa \in (0, 1/2)$ , there is  $\mu_0(\kappa) > 0$  such that, for each integer  $m \geq 2$ , there is a function  $e_{\text{col}}^{m, \kappa} : (0, \infty) \rightarrow [0, \infty)$  tending to zero at zero with the following property. For every integer  $R \geq 1$  and  $t > 0$ , let  $F, G : [m]^R \rightarrow [0, 1]$  have the same mean  $0 < \mu \leq \mu_0(\kappa)$  and satisfy*

$$\text{Inf}_\ell(T_{1-\kappa} F) \leq t, \quad \text{Inf}_\ell(T_{1-\kappa} G) \leq t \quad (\ell \in [R]).$$

*For every coupling  $(Z, Z')$  with both marginals uniform on  $[m]^R$ ,*

$$\mathbb{E}[(T_{1-2\kappa} F)(Z)(T_{1-2\kappa} G)(Z')] \leq \mu^{1+\kappa/2} + e_{\text{col}}^{m, \kappa}(t). \quad (\text{A.9})$$

*The error function is independent of  $R, \mu, F, G$  and the coupling.*

Its complete proof is in Section A.3.

Write  $\mathcal{S}_q = \{p \in [0, 1]^q : \sum_j p_j = 1\}$ . For  $p_i \in \mathcal{S}_q$ , denote its  $j$ th coordinate by  $p_i(j)$  and define the multilinear extension of  $\bar{P}_0$  on  $[q]^k$  by

$$P_0^{\text{ml}}(p_1, \dots, p_k) = \sum_{r \in [q]^k} \bar{P}_0(r) \prod_{i=1}^k p_i(r_i), \quad p_i \in \mathcal{S}_q. \quad (\text{A.10})$$

**Lemma A.5 (A simplex-valued local estimate).** *Fix the local data  $J, \mu_c, \alpha$  of Lemma A.3 and  $\lambda \in (0, 1)$ . There is a function  $e_{\text{inv}} : (0, 1) \rightarrow [0, \infty)$ , depending only on these data and  $k, q, \lambda$ , with  $e_{\text{inv}}(u) \rightarrow 0$  as  $u \downarrow 0$ , for which the following holds for every integer  $R \geq 1$ .*

*Choose one occurrence  $c$  with probability  $\omega_c$ , and then choose  $R$  independent tuples from its specified law  $\mu_c$ , giving strings  $Z_1, \dots, Z_k \in [m]^R$ . Independently at every position and coordinate, resample a uniform symbol with probability  $\lambda$  to obtain  $\tilde{Z}_i$ . Let  $F = (F^1, \dots, F^q) : [m]^R \rightarrow \mathcal{S}_q$  be a single function used for every  $c$  and every position. If  $0 < t < 1/(kq)$  and*

$$\text{Inf}_\ell(T_{1-\lambda} F^j) \leq t \quad (\ell \in [R], j \in [q]),$$

*then*

$$\mathbb{E}_{c, Z, \tilde{Z}} P_0^{\text{ml}}(F(\tilde{Z}_1), \dots, F(\tilde{Z}_k)) \leq \text{opt}_q(J) + e_{\text{inv}}(kqt). \quad (\text{A.11})$$

Its complete proof is in Section A.3.

*Proof of Proposition A.2.* We sample neighbors independently and use the local laws from Lemma A.3 to correlate their query strings. Lemmas A.4 and A.5 will control the loss from coarse ranks; the high-influence alternative will give a game labeling. All parameters for these estimates are chosen independently of  $R$ . Begin with  $0 < \kappa < 1/2$ ,  $q$  with  $1/q \leq \mu_0(\kappa)$ , and local data from Lemma A.3. Use the local experiment of Lemma A.5 with resampling probability  $\lambda = 2\kappa$ . Also fix  $0 < t < 1/(kq)$  and  $\chi > 0$ , and abbreviate  $e_{\text{col}} = e_{\text{col}}^{m, \kappa}$ .

For an input graph  $\Phi$ , let  $\pi_{a \rightarrow b}$  denote the permutation on edge  $ab$ . Its action on strings is fixed by

$$(\pi_{a \rightarrow b} z)_{\pi_{a \rightarrow b}(\ell)} = z_\ell \quad (\ell \in [R]).$$

The output variable set will be  $\mathcal{V} = B \times [m]^R$ . Consider the following independent neighbor experiment: choose  $a$  uniformly from  $A$ , choose  $b_1, \dots, b_k$  independently and uniformly from  $N(a)$ , and independently perform the local experiment, obtaining  $\tilde{Z}_1, \dots, \tilde{Z}_k$ . Its raw output positions are

$$w_i = (b_i, \pi_{a \rightarrow b_i} \tilde{Z}_i) \in \mathcal{V}. \quad (\text{A.12})$$

For any rank map  $f : \mathcal{V} \rightarrow \mathbb{Z}$ , define

$$V(f) = \mathbb{E}\bar{P}_0(f(w_1), \dots, f(w_k)), \quad V_a(f) = \mathbb{E}[\bar{P}_0(f(w_1), \dots, f(w_k)) \mid a].$$

The bounded payoff (A.4) is used even if the experiment repeats an output ID. The emitted instance conditions this experiment on distinct neighbors and permutes its tuple positions by (A.5). Its complete enumeration is given at the end of the proof. Put  $\delta_h = 1 - (h)_k/h^k \leq \binom{k}{2}/h$  for  $h \geq k$ , and define

$$S_{\text{loc}} = \text{opt}_q(J) + \binom{k}{2}(q^{-\kappa/2} + qe_{\text{col}}^{m,\kappa}(t)) + e_{\text{inv}}(kqt).$$

The argument below proves the two estimates collected in (A.23): a labeling strongly satisfying a  $1 - \zeta$  fraction of left vertices gives

$$\text{OPT}_{\Pi}(I) \geq (1 - \zeta)(1 - 2\kappa)^k \sum_c \omega_c \mathbb{E}_{X \sim \mu_c} \bar{P}_0(X) - \delta_h,$$

whereas  $\text{val}(\Phi) \leq \nu < \chi \kappa t^4/6400$  gives

$$\text{OPT}_{\Pi}(I) \leq S_{\text{loc}} + \chi + \delta_h.$$

The first estimate uses the coordinate read by a satisfying label. For the second, Lemmas A.4 and A.5 control the payoff at vertices with small influences, and a large influence yields a game label; (A.18)–(A.20) give the exact decoding bound. Conditioning on distinct neighbors changes either payoff by at most  $\delta_h$ . These estimates suffice for the proposition, as the following choices show.

Choose the parameters for the prescribed  $\delta$  as follows. Choose a positive rational  $\eta$  with  $10\eta < \delta$ , and a rational  $0 < \kappa < 1/2$  with  $2\kappa k < \eta$ . Put  $K = \binom{k}{2}$ . Choose an integer  $q$  so large that

$$\frac{1}{q} \leq \mu_0(\kappa), \quad Kq^{-\kappa/2} < \eta.$$

Choose the data  $J, m, \alpha$  from Lemma A.3 at error  $\eta$ . The two analytic error functions now have fixed parameters and are uniform in  $R$ . Choose a positive rational  $t < 1/(kq)$  so small that

$$Kqe_{\text{col}}(t) < \eta, \quad e_{\text{inv}}(kqt) < \eta.$$

Set  $\chi = \eta$ . Choose positive rationals  $\zeta, \nu < 1$  with

$$\zeta < \eta, \quad \nu < \frac{\eta \kappa t^4}{6400},$$

and choose an integer  $H \geq \max\{k, q\}$  with  $K/H < \eta$ . These choices depend only on  $k, \Pi, \delta$ . In particular, they precede both the choice of  $h$  in Lemma A.1 and the alphabet supplied by Theorem 1.1.

For every  $h \geq H$ ,  $\delta_h < \eta$ . The local gap (A.6), Bernoulli's inequality, and (A.23) give

$$\text{YES : } \text{OPT}_{\Pi}(I) \geq 1 - \zeta - 2\kappa k - \eta - \delta_h > 1 - 4\eta > 1 - \delta,$$

whereas the displayed definition of  $S_{\text{loc}}$ , verified in (A.17), gives

$$\text{NO : } \text{OPT}_{\Pi}(I) \leq S_{\text{loc}} + \chi + \delta_h < \rho_{\Pi} + 6\eta < \rho_{\Pi} + \delta.$$

This proves the numerical claims simultaneously for every  $R$ .

We prove the two estimates for these fixed parameters next. Their proofs also apply to any of the parameters satisfying the initial inequalities above.

Suppose a labeling  $L$  strongly satisfies at least  $1 - \zeta$  of the left vertices. Define the coarse rank map  $f(b, z) = z_{L(b)}$ . At every strongly satisfied  $a$  and every neighbor  $b$ ,

$$f(b, \pi_{a \rightarrow b} z) = z_{\pi_{a \rightarrow b}^{-1}(L(b))} = z_{L(a)}.$$

Thus all  $k$  positions read the same label coordinate, for every choice of neighbors. With probability  $(1 - 2\kappa)^k$  none of these  $k$  symbols is resampled; conditional on that event their tuple has law  $\mu_c$ . Nonnegativity of the payoff gives

$$V(f) \geq (1 - \zeta)(1 - 2\kappa)^k \sum_c \omega_c \mathbb{E}_{X \sim \mu_c} \bar{P}_0(X). \quad (\text{A.13})$$

For soundness fix an injective order  $O : \mathcal{V} \rightarrow [|\mathcal{V}|]$ . At a fixed  $a$ , its local set  $N(a) \times [m]^R$  has exactly  $M = hm^R$  distinct IDs. Sort this set by  $O$  and divide the sorted list into  $q$  consecutive blocks  $C_{a,1}, \dots, C_{a,q}$  of size  $M/q$ . These sizes are integers because  $q \mid h$ . Each block is the intersection of the local set with an interval of global ranks. Define

$$F_a^j(z) = \frac{1}{h} \sum_{b \in N(a)} \mathbf{1}_{\{(b, \pi_{a \rightarrow b} z) \in C_{a,j}\}}, \quad F_a = (F_a^1, \dots, F_a^q). \quad (\text{A.14})$$

For every  $z$ ,  $F_a(z) \in \mathcal{S}_q$ . Each edge permutation is a bijection on strings, so

$$\mathbb{E}_z F_a^j(z) = \frac{|C_{a,j}|}{hm^R} = \frac{1}{q}. \quad (\text{A.15})$$

For any interval  $I$  of integer ranks put

$$F_a^I(z) = \frac{1}{h} \sum_{b \in N(a)} \mathbf{1}_{\{O(b, \pi_{a \rightarrow b} z) \in I\}}.$$

Let  $A_t \subseteq A$  consist of the vertices  $a$  for which  $\text{Inf}_\ell(T_{1-\kappa} F_a^I) \leq t$  for every rank interval  $I$  and every  $\ell \in [R]$ . This is the condition on averaged interval functions used in [8, Definition 7.2], with the boundary value  $t$  included here. At each  $a \in A_t$ , the bound holds for every block function  $F_a^j$ .

Let  $j_a(w)$  be the block index of a local ID  $w$ . Conditional on the local query strings, independence of the neighbor choices gives the exact identity

$$\mathbb{E}_{b_1, \dots, b_k} \bar{P}_0(j_a(w_1), \dots, j_a(w_k)) = P_0^{\text{ml}}(F_a(\tilde{Z}_1), \dots, F_a(\tilde{Z}_k)). \quad (\text{A.16})$$

The identity includes outcomes with equal neighbors. Further smoothing decreases the block influences, so  $\text{Inf}_\ell(T_{1-2\kappa} F_a^j) \leq t$ . Lemma A.5, applied to the single function  $F_a$ , bounds the expected coarse payoff in (A.16) by  $\text{opt}_q(J) + e_{\text{inv}}(kqt)$ .

When the  $k$  coarse ranks are distinct, their induced permutation is the one induced by the fine ranks  $O(w_i)$ . Otherwise the two payoffs differ by at most one. This also covers a repeated output ID, which necessarily produces a coarse tie. For two positions  $i \neq i'$ , independence of their noise and neighbor choices gives their coarse tie probability as

$$\sum_{j=1}^q \mathbb{E}_c \mathbb{E}_{Z \sim \mu_c^{\otimes R}} [(T_{1-2\kappa} F_a^j)(Z_i)(T_{1-2\kappa} F_a^j)(Z_{i'})].$$

Conditional on  $c$ , both strings have uniform product marginals. Their correlation is permitted by Lemma A.4. Use that lemma with the fixed mean  $1/q$  from (A.15), and write  $e_{\text{col}} = e_{\text{col}}^{m, \kappa}$ . The last display is at most  $q^{-\kappa/2} + qe_{\text{col}}(t)$ . A union bound over pairs therefore gives, at every  $a \in A_t$ ,

$$V_a(O) \leq S_{\text{loc}} := \text{opt}_q(J) + \binom{k}{2} (q^{-\kappa/2} + qe_{\text{col}}(t)) + e_{\text{inv}}(kqt). \quad (\text{A.17})$$

We next turn an excess over this bound into a labeling of  $\Phi$ . For each  $b \in B$ , restrict  $O$  to  $\{b\} \times [m]^R$  and compress its ranks to obtain a strict order  $o_b$  of  $[m]^R$ . Let

$$\mathcal{L}_b = \{\ell \in [R] : \text{Inf}_\ell(T_{1-\kappa} \mathbf{1}_{\{o_b(z) \in I\}}) > t/2 \text{ for some rank interval } I\}.$$

The bound on the number of influential coordinates [8, Definition 4.2 and Lemma 4.3] gives

$$|\mathcal{L}_b| \leq \frac{400}{\kappa(t/2)^3} = \frac{3200}{\kappa t^3}. \quad (\text{A.18})$$

This bound is uniform in  $R$ .

If  $a \notin A_t$ , choose one witnessing interval  $I_a$  and coordinate  $\ell_a$  with  $\text{Inf}_{\ell_a}(T_{1-\kappa} F_a^{I_a}) > t$ . For  $g_b(z) = \mathbf{1}_{\{O(b,z) \in I_a\}}$ , the averaging identity is  $F_a^{I_a} = h^{-1} \sum_b g_b \circ \pi_{a \rightarrow b}$ . Uniform resampling is unchanged by a coordinate permutation. Thus, for every  $0 \leq \theta \leq 1$  and every coordinate  $\ell$ ,

$$T_\theta(g_b \circ \pi_{a \rightarrow b}) = (T_\theta g_b) \circ \pi_{a \rightarrow b}, \quad \text{Inf}_\ell(g_b \circ \pi_{a \rightarrow b}) = \text{Inf}_{\pi_{a \rightarrow b}(\ell)}(g_b).$$

Influence is the squared  $L^2$  norm of the projection onto terms using that coordinate. Applying convexity of this norm to the averaged functions gives

$$\text{Inf}_{\ell_a}(T_{1-\kappa} F_a^{I_a}) \leq \frac{1}{h} \sum_{b \in N(a)} \text{Inf}_{\pi_{a \rightarrow b}(\ell_a)}(T_{1-\kappa} \mathbf{1}_{\{O(b,z) \in I_a\}}). \quad (\text{A.19})$$

Each term is at most one. An average greater than  $t$  forces more than a  $t/2$  fraction of the terms to exceed  $t/2$ . For a fixed right vertex  $b$ , intersecting a global-rank interval with  $\{b\} \times [m]^R$  gives an interval in the relative order  $o_b$ , so these neighbors satisfy  $\pi_{a \rightarrow b}(\ell_a) \in \mathcal{L}_b$ .

Give each such  $a$  label  $\ell_a$ , and label every other left vertex arbitrarily. Independently at each right vertex with  $\mathcal{L}_b \neq \emptyset$ , choose one uniform label from  $\mathcal{L}_b$ ; use any label when the set is empty. By (A.18), the expected satisfied fraction among the edges at each  $a \notin A_t$  is greater than

$$\frac{t}{2} \frac{\kappa t^3}{3200} = \frac{\kappa t^4}{6400}.$$

One random right labeling is used for all left vertices. Linearity of expectation and the common left degree identify the resulting average with the ordinary uniform edge fraction.

If  $V(O) > S_{\text{loc}} + \chi$ , more than a  $\chi$  fraction of left vertices lie outside  $A_t$ : if that fraction is  $g$ , the bound  $0 \leq V_a(O) \leq 1$  and (A.17) give  $V(O) \leq (1-g)S_{\text{loc}} + g \leq S_{\text{loc}} + g$ . The preceding decoder would then give a labeling of edge value greater than  $\chi \kappa t^4 / 6400$ . Consequently

$$\text{val}(\Phi) \leq \nu < \frac{\chi \kappa t^4}{6400} \implies V(O) \leq S_{\text{loc}} + \chi \quad \text{for every injective } O. \quad (\text{A.20})$$

It remains to turn the independent experiment into strict ordering constraints. Let  $D$  be the event that  $b_1, \dots, b_k$  are distinct. For  $h \geq k$ ,

$$p_D := \Pr(D) = \frac{(h)_k}{h^k}, \quad \delta_h := 1 - p_D \leq \frac{\binom{k}{2}}{h}. \quad (\text{A.21})$$

The probability is the same at every  $a$ , and  $D$  is independent of the local experiment. Conditional on  $D$ ,  $a$  remains uniform and its ordered distinct neighbor tuple is uniform among its  $(h)_k$  possibilities. All output IDs are then distinct because their first coordinates  $b_i$  differ. Let  $I_0$  denote this finite weighted distribution of strict  $P_0$  constraints, and write  $\text{OPT}_{P_0}(I_0)$  for its maximum strict payoff.

Let  $V_D(f)$  denote the expected tie-extended payoff after this conditioning, and use  $V_{-D}(f)$  for conditioning on its complement. For any rank map  $f$ , coarse or injective,

$$V(f) = p_D V_D(f) + (1 - p_D) V_{-D}(f), \quad |V_D(f) - V(f)| \leq \delta_h, \quad (\text{A.22})$$

because all payoffs lie in  $[0, 1]$ . Apply this first to the coarse map in (A.13). Within each finite fiber of that map, choose one uniform random permutation of the distinct IDs, independently between fibers, and concatenate the fibers in rank order. On every conditioned tuple, restriction of these permutations to its IDs is uniform within each tied rank class. Thus the expected strict  $P_0$ -payoff of this one random global order equals  $V_D(f)$ . Some global order attains at least that value. For soundness, an injective  $O$  has distinct ranks on every conditioned tuple, so its tie-extended payoff there is its strict  $P_0$ -payoff. Equations (A.13), (A.20), and (A.22) therefore give the strict bounds

$$\begin{aligned} \text{YES : } \quad \text{OPT}_{P_0}(I_0) &\geq (1 - \zeta)(1 - 2\kappa)^k \sum_c \omega_c \mathbb{E}_{X \sim \mu_c} \bar{P}_0(X) - \delta_h, \\ \text{NO : } \quad \text{OPT}_{P_0}(I_0) &\leq S_{\text{loc}} + \chi + \delta_h \quad \text{if } \nu < \chi \kappa t^4 / 6400. \end{aligned} \tag{A.23}$$

These bounds are for the conditioned  $P_0$  tuples. Applying (A.5) makes every occurrence use the original predicate  $\Pi$  with exactly the same values. Call this emitted instance  $I$ .

Finally we specify the encoded transformation. Once  $R$  is fixed, the choice of  $c$ , the product  $\mu_c^{\otimes R}$ , and the independent rational  $2\kappa$ -noise form one fixed finite rational law  $\theta_R$  on local outcomes  $\xi$ . These are the actual local outcomes above; the Gaussian variables used for the analytic comparison are not part of the output. Renumber  $B$  by  $[|B|]$  and explicitly list  $\mathcal{V} = B \times [m]^R$ . For each  $a \in A$ , each ordered distinct neighbor tuple, and each  $\xi \in \text{supp } \theta_R$ , list the tuple in (A.12), permuted by (A.5), with weight

$$\frac{\theta_R(\xi)}{|A|(h)_k}. \tag{A.24}$$

The weights sum to one and are nonnegative rationals. Every listed tuple has distinct IDs; the same tuple may occur in several entries. There are  $|B|m^R$  variables and at most  $|A|(h)_k |\text{supp } \theta_R|$  entries. For fixed  $k, m, R, \theta_R$ , these counts are polynomial in the explicit graph size, even when  $h$  varies. A fixed common denominator for  $\theta_R$  shows that each weight has  $O(\log |A| + k \log h)$  bits plus a fixed constant. The IDs have  $O(\log |B|)$  bits plus a fixed constant. Exact enumeration of these entries is therefore deterministic and polynomial for fixed  $R$ , as asserted.  $\square$

### A.3 Proofs of the local ordering estimates

Lemma A.3 realizes the finite local data, Lemma A.4 controls collisions under arbitrary couplings, and Lemma A.5 compares the common simplex-valued function with a random assignment of the local instance.

*Proof of Lemma A.3.* The explicit increasing-tuple construction in [8, Theorem 11.1, Definition 11.2, and the proof of Theorem 2.5], applied to  $P_0$ , gives a finite instance  $J^* = (U^*, \mathcal{C}^*, \omega^*)$  with  $\text{opt}_q(J^*) \leq \rho_\Pi + \eta$  and a linear order satisfying every constraint. To obtain the error  $\eta$  in this application, use Theorem 11.1 there with per-permutation error  $\eta/k!$ . For any  $q$ -assignment, sum its induced-permutation probabilities against  $P_0(\sigma)$ ; their sum is at most  $\rho_\Pi + \eta$ . The identity order satisfies every increasing tuple because  $P_0(\text{id}) = 1$ . Definition 11.2 samples a scale, an interval at that scale, and one point from each of its  $k$  successive subintervals. These are uniform finite choices, giving rational weights and distinct tuple positions.

Use the finite-copy smoothing from the proof of [8, Lemma 8.9], specialized to this complete order. Take  $L$  disjoint copies of  $J^*$ , each of weight  $1/L$ , and place their complete orders consecutively in an order  $\sigma : U \rightarrow [m]$ , where  $m = L|U^*|$ . The resulting instance  $J$  still has  $\text{opt}_q(J) \leq \rho_\Pi + \eta$ , since the value of any  $q$ -assignment is the average of its values on the copies. Choose  $L$  and a positive rational  $s_0 < 1$  so that  $1/L + s_0 < \eta$ . Define a random assignment  $X : U \rightarrow [m]$  as follows: with probability  $1 - s_0$ , choose a uniform cyclic shift of the ranks of  $\sigma$ ; with probability  $s_0$ , assign the variables independently and uniformly from  $[m]$ .

For a variable of rank  $r$ , the  $m$  cyclic shifts give each value in  $[m]$  exactly once; its independent-noise label is also uniform. A cyclic cut can split at most one of the  $L$  consecutive copies. In every other copy it preserves all within-copy comparisons, including the order of every constraint tuple, so every shift has payoff at least  $1 - 1/L$ . The mixture's payoff is therefore at least  $(1 - s_0)(1 - 1/L) > 1 - \eta$ . Let  $\mu_c$  be its restriction to the variables of  $c$ . If  $n_c(x)$  is the number of the  $m$  cyclic shifts with restriction  $x$ , then

$$\mu_c(x) = (1 - s_0) \frac{n_c(x)}{m} + \frac{s_0}{m^k}.$$

These atoms are rational and exceed  $\alpha := s_0/(2m^k)$ . The denominator  $m^k$  is valid because the variables in  $c$  are distinct.

Finally regard the finite mixture as a probability space and take  $I = 1$  and  $b_{u,j} = \mathbf{1}_{\{X(u)=j\}}$  in its  $L^2$  space. Uniform marginals, disjoint events for different labels of one variable, and restrictions to each  $c$  give all the displayed Gram identities. This constructs the required feasible data for this particular local instance.  $\square$

*Proof of Lemma A.4.* We bound each smoothed function's squared norm first; Cauchy–Schwarz will then handle every coupling at once. This is the dimension-uniform form of [8, Lemmas 3.4 and 3.5] needed here. For  $H = T_{1-\kappa}F$ , the Fourier formula gives

$$\|T_{1-2\kappa}F\|_2^2 \leq \text{Stab}_{1-\kappa}(H),$$

where  $\text{Stab}_\theta(H) = \sum_\sigma \theta^{|\sigma|} \widehat{H}(\sigma)^2$ . Indeed,  $(1 - \kappa)^3 - (1 - 2\kappa)^2 = \kappa(1 - \kappa - \kappa^2) > 0$ . Each full coordinate influence bound for  $H$  implies the corresponding low-degree influence bound in Theorem 4.4 of Mossel, O'Donnell, and Oleszkiewicz [18]. Apply that theorem with product dimension  $R$ , one-coordinate atom bound  $1/m$ , and correlation  $1 - \kappa$ . For all sufficiently small  $t$ , it gives

$$\text{Stab}_{1-\kappa}(H) \leq \Gamma_{1-\kappa}(\mu) + C \frac{\log(2m)}{\kappa} \frac{\log \log(1/t)}{\log(1/t)}.$$

The constant  $C > 0$  is absolute, and both the small- $t$  cutoff and the error are independent of  $R$  and  $\mu$ . The Gaussian comparison term is  $\Gamma_\theta(\mu) = \Pr[X \leq a_\mu, Y \leq a_\mu]$ , where  $X, Y$  are standard jointly Gaussian variables of correlation  $0 \leq \theta \leq 1$  and  $\Pr[X \leq a_\mu] = \mu$  for  $0 < \mu < 1$ . At the endpoints, set  $\Gamma_\theta(0) = 0$  and  $\Gamma_\theta(1) = 1$ .

For fixed  $\kappa$ , Theorem B.5 of the same source gives  $\Gamma_{1-\kappa}(\mu) = \mu^{2/(2-\kappa)+o(1)}$  as  $\mu \downarrow 0$ . Since

$$\frac{2}{2-\kappa} - \left(1 + \frac{\kappa}{2}\right) = \frac{\kappa^2}{2(2-\kappa)} > 0,$$

there is a cutoff  $\mu_0(\kappa) > 0$ , independent of  $m, R$ , below which  $\Gamma_{1-\kappa}(\mu) \leq \mu^{1+\kappa/2}$ . This proves the corresponding squared  $L^2$  bound for both  $F$  and  $G$ , with an error  $e_{\text{col}}^{m,\kappa}(t) \rightarrow 0$  independent of the stated variables. Choose a cutoff  $t_0 < e^{-e}$  within the range of the displayed bound, use its nonnegative error for  $0 < t < t_0$ , and set  $e_{\text{col}}^{m,\kappa}(t) = 1$  for  $t \geq t_0$ . The latter range follows trivially from the product being at most one. Cauchy–Schwarz under any coupling with the specified marginals proves (A.9).  $\square$

*Proof of Lemma A.5.* This is the simplex-valued estimate of [8, Lemma 10.5], with its sufficient scalar influence bounds and their dependence made explicit. Put  $H = T_{1-\lambda}F$ . Independence of the noise between positions and multilinearity turn the left side of (A.11) into

$$\mathbb{E}_c \mathbb{E}_{Z \sim \mu_c^{\otimes R}} P_0^{\text{ml}}(H(Z_1), \dots, H(Z_k)). \quad (\text{A.25})$$

For each coordinate, the integral ensemble consists of the indicators  $\mathbf{1}_{\{X_i=j\}}$  under  $\mu_c$ ; its atoms have probability greater than  $\alpha$ . The common Gram data (A.7) supply one Gaussian ensemble for all variables

of  $J$  with the same first and second moments on every constraint, as in [8, Lemma 10.4]. Explicitly, for a standard Gaussian vector  $g$  in the span of the Gram vectors, the variables

$$G_{u,j} = \langle I, b_{u,j} \rangle + \langle b_{u,j} - \langle I, b_{u,j} \rangle I, g \rangle$$

have those moments. Indeed, put  $a_{u,j} = \langle I, b_{u,j} \rangle = 1/m$ . For any variables  $u, v$  and symbols  $i, j$ ,

$$\mathbb{E}G_{u,i} = a_{u,i}, \quad \mathbb{E}[G_{u,i}G_{v,j}] = a_{u,i}a_{v,j} + \langle b_{u,i} - a_{u,i}I, b_{v,j} - a_{v,j}I \rangle = \langle b_{u,i}, b_{v,j} \rangle.$$

Equation (A.7) identifies the last expression with the integral pair moment on each constraint, including the diagonal case. Using the same Gaussian vector  $g$  for all  $u, j$  gives one joint family for every occurrence. For  $\ell \in [R]$ , let  $(G_{u,j}^{(\ell)})_{u,j}$  be independent copies of this whole Gaussian ensemble, and write  $G_u = (G_{u,j}^{(\ell)})_{\ell \in [R], j \in [m]}$  for the array used to evaluate a product polynomial at variable  $u$ . Each  $H^j$  is evaluated on these ensembles through its multilinear product expansion; equivalently, express an orthonormal basis on  $[m]$  as linear combinations of the symbol indicators and use the same polynomial for the integral and Gaussian ensembles.

For a fixed  $c$  and  $0 \leq \theta \leq 1$ , let  $\mathcal{T}_\theta^c$  be the averaging operator on functions of  $(Z_1, \dots, Z_k)$  that retains each whole coordinate tuple with probability  $\theta$  and otherwise resamples it from  $\mu_c$ , independently between coordinates. Uniform one-position marginals give, componentwise,

$$\mathcal{T}_{1-\lambda}^c[F^j(Z_i)] = (T_{1-\lambda}F^j)(Z_i) = H^j(Z_i).$$

Resampling a whole coordinate tuple resamples its symbol at position  $i$  from the uniform marginal, independently across coordinates. Thus the identity holds for every input tuple and every  $\lambda$ . Both operators multiply the component supported on  $S \subseteq [R]$  by  $(1-\lambda)^{|S|}$ , and the same coefficients define the smoothed polynomial on the Gaussian ensemble. These are precisely the smoothed components needed by the invariance theorem. The uniform marginals also give the following influence identity, whose superscripts specify the product measures:

$$\text{Inf}_\ell^{\mu_c^{\otimes R}}(H^j(Z_i)) = \text{Inf}_\ell^{[m]^R}(H^j). \quad (\text{A.26})$$

Indeed, conditioning on all coordinate tuples except the  $\ell$ th leaves the symbol at position  $i$  uniform; the other symbols at that position have the uniform product law. Vector influence is the sum of component influences [11, Definition 2.6]. Consequently each coordinate influence of  $(H^j(Z_i))_{i \in [k], j \in [q]}$  is at most  $kqt$ , and each component has variance at most one.

For the remaining tail hypothesis, let  $(F^j(Z_i))_S$ ,  $S \subseteq [R]$ , be the Hoeffding components on the tuple product  $\mu_c^{\otimes R}$ , and let the superscript  $> d$  retain the terms with  $|S| > d$ . The tuple-noise identity above multiplies component  $S$  by  $(1-\lambda)^{|S|}$ . Since the unsmoothed  $F^j$  takes values in  $[0, 1]$ , for every real  $d > 0$ ,

$$\begin{aligned} \text{Var}((H^j(Z_i))^{>d}) &= \sum_{|S|>d} (1-\lambda)^{2|S|} \|(F^j(Z_i))_S\|_2^2 \\ &\leq (1-\lambda)^{2d} \text{Var}(F^j(Z_i)) \leq (1-\lambda)^{2d}. \end{aligned}$$

All norms and variances here use  $\mu_c^{\otimes R}$ . This supplies the tail bound in [11, Theorem 3.6]. In the notation of that theorem, take the polynomial vector dimension to be  $kq$ , the smoothing parameter to be  $\lambda$ , and the influence threshold to be  $\tau = kqt < 1$ . The low-degree influence of any component, at its cutoff  $d = \log(1/\tau)/(18 \log(1/\alpha))$ , is no larger than its full influence. The displayed tail estimate holds at this cutoff as well. These facts supply the hypotheses of [11, Theorem 3.6], and hence of the vector-valued estimate in [8, Theorem 10.1].

Gaussian evaluations need not lie in  $\mathcal{S}_q$ . To compare the same payoff on both ensembles, use the nonexpansive Euclidean projection  $\mathbf{p} : \mathbb{R}^q \rightarrow \mathcal{S}_q$  and define  $\Psi : (\mathbb{R}^q)^k \rightarrow \mathbb{R}$  by

$$\Psi(x_1, \dots, x_k) = P_0^{\text{ml}}(\mathbf{p}(x_1), \dots, \mathbf{p}(x_k))$$

This function is Lipschitz with constant at most  $\sqrt{kq}$ : on the simplex, telescoping the positions bounds the payoff difference by  $\sum_i \|x_i - y_i\|_1$ , and projection and Cauchy–Schwarz give the claim. Apply the cited invariance estimate to this function, separately for each  $c$ , with vector dimension  $kq$  and influence threshold  $kqt$ . Its displayed error is at most

$$e_{\text{inv}}(kqt), \quad e_{\text{inv}}(u) = C_{kq} \sqrt{kq} u^{\lambda/(18 \log(1/\alpha))}.$$

Here  $C_{kq} > 0$  is the constant depending only on the vector dimension  $kq$  in [11, Theorem 3.6]. The error is independent of  $R$ .

On the integral ensemble,  $H(Z_i) \in \mathcal{S}_q$ , so  $\Psi$  equals the payoff in (A.25). On the common Gaussian ensemble, project  $H(G_u)$  to  $\mathcal{S}_q$  for every  $u \in U$ , and independently draw one  $q$ -label for each  $u$  with those probabilities. This is one random assignment of  $J$ . The variables in every local tuple are distinct, so its conditional expected payoff at that tuple is the multilinear payoff  $\Psi$ . Every realized assignment has value at most  $\text{opt}_q(J)$ . Averaging the invariance comparisons over  $c$  proves (A.11).  $\square$

## B Abundance of subspaces with few vanishing restrictions

Lemma 3.5 supplies the subspaces used to control rank loss in Section 3. The proof below establishes the required uniformity over all extension maps.

*Proof of Lemma 3.5.* The difficulty is to rule out every map  $g$  at once. We write Equation (3.8), on more than  $3r$  distinct nonzero binary inputs, as a linear system for the  $3r$  field coordinates of the values of  $g$  on a binary basis. When the basis entries are indeterminates, this system has no solution. After entrywise squaring, a nonzero minor of its augmented matrix gives a polynomial that must vanish whenever substituting finite-field entries makes the system solvable.

Use the notation  $F = \mathbb{F}_{2^d}$  and  $q = 2^d$  from Section 3. Choose independent uniform  $x_{ij} \in F$ , for  $1 \leq i \leq 3$  and  $1 \leq j \leq r$ , and parametrize a binary linear map by

$$z_i(t) = \sum_{j=1}^r x_{ij} t_j, \quad t \in \mathbb{F}_2^r.$$

Each nonzero input has uniform image in  $F^3$ , so the probability that this map is not injective is at most  $(2^r - 1)q^{-3}$ . Conditioned on injectivity, its image is a uniform binary  $r$ -subspace: every such subspace has the same number of ordered binary bases. Distinct nonzero binary inputs are linearly independent also over  $F$ , so their images are independent uniform elements of  $F^3$ . Conditional on the first image being nonzero, its  $F$ -line has  $q$  elements, and the second uniform image lies on that line with probability  $q/q^3 = q^{-2}$ . A union bound therefore shows that the nonproportionality condition fails with probability  $O_r(q^{-2})$ , apart from the event of noninjectivity.

It remains to prove the alignment condition simultaneously for all  $g$ . Fix a set  $E$  of more than  $3r$  distinct nonzero inputs in  $\mathbb{F}_2^r$ . On an injective parametrization, the values of a binary linear  $g$  on a basis can be arbitrary vectors in  $F^3$ . Thus alignment on all of  $E$  implies that the vector

$$c = (\sqrt{z_1(t)z_2(t)z_3(t)})_{t \in E}$$

lies in the  $F$ -column span of the matrix with  $3r$  columns

$$M_{t,(i,j)} = t_j z_i(t), \quad t \in E, \quad 1 \leq i \leq 3, \quad 1 \leq j \leq r.$$

Conversely, membership supplies the basis values of such a  $g$ . In particular, this column-membership condition already accounts for all possible maps  $g$ .

First regard the  $x_{ij}$  as independent indeterminates and set  $L = \mathbb{F}_2(x_{ij})$ . We claim that  $c$  is not in the column span of  $M$  even after extending scalars to an algebraic closure  $\bar{L}$ . The following  $r^3$  elements are linearly independent over  $L$ :

$$\sqrt{x_{1j}x_{2k}x_{3l}}, \quad 1 \leq j, k, l \leq r.$$

Indeed, after clearing denominators, a dependence would be a polynomial identity in the square roots of the indeterminates. The displayed monomials have distinct parity vectors, whereas coefficients polynomial in the  $x_{ij}$  have even exponents in those square roots. Terms from different parity vectors cannot cancel.

The expansion of  $c$  in these independent elements has coefficient vectors

$$v_{jkl} = (t_j t_k t_l)_{t \in E}.$$

Suppose that  $c$  belonged to the column span over  $\bar{L}$ . Every row vector over  $L$  annihilating  $M$  would then annihilate  $c$ . The independence just proved implies that it would annihilate every  $v_{jkl}$ . By double annihilation in the finite-dimensional space  $L^E$ , every  $v_{jkl}$  would belong to the column span over  $L$  itself.

For  $h \geq 1$ , let  $V_h \leq L^E$  be the evaluation span of all positive-degree monomials in the binary coordinates  $t_1, \dots, t_r$  of degree at most  $h$ . Repeated indices in  $v_{jkl}$  give all monomials of degrees one and two as well as three, because  $t_i^2 = t_i$  on the evaluation set. The preceding conclusion and the form of the columns therefore imply

$$V_3 \subseteq \text{col}_L(M) \subseteq V_2 \subseteq V_3.$$

These three spaces are equal. Multiplication by any  $t_i$  sends  $V_2$  into  $V_3$ , so preserves their common value. Since this space contains the degree-one monomials, it contains evaluations of every positive-degree monomial. These evaluations span all of  $L^E$ : for any nonzero binary point, its usual indicator polynomial is a product of factors  $t_i$  and  $1 + t_i$  with at least one factor  $t_i$ , and thus has no constant monomial. Its restriction to  $E$  gives the corresponding point indicator. Consequently the column span would have dimension  $|E| > 3r$ , a contradiction.

To pass from this algebraic statement to finite fields, square the entries of both  $M$  and  $c$ . The resulting matrices have polynomial entries, and the appended column has entries  $z_1(t)z_2(t)z_3(t)$ . Let  $h = \text{rank}_L M \leq 3r$ . Squaring preserves rank over  $\bar{L}$ , so the augmented squared matrix has a nonzero minor of size  $h+1$ . This minor is a nonzero polynomial of degree at most  $2h+3 \leq 6r+3$ . All unaugmented minors of size  $h+1$  vanish identically, so the unaugmented rank cannot increase on specialization. For a finite-field specialization with  $M\lambda = c$  for some coefficient vector  $\lambda$ , entrywise squaring gives  $M^{[2]}\lambda^{[2]} = c^{[2]}$ . The augmented squared matrix then has rank at most  $h$ , because the unaugmented  $(h+1)$ -minors are identically zero. Thus column membership forces the particular nonzero augmented minor to vanish. The polynomial defining this minor, including a nonzero coefficient over  $\mathbb{F}_2$ , is fixed before the extension degree  $d$  varies. For uniform independent  $x_{ij} \in F$ , this event has probability at most  $(6r+3)/q$  by the polynomial zero bound [24, Corollary 1]; see also the earlier sparse-polynomial formulation in [25, Section 3.1].

If the alignment condition fails, it fails on some set  $E$  of size  $3r+1$ . There are finitely many such sets depending only on  $r$ ; if  $2^r - 1 \leq 3r$  there are none and the condition is automatic. A union bound, followed by conditioning on injectivity, proves the lemma.  $\square$

## C Agreement from Grassmann expansion

The reduction in Section 7 requires computable dimension thresholds in Theorem 5.1.

*Proof of Theorem 5.1.* We derive this consequence from the Grassmann expansion theorem of Khot–Minzer–Safra [14, Theorem 1.12 in the 2018 revision], using the matrix chart of Barak–Kothari–Steurer [1, Definition 3.2 and Lemmas 3.3–3.7].

The Grassmann graph on an ambient binary space  $E$  has the  $\ell$ -dimensional subspaces as vertices; two vertices are adjacent when their intersection has dimension  $\ell - 1$ . The retention of a nonempty vertex

set is the probability that a uniform vertex in the set has a uniform neighbor still in the set. The cited theorem says that a set of retention at least a fixed  $\zeta > 0$  has positive relative density in some interval

$$\{L : A_0 \subseteq L \subseteq B_0\}, \quad \dim A_0 + \text{codim}_E B_0 \leq r.$$

The density bound and  $r$  depend only on  $\zeta$ . The theorem holds for all sufficiently large  $\ell$ , and then all sufficiently large  $\dim E$ .

Equality acceptance is the average of the retention probabilities of the nonempty fibers  $f^{-1}(y)$  under the matrix step, weighted by their uniform masses. Some fiber  $S$  therefore has retention at least  $\eta$ . The probability that either perturbation factor is zero is

$$p_0 = 2^{-\ell} + 2^{-m} - 2^{-\ell-m}.$$

Choose the dimension thresholds so that  $p_0 \leq \eta/2$ . Conditional on both factors being nonzero, the retention of  $S$  is at least  $(\eta - p_0)/(1 - p_0) \geq \eta/2$ .

For each matrix put

$$L_M = \{(x, M^\top x) : x \in \mathbb{F}_2^\ell\} \subseteq E := \mathbb{F}_2^\ell \oplus \mathbb{F}_2^m.$$

The map  $M \mapsto L_M$  is a bijection onto the chart  $\mathcal{C}$  of subspaces whose projection onto the first summand is invertible. Since

$$\dim(L_M \cap L_N) = \ell - \text{rank}(M - N),$$

the neighbors within the chart correspond exactly to nonzero rank-one matrix differences. Each chart vertex has  $(2^\ell - 1)(2^m - 1)$  such neighbors, whereas its full Grassmann degree is  $(2^\ell - 1)(2^{m+1} - 2)$ : choose a hyperplane in the vertex and then a one-dimensional extension in the quotient by that hyperplane, other than the vertex itself. Thus exactly half of its neighbors lie in the chart. Over  $\mathbb{F}_2$ , every nonzero rank-one matrix has a unique factorization  $al^\top$  with nonzero factors. The lifted set  $\tilde{S} = \{L_M : M \in S\}$  consequently has Grassmann retention at least  $\eta/4$ .

Apply the expansion theorem with  $\zeta = \eta/4$ , obtaining density  $\alpha$  on an interval with  $\dim A_0 + \text{codim}_E B_0 \leq r_s$ . Write  $\mathcal{I} = \{L : A_0 \subseteq L \subseteq B_0\}$ . Since  $\tilde{S} \subseteq \mathcal{C}$ ,

$$\frac{|\tilde{S} \cap \mathcal{I} \cap \mathcal{C}|}{|\mathcal{I} \cap \mathcal{C}|} = \frac{|\tilde{S} \cap \mathcal{I}|}{|\mathcal{I} \cap \mathcal{C}|} \geq \frac{|\tilde{S} \cap \mathcal{I}|}{|\mathcal{I}|} \geq \alpha.$$

The positive numerator ensures that the chart intersection is nonempty. Uniform matrices on its preimage correspond to uniform chart vertices, because  $M \mapsto L_M$  is a bijection. A basis  $(d_i, s_i)$  of  $A_0$  specifies  $A_0 \subseteq L_M$  by the equations  $d_i^\top M = s_i^\top$ . For the standard dot product on  $E$ , a basis  $(t_j, q_j)$  of  $B_0^\perp$  specifies  $L_M \subseteq B_0$  by  $Mq_j = t_j$ . Their simultaneous solution set is a slice of the required form, with at most  $r_s$  equations of either kind. On at least an  $\alpha$  fraction of this slice,  $f$  equals the value defining  $S$ . This proves the conclusion with  $z = 0$  and  $u$  equal to that value.

To choose the constants effectively, put  $\zeta = \eta/4$  and  $n = \dim E$ . The basis lift of a Grassmann set consists of all ordered bases of its members, viewed in  $(\mathbb{F}_2^n)^\ell$ . Membership is unchanged when one ordered basis is replaced by another basis of the same subspace, which verifies basis invariance. The tuple walk replaces  $(x_i)_{i=1}^\ell$  by  $(x_i + b_i y)_{i=1}^\ell$  for independent uniform  $y \in \mathbb{F}_2^n$  and  $b \in \mathbb{F}_2^\ell$ .

In the convention of [14, Section 2.4 and Theorem 2.13],  $(r, \varepsilon')$ -pseudorandomness bounds the density by  $\varepsilon'$  after prescribing any  $j$  initial vectors and sampling the rest independently in a subspace  $W$ , whenever  $j + \text{codim } W \leq r$ . For a basis-invariant set with this property, that theorem bounds retention under the tuple walk by

$$(r+1)2^{7r^3+3}(\varepsilon')^{1/4} + 2^{-(r+1)}.$$

The two error bounds below compare this retention with retention in the Grassmann graph.

Choose  $r \geq 1$  with  $2^{-(r+1)} < \zeta/4$ , and then positive rational  $\alpha$  so small that  $(r+1)2^{7r^3+3}(2\alpha)^{1/4} < \zeta/4$ . If a Grassmann set had density less than  $\alpha$  in every interval  $\{L : A_0 \subseteq L \subseteq B_0\}$  with  $\dim A_0 + \text{codim}_E B_0 \leq$

$r$ , Lemma 2.8 of that source would make its basis lift  $(r, \alpha + 2^{\ell+r-n})$ -pseudorandom, where  $n = \dim E$ . Lemma 2.7 changes retention by at most  $2^{-\ell} + 2^{\ell+1-n}$ . Choose  $\ell$  large enough that  $2^{-\ell} < \zeta/8$ , then  $n$  large enough that  $2^{\ell+r-n} \leq \alpha$  and  $2^{\ell+1-n} < \zeta/8$ , also meeting the source’s dimension requirements. The lifted set would have retention less than  $\zeta/2$ , whereas a Grassmann retention at least  $\zeta$  would give lifted retention greater than  $3\zeta/4$ , a contradiction. This supplies a rational density bound and computable thresholds, with  $r_s$  enlarged to at least  $r$  if needed.

Thus  $\alpha, r_s$  depend only on  $\eta$ ; after fixing  $\ell$ , the remaining ambient-dimension threshold becomes  $m_0(\eta, \ell)$  through  $n = \ell + m$ . No folding condition was used. The fiber argument is the equality consequence in [1, Lemma 2.3]; numbered references to that paper use its arXiv version 1.  $\square$

## References

- [1] Boaz Barak, Pravesh K. Kothari, and David Steurer. Small-set expansion in shortcode graph and the 2-to-2 conjecture. In *10th Innovations in Theoretical Computer Science Conference (ITCS 2019)*, volume 124 of *Leibniz International Proceedings in Informatics*, pages 9:1–9:12. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2019. First circulated in 2018. Section-based theorem locators refer to arXiv:1804.08662v1. doi:10.4230/LIPIcs.ITCS.2019.9. <https://arxiv.org/abs/1804.08662v1>.
- [2] Mihir Bellare, Oded Goldreich, and Madhu Sudan. Free bits, PCPs, and nonapproximability—towards tight results. *SIAM Journal on Computing*, 27(3):804–915, 1998. doi:10.1137/S0097539796302531. <https://people.seas.harvard.edu/~madhusudan/papers/1995/bgs-journ.pdf>.
- [3] Shuchi Chawla, Robert Krauthgamer, Ravi Kumar, Yuval Rabani, and D. Sivakumar. On the hardness of approximating multicut and sparsest-cut. *Computational Complexity*, 15(2):94–114, 2006. Numbered locators refer to the author manuscript. doi:10.1007/s00037-006-0210-9. <https://www.wisdom.weizmann.ac.il/~robi/papers/CKKRS-CutHardness-CompComp06.pdf>.
- [4] Erik D. Demaine, Dotan Emanuel, Amos Fiat, and Nicole Immorlica. Correlation clustering in general weighted graphs. *Theoretical Computer Science*, 361(2–3):172–187, 2006. Numbered locators refer to the author manuscript dated February 5, 2005. doi:10.1016/j.tcs.2006.05.008. [https://erikdemaine.org/papers/Clustering\\_TCS/paper.pdf](https://erikdemaine.org/papers/Clustering_TCS/paper.pdf).
- [5] Irit Dinur, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. Towards a proof of the 2-to-1 games conjecture? *Theory of Computing*, 21(11):1–50, 2025. Preliminary version in STOC 2018. doi:10.4086/toc.2025.v021a011. <https://theoryofcomputing.org/articles/v021a011/v021a011.pdf>.
- [6] Irit Dinur and David Steurer. Analytical approach to parallel repetition. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, pages 624–633. ACM, 2014. Numbered locators refer to the author manuscript dated June 10, 2014. doi:10.1145/2591796.2591884. <https://www.dsteurer.org/paper/productgames.pdf>.
- [7] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995. doi:10.1145/227683.227684.
- [8] Venkatesan Guruswami, Johan Håstad, Rajsekar Manokaran, Prasad Raghavendra, and Moses Charikar. Beating the random ordering is hard: Every ordering CSP is approximation resistant. *SIAM Journal on Computing*, 40(3):878–914, 2011. Numbered locators refer to ECCC TR11-027. doi:10.1137/090756144. <https://eccccc.weizmann.ac.il/report/2011/027/download>.

- [9] Johan Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, 2001. Theorem locators refer to the author manuscript dated February 18, 2002. doi:10.1145/502090.502098. <https://people.kth.se/~johanh/optimalinap.pdf>.
- [10] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi:10.1080/01621459.1963.10500830.
- [11] Marcus Isaksson and Elchanan Mossel. Maximally stable gaussian partitions with discrete applications. arXiv:0903.3362v3 [math.PR], 2009. Numbered locators refer to arXiv:0903.3362v3 (3 August 2009). <https://arxiv.org/pdf/0903.3362v3>.
- [12] Subhash Khot. On the Unique Games Conjecture. In *Proceedings of the 25th Annual IEEE Conference on Computational Complexity*, pages 99–121. IEEE Computer Society, 2010. Conjecture locators refer to the author manuscript. doi:10.1109/CCC.2010.19. <https://cs.nyu.edu/~khot/papers/UGCSurvey.pdf>.
- [13] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O’Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM Journal on Computing*, 37(1):319–357, 2007. Numbered theorem references use the author manuscript dated September 19, 2005. Author manuscript. doi:10.1137/S0097539705447372.
- [14] Subhash Khot, Dor Minzer, and Muli Safra. Pseudorandom sets in Grassmann graph have near-perfect expansion. *Annals of Mathematics*, 198(1):1–92, 2023. Theorem locators refer to ECCC TR18-006, revision 2, May 19, 2018; the expansion theorem is Theorem 1.12. doi:10.4007/annals.2023.198.1.1. <https://eccc.weizmann.ac.il/report/2018/006/revision/2/download>.
- [15] Subhash Khot, Dor Minzer, and Muli Safra. On independent sets, 2-to-2 games and Grassmann graphs. *Theory of Computing*, 21(10):1–55, 2025. Preliminary version in STOC 2017. doi:10.4086/toc.2025.v021a010.
- [16] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within  $2 - \epsilon$ . *Journal of Computer and System Sciences*, 74(3):335–349, 2008. Numbered locators refer to the linked author manuscript. doi:10.1016/j.jcss.2007.06.019. [https://cims.nyu.edu/~regev/papers/vc\\_hard.pdf](https://cims.nyu.edu/~regev/papers/vc_hard.pdf).
- [17] Subhash Khot and Muli Safra. A two-prover one-round game with strong soundness. *Theory of Computing*, 9(28):863–887, 2013. Preliminary version in FOCS 2011. doi:10.4086/toc.2013.v009a028.
- [18] Elchanan Mossel, Ryan O’Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Annals of Mathematics*, 171(1):295–341, 2010. doi:10.4007/annals.2010.171.295.
- [19] OpenAI. Constant-factor hardness of uniform sparsest cut. OpenAI Math Release preprint OAI:Constant-factor-hardness-of-uniform-sparsest-cut-September-24-2026, 2026.
- [20] OpenAI. Perfect completeness for 2-to-1 games. OpenAI Math Release preprint OAI:Perfect-completeness-for-2-to-1-games-September-23-2026, 2026.
- [21] OpenAI. The Gaussian propeller bound in every dimension. OpenAI Math Release preprint OAI:The-Gaussian-Propeller-Bound-in-Every-Dimension-September-24-2026, 2026.
- [22] Prasad Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, pages 245–254. ACM, 2008. doi:10.1145/1374376.1374414.

- [23] Prasad Raghavendra. *Approximating NP-hard Problems: Efficient Algorithms and Their Limits*. PhD thesis, University of Washington, 2009. <https://people.eecs.berkeley.edu/~prasad/Files/thesis.pdf>.
- [24] Jacob T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM*, 27(4):701–717, 1980. doi:10.1145/322217.322225.
- [25] Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Symbolic and Algebraic Computation (EUROSAM 1979)*, volume 72 of *Lecture Notes in Computer Science*, pages 216–226. Springer, 1979. doi:10.1007/3-540-09519-5\_73.