

Two-Sided Product Expanding Codes via Rademacher Matrices

Eshan Chattopadhyay*
Cornell University
eshan@cs.cornell.edu

Noam Ringach †
Cornell University
nomir@cs.cornell.edu

Nicholas Spooner
Cornell University
nspooner@cornell.edu

Abstract

Beginning with the work of Dinur, Lin, and Vidick (FOCS, 2024), tensor codes with constant product expansion have been foundational to recent advances in quantum locally testable codes (qLTCs) based on cubical complexes. Informally, product expansion says that any low-weight parity check of the tensor code can be written as the sum of few parity checks of each component code in each dimension.

We give a proof of the existence of two-sided product expanding codes which, unlike the earlier result of Kalachev and Panteleev (FOCS, 2025), does not rely on explicit constructions of asymptotically optimal locally testable codes (c^3 -LTCs). For every fixed number of component codes and dimensions whose rates are bounded away from zero and one, we show that independent random linear codes over a sufficiently large prime field have constant two-sided product expansion with probability tending to one. The tradeoff is that our proof requires the characteristic to grow with the block length, while [KP25] takes extensions of $\mathbb{F}_2$.

To replace the use of c^3 -LTCs, we develop several new techniques that we view as interesting in their own right. Instead of working directly over finite fields, we work over the reals and use random Rademacher matrices for the generator matrices of the component codes. From here, we show that the extendability of ε -closed sets can be reduced to controlling the operator norm of sparse restrictions of carefully chosen Gram matrices. Applying the trace power method to bound this norm reduces to bounding the number of possible labelings of certain closed walks on bipartite graphs, which we bound using the sparsity of the operators and bounds on the number of equivalence classes of weak Wigner words from Anderson and Zeitouni (Probab. Theory Relat. Fields., 2006), which were originally applied to band matrices. Transferring our result from the reals to finite fields is what causes our result to require a large field characteristic.

Due to our techniques not relying on explicit c^3 -LTCs, we believe that they form a promising starting point towards showing the existence of product expanding tensor codes where the component codes have non-trivial automorphism groups and coboundary expanding codes that could be used as the local codes of non-cubical complexes, such as simplicial complexes. Overall, we offer a different random model and novel techniques for studying the product expansion of random linear codes.

*Supported by NSF Award CCF-2514586.

†Supported by NSF GRFP grant DGE – 2139899 and NSF Award CCF-2514586.

Contents

1	Introduction	3
1.1	Motivation	4
1.2	Our results	6
2	Proof Overview	7
2.1	Reducing to showing a fixed ε -closed set is inner-generated	7
2.2	Proving an ε -closed set is inner-generated by the norm of its plane-interaction matrix	9
2.3	Bounding the norm of the plane-interaction matrix by the expected value of Rademacher chaos polynomials	11
2.4	Bounding the expected value of Rademacher chaos polynomials using weak Wigner words	13
3	Preliminaries	15
3.1	Lines, planes, and sparseness	15
3.2	Maximal planes	16
3.3	Anchor sets	17
3.4	Tensor codes and related notions	19
3.5	Miscellanea	20
3.5.1	Linear algebra facts	20
3.5.2	A useful estimate on the distribution of primes	20
4	(Co)Restriction Maps and Atoms	21
4.1	The tensor code, restrictions, and extendability	21
4.2	The check-product code, corestrictions, and inner-generatedness	23
4.3	Putting it all together	25
4.4	Atom decomposition	27
5	Two-Sided Product Expansion Over Random Codes	29
6	Injectivity of the Corestriction Map Via the Plane-Interaction Matrix	30
6.1	Interactions between the corestriction maps of planes	31
6.2	Constructing the plane-interaction matrix	34
6.3	Injectivity of τ_M over the reals	34
6.3.1	Injectivity of τ_M from the operator norm of K	35
6.4	From the reals to finite fields	36
7	Bounding the Operator Norm of the Plane-Interaction Matrix	41
7.1	Decomposition of the plane-interaction matrix	41
7.1.1	Directional block decomposition	42
7.1.2	Decomposition over sets of mutually fixed coordinates	42
7.2	Bounding the expected operator norm of restricted directional blocks	44
7.3	Deferred proofs	45
8	Bounding the Operator Norm of Sparse Tensor Gram Matrices	50
8.1	Derivation of the sign monomials	50
8.2	Graph view of the sign monomials	52
8.3	Weak Wigner words	53
8.3.1	Equivalence and canonical representatives	54
8.4	Weak Wigner words in the bipartite setting	55
8.4.1	Equivalence classes from a specified deficiency	55
8.5	Bounding the number of equivalence classes	56
8.6	Bounding the size of an equivalence class of a word using sparsity	57
8.7	Putting it all together	61
9	Open Questions	62

1 Introduction

A *linear error-correcting code* is a subspace $\mathcal{C} \subseteq \mathbb{F}_q^n$ in which any pair of distinct elements $c \neq c' \in \mathcal{C}$ satisfy $\frac{1}{n} |c - c'| \geq \delta$, where δ is known as the *relative distance* of the code. One desirable property of a linear error-correcting code is for it to be *locally testable*: a randomized function $T : \mathbb{F}_q^n \rightarrow \{\text{Accept}, \text{Reject}\}$ such that $T(c) = \text{Accept}$ for all $c \in \mathcal{C}$ and $\Pr[T(w) = \text{Reject}] \geq \frac{s}{n} \cdot \text{dist}(w, \mathcal{C})$, where $\text{dist}(w, \mathcal{C})$ is the Hamming distance of w to $\mathcal{C}$. The *locality* of T is how many coordinates of w it reads and s is its *soundness*. A family of locally testable codes $(\mathcal{C}_n)_{n \in \mathbb{N}}$ is known as a c^3 -LTC if its parameters are asymptotically optimal, i.e., its relative distance δ , rate $\rho = \dim(\mathcal{C})/n$, soundness s and locality are all constant. One can also define an analogous notion of local testability for *quantum* codes; codes satisfying this property are referred to as qLTCs. Classical LTCs and qLTCs have found important applications throughout theoretical computer science [GVSV06; BGHMRS15; EH17; KM25; NP25]; in particular, LTCs are a key component in the original proof of the PCP theorem.

Recent constructions of classical LTCs [DELLM22; PK22], qLTCs [DLV24], and quantum low-density parity-check codes (qLDPCs) [PK22; DHLV22], all use variants of the same strategy whereby a high-dimensional expander (HDX) is used to lift a small, constant-sized code with good product expansion, as in Definition 1.1, to create a large code with good parameters. These constructions fall into the framework of sheaf codes, a generalization of Tanner codes, first formally defined by [PK24]. Just like a Tanner code, a sheaf code only needs two pieces of information: a large, explicit (possibly high-dimensional) graph with structured local neighborhoods and a collection of constant-sized “local” codes that can be placed at these neighborhoods on the graph.

Sheaf codes have already been implicitly used extensively for constructing linear codes [DK17; BE21; Mes18; PK22; DLV24; DLZ23; FK24; KP23]. The main benefit of the sheaf code framework is that it is flexible enough to encompass many different constructions of linear codes while being endowed with a formal cohomology theory that allows for the study of the cosystolic distance and cocycle expansion of the resulting codes, which are directly related to the relative distance and soundness of the resulting locally testable code.

The HDXs used in the current qLDPC [PK22] and qLTC [DLV24; CHLT26; GJ26; BLN26] constructions is a cubical complex which requires that the local codes they utilize be tensor codes; consequently, there has been much recent work on the coboundary expansion of tensor codes. The equivalent notion of coboundary expansion used in these works, and that we will focus on here, is that of *product expansion*. For two component codes of linear distance, product expansion is equivalent, up to a change in parameters, to robust and agreement testability. For more than two codes, it is a strictly stronger requirement [KP23; Kal23].

Product expansion Let $\mathcal{C}_1, \dots, \mathcal{C}_D \subseteq \mathbb{F}_q^n$ be linear codes and view words of length n^D as arrays indexed by the grid $[n]^D$. For $i \in [D]$, let

$$\mathcal{C}^{(i)} = (\mathbb{F}_q^n)^{\otimes(i-1)} \otimes \mathcal{C}_i \otimes (\mathbb{F}_q^n)^{\otimes(D-i)}.$$

Thus, a word belongs to $\mathcal{C}^{(i)}$ precisely when its restriction to each line parallel to the i -th axis belongs to $\mathcal{C}_i$. We define the *tensor code* and the *check-product code*, respectively, as

$$\mathcal{C}_1 \otimes \dots \otimes \mathcal{C}_D = \bigcap_{i=1}^D \mathcal{C}^{(i)}, \quad \mathcal{C}_1 \boxplus \dots \boxplus \mathcal{C}_D = \sum_{i=1}^D \mathcal{C}^{(i)}.$$

An easy way to visualize the tensor of two codes $\mathcal{C}_1 \otimes \mathcal{C}_2$ is by labeling the grid $[n]^2$ by field elements of $\mathbb{F}_q$ such that each row is a codeword of $\mathcal{C}_2$ and each column is a codeword of $\mathcal{C}_1$. The check-product code is related to the tensor code by duality:

$$(\mathcal{C}_1 \otimes \dots \otimes \mathcal{C}_D)^\perp = \mathcal{C}_1^\perp \boxplus \dots \boxplus \mathcal{C}_D^\perp.$$

A *line-codeword* is a word supported on a single axis-parallel line whose restriction to that line belongs to the corresponding component code. Every word in the check-product code is created by a sum of such

words. Product expansion then asks whether every check-product codeword can be written using few nonzero line-codewords relative to its Hamming weight.

For $a \in \mathbb{F}_q^{[n]^D}$, let $|a|_H$ denote its Hamming weight and set $\|a\|_H = |a|/n^D$. Let $|a|_i$ be the number of lines parallel to the i -th axis on which a is nonzero, and set $\|a\|_i = |a|_i/n^{D-1}$.

Definition 1.1 (Product expansion). A collection $(\mathcal{C}_1, \dots, \mathcal{C}_D)$ is ν -product expanding if every $c \in \mathcal{C}_1 \boxplus \dots \boxplus \mathcal{C}_D$ has a representation $c = \sum_{i=1}^D a_i$, where $a_i \in \mathcal{C}^{(i)}$, such that

$$\nu \sum_{i=1}^D \|a_i\|_i \leq \|c\|_H.$$

The collection is *two-sided ν -product expanding* if both $(\mathcal{C}_1, \dots, \mathcal{C}_D)$ and $(\mathcal{C}_1^\perp, \dots, \mathcal{C}_D^\perp)$ are ν -product expanding.

Equivalently, a word of Hamming weight w in the check-product code has a decomposition using at most $w/(\nu n)$ nonzero line-codewords. Conceptually, product expansion controls how the component codes interact and forces them to act independently of each other. In applications, it is necessary for ν to remain a positive constant as the component length n grows, with D fixed.

Kalachev and Panteleev [KP23] proved that pairs of random linear codes have constant two-sided product expansion, while [DHLV22] showed that pairs of random linear codes are robust. However, the almost-good qLTC construction of [DLV24] and subsequent work on quantum codes [GG25; NP25] require product expansion for collections of more than two codes. The breakthrough work of Kalachev and Panteleev [KP25] showed that random linear codes over sufficiently large fields have constant two-sided product expansion for any fixed number of component codes. In particular, they obtained the following existence theorem.

Theorem 1.2 (Corollary 3 of [KP25]). Fix $D \geq 1$ and intervals $I_1, \dots, I_D \subseteq (0, 1)$ with nonempty interior. There exists $\nu > 0$ such that, for all sufficiently large n , there are codes $\mathcal{C}_1, \dots, \mathcal{C}_D \subseteq \mathbb{F}_{2^{t_n}}^n$, where $t_n = (n+1)^D$, with $\dim(\mathcal{C}_i)/n \in I_i$ such that the collection is two-sided ν -product expanding.

Curiously, Kalachev and Panteleev’s proof of this result relies on the existence of c^3 -LTCs, as established by the explicit, deeply structured constructions of [DELLM22; PK22]. For the reader who has seen many probabilistic existential proofs of combinatorial objects, it should be quite surprising (and interesting!) that the proof of [KP25] relies on an explicit construction to obtain an existential result. The main contribution of this work is a “purely” probabilistic proof of (a variant of) Theorem 1.2, thereby providing new insight into the phenomenon of product expansion for random codes, and new analytical tools that may allow us to “go beyond” random codes. We expand upon these motivations next.

1.1 Motivation

Besides pure curiosity, there are two main motivations for this work. First, to obtain good two-sided product expanding codes where the component codes have non-trivial automorphism groups, which are relevant in the constructions of qLTCs [PK24] and applicable whenever tensor codes are being used as local codes on cubical complexes. Second, to prove the existence of good two-sided coboundary expanders that can be used as the local codes of other complexes, like simplicial HDXs. We elaborate on both of these goals below, but, before doing so, we recall the strategy of [KP25] to explain the barriers in adapting their argument.

The technique of [KP25] We begin by briefly recalling the role of *extendability*. Given a tensor code $\mathcal{C} = \mathcal{C}_1 \otimes \dots \otimes \mathcal{C}_D$ and a set $M \subseteq [n]^D$, an assignment $x \in \mathbb{F}_q^M$ is locally valid if for every axis-parallel line ℓ entirely contained in M , $x|_\ell \in \mathcal{C}_i$ for the corresponding component code $\mathcal{C}_i$. We call M *extendable* in $\mathcal{C}$ if *every* such assignment extends (not necessarily uniquely) to a codeword of $\mathcal{C}$ on the whole grid. A dual view to this is that the parity checks contained in M impose all the conditions needed for a global extension. In particular, if M contains no entire line, all assignments are locally valid, so extendability means that any arbitrary assignment of elements from $\mathbb{F}_q$ on M extends.

The sets relevant to product expansion are the ε -closed sets: sets for which every axis-parallel line is either contained in M or meets it in at most εn points. By duality, extendability of M in $\mathcal{C}$ is equivalent to the following property: every word in $\mathcal{C}^\perp = \mathcal{C}_1^\perp \boxplus \dots \boxplus \mathcal{C}_D^\perp$ supported on M can be generated by dual line-codewords on lines entirely contained in M . Such sets are called *inner-generated*. Kalachev and Panteleev show that if every ε -closed set has this property, then the collection $(\mathcal{C}_1^\perp, \dots, \mathcal{C}_D^\perp)$ has constant product expansion, with a parameter depending only on ε and D [KP25]. Applying the statement to both a collection and its dual gives two-sided product expansion.

Their proof obtains the necessary extendability from *maximally extendable* (ME) codes. Within a specified family, an ME code has the property that every set extendable in any member of the family is also extendable in the ME code. The family used in [KP25] is parametrized by parity-check matrices $H_i \in \mathbb{F}_2[\mathbf{X}]^{(n-k_i) \times n}$, with a separate variable for every entry. We call a setting of the variables in $\mathbf{X}$ a *specialization*, which in turn gives concrete codes $\mathcal{C}_1, \dots, \mathcal{C}_D$ as defined by these parity check matrices. Specializations of full row rank give component codes of dimensions k_i , and random specializations over a sufficiently large extension field give an ME tensor code with high probability. While this setup is equivalent to simply taking a random parity-check matrix, we introduce this notion of parity-check matrices parametrized as polynomials from [PK24] since they allow us to write the determinants of compositions of these matrices as polynomials in $\mathbf{X}$ themselves, and are more flexible, as we elaborate below.

More generally, [PK24] proves an analogous statement for polynomially parametrized families of sheaf codes, where the polynomials used to define the parity check matrices can be more elaborate than simply having a distinct variable for each entry. Importantly, an ME tensor code can inherit, with a constant loss in parameters, the product expansion of any code in its parametrized family.

Nevertheless, one still needs a member of the family from which the ME code can inherit product expansion, which we call a *witness*. Kalachev and Panteleev obtain this witness from the explicit c^3 -LTCs of [DELLM22], namely LTCs with constant rate, relative distance, and query complexity. After adapting lengths and rates, they show that suitable collections derived from these LTCs have constant product expansion from which ME codes can inherit their own product expansion. Lastly, the observation that the dual of a uniformly random linear code is itself uniformly random lets them union bound the two failure probabilities to obtain two-sided product expansion.

Towards product expanding codes with symmetries Even with such two-sided product expanding codes, the qLTCs of [DLV24] have constant rate but inverse-polylogarithmic relative distance and soundness. In [PK24], the predecessor to [KP25], the authors suggest using the Ramanujan cubical complexes of Rungtanapirom, Stix, and Vdovina [RSV19] as a possible route to asymptotically good qLTCs. The proposed approach calls for local codes that are equivariant with respect to the symmetries of the cubical complex. That is, the collection of component codes must be two-sided product expanding, and the automorphism group of each component must also contain the corresponding local permutation group of the complex. The random model of [KP25] cannot enforce this symmetry requirement since random linear codes have trivial automorphism groups with high probability [LWZ26].

We encounter two immediate barriers to using this technique to show the existence of two-sided product expanding codes with prescribed nontrivial symmetries. First, the parametrization of the parity-check matrices allows arbitrary component codes $\mathcal{C}_i$ instead of just ones with the desired automorphism groups. One can try to restrict the parametrization so that the resulting codes have the required symmetries, and, when such a polynomial parametrization is available, [PK24] guarantees an ME code within that family. However, this brings us to the second barrier: the c^3 -LTC construction of [DELLM22] does not provide the prescribed automorphism groups. Consequently, the extension witness obtained from these LTCs is not known to lie in the restricted family, and maximal extendability in that family has no established witness from which to inherit the required properties.

In our work, we remove the dependence on c^3 -LTCs that gives rise to this second barrier. We prove the necessary extension properties directly, using random matrices over the reals and then transferring the result to finite fields. This gives a different random model and novel techniques for studying the product expansion of random linear codes.

Towards the coboundary expansion of general sheaf codes As we elaborate more in Section 9, our techniques are flexible enough that they provide a new starting point to analyze the coboundary expansion of more general sheaf codes, as defined by [PK24], instead of just tensor codes. In particular, recent constructions of classical [DLZ23] and quantum [GK25] codes on simplicial high-dimensional expanders require local codes that do not have a tensor product structure and currently must use Reed-Solomon or Reed-Müller codes. We believe our techniques could lend themselves to analyzing the sheaf codes for the required local codes of these constructions. We also anticipate that our new framework of analyzing over the reals the desired properties of a code generated by random Rademacher matrices will be useful more broadly.

1.2 Our results

We obtain the following theorem over a prime field. In the statement, a uniformly random code of dimension k means a uniformly random k -dimensional subspace of the ambient space.

Theorem 1.3 (Simplified version of Theorem 5.1). Fix $D \geq 1$ and $0 < \rho < 1/2$. There are constants $\nu > 0$ and n_0 , depending only on D and ρ , such that the following holds. For every $n \geq n_0$ and integers $k_1, \dots, k_D$ satisfying $\rho n \leq k_i \leq (1 - \rho)n$, there is a prime

$$p = O\left(2^{n^D} n^D \log n\right)$$

such that independent uniformly random codes $\mathcal{C}_i \subseteq \mathbb{F}_p^n$ of dimensions k_i are two-sided ν -product expanding with probability at least $1 - 1/\log(n + 1)$. The constant in the field-size bound depends only on D .

The full statement, including an explicit expansion parameter and a sufficient threshold for n , appears in Theorem 5.1. In particular, we obtain codes whose rates lie in any intervals with nonempty interior in $(0, 1)$. Thus, we recover the existence of two-sided product expanding codes without the use of c^3 -LTCs. Unlike Theorem 1.2, our proof requires the characteristic to grow with n and does not establish the result in characteristic two. In applications, such as [DLV24], this is not a serious issue: it simply results in a code over some constant-sized alphabet instead of the binary alphabet.

Our main task is to prove extendability for every ε -closed set. We first work over the reals, using independent Rademacher entries for the generator matrices. By dualizing and showing that the given ε -closed set is inner-generated, we reduce the problem to showing a certain map has norm less than one.¹

To bound its norm, we apply the trace power method. Leveraging the unique properties of Rademacher random variables turns the norm estimate into a problem of counting labelings of closed walks on bipartite graphs. The work of Anderson and Zeitouni [AZ06] bounds the number of equivalence classes of such walks, known as weak Wigner words. We then bound the size of each equivalence class using the fact the original set is ε -closed. This argument shows that each ε -closed set has a collection of real generator matrices for which it is inner-generated for the corresponding collection of codes. We transfer this result from the reals back to finite fields and to a single collection of codes, instead of a different collection of codes for each ε -closed set, by a Schwartz-Zippel argument similar to that of [KP25]. We explain these arguments in Section 2.

Concurrent work In the last stages of preparing our manuscript, a number of preprints have been posted that give constructions of asymptotically optimal qLTCs. The works [GJ26; BLN26; CHLT26] all use cubical complexes with local Reed-Solomon codes to construct asymptotically optimal qLTCs. [CLRTT26] constructs asymptotically optimal qLTCs under a conjecture on the existence of certain lossless cubical complexes while not employing local codes at all.

Organization. After the proof overview, Section 3 introduces necessary preliminaries and our framework of planes and anchor sets. In Section 4 we derive the restriction and corestriction maps of tensor and check-product codes, respectively, and define the atom decomposition. Section 5 proves our main result based on the injectivity of a universal corestriction map. Then, Section 6 reduces this injectivity to an operator norm

¹Specifically, this is the plane-interaction matrix as defined in Equation (1).

bound over the reals and gives the transfer to finite fields. Finally, Section 7 reduces this to a norm bound on sparse tensor matrices, which we prove in Section 8 using weak Wigner words. We finish by proposing some future directions in Section 9.

Acknowledgements

N.R. would like to thank Robert Kleinberg for conversations that inspired him to consider conducting the analysis in this paper over the reals instead of over a finite field. We thank the Simons Institute for the Theory of Computing at UC Berkeley for providing a stimulating research environment. Part of this work was done during visits by the authors to the Institute.

AI Methodology

The ideas, proofs and exposition in this work are almost entirely the product of human effort. ChatGPT Pro 5.6 and 6 were used to help check correctness and for copy editing, and suggested the connection to the works of [Wig55; FK81; AZ06].

The authors take responsibility for all results, content, and the presentation of this paper.

2 Proof Overview

We now explain the main ideas behind Theorem 1.3. Fix the number of codes D and a constant $\rho > 0$, and assume that each code has dimension between ρn and $(1 - \rho)n$. Our goal is to prove ν -product expansion with $\nu = \Omega_{D,\rho}(1)$.

Our proof has four main parts:

1. Using Lemma 19 and the existence of maximally extendable codes from [KP25], we reduce product expansion to showing that, for a fixed ε -closed set $M \subseteq [n]^D$, there is a collection of codes (over the reals) for which M is inner-generated.
2. We switch to working over $\mathbb{R}$ and observe that, for a fixed ε -closed M , it is inner-generated if the *plane-interaction* matrix K_M , a centered Gram matrix built using the structure of M , has norm less than 1.
3. By cleverly decomposing K_M into a constant number of blocks with easily describable norms and taking the generator matrices $G_i \in \mathbb{R}^{k_i \times n}$ of the codes $\mathcal{C}_1, \dots, \mathcal{C}_D$ to have Rademacher entries, we can bound the norm of K_M by the expected value of Rademacher chaos polynomials.²
4. Finally, we bound the expected value of these polynomials by using the fact that M is ε -closed and tools from [AZ06] for weak Wigner words, ultimately giving us that $\|K_M\|_{\text{op}} < 1$ so M is inner-generated.

We now elaborate on these steps in detail.

2.1 Reducing to showing a fixed ε -closed set is inner-generated

As mentioned in the introduction, two important results from [KP25] are that the product expansion of a collection of codes is implied by every ε -closed set being inner-generated (Lemma 19) and that ME codes exist. Kalachev and Panteleev used maximal extendability by showing the fact that a collection of c^3 -LTCs from [DELLM22] has every ε -closed set inner-generated, so an ME code would have each of these sets inner-generated as well.³ Instead of showing that there exists a single collection of codes with all ε -closed sets

²A *Rademacher chaos* is a homogeneous multivariate polynomial with domain $\{-1, 1\}$.

³While the definition of maximally extendable codes is with regards to the extendability of a set, since extendable sets and inner-generated sets are dual notions to each other, we can equivalently consider their result on the existence of maximally extendable codes as one for maximally inner-generated codes.

inner-generated, we show that for each fixed ε -closed set M there is a collection of codes for which M is inner-generated. Then an ME code can inherit the inner-generatedness of each ε -closed set from different collections of codes instead of a single one. The difference in these strategies is summarized in Figure 1

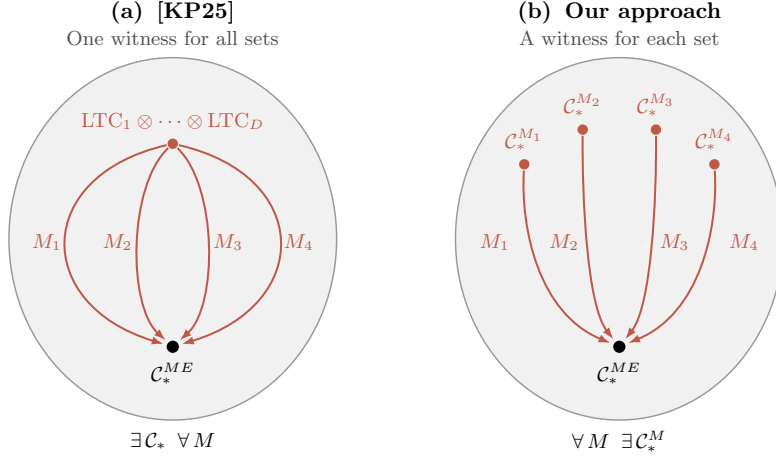


Figure 1: Comparison of strategies between [KP25] in (a) and our work in (b). Each gray ellipse represents the family of polynomially parametrized codes with C_*^{ME} as the maximally extendable code in the family. In (a), each ε -closed set M_i is extendable in the tensor product of LTCs, so C_*^{ME} can use this product as its witness of extendability for all ε -closed sets. In (b), each ε -closed M_i has a different collection of codes $C_*^{M_i}$ in which it is extendable. Here, C_*^{ME} still inherits the extendability of all ε -closed sets, but from different witnesses.

In reality, our proof technique has one major difference that does not allow us to apply the aforementioned results from [KP25] in a black-box manner. For a fixed ε -closed set $M \subseteq [n]^D$, the collection of codes that we show witness the fact that M is inner-generated is defined by real generator matrices $G_i \in \mathbb{R}^{k_i \times n}$ instead of over $\mathbb{F}_{2^t}$ as in [KP25]. Loosely speaking, to transfer this result from the reals to a finite field, we use the fact (shown in the next step) that M is inner-generated if and only if a certain matrix depending on the generator matrices $(G_i)_{i \in [D]}$ is full rank. By taking the determinant polynomial $f_M \in \mathbb{Z}[\mathbf{X}]$ of a non-singular minor in this matrix—which we know is not identically the zero polynomial since we have real generator matrices that are not a zero of f_M —we can condition M being inner-generated on choosing generator matrices $G_i \in \mathbb{F}_p^{k_i \times n}$ such that they are not a zero of f_M . Finally, bounding $\deg f_M$ as $O(n^D)$ and noticing that there are at most 2^{n^D} sets $M \subseteq [n]^D$, we use a Schwartz-Zippel argument to show that taking $p = O(2^{n^D} n^D)$ will guarantee some setting of the generator matrices for which $\prod_{\varepsilon\text{-closed } M} f_M((G_i)_{i \in [D]})$ is non-zero, meaning that all ε -closed sets are inner-generated.

We elide the details of this argument since the techniques and, in particular, the application of Schwartz-Zippel, are very similar to those of [KP25]. Our new observations and techniques are contained in the subsequent steps of our proof. The takeaway from this reduction is that, to prove Theorem 1.3, we just need to show the following:

Lemma 2.1. Let $D \geq 1$, $\rho \in (0, 1/2)$, $\varepsilon > 0$ be small enough, and $n \in \mathbb{N}$ be large enough. Take integers $k_1, \dots, k_D$ with $\rho n \leq k_i \leq (1 - \rho)n$ and fix an ε -closed set $M \subseteq [n]^D$. Then there exist full-rank generator matrices $G_1, \dots, G_D$ with $G_i \in \mathbb{R}^{k_i \times n}$ such that M is extendable for $\mathcal{C}_1 \otimes \dots \otimes \mathcal{C}_D$ or, equivalently, inner-generated for $\mathcal{C}_1^\perp \boxplus \dots \boxplus \mathcal{C}_D^\perp$.

2.2 Proving an ε -closed set is inner-generated by the norm of its plane-interaction matrix

Random model We begin by defining the random Rademacher model we will use in the proof of Lemma 2.1. To sample $G_i \in \mathbb{R}^{k_i \times n}$, we take it to have columns $g_i(j) \in \mathbb{R}^{k_i}$ indexed by $j \in [n]$ where $g_i(j) = k_i^{-1/2}(\sigma_{1,j}^{(i)}, \dots, \sigma_{k_i,j}^{(i)})^\top$ and each $\sigma_{r,j}^{(i)} \sim \{-1, 1\}$. By construction, each column has unit norm. While this setup allows for the generator matrices to not necessarily be full-rank, it is easy to show that they will all be full-rank with high probability, so we will just concern ourselves with showing that M is inner-generated.

Our goal in the rest of this step is to show that M is inner-generated for the codes from $G_1, \dots, G_D$ if *plane-interaction matrix* K_M , constructed later in this step, has $\|K_M\|_{\text{op}} < 1$.

Reducing to the injectivity of the corestriction map Take dual codes $\mathcal{B}_i = \mathcal{C}_i^\perp$ and define the check-product code $\mathcal{B} = \mathcal{B}_1 \boxplus \dots \boxplus \mathcal{B}_D$. Recall that M is inner-generated when every codeword of $\mathcal{B}$ supported only on M can be expressed as a sum of line-codewords on lines contained within M . Using the generator matrices $(G_i)_i$, we can define the natural map

$$\tau : \mathbb{K}^{[n]^D} \rightarrow \bigotimes_{i=1}^D \mathbb{K}^{k_i} =: U_{[D]}, \quad e_x \mapsto \bigotimes_{i=1}^D g_i(x_i)$$

where $\mathbb{K}$ is any field⁴ and e_x is the standard basis indexed by $x \in [n]^D$. By construction, the kernel of τ is exactly $\mathcal{B}$. This is close to the map we want to construct indicating whether M is inner-generated, but no information about M is present in τ , so it should not be surprising that we have to modify it to get our desired map. Observe that if we restrict τ to words only supported on M and M contains a line of $[n]^D$, say in the i -th direction, then τ is not injective because all the codewords from $\mathcal{B}_i$ on that line are supported on M and in the kernel of τ . However, for M to be inner-generated, we exactly want to allow words supported on M generated by the sum of line-codewords. Formally, this is the set $R_M^{\text{line}} = \sum_{\ell \in \mathcal{L}(M)} \mathcal{B}_\ell$ where $\mathcal{L}(M)$ is the set of axis-parallel lines contained in M and, for an i -directional line $\ell \in \mathcal{L}(M)$, we let $\mathcal{B}_\ell$ be $\mathcal{B}_i$ included into $\mathbb{K}^{[n]^D}$ by placing it on ℓ and padding with 0's elsewhere. We can then consider the induced quotient map

$$\tau_M : \mathcal{G}(M) := \mathbb{K}^M / R_M^{\text{line}} \rightarrow U_{[D]},$$

where we are using $\mathbb{K}^M$ as shorthand for words in $\mathbb{K}^{[n]^D}$ supported on M . Recalling that $\ker(\tau) = \mathcal{B}$, the kernel of τ_M is exactly

$$\ker(\tau_M) \cong \frac{\mathbb{K}^M \cap \ker(\tau)}{R_M^{\text{line}}} = \frac{\mathbb{K}^M \cap \mathcal{B}}{\sum_{\ell \in \mathcal{L}(M)} \mathcal{B}_\ell}.$$

Therefore, when τ_M is injective we get that $(\mathbb{K}^M \cap \mathcal{B}) / (\sum_{\ell \in \mathcal{L}(M)} \mathcal{B}_\ell) = \ker(\tau_M) = 0$, meaning that $\mathbb{K}^M \cap \mathcal{B} = \sum_{\ell \in \mathcal{L}(M)} \mathcal{B}_\ell$, which is exactly the definition of M being inner-generated.

An instructive special case to consider is when M is ε -closed but contains no full lines, a property that we call ε -sparse. Consequently, $R_M^{\text{line}} = \{0\}$ and our goal becomes showing that the columns of $G_1 \otimes \dots \otimes G_D$ indexed by M , that is $\otimes_{i \in [D]} g_i(x_i)$ for $x \in M$, are linearly independent. However, when M does contain full lines, we must handle the induced relations between them.

The local contributions of planes Instead of focusing on every single axis-parallel line in M individually, we will bundle these into planes when possible. Formally, an I -plane at direction $I \subseteq [D]$ anchored at $a \in [n]^{I^c}$

⁴While we will focus on the setting of $\mathbb{K} = \mathbb{R}$, these corestriction maps are well-defined over any field and, in particular, our ultimate goal of $\mathbb{K} = \mathbb{F}_p$.

is defined as $P = \mathcal{P}(I, a) = \{x \in [n]^D : \forall j \in I^c, x_j = a_j\}$. We assign to it a local space $U_P = U_I = \otimes_{i \in I} \mathbb{R}^{k_i}$. For a plane $Q = \mathcal{P}(J, b)$ containing P , so $J \supseteq I$, we define the corestriction map

$$\tau_{P \rightarrow Q} : U_P \rightarrow U_Q \quad u \mapsto u \otimes \bigotimes_{j \in J \setminus I} g_j(a_j),$$

with the tensor factors placed in the order of elements in $[D]$. In particular, we can always take $Q = [n]^D$ and consider the global corestriction map $\tau_P := \tau_{P \rightarrow [n]^D}$ which takes a local contribution from P and maps it to $U_{[n]^D}$.

The ε -closed set M can contain many planes—every I -plane in M is made up of n planes of dimension $|I| - 1$ also in M . We can simplify the structure of M by focusing on the *maximal* (by inclusion) planes, denoted as $\mathcal{P}(M)$, and the planes resulting from the intersection of maximal planes $\mathcal{P}^\cap(M) = \{\bigcap_{P \in \mathcal{S}} P : \emptyset \neq \mathcal{S} \subseteq \mathcal{P}(M), \bigcap_{P \in \mathcal{S}} P \neq \emptyset\}$. A useful combinatorial consequence of this viewpoint is that, for a fixed direction $I \subseteq [D]$, the set of anchors $\Omega_I(M) \subseteq [n]^{I^c}$ of the maximal I -planes in $\mathcal{P}(M)$ is ε -sparse because M is ε -closed. Indeed, if there were more than εn anchors from $\Omega_I(M)$ on a line in some direction $j \notin I$, then the ε -closure of M would force there to be a larger $(I \cup \{j\})$ -plane containing the anchored planes, contradicting their maximality. If we introduce intersections of planes to consider the I -planes in $\mathcal{P}^\cap(M)$, then the resulting anchor set $\Omega_I^\cap(M)$ is $(2^D \varepsilon)$ -sparse via a similar argument. In either case, we use the ε -closure of M to obtain sparseness guarantees for its anchor sets.

In particular, this viewpoint is helpful in dealing with the injectivity of τ_M because we can equivalently describe its domain $\mathcal{G}(M)$ as the direct sum of all U_P for $P \in \mathcal{P}^\cap(M)$ where every $u \in U_P$ is identified with its corestriction $\tau_{P \rightarrow Q}(u) \in U_Q$ whenever $P \subseteq Q$. Conceptually, this says that a contribution from a smaller plane contained in a larger one is equivalent to the same contribution from the larger one. This is shown formally in Lemma 4.8.

Removing interactions between planes via atoms In this equivalent definition of $\mathcal{G}(M)$, we see that a plane Q contained in another plane P contributes to the local space U_P by the image $\tau_{Q \rightarrow P}(U_Q) \subseteq U_P$. To tease apart how much of the local space U_P is controlled solely by P , we define the *atom* A_P of P as

$$U_{\mathcal{P}^\cap(M)} = \sum_{\substack{Q \in \mathcal{P}^\cap(M) \\ Q \subsetneq P}} \tau_{Q \rightarrow P}(U_Q) \quad \text{and} \quad A_P = U_{\mathcal{P}^\cap(M)}^\perp \subseteq U_P.$$

Importantly, since we repeat this procedure for every plane, the direct sum of all atoms $\bigoplus_{P \in \mathcal{P}^\cap(M)} A_P$ spans $\mathcal{G}(M)$, which we formally show in Lemma 4.14.

The main utility of this atom decomposition is that it makes clear the fact that non-intersecting planes are the only obstructions to M being inner-generated. In particular, we observe that the atoms of two distinct intersecting planes have orthogonal images in $U_{[D]}$, formally proved in Lemma 6.6. To see why, consider distinct intersecting planes P and Q with $R = P \cap Q$. For elements in their atoms $a_P \in A_P$ and $a_Q \in A_Q$, we can compute their inner product by using the adjoint of the corestriction map as

$$\langle \tau_{P \rightarrow [n]^D}(a_P), \tau_{Q \rightarrow [n]^D}(a_Q) \rangle = \langle \tau_{R \rightarrow P}^\dagger(a_P), \tau_{R \rightarrow Q}^\dagger(a_Q) \rangle.$$

Say, without loss of generality, that $R \subsetneq P$. Then, by the definition of atoms, $\tau_{R \rightarrow P}^\dagger(a_P) = 0$ making the entire inner product zero. In the case that P and Q are not distinct, so $P = Q$, it is not too hard to see that we get an isometry $\langle \tau_{P \rightarrow [n]^D}(a_P), \tau_{P \rightarrow [n]^D}(a'_P) \rangle = \langle a_P, a'_P \rangle$, as shown in Lemma 6.8.

The plane-interaction matrix We arrive at our main analytic tool, the plane-interaction matrix, by pairwise comparing the corestriction maps of every plane in $\mathcal{P}^\cap(M)$. In particular, we define $\Theta : \bigoplus_{P \in \mathcal{P}^\cap(M)} A_P \rightarrow U_{[D]}$ by $\Theta((a_P)_P) = \sum_P \tau_{P \rightarrow [n]^D}(a_P)$. Using our two observations about the inner product of intersecting planes above, we can decompose the Gram matrix of Θ as $\Theta^\dagger \Theta = I + K$ where the block matrix K is defined in its (P, Q) -th block by

$$K_{P,Q} = \begin{cases} \tau_{P \rightarrow [n]^D}^\dagger \tau_{Q \rightarrow [n]^D} & P \cap Q = \emptyset \\ 0 & P \cap Q \neq \emptyset. \end{cases} \quad (1)$$

This matrix K is exactly our plane-interaction matrix and visualized in Figure 2.

Technically, the domain of K is $\bigoplus_P A_P$, but expanding it to $\bigoplus_P U_P$ is convenient for later calculations and can only increase its operator norm, which we are trying to upper bound. From this construction, we get

$$\|\Theta((a_P)_P)\|^2 \geq (1 - \|K\|_{\text{op}}) \sum_P \|a_P\|^2,$$

meaning that $\Theta^\dagger \Theta$ is positive definite, and thus Θ injective, if $\|K\|_{\text{op}} < 1$. Finally, to conclude that τ_M is injective, meaning that M is inner-generated, we recall that the direct sum of all atoms spans $\mathcal{G}(M)$. With the observation that, after identifying $u \in U_P$ with $\tau_{P \rightarrow Q}(u) \in U_Q$ for $P \subseteq Q$, the resulting Θ on the quotient is equivalent to τ_M , we have that τ_M is injective if Θ is.

Consequently, our goal in the next couple steps is to bound the expected operator norm of K where our randomness comes from the Rademacher generator matrices $G_1, \dots, G_D$. We will eventually conclude that $\mathbb{E}[\|K\|_{\text{op}}] \leq O_{D,\rho}(1) \cdot \sqrt{2^D} \varepsilon$, so taking ε small enough with respect to the constants D and ρ can make this expectation less than $1/2$. Markov's inequality then gives us that there must exist some real setting of $G_1, \dots, G_D$ such that M is inner-generated for the resulting collection of codes.

2.3 Bounding the norm of the plane-interaction matrix by the expected value of Rademacher chaos polynomials

From the last step, we have reduced Lemma 2.1 to bounding $\mathbb{E}[\|K\|_{\text{op}}]$. A tempting first attempt at bounding $\|K\|_{\text{op}}$ is to use the convenient fact that it is defined via the blocks $K_{P,Q}$, so one could plausibly bound $\|K_{P,Q}\|_{\text{op}}$ for all pairs $P, Q \in \mathcal{P}^\cap(M)$ and then use the triangle inequality to bound $\|K\|_{\text{op}}$ by their sum. However, we would only be able to achieve a constant bound $\|K_{P,Q}\|_{\text{op}} = O_{D,\rho,\varepsilon}(1)$ for each block while it is possible to have $|\mathcal{P}^\cap(M)| = O_{D,\rho,\varepsilon}(n^D)$, meaning the triangle inequality would fail to provide the necessary bound.

Directional block decomposition To overcome this barrier, for two directions $I, J \subseteq [D]$, we chunk together all blocks $K_{P,Q}$ where P is an I -plane and Q is a J -plane to form the directional block $K_{I,J}$. There are at most 4^D many of these blocks, so bounding $\|K_{I,J}\|_{\text{op}}$ by a small enough constant and then using the triangle inequality now succeeds.

Focusing on a single directional block $K_{I,J}$, in Lemma 7.6 we again decompose it as the sum of matrices as $K_{I,J} = \sum_{\emptyset \neq E \subseteq F} K_{I,J}^E$ where $F = (I \cup J)^c$, the coordinates that are fixed by an anchor in both I -planes and J -planes. Notice that there are at most $2^{|F|} \leq 2^D$ terms in this sum, so bounding each term by a small enough constant and applying the triangle inequality is again a viable strategy. Using the product structure of the tensor code and some elementary, but tedious, algebraic manipulations that we omit here,

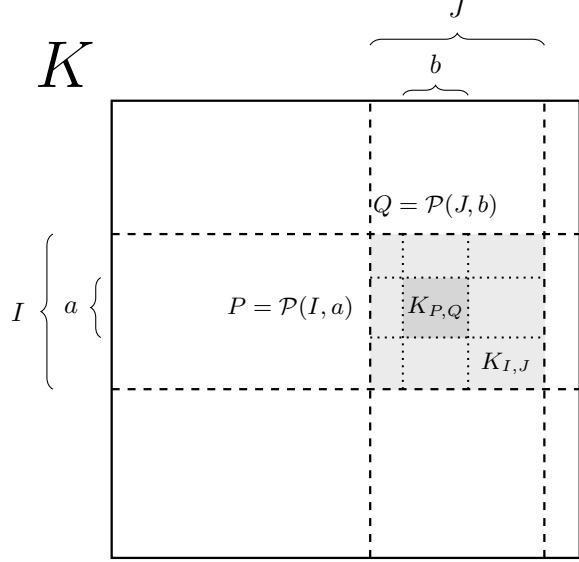


Figure 2: the plane-interaction matrix. The lighter gray box represents the directional block $K_{I,J}$, while the darker gray box is a plane block $K_{P,Q}$ with planes $P = \mathcal{P}(I, a)$ and $Q = \mathcal{P}(J, b)$.

in Lemma 7.8 we bound the operator norm of $K_{I,J}^E$ by

$$\|K_{I,J}^E\|_{\text{op}} \leq \left\| \bigotimes_{i \in I \setminus J} G_i \right\|_{\text{op}} \left\| \bigotimes_{j \in J \setminus I} G_j \right\|_{\text{op}} \max_{\lambda \in \Lambda} \left\| \Pi_{X_\lambda} \left(\bigotimes_{j \in E} D_j \right) \iota_{Y_\lambda} \right\|_{\text{op}},$$

where $D_j = G_j^\top G_j - \text{Id}_{\mathbb{R}^n}$, the indexing set Λ has cardinality at most n^ν where $\nu + |E| \leq D$, the sets X_λ, Y_λ are subsets of $[n]^E$, and Π_{X_λ} and ι_{Y_λ} are projection and inclusion (by 0's).

The first two operator norms are easy to bound by D and ρ using standard techniques from Gaussian matrices. The third factor is where we focus our attention on for the rest of this step and the next. The two sets X_λ, Y_λ are derived by taking an E -plane, indexed by λ , from $\Omega_I^\cap(M)$ and $\Omega_J^\cap(M)$, respectively. Consequently, they are both $(2^D \varepsilon)$ -sparse in $[n]^E$. Taking $T_\lambda = \Pi_{X_\lambda} \left(\bigotimes_{j \in E} D_j \right) \iota_{Y_\lambda}$, we will show, as stated in Proposition 8.1, that

$$\mathbb{E} \left[\max_{\lambda \in \Lambda} \|T_\lambda\|_{\text{op}} \right] \leq 2 \left(\frac{4}{\rho} \right)^{|E|} \sqrt{2^D \varepsilon}, \quad (2)$$

which we note can be made a small constant by taking ε small enough.

Deriving the Rademacher chaos polynomials To focus on a single $\lambda \in \Lambda$, we fix a pair of η -sparse subsets $X, Y \subseteq [n]^q$, where $\eta = 2^D \varepsilon$, and let $T = \Pi_X \left(\bigotimes_{j=1}^q D_j \right) \iota_Y$. We apply the trace-power method for an integer $m \in \mathbb{N}$, that we will set later, to get $\|T\|_{\text{op}}^{2m} \leq \text{tr}((TT^\top)^m)$. Expanding out this trace yields the expression

$$\text{tr}((TT^\top)^m) = \sum_{\substack{x_1, \dots, x_m \in X \\ y_1, \dots, y_m \in Y}} \prod_{j=1}^q \prod_{t=1}^m D_j(x_{t,j}, y_{t,j}) D_j(y_{t,j}, x_{t+1,j}), \quad x_{m+1} = x_1.$$

Recalling the definition of D_j , we can write a single entry of D_j as $D_j(a, b) = \mathbf{1}\{a \neq b\} k_j^{-1} \sum_{r=1}^{k_j} \sigma_{r,a}^{(j)} \sigma_{r,b}^{(j)}$. Plugging this in, as we do in Corollary 8.3 gives us

$$\text{tr}((TT^\top)^p) = \sum_{\substack{x_1, \dots, x_m \in X \\ y_1, \dots, y_m \in Y}} \sum_{\substack{r_j \in [k_j]^{2m} \\ j \in [q]}} \left(\prod_{j=1}^q k_j^{-2m} \right) \left(\prod_{j=1}^q \Phi_j(x, y, r_j) \right), \quad (3)$$

where $x = (x_1, \dots, x_m)$, $y = (y_1, \dots, y_m)$, and we define the monomials of signs

$$\Phi_j(x, y, r_j) = \prod_{t=1}^m \left(\sigma_{r_{j,2t-1}, x_{t,j}}^{(j)} \cdot \sigma_{r_{j,2t-1}, y_{t,j}}^{(j)} \cdot \sigma_{r_{j,2t}, y_{t,j}}^{(j)} \cdot \sigma_{r_{j,2t}, x_{t+1,j}}^{(j)} \cdot \mathbf{1}\{x_{t,j} \neq y_{t,j} \text{ and } y_{t,j} \neq x_{t+1,j}\} \right). \quad (4)$$

This expression for the trace is a sum of homogeneous multivariate polynomials with the domain of each variable being $\{-1, 1\}$, alternatively known as a Rademacher chaos.⁵ Now, we finally leverage the unique nature of Rademacher random variables to easily compute the expectation $\mathbb{E}_\sigma[\Phi_j(x, y, r_j)]$ as 1 whenever each sign appears an even number of times⁶ or 0 otherwise, since just one random sign is enough to randomize the entire monomial to zero mean.

Recall that our goal is to upper bound the expectation of Equation (3), so we would like to bound the number of $\mathbb{E}_\sigma[\Phi_j(x, y, r_j)]$ that are deterministically 1, which we do in the next step.

⁵The only slight deviation from the formal definition of a Rademacher chaos that we have here is the indicator functions in each monomial, but we do not use any off-the-shelf results on Rademacher chaoses in which this would be an issue; in fact, we significantly leverage them later on.

⁶While satisfying each indicator function $\mathbf{1}\{x_{t,j} \neq y_{t,j} \text{ and } y_{t,j} \neq x_{t+1,j}\}$, which we will see is equivalent to the corresponding bipartite walk being non-left-backtracking.

2.4 Bounding the expected value of Rademacher chaos polynomials using weak Wigner words

Sign monomials as bipartite walks As we've already observed, our sign monomials $\Phi_j(x, y, r_j)$ have mean 1 exactly when each sign appears an even number of times in the product of Equation (4). Thus, all we must do is count how many choices of $x \in X^m$, $y \in Y^m$, and $r_j \in [k_j]^{2m}$ for $j \in [q]$ there are that yield a sign monomial with each sign appearing an even number of times. A useful way to visualize this is by writing out the relevant indices in the order they are read for a fixed monomial $\Phi_j(x, y, r_j)$ as

$$x_{1,j}, r_{j,1}, y_{1,j}, r_{j,2}, x_{2,j}, \dots, y_{m,j}, r_{j,2m}, x_{1,j}. \quad (5)$$

Each sequential pair in this list defines one of the $4m$ signs in Equation (4). Equivalently, we could view each $x_{a,j}$, $y_{a,j}$, and $r_{j,a}$ index as a vertex and this list as giving us a walk on the resulting graph, as seen in Figure 3. Since the list from Equation (5) always alternates between $x_{a,j}$ or $y_{a,j}$ and a $r_{j,a}$, it is in fact a walk on a bipartite graph: the bipartite graph with $[n]$ left vertices and $[k_j]$ right vertices where each edge represents a single sign.

We now make three observations regarding this view of the signs in $\Phi_j(x, y, r_j)$ as a walk on a bipartite graph. First, it is a closed walk by definition since $x_{1,j} = x_{m+1,j}$. Second, it is non-left backtracking, meaning that each step along the walk must have different left neighbors when entering and exiting a right vertex. This is due to the indicator variables in Equation (4). Third, each sign appears an even number of times if each edge used in the walk is traversed an even number of times. Counting the number of such closed walks on this bipartite graph thus gives us an upper bound on the number of sign monomials with expectation 1.

There are two pieces of information that define such a bipartite walk: its shape and the labels on its vertices. For example, the shape of the bipartite walk, if it had length 8, could be tracing the bipartite clique $K_{2,2}$ twice. The labeling of the walk determines where exactly we place this clique on the left $[n]$ and right $[k_j]$ vertices.

Our strategy for bounding the number of non-left-backtracking closed walks that traverse each edge an even number of times then is to use the series of works [Wig55; FK81; AZ06] to bound the number of shapes such a walk could take and then to use the η -sparsity of X and Y to bound the number of labelings.

Bounding the number of shapes with weak Wigner words The walks on bipartite graphs that we have described here, which can alternatively be described by a string as in Equation (5), are a special case of weak Wigner words, which we treat formally in Section 8. Anderson and Zeitouni in [AZ06] give a combinatorial bound on the number of shapes such a walk can take. More concretely, they bound the number of equivalence classes of these bipartite walks, where two walks are equivalent if there is a bijection between the strings representing the walks.

To apply their bound, we define $\text{wt}_L(j)$ and $\text{wt}_R(j)$ to be the number of distinct left and right vertices in the walk of $\Phi_j(x, y, r_j)$, and $\text{wt}(j) = \text{wt}_L(j) + \text{wt}_R(j)$. Since there are $4m$ edge traversals, and each edge is used at least twice, we have that $\text{wt}(j) \leq 2m+1$. We define the *deficiency* of a walk to be $\Delta_j = 2m+1 - \text{wt}(j)$, which represents how far the walk is from using as many vertices as it possibly could. If we let $\mathfrak{W}_m(\Delta_j)$ be the set of equivalence classes of even, non-left-backtracking, closed bipartite walks that traverse $4m$ edges with deficiency Δ_j , then the bound from [AZ06], that we state in a form useful to us in Corollary 8.26, says that $|\mathfrak{W}_m(\Delta_j)| \leq 16^m (4m)^{6\Delta_j}$. For example, in Figure 4 we explicitly write out all equivalence classes in $\mathfrak{W}_2(1)$.

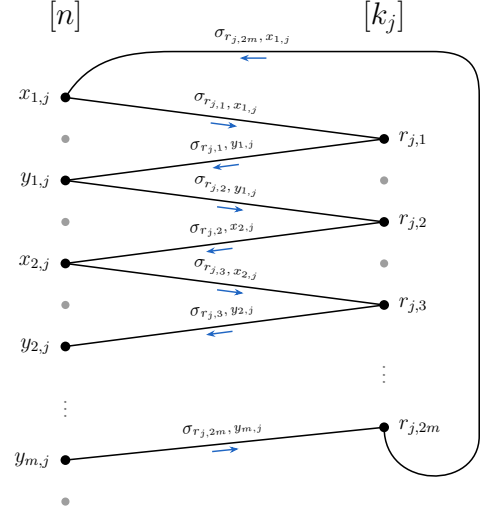


Figure 3: The closed walk on the bipartite graph for the sign monomial $\Phi_j(x, y, r_j)$. We omit the superscript (j) for legibility.

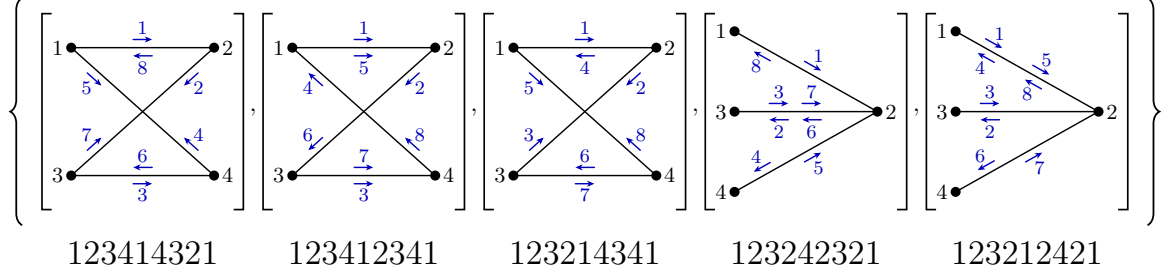


Figure 4: An example of the set of equivalence classes $\mathfrak{W}_{m=2}(\Delta = 1)$ with 8 edge traversals (i.e., word length 9) and 4 distinct vertices, giving deficiency $\Delta = 2m + 1 - \text{wt}(w) = 4 + 1 - 4 = 1$. All walks start from node 1 and the numbered blue arrows dictate the direction that the walk takes. The label beneath each walk is the string representing the walk on the vertices, formally the canonical representative of the equivalence class that we define in Definition 8.16. These are different equivalence classes because there is no bijective relabeling map from $[4] \rightarrow [4]$ that can take one of these strings to another.

All that remains now is to bound the number of ways we can label each shape using X and Y . That is, the size of an individual equivalence class $\mathcal{W} \in \mathfrak{W}_m(\Delta_j)$.

Using sparsity to bound labelings Fix deficiencies $\Delta_1, \dots, \Delta_q$ and let $\mathcal{W}_1, \dots, \mathcal{W}_q$ be equivalence classes of walks where $\mathcal{W}_j \in \mathfrak{W}_m(\Delta_j)$. We must bound how many choices of $x \in X^m$, $y \in Y^m$, and $r_j \in [k_j]^{2m}$ for $j \in [q]$ there are that yield sign monomials $\Phi_j(x, y, r_j)$ with their corresponding walks in each $\mathcal{W}_j$. See Figure 5 for explicit examples of elements in an equivalence class of $\mathfrak{W}_2(1)$.

There are no restrictions on how we can choose the right vertices $r_1, \dots, r_q$, so we take the naive bound of $\prod_{j=1}^q k_j^{\text{wt}_R(j)}$. For the left vertices in each walk, they must adhere to both the shapes we have chosen and the η -sparsity of X and Y .

To see how the shape $\mathcal{W}_j$ comes into play, consider the example where the path returns to its first vertex after four steps, as in Figure 5. Consequently, when we give a label to the first vertex, we are also labeling the fifth vertex in this walk.

For the general setting, we will consider choosing the vectors $x_1, y_1, x_2, y_2, \dots, x_m, y_m$ in order. Due to the shapes $\mathcal{W}_1, \dots, \mathcal{W}_q$ being fixed, whenever we choose a new vector, it's possible that some of its elements have already been fixed by the choices of previous vectors. Say that there are $f \geq 1$ free (unfixed) coordinates when we're choosing a vector. Since all of these vectors must lie in either of the η -sparse sets X or Y , a consequence of sparseness, shown in Fact 3.7, is that there are at most ηn^f ways to choose a valid vector.

Now, take $W_L = \sum_{j=1}^q \text{wt}_L(j)$. Each vector can have at most q unfixed coordinates, while at least W_L/q

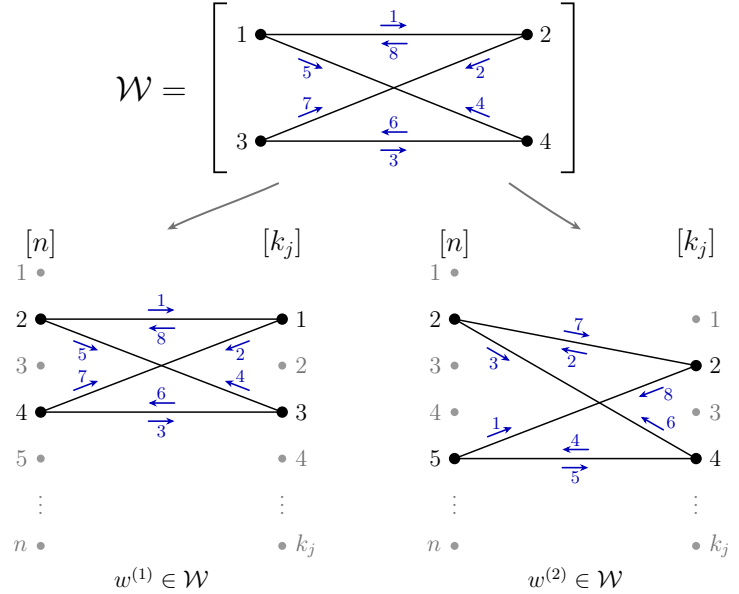


Figure 5: An example of an equivalence class $\mathcal{W}$ of a walk on a bipartite graph, which gives the shape of the walk, and elements $w^{(1)}$ and $w^{(2)}$ of the equivalence class, which are possible realizations of labelings.

of the vectors have an unfixed coordinate. Therefore, there are at most $n^{\text{wt}_L \eta^{\text{wt}_L/q}}$ possible ways to choose $x_1, y_1, x_2, y_2, \dots, x_m, y_m$ with $x_t \in X$ and $y_t \in Y$ that respect the fixed shapes $\mathcal{W}_1, \dots, \mathcal{W}_q$ of each closed walk. Including our bound on the right vertices $r_1, \dots, r_q$, we get that each choice of shapes $(\mathcal{W}_j)_{j \in [q]}$ has at most

$$n^{W_L \eta^{W_L/q}} \prod_{j=1}^q k_j^{\text{wt}_R(j)} \quad (6)$$

possible labelings. This is the only part of our bound that uses the sparsity of X and Y .

Finally, to achieve the desired bound in Equation (2), all that is left to do is to combine Equation (6) with the bound from [AZ06], sum over all possible deficiencies, and normalize for the expectation. After much simplification, this yields

$$\mathbb{E}_{\sigma}[\text{tr}((T_{\lambda} T_{\lambda}^{\top})^m)] \leq 2^q n^q \eta^{m+1} (16^q \rho^{-2q})^m,$$

as shown in Lemma 8.31. Then we can achieve Equation (2) by summing this bound over all n^{ν} choices of λ , taking the $(2m)$ -th root, and setting the power in our trace-power argument to $m = D \log n$.

3 Preliminaries

We introduce some basic definitions and facts that we will use throughout this work. In Section 3.1 we present the central objects of planes, ε -closed and ε -sparse sets, and consequences of sparseness. Section 3.2 focuses on the notion of maximal planes that, as shown in Section 3.3, are sparsely distributed if they come from an ε -closed set. We then shift to tensor codes and their related notions in Section 3.4 before finishing with some necessary facts about operator norms in Section 3.5.1.

3.1 Lines, planes, and sparseness

We begin by defining the central combinatorial objects that we will work on throughout this paper. Throughout, we take integers $D \geq 1$ and $n \geq 2$, so that the direction and anchor of a plane are uniquely determined by it as a subset of $[n]^D$.

Definition 3.1 (Axis-parallel planes). For a set of free coordinates $I \subseteq [D]$ and $a \in [n]^{I^c}$, we define the *axis-parallel plane* at the *anchor* a as

$$\mathcal{P}(I, a) = \{x \in [n]^D : \forall j \in I^c, x_j = a_j\}.$$

So points have $I = \emptyset$, lines have $|I| = 1$, and the whole cube is given when $I = [D]$.

It will be useful to focus on the specific setting of 1-dimensional planes, which we call axis-parallel lines.

Definition 3.2. For $i \in [D]$ let $I = \{i\}$ and $\mathcal{L}_i(n, D) = \{\mathcal{P}(I, a) : a \in [n]^{I^c}\}$. All axis-parallel lines are denoted as $\mathcal{L}(n, D) = \bigcup_{i \in [D]} \mathcal{L}_i(n, D)$.

Definition 3.3. Let $\mathcal{L}(M) = \{\ell \in \mathcal{L}(n, D) : \ell \subseteq M\}$ denote the set of axis-parallel lines contained in a set $M \subseteq [n]^D$.

There are two notions of sparseness that we will focus on. The first is similar to a gap property and requires that a subset of $[n]^D$ be either sparse along an axis-parallel line or completely contain it.

Definition 3.4 (ε -closed). A set $M \subseteq [n]^D$ is ε -closed if for all $\ell \in \mathcal{L}(n, D)$, we have that either $\ell \subseteq M$ or $|M \cap \ell| \leq \varepsilon n$.

Our second definition of sparseness requires that no axis-parallel line is fully contained in our set.

Definition 3.5 (ε -sparse). We say that an ε -closed set M is ε -sparse if no axis-parallel line is contained in M .

One immediate consequence of sparseness is a bound on the cardinality of M .

Lemma 3.6. For any ε -sparse set $M \subsetneq [n]^D$, we can bound its cardinality as

$$|M| \leq \varepsilon n^D.$$

Proof. Fix any direction $i \in [D]$. Then any axis-parallel line $\ell \in \mathcal{L}_i(n, D)$ has $|M \cap \ell| \leq \varepsilon n$. Since there are n^{D-1} axis-parallel lines in the i -th direction, we get

$$\begin{aligned} |M| &\leq \sum_{\ell \in \mathcal{L}_i(n, D)} |M \cap \ell| \\ &\leq n^{D-1} \cdot \varepsilon n \\ &= \varepsilon n^D, \end{aligned}$$

as claimed. $\square$

We will also make use of the following fact about sparse sets.

Fact 3.7. For a constant $\varepsilon \in (0, 1)$ and an ε -sparse set $A \subseteq [n]^D$, if we fix $I \subsetneq [D]$ and $a_I \in [n]^I$, then

$$\left| \{a_{I^c} \in [n]^{I^c} : a = (a_I, a_{I^c}) \in A\} \right| \leq \varepsilon n^{|I^c|}.$$

Proof. For $|I^c| - 1$ of the free coordinates to assign in a_{I^c} , give arbitrary assignments, of which there are $n^{|I^c|-1}$ many. The last coordinate not assigned in I^c then only has εn many options by the ε -sparseness of A . Thus, there are at most $n^{|I^c|-1} \cdot \varepsilon n = \varepsilon n^{|I^c|}$ many assignments of a_{I^c} . $\square$

3.2 Maximal planes

We first define the set of all planes contained in M .

Definition 3.8. For a set $M \subseteq [n]^D$, we define the set of all planes in M to be

$$\mathcal{P}^{\text{all}}(M) = \{\mathcal{P}(I, a) \subseteq M : I \subseteq [D], a \in [n]^{I^c}\}.$$

In particular, we do allow $I = \emptyset$.

Throughout our work, we will be particularly focused on the maximal planes of M because they give a natural way to decompose M into structured components.

Definition 3.9 (Maximal planes in M). For a set $M \subseteq [n]^D$, we define the set $\mathcal{P}(M)$ of *maximal planes* of M to contain $\mathcal{P}(I, a)$ for $I \subseteq [D]$ and $a \in [n]^{I^c}$ if

1. it is contained in M : $\mathcal{P}(I, a) \subseteq M$
2. and there is no other $\mathcal{P}(J, b)$ such that $\mathcal{P}(I, a) \subsetneq \mathcal{P}(J, b) \subseteq M$.

See Figure 6 for an example illustration.

As we will see later, these maximal planes interact at their intersections, so we will need to consider the set of planes generated as intersections of maximal planes.

Definition 3.10. For a set $M \subseteq [n]^D$, we define $\mathcal{P}^\cap(M) \subseteq \mathcal{P}^{\text{all}}(M)$ to be the closure by intersection of $\mathcal{P}(M)$, not including empty intersections. Formally,

$$\mathcal{P}^\cap(M) = \left\{ \bigcap_{P \in \mathcal{S}} P : \emptyset \neq \mathcal{S} \subseteq \mathcal{P}(M), \bigcap_{P \in \mathcal{S}} P \neq \emptyset \right\}.$$

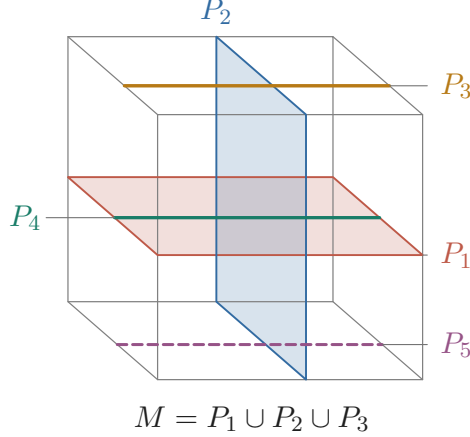


Figure 6: In this example, M is composed of three planes $M = P_1 \cup P_2 \cup P_3$, each of which is a maximal plane. On the other hand, P_4 is a plane of M but not maximal because it's contained in P_1 , and P_5 isn't a plane of M because it's not fully contained in M .

Using these definitions, we can show the simple fact that any $M \subseteq [n]^D$ can be covered by maximal planes.

Lemma 3.11. For any $M \subseteq [n]^D$, we have

$$M = \bigcup_{I \subseteq [D]} \bigcup_{a \in \Omega_I(M)} \mathcal{P}(I, a).$$

Proof. Every maximal plane is contained in M , which gives the inclusion $\supseteq$. Conversely, any $x \in M$ is itself a zero-dimensional plane $\{x\} = \mathcal{P}(\emptyset, x) \subseteq M$, so it belongs to some element of $\mathcal{P}^{\text{all}}(M)$. Since $\mathcal{P}^{\text{all}}(M)$ is a finite poset under inclusion, $\{x\}$ is contained in a maximal one, which is by definition an element of $\mathcal{P}(M)$ containing x . $\square$

3.3 Anchor sets

Next, given a set of free coordinates $I \subseteq [D]$, we can define the set of fixed coordinates that index all maximal planes.

Definition 3.12 (Indexing set of maximal planes in direction I). For a set $M \subseteq [n]^D$ and free coordinates $I \subseteq [D]$, we define the set of *anchors* indexing the maximal planes as

$$\Omega_I(M) = \{a \in [n]^{I^c} : \mathcal{P}(I, a) \in \mathcal{P}(M)\}.$$

Now, we will show that M being ε -closed makes $\Omega_I(M)$ ε -sparse.

Lemma 3.13. Let $0 < \varepsilon < 1$ and let $M \subseteq [n]^D$ be ε -closed. Then for any $I \subseteq [D]$, the set $\Omega_I(M) \subseteq [n]^{I^c}$ is ε -sparse. Consequently,

$$|\Omega_I(M)| \leq \varepsilon n^{D-|I|}.$$

Proof. This proof is visually simple, and we recommend that the reader follow along in Figure 7. Suppose, for the sake of contradiction, that there exist $j \in I^c$ and $b \in [n]^{I^c \setminus \{j\}}$ such that the line

$$\ell_b = \{s \in [n]^{I^c} : s_j \in [n], s_{I^c \setminus \{j\}} = b\}$$

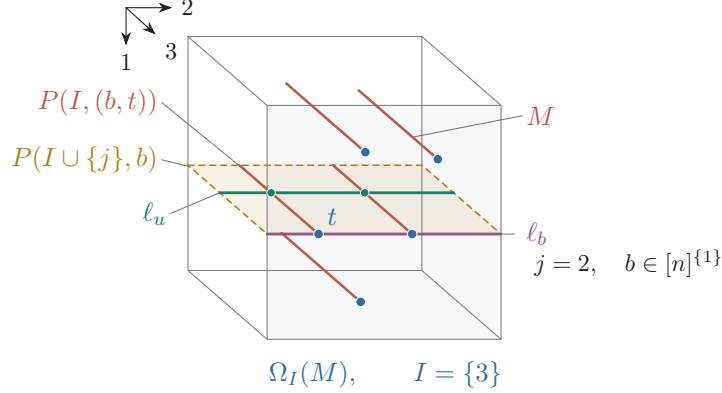


Figure 7: Visualization of the proof of Lemma 3.13 for $D = 3$ with $I = \{3\}$, $j = 2$, and an M with lines parallel to the third axis only.

in the j -th direction has $|\Omega_I(M) \cap \ell_b| > \varepsilon n$. For each $t \in [n]$ with $(b, t) \in \Omega_I(M) \cap \ell_b$, it must be that $P(I, (b, t)) \in \mathcal{P}(M)$ where $(b, t) \in [n]^{I^c}$ is the element with the j -th coordinate as t and the other coordinates set as b .

Now, consider any $u \in [n]^I$ (the empty tuple when $I = \emptyset$) and the j -parallel line

$$\ell_u = \{x \in [n]^D : x_I = u, x_{I^c \setminus \{j\}} = b, x_j \in [n]\}.$$

By construction, for any $t \in [n]$ with $(b, t) \in \Omega_I(M) \cap \ell_b$, we have that the point $y = (y_I = u, y_{I^c \setminus \{j\}} = b, y_j = t) \in \ell_u$ is contained in the maximal plane $P(I, (b, t))$. Because $y \in \ell_u \cap P(I, (b, t))$ and $P(I, (b, t)) \subseteq M$ and $|\Omega_I(M) \cap \ell_b| > \varepsilon n$, we have that $|\ell_u \cap M| > \varepsilon n$. Using the fact that M is ε -closed, we get that $\ell_u \subseteq M$.

Our choice of $u \in [n]^I$ was arbitrary, so we can take their union to conclude that the plane

$$P(I \cup \{j\}, b) = \bigcup_{u \in [n]^I} \ell_u$$

is fully contained in M . However, we then get that $P(I \cup \{j\}, b) \supsetneq P(I, (b, t))$ for any $t \in [n]$ with $(b, t) \in \Omega_I(M) \cap \ell_b$, contradicting the fact that $P(I, (b, t))$ is maximal. Therefore, for all $j \in I^c$ and $b \in [n]^{I^c \setminus \{j\}}$, the line ℓ_b has $|\Omega_I(M) \cap \ell_b| \leq \varepsilon n < n$, meaning that $\Omega_I(M)$ is ε -sparse.

Finally, now that we have that $\Omega_I(M)$ is ε -sparse in $[n]^{I^c}$, by Lemma 3.6 we can bound its cardinality as $|\Omega_I(M)| \leq \varepsilon n^{D-|I|}$, as claimed. $\square$

We can get a similar result for the indexing set $\Omega_I^\cap(M) = \{a \in [n]^{I^c} : \mathcal{P}(I, a) \in \mathcal{P}^\cap(M)\}$ of $\mathcal{P}^\cap(M)$. This is a loose bound and introduces a factor of 2^D , but since D is a constant, this is not an issue.

Corollary 3.14. Let $M \subseteq [n]^D$ be ε -closed with $0 < 2^D \varepsilon < 1$. Then for every $I \subseteq [D]$, the set $\Omega_I^\cap(M) \subseteq [n]^{I^c}$ is $(2^D \varepsilon)$ -sparse.

Proof. Consider a direction $j \in I^c$ and an axis-parallel line ℓ_j in $[n]^{I^c}$ in the j -th direction. For every anchor $a \in \ell_j \cap \Omega_I^\cap(M)$, write $\mathcal{P}(I, a)$ as an intersection of maximal planes. At least one of these planes, say $\mathcal{P}(J, b)$, must fix coordinate j , so $I \subseteq J$, $j \notin J$, and $b = a_{J^c}$. For a fixed such J , all coordinates of b except j are fixed by ℓ_j . By Lemma 3.13, $\Omega_J(M)$ is ε -sparse, meaning that there are at most εn possible values of $b_j = a_j$. There are at most 2^D choices of J , so a union bound gives at most $2^D \varepsilon n$ anchors in $\ell_j \cap \Omega_I^\cap(M)$. $\square$

3.4 Tensor codes and related notions

We now introduce the notions of inner-generated and extendable sets. In the next section, we will go over some of these definitions again and see how they are naturally derived in a slightly different context, but we introduce their standard definitions here as they show up in the literature. Throughout this section, we will let $\mathbb{K}$ represent a generic field, possibly finite or infinite. Throughout, we identify a line $\ell = \mathcal{P}(\{i\}, a) \in \mathcal{L}_i(n, D)$ with $[n]$ via its i -th coordinate, so that the restriction $x|_\ell$ of a word is an element of $\mathbb{K}^n$ and can be tested for membership in the component code $\mathcal{C}_i$.

Definition 3.15 (Codes along a line ℓ). For $\ell \in \mathcal{L}_i(n, D)$ and a code $\mathcal{C}_i \subseteq \mathbb{K}^n$, define

$$\mathcal{C}_\ell = \{c \in \mathbb{K}^{[n]^D} : \text{supp}(c) \subseteq \ell, c|_\ell \in \mathcal{C}_i\}.$$

Definition 3.16 (Check-product code). Given a collection of codes $\mathcal{C}_1, \dots, \mathcal{C}_D$, the dual of their tensor code is the check-product code $\mathcal{C}_1^\perp \boxplus \dots \boxplus \mathcal{C}_D^\perp = (\mathcal{C}_1 \otimes \dots \otimes \mathcal{C}_D)^\perp$.

An easy way of seeing the relation between the check-product code and the tensor product code is to define $\mathcal{C}^{(i)} = \mathbb{K}^n \otimes \dots \otimes \mathcal{C}_i \otimes \dots \otimes \mathbb{K}^n \subseteq \mathbb{K}^{[n]^D}$ which allows us to write

$$\begin{aligned} \bigotimes_{i=1}^D \mathcal{C}_i &= \bigcap_{i=1}^D \mathcal{C}^{(i)} \\ \bigboxplus_{i=1}^D \mathcal{C}_i &= \sum_{i=1}^D \mathcal{C}^{(i)}. \end{aligned}$$

Instead of considering the entire tensor code at once, we can also consider words that satisfy local checks contained in a set of coordinates.

Definition 3.17 (Words accepted by checks in M). For a subset of coordinates $M \subseteq [n]^D$ of a tensor code $\mathcal{C} = \mathcal{C}_1 \otimes \dots \otimes \mathcal{C}_D$, let

$$\mathcal{F}_\mathcal{C}(M) = \{x \in \mathbb{K}^M : \forall i \in [D], \forall \ell \in \mathcal{L}_i(n, D), \ell \subseteq M \Rightarrow x|_\ell \in \mathcal{C}_i\}$$

be the set of words that satisfy the local checks of the corresponding component code for each axis-parallel line contained in M .

Definition 3.18 (Extendability). We say that M is *extendable* in $\mathcal{C}$ if the restriction map

$$\pi_M : \mathcal{C} \rightarrow \mathcal{F}(M)$$

is surjective. That is, each local codeword in $\mathcal{F}(M)$ can be extended to a global codeword of $\mathcal{C}$.

Definition 3.19 (Inner-generated sets for the check-product code). A set $M \subseteq [n]^D$ is *inner-generated* for a code $\mathcal{C}_1 \boxplus \dots \boxplus \mathcal{C}_D$ if for each codeword $c \in \mathcal{C}_1 \boxplus \dots \boxplus \mathcal{C}_D$ with $\text{supp}(c) \subseteq M$ we have $c \in \sum_{\ell \in \mathcal{L}(M)} \mathcal{C}_\ell$, where $\mathcal{L}(M)$ is the set of axis-parallel lines contained in M .

We can now define the central notion of expansion used in this paper. For a word $a \in \mathbb{K}^{[n]^D}$ we write $|a|$ for its Hamming weight, and for a direction $i \in [D]$ we write

$$|a|_i = |\{\ell \in \mathcal{L}_i(n, D) : a|_\ell \neq 0\}|$$

for the number of i -parallel lines on which a does not vanish.

Definition 3.20 (Product expansion). A tuple of codes $(\mathcal{C}_i)_{i \in [D]}$ is *ν -product expanding* if every $x \in \bigboxplus_{i=1}^D \mathcal{C}_i$ admits a decomposition $x = \sum_{i=1}^D a_i$ with $a_i \in \mathcal{C}^{(i)}$ such that

$$\nu n \sum_{i=1}^D |a_i|_i \leq |x|.$$

It is *two-sided ν -product expanding* if both $(C_i)_{i \in [D]}$ and $(C_i^\perp)_{i \in [D]}$ are ν -product expanding, where duals are taken with respect to the standard coordinate pairing on $\mathbb{K}^n$.

Finally, for a prime p we write $\text{Gr}_{p^t}(n, k)$ for the Grassmannian of k -dimensional subspaces of $\mathbb{F}_{p^t}^n$. Sampling a random code of dimension k always means sampling uniformly from $\text{Gr}_p(n, k)$.

3.5 Miscellanea

3.5.1 Linear algebra facts

We record some useful linear algebra facts that we will use at various points later on in this work. Here all spaces are finite-dimensional real or complex Hilbert spaces, and we use the Euclidean operator norm.

Fact 3.21 (Equation (2.79) and Exercise 2.3.16 of [Tao12]). For a linear map T and integer $p \geq 1$,

$$\|T\|_{\text{op}}^{2p} \leq \text{tr}((TT^\dagger)^p).$$

Fact 3.22 (Equation (1.175) of [Wat18]). For two compatible linear maps A, B , we have that the operator norm is submultiplicative

$$\|AB\|_{\text{op}} \leq \|A\|_{\text{op}} \|B\|_{\text{op}}. \quad (7)$$

Fact 3.23 (Equation (2.2.10) of [Tro15] or Problem 6(b) of [JH91b]). For a family of linear maps $(A_s)_{s \in S}$,

$$\left\| \bigoplus_{s \in S} A_s \right\|_{\text{op}} = \max_{s \in S} \|A_s\|_{\text{op}}.$$

Fact 3.24 (Theorem 4.2.15 and Problem 27 of [JH91a]). For a family of linear maps $(A_s)_{s \in S}$,

$$\left\| \bigotimes_{s \in S} A_s \right\|_{\text{op}} = \prod_{s \in S} \|A_s\|_{\text{op}}.$$

3.5.2 A useful estimate on the distribution of primes

Here we record a consequence of Chebyshev's estimates that bounds the value of a prime not dividing a given integer. We begin by recalling Chebyshev's estimate.

Theorem 3.25 (Corollary 2.6 of [MV06]). For any $x \geq 2$, the first Chebyshev function $\vartheta(x) := \sum_{\substack{p \leq x \\ p \text{ prime}}} \log p$ is bounded as $\vartheta(x) = \Theta(x)$.

We now derive our required corollary.

Corollary 3.26. There exists a universal constant C such that for all $q \geq 2$ and every $m \in \mathbb{N}$, there exists a prime p *not* dividing m such that

$$q < p \leq C(q + \log(2m)).$$

Proof. By Theorem 3.25, there exist universal constants $0 < A \leq B$ such that for all $x \geq 2$ we have

$$Ax \leq \vartheta(x) \leq Bx. \quad (8)$$

Let $\ell = \log(2m)$ and $s = \frac{B+1}{A}(q + \ell)$. Now, for the sake of contradiction, assume that all primes $p \in (q, s]$ divide m . These primes are of course distinct, so $\prod_{\substack{p \in (q, s] \\ p \text{ prime}}} p$ must divide m . Consequently,

$$\begin{aligned} \vartheta(s) - \vartheta(q) &= \sum_{\substack{p \leq s \\ p \text{ prime}}} \log p - \sum_{\substack{p \leq q \\ p \text{ prime}}} \log p \\ &= \log \prod_{\substack{q < p \leq s \\ p \text{ prime}}} p \\ &\leq \log m. \end{aligned} \tag{9}$$

However, from Equation (8) we obtain

$$\begin{aligned} \vartheta(s) - \vartheta(q) &\geq As - Bq \\ &= (B+1)(q + \ell) - Bq \\ &= q + (B+1)\ell \\ &> \log m, \end{aligned}$$

which contradicts Equation (9). Therefore, there must exist some prime $p \in (q, s]$ not dividing m such that

$$q < p \leq s = \frac{B+1}{A}(q + \log(2m)),$$

as claimed. $\square$

4 (Co)Restriction Maps and Atoms

In this section, we will give intuition for the duality between the properties of extendability and inner-generatedness. We will detail how exactly the tensor product code and check-product codes are dual to each other along with how surjectivity of the restriction map of the tensor code, which is the condition for extendability, is equivalent to injectivity of the corestriction map for the check-product code, which is the condition for inner-generatedness. Since many of these definitions can be defined over any arbitrary field, and we will use them over both finite fields and the reals, we will use the generic field $\mathbb{K}$, which may be finite or infinite, unless we specify otherwise.

In Section 4.1 we introduce the restriction map on tensor codes to subsets of coordinates and its relation to extendability, and in Section 4.2 we define the dual notion of the corestriction map on the check-product code and its relation to inner-generatedness.⁷ We combine these notions in Section 4.3 to show precisely how they are related and give an explicit formula for the corestriction maps. Finally, in Section 4.4 we present the notion of the atom decomposition of the check-product code.

4.1 The tensor code, restrictions, and extendability

We begin by choosing generator matrices

$$G_i = \begin{bmatrix} | & | & \cdots & | \\ g_i(1) & g_i(2) & \cdots & g_i(n) \\ | & | & & | \end{bmatrix} \in \mathbb{K}^{k_i \times n}.$$

for codes

$$\mathcal{C}_1, \dots, \mathcal{C}_D \subseteq \mathbb{K}^n$$

⁷The reader familiar with tensor codes may skip these two sections.

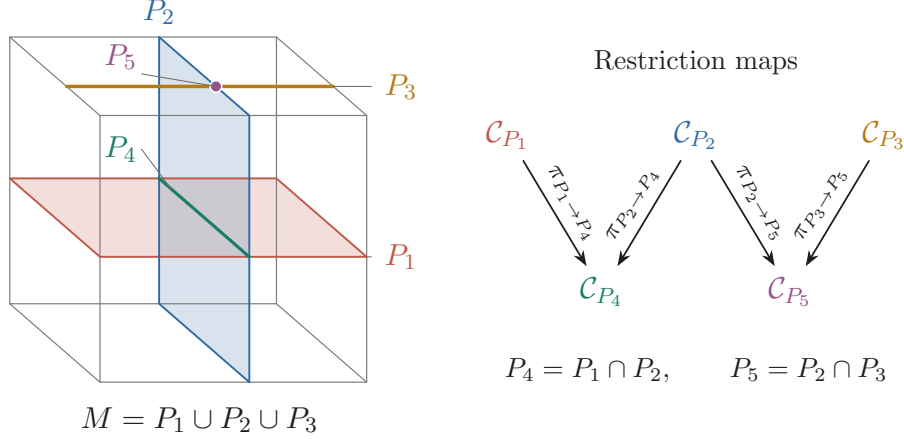


Figure 8: Example of planes in $M = P_1 \cup P_2 \cup P_3$ and part of the diagram that is ultimately being used for the limit over the $\mathcal{C}_P$'s.

such that $\mathcal{C}_i = \text{Im}(G_i^\top)$ and $\dim \mathcal{C}_i = k_i$. Their tensor code is denoted as $\mathcal{C} = \mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_D \subseteq \mathbb{K}[n]^D$.

For $P = \mathcal{P}(I, a) \in \mathcal{P}^{\text{all}}(M)$, define

$$\mathcal{C}_P = \mathcal{C}_I = \bigotimes_{i \in I} \mathcal{C}_i \subseteq \mathbb{K}[n]^I.$$

If $I = \emptyset$, then we define $\mathcal{C}_I = \mathbb{K}$. Between these vector spaces, we can define restriction maps.

Definition 4.1 (Restriction maps). For any $P, Q \in \mathcal{P}^{\text{all}}(M)$ such that $P = \mathcal{P}(I, a) \subseteq Q = \mathcal{P}(J, b)$, so $I \subseteq J \subseteq [D]$, define the restriction map

$$\pi_{Q \rightarrow P} = \left(\bigotimes_{j \in J \setminus I} \text{ev}_{a_j} \right) \otimes \left(\bigotimes_{i \in I} \text{id}_{\mathcal{C}_i} \right),$$

where tensor factors are placed in increasing coordinate order and ev_{a_j} is the evaluation map at coordinate a_j . Equivalently, we can define the restriction map on pure tensors $c = \bigotimes_{j \in J} c_j$ as

$$\pi_{Q \rightarrow P} : \mathcal{F}(Q) \rightarrow \mathcal{F}(P)$$

$$c \mapsto \left(\prod_{j \in J \setminus I} c_j(a_j) \right) \left(\bigotimes_{i \in I} c_i \right),$$

and extend linearly to the whole space.

In particular, one can check that for planes $P \subseteq Q \subseteq R$, we have

$$\pi_{R \rightarrow P} = \pi_{Q \rightarrow P} \circ \pi_{R \rightarrow Q}.$$

See Figure 8 for an example. Therefore, because Lemma 3.11 shows that M is covered by its maximal planes, we have that for any $M \subseteq [n]^D$ we can order $\mathcal{P}^{\text{all}}(M)$ by inclusion and define the local sections on

M as limits:⁸

$$\begin{aligned}\mathcal{F}(M) &= \lim_{P \in \mathcal{P}^\cap(M)} \mathcal{C}_P \\ &= \{(c_P)_{P \in \mathcal{P}^\cap(M)} : c_P \in \mathcal{C}_P, \text{ s.t. for all } P \subseteq Q, \pi_{Q \rightarrow P}(c_Q) = c_P\} \\ &\cong \{x \in \mathbb{K}^M : \forall i \in [D], \forall \ell \in \mathcal{L}_i(n, D), \ell \subseteq M \Rightarrow x|_\ell \in \mathcal{C}_i\}.\end{aligned}$$

The last line is an isomorphism rather than an equality because a compatible family is indexed by planes and not by the points of M . It is a bijection because M is covered by its maximal planes by Lemma 3.11 (some of which may be zero-dimensional), and any two planes of $\mathcal{P}^\cap(M)$ agree on their intersection whenever it is nonempty, since that intersection again lies in $\mathcal{P}^\cap(M)$. So a compatible family glues to a single word $x \in \mathbb{K}^M$, and conversely every such x restricts to a compatible family. Since every line contained in M lies in a maximal plane, the constraints on x are exactly the local checks of Definition 3.17, and taking $M = [n]^D$ gives $\mathcal{C} = \mathcal{F}([n]^D)$. Finally, to interpret extendability in this context, we consider the global restriction map

$$\pi_M : \mathcal{C} \rightarrow \mathcal{F}(M).$$

Fact 4.2. A set $M \subseteq [n]^D$ is extendable if and only if π_M is surjective.

Proof. The condition that a codeword $c_M \in \mathcal{F}(M)$ can extend to a global word $c \in \mathcal{C}$ such that $c|_M = c_M$ is exactly the condition that π_M is surjective, which allows us to “lift” c_M to c (possibly non-uniquely). $\square$

4.2 The check-product code, corestrictions, and inner-generatedness

In this subsection, we will essentially take the dual of the previous subsection and derive the corresponding notions. Let $\mathcal{B}_i = \mathcal{C}_i^\perp$ for $i \in [D]$, and define

$$\mathcal{B} = \mathcal{B}_1 \boxplus \dots \boxplus \mathcal{B}_D \subseteq \mathbb{K}^{[n]^D}.$$

Similarly, for $P = \mathcal{P}(I, a) \in \mathcal{P}^{\text{all}}(M)$, let

$$\mathcal{B}_P = \mathcal{B}_I = \bigsqcup_{i \in I} \mathcal{B}_i \subseteq \mathbb{K}^{[n]^I}.$$

For a point P we have $I = \emptyset$, and we take the empty check-product to be $\mathcal{B}_\emptyset = 0$, matching the empty tensor product $\mathcal{C}_\emptyset = \mathbb{K}$. As we saw in Definition 3.16, $\mathcal{C}^\perp = \mathcal{B}$, and similarly for any plane $P = \mathcal{P}(I, a) \in \mathcal{P}^{\text{all}}(M)$, it is easy to verify that

$$\begin{aligned}\mathcal{C}_P^\perp &= \left(\bigotimes_{i \in I} \mathcal{C}_i \right)^\perp \\ &= \bigsqcup_{i \in I} \mathcal{C}_i^\perp \\ &= \bigsqcup_{i \in I} \mathcal{B}_i \\ &= \mathcal{B}_P.\end{aligned}$$

⁸We will not use (co)limits outside of this section and they do not play a central role in the rest of this paper, so we do not define them formally and instead simply expand out their definitions in this context.

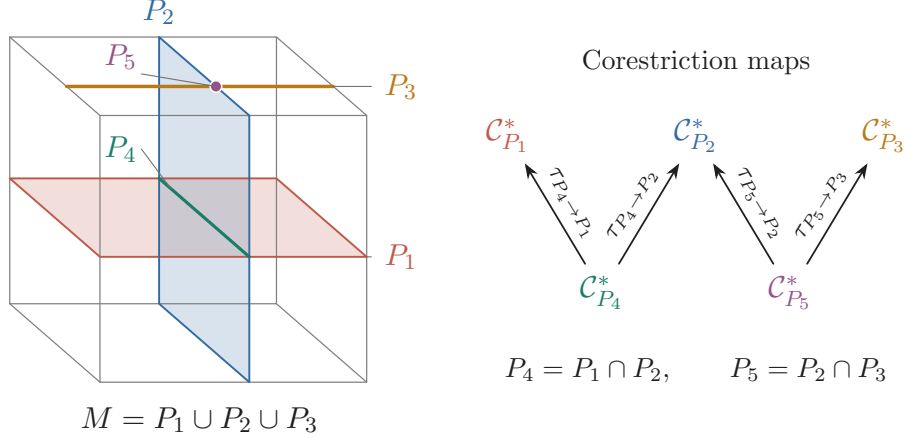


Figure 9: Example of planes in $M = P_1 \cup P_2 \cup P_3$ and part of the diagram that is ultimately being used for the colimit over the $\mathcal{C}_P^*$'s.

Therefore, to dualize the construction of tensor codes above, we assign to each $P \in \mathcal{P}^{\text{all}}(M)$ the space

$$\begin{aligned}
 \mathcal{G}(P) &= \mathcal{F}(P)^* \\
 &= \mathcal{C}_P^* \\
 &= \text{Hom}_{\mathbb{K}}(\mathcal{C}_P, \mathbb{K}) \\
 &\cong \mathbb{K}^P / \mathcal{C}_P^\perp \\
 &= \mathbb{K}^P / \mathcal{B}_P.
 \end{aligned}$$

Between these local spaces, we can define the dual of the restriction map, the corestriction map.

Definition 4.3 (Corestriction map). For $P, Q \in \mathcal{P}^{\text{all}}(M)$ with $P \subseteq Q$, restriction goes from Q to P , so its dual goes from P to Q . We define the corestriction map

$$\begin{aligned}
 \tau_{P \rightarrow Q} &= \pi_{Q \rightarrow P}^* : \mathcal{C}_P^* \rightarrow \mathcal{C}_Q^* \\
 \varphi &\mapsto \varphi \circ \pi_{Q \rightarrow P}.
 \end{aligned}$$

Now, it is clear that for planes $P \subseteq Q \subseteq R$, their corestriction maps are compatible via

$$\tau_{P \rightarrow R} = \tau_{Q \rightarrow R} \circ \tau_{P \rightarrow Q}$$

exactly because $\pi_{R \rightarrow P} = \pi_{Q \rightarrow P} \circ \pi_{R \rightarrow Q}$. See Figure 9 for an example. Therefore, for any $M \subseteq [n]^D$ we can define

$$\begin{aligned}
 \mathcal{G}(M) &= \mathcal{F}(M)^* \\
 &= \left(\lim_{P \in \mathcal{P}^{\cap}(M)} \mathcal{C}_P \right)^* \\
 &\cong \text{colim}_{P \in \mathcal{P}^{\cap}(M)} \mathcal{C}_P^* \\
 &\cong \text{colim}_{P \in \mathcal{P}^{\cap}(M)} \mathbb{K}^P / \mathcal{B}_P \\
 &\cong \frac{\mathbb{K}^M}{\sum_{\ell \in \mathcal{L}(M)} \mathcal{B}_\ell},
 \end{aligned} \tag{10}$$

where we are including each $\mathcal{B}_\ell$ into $\mathbb{K}^M$ by extending by 0's outside of the line ℓ . For the whole space, this gives $\mathcal{G}([n]^D) = \frac{\mathbb{K}^{[n]^D}}{\mathcal{B}}$. We also get the global corestriction map

$$\tau_M : \mathcal{G}(M) \cong \frac{\mathbb{K}^M}{\sum_{\ell \in \mathcal{L}(M)} \mathcal{B}_\ell} \rightarrow \frac{\mathbb{K}^{[n]^D}}{\mathcal{B}} \cong \mathcal{G}([n]^D).$$

We now show that the kernel of τ_M is exactly the obstruction to M being inner-generated.

Fact 4.4. A set $M \subseteq [n]^D$ is inner-generated in $\mathcal{B}$ iff τ_M is injective iff $\ker(\tau_M) = 0$.

Proof. To be explicit, we define the inclusion $\iota_M : \mathbb{K}^M \rightarrow \mathbb{K}^{[n]^D}$ that extends an element of $\mathbb{K}^M$ by 0's to a full element of $\mathbb{K}^{[n]^D}$. Thus, an element $[z] \in \frac{\mathbb{K}^M}{\sum_{\ell \in \mathcal{L}(M)} \mathcal{B}_\ell}$ is in $\ker(\tau_M)$ iff $[\iota_M(z)] = 0 \in \frac{\mathbb{K}^{[n]^D}}{\mathcal{B}}$, meaning that $\iota_M(z) \in \mathcal{B}$. Equivalently, this means that we need $z \in \mathbb{K}^M \cap \mathcal{B}$, so

$$\ker(\tau_M) \cong \frac{\mathbb{K}^M \cap \mathcal{B}}{\sum_{\ell \in \mathcal{L}(M)} \mathcal{B}_\ell},$$

where we are implicitly applying ι_M to $\mathbb{K}^M$. The definition of M being inner-generated from Definition 3.19 is exactly when $\mathbb{K}^M \cap \mathcal{B} = \sum_{\ell \in \mathcal{L}(M)} \mathcal{B}_\ell$, which is exactly $\ker(\tau_M) = 0$, as claimed. $\square$

4.3 Putting it all together

Using the fact that a linear map is injective if and only if its dual is surjective, we get the following lemma.

Lemma 4.5. For any set $M \subseteq [n]^D$, the following are equivalent:

1. M is extendable in $\mathcal{C} = \mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_D$,
2. M is inner-generated in $\mathcal{B} = \mathcal{B}_1 \boxplus \cdots \boxplus \mathcal{B}_D$,
3. π_M is surjective, and
4. τ_M is injective.

Proof. We just use Facts 4.2 and 4.4 together with the fact that a linear map is injective if and only if its dual is surjective. $\square$

In the rest of this paper, our goal will be to show that τ_M is injective with high probability for a selection of random linear codes. In this vein, we show how the injectivity of τ_M depends on the generator matrices $(G_i)_{i \in [D]}$.

Definition 4.6 (Dual message space). For $i \in [D]$, let U_i be a k_i -dimensional vector space over $\mathbb{K}$. This is the *dual message space* of $\mathcal{C}_i$ since we can consider $\mathcal{C}_i \cong U_i^*$ by taking the columns $g_i(1), \dots, g_i(n)$ of the generator matrix G_i and writing down a message $\mu_i \in U_i^*$ as the codeword

$$c_{\mu_i} = (\mu_i(g_i(1)), \mu_i(g_i(2)), \dots, \mu_i(g_i(n))) \in \mathcal{C}_i \subseteq \mathbb{K}^n.$$

For any plane $P = \mathcal{P}(I, a)$, define

$$U_P = U_I = \bigotimes_{i \in I} U_i.$$

We take $U_\emptyset = \mathbb{K}$ for a point.

By definition of the U_i 's, we have that for a plane $P = \mathcal{P}(I, a)$,

$$\begin{aligned}\mathcal{C}_P^* &= \left(\bigotimes_{i \in I} \mathcal{C}_i \right)^* \\ &\cong \bigotimes_{i \in I} \mathcal{C}_i^* \\ &\cong \bigotimes_{i \in I} U_i \\ &= U_I = U_P.\end{aligned}$$

Thus, we can make the computation of $\tau_{P \rightarrow Q}$ explicit.

Fact 4.7. For planes $P = \mathcal{P}(I, a) \subseteq Q = \mathcal{P}(J, b)$ with $I \subseteq J$

$$\begin{aligned}\tau_{P \rightarrow Q} &= \pi_{Q \rightarrow P}^* : U_I \rightarrow U_J \\ u &\mapsto u \otimes \left(\bigotimes_{j \in J \setminus I} g_j(a_j) \right),\end{aligned}\tag{11}$$

where we order the tensor products by the ordering of $[D]$.

From this, we can derive a nice structural result on $\mathcal{G}(M)$.

Lemma 4.8. We have that $\mathcal{G}(M) \cong U_M / R_M$ where $U_M = \bigoplus_{P \in \mathcal{P}^\cap(M)} U_P$ and

$$R_M = \text{span}\{e_P(u) - e_Q(\tau_{P \rightarrow Q}(u)) : P, Q \in \mathcal{P}^\cap(M), P \subseteq Q, u \in U_P\}$$

where $e_P : U_P \rightarrow U_M$ is the natural inclusion into the direct sum of vector spaces and e_Q is analogous.

Proof. We start from Equation (10) and use the definition of a colimit to compute

$$\begin{aligned}\mathcal{G}(M) &\cong \text{colim}_{P \in \mathcal{P}^\cap(M)} \mathcal{C}_P^* \\ &= \text{colim}_{P \in \mathcal{P}^\cap(M)} U_P \\ &\cong U_M / R_M,\end{aligned}$$

where the last step is just the definition of the colimit on vector spaces. $\square$

Later on we will need to treat the entries of the generator matrices as free parameters and allow arbitrary matrices G_i , including ones of deficient row rank that do not generate a k_i -dimensional code. Since Equation (11) only inserts columns of the G_i , it continues to make sense in this generality.

Definition 4.9 (Polynomially parametrized insertion maps). Let $\mathbf{X} = (X_{r,j}^{(i)})_{i \in [D], r \in [k_i], j \in [n]}$ be independent indeterminates, let X_i be the $k_i \times n$ matrix over $\mathbb{Z}[\mathbf{X}]$ with entries $X_{r,j}^{(i)}$, and write $x_i(j)$ for its j -th column. We take $U_i = \mathbb{Z}[\mathbf{X}]^{k_i}$ and $U_P = U_I = \bigotimes_{i \in I} U_i$ for a plane $P = \mathcal{P}(I, a)$, with $U_\emptyset = \mathbb{Z}[\mathbf{X}]$ for a point. For planes $P = \mathcal{P}(I, a) \subseteq Q = \mathcal{P}(J, b)$, the corestriction map $\tau_{P \rightarrow Q}^{\mathbf{X}}$ is a $\mathbb{Z}[\mathbf{X}]$ -linear map $U_P \rightarrow U_Q$ defined by the same formula as in Equation (11) but using X_i in place of G_i .

A *specialization* of these maps substitutes matrices $G_i \in \mathbb{K}^{k_i \times n}$ over a field $\mathbb{K}$ for the X_i , and recovers the corestriction maps of Equation (11) for those G_i .

No rank requirement is involved here, and compatibility is immediate from the formula: for $P \subseteq Q \subseteq R$, both $\tau_{Q \rightarrow R}^{\mathbf{X}} \circ \tau_{P \rightarrow Q}^{\mathbf{X}}$ and $\tau_{P \rightarrow R}^{\mathbf{X}}$ insert the columns $x_m(a_m)$ for the coordinates m freed between P and R ,

since the anchors of $P \subseteq Q$ agree on the coordinates that Q fixes. This is a polynomial identity, so it survives every specialization.

Taking these maps with the indeterminates $\mathbf{X}$ over all planes of M gives the diagram we will actually compute with later on.

Definition 4.10. For $M \subseteq [n]^D$, keep $U_M = \bigoplus_{P \in \mathcal{P}^\cap(M)} U_P$ from Lemma 4.8 and set

$$\mathcal{E}_M = \bigoplus_{\substack{P, Q \in \mathcal{P}^\cap(M) \\ P \subsetneq Q}} U_P,$$

and let $e_{P,Q} : U_P \rightarrow \mathcal{E}_M$ and $e_P : U_P \rightarrow U_M$ denote the inclusions of the corresponding summands. The maps

$$\begin{aligned} \alpha_M^{\mathbf{X}} : \mathcal{E}_M &\rightarrow U_M & \beta_M^{\mathbf{X}} : U_M &\rightarrow U_{[D]} \\ e_{P,Q}(u) &\mapsto e_P(u) - e_Q(\tau_{P \rightarrow Q}^{\mathbf{X}}(u)) & (u_P)_P &\mapsto \sum_{P \in \mathcal{P}^\cap(M)} \tau_{P \rightarrow [n]^D}^{\mathbf{X}}(u_P), \end{aligned} \quad (12)$$

are defined on generators and extend linearly. Compatibility gives $\beta_M^{\mathbf{X}} \circ \alpha_M^{\mathbf{X}} = 0$, so $\beta_M^{\mathbf{X}}$ descends to the quotient map

$$\bar{\beta}_M^{\mathbf{X}} : \text{Coker } \alpha_M^{\mathbf{X}} \rightarrow U_{[D]}.$$

The $P = Q$ terms in the spanning set of R_M contribute $e_P(u) - e_P(u) = 0$, so $\text{Im } \alpha_M = R_M$. Hence whenever every G_i has full row rank, Lemma 4.8 gives $\text{Coker } \alpha_M \cong \mathcal{G}(M)$ and $\bar{\beta}_M = \tau_M$. For arbitrary G_i there is no such interpretation in terms of codes, but the diagram and its cokernel are still well defined.

The specializations of this diagram that we will need are the ones witnessing injectivity over $\mathbb{R}$.

Definition 4.11 (Real sign witness). Fix $M \subseteq [n]^D$ and dimensions $(k_i)_{i \in [D]}$. A *real sign witness* is a tuple of sign matrices $\sigma = (\sigma^{(i)})_{i \in [D]}$ with $\sigma^{(i)} \in \{-1, 1\}^{k_i \times n}$ at whose specialization over $\mathbb{K} = \mathbb{R}$ the map $\bar{\beta}_M^{\mathbf{X}}$ of Definition 4.10 is injective. We emphasize that the $\sigma^{(i)}$ are *not* required to have full row rank.

4.4 Atom decomposition

Using these definitions so far, we give an even more precise decomposition of $\mathcal{G}(M)$ into what we call “atoms”. For the remainder of this section, we will take our base field $\mathbb{K} = \mathbb{R}$ and assign the standard inner product to all U_i .

Definition 4.12. For $P \in \mathcal{P}^\cap(M)$, define the vector space

$$U_{\subsetneq P} = \sum_{\substack{P' \in \mathcal{P}^\cap(M) \\ P' \subsetneq P}} \tau_{P' \rightarrow P}(U_{P'}) \subseteq U_P.$$

The *atom* A_P of P is then defined as the orthogonal complement

$$A_P = U_{\subsetneq P}^\perp \subseteq U_P,$$

which is naturally isomorphic to the quotient

$$A_P \cong U_P / U_{\subsetneq P}.$$

Unlike in Lemma 4.8 where we describe $\mathcal{G}(M)$ exactly as the quotient of a direct sum, the direct sum of these atoms covers $\mathcal{G}(M)$; that is, there is a natural surjection from this direct sum onto $\mathcal{G}(M)$.

Definition 4.13. For a plane $P \in \mathcal{P}^\cap(M)$, let $\bar{e}_P : U_P \rightarrow \left(\bigoplus_{P' \in \mathcal{P}^\cap(M)} U_{P'} \right) / R_M$ be the natural inclusion e_P composed with quotienting by R_M .

Lemma 4.14. Let $M \subseteq [n]^D$. Then the natural map

$$\begin{aligned} \Psi : \bigoplus_{P \in \mathcal{P}^\cap(M)} A_P &\rightarrow \mathcal{G}(M) \\ (a_P)_{P \in \mathcal{P}^\cap(M)} &\mapsto \sum_{P \in \mathcal{P}^\cap(M)} \bar{e}_P(a_P) \end{aligned}$$

is always surjective.

Proof. If $M = \emptyset$, then $\Psi : 0 \rightarrow 0$ is surjective, so assume $M \neq \emptyset$. Point planes are also included: for $P = \{x\}$, we have $U_P = \mathbb{R}$ and $U_{\subsetneq P} = 0$, so $A_P = \mathbb{R}$. Ordering $\mathcal{P}^\cap(M)$ by inclusion gives a finite poset, so we can find a labeling $\mathcal{P}^\cap(M) = \{P_1, \dots, P_s\}$ such that $P_a \subsetneq P_b$ implies $a < b$. Define the subset $\mathcal{P}_i(M) = \{P_1, \dots, P_i\}$ and the colimit of the subdiagram using these planes

$$\mathcal{G}_i(M) = \operatorname{colim}_{P \in \mathcal{P}_i(M)} U_P,$$

so $\mathcal{G}_s(M) = \mathcal{G}(M)$. We will now induct on $i \in [s]$ by showing that, at each step,

$$\begin{aligned} \Psi_i : \bigoplus_{j=1}^i A_{P_j} &\rightarrow \mathcal{G}_i(M) \\ (a_j)_{j \in [i]} &\mapsto \sum_{j=1}^i \bar{e}_{P_j}^{(i)}(a_j) \end{aligned}$$

is surjective, where $\bar{e}_{P_j}^{(i)} : U_{P_j} \rightarrow \mathcal{G}_i(M)$ denotes the canonical map from U_{P_j} into the partial colimit. By definition, $A_{P_i} = U_{\subsetneq P_i}^\perp$, so over $\mathbb{R}$ we have

$$U_{P_i} = U_{\subsetneq P_i} \oplus A_{P_i}. \quad (13)$$

For the base case of $i = 1$, there are no proper subplanes of P_1 by our labeling, so $U_{\subsetneq P_1} = 0 \subseteq U_{P_1}$ and

$$\begin{aligned} \mathcal{G}_1(M) &= U_{P_1} \\ &= U_{\subsetneq P_1} \oplus A_{P_1} \\ &= 0 \oplus A_{P_1} \\ &= A_{P_1}, \end{aligned}$$

as required.

For $i > 1$, our inductive hypothesis is that Ψ_{i-1} is surjective. Let $\eta_i : \mathcal{G}_{i-1}(M) \rightarrow \mathcal{G}_i(M)$ be the canonical map induced by the inclusion of diagrams $\mathcal{P}_{i-1}(M) \subseteq \mathcal{P}_i(M)$. For every $j < i$, the canonical maps satisfy

$$\bar{e}_{P_j}^{(i)} = \eta_i \circ \bar{e}_{P_j}^{(i-1)}.$$

Therefore, by our inductive hypothesis,

$$\begin{aligned} \sum_{j=1}^{i-1} \bar{e}_{P_j}^{(i)}(U_{P_j}) &= \eta_i \left(\sum_{j=1}^{i-1} \bar{e}_{P_j}^{(i-1)}(U_{P_j}) \right) \\ &\subseteq \eta_i(\operatorname{Im}(\Psi_{i-1})) \\ &\subseteq \operatorname{Im}(\Psi_i). \end{aligned} \quad (14)$$

Now it remains to analyze the image of U_{P_i} in $\mathcal{G}_i(M)$. We first show that the image of $U_{\subsetneq P_i}$ in $\mathcal{G}_i(M)$ is contained in $\text{Im}(\Psi_i)$. Take any $u \in U_{\subsetneq P_i}$. By the definition of $U_{\subsetneq P_i}$, we can write

$$u = \sum_{\substack{P' \in \mathcal{P}^\cap(M) \\ P' \subsetneq P_i}} \tau_{P' \rightarrow P_i}(u_{P'})$$

for some vectors $u_{P'} \in U_{P'}$. In the colimit $\mathcal{G}_i(M)$, the colimit relations (from R_M) give

$$\bar{e}_{P_i}^{(i)}(\tau_{P' \rightarrow P_i}(u_{P'})) = \bar{e}_{P'}^{(i)}(u_{P'}).$$

Therefore, using Equation (14) we get

$$\begin{aligned} \bar{e}_{P_i}^{(i)}(u) &= \sum_{\substack{P' \in \mathcal{P}^\cap(M) \\ P' \subsetneq P_i}} \bar{e}_{P'}^{(i)}(u_{P'}) \\ &\in \text{Im}(\Psi_i), \end{aligned}$$

showing that $\bar{e}_{P_i}^{(i)}(U_{\subsetneq P_i}) \subseteq \text{im}(\Psi_i)$. The surjectivity of Ψ_i now follows from Equation (13) and the fact that, by definition, $\bar{e}_{P_i}^{(i)}(A_{P_i}) \subseteq \text{Im}(\Psi_i)$. Finally, taking $i = s$ gives us that Ψ is surjective. This proof uses only the insertion maps and colimit relations, so it also applies to the formal cokernel of Definition 4.10 after any real specialization. $\square$

5 Two-Sided Product Expansion Over Random Codes

In this section, we prove our main result in Theorem 5.1 by reducing it to a single statement: that over one prime field, the corestriction maps of a random tuple of codes and of its dual are simultaneously injective for every ε -closed set. That statement is Theorem 6.1, which we prove in the next section.

Before stating the theorem, we fix the parameters at which the real input will be applied. Let $0 < \rho < 1/2$, and define $C_\rho = 3(1 + \rho^{-1/2})$ and $A_{D,\rho} = 2(4C_\rho)^D(1 + 4/\rho)^D$ so that we can set

$$\varepsilon_0(D, \rho) = \frac{1}{2^{D+2}A_{D,\rho}^2}, \quad N_{D,\varepsilon}^* = \left\lceil \max \left\{ 16 \left(\frac{192D}{e} \right)^{24}, (2^D \varepsilon)^{-2} \right\} \right\rceil. \quad (15)$$

We will see later in Section 6 that these parameters are sufficient to guarantee the existence of a real witness for every ε -closed set, which is the key to proving injectivity of the corestriction maps.

With the parameters fixed, our main result follows.

Theorem 5.1. Let $D \geq 1$, let $0 < \rho < 1/2$, let $0 < \varepsilon \leq \varepsilon_0(D, \rho)$, let $n \geq N_{D,\varepsilon}^*$, and let $k_1, \dots, k_D$ be integers with $\rho n \leq k_i \leq (1 - \rho)n$. Then there is a prime

$$p = O_D \left(2^{n^D} n^D \log n \right)$$

such that independent uniformly random codes $\mathcal{C}_i \in \text{Gr}_p(n, k_i)$ are two-sided $\gamma(\varepsilon, D)$ -product expanding with probability at least

$$1 - \frac{1}{\log(n+1)}, \quad \text{where} \quad \gamma(\varepsilon, D) = \frac{\varepsilon^D}{D(2^D + 1)^D}.$$

The one ingredient we need besides Theorem 6.1 is the implication from inner-generatedness to product expansion, which we take from [KP25].

Lemma 5.2 (Lemma 19 of [KP25]). Let $\mathcal{C}_1, \dots, \mathcal{C}_D \subseteq \mathbb{K}^n$ be codes over any field, and let $0 < \varepsilon \leq 1$. If every ε -closed subset of $[n]^D$ is inner-generated in $\boxplus_i \mathcal{C}_i$, then $(\mathcal{C}_i)_{i \in [D]}$ is $\gamma(\varepsilon, D)$ -product expanding.

Although [KP25] works in characteristic two throughout, the proof of their Lemma 19 never uses the characteristic of the base field and immediately generalizes to any finite field.

Proof of Theorem 5.1. Let p be the prime given by Theorem 6.1 and apply that theorem with $t = 1$. Then with probability at least $1 - 1/\log(n+1)$, every ε -closed set M has injective corestriction maps for both $(\mathcal{C}_i)_{i \in [D]}$ and $(\mathcal{C}_i^\perp)_{i \in [D]}$.

Conditioned on this event, Fact 4.4 to $(\mathcal{C}_i)_{i \in [D]}$, whose dual codes are the $\mathcal{C}_i^\perp$, makes every ε -closed set inner-generated in $\boxplus_i \mathcal{C}_i^\perp$. Applying the same fact to $(\mathcal{C}_i^\perp)_{i \in [D]}$ makes every ε -closed set inner-generated in $\boxplus_i \mathcal{C}_i$. So Lemma 5.2 applies to each of the two tuples in turn and gives $\gamma(\varepsilon, D)$ -product expansion of both, which is exactly two-sided expansion. $\square$

6 Injectivity of the Corestriction Map Via the Plane-Interaction Matrix

As we saw in the previous section, our remaining goal is to show the simultaneous injectivity of τ_M over a single prime field for all ε -closed sets M and both $(\mathcal{C}_i)_{i \in [D]}$ and $(\mathcal{C}_i^\perp)_{i \in [D]}$. In particular, we show the following.

Theorem 6.1. Let $D \geq 1$, $0 < \rho < 1/2$, $0 < \varepsilon \leq \varepsilon_0(D, \rho)$, and $n \geq N_{D, \varepsilon}^*$, with the parameters from Equation (15), and let $k_1, \dots, k_D$ be integers with $\rho n \leq k_i \leq (1 - \rho)n$. Write $k_i^+ = k_i$ and $k_i^- = n - k_i$ for the dimensions of $(\mathcal{C}_i)_{i \in [D]}$ and $(\mathcal{C}_i^\perp)_{i \in [D]}$, respectively. Then there is a prime

$$p = O_D \left(2^{n^D} n^D \log n \right)$$

such that the following holds: for every integer $t \geq 1$, if $\mathcal{C}_i \sim \text{Gr}_{p^t}(n, k_i)$ independently for $i \in [D]$, then, with probability at least

$$1 - \frac{p^{1-t}}{\log(n+1)},$$

the corestriction map τ_M is injective for every ε -closed set $M \subseteq [n]^D$, simultaneously for $(\mathcal{C}_i)_{i \in [D]}$ and $(\mathcal{C}_i^\perp)_{i \in [D]}$.

Recall from Definition 4.11 that a witness is not required to have full row rank. Our analysis of the *plane-interaction matrix* over the reals is what we use to show the existence of such witnesses. The rest of our argument involves going back to a finite field from the reals, which we give at the end of this section. We begin by defining the random model over the reals that we will use.

Setup Our random model over $\mathbb{R}$ samples matrices with independent Rademacher entries and then normalizes the columns to have norm one.

Definition 6.2 (Random model). For fixed integers $1 \leq k_i \leq n$ for $i \in [D]$, take $G_i \in \mathbb{R}^{k_i \times n}$ to have columns $(g_i(j))_{j \in [n]} \in \mathbb{R}^{k_i}$ where

$$g_i(j) = k_i^{-1/2} (\sigma_{1,j}^{(i)}, \sigma_{2,j}^{(i)}, \dots, \sigma_{k_i,j}^{(i)}) \in \mathbb{R}^{k_i}$$

and each $\sigma_{r,j}^{(i)} \sim \{-1, 1\}$ is sampled identically and independently. By construction, $\|g_i(j)\| = 1$. The row-space code $\mathcal{C}_i$ has dimension $\text{rank } G_i$, which may be smaller than k_i .

We set the formal variables $\mathbf{X}$ of Definition 4.10 to $G = (G_i)_{i \in [D]}$, keeping $U_i = \mathbb{R}^{k_i}$ and $U_\emptyset = \mathbb{R}$ regardless of the ranks of the matrices. For the real argument up to Section 6.4, we use the notation

$$\mathcal{G}(M) := \text{Coker } \alpha_M^G, \quad \tau_M := \bar{\beta}_M^G : \mathcal{G}(M) \rightarrow U_{[D]}.$$

These agree with the cosheaf and corestriction map of the row-space codes when every G_i has full row rank. We do not condition the signs on this event. The atoms are defined from these insertion maps as in Definition 4.12.

With this setup, our first ingredient is the following bound on the probability that τ_M is injective over the reals.

Lemma 6.3. For constants $D \in \mathbb{N}$, $\rho \in (0, 1)$, and $\varepsilon \in (0, 2^{-D})$, let

$$n \geq N_{D,\varepsilon} := \max \left\{ 16 \left(\frac{192D}{e} \right)^{24}, (2^D \varepsilon)^{-2} \right\}.$$

Fix an ε -closed set $M \subseteq [n]^D$, and let $\rho n \leq k_i \leq n$ for every $i \in [D]$. Then

$$\Pr_\sigma[\tau_M \text{ injective}] \geq 1 - C_{D,\rho,\varepsilon},$$

where

$$C_{D,\rho,\varepsilon} = 2(4C_\rho)^D \left(1 + \frac{4}{\rho} \right)^D \sqrt{2^D \varepsilon} \quad \text{and} \quad C_\rho := 3 \left(1 + \rho^{-1/2} \right).$$

For the parameters in Equation (15), this lower bound is at least $1/2$.

Our second ingredient transfers this result over the reals to a finite field. We first state the transfer for a single fixed set M , where the resulting prime is allowed to depend on M . Then, the proof of Theorem 6.1 at the end of this section upgrades it to one prime for every ε -closed set and both $(\mathcal{C}_i)_{i \in [D]}$ and $(\mathcal{C}_i^\perp)_{i \in [D]}$ simultaneously.

Lemma 6.4. Fix constants $D \in \mathbb{N}$ and $\varepsilon, \rho \in (0, 1)$, let $k_i \geq \rho n$ for every $i \in [D]$, and let $\emptyset \neq M \subsetneq [n]^D$ be ε -closed. Suppose that there is a real sign witness for $(M, (k_i)_{i \in [D]})$. Then, for all sufficiently large n , there is a prime

$$p = O_{D,\varepsilon}(n^D \log n)$$

such that, for independent uniformly random k_i -dimensional codes $\mathcal{C}_i \subseteq \mathbb{F}_p^n$,

$$\Pr[\tau_M \text{ is injective}] \geq 1 - \frac{2}{\log(n+1)}.$$

To obtain the real sign witnesses for Theorem 6.1, we introduce the plane-interaction matrix K which will be the main tool we use to prove the injectivity of τ_M over $\mathbb{R}$. In particular, the construction of the plane-interaction matrix will allow us to make explicit the obstructions to the injectivity of τ_M .

In Section 6.1 we observe how the corestriction maps of intersecting planes interact to motivate our derivation of the definition of the plane-interaction matrix over $\mathbb{R}$ in Section 6.2. Then, in Section 6.3, we prove Lemma 6.3 by showing that if the operator norm of K is less than one, then τ_M is injective. Finally, in Section 6.4 we prove the fixed-set transfer in Lemma 6.4 and the simultaneous statement in Theorem 6.1.

6.1 Interactions between the corestriction maps of planes

We begin by analyzing the interactions between different planes under our random Rademacher model to motivate the definition of the plane-interaction matrix. We first make a quick notational definition for succinctness.

Definition 6.5. For a plane $P \in \mathcal{P}^\cap(M)$, we write its global corestriction map $\tau_{P \rightarrow [n]^D} : U_P \rightarrow U_{[D]}$ as simply τ_P .

Now we show that the inner product of the global corestriction maps of incident planes vanishes when restricted to atoms.

Lemma 6.6. For distinct intersecting $P, Q \in \mathcal{P}^\cap(M)$, we have that for all $x \in A_P$ and $y \in A_Q$:

$$\langle \tau_P x, \tau_Q y \rangle = 0.$$

The main technical tool required for this is the fact that we can work with the inner product of the adjoint of the corestriction maps instead of the corestriction maps themselves.

Lemma 6.7. For two intersecting planes $P = \mathcal{P}(I, a), Q = \mathcal{P}(J, b) \in \mathcal{P}^\cap(M)$, let $K = I \cap J$ and define the plane at their intersection $R = \mathcal{P}(K, c)$ where

$$c_i = \begin{cases} a_i = b_i & i \in (I \cup J)^c \\ a_i & i \in J \setminus I \\ b_i & i \in I \setminus J. \end{cases}$$

Then we have that for all $x \in U_P$ and $y \in U_Q$

$$\langle \tau_P x, \tau_Q y \rangle = \langle \gamma_{P \rightarrow R} x, \gamma_{Q \rightarrow R} y \rangle$$

where $\gamma_{P \rightarrow R}$ and $\gamma_{Q \rightarrow R}$ are the adjoints of $\tau_{R \rightarrow P}$ and $\tau_{R \rightarrow Q}$, respectively, given on pure tensors by⁹

$$\begin{aligned} \gamma_{P \rightarrow R} &= \tau_{R \rightarrow P}^\dagger : U_P \rightarrow U_R & \gamma_{Q \rightarrow R} &= \tau_{R \rightarrow Q}^\dagger : U_Q \rightarrow U_R \\ x &\mapsto \left(\prod_{i \in I \setminus J} \langle x_i, g_i(c_i) \rangle \right) \bigotimes_{k \in K} x_k & y &\mapsto \left(\prod_{j \in J \setminus I} \langle y_j, g_j(c_j) \rangle \right) \bigotimes_{k \in K} y_k. \end{aligned}$$

We now prove Lemma 6.6.

Proof of Lemma 6.6. Let $R = P \cap Q = \mathcal{P}(I \cap J, c)$, using the same notation as in Lemma 6.7. Using the assumption on P and Q that they have non-empty intersection, at least one of $R \subsetneq P$ or $R \subsetneq Q$ must hold.

Case $R \subsetneq P$: In this case, by Definition 4.12 we have that $\tau_{R \rightarrow P}(U_R) \subseteq U_{\subsetneq P}$. Since $A_P = U_{\subsetneq P}^\perp$, we have that for any $x \in A_P$ and $z \in U_R$

$$\begin{aligned} 0 &= \langle x, \tau_{R \rightarrow P} z \rangle \\ &= \langle \gamma_{P \rightarrow R} x, z \rangle, \end{aligned}$$

by the definition of adjoints. However, since this holds for any $z \in U_R$, it must be that $\gamma_{P \rightarrow R} x = 0$. Consequently, using Lemma 6.7 we get

$$\begin{aligned} \langle \tau_P x, \tau_Q y \rangle &= \langle \gamma_{P \rightarrow R} x, \gamma_{Q \rightarrow R} y \rangle \\ &= \langle 0, \gamma_{Q \rightarrow R} y \rangle \\ &= 0, \end{aligned}$$

as claimed.

⁹We use the standard Euclidean inner products on the real spaces U_P , U_Q , and U_R .

Case $R \subsetneq Q$: Symmetrically in this case we can derive that $\gamma_{Q \rightarrow R}y = 0$, so

$$\begin{aligned}\langle \tau_P x, \tau_Q y \rangle &= \langle \gamma_{P \rightarrow R} x, \gamma_{Q \rightarrow R} y \rangle \\ &= \langle \gamma_{P \rightarrow R} x, 0 \rangle \\ &= 0,\end{aligned}$$

as required. $\square$

On the other hand, if instead of distinct planes P and Q we have that $P = Q$, then the corestriction map is an isometry when restricted to atoms.

Lemma 6.8. If instead $P = Q$, then for all $x, y \in A_P$ we have that

$$\langle \tau_P x, \tau_P y \rangle_{U_{[D]}} = \langle x, y \rangle_{U_P}.$$

Proof. We directly apply Lemma 6.7 to compute

$$\begin{aligned}\langle \tau_P x, \tau_P y \rangle_{U_{[D]}} &= \langle \gamma_{P \rightarrow P} x, \gamma_{P \rightarrow P} y \rangle_{U_P} \\ &= \langle x, y \rangle_{U_P}.\end{aligned}$$

$\square$

We end by providing the proof of Lemma 6.7.

Proof of Lemma 6.7. To examine the result of the desired inner product $\langle \tau_P x, \tau_Q y \rangle$, we begin by taking pure tensors

$$x = \bigotimes_{i \in I} x_i \in U_P = U_I \quad y = \bigotimes_{j \in J} y_j \in U_Q = U_J$$

and analyzing $\tau_P x$ and $\tau_Q y$. Recalling Equation (11), we get that

$$\tau_P x = x \otimes \left(\bigotimes_{i \in I^c} g_i(a_i) \right) \quad \tau_Q y = y \otimes \left(\bigotimes_{j \in J^c} g_j(b_j) \right).$$

This allows us to compute their inner product explicitly as

$$\begin{aligned}\langle \tau_P x, \tau_Q y \rangle &= \left\langle x \otimes \left(\bigotimes_{i \in I^c} g_i(a_i) \right), y \otimes \left(\bigotimes_{j \in J^c} g_j(b_j) \right) \right\rangle \\ &= \left(\prod_{k \in K} \langle x_k, y_k \rangle \right) \left(\prod_{i \in I \setminus J} \langle x_i, g_i(c_i) \rangle \right) \left(\prod_{j \in J \setminus I} \langle g_j(c_j), y_j \rangle \right) \left(\prod_{i \in (I \cup J)^c} \langle g_i(c_i), g_i(c_i) \rangle \right) \\ &= \left(\prod_{k \in K} \langle x_k, y_k \rangle \right) \left(\prod_{i \in I \setminus J} \langle x_i, g_i(c_i) \rangle \right) \left(\prod_{j \in J \setminus I} \langle g_j(c_j), y_j \rangle \right) \left(\prod_{i \in (I \cup J)^c} \|g_i(c_i)\|^2 \right) \\ &= \left(\prod_{k \in K} \langle x_k, y_k \rangle \right) \left(\prod_{i \in I \setminus J} \langle x_i, g_i(c_i) \rangle \right) \left(\prod_{j \in J \setminus I} \langle g_j(c_j), y_j \rangle \right) \tag{16}\end{aligned}$$

$$\begin{aligned}&= \left\langle \left(\prod_{i \in I \setminus J} \langle x_i, g_i(c_i) \rangle \right) \bigotimes_{k \in K} x_k, \left(\prod_{j \in J \setminus I} \langle y_j, g_j(c_j) \rangle \right) \bigotimes_{k \in K} y_k \right\rangle \\ &= \langle \gamma_{P \rightarrow R} x, \gamma_{Q \rightarrow R} y \rangle,\end{aligned} \tag{17}$$

where to get Equation (16) we used the fact that the columns are normalized as $\|g_i(c_i)\| = 1$. Moreover, we have that Equation (17) holds for any $x \in U_P$ and $y \in U_Q$ because pure tensors span U_P and U_Q . $\square$

6.2 Constructing the plane-interaction matrix

With these observations of the interactions between the corestriction maps of intersecting planes, we move on to defining the plane-interaction matrix K . We begin by introducing some useful notation.

Definition 6.9. Recall that for $P \in \mathcal{P}^\cap(M)$ we have $A_P := U_{\bar{C}P}^\perp \subseteq U_P$, and let $\Pi_P : U_P \rightarrow A_P$ be the orthogonal projection map. Define

$$\begin{aligned} U = U_M &= \bigoplus_{P \in \mathcal{P}^\cap(M)} U_P \\ A &= \bigoplus_{P \in \mathcal{P}^\cap(M)} A_P \\ \Pi_A &= \bigoplus_{P \in \mathcal{P}^\cap(M)} \Pi_P : U \rightarrow A. \end{aligned}$$

We drop the subscript M when it is clear from context.

Next, we examine all interactions between the corestriction maps.

Definition 6.10. For $P \in \mathcal{P}^\cap(M)$, define the map $\theta_P := \tau_P|_{A_P} : A_P \rightarrow U_{[D]}$ and the global operator

$$\begin{aligned} \Theta : A &\rightarrow U_{[D]} \\ (a_P)_{P \in \mathcal{P}^\cap(M)} &\mapsto \sum_{P \in \mathcal{P}^\cap(M)} \theta_P a_P. \end{aligned}$$

Moreover, we define its Gram operator

$$\Upsilon := \Theta^\dagger \Theta : A \rightarrow A.$$

We can immediately observe that for $P, Q \in \mathcal{P}^\cap(M)$, the (P, Q) -th block of Υ is $\Upsilon_{P,Q} = \theta_P^\dagger \theta_Q : A_Q \rightarrow A_P$. Consequently, $\Upsilon_{P,P} = \theta_P^\dagger \theta_P = \text{Id}_{A_P}$ by Lemma 6.8 and $\Upsilon_{P,Q} = 0$ when $P \neq Q$, and $P \cap Q \neq \emptyset$ by Lemma 6.6. Since the diagonal of Υ is only the identity, only the off-diagonal blocks are interesting. This motivates the following definition.

Definition 6.11 (Plane-interaction matrix). Let $H := \Upsilon - \text{Id}_A : A \rightarrow A$, so by Lemma 6.6 and Lemma 6.8 for $P, Q \in \mathcal{P}^\cap(M)$ we have

$$H_{P,Q} = \begin{cases} 0 & P \cap Q \neq \emptyset \\ \theta_P^\dagger \theta_Q & P \cap Q = \emptyset. \end{cases}$$

Similarly, on the entire space $U \supseteq A$, we will define the *plane-interaction matrix* $K : U \rightarrow U$ via blocks as

$$K_{P,Q} = \begin{cases} 0 & P \cap Q \neq \emptyset \\ \tau_P^\dagger \tau_Q & P \cap Q = \emptyset. \end{cases}$$

6.3 Injectivity of τ_M over the reals

In this subsection, we will prove that τ_M is injective with probability at least $1 - C_{D,\rho,\varepsilon}$ under our real Rademacher model.

Lemma 6.3. For constants $D \in \mathbb{N}$, $\rho \in (0, 1)$, and $\varepsilon \in (0, 2^{-D})$, let

$$n \geq N_{D,\varepsilon} := \max \left\{ 16 \left(\frac{192D}{e} \right)^{24}, (2^D \varepsilon)^{-2} \right\}.$$

Fix an ε -closed set $M \subseteq [n]^D$, and let $\rho n \leq k_i \leq n$ for every $i \in [D]$. Then

$$\Pr_{\sigma}[\tau_M \text{ injective}] \geq 1 - C_{D,\rho,\varepsilon},$$

where

$$C_{D,\rho,\varepsilon} = 2(4C_{\rho})^D \left(1 + \frac{4}{\rho}\right)^D \sqrt{2^D \varepsilon} \quad \text{and} \quad C_{\rho} := 3 \left(1 + \rho^{-1/2}\right).$$

This follows from two facts. First, that if K has operator norm less than one, then τ_M is injective.

Lemma 6.12. If $\|K\|_{\text{op}} < 1$, then τ_M is injective.

Second, that the operator norm of K is less than one with probability at least $1 - C_{D,\rho,\varepsilon}$. Proving this theorem is the main focus of Section 7.

Theorem 6.13. For constants $D \in \mathbb{N}$, $\rho \in (0, 1)$, and $\varepsilon \in (0, 2^{-D})$, let

$$n \geq N_{D,\varepsilon} := \max \left\{ 16 \left(\frac{192D}{e} \right)^{24}, (2^D \varepsilon)^{-2} \right\}.$$

Fix an ε -closed set $M \subseteq [n]^D$, and let $\rho n \leq k_i \leq n$ for every $i \in [D]$. Then

$$\Pr_{\sigma}[\|K\|_{\text{op}} < 1] \geq 1 - C_{D,\rho,\varepsilon},$$

where

$$C_{D,\rho,\varepsilon} = 2(4C_{\rho})^D \left(1 + \frac{4}{\rho}\right)^D \sqrt{2^D \varepsilon} \quad \text{and} \quad C_{\rho} := 3 \left(1 + \rho^{-1/2}\right).$$

The proof of Lemma 6.3 now follows directly.

Proof of Lemma 6.3. Apply Lemma 6.12 to Theorem 6.13. □

6.3.1 Injectivity of τ_M from the operator norm of K

The rest of this subsection is dedicated to proving Lemma 6.12. The only tool we need is that the operator norm of K bounds the operator norm of H .

Claim 6.14. We have $\|H\|_{\text{op}} \leq \|K\|_{\text{op}}$.

Proof. Let $\iota_A : A \hookrightarrow U$ be the inclusion. By construction, we have that $H = \Pi_A K \iota_A$, so we can compute

$$\begin{aligned} \|H\|_{\text{op}} &= \|\Pi_A K \iota_A\|_{\text{op}} \\ &\leq \|\Pi_A\|_{\text{op}} \|K\|_{\text{op}} \|\iota_A\|_{\text{op}} \\ &\leq \|K\|_{\text{op}}, \end{aligned}$$

since both Π_A and ι_A have operator norm at most 1. □

We can now condition the injectivity of τ_M on the operator norm of H and K .

Proof of Lemma 6.12. By Claim 6.14, we have that $\|H\|_{\text{op}} \leq \|K\|_{\text{op}} < 1$. Since H is self-adjoint, for any $a \in A$ we can lower bound its quadratic form as

$$\langle a, Ha \rangle \geq -\|H\|_{\text{op}} \|a\|^2. \tag{18}$$

Therefore, Θ is injective because $1 - \|H\|_{\text{op}} > 0$ and

$$\begin{aligned}
\|\Theta a\|^2 &= \langle \Theta a, \Theta a \rangle \\
&= \langle a, \Upsilon a \rangle \\
&= \langle a, (\text{Id}_A + H)a \rangle \\
&= \|a\|^2 + \langle a, Ha \rangle \\
&\geq \|a\|^2 - \|H\|_{\text{op}} \|a\|^2 && \text{By Equation (18)} \\
&= (1 - \|H\|_{\text{op}}) \|a\|^2.
\end{aligned}$$

To conclude that τ_M is injective, we consider any $z \in \text{Ker}(\tau_M) \subseteq \mathcal{G}(M)$. The surjectivity of Ψ gives an $a \in A$ such that $\Psi(a) = z$. Finally, compatibility of the maps in the colimit defining $\mathcal{G}(M)$ gives

$$\begin{aligned}
\Theta(a) &= \tau_M(\Psi(a)) \\
&= \tau_M(z) \\
&= 0.
\end{aligned}$$

This forces $a = 0$ since Θ is injective, so $z = \Psi(a) = \Psi(0) = 0$, meaning that τ_M is injective, as desired. $\square$

This leads to a nice observation that τ_M is injective if all maximal planes pairwise intersect, showing that the barrier to τ_M being injective is exactly pairs of maximal planes with empty intersection.

Corollary 6.15. If all planes in $\mathcal{P}(M)$ have pairwise nonempty intersections, then $K = 0$ so τ_M is injective.

6.4 From the reals to finite fields

We finish this section by proving that we can transfer our result about the injectivity of τ_M over the reals to a finite field of large enough characteristic.

To simplify analysis, we note that we can remove the normalization of the sign matrices without affecting the injectivity of τ_M since they're just a rescaling. In particular, we can write $\sigma^{(i)} = \sqrt{k_i} G_i$ and, for $I \subseteq [D]$, set

$$S_I = \left(\prod_{i \in I} \sqrt{k_i} \right) \text{Id}_{U_I}, \quad S_\emptyset = \text{Id}_{\mathbb{R}}.$$

Then for $P = \mathcal{P}(I, a) \subseteq Q = \mathcal{P}(J, b)$, we can relate the corestriction map with the unnormalized signs $\tau_{P \rightarrow Q}^\sigma$ to the corestriction map with the normalized signs $\tau_{P \rightarrow Q}^G$ as

$$\tau_{P \rightarrow Q}^\sigma = S_J \tau_{P \rightarrow Q}^G S_I^{-1}.$$

Thus, these maps give an isomorphism of the normalized and unnormalized diagrams. Consequently, $\overline{\beta}_M^G$ is injective if and only if $\overline{\beta}_M^\sigma$ is injective, so any injective normalized instance gives a real sign witness in the sense of Definition 4.11. No full-row-rank assumption is needed here.

Lemma 6.4. Fix constants $D \in \mathbb{N}$ and $\varepsilon, \rho \in (0, 1)$, let $k_i \geq \rho n$ for every $i \in [D]$, and let $\emptyset \neq M \subsetneq [n]^D$ be ε -closed. Suppose that there is a real sign witness for $(M, (k_i)_{i \in [D]})$. Then, for all sufficiently large n , there is a prime

$$p = O_{D, \varepsilon}(n^D \log n)$$

such that, for independent uniformly random k_i -dimensional codes $\mathcal{C}_i \subseteq \mathbb{F}_p^n$,

$$\Pr[\tau_M \text{ is injective}] \geq 1 - \frac{2}{\log(n+1)}.$$

We will prove Lemma 6.4 in two steps. First, we will show that a real sign witness gives a non-identically zero determinant polynomial over the integers with bounded degree and a specific value when evaluated at the witness. We state this for an arbitrary set and arbitrary dimensions, since the proof of Theorem 6.1 will need to apply it to both $(\mathcal{C}_i)_{i \in [D]}$ and $(\mathcal{C}_i^\perp)_{i \in [D]}$.

Lemma 6.16. Let $M \subseteq [n]^D$, let $(k_i)_{i \in [D]}$ be dimensions, and let σ be a real sign witness for $(M, (k_i)_{i \in [D]})$. Put

$$N_M := \dim U_M = \sum_{P=\mathcal{P}(I,a) \in \mathcal{P}^\cap(M)} \prod_{i \in I} k_i.$$

There is a polynomial $f_\sigma \in \mathbb{Z}[\mathbf{X}]$ such that

$$1 \leq |f_\sigma(\sigma)| \leq N_M^{N_M/2}, \quad \deg f_\sigma \leq DN_M, \quad (19)$$

and such that, over any field, every specialization x of $\mathbf{X}$ with $f_\sigma(x) \neq 0$ makes the induced map $\bar{\beta}_M^x$ of Definition 4.10 injective. If, in addition, every component of x has full row rank, then $\bar{\beta}_M^x$ is the corestriction map τ_M of the corresponding row-space codes, so τ_M is injective.

Proof. Throughout the proof M is fixed, so we drop it from the notation and write $\alpha^{\mathbf{X}}, \beta^{\mathbf{X}}$ for the maps $\alpha_M^{\mathbf{X}}, \beta_M^{\mathbf{X}}$ of Definition 4.10, and α^x, β^x for their specializations at a tuple of matrices x . Recall that $\beta^{\mathbf{X}} \alpha^{\mathbf{X}} = 0$ is a polynomial identity and so holds after every specialization.

We work first at the witness σ and set $a := \text{rank } \alpha^\sigma$ and $b := \text{rank } \beta^\sigma$. Because $\ker \bar{\beta}^\sigma = \ker \beta^\sigma / \text{Im } \alpha^\sigma$, the injectivity of $\bar{\beta}^\sigma$ says exactly that $\ker \beta^\sigma \subseteq \text{Im } \alpha^\sigma$, while $\beta^\sigma \alpha^\sigma = 0$ gives the reverse inclusion. Hence $\text{Im } \alpha^\sigma = \ker \beta^\sigma$, and the rank-nullity theorem gives $a + b = N_M$.

Choose a independent columns of α^σ and b independent rows of β^σ , and let $\tilde{\alpha}^{\mathbf{X}}$ and $\tilde{\beta}^{\mathbf{X}}$ be the submatrices of $\alpha^{\mathbf{X}}$ and $\beta^{\mathbf{X}}$ on those indices. The columns of $\tilde{\alpha}^\sigma$ span $\text{Im } \alpha^\sigma = \ker \beta^\sigma$, while the columns of $(\tilde{\beta}^\sigma)^\top$ span the row space of β^σ , which is the orthogonal complement of $\ker \beta^\sigma$ in $\mathbb{R}^{N_M}$. Consequently, the square matrix

$$\left[\tilde{\alpha}^\sigma \mid (\tilde{\beta}^\sigma)^\top \right]$$

is invertible, so the determinant polynomial

$$f_\sigma(\mathbf{X}) := \det \left[\tilde{\alpha}^{\mathbf{X}} \mid (\tilde{\beta}^{\mathbf{X}})^\top \right] \in \mathbb{Z}[\mathbf{X}]$$

satisfies $f_\sigma(\sigma) \neq 0$. This is the only place where the real inner product is used.

Now let x be any specialization over any field with $f_\sigma(x) \neq 0$. Then the selected columns of α^x and rows of β^x are still independent, so

$$\text{rank } \alpha^x \geq a \quad \text{and} \quad \text{rank } \beta^x \geq b.$$

On the other hand $\beta^x \alpha^x = 0$ forces $\text{Im } \alpha^x \subseteq \ker \beta^x$, so $\text{rank } \alpha^x + \text{rank } \beta^x \leq N_M$. Since $a + b = N_M$, we can conclude that $\text{Im } \alpha^x = \ker \beta^x$ and $\bar{\beta}^x$ is injective. Moreover, if every component of x has full row rank, then Lemma 4.8 identifies $\text{Coker } \alpha^x$ with $\mathcal{G}(M)$ and $\bar{\beta}^x$ with τ_M , so τ_M is injective.

It remains to prove the bounds in (19). By Equation (11), each corestriction map inserts at most D columns, so every entry of $\alpha^{\mathbf{X}}$ and of $\beta^{\mathbf{X}}$ is either zero or a signed monomial of degree at most D . Expanding the $N_M \times N_M$ determinant gives $\deg f_\sigma \leq DN_M$. At the sign instance σ , the matrix defining f_σ has entries in $\{0, \pm 1\}$, so each of its N_M columns has Euclidean norm at most $\sqrt{N_M}$, and Hadamard's inequality gives $|f_\sigma(\sigma)| \leq N_M^{N_M/2}$. Finally, $f_\sigma(\sigma)$ is a nonzero integer, so $|f_\sigma(\sigma)| \geq 1$. $\square$

Second, we will show that we can upper bound the characteristic of the finite field. We need to ensure that, when modding by this characteristic, the determinant polynomial does not become zero.

Lemma 6.17. Let $d_\sigma := \deg f_\sigma$ and $m_\sigma := |f_\sigma(\sigma)|$. There is a prime $p \nmid m_\sigma$ satisfying

$$\max\{2D + 1, d_\sigma \log(n + 1)\} < p, \quad p = O_{D,\varepsilon}(n^D \log n). \quad (20)$$

In particular, f_σ modulo p is not identically zero.

Proof. From Corollary 3.26, there exists a universal constant C such that for $q = \max\{2D + 1, d_\sigma \log(n + 1)\}$ there exists a prime $p \nmid m_\sigma$ satisfying

$$q < p \leq C(q + \log(2m_\sigma)). \quad (21)$$

The lower bound of Equation (21) is the first claim, so it remains only to bound its upper endpoint. We first bound N_M by computing

$$\begin{aligned} N_M &= \dim \left(\bigoplus_{P \in \mathcal{P}^\cap(M)} U_P \right) \\ &= \sum_{I \subsetneq [D]} |\Omega_I^\cap(M)| \prod_{i \in I} k_i \\ &\leq 2^D \varepsilon n^D \sum_{I \subsetneq [D]} \prod_{i \in I} \frac{k_i}{n} && \text{By Corollary 3.14 and lemma 3.6} \\ &\leq 2^D (2^D - 1) \varepsilon n^D. && \text{By } k_i \leq n \end{aligned}$$

From Lemma 6.16 we have that $d_\sigma \leq DN_M$ and $m_\sigma \leq N_M^{N_M/2}$, so $\log(2m_\sigma) \leq \log 2 + \frac{N_M}{2} \log N_M$. Thus,

$$\begin{aligned} C(q + \log(2m_\sigma)) &\leq C(\max\{2D + 1, d_\sigma \log(n + 1)\} + \log(2m_\sigma)) \\ &= O_{D,\varepsilon}(n^D \log(n)), \end{aligned}$$

as required.

Since $p \nmid m_\sigma$, the reduction of f_σ modulo p evaluates nontrivially at σ and is consequently not the zero polynomial. $\square$

Now, we can prove Lemma 6.4. Our use of the Schwartz-Zippel lemma here is similar to that of Lemma 21 of [KP25].

Proof of Lemma 6.4. Let σ be a real sign witness for $(M, (k_i)_{i \in [D]})$, let f_σ be as in Lemma 6.16, and choose p as in Lemma 6.17. Sample independent uniform matrices $G_i \in \mathbb{F}_p^{k_i \times n}$. Since $f_\sigma \bmod p$ is nonzero, the Schwartz-Zippel lemma and Lemma 6.17 give

$$\Pr[f_\sigma(G_1, \dots, G_D) = 0] \leq \frac{d_\sigma}{p} < \frac{1}{\log(n + 1)}.$$

Let E_{full} be the event that every G_i has full row rank. For each G_i individually, we have

$$\Pr[\text{rank } G_i < k_i] \leq \sum_{j=0}^{k_i-1} p^{j-n} < \frac{1}{p-1}.$$

Consequently, since $p > 2D + 1$ we can use a union bound to get

$$\begin{aligned} \Pr[E_{\text{full}}] &\geq 1 - \sum_{i=1}^D \Pr[\text{rank } G_i < k_i] \\ &> 1 - \frac{D}{p-1} \\ &> \frac{1}{2}. \end{aligned}$$

By Lemma 6.16,

$$\Pr[\tau_M \text{ is not injective} \mid E_{\text{full}}] \leq \frac{\Pr[f_\sigma(G_1, \dots, G_D) = 0]}{\Pr[E_{\text{full}}]} < \frac{2}{\log(n+1)}.$$

Conditioned on E_{full} , the row spaces of the G_i are independent uniformly random k_i -dimensional codes. Every such code has exactly $|\text{GL}_{k_i}(\mathbb{F}_p)|$ ordered generator matrices, so choosing random generator matrices gives the same distribution. $\square$

We can now choose a single characteristic that works for every ε -closed set and both $(\mathcal{C}_i)_{i \in [D]}$ and $(\mathcal{C}_i^\perp)_{i \in [D]}$ at once. The strategy is the product-of-polynomials method from the proof of Lemma 21 of [KP25], but applied to the witnesses of Lemma 6.16 while taking care to choose a large enough characteristic.

Proof of Theorem 6.1. The argument has five steps. We consider one set of dimensions for $(\mathcal{C}_i)_{i \in [D]}$ and a fixed ε -closed M , then produce one real witness for each of them, then choose a single prime at which every witness survives, and finally sample the codes and take a union bound over both the codes and their duals.

Step 1: A uniform dimension bound. For $M = \emptyset$ the map τ_M has zero domain, and for $M = [n]^D$ it is the identity, so both are injective for every choice of codes. We may therefore restrict attention to the family $\mathcal{M}_\varepsilon$ of nonempty proper ε -closed subsets of $[n]^D$, and we write $s := |\mathcal{M}_\varepsilon| \leq 2^{n^D}$.

Fix $M \in \mathcal{M}_\varepsilon$ and $\eta \in \{+, -\}$, and let N_M^η denote the dimension N_M of Lemma 6.16 for the profile $(k_i^\eta)_{i \in [D]}$. Using the fact that, for a given $I \subseteq [D]$, we have $|\Omega_I^\cap| \leq n^{D-|I|}$, and that $k_i^\eta \leq n$,

$$N_M^\eta = \sum_{P=\mathcal{P}(I,a) \in \mathcal{P}^\cap(M)} \prod_{i \in I} k_i^\eta \leq \sum_{I \subsetneq [D]} n^{D-|I|} \prod_{i \in I} n = (2^D - 1)n^D =: R.$$

This loose count suffices, and we do not need to use any sparsity properties for the anchor sets here. Note also that $R \geq n \geq 2$.

Step 2: an integer certificate for each set and profile. Both dimension profiles satisfy $k_i^\eta \geq \rho n$, and $n \geq N_{D,\varepsilon}^* \geq N_{D,\varepsilon}$. Moreover, Equation (15) gives $2^D \varepsilon \leq 1/(4A_{D,\rho}^2) < 1$ and $C_{D,\rho,\varepsilon} = A_{D,\rho} \sqrt{2^D \varepsilon} \leq 1/2$. Thus, Lemma 6.3 gives an injective normalized instance with probability at least $1/2$ for this fixed M and profile $(k_i^\eta)_{i \in [D]}$. Removing the normalization as above gives a real sign witness σ_M^η for $(M, (k_i^\eta)_{i \in [D]})$. Applying Lemma 6.16 to it produces a polynomial $f_M^\eta \in \mathbb{Z}[\mathbf{X}]$ with

$$1 \leq |f_M^\eta(\sigma_M^\eta)| \leq (N_M^\eta)^{N_M^\eta/2} \leq R^{R/2} \quad \text{and} \quad \deg f_M^\eta \leq DN_M^\eta,$$

using in the third inequality that $y \mapsto y^{y/2}$ is nondecreasing for $y \geq 1$. Over any field, any specialization x at which f_M^η does not vanish makes $\bar{\beta}_M^x$ injective, and makes it equal to τ_M as soon as x has full row rank. The witnesses are allowed to depend on both M and η . However, we need one characteristic in which all of the f_M^η survive.

Step 3: one prime avoiding all of the certificates. Set

$$m_* := \prod_{\eta \in \{+, -\}} \prod_{M \in \mathcal{M}_\varepsilon} |f_M^\eta(\sigma_M^\eta)|, \quad \text{and} \quad L := Dn + 2DsR,$$

with the convention that an empty product is 1, so that $m_* \geq 1$ is an integer. There are $2s$ factors, each at most $R^{R/2}$ by Step 2, so $\log m_* \leq sR \log R$. Applying Corollary 3.26 with $m = m_*$ and $q = 2 + L \log(n+1)$ gives a prime $p \nmid m_*$ with

$$L \log(n+1) < p \leq C(2 + L \log(n+1) + \log 2 + sR \log R) = O_D(2^{n^D} n^D \log n),$$

where the last estimate uses $s \leq 2^{n^D}$ and $R = (2^D - 1)n^D$. Since p divides none of the integers $f_M^\eta(\sigma_M^\eta)$, each f_M^η reduces to a nonzero polynomial over $\mathbb{F}_p$, and hence over each of its extensions $\mathbb{F}_{p^t}$. Observe that this step fixes p before any code is sampled, and that the same p works for every M and both dimension profiles.

Step 4: full rank and simultaneous injectivity over $\mathbb{F}_{p^t}$. Fix an integer $t \geq 1$ and a profile $\eta \in \{+, -\}$, and let $X_1, \dots, X_D$ with $X_i \in \mathbb{F}_{p^t}^{k_i^\eta \times n}$ have independent uniformly random entries. Writing $\Delta_i(\mathbf{X})$ for the determinant of the formal matrix $(X_{r,j}^{(i)})_{r,j \in [k_i^\eta]}$, set

$$h^\eta := \prod_{i=1}^D \Delta_i, \quad F^\eta := h^\eta \prod_{M \in \mathcal{M}_\varepsilon} f_M^\eta, \quad d_\eta := \sum_{i=1}^D k_i^\eta + D \sum_{M \in \mathcal{M}_\varepsilon} N_M^\eta,$$

where all of the f_M^η are formed in the same variables $\mathbf{X}$, so that a single sample specializes all of them at once. Each Δ_i contains the diagonal monomial $\prod_{r \in [k_i^\eta]} X_{r,r}^{(i)}$ with coefficient 1 and so is a nonzero polynomial in $\mathbb{F}_{p^t}[\mathbf{X}]$. Combined with Step 3 this shows that F^η is a nonzero polynomial in $\mathbb{F}_{p^t}[\mathbf{X}]$, and $\deg F^\eta \leq d_\eta$ by Step 2. The Schwartz–Zippel lemma therefore gives

$$\Pr[F^\eta(X_1, \dots, X_D) \neq 0] \geq 1 - \frac{d_\eta}{p^t}.$$

Let E_η be the event that every X_i has full row rank, and let A_η be the event that τ_M is injective for every $M \in \mathcal{M}_\varepsilon$ in the diagram of the row spaces of $X_1, \dots, X_D$. A nonzero evaluation $\Delta_i(X_1, \dots, X_D)$ forces X_i to have full row rank, so by Step 2 we have $\{F^\eta(X_1, \dots, X_D) \neq 0\} \subseteq A_\eta \cap E_\eta$. Since $k_i^\eta \leq n$, full-row-rank matrices exist and $\Pr[E_\eta] > 0$, so we may condition on E_η ; using $\Pr[E_\eta] \leq 1$ in the denominator,

$$\Pr[A_\eta \mid E_\eta] = \frac{\Pr[A_\eta \cap E_\eta]}{\Pr[E_\eta]} \geq \Pr[F^\eta(X_1, \dots, X_D) \neq 0] \geq 1 - \frac{d_\eta}{p^t}.$$

Moreover, conditioned on E_η the row spaces of $X_1, \dots, X_D$ are independent and uniformly random in $\text{Gr}_{p^t}(n, k_i^\eta)$, because every k_i^η -dimensional code has exactly $|\text{GL}_{k_i^\eta}(\mathbb{F}_{p^t})|$ ordered generator matrices. So this conditional sampling is the same as choosing a random full rank generator matrix for the profile $(k_i^\eta)_{i \in [D]}$.

Step 5: the union bound over the tuple and its dual. Let $\mathcal{C}_1, \dots, \mathcal{C}_D$ be independent and uniformly random with $\mathcal{C}_i \in \text{Gr}_{p^t}(n, k_i)$. Applying Step 4 with $\eta = +$, the maps τ_M for $(\mathcal{C}_i)_{i \in [D]}$ fail to be simultaneously injective with probability at most d_+/p^t . Since the dual of a random k_i dimensional space is a random $(n - k_i)$ dimensional space, the tuple $(\mathcal{C}_i^\perp)_{i \in [D]}$ is itself uniformly random with the dimensions $(k_i^-)_{i \in [D]}$. Applying Step 4 with $\eta = -$, its global corestriction maps fail to be injective with probability at most d_-/p^t . A union bound gives a total failure probability of at most

$$\frac{d_+ + d_-}{p^t} \leq \frac{L}{p^t} < \frac{p^{1-t}}{\log(n+1)}.$$

Here the first inequality is the computation

$$d_+ + d_- = \sum_{i=1}^D (k_i + (n - k_i)) + D \sum_{M \in \mathcal{M}_\varepsilon} (N_M^+ + N_M^-) \leq Dn + 2DsR = L,$$

using $N_M^\eta \leq R$ from Step 1, and the second holds because $p > L \log(n+1)$ by Step 3. Together with Step 1, this is exactly the claim. $\square$

7 Bounding the Operator Norm of the Plane-Interaction Matrix

In this section, we will take the same setup as the previous section and continue to work over $\mathbb{R}$ to lower bound the probability that the operator norm of the plane-interaction matrix K of M is less than one.

Theorem 6.13. For constants $D \in \mathbb{N}$, $\rho \in (0, 1)$, and $\varepsilon \in (0, 2^{-D})$, let

$$n \geq N_{D,\varepsilon} := \max \left\{ 16 \left(\frac{192D}{e} \right)^{24}, (2^D \varepsilon)^{-2} \right\}.$$

Fix an ε -closed set $M \subseteq [n]^D$, and let $\rho n \leq k_i \leq n$ for every $i \in [D]$. Then

$$\Pr_{\sigma}[\|K\|_{\text{op}} < 1] \geq 1 - C_{D,\rho,\varepsilon},$$

where

$$C_{D,\rho,\varepsilon} = 2(4C_{\rho})^D \left(1 + \frac{4}{\rho}\right)^D \sqrt{2^D \varepsilon} \quad \text{and} \quad C_{\rho} := 3 \left(1 + \rho^{-1/2}\right).$$

Since we will do this via Markov's inequality, it suffices to bound the expected operator norm of K . Thus, the main result that we will derive in this section is the following upper bound.

Lemma 7.1. For constants $D \in \mathbb{N}$, $\rho \in (0, 1)$, and $\varepsilon \in (0, 2^{-D})$, let

$$n \geq N_{D,\varepsilon} := \max \left\{ 16 \left(\frac{192D}{e} \right)^{24}, (2^D \varepsilon)^{-2} \right\}.$$

Fix an ε -closed set $M \subseteq [n]^D$, and let $\rho n \leq k_i \leq n$ for every $i \in [D]$. Then

$$\mathbb{E}_{\sigma}[\|K\|_{\text{op}}] \leq C_{D,\rho,\varepsilon}.$$

A single application of Markov's inequality gives us the necessary result for the previous section.

Proof of Theorem 6.13. We directly compute

$$\begin{aligned} \Pr_{\sigma}[\|K\|_{\text{op}} < 1] &\geq 1 - \mathbb{E}_{\sigma}[\|K\|_{\text{op}}] && \text{By Markov} \\ &\geq 1 - C_{D,\rho,\varepsilon} && \text{By Lemma 7.1} \end{aligned}$$

□

Our path to proving Lemma 7.1 begins in Section 7.1 by rearranging K into “superblocks” of the blocks $K_{P,Q}$ used to define it and then decomposing these blocks as sums of restricted operators whose operator norm we can cleanly express as a product of operator norms. Then, in Section 7.2, we bound the expected operator norm of these restricted operators. Finally, we complete several deferred proofs in Section 7.3.

7.1 Decomposition of the plane-interaction matrix

To bound the expected operator norm of K , we proceed via two decomposition steps. In Section 7.1.1, we simply change our view from the block definition of K via the plane blocks $K_{P,Q}$ for planes P and Q to what we call *directional blocks* $K_{I,J}$ defined by two directions $I, J \subseteq [D]$. There are only 4^D such pairs (I, J) , so bounding each $\|K_{I,J}\|_{\text{op}}$ lets us easily bound $\|K\|_{\text{op}}$ via the triangle inequality. To work on these $K_{I,J}$'s, however, we perform one more decomposition in Section 7.1.2 over subsets of $(I \cup J)^c$. This latter decomposition is where we use both the fact that M is ε -closed and that the entries of our generator matrices are random Rademacher variables.

7.1.1 Directional block decomposition

This first decomposition combines all plane blocks $K_{P,Q}$, where P is an I -plane and Q is a J -plane, together into one direction block $K_{I,J}$.

Definition 7.2. Since each plane block $K_{P,Q} : U_Q \rightarrow U_P$ and for two directions $I, J \subseteq [D]$ the planes in $\mathcal{P}^\cap(M)$ are indexed by $\Omega_I^\cap(M)$ and $\Omega_J^\cap(M)$, respectively, we can define the block $K_{I,J}$ combining all blocks $K_{P,Q}$ for an I -plane P and a J -plane Q as an operator $K_{I,J} : U_J \otimes \mathbb{R}^{\Omega_J^\cap} \rightarrow U_I \otimes \mathbb{R}^{\Omega_I^\cap}$ where for $a \in \Omega_I^\cap$ and $b \in \Omega_J^\cap$:

$$(K_{I,J})_{a,b} = K_{\mathcal{P}(I,a), \mathcal{P}(J,b)}.$$

Thus, $K_{I,J}$ is $(|\Omega_I^\cap| \prod_{i \in I} k_i) \times (|\Omega_J^\cap| \prod_{j \in J} k_j)$ dimensional.

Later, we will prove the following bound on the expected operator norm of each $K_{I,J}$.

Lemma 7.3. For the direction block $K_{I,J}$ defined as above and $n \geq N_{D,\varepsilon} := \max \left\{ 16 \left(\frac{192D}{e} \right)^{24}, (2^D \varepsilon)^{-2} \right\}$, we have

$$\mathbb{E}_\sigma [\|K_{I,J}\|_{\text{op}}] \leq 2C_\rho^D \left(1 + \frac{4}{\rho} \right)^D \sqrt{2^D \varepsilon}.$$

With this bound in hand, we can use the triangle inequality as mentioned previously to finish off Lemma 7.1.

Proof of Lemma 7.1. If $M = \emptyset$ or $M = [n]^D$, then $K = 0$, so assume $\emptyset \neq M \subsetneq [n]^D$. Since both $I, J \subseteq [D]$, there are at most 4^D possible choices of (I, J) . For each $K_{I,J}$, we can pad it with zeros, which doesn't change the operator norm, to be the same dimension as K . Thus, we can compute

$$\begin{aligned} \mathbb{E}_\sigma [\|K\|_{\text{op}}] &= \mathbb{E}_\sigma \left[\left\| \sum_{I, J \subseteq [D]} K_{I,J} \right\|_{\text{op}} \right] \\ &\leq \sum_{I, J \subseteq [D]} \mathbb{E}_\sigma [\|K_{I,J}\|_{\text{op}}] && \text{By triangle ineq} \\ &\leq \sum_{I, J \subseteq [D]} 2C_\rho^D \left(1 + \frac{4}{\rho} \right)^D \sqrt{2^D \varepsilon} && \text{By Lemma 7.3} \\ &= 4^D 2C_\rho^D \left(1 + \frac{4}{\rho} \right)^D \sqrt{2^D \varepsilon} \\ &= C_{D,\rho,\varepsilon}, \end{aligned}$$

as claimed. □

7.1.2 Decomposition over sets of mutually fixed coordinates

Next, to achieve Lemma 7.3, we write out explicitly the action of $K_{I,J}$ and observe a clean way to write it as a sum of matrices. This will require some new notation.

Definition 7.4. With $i \in [D]$ and $G_i \in \mathbb{R}^{k_i \times n}$ as the generator matrix of the component code $\mathcal{C}_i$, we define the $n \times n$ real matrices

$$B_i = G_i^\top G_i \quad D_i = B_i - \text{Id}_{\mathbb{R}^n},$$

where we can see that $D_i(a, a) = 0$ and $D_i(a, b) = \frac{1}{k_i} \sum_{r=1}^{k_i} \sigma_{r,a}^{(i)} \sigma_{r,b}^{(i)}$ for $a \neq b$. Moreover, for a set $S \subseteq [D]$ we let

$$G_S = \bigotimes_{i \in S} G_i, \quad B_S = \bigotimes_{i \in S} B_i, \quad D_S = \bigotimes_{i \in S} D_i.$$

We write $g_S(a) = \bigotimes_{i \in S} g_i(a_i)$ for $a \in [n]^S$, with $g_\emptyset = 1$ and empty tensor products of operators interpreted as the identity on $\mathbb{R}$.

Definition 7.5. For a subset $S \subseteq X$ of some set X , we define the canonical inclusion and projection maps

$$\begin{aligned} \iota_S : \mathbb{R}^S &\rightarrow \mathbb{R}^X \\ \Pi_S : \mathbb{R}^X &\rightarrow \mathbb{R}^S. \end{aligned}$$

We now state the decomposition.

Lemma 7.6. For fixed $I, J \subseteq [D]$, define

$$C = I \cap J \quad L = I \setminus J \quad R = J \setminus I \quad F = (I \cup J)^c.$$

Then, after ordering the tensor coordinates in order as (C, L, R, F) , we can decompose $K_{I,J}$ as

$$K_{I,J} = (\text{Id}_{U_I} \otimes \Pi_{\Omega_I^\Gamma}) (\text{Id}_{U_C} \otimes G_L \otimes G_R^\top \otimes (B_F - \text{Id}_{\mathbb{R}^{[n]^F}})) (\text{Id}_{U_J} \otimes \iota_{\Omega_J^\Gamma}) \quad (22)$$

$$= \sum_{\emptyset \neq E \subseteq F} K_{I,J}^E \quad (23)$$

where

$$K_{I,J}^E = (\text{Id}_{U_I} \otimes \Pi_{\Omega_I^\Gamma}) (\text{Id}_{U_C} \otimes G_L \otimes G_R^\top \otimes D_E \otimes \text{Id}_{\mathbb{R}^{[n]^{F \setminus E}}}) (\text{Id}_{U_J} \otimes \iota_{\Omega_J^\Gamma}). \quad (24)$$

We defer the proof of Lemma 7.6 to Section 7.3 and provide a bound on $\mathbb{E} [\|K_{I,J}^E\|_{\text{op}}]$.

Lemma 7.7. For a non-empty subset $E \subseteq F$ and $n \geq N_{D,\varepsilon} := \max \left\{ 16 \left(\frac{192D}{e} \right)^{24}, (2^D \varepsilon)^{-2} \right\}$, we have that

$$\mathbb{E}_\sigma [\|K_{I,J}^E\|_{\text{op}}] \leq 2C_\rho^{|L|+|R|} \left(\frac{4}{\rho} \right)^{|E|} \sqrt{2^D \varepsilon}$$

Given this bound on each $\mathbb{E} [\|K_{I,J}^E\|_{\text{op}}]$, we can use the binomial theorem to sum over all subsets E to bound $\mathbb{E} [\|K_{I,J}\|_{\text{op}}]$.

Proof of Lemma 7.3. We start from Lemma 7.6 to compute

$$\begin{aligned} \mathbb{E}_\sigma [\|K_{I,J}\|_{\text{op}}] &= \mathbb{E}_\sigma \left[\left\| \sum_{\emptyset \neq E \subseteq F} K_{I,J}^E \right\|_{\text{op}} \right] && \text{By Equation (23)} \\ &\leq \sum_{\emptyset \neq E \subseteq F} \mathbb{E}_\sigma [\|K_{I,J}^E\|_{\text{op}}] && \text{By triangle ineq} \\ &\leq \sum_{\emptyset \neq E \subseteq F} 2C_\rho^{|L|+|R|} \left(\frac{4}{\rho} \right)^{|E|} \sqrt{2^D \varepsilon} && \text{By Lemma 7.7} \\ &\leq 2C_\rho^D \sqrt{2^D \varepsilon} \sum_{\emptyset \neq E \subseteq F} \left(\frac{4}{\rho} \right)^{|E|} \\ &\leq 2C_\rho^D \sqrt{2^D \varepsilon} \left(1 + \frac{4}{\rho} \right)^D && \text{By binomial thm} \end{aligned}$$

□

7.2 Bounding the expected operator norm of restricted directional blocks

The proof of Lemma 7.7 requires two main ingredients: First, we will upper bound $\mathbb{E}[\|K_{I,J}^E\|_{\text{op}}]$ by the product of the expected norm of the tensor of our generator matrices and the maximum operator norm of a sparsely applied operator. Second, we upper bound the expected operator norm of such sparsely applied operators.

We begin by using the definition of $K_{I,J}^E$ from Equation (24) to derive an upper bound on its operator norm.

Lemma 7.8. For a non-empty subset $E \subseteq F = (I \cup J)^c$, let $Z = F \setminus E$ so that we can decompose $[n]^F = [n]^R \times [n]^E \times [n]^Z$ and $[n]^{J^c} = [n]^L \times [n]^E \times [n]^Z$. Then define

$$\begin{aligned} X_{r,z} &= \{a \in [n]^E : (r, a, z) \in \Omega_I^\cap\} \\ Y_{\ell,z} &= \{b \in [n]^E : (\ell, b, z) \in \Omega_J^\cap\}. \end{aligned}$$

Then we can bound the operator norm of $K_{I,J}^E$ as

$$\|K_{I,J}^E\|_{\text{op}} \leq \|G_L\|_{\text{op}} \|G_R\|_{\text{op}} \max_{\substack{\ell \in [n]^L \\ r \in [n]^R \\ z \in [n]^Z}} \|\Pi_{X_{r,z}} D_{E \cup Y_{\ell,z}}\|_{\text{op}}. \quad (25)$$

Again, we defer the proof to Section 7.3 since it is mainly derived via careful algebraic manipulation of the definition of $K_{I,J}^E$. All that is left to obtain Lemma 7.7 then is to bound, in expectation, $\|G_L\|_{\text{op}}$, $\|G_R\|_{\text{op}}$, and $\max_{\ell,r,z} \|\Pi_{X_{r,z}} D_{E \cup Y_{\ell,z}}\|_{\text{op}}$. For G_L and G_R , we use standard results relating the operator norm of random Rademacher matrices to Gaussian matrices to get the following bounds, whose proofs we also defer.

Proposition 7.9. For a normalized (each column is scaled to have norm 1) random Rademacher matrix

$$G = \frac{1}{\sqrt{k}} (\sigma_{r,c})_{r \in [k], c \in [n]} \in \mathbb{R}^{k \times n},$$

where $k \geq \rho n$, its expected operator norm over the random signs is at most

$$\mathbb{E}[\|G\|_{\text{op}}] \leq C_\rho = 3 \left(1 + \rho^{-1/2}\right).$$

Using the independence of the signs of our generator matrices, we immediately get the following corollary.

Corollary 7.10. For a set $S \subseteq [D]$ and normalized random Rademacher matrices $(G_i)_{i \in [D]}$ with $k_i \geq \rho n$, we have that

$$\mathbb{E}[\|G_S\|_{\text{op}}] \leq C_\rho^{|S|}.$$

Proof. Apply Fact 3.24 to Proposition 7.9. □

To deal with $\max_{\ell,r,z} \|\Pi_{X_{r,z}} D_{E \cup Y_{\ell,z}}\|_{\text{op}}$, we state a general proposition about the expected maximum operator norm of such sparsely applied operators with the maximum taken over bounded indexing sets. The proof of this proposition, which comprises all of Section 8, crucially utilizes both the fact that M is ε -closed and properties of random Rademacher matrices.

Proposition 7.11 (Verbose version of Proposition 8.1). Take fixed integers $\nu \geq 0$ and $q \geq 1$ such that $\nu + q \leq D$. Let Λ be a fixed indexing set such that $|\Lambda| \leq n^\nu$, and take normalized random Rademacher matrices $(G_i)_{i \in [q]}$ with $k_i \geq \rho n$ and $D_i = G_i^\top G_i - \text{Id}_{\mathbb{R}^{[n]}}$. For all indices $\lambda \in \Lambda$, take $X_\lambda, Y_\lambda \subseteq [n]^q$ to be η -sparse for some constant $\eta \in (0, 1)$, fixed independently of the signs. Letting

$$T_\lambda = \Pi_{X_\lambda} \left(\bigotimes_{j=1}^q D_j \right) \iota_{Y_\lambda} : \mathbb{R}^{Y_\lambda} \rightarrow \mathbb{R}^{X_\lambda},$$

we have that for $n \geq N_{D,q,\eta} := \max \left\{ 16 \left(\frac{192D}{e} \right)^{24}, \eta^{-2/q} \right\}$,

$$\mathbb{E}[\max_{\lambda \in \Lambda} \|T_\lambda\|_{\text{op}}] \leq 2 \left(\frac{4}{\rho} \right)^q \sqrt{\eta}.$$

If $\Lambda = \emptyset$, then this expectation is just 0.

Having bounded each of the components of Equation (25), we can now prove Lemma 7.7.

Proof of Lemma 7.7. From Lemma 7.8 we have that

$$\|K_{I,J}^E\|_{\text{op}} \leq \|G_L\|_{\text{op}} \|G_R\|_{\text{op}} \max_{\substack{\ell \in [n]^L \\ r \in [n]^R \\ z \in [n]^Z}} \|\Pi_{X_{r,z}} D_{E \iota_{Y_{\ell,z}}}\|_{\text{op}}$$

where $\emptyset \neq E \subseteq F = (I \cup J)^c$ and $Z = F \setminus E$ so that we can decompose $[n]^{I^c} = [n]^R \times [n]^E \times [n]^Z$ and $[n]^{J^c} = [n]^L \times [n]^E \times [n]^Z$ to define

$$\begin{aligned} X_{r,z} &= \{a \in [n]^E : (r, a, z) \in \Omega_I^\cap\} \\ Y_{\ell,z} &= \{b \in [n]^E : (\ell, b, z) \in \Omega_J^\cap\}. \end{aligned}$$

To apply Proposition 7.11 with $q = |E|$ and $\nu = |L| + |R| + |Z|$, we note that number of triples (ℓ, r, z) is at most $n^{|L|+|R|+|Z|}$ and that $|L| + |R| + |Z| + |E| = D - |C| \leq D$.¹⁰ Furthermore, since $\Omega_I^\cap$ and $\Omega_J^\cap$ are $2^D \varepsilon$ -sparse by Corollary 3.14, $X_{r,z}$ and $Y_{\ell,z}$ are as well, so we can take $\eta = 2^D \varepsilon$. Our choice $n \geq N_{D,\varepsilon}$ also gives $n \geq N_{D,q,\eta}$, since $\eta^{-2/q} \leq \max\{1, \eta^{-2}\}$ for $q \geq 1$. The three factors in Equation (25) depend on disjoint collections of signs indexed by L , R , and E , respectively, since the sets $X_{r,z}$ and $Y_{\ell,z}$ depend only on the fixed set M . Thus, independence and Corollary 7.10 give

$$\begin{aligned} \mathbb{E} \left[\|K_{I,J}^E\|_{\text{op}} \right] &\leq \mathbb{E}_\sigma \left[\|G_L\|_{\text{op}} \right] \mathbb{E}_\sigma \left[\|G_R\|_{\text{op}} \right] \mathbb{E}_\sigma \left[\max_{\ell,r,z} \|\Pi_{X_{r,z}} D_{E \iota_{Y_{\ell,z}}}\|_{\text{op}} \right] \\ &\leq C_\rho^{|L|+|R|} \cdot 2 \left(\frac{4}{\rho} \right)^{|E|} \sqrt{2^D \varepsilon}, \end{aligned}$$

as claimed. $\square$

7.3 Deferred proofs

Finally, we present the deferred proofs of Lemmas 7.6 and 7.8 and Proposition 7.9. For convenience, we restate the relevant lemma and propositions.

Lemma 7.6. For fixed $I, J \subseteq [D]$, define

$$C = I \cap J \quad L = I \setminus J \quad R = J \setminus I \quad F = (I \cup J)^c.$$

Then, after ordering the tensor coordinates in order as (C, L, R, F) , we can decompose $K_{I,J}$ as

$$K_{I,J} = (\text{Id}_{U_I} \otimes \Pi_{\Omega_I^\cap}) (\text{Id}_{U_C} \otimes G_L \otimes G_R^\top \otimes (B_F - \text{Id}_{\mathbb{R}^{[n]^F}})) (\text{Id}_{U_J} \otimes \iota_{\Omega_J^\cap}) \quad (22)$$

$$= \sum_{\emptyset \neq E \subseteq F} K_{I,J}^E \quad (23)$$

where

$$K_{I,J}^E = (\text{Id}_{U_I} \otimes \Pi_{\Omega_I^\cap}) (\text{Id}_{U_C} \otimes G_L \otimes G_R^\top \otimes D_E \otimes \text{Id}_{\mathbb{R}^{[n]^{F \setminus E}}}) (\text{Id}_{U_J} \otimes \iota_{\Omega_J^\cap}). \quad (24)$$

¹⁰In this application of Proposition 7.11, we are considering the q generator matrices used the proposition to be exactly $(G_i)_{i \in E}$ and not necessarily the first q generator matrices of $(G_i)_{i \in [D]}$.

Proof of Lemma 7.6. We begin by formalizing what we mean by ordering the tensor coordinates in order as (C, L, R, F) . Since $I = C \sqcup L$, $I^c = R \sqcup F$, $J = C \sqcup R$ and $J^c = L \sqcup F$, we can decompose and order the ambient domain and codomain of $K_{I,J}$ as

$$U_J \otimes \mathbb{R}^{[n]^{J^c}} = U_{C \sqcup R} \otimes \mathbb{R}^{[n]^{L \sqcup F}} \cong U_C \otimes \mathbb{R}^{[n]^L} \otimes U_R \otimes \mathbb{R}^{[n]^F} \quad (26)$$

$$U_I \otimes \mathbb{R}^{[n]^{I^c}} = U_{C \sqcup L} \otimes \mathbb{R}^{[n]^{R \sqcup F}} \cong U_C \otimes U_L \otimes \mathbb{R}^{[n]^R} \otimes \mathbb{R}^{[n]^F}, \quad (27)$$

respectively. Recall that by convention, $U_\emptyset = \mathbb{R}$ and $\mathbb{R}^{[n]^\emptyset} = \mathbb{R}$. Moreover, ordering the tensor coordinates is equivalent to multiplying by a unitary matrix and doesn't change the operator norm.

To show Equation (22), we will let $\hat{K}_{I,J}$ be defined as the right hand side of Equation (22) and prove that $K_{I,J}$ and $\hat{K}_{I,J}$ agree on all of their plane-interaction blocks. Take

$$\begin{aligned} a &= (a_R, a_F) \in \Omega_I^\cap \subset [n]^R \times [n]^F \\ b &= (b_L, b_F) \in \Omega_J^\cap \subseteq [n]^L \times [n]^F \end{aligned}$$

and consider the planes $P = \mathcal{P}(I, a)$ and $Q = \mathcal{P}(J, b)$, so $(K_{I,J})_{a,b} = K_{P,Q}$ and $(\hat{K}_{I,J})_{a,b} = \hat{K}_{P,Q}$. For a pure tensor $u_C \otimes u_R \in U_C \otimes U_R \cong U_J$, we apply the first part $\text{Id}_{U_J} \otimes \iota_{\Omega_J^\cap}$ of $\hat{K}_{I,J}$ and use Equation (26) to perform the identification

$$(u_C \otimes u_R) \otimes e_b \mapsto u_C \otimes e_{b_L} \otimes u_R \otimes e_{b_F}. \quad (28)$$

Applying the second operator of $\hat{K}_{I,J}$ to Equation (28) gives

$$\begin{aligned} & (\text{Id}_{U_C} \otimes G_L \otimes G_R^\top \otimes (B_F - \text{Id}_{\mathbb{R}^{[n]^F}})) (u_C \otimes e_{b_L} \otimes u_R \otimes e_{b_F}) \\ &= u_C \otimes g_L(b_L) \otimes G_R^\top u_R \otimes (B_F - \text{Id}_{\mathbb{R}^{[n]^F}}) e_{b_F} \\ &= u_C \otimes g_L(b_L) \otimes \left(\sum_{r \in [n]^R} (G_R^\top u_R)(r) e_r \right) \otimes \left(\sum_{f \in [n]^F} (B_F(f, b_F) - \mathbf{1}\{f = b_F\}) e_f \right). \end{aligned}$$

Thus, after applying the third operator, we get

$$\begin{aligned} (\hat{K}_{I,J})_{a,b}(u_C \otimes u_R) &= \hat{K}_{P,Q}(u_C \otimes u_R) \\ &= ((G_R^\top u_R)(a_R)) (B_F(a_F, b_F) - \mathbf{1}\{a_F = b_F\}) (u_C \otimes g_L(b_L)). \end{aligned} \quad (29)$$

We now observe what $K_{P,Q}$ does to $u_C \otimes u_R$ by computing

$$\begin{aligned} \tau_Q(u_C \otimes u_R) &= u_C \otimes g_L(b_L) \otimes u_R \otimes g_F(b_F) \\ \tau_P^\top \tau_Q(u_C \otimes u_R) &= \langle g_R(a_R), u_R \rangle \langle g_F(a_F), g_F(b_F) \rangle (u_C \otimes g_L(b_L)) \\ &= ((G_R^\top u_R)(a_R)) (B_F(a_F, b_F)) (u_C \otimes g_L(b_L)). \end{aligned}$$

Recall though that if P and Q intersect, then $K_{P,Q} = 0$ instead of $\tau_P^\top \tau_Q$. Since $P \cap Q \neq \emptyset$ if and only if $a_F = b_F$, then

$$B_F(a_F, b_F) - \mathbf{1}\{a_F = b_F\} = \begin{cases} B_F(a_F, b_F) & P \cap Q = \emptyset \\ 0 & P \cap Q \neq \emptyset \end{cases} \quad (30)$$

Therefore, combining the expression for $\hat{K}_{P,Q}$ from Equation (29) with Equation (30) shows us that

$$\begin{aligned} \hat{K}_{P,Q} &= \begin{cases} \tau_P^\top \tau_Q & P \cap Q = \emptyset \\ 0 & P \cap Q \neq \emptyset \end{cases} \\ &= K_{P,Q}. \end{aligned}$$

Since this holds for any $a \in \Omega_I^\cap$ and $b \in \Omega_J^\cap$, we can conclude that $K_{I,J} = \widehat{K}_{I,J}$, as claimed in Equation (22).

Finally, to derive Equation (23), we expand $B_F - \text{Id}_{\mathbb{R}[n]^F}$ as

$$\begin{aligned}
B_F - \text{Id}_{\mathbb{R}[n]^F} &= \bigotimes_{f \in F} B_f - \text{Id}_{\mathbb{R}[n]^F} \\
&= \bigotimes_{f \in F} (D_f + \text{Id}_{\mathbb{R}[n]^{\{f\}}}) - \text{Id}_{\mathbb{R}[n]^F} \\
&= \left[\sum_{E \subseteq F} \left(\bigotimes_{e \in E} D_e \otimes \bigotimes_{f \in F \setminus E} \text{Id}_{\mathbb{R}[n]^{\{f\}}} \right) \right] - \text{Id}_{\mathbb{R}[n]^F} \\
&= \sum_{\emptyset \neq E \subseteq F} (D_E \otimes \text{Id}_{\mathbb{R}[n]^{F \setminus E}}).
\end{aligned} \tag{31}$$

Plugging Equation (31) into Equation (22) then yields

$$\begin{aligned}
K_{I,J} &= (\text{Id}_{U_I} \otimes \Pi_{\Omega_I^\cap}) (\text{Id}_{U_C} \otimes G_L \otimes G_R^\top \otimes (B_F - \text{Id}_{\mathbb{R}[n]^F})) (\text{Id}_{U_J} \otimes \iota_{\Omega_J^\cap}) \\
&= \sum_{\emptyset \neq E \subseteq F} (\text{Id}_{U_I} \otimes \Pi_{\Omega_I^\cap}) (\text{Id}_{U_C} \otimes G_L \otimes G_R^\top \otimes D_E \otimes \text{Id}_{\mathbb{R}[n]^{F \setminus E}}) (\text{Id}_{U_J} \otimes \iota_{\Omega_J^\cap}) \\
&= \sum_{\emptyset \neq E \subseteq F} K_{I,J}^E,
\end{aligned}$$

as claimed in Equation (23). $\square$

Lemma 7.8. For a non-empty subset $E \subseteq F = (I \cup J)^\complement$, let $Z = F \setminus E$ so that we can decompose $[n]^F = [n]^R \times [n]^E \times [n]^Z$ and $[n]^{J^\complement} = [n]^L \times [n]^E \times [n]^Z$. Then define

$$\begin{aligned}
X_{r,z} &= \{a \in [n]^E : (r, a, z) \in \Omega_I^\cap\} \\
Y_{\ell,z} &= \{b \in [n]^E : (\ell, b, z) \in \Omega_J^\cap\}.
\end{aligned}$$

Then we can bound the operator norm of $K_{I,J}^E$ as

$$\|K_{I,J}^E\|_{\text{op}} \leq \|G_L\|_{\text{op}} \|G_R\|_{\text{op}} \max_{\substack{\ell \in [n]^L \\ r \in [n]^R \\ z \in [n]^Z}} \|\Pi_{X_{r,z}} D_E \iota_{Y_{\ell,z}}\|_{\text{op}}. \tag{25}$$

Proof of Lemma 7.8. By definition of $X_{r,z}$ and $Y_{\ell,z}$, we can decompose $\Omega_I^\cap$ and $\Omega_J^\cap$ as

$$\begin{aligned}
\Omega_J^\cap &= \bigsqcup_{\substack{\ell \in [n]^L \\ z \in [n]^Z}} \{(\ell, b, z) : b \in Y_{\ell,z}\} \\
\Omega_I^\cap &= \bigsqcup_{\substack{r \in [n]^R \\ z \in [n]^Z}} \{(r, a, z) : a \in X_{r,z}\}.
\end{aligned}$$

Therefore, we can decompose the domain $U_J \otimes \mathbb{R}^{\Omega_J^\cap}$ and the codomain $U_I \otimes \mathbb{R}^{\Omega_I^\cap}$ of $K_{I,J}^E$ as

$$U_J \otimes \mathbb{R}^{\Omega_J^\cap} \cong U_C \otimes U_R \otimes \mathbb{R}^{\Omega_J^\cap} \cong \bigoplus_{\substack{\ell \in [n]^L \\ z \in [n]^Z}} U_C \otimes U_R \otimes \mathbb{R}^{Y_{\ell,z}} \tag{32}$$

$$U_I \otimes \mathbb{R}^{\Omega_I^\cap} \cong U_C \otimes U_L \otimes \mathbb{R}^{\Omega_I^\cap} \cong \bigoplus_{\substack{r \in [n]^R \\ z \in [n]^Z}} U_C \otimes U_L \otimes \mathbb{R}^{X_{r,z}}. \tag{33}$$

Using these decompositions, we'll write $K_{I,J}^E$ as the product of three operators that will make computing its norm in the form of Equation (25) immediate. The first operator applies $G_R^\top : U_R \rightarrow \mathbb{R}^{[n]^R}$ in every summand of the decomposition from Equation (32):

$$\Psi_R := \bigoplus_{\substack{\ell \in [n]^L \\ r \in [n]^R \\ z \in [n]^Z}} (\text{Id}_{U_C} \otimes G_R^\top \otimes \text{Id}_{\mathbb{R}^{Y_{\ell,z}}}) : U_J \otimes \mathbb{R}^{\Omega_J^\cap} \rightarrow \bigoplus_{\substack{\ell \in [n]^L \\ r \in [n]^R \\ z \in [n]^Z}} U_C \otimes \mathbb{R}^{Y_{\ell,z}}.$$

Here, to get the codomain as above, we use the identification

$$\begin{aligned} \mathbb{R}^{[n]^R} \otimes \mathbb{R}^{Y_{\ell,z}} &\cong \bigoplus_{r \in [n]^R} \mathbb{R}^{Y_{\ell,z}} \\ e_r \otimes e_b &\mapsto [e_b]_r, \end{aligned}$$

where $[e_b]_r$ denotes placing e_b in the r -th summand.

Next, we define our second operator in parts by taking $\ell \in [n]^L$, $r \in [n]^R$ and $z \in [n]^Z$ and letting

$$T_{\ell,r,z} := \Pi_{X_{r,z}} D_E \iota_{Y_{\ell,z}} : \mathbb{R}^{Y_{\ell,z}} \rightarrow \mathbb{R}^{X_{r,z}},$$

where $\iota_{Y_{\ell,z}} : \mathbb{R}^{Y_{\ell,z}} \rightarrow \mathbb{R}^{[n]^E}$ is inclusion by 0's (so padding in the coordinates $[n]^E \setminus Y_{\ell,z}$) and $\Pi_{X_{r,z}} : \mathbb{R}^{[n]^E} \rightarrow \mathbb{R}^{X_{r,z}}$ is projection. The direct sum of all $T_{\ell,r,z}$'s gives us our second operator

$$T_E := \bigoplus_{\substack{\ell \in [n]^L \\ r \in [n]^R \\ z \in [n]^Z}} (\text{Id}_{U_C} \otimes T_{\ell,r,z}) : \bigoplus_{\substack{\ell \in [n]^L \\ r \in [n]^R \\ z \in [n]^Z}} U_C \otimes \mathbb{R}^{Y_{\ell,z}} \rightarrow \bigoplus_{\substack{\ell \in [n]^L \\ r \in [n]^R \\ z \in [n]^Z}} U_C \otimes \mathbb{R}^{X_{r,z}} \cong \bigoplus_{\substack{r \in [n]^R \\ z \in [n]^Z}} U_C \otimes \mathbb{R}^{[n]^L} \otimes \mathbb{R}^{X_{r,z}}. \quad (34)$$

Finally, our third operator is similar to the first and applies $G_L : \mathbb{R}^{[n]^L} \rightarrow U_L$ to each summand in the codomain of T_E for all pairs (r, z) :

$$\Psi_L := \bigoplus_{\substack{r \in [n]^R \\ z \in [n]^Z}} (\text{Id}_{U_C} \otimes G_L \otimes \text{Id}_{\mathbb{R}^{X_{r,z}}}) : \bigoplus_{\substack{r \in [n]^R \\ z \in [n]^Z}} U_C \otimes \mathbb{R}^{[n]^L} \otimes \mathbb{R}^{X_{r,z}} \rightarrow U_I \otimes \mathbb{R}^{\Omega_I^\cap},$$

where we get the desired codomain by the decomposition in Equation (33). All together, our proposed factorization of $K_{I,J}^E$ looks like

$$U_J \otimes \mathbb{R}^{\Omega_J^\cap} \xrightarrow{\Psi_R} \bigoplus_{\substack{\ell \in [n]^L \\ r \in [n]^R \\ z \in [n]^Z}} U_C \otimes \mathbb{R}^{Y_{\ell,z}} \xrightarrow{T_E} \bigoplus_{\substack{\ell \in [n]^L \\ r \in [n]^R \\ z \in [n]^Z}} U_C \otimes \mathbb{R}^{X_{r,z}} \xrightarrow{\Psi_L} U_I \otimes \mathbb{R}^{\Omega_I^\cap}. \quad (35)$$

All that remains is to verify that this factorization is correct. To do so, we observe what Equation (35) does to pure tensors. Take $u_C \in U_C$, $u_R \in U_R$, and $e_{\ell,b,z} \in \mathbb{R}^{\Omega_J^\cap}$, the indicator vector at $(\ell, b, z) \in \Omega_J^\cap$. We then denote $[v]_{\ell,r,z}$ to be the vector v placed in the (ℓ, r, z) summand of the domain or codomain of T_E as defined in Equation (34) to form the pure tensor $u_C \otimes u_R \otimes e_{\ell,b,z} \in U_J \otimes \mathbb{R}^{\Omega_J^\cap}$ (as shown by the decomposition in Equation (32)) to compute

$$\begin{aligned} \Psi_R(u_C \otimes u_R \otimes e_{\ell,b,z}) &= \sum_{r \in [n]^R} (G_R^\top u_R)(r) [u_C \otimes e_b]_{\ell,r,z} \\ T_E \Psi_R(u_C \otimes u_R \otimes e_{\ell,b,z}) &= \sum_{\substack{r \in [n]^R \\ a \in X_{r,z}}} (G_R^\top u_R)(r) D_E(a, b) [u_C \otimes e_a]_{\ell,r,z} \\ \Psi_L T_E \Psi_R(u_C \otimes u_R \otimes e_{\ell,b,z}) &= \sum_{\substack{r \in [n]^R \\ a \in X_{r,z}}} (G_R^\top u_R)(r) D_E(a, b) (u_C \otimes G_L e_\ell \otimes e_{r,a,z}). \end{aligned} \quad (36)$$

From here, it is straightforward to see that Equation (36) is the same operation as $K_{I,J}^E$ as defined in Equation (24). Since pure tensors span the domain, we can conclude that $K_{I,J}^E = \Psi_L T_E \Psi_R$. A quick computation on this factorization yields Equation (25):

$$\begin{aligned}
\|K_{I,J}^E\|_{\text{op}} &= \|\Psi_L T_E \Psi_R\|_{\text{op}} \\
&\leq \|\Psi_L\|_{\text{op}} \|T_E\|_{\text{op}} \|\Psi_R\|_{\text{op}} && \text{By submultiplicativity} \\
&\leq \|G_L\|_{\text{op}} \left\| \bigoplus_{\ell,r,z} \text{Id}_{U_C} \otimes T_{\ell,r,z} \right\|_{\text{op}} \|G_R^\top\|_{\text{op}} \\
&\leq \|G_L\|_{\text{op}} \left(\max_{\ell,r,z} \|T_{\ell,r,z}\|_{\text{op}} \right) \|G_R\|_{\text{op}}, && \text{By Facts 3.23 and 3.24}
\end{aligned}$$

as desired. $\square$

We will need two standard facts about Gaussian matrices to prove Proposition 7.9. The first allows us to symmetrize the norm of a mean-zero matrix using standard Gaussian random variables.

Lemma 7.12 (Lemma 6.7.4 of [Ver18]). Let $X_1, \dots, X_N$ be independent mean-zero random vectors in a normed space. Let $g_1, \dots, g_N \sim \mathcal{N}(0, 1)$ be independent Gaussian random variables which are also independent of X_i . Then

$$\mathbb{E} \left\| \sum_{i=1}^N X_i \right\| \leq 3 \mathbb{E} \left\| \sum_{i=1}^N g_i X_i \right\|.$$

The second bounds the norm of a random Gaussian matrix.

Lemma 7.13 (Theorem 7.3.1 of [Ver18]). Let A be a $k \times n$ matrix with independent $\mathcal{N}(0, 1)$ entries. Then

$$\mathbb{E} \|A\|_{\text{op}} \leq \sqrt{k} + \sqrt{n}.$$

We can now prove Proposition 7.9.

Proposition 7.9. For a normalized (each column is scaled to have norm 1) random Rademacher matrix

$$G = \frac{1}{\sqrt{k}} (\sigma_{r,c})_{r \in [k], c \in [n]} \in \mathbb{R}^{k \times n},$$

where $k \geq \rho n$, its expected operator norm over the random signs is at most

$$\mathbb{E}[\|G\|_{\text{op}}] \leq C_\rho = 3 \left(1 + \rho^{-1/2}\right).$$

Proof of Proposition 7.9. For a row $r \in [k]$ and column $c \in [n]$, let $X_{r,c} \in \mathbb{R}^{k \times n}$ be the matrix of all 0's except with a 1 in the (r, c) position. Taking $\tilde{G}$ to be the un-normalized version of G , we have

$$\tilde{G} := \sqrt{k} G = \sum_{r \in [k], c \in [n]} \sigma_{r,c} X_{r,c}.$$

Applying Lemma 7.12 to the independent mean-zero matrices $\sigma_{r,c} X_{r,c}$, we have that

$$\mathbb{E}_\sigma \left[\|\tilde{G}\|_{\text{op}} \right] \leq 3 \mathbb{E}_{\sigma, \gamma} \left[\left\| \sum_{r \in [k], c \in [n]} \gamma_{r,c} \sigma_{r,c} X_{r,c} \right\|_{\text{op}} \right] \quad (37)$$

where the $\gamma_{r,c} \sim \mathcal{N}(0,1)$ are jointly independent and independent of the signs. Thus, the products $\gamma_{r,c}\sigma_{r,c}$ are jointly independent standard Gaussians, so we can use Lemma 7.13 to get

$$\mathbb{E}_{\sigma,\gamma} \left[\left\| \sum_{r \in [k], c \in [n]} \gamma_{r,c} \sigma_{r,c} X_{r,c} \right\|_{\text{op}} \right] \leq \sqrt{k} + \sqrt{n}. \quad (38)$$

Putting Equations (37) and (38) together yields

$$\begin{aligned} \mathbb{E}_{\sigma} [\|G\|_{\text{op}}] &= \frac{1}{\sqrt{k}} \mathbb{E}_{\sigma} \left[\|\tilde{G}\|_{\text{op}} \right] \\ &\leq \frac{1}{\sqrt{k}} 3 (\sqrt{k} + \sqrt{n}) \\ &\leq 3 (1 + \rho^{-1/2}), \end{aligned}$$

as desired. $\square$

8 Bounding the Operator Norm of Sparse Tensor Gram Matrices

In this section, we will prove the standalone proposition regarding the expected max operator norm over a sufficiently small family of sparse operators formed by the tensor of operators with random Rademacher entries. Since the main result of this section should be thought of as an independent result on the operator norm of such operators, we will redefine the objects we're working on.

Setup Throughout this section, we will use the following notation. Let $D > 0$ be a fixed integer, $q \geq 1$ and $\nu \geq 0$ be fixed integers such that $q + \nu \leq D$, $\rho \in (0,1)$ be constant, and for any $n \in \mathbb{N}$ and $i \in [D]$ let $k_i \in [n]$ be such that $k_i \geq \rho n$. With these parameters, we define q random matrices $G_i \in \mathbb{R}^{k_i \times n}$ as $(G_i)_{r,a} = k_i^{-1/2} \sigma_{r,a}^{(i)}$ for $r \in [k_i]$ and $a \in [n]$ where all $\sigma_{r,a}^{(i)} \sim \{-1,1\}$ are sampled independently. For $i \in [q]$, we let $D_i = G_i^\top G_i - \text{Id}_{\mathbb{R}^{[n]}}$. Lastly, for a set $X \subseteq [n]^q$, we let $\Pi_X : \mathbb{R}^{[n]^q} \rightarrow \mathbb{R}^X$ be the standard projection map and $\iota_X : \mathbb{R}^X \rightarrow \mathbb{R}^{[n]^q}$ be the standard inclusion map by 0's.

Proposition 8.1 (Proposition 7.11 restated). Take fixed integers $\nu \geq 0$ and $q \geq 1$ such that $\nu + q \leq D$. Let Λ be a fixed indexing set such that $|\Lambda| \leq n^\nu$. For all indices $\lambda \in \Lambda$, take $X_\lambda, Y_\lambda \subseteq [n]^q$ to be η -sparse for some constant $\eta \in (0,1)$, fixed independently of the signs. Letting

$$T_\lambda = \Pi_{X_\lambda} \left(\bigotimes_{j=1}^q D_j \right) \iota_{Y_\lambda} : \mathbb{R}^{Y_\lambda} \rightarrow \mathbb{R}^{X_\lambda},$$

we have that for $n \geq N_{D,q,\eta} := \max \left\{ 16 \left(\frac{192D}{e} \right)^{24}, \eta^{-2/q} \right\}$,

$$\mathbb{E} [\max_{\lambda \in \Lambda} \|T_\lambda\|_{\text{op}}] \leq 2 \left(\frac{4}{\rho} \right)^q \sqrt{\eta}.$$

If $\Lambda = \emptyset$, then this expectation is just 0.

8.1 Derivation of the sign monomials

Our proof of Proposition 8.1 will rely on the trace power method, so we begin by deriving an explicit expression for the trace of the power of these T_λ 's, which we will eventually decompose into monomials of the signs $\sigma_{r,a}^{(j)}$.

Lemma 8.2. For $X, Y \subseteq [n]^q$ and $D_1, \dots, D_q$ as above and

$$T := T_{X,Y} := \Pi_X \left(\bigotimes_{j=1}^q D_j \right) \iota_Y : \mathbb{R}^Y \rightarrow \mathbb{R}^X,$$

we have that for all integers $p \geq 1$:

$$\text{tr}((TT^\top)^p) = \sum_{\substack{x_1, \dots, x_p \in X \\ y_1, \dots, y_p \in Y}} \prod_{j=1}^q \prod_{t=1}^p D_j(x_{t,j}, y_{t,j}) D_j(y_{t,j}, x_{t+1,j})$$

where we define $x_{p+1} := x_1$.

Proof. For $x \in X$ and $y \in Y$ we have $T(x, y) = \prod_{j=1}^q D_j(x_j, y_j)$. Using the fact that each D_j is symmetric, we get that for $x, x' \in X$,

$$\begin{aligned} (TT^\top)(x, x') &= \sum_{y \in Y} T(x, y) T(x', y) \\ &= \sum_{y \in Y} \prod_{j=1}^q D_j(x_j, y_j) D_j(y_j, x'_j). \end{aligned} \tag{39}$$

Thus, taking the trace of $(TT^\top)^p$ yields

$$\begin{aligned} \text{tr}((TT^\top)^p) &= \sum_{x_1, \dots, x_p \in X} \prod_{t=1}^p (TT^\top)(x_t, x_{t+1}) \\ &= \sum_{x_1, \dots, x_p \in X} \prod_{t=1}^p \sum_{y \in Y} \prod_{j=1}^q D_j(x_{t,j}, y_j) D_j(y_j, x_{t+1,j}) \quad \text{By Equation (39)} \\ &= \sum_{\substack{x_1, \dots, x_p \in X \\ y_1, \dots, y_p \in Y}} \prod_{j=1}^q \prod_{t=1}^p D_j(x_{t,j}, y_{t,j}) D_j(y_{t,j}, x_{t+1,j}), \end{aligned}$$

as claimed. $\square$

Expanding out the result of Lemma 8.2 to contain the individual signs $\sigma_{r,a}^{(j)}$ will let us eventually count how many terms in this sum of products could possibly be 1.

Corollary 8.3. With the setup of Lemma 8.2,

$$\text{tr}((TT^\top)^p) = \sum_{\substack{x_1, \dots, x_p \in X \\ y_1, \dots, y_p \in Y}} \sum_{\substack{r_j \in [k_j]^{2p} \\ j \in [q]}} \left(\prod_{j=1}^q k_j^{-2p} \right) \left(\prod_{j=1}^q \Phi_j(x, y, r_j) \right),$$

where we define the monomials of signs

$$\Phi_j(x, y, r_j) = \prod_{t=1}^p \left(\sigma_{r_{j,2t-1}, x_{t,j}}^{(j)} \cdot \sigma_{r_{j,2t}, y_{t,j}}^{(j)} \cdot \sigma_{r_{j,2t}, y_{t,j}}^{(j)} \cdot \sigma_{r_{j,2t}, x_{t+1,j}}^{(j)} \cdot \mathbf{1}\{x_{t,j} \neq y_{t,j} \text{ and } y_{t,j} \neq x_{t+1,j}\} \right). \tag{40}$$

Proof. We use the definition of each D_j to directly compute

$$\begin{aligned}
& \prod_{t=1}^p D_j(x_{t,j}, y_{t,j}) D_j(y_{t,j}, x_{t+1,j}) \\
&= \prod_{t=1}^p \left[\left(\frac{1}{k_j} \sum_{r=1}^{k_j} \sigma_{r,x_{t,j}}^{(j)} \cdot \sigma_{r,y_{t,j}}^{(j)} \cdot \mathbf{1}\{x_{t,j} \neq y_{t,j}\} \right) \left(\frac{1}{k_j} \sum_{r=1}^{k_j} \sigma_{r,y_{t,j}}^{(j)} \cdot \sigma_{r,x_{t+1,j}}^{(j)} \cdot \mathbf{1}\{y_{t,j} \neq x_{t+1,j}\} \right) \right] \\
&= k_j^{-2p} \sum_{r_{j,1}, \dots, r_{j,2p} \in [k_j]} \prod_{t=1}^p \left(\sigma_{r_{j,2t-1}, x_{t,j}}^{(j)} \cdot \sigma_{r_{j,2t-1}, y_{t,j}}^{(j)} \cdot \sigma_{r_{j,2t}, y_{t,j}}^{(j)} \cdot \sigma_{r_{j,2t}, x_{t+1,j}}^{(j)} \cdot \mathbf{1}\{x_{t,j} \neq y_{t,j} \text{ and } y_{t,j} \neq x_{t+1,j}\} \right) \\
&= k_j^{-2p} \sum_{r_{j,1}, \dots, r_{j,2p} \in [k_j]} \Phi_j(x, y, r_j). \tag{41}
\end{aligned}$$

Plugging in Equation (41) to the result of Lemma 8.2 gives us the desired claim. $\square$

8.2 Graph view of the sign monomials

With the decomposition of $\text{tr}((TT^\top))$ into our sign monomials from Equation (40), the next natural step is to bound how many of these monomials could possibly have non-zero expectation. Our main tool to achieve this is to view the list of indices (x, y, r_j) as a walk along a graph, specifically a weak Wigner word from [Wig55], and use the tools from [FK81; AZ06] to bound how many monomials could have expectation 1. Overall, the idea is to remove all monomials with expectation 0, split the remaining monomials into equivalence classes, bound the size of each equivalence class using the sparsity of our indexing sets, and finally bound the number of equivalence classes using bounds on the number of equivalence classes of weak Wigner words from [AZ06].

To motivate this connection to weak Wigner words, we begin by repackaging the information contained in (x, y, r_j) into a single tuple which we will interpret as a walk along a bipartite graph.

Definition 8.4. For a fixed $j \in [q]$ and $x \in X^p$, $y \in Y^p$, and $r_j \in [k_j]^{2p}$, define the string $w_j = (w_{j,0}, \dots, w_{j,4p})$ as

$$\begin{aligned}
w_{j,4t-4} &= (\mathbf{L}, x_{t,j}) & w_{j,4t-3} &= (\mathbf{R}, r_{j,2t-1}) \\
w_{j,4t-2} &= (\mathbf{L}, y_{t,j}) & w_{j,4t-1} &= (\mathbf{R}, r_{j,2t}) \\
w_{j,4t} &= (\mathbf{L}, x_{t+1,j})
\end{aligned}$$

for all $t \in [p]$ where $\mathbf{L}$ and $\mathbf{R}$ are formal symbols indicating “left” and “right”. We will view this as a walk on the bipartite graph $\Gamma_{w_j} = (\mathbf{V} = \mathbf{V}_\mathbf{L} \sqcup \mathbf{V}_\mathbf{R}, \mathbf{E}_j)$ where $\mathbf{V}_\mathbf{L} = \{w_{j,4t-4}, w_{j,4t-2}, w_{j,4t} : t \in [p]\} \subseteq \{\mathbf{L}\} \times [n]$ and $\mathbf{V}_\mathbf{R} = \{w_{j,4t-3}, w_{j,4t-1} : t \in [p]\} \subseteq \{\mathbf{R}\} \times [k_j]$, and the distinct undirected edges are $\mathbf{E}_j = \{\{w_{j,t}, w_{j,t+1}\} : 0 \leq t < 4p\}$.

Remark 8.5. It’s useful to observe that, via this construction, there is a one-to-one correspondence between the q -tuple $(w_1, \dots, w_q)$ and the list (x, y, r) for $r \in \prod_{j=1}^q [k_j]^{2p}$. That is, these are just different ways of indexing the same $q \times 4p$ elements, which we will more clearly see in the proof of Lemma 8.29.

Next, we make a simple observation about when Φ_j is non-zero.

Lemma 8.6. With the setup from Definition 8.4, $\Phi_j(x, y, r_j)$ is non-zero if and only if the walk w_j is a non-left-backtracking closed walk.¹¹

Proof. From the definition of $\Phi_j(x, y, r_j)$ in Equation (40), we see that $\Phi_j(x, y, r_j)$ is non-zero if and only if $x_{t,j} \neq y_{t,j}$ and $y_{t,j} \neq x_{t+1,j}$ for all $t \in [p]$, meaning that $w_{j,4t-4} \neq w_{j,4t-2} \neq w_{j,4t}$. This gives us that

¹¹By non-left-backtracking, we mean that at each visit to a right vertex, the preceding and following left vertices differ. Backtracking at left vertices is allowed.

the walk is non-left-backtracking. To see that the walk is closed, we recall from Lemma 8.2 that we defined $x_{p+1} = x_1$, giving us $w_{j,0} = (\mathbf{L}, x_{1,j}) = (\mathbf{L}, x_{p+1,j}) = w_{j,4p}$ $\square$

In this setting, we can express $\Phi_j(x, y, r_j)$ directly by the edges of Γ_{w_j} .

Lemma 8.7. For a non-left-backtracking closed walk w_j from (x, y, r_j) , we can compute $\Phi_j(x, y, r_j)$ from w_j as

$$\Phi_j(x, y, r_j) = \prod_{\substack{r \in [k_j] \\ a \in [n]}} \left(\sigma_{r,a}^{(j)} \right)^{m_j(r,a)}, \quad (42)$$

where $m_j(r, a)$ is the number of times w_j traverses the edge $\{(\mathbf{L}, a), (\mathbf{R}, r)\}$ in either direction.

Proof. Since w_j is a non-left-backtracking closed walk, the indicators $\mathbf{1}\{x_{t,j} \neq y_{t,j} \text{ and } y_{t,j} \neq x_{t+1,j}\}$ from Equation (40) are always 1, so $\Phi_j(x, y, r_j)$ is always just a product of signs. The expression in Equation (42) then exactly recovers the signs used in $\Phi_j(x, y, r_j)$ because each occurrence is counted by a traversal of the corresponding edge. $\square$

An immediate consequence of this formulation of Φ_j is an easy expression for its expectation.

Corollary 8.8. For a non-left-backtracking closed walk w_j , we have

$$\mathbb{E}_{\sigma}[\Phi_j(x, y, r_j)] = \begin{cases} 1 & m_j(r, a) \text{ is even for all } r \in [k_j], a \in [n] \\ 0 & \text{otherwise.} \end{cases}$$

Proof. If at least one of the $m_j(r, a)$'s is odd, then since all of the signs $\sigma_{r,a}^{(j)}$ are independent Rademacher variables, the whole product is randomized as a Rademacher variable and has expectation 0. If all the exponents are even, then $\Phi_j(x, y, r_j) = 1$ deterministically. $\square$

Taking the expectation over all of $(TT^{\top})^p$ then gives us a nice expression.

Corollary 8.9. Randomizing over the signs, the expectation of $\text{tr}((TT^{\top})^p)$ is

$$\mathbb{E}_{\sigma}[\text{tr}((TT^{\top})^p)] = \left(\prod_{j=1}^q k_j^{-2p} \right) \cdot |\mathcal{W}|$$

where

$$\mathcal{W}(X, Y) = \left\{ (x, y, r) \in X^p \times Y^p \times \prod_{j=1}^q [k_j]^{2p} : \forall j \in [q], w_j \text{ is even and non-left-backtracking} \right\}$$

Therefore, to bound the expectation of $\text{tr}((TT^{\top})^p)$, we want to bound the number of monomials $\Phi_j(x, y, r_j)$ where all exponents are even. Equivalently, we want to bound the number of possible non-left-backtracking closed walks w_j that traverse each edge of Γ_{w_j} an even number of times—the cardinality of $\mathcal{W}$. This brings us to weak Wigner words.

8.3 Weak Wigner words

One can naturally view a walk over a graph as just a list enumerating the labels of the vertices visited along that walk. That intuition forms the basis for the definition of words and sentences in the context of Wigner words.

Definition 8.10 (Words and sentences). [[Wig55; AZ06]] A *word* w is a non-empty finite sequence of letters from some alphabet, where we view letters as length-1 words. We call w *closed* if the first and last letters are the same. A *sentence* a is a non-empty sequence of words, so we may view words as one-word sentences.

There are many properties of words and sentences that we will work with.

Definition 8.11. For a word w and sentence a , we define:

- The *length* $\text{len}(w)$ is the number of letters in w ,
- The *support* $\text{supp}(a)$ is the set of letters appearing in a , and
- The *weight* $\text{wt}(a)$ is the cardinality of $\text{supp}(a)$.

Given a word, we can translate back and forth between its associated graph.

Definition 8.12 (Graph of a word). Given a word $w = \alpha_1 \cdots \alpha_p$, we define its *graph* $\Gamma_w = (\mathbf{V}, \mathbf{E})$ to have vertices $\mathbf{V} = \text{supp}(w)$ and distinct undirected edges $\mathbf{E} = \{\{\alpha_i, \alpha_{i+1}\} : i \in [p-1]\}$.

We can now finally define weak Wigner words.

Definition 8.13 (weak Wigner word). We say that a word w is a *weak Wigner word* if w is closed and traverses each edge of Γ_w at least twice.

8.3.1 Equivalence and canonical representatives

Given that words define walks on graphs, it is natural to define an equivalence relation between them encoding the shape of the walk.

Definition 8.14 (Equivalence). Two words w and w' are *equivalent*, denoted as $w \sim w'$, if $\text{len}(w) = \text{len}(w')$ and there exists a bijection $\psi : \text{supp}(w) \rightarrow \text{supp}(w')$ such that $\psi(w_i) = w'_i$ for all $i \in [\text{len}(w)]$.

Remark 8.15. It is worthwhile to note that $w \sim w'$ isn't the same as the graphs Γ_w and $\Gamma_{w'}$ being isomorphic. For example, $w = \{12321\}$ and $w' = \{12131\}$ have isomorphic graphs but aren't equivalent. On the other hand, if $w \sim w'$ then Γ_w and $\Gamma_{w'}$ are clearly isomorphic where the isomorphism is given by the relabeling of their letters, ψ .

With an equivalence relation, we can talk about the equivalence class $\mathcal{W}$ of a word w . Intuitively, an equivalence class encodes the shape of a walk, but not the actual values of the vertices. Since we will need to reference the shape of the walk of an equivalence class, we define the canonical representative of an equivalence class—this is not necessarily a word generated as in Definition 8.4, but it clearly encodes the shape of the path of the elements of $\mathcal{W}$.

Definition 8.16 (Canonical representative). Let $\mathcal{W}$ be an equivalence class and $w = (w_1, \dots, w_m) \in \mathcal{W}$. Define $\psi_w : \text{supp}(w) \rightarrow [\text{wt}(w)]$ by ordering the distinct letters by their first appearance in w . The canonical representative of $\mathcal{W}$ is the word $\bar{w} = (\psi_w(w_1), \dots, \psi_w(w_m))$.

Example 8.17. For the word $w = \{421314214\}$, the canonical representative of its equivalence class is $\bar{w} = \{123431231\}$.

It is easy to check that this notion of canonical representative is well-defined.

Fact 8.18. If $w \sim w'$, then their canonical representatives satisfy $\bar{w} = \bar{w}'$.

Proof. Equivalence preserves the equality of letters at any two positions. Thus first appearances occur at the same positions in w and w' , and every repeated letter receives the same canonical label in both words. $\square$

8.4 Weak Wigner words in the bipartite setting

Since the words we will work on are generated as in Definition 8.4 and represent walks on a bipartite graph, we will introduce some additional definitions and consequences in this context. We call a word w *bipartite* if Γ_w is a bipartite graph.

Definition 8.19 (Properties of bipartite walks). Given a bipartite walk w on $\Gamma_w = (\mathbf{V} = \mathbf{V}_L \sqcup \mathbf{V}_R, \mathbf{E})$, we define

- The left weight $\text{wt}_L(w) = |\mathbf{V}_L|$, and
- The right weight $\text{wt}_R(w) = |\mathbf{V}_R|$.

Moreover, we can measure the loss of vertices from the maximum possible weight of a weak Wigner word from Definition 8.4. Such words have length $4p+1$ and traverse each edge at least twice, so their connected graphs have at most $2p+1$ vertices.

Definition 8.20 (Deficiency). For a weak Wigner word w on $\Gamma_w = (\mathbf{V} = \mathbf{V}_L \sqcup \mathbf{V}_R, \mathbf{E})$ constructed as in Definition 8.4, we define its *deficiency* to be

$$\Delta(w) = 2p + 1 - \text{wt}(w).$$

Lastly, we define a specific notion of equivalence on the words from Definition 8.4 and show that it agrees with the standard notion of equivalence.

Definition 8.21 (Bipartite equivalence). Given two walks w on $\Gamma_w = (\mathbf{V} = \mathbf{V}_L \sqcup \mathbf{V}_R, \mathbf{E})$ and w' on $\Gamma_{w'} = (\mathbf{V}' = \mathbf{V}'_L \sqcup \mathbf{V}'_R, \mathbf{E}')$ of equal length, we say that w is *bipartite-equivalent* to w' if there exist bijections $\psi_L : \mathbf{V}_L \rightarrow \mathbf{V}'_L$ and $\psi_R : \mathbf{V}_R \rightarrow \mathbf{V}'_R$ such that $\psi_L(w_i) = w'_i$ for all even i , and $\psi_R(w_i) = w'_i$ for all odd i . This is denoted as $w \sim_{\text{bi}} w'$.

Fact 8.22. For two walks w on $\Gamma_w = (\mathbf{V} = \mathbf{V}_L \sqcup \mathbf{V}_R, \mathbf{E})$ and w' on $\Gamma_{w'} = (\mathbf{V}' = \mathbf{V}'_L \sqcup \mathbf{V}'_R, \mathbf{E}')$ of equal length produced as in Definition 8.4,

$$w \sim_{\text{bi}} w' \iff w \sim w'.$$

Proof. The forward direction is immediate since we just define $\psi : \mathbf{V} \rightarrow \mathbf{V}'$ to act on $\mathbf{V}_L$ via ψ_L and on $\mathbf{V}_R$ via ψ_R . The reverse direction comes from noticing that in Definition 8.4, it must be that each even index of w is an element of $\mathbf{V}_L$ and every odd index is an element of $\mathbf{V}_R$, and similarly for w' . Thus, for an even index i , we have that $\psi(w_i) = \psi((\mathbf{L}, x)) = (\mathbf{L}, x') = w'_i$. Consequently, on even indices ψ restricts to a bijection from $\mathbf{V}_L$ to $\mathbf{V}'_L$, which allows us to define ψ_L . Similarly, on odd indices ψ restricts to a bijection from $\mathbf{V}_R$ to $\mathbf{V}'_R$, allowing us to define ψ_R . This gives us the necessary bipartite-equivalence of w and w' . $\square$

8.4.1 Equivalence classes from a specified deficiency

Since all words in an equivalence class have the same weight and length, which are the only two parameters used in the deficiency of a word, they have the same deficiency as well. Thus, we can parametrize a set of equivalence classes by only their deficiency.

Definition 8.23. For integers $p \geq 1$ and $\Delta \geq 0$, we let $\mathfrak{W}_p(\Delta)$ be the set of equivalence classes of even non-left-backtracking closed walks with length $4p+1$, deficiency Δ , and left and right positions as in Definition 8.4. For a tuple of equivalence classes $(\mathcal{W}_1, \dots, \mathcal{W}_q)$, define its admissible part by

$$\mathscr{W}_{X,Y}(\mathcal{W}_1, \dots, \mathcal{W}_q) = \{(x, y, r) \in \mathscr{W}(X, Y) : w_j(x, y, r_j) \in \mathcal{W}_j \text{ for all } j \in [q]\}.$$

With this notation, we can cleanly decompose $\mathscr{W}$.

Lemma 8.24.

$$\mathcal{W}(X, Y) = \bigsqcup_{(\Delta_1, \dots, \Delta_q) \in \{0, \dots, 2p\}^q} \bigsqcup_{(\mathcal{W}_1, \dots, \mathcal{W}_q) \in \mathfrak{W}_p(\Delta_1) \times \dots \times \mathfrak{W}_p(\Delta_q)} \mathcal{W}_{X, Y}(\mathcal{W}_1, \dots, \mathcal{W}_q).$$

Here, each $\mathcal{W}_j$ is one of the equivalence classes of $\mathfrak{W}_p(\Delta_j)$.

Proof. Each $(x, y, r) \in \mathcal{W}(X, Y)$ determines a unique equivalence class $\mathcal{W}_j$ for each $w_j(x, y, r_j)$ and a unique deficiency $\Delta_j \in \{0, \dots, 2p\}$. Thus it belongs to exactly one of the admissible parts, while every member of an admissible part belongs to $\mathcal{W}(X, Y)$ by definition. $\square$

8.5 Bounding the number of equivalence classes

The first step in bounding the number of non-left-backtracking closed walks with even multiplicity on each edge is to count the number of equivalence classes of such closed walks.

Proposition 8.25 (Proposition 7.9 of [AZ06]). For all positive integers N and W satisfying $N \geq 2W - 2$, there are at most

$$2^N N^{3(N-2W+2)}$$

equivalence classes of weak Wigner words of length $N + 1$ and weight W .

Corollary 8.26. For all positive integers N and W satisfying $N \geq 2W - 2$, there are at most

$$2^N N^{3(N-2W+2)}$$

bipartite-equivalence classes of weak Wigner words produced as in Definition 8.4 of length $N + 1$ and weight W .

Proof. The parity argument in Fact 8.22 applies to arbitrary left and right labels, so ordinary and bipartite equivalence agree for these words. We then apply Proposition 8.25. $\square$

Fact 8.27. For a non-left-backtracking closed walk w on $\Gamma_w = (\mathcal{V} = \mathcal{V}_L \sqcup \mathcal{V}_R, \mathcal{E})$ from Definition 8.4 such that every edge is traversed an even number of times, we have

$$\text{wt}(w) \leq 2p + 1 \tag{43}$$

$$\Delta(w) \geq 0 \tag{44}$$

$$\Delta(w) \leq 2p \tag{45}$$

$$\text{wt}_R(w) \leq p \tag{46}$$

$$\text{wt}_L(w) \geq p + 1 - \Delta(w). \tag{47}$$

Proof. Let e be the number of distinct undirected edges traversed by w . By construction, the walk makes $4p$ edge traversals. Since each edge is traversed at least twice, $e \leq 2p$. The graph traced by the walk is connected, so $\text{wt}(w) \leq e + 1 \leq 2p + 1$, giving Equation (43). Consequently, $\Delta(w) = 2p + 1 - \text{wt}(w) \geq 0$, yielding Equation (44). Since w is nonempty, $\text{wt}(w) \geq 1$ also gives $\Delta(w) \leq 2p$. Each right vertex is entered from and exited to distinct left vertices by the non-left-backtracking condition, so it has at least two distinct neighbors. Counting each distinct edge at its right endpoint gives $2 \text{wt}_R(w) \leq e \leq 2p$, so $\text{wt}_R(w) \leq p$. Finally, $\text{wt}_L(w) = 2p + 1 - \Delta(w) - \text{wt}_R(w) \geq p + 1 - \Delta(w)$, giving the final two inequalities. $\square$

This gives us a convenient rewording of Proposition 7.9 of [AZ06].

Corollary 8.28. For an integer $p \geq 1$ and deficiency $\Delta \geq 0$, the number of equivalence classes of even non-left-backtracking closed walks from Definition 8.4 is at most

$$|\mathfrak{W}_p(\Delta)| \leq 16^p (4p)^{6\Delta}.$$

Proof. If $\Delta(w) > 2p$, then $\text{wt}(w) = 2p+1-\Delta(w) \leq 0$, which would give us an empty word, so we may assume that $\Delta(w) \in \{0, 1, \dots, 2p\}$ since $\Delta(w) \geq 0$ by Equation (44). Then we set the variables in Corollary 8.26 as $N = \text{len}(w) - 1 = 4p$ and $W = 2p + 1 - \Delta(w)$. The requirement of Corollary 8.26 is satisfied since $N - (2W - 2) = 2\Delta(w) \geq 0$, so there are at most

$$\begin{aligned} 2^N N^{3(N-2W+2)} &= 2^{4p} (4p)^{3(4p-2(2p+1-\Delta(w))+2)} \\ &= 16^p (4p)^{6\Delta(w)} \end{aligned}$$

equivalence classes of non-left-backtracking closed walks from Definition 8.4 with even multiplicity on each edge. $\square$

8.6 Bounding the size of an equivalence class of a word using sparsity

Now that we've bounded the number of equivalence classes of words arising from Definition 8.4 with even multiplicity on each edge, we move on to bounding the size of their admissible part, the term $\mathcal{W}_{X,Y}(\mathcal{W}_1, \dots, \mathcal{W}_q)$ showing up in Lemma 8.24. This is the key step where the sparsity of the indexing sets is used.

Lemma 8.29. For a constant $\eta \in (0, 1)$ and η -sparse subsets $X, Y \subseteq [n]^q$, take integers $\Delta_1, \dots, \Delta_q \in \{0, \dots, 2p\}$. Then, for any $(\mathcal{W}_1, \dots, \mathcal{W}_q) \in \mathfrak{W}_p(\Delta_1) \times \dots \times \mathfrak{W}_p(\Delta_q)$, we have that

$$|\mathcal{W}_{X,Y}(\mathcal{W}_1, \dots, \mathcal{W}_q)| \leq n^{W_L} \eta^{W_L/q} \prod_{j=1}^q k_j^{\text{wt}_R(\mathcal{W}_j)} \quad (48)$$

where $\text{wt}_L(\mathcal{W}_j) = \text{wt}_L(w_j)$ for any $w_j \in \mathcal{W}_j$ since all elements in the same equivalence class have the same weight, $\text{wt}_R(\mathcal{W}_j)$ is similar, and

$$W_L = \sum_{j=1}^q \text{wt}_L(\mathcal{W}_j) \quad \text{and} \quad W_R = \sum_{j=1}^q \text{wt}_R(\mathcal{W}_j).$$

Proof. If $\mathcal{W}_{X,Y}(\mathcal{W}_1, \dots, \mathcal{W}_q)$ is empty, the bound is immediate. Otherwise, consider a tuple $(x, y, r) \in \mathcal{W}_{X,Y}(\mathcal{W}_1, \dots, \mathcal{W}_q)$ and its words $w_1, \dots, w_q$. Since w_j has $\text{wt}_R(w_j) \leq k_j$ distinct right vertices and there are k_j possible right vertices to choose from, there are $k_j(k_j - 1) \cdots (k_j - \text{wt}_R(w_j) + 1) \leq k_j^{\text{wt}_R(w_j)}$ possible labelings of right vertices for w_j . Over all $j \in [q]$, this means there are at most

$$\prod_{j=1}^q k_j^{\text{wt}_R(\mathcal{W}_j)} \quad (49)$$

many assignments of all the right vertices of $w_1, \dots, w_q$.

For the left vertices, we must use the sparsity of X and Y . This is where it becomes apparent why we must count admissible tuples rather than choose each w_j separately: fixing one w_j fixes only the j -th coordinates of all x_t, y_t , together with r_j . The remaining coordinates must still be chosen so that each complete x_t belongs to X and each complete y_t belongs to Y . Put more formally, we will write out the

entries of $x \in X^p$ and $y \in Y^p$ as a matrix. Let $Z \in [n]^{2p \times q}$ be defined as

$$\begin{aligned} Z &= \begin{bmatrix} - & x_1 & - \\ - & y_1 & - \\ - & x_2 & - \\ - & y_2 & - \\ & \vdots & \\ - & x_p & - \\ - & y_p & - \end{bmatrix} = \begin{bmatrix} - & Z_1 & - \\ - & Z_2 & - \\ - & Z_3 & - \\ - & Z_4 & - \\ & \vdots & \\ - & Z_{2p-1} & - \\ - & Z_{2p} & - \end{bmatrix} \\ &= \begin{bmatrix} z_{1,1} & z_{1,2} & \cdots & z_{1,q} \\ z_{2,1} & z_{2,2} & \cdots & z_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ z_{2p,1} & z_{2p,2} & \cdots & z_{2p,q} \end{bmatrix} \\ &= \begin{bmatrix} | & | & & | \\ w_1^{\mathsf{L}} & w_2^{\mathsf{L}} & \cdots & w_q^{\mathsf{L}} \\ | & | & & | \end{bmatrix}. \end{aligned}$$

where for $t \in [p]$ and $j \in [q]$ we are letting $z_{2t-1,j} = x_{t,j}$ and $z_{2t,j} = y_{t,j}$ and also removing the “L” label to just get elements of $[n]$. Here w_j^{L} is the column vector of these $2p$ left entries of w_j , with the repeated last entry omitted. Now, it is clear that each row of Z is an element of X (if it is at an odd index) or Y (if it is at an even index), and that each column of Z gives the remaining information to define $w_j(x, y, r_j)$ since we have already fixed r_j . Consequently, once we have fixed $r \in \prod_{j=1}^q [k_j]^{2p}$, counting the members of $\mathscr{W}_{X,Y}(\mathcal{W}_1, \dots, \mathcal{W}_q)$ with this r reduces to counting the number of possible matrices Z with each odd row in X and each even row in Y , via Remark 8.5. However, each column must also respect the shape of the path given by the equivalence classes $\mathcal{W}_1, \dots, \mathcal{W}_q$. To enforce this, we take the canonical representative $\bar{w}_1$ of $\mathcal{W}_j$ for each $j \in [q]$, write $\bar{w}_j^{\mathsf{L}}$ for the corresponding column of $2p$ left entries, and define

$$\begin{aligned} \bar{Z} &= \begin{bmatrix} | & | & & | \\ \bar{w}_1^{\mathsf{L}} & \bar{w}_2^{\mathsf{L}} & \cdots & \bar{w}_q^{\mathsf{L}} \\ | & | & & | \end{bmatrix} \\ &= \begin{bmatrix} \bar{z}_{1,1} & \bar{z}_{1,2} & \cdots & \bar{z}_{1,q} \\ \bar{z}_{2,1} & \bar{z}_{2,2} & \cdots & \bar{z}_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{z}_{2p,1} & \bar{z}_{2p,2} & \cdots & \bar{z}_{2p,q} \end{bmatrix} \end{aligned}$$

These canonical labels specify equality patterns—the rows of $\bar{Z}$ need not belong to X or Y .

To go from columns, which define each w_j , to rows, which are chosen by x_t and y_t , we define for a row index $s \in [2p]$ the set

$$F_s = \{j \in [q] : \bar{z}_{s,j} \notin \{\bar{z}_{1,j}, \dots, \bar{z}_{s-1,j}\}\},$$

which gives the column indices in row s of elements whose symbol first appears at position s in their column. This allows us to rewrite W_{L} as

$$\begin{aligned} W_{\mathsf{L}} &= \sum_{j=1}^q \text{wt}_{\mathsf{L}}(\mathcal{W}_j) \\ &= \sum_{j=1}^q \sum_{s=1}^{2p} \mathbf{1}\{\bar{z}_{s,j} \notin \{\bar{z}_{1,j}, \dots, \bar{z}_{s-1,j}\}\} \\ &= \sum_{s=1}^{2p} |F_s|. \end{aligned} \tag{50}$$

Thus, once the rows $Z_1, \dots, Z_{s-1}$ have been fixed, the coordinates of Z_s in $[q] \setminus F_s$ are also fixed because $w_j \sim \bar{w}_j$ for all $j \in [q]$, so we are only free to choose the value of the coordinates of Z_s in F_s . If $|F_s| = 0$, then Z_s is already fixed and there is at most one choice. If $|F_s| > 0$, then since both X and Y are η -sparse, from Fact 3.7 we have that there are at most $\eta n^{|F_s|}$ possible values for Z_s . Note that in this counting argument, we neither required the non-left-backtracking requirement that $x_{t,j} \neq y_{t,j}$ and $y_{t,j} \neq x_{t+1,j}$ nor that unfixed indices in F_s actually be chosen to not match the previous elements in their column. Loosening our requirements just gives us a weaker upper bound, which is acceptable for our argument.

Finally, multiplying the number of choices for each row $s \in [2p]$ gives us at most

$$n^{\sum_{s=1}^{2p} |F_s|} \eta^{\sum_{s=1}^{2p} \mathbf{1}\{|F_s| > 0\}}$$

many assignments. To simplify this expression, we notice that trivially $|F_s| \leq q$ so by Equation (50) we have $W_L \leq q \sum_{s=1}^{2p} \mathbf{1}\{|F_s| > 0\}$, meaning that $\sum_{s=1}^{2p} \mathbf{1}\{|F_s| > 0\} \geq W_L/q$. Along with another application of Equation (50) this gives us the upper bound

$$n^{\sum_{s=1}^{2p} |F_s|} \eta^{\sum_{s=1}^{2p} \mathbf{1}\{|F_s| > 0\}} \leq n^{W_L} \eta^{W_L/q} \quad (51)$$

Taking the product of the upper bound on the number of assignments of right vertices in Equation (49) and the upper bound on the number of assignments of left vertices from Equation (51) gives us Equation (48). $\square$

We take one step further and multiply by the coefficient $\prod_{j=1}^q k_j^{-2p}$.

Lemma 8.30. With the assumptions of Proposition 8.1 and η -sparse $X, Y \subseteq [n]^q$ and integers $\Delta_1, \dots, \Delta_q \in \{0, \dots, 2p\}$, for any equivalence classes $(\mathcal{W}_1, \dots, \mathcal{W}_q) \in \mathfrak{W}_p(\Delta_1) \times \dots \times \mathfrak{W}_p(\Delta_q)$, we have that

$$\left(\prod_{j=1}^q k_j^{-2p} \right) |\mathscr{W}_{X,Y}(\mathcal{W}_1, \dots, \mathcal{W}_q)| \leq \rho^{-2pq} n^q \eta^{p+1} (n \eta^{1/q})^{-\Delta} \quad (52)$$

where $\Delta = \sum_{j=1}^q \Delta_j$.

Proof. By Lemma 8.29, we have

$$\begin{aligned} \left(\prod_{j=1}^q k_j^{-2p} \right) |\mathscr{W}_{X,Y}(\mathcal{W}_1, \dots, \mathcal{W}_q)| &\leq \left(\prod_{j=1}^q k_j^{-2p} \right) n^{W_L} \eta^{W_L/q} \prod_{j=1}^q k_j^{\text{wt}_R(\mathcal{W}_j)} \\ &= n^{W_L} \eta^{W_L/q} \prod_{j=1}^q k_j^{\text{wt}_R(\mathcal{W}_j) - 2p} \end{aligned} \quad (53)$$

Equation (46) gives us that $\text{wt}_R(\mathcal{W}_j) \leq p$, so using the fact that $k_j \geq \rho n$ and $\rho \in (0, 1)$ we can compute

$$k_j^{\text{wt}_R(\mathcal{W}_j) - 2p} \leq (\rho n)^{\text{wt}_R(\mathcal{W}_j) - 2p} \leq \rho^{-2p} n^{\text{wt}_R(\mathcal{W}_j) - 2p}. \quad (54)$$

Therefore, over all $j \in [q]$ we get

$$\begin{aligned} \prod_{j=1}^q k_j^{\text{wt}_R(\mathcal{W}_j) - 2p} &\leq \prod_{j=1}^q \rho^{-2p} n^{\text{wt}_R(\mathcal{W}_j) - 2p} && \text{By Equation (54)} \\ &= \rho^{-2pq} n^{W_R - 2pq} \end{aligned} \quad (55)$$

Moreover, by Equation (47) we have that $W_L \geq q(p+1) - \Delta$ so

$$\eta^{W_L/q} \leq \eta^{p+1-\Delta/q} \quad (56)$$

since $\eta \in (0, 1)$. By the definition of the deficiencies, we also have that

$$W_L + W_R \leq q(2p + 1) - \Delta \quad (57)$$

which allows us to simplify Equation (53) as

$$\begin{aligned} \left(\prod_{j=1}^q k_j^{-2p} \right) |\mathcal{W}_{X,Y}(\mathcal{W}_1, \dots, \mathcal{W}_q)| &\leq n^{W_L} \eta^{W_L/q} \prod_{j=1}^q k_j^{\text{wt}_R(\mathcal{W}_j) - 2p} \\ &\leq n^{W_L} \cdot \eta^{W_L/q} \cdot \rho^{-2pq} n^{W_R - 2pq} && \text{By Equation (55)} \\ &\leq n^{W_L + W_R - 2pq} \cdot \eta^{p+1 - \Delta/q} \cdot \rho^{-2pq} && \text{By Equation (56)} \\ &\leq n^{q - \Delta} \cdot \eta^{p+1 - \Delta/q} \cdot \rho^{-2pq} && \text{By Equation (57)} \\ &= \rho^{-2pq} \cdot n^q \cdot \eta^{p+1} (n\eta^{1/q})^{-\Delta}, \end{aligned}$$

matching Equation (52). $\square$

We are now ready to bound $\mathbb{E}_\sigma[\text{tr}((TT^\top)^p)]$.

Lemma 8.31. With the setup of Proposition 8.1, for any $\lambda \in \Lambda$ let $X = X_\lambda, Y = Y_\lambda \subseteq [n]^q$ be η -sparse subsets and set $p = \lceil D \log(n) \rceil$. Then for all $n \geq N_{D,q,\eta}$ we have

$$\mathbb{E}_\sigma[\text{tr}((T_\lambda T_\lambda^\top)^p)] \leq 2^q n^q \eta^{p+1} (16^q \rho^{-2q})^p.$$

Note that, while we set p here, we do not fully eliminate it to make cancellations in subsequent computations cleaner.

Proof. Let $T = T_\lambda$. Corollary 8.9 and Lemma 8.24 together give us the expression

$$\mathbb{E}_\sigma[\text{tr}((TT^\top)^p)] = \sum_{(\Delta_1, \dots, \Delta_q) \in \{0, \dots, 2p\}^q} \sum_{(\mathcal{W}_1, \dots, \mathcal{W}_q) \in \mathfrak{W}_p(\Delta_1) \times \dots \times \mathfrak{W}_p(\Delta_q)} \left(\prod_{j=1}^q k_j^{-2p} \right) |\mathcal{W}_{X,Y}(\mathcal{W}_1, \dots, \mathcal{W}_q)|.$$

Using the result of Lemma 8.30 yields

$$\mathbb{E}_\sigma[\text{tr}((TT^\top)^p)] \leq \sum_{(\Delta_1, \dots, \Delta_q) \in \{0, \dots, 2p\}^q} \sum_{(\mathcal{W}_j)_j \in \prod_{j=1}^q \mathfrak{W}_p(\Delta_j)} \rho^{-2pq} n^q \eta^{p+1} (n\eta^{1/q})^{-\Delta}, \quad (58)$$

where we recall that $\Delta = \sum_{j=1}^q \Delta_j$. Because the summand in Equation (58) does not depend on $\mathcal{W}_1, \dots, \mathcal{W}_q$, the second sum just contributes a factor of

$$\prod_{j=1}^q |\mathfrak{W}_p(\Delta_j)| \leq \prod_{j=1}^q 16^p (4p)^{6\Delta_j} \quad (59)$$

$$= 16^{pq} (4p)^{6\Delta}. \quad (60)$$

All that is left now is to deal with the first sum of Equation (58). We first observe some useful inequalities by our choice of parameters. Because $p = \lceil D \log(n) \rceil$, we have $4p \leq 8D \log(n)$. Thus, using the inequality

$$\log(n) \leq \frac{n^\alpha}{\alpha e}$$

for $n \geq 1$ and $\alpha > 0$ we can set $\alpha = 1/24$ to compute

$$\begin{aligned}
(4p)^6 &\leq (8D \log(n))^6 \\
&\leq \left(8D \cdot \frac{24n^{1/24}}{e}\right)^6 \\
&= \left(\frac{192D}{e}\right)^6 n^{1/4} \\
&\leq \frac{n^{1/2}}{2},
\end{aligned} \tag{61}$$

where the last inequality is due to our choice of $n \geq N_{D,q,\eta} \geq 16 \left(\frac{192D}{e}\right)^{24}$. Consequently, using the fact that $\eta \geq n^{-q/2}$ so $\eta^{1/q} \geq n^{-1/2}$ we can use Equation (61) to get

$$\frac{(4p)^6}{n\eta^{1/q}} \leq \frac{1}{2}. \tag{62}$$

Finally, we put all these inequalities to use to reduce Equation (58) as follows:

$$\begin{aligned}
\mathbb{E}[\text{tr}((TT^\top)^p)] &\leq \sum_{(\Delta_1, \dots, \Delta_q) \in \{0, \dots, 2p\}^q} \sum_{(\mathcal{W}_j)_j \in \prod_{j=1}^q \mathfrak{W}_p(\Delta_j)} \rho^{-2pq} n^q \eta^{p+1} (n\eta^{1/q})^{-\Delta} \\
&\leq \sum_{(\Delta_1, \dots, \Delta_q) \in \{0, \dots, 2p\}^q} (16^{pq} (4p)^{6\Delta}) \left(\rho^{-2pq} n^q \eta^{p+1} (n\eta^{1/q})^{-\Delta} \right) \quad \text{By Equation (60)} \\
&= n^q \eta^{p+1} (16^q \rho^{-2q})^p \prod_{j=1}^q \sum_{\Delta_j=0}^{2p} \left(\frac{(4p)^6}{n\eta^{1/q}} \right)^{\Delta_j} \\
&\leq n^q \eta^{p+1} (16^q \rho^{-2q})^p \prod_{j=1}^q \sum_{\Delta_j=0}^{\infty} \left(\frac{(4p)^6}{n\eta^{1/q}} \right)^{\Delta_j} \\
&\leq n^q \eta^{p+1} (16^q \rho^{-2q})^p \prod_{j=1}^q \sum_{\Delta_j=0}^{\infty} \left(\frac{1}{2} \right)^{\Delta_j} \quad \text{By Equation (62)} \\
&= 2^q n^q \eta^{p+1} (16^q \rho^{-2q})^p,
\end{aligned}$$

as claimed. $\square$

8.7 Putting it all together

We can finally prove Proposition 8.1.

Proposition 8.1 (Proposition 7.11 restated). Take fixed integers $\nu \geq 0$ and $q \geq 1$ such that $\nu + q \leq D$. Let Λ be a fixed indexing set such that $|\Lambda| \leq n^\nu$. For all indices $\lambda \in \Lambda$, take $X_\lambda, Y_\lambda \subseteq [n]^q$ to be η -sparse for some constant $\eta \in (0, 1)$, fixed independently of the signs. Letting

$$T_\lambda = \Pi_{X_\lambda} \left(\bigotimes_{j=1}^q D_j \right) \iota_{Y_\lambda} : \mathbb{R}^{Y_\lambda} \rightarrow \mathbb{R}^{X_\lambda},$$

we have that for $n \geq N_{D,q,\eta} := \max \left\{ 16 \left(\frac{192D}{e} \right)^{24}, \eta^{-2/q} \right\}$,

$$\mathbb{E}[\max_{\lambda \in \Lambda} \|T_\lambda\|_{\text{op}}] \leq 2 \left(\frac{4}{\rho} \right)^q \sqrt{\eta}.$$

If $\Lambda = \emptyset$, then this expectation is just 0.

Proof. As in Lemma 8.31, we let $p = \lceil D \log(n) \rceil$. By Hölder's inequality,

$$\begin{aligned}
\mathbb{E}_\sigma \left[\max_{\lambda \in \Lambda} \|T_\lambda\|_{\text{op}} \right] &= \mathbb{E}_\sigma \left[1 \cdot \max_{\lambda \in \Lambda} \|T_\lambda\|_{\text{op}} \right] \\
&\leq (\mathbb{E}_\sigma [1^{2p/(2p-1)}])^{(2p-1)/2p} \cdot \left(\mathbb{E}_\sigma \left[\left(\max_{\lambda \in \Lambda} \|T_\lambda\|_{\text{op}} \right)^{2p} \right] \right)^{1/(2p)} \\
&= \left(\mathbb{E}_\sigma \left[\max_{\lambda \in \Lambda} \|T_\lambda\|_{\text{op}}^{2p} \right] \right)^{1/(2p)}.
\end{aligned} \tag{63}$$

Next, we use a summation and Fact 3.21 to bound the maximum as

$$\begin{aligned}
\max_{\lambda \in \Lambda} \|T_\lambda\|_{\text{op}}^{2p} &\leq \sum_{\lambda \in \Lambda} \|T_\lambda\|_{\text{op}}^{2p} \\
&\leq \sum_{\lambda \in \Lambda} \text{tr}((T_\lambda T_\lambda^\top)^p).
\end{aligned} \tag{64}$$

We now put Equations (63) and (64) together with the assumption that $|\Lambda| \leq n^\nu$ to compute

$$\begin{aligned}
\mathbb{E}_\sigma \left[\max_{\lambda \in \Lambda} \|T_\lambda\|_{\text{op}} \right] &\leq \left(\mathbb{E}_\sigma \left[\sum_{\lambda \in \Lambda} \text{tr}((T_\lambda T_\lambda^\top)^p) \right] \right)^{1/(2p)} \\
&= \left(\sum_{\lambda \in \Lambda} \mathbb{E}_\sigma [\text{tr}((T_\lambda T_\lambda^\top)^p)] \right)^{1/(2p)} \\
&\leq \left(\sum_{\lambda \in \Lambda} 2^q n^q \eta^{p+1} (16^q \rho^{-2q})^p \right)^{1/(2p)} && \text{By Lemma 8.31} \\
&\leq (n^\nu 2^q n^q \eta^{p+1} (16^q \rho^{-2q})^p)^{1/(2p)} \\
&= 2^{q/(2p)} n^{(q+\nu)/(2p)} \eta^{(p+1)/(2p)} \left(\frac{4}{\rho} \right)^q
\end{aligned} \tag{65}$$

Using our assumption that $q + \nu \leq D$, $q \leq D$, and $p \geq D \log n$ we can simplify the first two terms of Equation (65) as

$$\begin{aligned}
2^{q/(2p)} n^{(q+\nu)/(2p)} &\leq 2^{\frac{D}{2D \log n}} n^{\frac{D}{2D \log n}} \\
&= \exp \left(\frac{\log 2}{2 \log n} + \frac{1}{2} \right) \\
&< 2
\end{aligned} \tag{66}$$

where the last inequality uses $n \geq N_{D,q,\eta} \geq 16$. To bound the third term of Equation (65), we note that $\frac{p+1}{2p} \geq 1/2$ and $0 < \eta \leq 1$ so

$$\eta^{(p+1)/(2p)} \leq \sqrt{\eta}. \tag{67}$$

Plugging Equations (67) and (66) into Equation (65) yields our desired claim. $\square$

9 Open Questions

There are several immediate directions that we think would be interesting and worthwhile to explore.

1. The transference of our result from the reals to finite fields in the proof of Theorem 6.1 requires our base field to have characteristic $O(2^{n^D} n^D \log n)$ instead of two as obtained by [KP25]. Can this be avoided and our method be made to yield the same result over an extension field of characteristic two?
2. Different parametrizations of the generator matrices used to define a family of collections of codes can force the component codes to have desired groups contained in their automorphism groups. Can our techniques be used to analyze such a model? And if so, what automorphism groups can be achieved? The main difficulty here is that we explicitly use the independence of the signs in our random Rademacher generator matrices in Sections 7 and 8, and non-trivial parametrizations might introduce relations amongst these signs.
3. Besides analyzing the product expansion of tensor codes, one could analyze the coboundary expansion of more general sheaf codes that could be used as the local codes on other high-dimensional expanders, such as simplicial HDXs. In the formalization of [PK24], all of the main components of our analysis generalize. Corestriction maps and extendable/inner-generated sets are still well-defined, planes become the set of top-level neighbors of an element in a poset, and we can still use our random Rademacher model for the generator matrices. However, the clean tensor decomposition of blocks of the plane-interaction matrix from Section 7 is no longer possible, and any other decomposition will depend on the exact sheaf codes being studied. It is especially tempting to combine this question with the last one to ask what automorphisms we can require these codes to have.
4. Similar to the last couple questions, can our techniques be used to analyze polynomially parametrized families of codes that result in instances having the multiplication property? This would be interesting in both the cubical and simplicial settings to achieve global codes with the multiplication property. As previously mentioned, this is relevant when trying to construct quantum codes that are useful for computation [Ngu25].

References

- [AZ06] Greg W. Anderson and Ofer Zeitouni. “A CLT for a Band Matrix Model”. en. In: *Probability Theory and Related Fields* 134.2 (Feb. 2006), pp. 283–338. ISSN: 1432-2064. DOI: [10.1007/s00440-004-0422-3](#) (cit. on pp. 6, 7, 13, 15, 52, 54, 56).
- [BLN26] Mitali Bafna, Anqi Li, and Quynh Nguyen. *Good Quantum Locally Testable Codes from Product Expansion*. en. ISSN: 1433-8092. Sept. 2026 (cit. on pp. 3, 6).
- [BGHMRS15] Boaz Barak, Parikshit Gopalan, Johan Håstad, Raghu Meka, Prasad Raghavendra, and David Steurer. “Making the Long Code Shorter”. In: *SIAM Journal on Computing* 44.5 (Jan. 2015), pp. 1287–1324. ISSN: 0097-5397. DOI: [10.1137/130929394](#) (cit. on p. 3).
- [BE21] Nikolas P. Breuckmann and Jens Niklas Eberhardt. “Balanced Product Quantum Codes”. In: *IEEE Trans. Inf. Theory* 67.10 (2021), pp. 6653–6674. DOI: [10.1109/TIT.2021.3097347](#) (cit. on p. 3).
- [CHLT26] Yeyuan Chen, Miryam Mi-Ying Huang, Yincheng Liu, and Er-Cheng Tang. *Cubical Sheaf Complexes with Constant Expansion with Applications to Asymptotically Good qLTCs*. arXiv:2609.28028 [cs.IT]. Sept. 2026. DOI: [10.48550/arXiv.2609.28028](#) (cit. on pp. 3, 6).
- [CLRTT26] Itay Cohen, Itai Leigh, Assaf Reiner, Amnon Ta-Shma, and Elad Tzalik. *Good Quantum Locally Testable Codes from Lossless Cubical Complexes*. en. ISSN: 1433-8092. Sept. 2026 (cit. on p. 6).
- [DELLM22] Irit Dinur, Shai Evra, Ron Livne, Alexander Lubotzky, and Shahar Mozes. “Locally testable codes with constant rate, distance, and locality”. In: *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*. ACM Conferences. June 2022, pp. 357–374. ISBN: 978-1-4503-9264-8. DOI: [10.1145/3519935.3520024](#) (cit. on pp. 3–5, 7).

- [DHLV22] Irit Dinur, Min-Hsiu Hsieh, Ting-Chun Lin, and Thomas Vidick. *Good Quantum LDPC Codes with Linear Time Decoders*. arXiv:2206.07750 [quant-ph]. June 2022. DOI: [10.48550/arXiv.2206.07750](#) (cit. on pp. 3, 4).
- [DK17] Irit Dinur and Tali Kaufman. “High Dimensional Expanders Imply Agreement Expanders”. In: *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*. Ed. by Chris Umans. IEEE Computer Society, 2017, pp. 974–985. DOI: [10.1109/FOCS.2017.94](#) (cit. on p. 3).
- [DLV24] Irit Dinur, Ting-Chun Lin, and Thomas Vidick. “Expansion of High-Dimensional Cubical Complexes: with Application to Quantum Locally Testable Codes”. In: *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*. ISSN: 2575-8454. Oct. 2024, pp. 379–385. DOI: [10.1109/FOCS61266.2024.00031](#) (cit. on pp. 3–6).
- [DLZ23] Irit Dinur, Siqi Liu, and Rachel Yun Zhang. *New Codes on High Dimensional Expanders*. arXiv:2308.15563 [cs]. Aug. 2023. DOI: [10.48550/arXiv.2308.15563](#) (cit. on pp. 3, 6).
- [EH17] Lior Eldar and Aram W. Harrow. “Local Hamiltonians Whose Ground States Are Hard to Approximate”. In: *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*. Ed. by Chris Umans. IEEE Computer Society, 2017, pp. 427–438. DOI: [10.1109/FOCS.2017.46](#) (cit. on p. 3).
- [FK24] Uriya A. First and Tali Kaufman. *On Good $\$2\$$ -Query Locally Testable Codes from Sheaves on High Dimensional Expanders*. arXiv:2208.01778 [math]. May 2024. DOI: [10.48550/arXiv.2208.01778](#) (cit. on p. 3).
- [FK81] Z. Füredi and J. Komlós. “The eigenvalues of random symmetric matrices”. en. In: *Combinatorica* 1.3 (Sept. 1981), pp. 233–241. ISSN: 1439-6912. DOI: [10.1007/BF02579329](#) (cit. on pp. 7, 13, 52).
- [GJ26] William Gay and Fernando Granha Jeronimo. *Asymptotically Good Quantum Locally Testable Codes*. arXiv:2609.20780 [quant-ph]. Sept. 2026. DOI: [10.48550/arXiv.2609.20780](#) (cit. on pp. 3, 6).
- [GVSV06] Oded Goldreich, View Profile, Madhu Sudan, and View Profile. “Locally testable codes and PCPs of almost-linear length”. In: *Journal of the ACM* 53.4 (July 2006), pp. 558–655. DOI: [10.1145/1162349.1162351](#) (cit. on p. 3).
- [GG25] Louis Golowich and Venkatesan Guruswami. “Quantum LDPC Codes of Almost Linear Distance via Iterated Homological Products”. In: *40th Computational Complexity Conference (CCC 2025)*. Ed. by Srikanth Srinivasan. Vol. 339. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2025, 25:1–25:11. ISBN: 978-3-95977-379-9. DOI: [10.4230/LIPIcs.CCC.2025.25](#) (cit. on p. 4).
- [GK25] Kyle Gulshen and Tali Kaufman. *Quantum Tanner Color Codes on Qubits with Transversal Gates*. arXiv:2510.07864 [quant-ph]. Oct. 2025. DOI: [10.48550/arXiv.2510.07864](#) (cit. on p. 6).
- [JH91a] “Matrix equations and the Kronecker product”. In: *Topics in Matrix Analysis*. Ed. by Charles R. Johnson and Roger A. Horn. Cambridge: Cambridge University Press, 1991, pp. 239–297. ISBN: 978-0-521-46713-1. DOI: [10.1017/CB09780511840371.005](#) (cit. on p. 20).
- [JH91b] “Singular value inequalities”. In: *Topics in Matrix Analysis*. Ed. by Charles R. Johnson and Roger A. Horn. Cambridge: Cambridge University Press, 1991, pp. 134–238. ISBN: 978-0-521-46713-1. DOI: [10.1017/CB09780511840371.004](#) (cit. on p. 20).
- [Kal23] Gleb Kalachev. *High-dimensional Expansion of Product Codes is Stronger than Robust and Agreement Testability*. arXiv:2308.02889 [cs]. Aug. 2023. DOI: [10.48550/arXiv.2308.02889](#) (cit. on p. 3).

- [KP23] Gleb Kalachev and Pavel Panteleev. *Two-sided Robustly Testable Codes*. arXiv:2206.09973 [cs]. Aug. 2023. DOI: [10.48550/arXiv.2206.09973](https://doi.org/10.48550/arXiv.2206.09973) (cit. on pp. 3, 4).
- [KP25] Gleb Kalachev and Pavel Panteleev. “Maximally Extendable Product Codes are Good Coboundary Expanders”. In: *2025 IEEE 66th Annual Symposium on Foundations of Computer Science (FOCS)*. ISSN: 2575-8454. Dec. 2025, pp. 1512–1524. DOI: [10.1109/FOCS63196.2025.00079](https://doi.org/10.1109/FOCS63196.2025.00079) (cit. on pp. 4–8, 29, 30, 38, 39, 63).
- [KM25] Vinayak M. Kumar and Geoffrey Mon. “Relaxed Local Correctability from Local Testing”. In: *SIAM Journal on Computing* (Sept. 2025). Num Pages: STOC24-119, STOC24-100. ISSN: 0097-5397. DOI: [10.1137/24M1687169](https://doi.org/10.1137/24M1687169) (cit. on p. 3).
- [LWZ26] Xiaoru Li, Qi Wang, and Yue Zhou. *The automorphism groups of random linear codes*. arXiv:2607.23936 [cs.IT]. Aug. 2026. DOI: [10.48550/arXiv.2607.23936](https://doi.org/10.48550/arXiv.2607.23936) (cit. on p. 5).
- [Mes18] Roy Meshulam. *Graph codes and local systems*. arXiv:1803.05643 [math.CO]. Mar. 2018. DOI: [10.48550/arXiv.1803.05643](https://doi.org/10.48550/arXiv.1803.05643) (cit. on p. 3).
- [MV06] Hugh L. Montgomery and Robert C. Vaughan. “The elementary theory of arithmetic functions”. In: *Multiplicative Number Theory I: Classical Theory*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2006, pp. 35–75 (cit. on p. 20).
- [Ngu25] Quynh T. Nguyen. “Good Binary Quantum Codes with Transversal CCZ Gate”. In: *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*. ACM Conferences. June 2025, pp. 697–706. ISBN: 979-8-4007-1510-5. DOI: [10.1145/3717823.3718186](https://doi.org/10.1145/3717823.3718186) (cit. on p. 63).
- [NP25] Quynh T. Nguyen and Christopher A. Pattison. “Quantum Fault Tolerance with Constant-Space and Logarithmic-Time Overheads”. In: *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*. ACM Conferences. June 2025, pp. 730–737. ISBN: 979-8-4007-1510-5. DOI: [10.1145/3717823.3718318](https://doi.org/10.1145/3717823.3718318) (cit. on pp. 3, 4).
- [PK22] Pavel Panteleev and Gleb Kalachev. *Asymptotically Good Quantum and Locally Testable Classical LDPC Codes*. arXiv:2111.03654 [cs]. Jan. 2022. DOI: [10.48550/arXiv.2111.03654](https://doi.org/10.48550/arXiv.2111.03654) (cit. on pp. 3, 4).
- [PK24] Pavel Panteleev and Gleb Kalachev. *Maximally Extendable Sheaf Codes*. arXiv:2403.03651 [cs]. Mar. 2024. DOI: [10.48550/arXiv.2403.03651](https://doi.org/10.48550/arXiv.2403.03651) (cit. on pp. 3–6, 63).
- [RSV19] Nithi Rungtanapirom, Jakob Stix, and Alina Vdovina. “Infinite series of quaternionic 1-vertex cube complexes, the doubling construction, and explicit cubical Ramanujan complexes”. In: *International Journal of Algebra and Computation* 29.06 (Sept. 2019). arXiv:1808.03290 [math], pp. 951–1007. ISSN: 0218-1967, 1793-6500. DOI: [10.1142/S0218196719500371](https://doi.org/10.1142/S0218196719500371) (cit. on p. 5).
- [Tao12] Terence Tao. *Topics in random matrix theory*. en. Vol. 132. Graduate Studies in Mathematics. American Mathematical Society, 2012. ISBN: 978-1-4704-7459-1 (cit. on p. 20).
- [Tro15] Joel A. Tropp. “An Introduction to Matrix Concentration Inequalities”. In: *Foundations and Trends® in Machine Learning* 8.1-2 (May 2015), pp. 1–230. ISSN: 1935-8237. DOI: [10.1561/22000000048](https://doi.org/10.1561/22000000048) (cit. on p. 20).
- [Ver18] Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge University Press, 2018. DOI: [10.1017/9781108231596](https://doi.org/10.1017/9781108231596) (cit. on p. 49).
- [Wat18] John Watrous. *The Theory of Quantum Information*. Cambridge: Cambridge University Press, 2018. ISBN: 978-1-107-18056-7. DOI: [10.1017/9781316848142](https://doi.org/10.1017/9781316848142) (cit. on p. 20).
- [Wig55] Eugene P. Wigner. “Characteristic Vectors of Bordered Matrices With Infinite Dimensions”. In: *Annals of Mathematics* 62.3 (1955), pp. 548–564. ISSN: 0003-486X. DOI: [10.2307/1970079](https://doi.org/10.2307/1970079) (cit. on pp. 7, 13, 52, 54).