Electronic Colloquium on Computational Complexity, Revision 2 of Report No. 119 (2021)

Visible Rank and Codes with Locality*

Omar Alrabiah[†] Venkatesan Guruswami[‡]

Abstract

We propose a framework to study the effect of local recovery requirements of codeword symbols on the dimension of linear codes, based on a combinatorial proxy that we call *visible rank*. The locality constraints of a linear code are stipulated by a matrix H of 0's and $\star$'s (which we call a "stencil"), whose rows correspond to the local parity checks (with the $\star$'s indicating the support of the check). The visible rank of H is the largest r for which there is a $r \times r$ submatrix in H with a unique generalized diagonal of $\star$'s. The visible rank yields a field-independent combinatorial lower bound on the rank of any matrix obtained by replacing the stars in H by nonzero field elements and thus the co-dimension of the code.

We point out connections of the visible rank to other notions in the literature such as unique restricted graph matchings, matroids, spanoids, and min-rank. In particular, we prove a rank-nullity type theorem relating visible rank to the rank of an associated construct called *symmetric spanoid*, which was introduced by Dvir, Gopi, Gu, and Wigderson [DGGW20]. Using this connection and a construction of appropriate stencils, we answer a question posed by [DGGW20] and demonstrate that the symmetric spanoid rank cannot improve the currently best known $\widetilde{O}(n^{(q-2)/(q-1)})$ upper bound on the dimension of q-query locally correctable codes (LCCs) of length n. This also pins down the efficacy of visible rank as a proxy for the dimension of LCCs.

We also study the t-Disjoint Repair Group Property (t-DRGP) of codes where each codeword symbol must belong to t disjoint check equations. It is known that linear codes with 2-DRGP must have co-dimension $\Omega(\sqrt{n})$ (which is matched by a simple product code construction). We show that there are stencils corresponding to 2-DRGP codes with visible rank as small as $O(\log n)$. However, we show that the second tensor power of any 2-DRGP stencil has visible rank $\Omega(n)$, thus recovering the $\Omega(\sqrt{n})$ lower bound on the co-dimension of 2-DRGP codes. More broadly, we show that visible ranks of tensor powers yield sharper lower bounds on the co-dimension of the code. For q-LCCs, however, the k'th tensor power for $k \leqslant n^{o(1)}$ is unable to improve the $\widetilde{O}(n^{(q-2)/(q-1)})$ upper bound on the dimension of q-LCCs by a polynomial factor. Inspired by this and as a notion of intrinsic interest, we define the notion of *visible capacity* of a stencil as the limiting visible rank of high tensor powers, analogous to Shannon capacity, and pose the question of whether there can be large gaps between visible capacity and algebraic rank.

*An earlier version of this work was presented at 2021 RANDOM/APPROX [AG21]. The current version includes Theorem 5, which is a solution to Question 2 that was asked in [AG21]. This work was done when both authors were affiliated with Carnegie Mellon University.

[†]Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94704. Email: `oalrabiah@berkeley.edu`. Research supported in part by NSF grants CCF-1814603 and CCF-1908125 and V. Guruswami's Simons Investigator Award.

[‡]Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94704. Email: `venkatg@berkeley.edu`. Research supported in part by NSF grants CCF-1814603 and CCF-1908125 and a Simons Investigator Award.

ISSN 1433-8092

1 Introduction

The notion of *locality* in error-correcting codes refers to the concept of recovering codeword symbols as a function of a small number of other codeword symbols. Local decoding requirements of various kinds have received a lot of attention in coding theory, due to both their theoretical and practical interest. For instance, q-query locally correctable codes (LCCs) aim to recover any codeword symbol as a function of q other codeword symbols in a manner robust to a constant fraction of errors. On the other hand, locally recoverable codes (LRCs), in their simplest incarnation, require each codeword symbol to be a function of some ℓ other codeword symbols, allowing local recovery from any single erasure.[1]

LCCs have been extensively studied in theoretical computer science, and have connections beyond coding theory to topics such as probabilistically checkable proofs and private information retrieval. We refer the reader to [Yek12] and the introduction of [Gop18] for excellent surveys on LCCs and their connections. On the other hand, LRCs were motivated by the need to balance global fault tolerance with extremely efficient repair of a small number of failed storage nodes in modern large-scale distributed storage systems [GHSY12]. They have led to intriguing new theoretical questions, and have also had a significant practical impact with adoption in large-scale systems such as Microsoft Azure [HSX+12] and Hadoop [SAP+13].

Let us define the above notions formally, in a convenient form that sets up this work. Throughout this work, we will restrict our attention to linear codes, i.e., subspaces C of $\mathbb{F}^n$ for some finite field $\mathbb{F}$. In this case, any nontrivial repair scheme can be reduced to one that outputs a linear combination of other symbols. More formally, the i'th symbol c_i of every codeword $c = (c_1, \ldots, c_n) \in C$ can be recovered as a function of the symbols $\{c_j \mid j \in R_i\}$, for a (minimal) subset $R_i \subset [n] \setminus \{i\}$, if and only if c_i and $\{c_j \mid j \in R_i\}$ satisfy a linear check equation. In such a case, we say that R_i is a *repair group* for the i'th codeword symbol.[2] Now, the codeword symbols $\{c_i\} \cup \{c_j \mid j \in R_i\}$ satisfy a linear check equation if and only if there exists a dual codeword whose support equals $\{i\} \cup R_i$. Thus a linear code possessing locality is equivalent to having numerous dual codewords whose supports satisfy given combinatorial properties This specific formulation of code locality is the one that we will primarily focus on in this work.

With this setup laid out, we can now elucidate the notions of LCCs and LRCs. The q-LCC property, for a fixed number of queries q and growing n, corresponds to having $\Omega(n)$ disjoint repair groups of size $\leqslant q$ for each position $i \in [n]$, or equivalently $\Omega(n)$ dual codewords of Hamming weight at most $(q+1)$ whose support includes i and are otherwise disjoint. On the other hand, the ℓ-LRC property corresponds to having a dual codeword of Hamming weight at most $(\ell+1)$ whose support includes i, for each $i \in [n]$.

Another property that interpolates between these extremes of a single repair group and $\Omega(n)$ disjoint repair groups is the Disjoint Repair Group Property (t-DRGP) where we require t disjoint repair groups for each position $i \in [n]$ (equivalently t dual codewords whose support includes i but are otherwise disjoint). Unlike the q-LCC property and the ℓ-LRC property, the t-DRGP property imposes no size constraints on its repair groups.

Turning to known code constructions and bounds, we know that 1-LCCs do not exist [KT00], and the optimal 2-LCC dimension is known to be $\Theta(\log n)$ [GKST06, KW04, DS05, BGT17]. As for $q \geqslant 3$, there is currently an exponentially large gap between upper and lower bounds on the

[1]There is also a distance requirement on LRCs to provide more global error/erasure resilience.

[2]Other terminology used in the literature includes regenerating sets and recovery sets.

trade-off between code dimension and code length for q-LCCs. The best known code constructions have dimension only $O((\log n)^{q-1})$ (achieved by generalized Reed-Muller codes or certain lifted codes [GKS13]), whereas the best known upper bound on the dimension of q-LCCs is much larger and equals $\widetilde{O}(n^{(q-2)/(q-1)})$ [Woo12, IS20, DGGW20].[3] Narrowing this huge gap has remained open for over two decades.

In contrast, the best possible dimension of a ℓ-LRC is easily determined to be $\lfloor \frac{\ell n}{\ell+1} \rfloor$.[4] However, for t-DRGP, there are again some intriguing mysteries. For 2-DRGP, we have tight bounds—the minimum possible redundancy (co-dimension) equals $\sqrt{2n} \pm \Theta(1)$. The lower bound is established via very elegant proofs based on the polynomial method [Woo16] or rank arguments [RV16]. However, for fixed $t > 2$, we do not know better lower bounds, and the best known constructions have co-dimension $\approx t\sqrt{n}$ [FVY15]. There are better constructions known for some values of $t = n^{\Theta(1)}$ [FGW17, LW21]. A lower bound on the co-dimension of $c(t)\sqrt{n}$ for some function $c(t)$ that grows with t seems likely, but has been elusive despite various attempts, and so far for any fixed t, the bound for $t = 2$ is the best known.

This work was motivated in part by these major gaps in our knowledge concerning q-LCCs and t-DRGPs. Our investigation follows a new perspective based on *visible rank* (to be defined soon), which is a combinatorial proxy for (linear-algebraic) rank that we believe is of broader interest. This is similar in spirit to a thought-provoking recent work [DGGW20] that introduced a combinatorial abstraction of spanning structures called *spanoids*[5] to shed light on the limitations of current techniques to prove better upper bounds on the dimension of q-LCCs. They noted that current techniques to bound q-LCC dimension apply more generally to the associated spanoids, which they showed could have rank as large as $\widetilde{\Omega}(n^{(q-2)/(q-1)})$. Therefore, to improve the current LCC bounds, one needs techniques that are more specific than spanoids and better tailored to the LCC setting. One such possibility mentioned in [DGGW20] is to restrict attention to *symmetric spanoids*, which have a natural symmetry property that linear LCCs imply.

Our visible rank notion turns out to be intimately related to symmetric spanoids via a rank-nullity type theorem (Theorem 10). While technically simple in hindsight, it offers a powerful viewpoint on symmetric spanoids which in particular resolves a question posed by [DGGW20]—we show that symmetric spanoids are also too coarse a technique to beat the $\widetilde{O}(n^{(q-2)/(q-1)})$ upper bound on q-LCC dimension.

Before diving into our setup, we must note that our framework and the framework of [DGGW20] make no assumption on the underlying field. Indeed, if we were to restrict our attention to constant-sized fields, then the works of [KW04, Woo07, AGKM23] show an improved q-LCC dimension upper bound of $\widetilde{O}(n^{(q-2)/q})$ for $q = 3$ and even $q \geqslant 4$. In fact, their lower bounds also hold for the weaker notion of q-query Locally Decodable Codes (LDCs), wherein one can recover any message symbol (rather than a codeword symbol) by randomly querying q codewords and recovering the message symbol with a constant advantage against a constant fraction of adversarial errors. By considering a systematic encoding of a linear q-LCC, we immediately see that the q-LCC property implies the q-LDC property. On the other hand, if our field was the real numbers, then the works of [BDYW11, DSW17] show that 2-LCCs over the reals do not exist, and the 3-LCC dimension over the reals is at most $O(n^{1/2-c})$ for some universal constant $c > 0$.

[3]The $\widetilde{O}(\cdot)$ and $\widetilde{\Omega}(\cdot)$ are used to suppress factors poly-logarithmic in n.

[4]In this case, a more interesting trade-off is a Singleton-type bound that also factors in the distance of the code [GHSY12].

[5]We defer a precise description of spanoids, along with their strong connection to visible rank, to Section 2.4.

1.1 Stencils and visible rank

With the above backdrop, we now proceed to describe the setup we use to study these questions, based on the rank of certain matrix templates which we call "stencils," which are simply $\{0, \star\}$ matrices capturing the support of a matrix. We can represent the support structure of the check equations (i.e., dual codewords) governing a locality property by an n-column matrix of 0's and $\star$'s. For each check equation involving the i'th symbol and a repair group $R_i \subset [n] \setminus \{i\}$, we place a row in the stencil with $\star$'s precisely at $R_i \cup \{i\}$ (i.e., with $\star$'s at the support of the associated dual codeword). For the ℓ-LRC property for instance, an associated stencil would be an $n \times n$ matrix with $\star$'s on the diagonal and at most ℓ other $\star$'s in each row. For q-LCC, we would have a $\delta n^2 \times n$ matrix (for some fixed constant $\delta > 0$) whose rows are split into n groups with the rows in the i'th group having a $\star$ in the i'th column and at most q other $\star$'s in disjoint columns.

The smallest co-dimension of linear codes over a field $\mathbb{F}$ with certain locality property is, by design, the minimum rank $\mathrm{rk}_{\mathbb{F}}(H)$ of the associated stencil H when the $\star$'s are replaced by arbitrary nonzero entries from $\mathbb{F}$. In this work, our goal is to understand this quantity via field oblivious methods based only on the combinatorial structure of the stencil of $\star$'s.

The tool we put forth for this purpose is the *visible rank* of H, denoted $\mathrm{vrk}(H)$, and defined to be the largest r for which there is a $r \times r$ submatrix of H that has exactly one generalized diagonal whose entries are all $\star$'s. By the Leibniz formula, the determinant of such a submatrix is nonzero for any substitution of nonzero entries for the $\star$'s. Thus $\mathrm{rk}_{\mathbb{F}}(H) \geqslant \mathrm{vrk}(H)$ for every field $\mathbb{F}$.

Our goal in this work is to understand the interrelationship between visible rank and the co-dimension of linear codes under various locality requirements. This can shed further light on the bottleneck in known techniques to study trade-offs between locality and code dimension, and optimistically could also lead to better constructions.

1.2 Visible rank and locality

For ℓ-LRCs, a simple greedy argument shows that its associated parity-check stencil H satisfies $\mathrm{vrk}(H) \geqslant n/(\ell+1)$. Thus visible rank captures the optimal trade-off between code dimension and locality ℓ.

For q-LCCs with $q \geqslant 3$, an argument similar to (in fact a bit simpler than and implied by) the one for spanoids in [DGGW20] shows that the stencil corresponding to q-LCCs has visible rank at least $n - \widetilde{O}(n^{(q-2)/(q-1)})$, showing an upper bound of $\widetilde{O}(n^{(q-2)/(q-1)})$ on the dimension of q-LCCs. We show that visible rank suffers the same bottleneck as spanoids in terms of bounding the dimension of q-LCCs.

Theorem 1. *For $q \geqslant 3$, there exist n-column stencils H with $\star$'s structure compatible with q-LCCs for which $vrk(H) \leqslant n - \widetilde{\Omega}(n^{(q-2)/(q-1)})$.*

Through the precise connection we establish between visible rank and symmetric spanoids, this shows the same limitation for symmetric spanoids, thus answering a question posed by [DGGW20].

For the t-DRGP property, we focus on the $t = 2$ case, with the goal of finding a combinatorial substitute for the currently known $\Omega(\sqrt{n})$ lower bounds on co-dimension [Woo16, FGW17] which are algebraic. Unfortunately, we show that visible rank, in its basic form, is too weak in this context.

Theorem 2. *There exist $2n \times n$ stencils H with $\star$'s structure compatible with 2-DRGP for which $vrk(H) \leqslant O(\log n)$.*

1.3 Visible rank and tensor powers

In view of Theorem 2, we investigate avenues to get better bounds out of the visible rank approach. Specifically, we study the visible rank of tensor powers of the matrix. It turns out that the visible rank is super-multiplicative: $\mathrm{vrk}(H \otimes H) \geqslant \mathrm{vrk}(H)^2$, while on the other hand, algebraic rank is sub-multiplicative, so higher tensor powers could yield better lower bounds on the rank. Indeed, we are able to show precisely this for 2-DRGP:

Theorem 3. *For every $2n \times n$ stencil H with $\star$'s structure compatible with 2-DRGP, we have $vrk(H \otimes H) \geqslant \Omega(n)$, and thus $rk_{\mathbb{F}}(H) \geqslant \Omega(\sqrt{n})$ for every field $\mathbb{F}$.*

On the other hand, for q-LCCs with $q \geqslant 3$, we show that higher tensor powers suffer the same bottleneck as Theorem 1.

Theorem 4. *For $q \geqslant 3$, there exist n-column stencils H with $\star$'s structure compatible with q-LCCs for which $vrk(H^{\otimes k})^{1/k} \leqslant n - \widetilde{\Omega}(n^{(q-2)/(q-1)})/k$ for any integer k. In particular, even for $k = n^{o(1)}$, we get no polynomial improvements to the current upper bounds on the dimension of q-LCCs.*

1.4 Visible capacity

Given the super-multiplicativity of visible rank under tensor powers, and drawing inspiration from the Shannon capacity of graphs, we put forth the notion of *visible capacity* of a matrix H of 0's and $\star$'s, defined as

$$\Upsilon(H) := \sup_k \mathrm{vrk}(H^{\otimes k})^{1/k} .$$

The visible capacity is also a field oblivious lower bound on algebraic rank $\mathrm{rk}_{\mathbb{F}}(H)$ for any field $\mathbb{F}$. It is not known whether there are stencils that exhibit a gap between visible capacity and its minimum possible $\mathrm{rk}_{\mathbb{F}}(H)$ over all fields $\mathbb{F}$.

In the spirit of visible capacity, we extend Theorems 2 and 3 to show that for any constant $t \in \mathbb{N}$, one can attain an exponentially better lower bound on the algebraic rank of a stencil H by examining the visible rank of $H^{\otimes t}$ rather than examining the visible rank of $H^{\otimes k}$ for any $k < t$. In other words:

Theorem 5. *For any fixed natural number $t \geqslant 2$, there exists a $tn \times n$ stencil H such that $vrk(H^{\otimes k}) = O_t((\log n)^k)$ for any $k = 1, \ldots, t-1$ and $vrk(H^{\otimes t}) = \Omega(n)$.*

The proofs of our results are technically simple, once the framework is set up. Our contributions are more on the conceptual side, via the introduction and initial systematic study of visible rank and its diverse connections. Our inquiry also raises interesting questions and directions for future work, some of which are outlined in Section 8, including the relationship between visible capacity and algebraic rank.

1.5 Connections and related work

Studying the interplay between the combinatorial structure of a matrix and its rank is a natural quest that arises in several contexts. See Chapter 3 of [Tré16] for a survey of works on lower bounding the algebraic rank. For works specific to codes with locality, the work of [BDYW11, DSW14] analyzed the combinatorial properties of design matrices over the reals to improve bounds on LCCs over the real numbers, although the methods used are particular to the field of reals and do not carry over to any field.

Visible rank in particular turns out to have a diverse array of connections, some of which we briefly discuss here. The connection to spanoids, that we already mentioned, is described in more detail in Section 2.4.

Uniquely restricted matchings. Given a stencil $H \in \{0, \star\}^{m \times n}$, there is a *canonical bipartite graph* G between the rows and columns of H, where a row connects to a column if and only if their shared entry has a $\star$. Visible rank has a nice graph-theoretic formulation: a submatrix of H has a unique generalized diagonal of $\star$'s iff the corresponding induced subgraph of G has a unique perfect matching. Such induced bipartite graphs are known in the literature as *Uniquely Restricted Matchings (URMs)* and have been extensively studied [GHL01, GKT01, HMT06, Mis11, TD13, FJJ18]. They were first introduced in [GHL01], wherein they proved that computing the maximum URM of a bipartite graph is NP-complete. It was later shown in [Mis11] that $n^{1/3-o(1)}$ approximations of the maximum URM is also NP-hard unless $\mathsf{NP} = \mathsf{ZPP}$ and additionally that the problem of finding the maximum URM is APX-complete. Given the graph-theoretic formulation of visible rank, the known computational hardness results in computing the maximum URM of a given bipartite graph can therefore be translated as computational hardness results in computing the visible rank of a given stencil.

Matroids. One can encode any matroid into a stencil. Recall that a circuit of a matroid is a minimal dependent set—that is, a dependent set whose proper subsets are all independent (the terminology reflects the fact that in a graphic matroid, the circuits are simple cycles of the graph). Given a matroid $\mathcal{M}$ on universe $[n]$ and a set $\mathcal{C} = \{C_1, \ldots, C_m\}$ of circuits of $\mathcal{M}$, we consider an $m \times n$ stencil H where the entry at (i, j) is a $\star$ if and only if $j \in C_i$. For this matrix, one can show that any collection of columns is visibly independent (see Section 2.2 for the definition of visible independence) if and only if it is an independent set in the dual matroid. Therefore, we have $\mathrm{rk}(\mathcal{M}) + \mathrm{vrk}(H) = n$—this also follows from our rank-nullity theorem for symmetric spanoids as one can associate a symmetric spanoid with any matroid (the collection of sets in Definition 5 will just be the circuits of the matroid).

Min-rank. The minimum possible rank over a given field $\mathbb{F}$ of a square 0-$\star$ stencil over assignments to the $\star$'s from $\mathbb{F}$ has been well studied in combinatorics. For example, we have Haemers' classic bound on the independence number of a graph and its applications to Shannon capacity [Hae79]. But while the purpose of min-rank is to use a linear-algebraic tool to understand a combinatorial quantity, our notion of visible rank goes the other way, serving as a combinatorial proxy for a linear-algebraic quantity. Furthermore, The setup of the $\star$'s differ slightly. For min-rank, the off-diagonal $\star$'s are allowed to be zero, while the diagonal entries are necessarily nonzero. For visible rank, every $\star$ is necessarily nonzero. Recent interest in min-rank has included their characterization of the most efficient linear index codes [BBJK11]. The min-rank of stencils corresponding to n-vertex random Erdös-Rényi graphs was recently shown to be $\Theta(n/\log n)$ with high probability over any field that is polynomially bounded [GRW18]. This was later then extended to arbitrary fields [ABG+20].

Matrix Rigidity. Given a square matrix $A \in \mathbb{F}^{n \times n}$ and a natural number $r \leqslant n$, the *rigidity* of A is the minimal number of entries that one can perturb in A so that its rank becomes at most r. Matrix rigidity was introduced in the seminal work [Val77] and since then has had expansive research on constructing explicit rigid matrices. See [Ram20] for a recent survey on matrix rigidity and related connections. Visible rank provides a combinatorial guarantee on the rank of a matrix, which opens up the possibility of constructing explicit rigid matrices by finding explicit stencils whose visible rank is robust to small amounts of corruption of its entries. We pose this approach as Question 5 in Section 8.

Incidence Theorems. Given an $m \times n$ matrix A over the field $\mathbb{F}$ of rank r, one can decompose $A = MN$ where M and N are $m \times r$ and $r \times n$ matrices over $\mathbb{F}$. If we consider the rows of M as hyperplanes over the $(r-1)$-dimensional projective plane $\mathbb{PF}^{r-1}$ and the columns of N as points in $\mathbb{PF}^{r-1}$, then the stencil of A defines a point-hyperplane incidence over $\mathbb{PF}^{r-1}$. In particular, when $r = 3$, the stencil of A defines a point-line incidence over the field $\mathbb{F}$. Thus studying the combinatorial properties of a stencil whose $\mathbb{F}$-rank (see Definition 2) is at most 3 is equivalent to studying the combinatorics of point-line incidences over the field $\mathbb{F}$. For more on incidence theorems, see [D$^+$12] for an excellent survey in the area.

Communication complexity. The visible rank provides a connection between deterministic and nondetereministic communication complexity [Lov89]. For a communication problem $f : X \times Y \to \{0, 1\}$, define the stencil $H_f \in \{0, \star\}^{X \times Y}$ by $H_f[x, y] = \star$ if $f(x, y) = 0$ and $H_f[x, y] = 0$ if $f(x, y) = 1$. Then it is known that $D(f) \leqslant (\log_2 \mathrm{vrk}(H_f)) \cdot (N(f) + 1)$ where $D(f)$ and $N(f)$ are respectively the deterministic and nondeterministic communication complexity of f [Lov89, Thm 3.5].

Distance of expander codes. We can associate a stencil with the parity-check matrix H of a code C in the natural way, by placing $\star$'s at the positions of the nonzero field elements. Suppose that every subset of d columns of the stencil is visibly independent. Then the minimum distance of C is at least d. This is in fact the lower bound on the distance of expander codes established in [SS96]. The visible independence of subsets of d columns is argued via *unique expansion* of the canonical bipartite graph associated with H (also called the factor graph in coding theory). Namely, for every subset S of d vertices on the left, there is a vertex on the right that's adjacent to exactly one vertex in S. Thus, linear independence is really argued via visible independence.

1.6 Organization

We begin in Section 2 by formally introducing the notations and terminology for stencils and establishing some simple but very useful combinatorial facts about visible rank. We use these to show that there are q-LCC stencils for $q \geqslant 3$ with visible rank at most $n - \widetilde{\Omega}(n^{(q-2)/(q-1)})$ (Section 3), and the existence of a 2-DRGP stencil with visible rank of at most $O(\log n)$ (Section 4). In Section 5, we introduce a tensor product operation on stencils and prove various properties about them. In Section 6, we utilize tensor powers to show that the rank of a 2-DRGP over any field $\mathbb{F}$ is at least $\sqrt{n}$, which asymptotically matches the current best lower bounds on t-DRGP codes. We also show that for q-LCC stencils, the tensor powers at the k'th level for $k \leqslant n^{o(1)}$ do not yield better lower bounds on the rank than the ones obtained from the visible rank. We then demonstrate in Section 7 how the visible rank of any fixed tensor power can yield an exponentially better lower bound on the algebraic rank. Finally, in Section 8, we discuss further directions and questions inspired by this work.

1.7 Later Works

After the completion of this work, Li and Wootters [LW22] proved a $\Omega(\sqrt{nk})$ redundancy lower bound for linear k-batch codes with a systematic encoding. A linear code $C : \mathbb{F}^m \to \mathbb{F}^n$ is said to be a k-batch code if every multiset of message indices $\{i_1, \ldots, i_k\} \subseteq [m]$ has k pairwise disjoint repair groups $R_1, \ldots, R_k \subseteq [n]$ such that for each $j \in [p]$, there is a linear combination of the codeword symbols at R_j that is equal to the i_j'th message symbol. Since [LW22] assumes a systematic encoding, the k-batch constraint on the code can be recast as the code being in the kernel of a matrix M with a particular support structure. They then proceed to show that the rank of M is high by lower bounding the rank of a $\Theta(k)$ tensor power of M through a careful analysis of its support structure. Given that their argument depended solely on the support structure of M, it can in fact be captured by our stencil framework by considering the corresponding stencil H of M, lower bounding $\mathrm{vrk}(H^{\otimes t})$ for $t = \Theta(k)$, and then applying Corollary 18 to conclude a lower bound on $\mathrm{rk}_{\mathbb{F}}(H)$.

2 Stencils and their visible rank

In this section, we will be formally setting up the model of stencils and all the associated definitions and notations.

We denote $[n]$ to be the set $\{1, 2, \ldots, n\}$. For any matrix $H \in \{0, \star\}^{m \times n}$, we denote it as a *stencil*. For an $m \times n$ stencil H, we denote its entry in the i'th row and j'th column by $H[i, j]$. Any restriction to the specific sub-collection of the rows and columns of H is said to be a *sub-stencil* of H. For given sets A and B, a stencil H is said to be an $A \times B$ stencil if it is an $|A| \times |B|$ stencil with an associated indexing of the rows by A and the columns by B. Given a square stencil $M \in \{0, \star\}^{n \times n}$, a *generalized diagonal* of M is a tuple of entries $(M[1, \pi(1)], \ldots, M[n, \pi(n)])$ where π is a permutation on $[n]$. We say that a generalized diagonal is a *star diagonal* if all n entries are $\star$'s.

2.1 Algebraic witnesses of stencils

Instantiating a code with the locality properties stipulated by a stencil amounts to filling its $\star$'s with nonzero field entries, or realizing an algebraic witness as defined below.

Definition 1 (Algebraic witness)**.** *For field $\mathbb{F}$ and stencil $H \in \{0, \star\}^{m \times n}$, a matrix $W \in \mathbb{F}^{m \times n}$ is said to be an $\mathbb{F}$-witness of H if it satisfies the property that $W[i, j] \neq 0$ if and only if $H[i, j] = \star$. More generally, any $\mathbb{F}$-witness of H is said to be an algebraic witness of H.*

We stress that every $\star$ in the stencil H must be replaced by a *nonzero* entry from $\mathbb{F}$ and cannot be zero. Of the possible algebraic witnesses for H, we will be primarily focused in this paper on the algebraic witnesses that attain the smallest feasible rank over some field $\mathbb{F}$, which leads us to the following definition.

Definition 2 (Rank)**.** *Given an $m \times n$ stencil H, the $\mathbb{F}$-rank of H is the smallest natural number r such that there exists an $\mathbb{F}$-witness $W \in \mathbb{F}^{m \times n}$ whose rank is equal to r. We denote the value r by $\mathrm{rk}_{\mathbb{F}}(H)$.*

2.2 Visible Rank

In this section, we introduce our notion of the *visible rank* of a stencil. The main motivation for introducing the visible is to be able to determine the optimal lower bound on the rank of a matrix with only the knowledge of the support of a matrix and nothing else about the values of that support.

Consider a square matrix $A \in \mathbb{F}^{n\times n}$, and suppose we are interested in determining if it is full rank. A natural approach would be to inspect its determinant. From the Leibniz formula, we know that $\det(A) = \sum_{\pi\in S_n} (-1)^{\text{sgn}(\pi)} \prod_{i=1}^{n} A_{i,\pi(i)}$, where S_n denotes the symmetric group of order n and $\text{sgn}(\pi)$ denotes the sign of a permutation π. From the Leibniz formula, notice that $\det(A)$ is a linear combination of the nonzero generalized diagonals of A. If our hope is to obtain $\det(A) \neq 0$ without any knowledge of the values of the support of A, then one way to guarantee it is to say that A has exactly one nonzero generalized diagonal. In such a case, we can guarantee that $\det(A) \neq 0$. As for when A has more than one generalized diagonal, there is no guarantee if $\det(A) \neq 0$ without inspecting the values of the support of A.

Given the previous discussion, it seems natural to define a notion of rank on stencils as follows.

Definition 3 (Visibly Full Rank)**.** *For a square stencil $M \in \{0,\star\}^{n\times n}$, we say that M is visibly full rank if M has exactly one star diagonal. That is, a generalized diagonal whose entries are all $\star$'s.*

Of course, determining whether our matrix is full rank or not is unsatisfactory for our purposes. In most cases, we are interested in determining the rank of a matrix $A \in \mathbb{F}^{m\times n}$. Thus we need to develop an extension to our notion of being visibly full rank. Observe that the rank of A can be stated as the size of the largest square submatrix in A that is full rank. From this perspective, it becomes natural to define the rank of a stencil as follows.

Definition 4 (Visible Rank)**.** *For a stencil $H \in \{0,\star\}^{m\times n}$, the visible rank of H, denoted $vrk(H)$, is the largest square sub-stencil in H that is visibly full rank.*

We extend our notion of being visibly full rank to non-square stencils to be when $\text{vrk}(H) = \min\{m,n\}$. We also say that a set of k columns in H is *visibly independent* if there exists a $k\times k$ sub-stencil within these k columns that is visibly full rank.

Evidently, not all full-rank square matrices have exactly one nonzero generalized diagonal, but all square matrices with exactly one nonzero generalized diagonal are full rank. Thus if we are interested in determining $\text{rank}(A)$ by finding the size of the largest full rank square submatrix in A, we can instead search for the largest square submatrix in A with exactly one nonzero generalized diagonal, and that yields us lower bounds on the rank of A.

Proposition 6. *Given a field $\mathbb{F}$ and stencil $H \in \{0,\star\}^{m\times n}$, we have that $rk_{\mathbb{F}}(H) \geqslant vrk(H)$.*

Proof. Let $W \in \mathbb{F}^{m\times n}$ be an algebraic witness of H. Let H_0 be a $k\times k$ square sub-stencil of H that is visibly full rank. Let W_0 be the corresponding $k\times k$ submatrix in W. Because H_0 has exactly one generalized diagonal of $\star$'s, W_0 will have exactly one generalized diagonal of nonzero entries, which implies that $\det(W_0) \neq 0$ by the Leibniz formula. This proves that $\text{rank}(W) \geqslant \text{rank}(W_0) = k$. By picking the largest such H_0, we deduce that $\text{rank}(W) \geqslant \text{vrk}(H)$. Since W was an arbitrary $\mathbb{F}$-witness, this proves $\text{rk}_{\mathbb{F}}(H) \geqslant \text{vrk}(H)$. □

2.3 Combinatorial properties of visible rank

In this subsection, we will be proving some properties about visible rank. In particular, we will show that any visibly full rank square stencil M is permutationally equivalent to an upper triangular stencil. From this observation, we can then upper bound the visible rank by the largest rectangle of zeros in the stencil. This idea will be our main tool in proving our constructions of q-LCC and t-DRGP stencils. We additionally show an upper bound on the $\mathbb{F}$-rank of a stencil by the maximum number of zeros in each row over any field $\mathbb{F}$ satisfying $|\mathbb{F}| \geqslant n$, where n is the number of columns.

Given two stencils $H_1, H_2 \in \{0, \star\}^{m\times n}$, we say that H_1 is *permutationally equivalent* to H_2 if there are permutations $\pi : [m] \to [m]$ and $\sigma : [n] \to [n]$ such that $H_1[i, j] = H_2[\pi(i), \sigma(j)]$ for all $i \in [m]$ and $j \in [n]$. For such H_1 and H_2, we write $H_2 = (H_1)_{\pi,\sigma}$ to say that H_2 is obtained from H_1 by applying the permutation π on the rows and the permutation σ on the columns. Note that the order in which we apply them does not matter as row permutations and column permutations commute with each other.

Lemma 7. *Let $M \in \{0, \star\}^{n\times n}$ be visibly full rank. Then there exist permutations π and σ on $[n]$ such that $N := M_{\pi,\sigma}$ is an upper triangular stencil. That is, $N[i, i] = \star$ and $N[i, j] = 0$ for all $i, j \in [n]$ with $i > j$.*

Proof. We first prove the following claim.

Claim 1. *Let $M \in \{0, \star\}^{n\times n}$ be visibly full rank. Then there exists a row in M with exactly one $\star$.*

Proof: Assume (for the sake of a contradiction) that such a row doesn't exist. Index the rows of M by $R = \{r_1, \dots, r_n\}$ and the columns by $C = \{c_1, \dots, c_n\}$. Let $G = (R, C, E)$ be a bipartite graph on the rows and columns of M with edges E, where E is the set of $\star$'s in M. Because M is visibly full rank, G has a unique perfect matching, which we will denote by T. Moreover, by our initial assumption, $d_G(v) \geqslant 2$ for all $v \in R$. Here, $d_G(\cdot)$ denotes the degree of a vertex in G.

Color the edges in T red and all remaining edges blue. We will show that G has a simple cycle S with alternating edge colors, which will suffice to prove our contradiction. Indeed, let R_S and B_S be the red and blue edges of S, respectively. Notice that both R_S and B_S match the same vertex sets, and since T is a perfect matching in G, then we deduce that $(T \setminus R_S) \cup B_S$ is also a perfect matching in G, which is a contradiction as T is the unique perfect matching in G.

Now, we will proceed to show the existence of an alternating cycle. Fix a vertex $v_1 \in C$ and $v_2 \in R$ such that the edge $\{v_1, v_2\}$ is red. Extend (v_1, v_2) to a maximal alternating path $P := (v_1, v_2, \dots, v_p)$. Then for all $i \in [p - 1]$, we know that the edge $\{v_i, v_{i+1}\}$ is red if $v_i \in C$ and blue if $v_i \in R$. Now, consider two cases on v_p:

(i) If $v_p \in C$, then because the red edges form a perfect matching in G, we can find a vertex $v_{p+1} \in R$ such that $\{v_p, v_{p+1}\}$ is a red edge. Since P is a maximal alternating path and $\{v_{p-1}, v_p\}$ is a blue edge, there exists an $i \in [p]$ such that $v_{p+1} = v_i$. Thus $(v_i, v_{i+1}, \dots, v_p, v_{p+1} = v_i)$ is an alternating cycle in G.

(ii) If $v_p \in R$, then because $d_G(v) \geqslant 2$ for $v \in R$ and the red edges form a perfect matching, we can find a vertex $v_{p+1} \in C$ such that $\{v_p, v_{p+1}\}$ is a blue edge. Since P is a maximal alternating path and $\{v_{p-1}, v_p\}$ is a red edge, there exists an $i \in [p]$ such that $v_{p+1} = v_i$. Thus $(v_i, v_{i+1}, \dots, v_p, v_{p+1} = v_i)$ is an alternating cycle in G.

Thus in both cases, we can find an alternating cycle. Since these are the only two possible cases, we therefore conclude the existence of an alternating cycle. ■

With Claim 1 at hand, we can now proceed to show our lemma by inducting on n. The base case $n = 1$ is immediate to see. As for the induction step, index the rows and columns of M by $[n]$. We know by Claim 1 that there exists a row i in M which has exactly a single $\star$, and suppose that $\star$ was in column j. Let M' be the $n \times n$ matrix obtained by transposing row i with row n and column

j with column n in M. Because M is visibly full rank, then so is M'. Moreover, the n'th row of M' has exactly one $\star$ at column n, meaning that any star diagonal in M' must contain $M[n,n]$. Thus the $(n-1)\times(n-1)$ stencil M'_0, which is obtained by deleting row n and column n of M', has the same number of star diagonals as M'. This implies that M'_0 is visibly full rank. By our induction hypothesis, we can permute the rows and columns of M'_0 to make it upper triangular. Given how M'_0 was constructed from M', we can therefore apply the same permutations to M' and turn it into an upper triangular stencil. Since M' is a permutation of M, we therefore conclude that M is permutationally equivalent to an upper triangular stencil. This finishes our induction step. □

Thus we can structurally characterize all visibly full rank square stencils as being permutationally equivalent to upper triangular stencils. From this, we can exploit the structure of upper triangular stencils to deduce the following upper bound on the visible rank.

Lemma 8. *Given an $m\times n$ stencil H, if there exists $a,b\in\mathbb{N}$ such that H has no $a\times b$ sub-stencil of zeros, then $vrk(H) < a+b$.*

Proof. Assume (for the sake of a contradiction) that H has an $(a+b)\times(a+b)$ sub-stencil H_0 that is visibly full rank. By Lemma 7, we know that H_0 is permutationally equivalent to an $(a+b)\times(a+b)$ upper triangular stencil. Since such an upper triangular stencil has a $a\times b$ sub-stencil of zeros (namely the last a rows and the first b columns), we arrive at a contradiction. □

Given the upper bound on the visible rank given above, one might wonder if analogous upper bounds hold for the algebraic rank of a stencil. In what follows, we show an upper bound on the algebraic rank of a stencil by the maximum number of zeros in each row.

Proposition 9. *For any $m\times n$ stencil H, if each row of H has at most d zeros, then $rk_{\mathbb{F}}(H)\leqslant d+1$ for all fields $\mathbb{F}$ satisfying $|\mathbb{F}|\geqslant n$.*

Proof. For each $i\in[m]$, let $Z_i\subseteq[n]$ be the set of columns that are zero at row i. Since $|\mathbb{F}|\geqslant n$, we can find a set of pairwise distinct entries $a_1,\ldots,a_n\in\mathbb{F}$. For each $i\in[m]$, define the polynomial $p_i(x) := \prod_{j\in Z_i}(x-a_j)$. By construction, we have that $p_i(a_j)=0$ if and only if $j\in Z_i$. This means that the matrix $E\in\mathbb{F}^{m\times n}$ defined by $E_{ij}=p_i(a_j)$ is an $\mathbb{F}$-witness of H. Moreover, since $|Z_i|\leqslant d$ for each $i\in[m]$, then we know that the monomials $\{1,x,\ldots,x^d\}$ span the polynomials $\{p_1,\ldots,p_m\}$. Since the dimension of the row space of E is at most the dimension of the span of $\{p_1,\ldots,p_m\}$, we conclude that $\text{rank}(E)\leqslant d+1$. Thus $\text{rk}_{\mathbb{F}}(H)\leqslant d+1$. □

We remark that the bound $|\mathbb{F}|\geqslant n$ is crucial for Proposition 9. Consider the stencil $D\in\{0,\star\}^{n\times n}$ that has $\star$'s everywhere except on the diagonal. Such a stencil has a visible rank of 2, but its rank over $\mathbb{F}_2$ is at least $n-1$. We will revisit this remark in Question 3 in Section 8.

2.4 A Rank-Nullity Type Theorem Between Stencils and Symmetric Spanoids

In this subsection, we formally set up spanoids and prove a rank-nullity type theorem between symmetric spanoids and stencils.

A *spanoid* $\mathcal{S}$ consists of a ground set $[n]$, which can be thought of as a collection of logical statements, and a collection of inference rules in the form of pairs (S,i), which are written as $S\to i$, for $S\subseteq[n]$ and $i\in[n]$. Naturally, given any collection $\mathcal{S}$ of logical statements and inference rules among them, one would be interested in determining the minimum number of statements needed to infer all other statements, which is defined to be the *rank* of $\mathcal{S}$. Spanoids were introduced in [DGGW20] as an abstraction of LCCs, where the ground set $[n]$ can be thought of as the codeword symbols of the LCC and the inference rules as its locality constraints. They proved that the spanoid analog of q-LCCs satisfies a rank upper bound of $\widetilde{O}(n^{(q-2)/(q-1)})$. Moreover, they also showed that there are q-LCC spanoids for which their rank is $\widetilde{\Omega}(n^{(q-2)/(q-1)})$. Thus within the proof framework of spanoids, one cannot hope to obtain a polynomial improvement on the currently known dimension upper bounds on LCCs.

Let us formally set up the definitions needed for spanoids. A derivation in $\mathcal{S}$ of $i \in [n]$ from a set $T \subseteq [n]$ is a sequence of sets $T_0 = T, T_1, \ldots, T_r$ satisfying $T_j = T_{j-1} \cup \{i_j\}$ for some $i_j \in [n]$, $j \in [r]$, and with $i_r = i$. Further, for every $j \in [r]$, there is a rule (S_{j-1}, i_j) in $\mathcal{S}$ for some $S_{j-1} \subseteq T_{j-1}$. The *span* of a set $T \subseteq [n]$, denoted $\mathsf{span}_{\mathcal{S}}(T)$, is the set of all $i \in [n]$ for which there is a derivation of i from T. The rank of a spanoid, denoted $\mathsf{rank}(\mathcal{S})$, is the size of the smallest set $T \subseteq [n]$ such that $\mathsf{span}_{\mathcal{S}}(T) = [n]$. Finally, we define symmetric spanoids below.

Definition 5 (Symmetric Spanoids)**.** *A spanoid $\mathcal{S}$ over $[n]$ is a symmetric spanoid if there is a collection of sets $\{S_1, \ldots, S_m\}$ so that the inference rules of $\mathcal{S}$ are of the form $S_j \setminus \{i\} \to \{i\}$ for any $i \in S_j$ and $j \in [m]$.*

Now we may proceed to prove our theorem that relates the rank of symmetric spanoids with the visible rank of an associated stencil.

Theorem 10. *For any symmetric spanoid $\mathcal{S}$ over $[n]$ with m sets, there exists a canonical stencil H of size $m \times n$ such that for any collection of columns $C \subseteq [n]$ in H, they are visibly independent if and only if $\mathsf{span}_{\mathcal{S}}([n] \setminus C) = [n]$. Moreover, we have $\mathrm{vrk}(H) + \mathsf{rank}(\mathcal{S}) = n$.*

Proof. Define $H[i,j] = \star$ if $j \in S_i$ and zero otherwise. We claim that the stencil H satisfies the conditions stated above. Indeed, suppose that the columns $C = \{c_1, \ldots, c_k\}$ are visibly independent. Then by definition, we can find rows $R = \{r_1, \ldots, r_k\}$ so that the $R \times C$ sub-stencil H_C is visibly full rank. By Lemma 7, we can find permutations π, σ over $[k]$ such that the stencil $H'_C := (H_C)_{\pi,\sigma}$ is an upper triangular stencil. In terms of spanoids, that means $c_{\sigma(i)} \in S_{r_{\pi(i)}}$ and $S_{r_{\pi(i)}} \subseteq [n] \setminus \{c_{\sigma(1)}, \ldots, c_{\sigma(i-1)}\}$ for all $i \in [k]$, which can be rewritten as $S_{r_{\pi(i)}} \subseteq ([n] \setminus C) \cup \{c_{\sigma(i)}, \ldots, c_{\sigma(k)}\}$ for all $i \in [k]$. Thus starting with the set $[n] \setminus C$, we can apply the inference rules $S_{r_{\pi(i)}} \setminus \{c_{\sigma(i)}\} \to c_{\sigma(i)}$ for $i = k, k-1, \ldots, 1$ to derive all of $[n]$. Thus $\mathsf{span}_{\mathcal{S}}([n] \setminus C) = [n]$.

Now, suppose that $\mathsf{span}_{\mathcal{S}}([n] \setminus C) = [n]$. Then that means we can find sets $S_{r_1}, \ldots, S_{r_k}$ and a permutation σ over $[k]$ such that $c_{\sigma(i)} \in S_{r_i}$, and starting with the set $[n] \setminus C$, we can apply the inference rules $S_{r_i} \setminus \{c_{\sigma(i)}\} \to c_{\sigma(i)}$ for $i = 1, 2, \ldots, k$ to derive all of $[n]$. For this to be possible, it must be the case that $S_{r_i} \subseteq ([n] \setminus C) \cup \{c_{\sigma(1)}, \ldots, c_{\sigma(i)}\}$ for each $i \in [k]$, which can be rewritten as $S_{r_i} \subseteq [n] \setminus \{c_{\sigma(i+1)}, \ldots, c_{\sigma(k)}\}$ for all $i \in [k]$. Now, in terms of the stencil H, this means $H[r_i, c_{\sigma(i)}] = \star$ and $H[r_i, c_{\sigma(\ell)}] = 0$ for $i < \ell$. Thus the $k \times k$ sub-stencil H' obtained by restricting to the columns $c_{\sigma(k)}, \ldots, c_{\sigma(1)}$ and rows $r_k, \ldots, r_1$ in that order forms an upper triangular stencil, which is visibly full rank. Thus the set of columns C is visibly independent.

Now, for any set $S \subseteq [n]$ such that $\mathsf{span}_{\mathcal{S}}(S) = [n]$, we know that $[n] \setminus S$ is visibly independent in H. Thus $n - |S| \leqslant \mathrm{vrk}(H)$. Since this holds for any such set S, then we deduce that $\mathsf{rank}(\mathcal{S}) + \mathrm{vrk}(H) \geqslant n$. On the other hand, for any collection of columns C in H that is visibly independent, we know that $\mathsf{span}_{\mathcal{S}}([n] \setminus C) = [n]$. This implies $n - |C| \geqslant \mathsf{rank}(\mathcal{S})$. Since this holds for any visibly independent set of columns C in H, then we find that $n \geqslant \mathsf{rank}(\mathcal{S}) + \mathrm{vrk}(H)$. Hence $\mathsf{rank}(\mathcal{S}) + \mathrm{vrk}(H) = n$. $\square$

3 Constructing q-LCC Stencils

In this section, we define and construct q-LCC stencils whose visible rank achieves known lower bounds up to polylog factors. Because q-LCCs are completely defined by the existence of $\Omega(n^2)$ linear dependencies satisfying a special combinatorial structure, we can therefore distill q-LCCs as being the kernel of a structured matrix. Such a structure can be neatly captured in the following definition.[6]

Definition 6 (q-LCC Stencils)**.** *For some fixed constant $\delta > 0$, a $\delta n^2 \times n$ stencil H whose rows and columns are indexed by $[n] \times [\delta n]$ and $[n]$, respectively, is said to be a q-LCC stencil if the following conditions hold:*

1. *$H[(i,j), i] = \star$ for all $(i,j) \in [n] \times [\delta n]$,*
2. *for all $i, k \in [n]$ with $k \neq i$, there is at most one $\star$ in $\{H[(i,1),k], \ldots, H[(i,\delta n),k]\}$,*
3. *the number of $\star$'s in each row of H is at most $q+1$.*

From the work of [DGGW20], we know that any q-LCC spanoid has rank at most $\widetilde{O}(n^{(q-2)/(q-1)})$ for $q \geqslant 3$. By Theorem 10, this implies that the visible rank of any q-LCC stencil is at least $n - \widetilde{O}(n^{(q-2)/(q-1)})$ for $q \geqslant 3$. As for upper bounds, [DGGW20] proved the existence of q-LCC spanoids with rank at least $\widetilde{\Omega}(n^{(q-2)/(q-1)})$ for $q \geqslant 3$. However, the spanoids they constructed were not necessarily symmetric, and thus Theorem 10 could not apply in this case.

Our main theorem for this section shows the existence of q-LCC stencils whose visible rank matches known lower bounds up to polylog factors. By Theorem 10, this implies the existence of q-LCC symmetric spanoids whose rank is at least $\widetilde{\Omega}(n^{(q-2)/(q-1)})$ for $q \geqslant 3$, thereby addressing a question posed by [DGGW20].

Theorem 11. *For any fixed $q \geqslant 3$, there exists a constant $\delta = \delta(q) > 0$ and a $\delta n^2 \times n$ q-LCC stencil with visible rank at most $n - \widetilde{\Omega}(n^{(q-2)/(q-1)})$.*

Proof. Let $\delta \in (0, 1/q)$ be a constant that we will choose later. For $(i,s) \in [n] \times [\delta n]$. Define $r_s^i := \{j \in [n] : H[(i,s),j] = \star\}$ to be the support of row (i,s), and let $G_i := \{r_s^i : s \in [\delta n]\}$ be the δn repair groups for column i. Considering picking the groups G_i uniformly at random by considering a bijection $\pi_i : [n-1] \to [n] \setminus \{i\}$ chosen uniformly at random and setting $r_s^i = \{i, \pi_i(q(s-1)+1), \ldots, \pi_i(q(s-1)+q)\}$ for each $s \in [\delta n]$. We shall show that the visible rank of H will at most be $n - \Omega(n^{(q-2)/(q-1)}/\log n)$ with high probability.

Consider natural numbers $\ell_0 > \ell_1$ where $\ell_0 = n^{(q-2)/(q-1)}$ and $\ell_1 = \ell_0/\log n = n^{(q-2)/(q-1)}/\log n$. Let E be the event that there is no $(\ell_0 - \ell_1) \times (n - \ell_0)$ sub-stencil in H whose entries are all zeros. It suffices to show that E occurs with high probability, as Lemma 8 would then imply an upper bound of $n - \ell_1$ on the visible rank of H.

Notice that the event E is equivalent to the event that there are no $\ell_0 - \ell_1$ rows in H whose joint support is of size at most ℓ_0. Now, consider any two collections of ℓ_0 columns $C \subseteq [n]$ and $\ell_0 - \ell_1$ rows $R \subseteq [n] \times [\delta n]$. For $i \in [n]$, define $R_i := \{s \in [\delta n] \mid (i,s) \in R\}$. Without loss of generality, write $R_i = [a_i]$. Then we see that $\sum_{i=1}^n a_i = |R| = \ell_0 - \ell_1$. Furthermore, if the supports of the rows at R were contained in C, then for any $(i,s) \in R$, we must have that $i \in C$ as $i \in r_s^i$. Under such a

[6] Throughout this section, one can appropriately add floors and ceilings to resolve all integrality concerns, so we will not worry about such issues.

case, by using the chain rule, the pairwise disjointedness of the sets $\{r_s^i \setminus \{i\}\}_{s=1}^{a_i}$, and the definition of the G_i's, we find that

$$\begin{aligned}
\mathbf{Pr}\left[\bigwedge_{(i,s)\in R} (r_s^i \subseteq C)\right] &= \prod_{i=1}^{n} \mathbf{Pr}\left[\bigwedge_{s=1}^{a_i} (r_s^i \subseteq C)\right] \\
&= \prod_{i=1}^{n}\prod_{s=1}^{a_i} \mathbf{Pr}\left[r_s^i \subseteq C \;\middle|\; r_k^i \subseteq C \text{ for all } k \in [s-1]\right] \\
&= \prod_{i=1}^{n}\prod_{s=1}^{a_i} \mathbf{Pr}\left[r_s^i \setminus \{i\} \subseteq C \setminus \{i\} \;\middle|\; r_k^i \setminus \{i\} \subseteq C \setminus \{i\} \text{ for all } k \in [s-1]\right] \\
&= \prod_{i=1}^{n}\prod_{s=1}^{a_i} \mathbf{Pr}\left[r_s^i \setminus \{i\} \subseteq (C \setminus \{i\}) \setminus \cup_{k=1}^{s-1}(r_k^i \setminus \{i\}) \;\middle|\; r_k^i \setminus \{i\} \subseteq C \setminus \{i\} \text{ for all } k \in [s-1]\right] \\
&= \prod_{i=1}^{n}\prod_{s=1}^{a_i} \frac{\binom{\ell_0-(s-1)q-1}{q}}{\binom{n-(s-1)q-1}{q}} \leqslant \prod_{i=1}^{n}\prod_{s=1}^{a_i} \frac{\binom{\ell_0-1}{q}}{\binom{n-1}{q}} = \left(\frac{\binom{\ell_0-1}{q}}{\binom{n-1}{q}}\right)^{\ell_0-\ell_1} \leqslant \left(\frac{\ell_0}{n}\right)^{q(\ell_0-\ell_1)} .
\end{aligned}$$

Therefore, by applying a union bound over all collections of columns $C \subseteq [n]$ of size ℓ_0 and collections of rows $R \subseteq [\delta n] \times [n]$ of size $\ell_0 - \ell_1$ satisfying $i \in C$ for all $(i, s) \in R$ and using the bound $\binom{a}{b} \leqslant \left(\frac{ea}{b}\right)^b$, we deduce that

$$\begin{aligned}
\mathbf{Pr}[E] \leqslant \binom{n}{\ell_0}\binom{\delta n \ell_0}{\ell_0-\ell_1}\left(\frac{\ell_0}{n}\right)^{q(\ell_0-\ell_1)} &\leqslant \left(\frac{en}{\ell_0}\right)^{\ell_0}\left(\frac{e\delta n\ell_0}{\ell_0-\ell_1}\right)^{\ell_0-\ell_1}\left(\frac{\ell_0}{n}\right)^{q(\ell_0-\ell_1)} \\
&= \left(\frac{en}{\ell_0}\right)^{\ell_1}\left(\frac{en}{\ell_0}\cdot\frac{e\delta n\ell_0}{\ell_0-\ell_1}\cdot\frac{\ell_0^q}{n^q}\right)^{\ell_0-\ell_1} \\
&= \left(\frac{en}{\ell_0}\right)^{\ell_1}\left(\frac{e^2\delta\ell_0^{q-1}}{\left(1-\frac{\ell_1}{\ell_0}\right)n^{q-2}}\right)^{\ell_0-\ell_1} \\
&= \left(en^{1/(q-1)}\right)^{n^{\frac{q-2}{q-1}}\cdot\frac{1}{\log n}}\left(\frac{e^2\delta}{\left(1-\frac{1}{\log n}\right)}\right)^{n^{\frac{q-2}{q-1}}\left(1-\frac{1}{\log n}\right)} \\
&= \exp\left(\left(\frac{1}{q-1}+2+\log(\delta)+o(1)\right)n^{\frac{q-2}{q-1}}\right) .
\end{aligned}$$

Thus for $\delta = \min\{e^{-3}, 1/(2q)\}$, the quantity above becomes $\exp\left(-\Omega(n^{(q-2)/(q-1)})\right)$. Therefore, for sufficiently large n, there exists a q-LCC stencil H such that any $\ell_0 - \ell_1$ rows of it have joint support of size more than ℓ_0. This is equivalent to having no $(\ell_0 - \ell_1) \times (n - \ell_0)$ sub-stencil whose entries are all zeros. By Lemma 8, we therefore conclude that $\mathrm{vrk}(H) < n - \ell_1 = n - \Omega(n^{(q-2)/(q-1)}/\log n)$. $\square$

4 t-DRGP Stencils

In this section, we define and construct q-DRGP stencils whose visible rank is exponentially smaller than known rank lower bounds for t-DRGP codes. Because t-DRGP codes are completely defined by the existence of tn linear dependencies satisfying a specific combinatorial structure, we can therefore distill t-DRGP codes as being the kernel of a structured matrix. Such a structure can be neatly captured in the following class of stencils.

Definition 7 (t-DRGP Stencils)**.** *A $tn \times n$ stencil H whose rows and columns are indexed by $[n] \times [t]$ and $[n]$, respectively, is said to be a t-DRGP stencil if the following conditions hold:*

1. *$H[(i,j),i] = \star$ for all $(i,j) \in [n] \times [t]$,*
2. *for all $i,k \in [n]$ with $k \neq i$, there is at most one $\star$ in $\{H[(i,1),k], \ldots, H[(i,t),k]\}$.*

By the works of [RV16, Woo16], we deduce that $\mathrm{rk}_{\mathbb{F}}(H) \geqslant 2\sqrt{n} - O(1)$ for all fields $\mathbb{F}$ and t-DRGP stencils H with $t \geqslant 2$. Naturally, given the notion of visible rank, one might wonder if the rank lower bounds of [RV16, Woo16] can be captured as a lower bound on the visible rank. Our main theorem for this section shows that visible rank, in its simplest version, cannot recover such a rank lower bound for $t = O(1)$. In fact, it can at best yield an exponentially smaller lower bound than the currently known lower bound.

Theorem 12. *For any fixed natural number $t \geqslant 2$, there exists a $tn \times n$ t-DRGP stencil with visible rank at most $O(t \log n)$.*

Proof. Consider a random $tn \times n$ t-DRGP stencil H as follows. For $i,j \in [n]$, define the set of entries $S_{i,j} := \{((i,s),j) \mid s \in [t]\}$. For each $i,j \in [n]$ with $i \neq j$, set all of the entries $S_{i,i}$ to be $\star$'s, and uniformly sample an entry from $S_{i,j}$ to be a $\star$, while everything else in $S_{i,j}$ is set to zero. We will show that $\mathrm{vrk}(H) \leqslant O(t \log n)$ occurs with high probability.

Fix a natural number ℓ that we will choose later. In order for us to show that $\mathrm{vrk}(H) < 2\ell$, it suffices by Lemma 8 to show that H has no $\ell \times \ell$ sub-stencil whose entries are all zeros. Now, consider any collection of rows $R \subseteq [n] \times [t]$ and columns $C \subseteq [n]$ such that $|R| = |C| = \ell$. Let H_0 be the $R \times C$ sub-stencil in H. By the randomness structure of H and the inequality $1 - z \leqslant e^{-z}$ for $z \in \mathbb{R}$, we find that

$$\begin{aligned}
\mathbf{Pr}\left[H_0 \text{ is all-zeroes}\right] &= \prod_{i,j \in [n]} \mathbf{Pr}\left[H[(i,s),j)] = 0 \text{ for all } ((i,s),j) \in S_{i,j} \cap (R \times C)\right] \\
&= \prod_{i,j \in [n]} \left(1 - \frac{\left|S_{i,j} \cap (R \times C)\right|}{t}\right) \\
&\leqslant \prod_{i,j \in [n]} \exp\left(-\frac{\left|S_{i,j} \cap (R \times C)\right|}{t}\right) \\
&= \exp\left(-\frac{1}{t} \sum_{i,j \in [n]} \left|S_{i,j} \cap (R \times C)\right|\right) \\
&= \exp\left(-\frac{\ell^2}{t}\right) .
\end{aligned}$$

Thus by applying a union bound over all possible choices of rows $R \subseteq [n] \times [t]$ and columns $C \subseteq [n]$, we find that

$$\mathbf{Pr}[\mathrm{vrk}(H) \geqslant 2\ell] \leqslant \binom{tn}{\ell}\binom{n}{\ell}\exp\left(-\frac{\ell^2}{t}\right) \leqslant \left(tn^2 e^{-\frac{\ell}{t}}\right)^{\ell} .$$

Picking $\ell = 4t\log n$, we conclude that $\mathbf{Pr}[\mathrm{vrk}(H) \geqslant 2\ell] \leqslant \exp(-\Omega(t(\log n)^2))$. Thus by Lemma 8, for large enough n, there exists a t-DRGP stencil H with visible rank is less than $2\ell = O(t\log n)$. □

5 Tensor Products

In this section, we introduce a tensor product operation on stencils and then proceed to explore its properties. Given the natural tensor product $A \otimes B$ for matrices A and B, notice that the support of $A \otimes B$ is determined completely by the support of the matrices A and B. As a consequence of this observation, we are able to define the stencil of $A \otimes B$ based solely on the stencils of A and B. This leads us to our definition of a tensor product between stencils.

Definition 8 (Tensor product)**.** *Given an $A_1 \times B_1$ stencil H_1 and an $A_2 \times B_2$ stencil H_2, let $H_1 \otimes H_2$ be an $(A_1 \times A_2) \times (B_1 \times B_2)$ stencil defined by*

$$(H_1 \otimes H_2)[(a_1, a_2), (b_1, b_2)] = \begin{cases} \star & \text{if } H_1[a_1, b_1] \text{ and } H_2[a_2, b_2] \text{ both equal } \star, \\ 0 & \text{if at least one of } H_1[a_1, b_1] \text{ and } H_2[a_2, b_2] \text{ equals } 0. \end{cases}$$

We remark that our tensor product follows similar properties as the natural tensor product for matrices, such as associativity and non-commutativity.

5.1 Algebraic witnesses of tensor products

In this subsection, we prove that the tensor product of any algebraic witnesses of the stencils H_1 and H_2 is an algebraic witness of $H_1 \otimes H_2$. This therefore shows us that the $\mathbb{F}$-rank is a *sub-multiplicative* function with respect to the tensor product.

Proposition 13. *Let M and N be matrices over a field $\mathbb{F}$ who are $\mathbb{F}$-witnesses to stencils H_1, H_2, respectively. Then $M \otimes N$ is an $\mathbb{F}$-witness of $H_1 \otimes H_2$.*

Proof. For every entry in $H_1 \otimes H_2$, we know that $(H_1 \otimes H_2)[(i_1, i_2), (j_1, j_2)]$ is a $\star$ if and only if $H_1[i_1, j_1]$ and $H_2[i_2, j_2]$ are both $\star$'s. This holds if and only if $M_{i_1 j_1}$ and $N_{i_2 j_2}$ are both nonzero. Because $(M \otimes N)_{(i_1,i_2),(j_1,j_2)} = M_{i_1 j_1} N_{i_2 j_2}$, then the entry $(M \otimes N)(i_1, i_2), (j_1, j_2)$ is nonzero if and only if $M_{i_1 j_1}$ and $N_{i_2 j_2}$ are both nonzero. Thus $(H_1 \otimes H_2)[(i_1, i_2), (j_1, j_2)]$ is a $\star$ if and only if $(M \otimes N)_{(i_1,i_2),(j_1,j_2)}$ is nonzero. Hence we conclude that $M \otimes N$ is an $\mathbb{F}$-witness of $H_1 \otimes H_2$. □

By applying Proposition 13 on the $\mathbb{F}$-witnesses of H_1 and H_2 with the smallest ranks, we deduce the following corollary.

Corollary 14. *For any field $\mathbb{F}$, we have that $rk_{\mathbb{F}}(H_1) rk_{\mathbb{F}}(H_2) \geqslant rk_{\mathbb{F}}(H_1 \otimes H_2)$.*

5.2 Visible rank and tensor products

In this subsection, we show that the tensor product of two visibly full rank stencils is also visibly full rank. This therefore shows that the visible rank is a *super-multiplicative* function with respect to the tensor product. We additionally prove an upper bound on the visible rank of the tensor product with respect to the visible rank of one of the stencils.

Proposition 15. *Given visibly full rank square stencils A and B, their tensor $A \otimes B$ is also visibly full rank.*

Proof. Let m and n be the orders of A and B, respectively. Then by Lemma 7, we know that there are permutations π_A and σ_A over $[m]$ and permutations π_B and σ_B over $[n]$ such that the stencils $A_0 = (A)_{\pi_A,\sigma_A}$ and $B_0 = (B)_{\pi_B,\sigma_B}$ are both upper triangular stencils. Moreover, we see that $A_0 \otimes B_0 = (A \otimes B)_{(\pi_A,\pi_B),(\sigma_A,\sigma_B)}$. Therefore, it suffices to show that $A_0 \otimes B_0$ is an upper triangular stencil.

Consider the lexicographical ordering on $[m] \times [n]$. When $(i_1, i_2) > (j_1, j_2)$, we know that one of the inequalities $i_1 > j_1$ and $i_2 > j_2$ must hold, which means that one of $A_0[i_1, j_1]$ or $B_0[i_2, j_2]$ must be zero. This proves that $(A_0 \otimes B_0)[(i_1, i_2), (j_1, j_2)] = 0$ whenever $(i_1, i_2) > (j_1, j_2)$. As for when $(i_1, i_2) = (j_1, j_2)$, we immediately know that $(A_0 \otimes B_0)[(i_1, i_2), (i_1, i_2)] = \star$ as $A_0[i_1, i_1] = \star$ and $B_0[i_2, i_2] = \star$. Hence $A_0 \otimes B_0$ is an upper triangular stencil with respect to the lexicographical ordering. □

Given any two stencils H_1 and H_2, we deduce from Proposition 15 that the tensor product of any two visibly full rank square sub-stencils of H_1 and H_2 will also be a visibly full rank square sub-stencil of $H_1 \otimes H_2$. By considering the largest visibly full rank square sub-stencils in H_1 and H_2, we therefore deduce the following corollary.

Corollary 16. *For stencils H_1 and H_2, We have the inequality $vrk(H_1 \otimes H_2) \geqslant vrk(H_1)vrk(H_2)$.*

Lastly, we end this subsection with an upper bound on the visible rank of $H_1 \otimes H_2$ in terms of the visible rank of H_1 and the dimensions of H_2.

Proposition 17. *Given any two stencils H_1 and H_2 of dimensions $m_1 \times n_1$ and $m_2 \times n_2$, respectively, we have that $vrk(H_1 \otimes H_2) \leqslant vrk(H_1)n_2$.*

Proof. Consider a visibly full rank sub-stencil M in $H_1 \otimes H_2$ of size $k \times k$. By Lemma 7, we can find permutations π and σ on the rows and columns of M, respectively, so that the stencil $M_0 := M_{\pi,\sigma}$ is an upper triangular stencil. Let the indices of the rows of M_0 from top to bottom be $(a_1, b_1), \ldots, (a_k, b_k)$, and let the indices of the columns of M_0 from left to right be $(c_1, d_1), \ldots, (c_k, d_k)$. For $d \in [n_2]$, define $I_d := \{i \in [k] \ : \ d_i = d\}$. Define $d_{max} := \arg\max_{d\in[n_2]}\{|I_d|\}$, and let $s := |I_{d_{max}}|$. Then from these definitions, it follows that $s \geqslant k/n_2$.

Now, write $I_{d_{max}} = \{i_1 < \ldots < i_s\}$, and consider the $s \times s$ sub-stencil M_1 in M_0 with rows $(a_{i_1}, b_{i_1}), \ldots, (a_{i_s}, b_{i_s})$ and columns $(c_{i_1}, d_{i_1}), \ldots, (c_{i_s}, d_{i_s})$. Since M_0 is an upper triangular stencil, then so is M_1. Moreover, since $d_{max} = d_{i_1} = \ldots = d_{i_s}$, then the indices of the columns of M_1 are $(c_{i_1}, d_{max}), \ldots, (c_{i_s}, d_{max})$. We claim that the $s \times s$ sub-stencil H_1' in H_1 formed by rows $a_{i_1}, \ldots, a_{i_s}$ and columns $c_{i_1}, \ldots, c_{i_s}$ is an upper triangular stencil, which will imply that $\mathrm{vrk}(H_1) \geqslant s \geqslant k/n_2$. Because M is an arbitrary visibly full rank square sub-stencil of $H_1 \otimes H_2$, we deduce that $\mathrm{vrk}(H_1) \geqslant \mathrm{vrk}(H_1 \otimes H_2)/n_2$. Thus it suffices for us to show that the stencil H_1' is upper triangular.

Since M_1 is an upper triangular stencil, we know that $M_1[(a_i, b_i), (c_i, d_{max})] = \star$, which is equivalent to $H_1[a_i, c_i] = H_2[b_i, d_{max}] = \star$ for all $i \in [k]$. Moreover, for any $i, j \in [k]$ with $i > j$, we know that $M_1[(a_i, b_i), (c_j, d_{max})] = 0$, which is equivalent to one of the entries $H_1[a_j, c_i]$ and $H_2[b_i, d_{max}]$ being zero. However, since $H_2[b_i, d_{max}] = \star$, then we deduce that $H_1[a_i, c_j] = 0$ for all $i, j \in [k]$ with $i > j$. Thus we conclude that H'_1 is an upper triangular stencil. □

5.3 Visible rank of the tensor powers

From Corollary 14, we found that the algebraic rank of stencils is sub-multiplicative with respect to tensoring. On the other hand, Corollary 16 tells us that the visible rank of stencils is super-multiplicative with respect to tensoring. Since the visible rank lower bounds the algebraic rank (Proposition 6), this therefore opens up the possibility of obtaining better lower bounds on the algebraic rank by analyzing the visible rank of the tensor powers of a stencil, which is what we will formally set up in this subsection. First, we define the tensor power of a stencil.

Definition 9 (Tensor power)**.** *Given an $m \times n$ stencil H, the k'th tensor of H is the $m^k \times n^k$ stencil $H^{\otimes k}$ defined as*

$$H^{\otimes k} := \underbrace{H \otimes H \otimes \ldots \otimes H}_{k \text{ times}}$$

By combining Corollary 14, Proposition 6, Corollary 16, and Proposition 17, we obtain the following corollary.

Corollary 18. *Given any field $\mathbb{F}$ and $k \in \mathbb{N}$, and an $m \times n$ stencil H, we have that*

$$rk_{\mathbb{F}}(H) \geqslant rk_{\mathbb{F}}(H^{\otimes k})^{1/k} \geqslant vrk(H^{\otimes k})^{1/k} \geqslant vrk(H)$$

Moreover, we also have the inequality $vrk(H^{\otimes k}) \leqslant n^{k-1} vrk(H)$.

From Corollary 18, we see that the visible rank of the higher tensor powers of H can attain better lower bounds on the $\mathbb{F}$-rank of H. More specifically, $\mathrm{rk}_{\mathbb{F}}(H)$ is lower bounded by the sequence $\{\mathrm{rk}_{\mathbb{F}}(H^{\otimes k})^{1/k}\}_{k=1}^{\infty}$. Naturally, we could want to formalize the largest lower bound on $\mathrm{rk}_{\mathbb{F}}(H)$ that the sequence $\{\mathrm{rk}_{\mathbb{F}}(H^{\otimes k})^{1/k}\}_{k=1}^{\infty}$ can attain, which we define below.

Definition 10 (Visible Capacity)**.** *The visible capacity of a stencil H, denoted as $\Upsilon(H)$, is defined as $\Upsilon(H) := \sup_{k \in \mathbb{N}} vrk(H^{\otimes k})^{1/k}$.*

By Corollary 18, we deduce that $\mathrm{rk}_{\mathbb{F}}(H) \geqslant \Upsilon(H)$ over any field $\mathbb{F}$. It is not known to us if there are stencils for which there is a gap between its visible capacity and all its $\mathbb{F}$-ranks. We defer the discussion of this point to Question 1 in Section 8.

6 Tensor Powers of Stencils for 2-DRGP Codes and q-LCCs

In this section, we will apply the tensor product tools we developed in the previous section to analyze the visible rank of the tensor powers of 2-DRGP and q-LCC stencils in hopes of obtaining better lower bounds on the algebraic rank through Corollary 18.

6.1 2-DRGP stencils

In this subsection, we prove that the second tensor power of an arbitrary $2n \times n$ 2-DRGP stencil is at least n. This proof follows the ones given in [RV16, Woo16] but is rewritten in terms of stencils. While both proofs show that $\mathrm{rk}_{\mathbb{F}}(H) \geqslant \sqrt{2n} - O(1)$, we will instead prove that $\mathrm{rk}_{\mathbb{F}}(H) \geqslant \sqrt{n}$ by showing that $\mathrm{vrk}(H \otimes H) \geqslant n$ and then applying Corollary 18.

Theorem 19. *For any 2-DRGP stencil H, we have $vrk(H \otimes H) \geqslant n$.*

Proof. Consider the $n \times n$ sub-stencil D in $H \otimes H$ whose rows are $\{((i,1),(i,2))\}_{i=1}^{n}$ and whose columns are $\{(i,i)\}_{i=1}^{n}$. We claim that D has $\star$'s along the diagonal and zero everywhere else, which immediately implies that it is visibly full rank. Indeed, the entry $D[((i,1),(i,2)),(j,j)]$ equals $\star$ if and only if $H[(i,1),j]$ and $H[(i,2),j]$ are both $\star$'s. Since H is a 2-DRGP stencil, this happens precisely when $i = j$. Thus D is a diagonal stencil and therefore visibly full rank. ◻

Thus by Corollary 18, we obtain a lower bound $\mathrm{rk}_{\mathbb{F}}(H) \geqslant \sqrt{n}$ for any field $\mathbb{F}$. On the other hand, the best-known lower bounds [RV16, Woo16] obtain $\mathrm{rk}_{\mathbb{F}}(H) \geqslant \sqrt{2n} - O(1)$, which is better by a multiplicative factor of $\sqrt{2}$. To see how we can gain this improvement, we will translate our proof of Theorem 19 to linear-algebraic terms.

Given a field $\mathbb{F}$, suppose that we have an $\mathbb{F}$-witness A of the 2-DRGP stencil H whose rank is r. Decompose $A = MN$ where M is a $2n \times r$ matrix and N is a $r \times n$ matrix. Denote the i'th column of N by w_i. In the proof of Theorem 19, we showed that the columns $\{(i,i)\}_{i=1}^{n}$ in $H \otimes H$ are visibly independent. In linear-algebraic terms, this implies that the tensors $\{w_i \otimes w_i\}_{i=1}^{n}$ are linearly independent. Because these vectors are in $\mathbb{F}^r \otimes \mathbb{F}^r$, which is of dimension r^2, we deduce the inequality $r^2 \geqslant n$, which is the same lower bound as in Theorem 19. Now, a more careful inspection would show that the tensors $\{w_i \otimes w_i\}_{i=1}^{n}$ are in fact contained in the space of *symmetric* tensors, the space of vectors that are invariant under transposing the first and second entries. The space of symmetric tensors has dimension $\binom{r+1}{2}$, and as a consequence, we obtain the sharper inequality $\binom{r+1}{2} \geqslant n$, which gives us $r \geqslant \sqrt{2n} - O(1)$.

6.2 q-LCC stencils

In this subsection, we show that for $k \leqslant n^{o(1)}$, the visible rank of the k'th tensor power of a q-LCC stencil suffers the same upper bound of $n - \widetilde{\Omega}(n^{(q-2)/(q-1)})$ as in Theorem 11. More generally, we show that the visible rank of small tensor powers is not significantly bigger than the visible rank for the regime of high-visible-rank stencils, which we formally state and prove below.

Proposition 20. *Let H be an $m \times n$ stencil whose visible rank is at most $n - s$. For any natural number k, we have that $vrk(H^{\otimes k})^{1/k} \leqslant n - \frac{s}{k}$.*

Proof. By Corollary 18 and the inequality $(1-x)^{1/k} \leqslant 1 - \frac{x}{k}$ for $x \geqslant 0$, we have that

$$\mathrm{vrk}(H^{\otimes k})^{1/k} \leqslant \left(n^{k-1}\mathrm{vrk}(H)\right)^{1/k} \leqslant \left(n^{k-1}(n-s)\right)^{1/k} = n\left(1 - \frac{s}{n}\right)^{1/k} \leqslant n\left(1 - \frac{s}{kn}\right) = n - \frac{s}{k}\,. \qquad ◻$$

By directly applying Proposition 20 to the q-LCC stencil found in Theorem 11, we deduce the following corollary.

Corollary 21. *For $q \geqslant 3$, let H be a q-LCC stencil whose visible rank is at most $n - \widetilde{\Omega}(n^{(q-2)/(q-1)})$. For any natural number k, we have that $vrk(H^{\otimes k})^{1/k} \leqslant n - \widetilde{\Omega}(n^{(q-2)/(q-1)})/k$.*

7 Exponential Gaps between Tensor Powers

In this section, we prove Theorem 5, which shows that any fixed tensor power of a stencil can yield an exponentially improved lower bound on its algebraic rank. Such exponential gaps have previously been shown in [AL06] in the context of Shannon capacities of graphs, wherein they proved that the independence number of any fixed graph power can be exponentially larger than the independence numbers of the smaller graph powers. By following the same proof method given there, we are able to attain the same exponential gaps for the visible ranks of the tensor powers.

Before proceeding with the proof, we will first setup some necessary definitions and notation for it. For any tuple of indices $x \in X^k$ and $a \in [k]$, let $p_a(x) \in X$ be the projection of x onto its a'th coordinate. Define $S(x) := \{p_1(x), \dots, p_k(x)\}$. Let $x^{(-a)} \in X^{k-1}$ denote the punctured vector $(p_1(x), \dots, p_{a-1}(x), p_{a+1}(x), \dots, p_k(x))$. Given a $X^k \times Y^k$ stencil $H^{\otimes k}$ and subsets $R \subseteq X^k$ and $C \subseteq Y^k$ with $|R| = |C|$, we say that the $R \times C$ sub-stencil H_d is *distinctively full rank* if H_d is visibly full rank and $S(r_1) \cap S(r_2) = S(c_1) \cap S(c_2) = \emptyset$ for any distinct $r_1, r_2 \in R$ and distinct $c_1, c_2 \in C$. We define the *distinctive rank* of $H^{\otimes k}$, denoted $\mathrm{drk}(H)$, to be the size of the largest distinctively full rank sub-stencil in $H^{\otimes k}$.

We introduce this intermediate notion of distinctive rank as a method to decouple the random events appearing in our forthcoming analysis in Theorem 23. In preparation for that proof, we first prove the following lemma, which will help reduce the analysis of the visible ranks in Theorem 23 to their respective distinctive ranks.

Lemma 22. *For any stencil H and natural number $k \geqslant 2$, we have $vrk(H^{\otimes k}) \leqslant 2k^2 vrk(H^{\otimes (k-1)}) drk(H^{\otimes k})$.*

Proof. Index the rows and columns of H by the sets X and Y, respectively. Consider any visibly full rank sub-stencil H_0 in $H^{\otimes k}$ of size $\ell \times \ell$. By Lemma 7, we can assume without loss of generality that H_0 is an upper triangular stencil. Let the indices of the rows of H_0 from top to bottom be $r_1, \dots, r_\ell$ for $r_i \in X^k$, and let the indices of the columns of H_0 from left to right be $c_1, \dots, c_\ell$ for $c_i \in Y^k$. Through a greedy process, we are going to select a distinctively full rank $\ell_d \times \ell_d$ sub-stencil H_d from H_0 such that $\ell_d \geqslant \ell/(2k^2 \mathrm{vrk}(H^{\otimes(k-1)}))$. Since H_0 was an arbitrary visibly full rank sub-stencil of $H^{\otimes k}$, this proves our lemma.

We now proceed with the construction of H_d. Start with the sets $I_d = \emptyset$ and $I = [\ell]$, and repeat the following process: pick an arbitrary index $i \in I$, add i to I_d, and remove all indices $j \in I$ from I for which we either have $S(r_i) \cap S(r_j) \neq \emptyset$ or $S(c_i) \cap S(c_j) \neq \emptyset$. Since $|I|$ will be strictly decreasing in this process, then this process will terminate. Once it ends, the set I_d will satisfy the property that $S(r_i) \cap S(r_j) = \emptyset$ and $S(c_i) \cap S(c_j) = \emptyset$ for any distinct indices $i, j \in I_d$. Now, let H_d be the sub-stencil in H_0 formed by the rows $\{r_i \mid i \in I_d\}$ and the columns $\{c_i \mid i \in I_d\}$. Since H_d is a principal sub-stencil of H_0 (i.e. it picks the corresponding column for each row and vice versa), then H_d will be an upper triangular stencil and thus visibly full rank. Furthermore, given how I_d was constructed, H_d will moreover be distinctively full rank. To finish this proof, it therefore suffices for us to show that $|I_d| \geqslant \ell/(2k^2 \mathrm{vrk}(H^{\otimes(k-1)}))$.

Next, we will show that every step of the process from the preceding paragraph deletes at most $2k^2 \mathrm{vrk}(H^{\otimes(k-1)})$ indices from I, from which then we conclude that $|I_d| \geqslant \ell/(2k^2 \mathrm{vrk}(H^{\otimes(k-1)}))$. The key observation in this analysis is that any visibly full rank square sub-stencil in $H^{\otimes k}$ for which its row indices all agree on some fixed coordinate has a corresponding visibly full rank square sub-stencil in $H^{\otimes(k-1)}$ of the same order. More specifically, for any $a, b \in [k]$ and $s \in [\ell]$, consider the collection of indices $I_{a,b,s} := \{i \in [\ell] \mid p_a(r_i) = p_b(r_s)\}$. We claim that $|I_{a,b,s}| \leqslant \mathrm{vrk}(H^{\otimes(k-1)})$. Assuming this claim, we deduce that the constraint $S(r_i) \cap S(r_j) = \varnothing$ will delete at most $k^2\mathrm{vrk}(H^{\otimes(k-1)})$ pairs. Similarly for c_i, the constraint $S(c_i) \cap S(c_j) = \varnothing$ will also delete at most $k^2\mathrm{vrk}(H^{\otimes(k-1)})$ pairs, totaling to at most $2k^2\mathrm{vrk}(H^{\otimes(k-1)})$ deletions per iteration of the process.

Thus it remains for us to show that $|I_{a,b,s}| \leqslant \mathrm{vrk}(H^{\otimes(k-1)})$. Indeed, let $H_0^{a,b,s}$ be the sub-stencil in H_0 whose rows are $\{r_i \mid i \in I_{a,b,s}\}$ and whose columns are $\{c_i \mid i \in I_{a,b,s}\}$. Since H_0 is an upper triangular stencil and $H_0^{a,b,s}$ is a principal sub-stencil of H_0, then $H_0^{a,b,s}$ is also an upper triangular stencil. Now, because H_0 is an upper triangular stencil, then we know that $H^{\otimes k}[r_i, c_i] = H_0[r_i, c_i] = \star$ for all $i \in I_{a,b,s}$. By definition of $H^{\otimes k}$, we deduce that $H[p_a(r_i), p_a(c_i)] = \star$ for all $i \in I_{a,b,s}$. Now, by definition of $I_{a,b,s}$, this implies that $H[p_a(r_j), p_a(c_i)] = H[p_b(r_s), p_a(c_i)] = H[p_a(r_i), p_a(c_i)] = \star$ for all $i, j \in I_{a,b,s}$. By definition of the tensor product, we deduce that $H^{\otimes(k-1)}[r_j^{(-a)}, c_i^{(-a)}] = H_0[r_j, c_i]$ for all $i, j \in I_{a,b,s}$. Since $H_0^{a,b,s}$ is an upper triangular stencil, then this implies that the sub-stencil in $H^{\otimes(k-1)}$ whose rows are $\{r_i^{(-a)} \mid i \in I_{a,b,s}\}$ and whose columns are $\{c_i^{(-a)} \mid i \in I_{a,b,s}\}$ is an upper triangular stencil. Moreover, this sub-stencil is of the order $|I_{a,b,s}|$. Since it is an upper triangular stencil, then it is visibly full rank, yielding us the bound $|I_{a,b,s}| \leqslant \mathrm{vrk}(H^{\otimes(k-1)})$ for any $a, b \in [k]$ and $i \in [\ell]$. This finishes the proof. ◻

Now, we proceed to prove our main theorem for this section.

Theorem 23. *For any fixed natural number $t \geqslant 2$, there exists a $tn \times n$ stencil H such that vrk$(H^{\otimes k}) = O_t((\log n)^k)$ for any $k = 1, \ldots, t-1$ and vrk$(H^{\otimes t}) \geqslant n$.*

Proof. We proceed to prove this theorem by showing that a random stencil with a specified structure satisfies those properties with high probability. Index the rows by $[n] \times [t]$ and the columns by $[n]$. For $i, j \in [n]$, define the set of entries $S_{i,j} := \{(i, s), j) \mid s \in [t]\}$. For each $i, j \in [n]$ with $i \neq j$, set all of the entries $S_{i,i}$ to be $\star$'s, and uniformly sample an entry from $S_{i,j}$ to be a zero, while everything else in $S_{i,j}$ is set to a $\star$. We are going to show that such a random stencil H satisfies the visible rank bounds in the theorem statement with high probability.

First, let us show that $\mathrm{vrk}(H^{\otimes t}) \geqslant n$ for any such stencil H. Consider the $n \times n$ sub-stencil H_0 in $H^{\otimes t}$ whose rows are $\{((i, 1), \ldots, (i, t)) \in ([n] \times [t])^t \mid i \in [n]\}$ and whose columns are $\{(i, \ldots, i) \in [n]^t \mid i \in [n]\}$. For any row $x_i := ((i, 1), \ldots, (i, t)) \in ([n] \times [t])^t$ and column $y_j := (j, \ldots, j) \in [n]^t$, we know that $H_0[x_i, y_j] = \star$ if and only if $H[(i, a), j] = \star$ for all $a \in [t]$. By construction of H, that is true if and only if $i = j$. Thus H_0 is a star diagonal and therefore visibly full rank. Hence $\mathrm{vrk}(H^{\otimes t}) \geqslant n$.

Now, as for the smaller tensor powers, we claim that $\mathrm{drk}(H^{\otimes k}) < 8k^2t^2 \log n$ holds with high probability for all $k = 1, \ldots, t-1$. Assuming this claim, the theorem then follows. Indeed, from the definition of $\mathrm{drk}(\cdot)$, we see that $\mathrm{vrk}(H) = \mathrm{drk}(H)$. Thus by inductively applying Lemma 22 on $\mathrm{vrk}(H^{\otimes k})$, we deduce that

$$\mathrm{vrk}(H^{\otimes k}) \leqslant 2^{k-1}(k!)^2 \prod_{a=1}^{k} \mathrm{drk}(H^{\otimes a}) < 2^{k-1}(k!)^2 \prod_{a=1}^{k} (8a^2t^2 \log n) = 2^{4k-1}(k!)^4 t^{2k} (\log n)^k$$

holds with high probability for all $k = 1, \ldots, t-1$. Thus it suffices for us to show that $\mathrm{drk}(H^{\otimes k}) < 8k^2t^2 \log n$ holds with high probability for all $k = 1, \ldots, t-1$.

Now, fix a natural number $k \leqslant t-1$. We will bound $\mathrm{drk}(H^{\otimes k})$ using a distinctive rank analog of Lemma 8. Fix $\ell_d \in \mathbb{N}$ that we will choose later. For any distinctively full rank $2\ell_d \times 2\ell_d$ sub-stencil

$H_{0,d}$ of $H^{\otimes k}$, we know by Lemma 7 that there is an ordering $r_1, \ldots, r_{2\ell_d}$ of the rows of $H_{0,d}$ from top to bottom (where $r_i \in ([n] \times [t])^k$) and an ordering $c_1, \ldots, c_{2\ell_d}$ of the columns of $H_{0,d}$ from left to right (where $c_i \in [n]^k$) such that $H_{0,d}$ is an upper triangular stencil. Thus if we consider the $\ell_d \times \ell_d$ sub-stencil H_d whose rows and columns are $r_{\ell_d+1}, \ldots, r_{2\ell_d}$ and $c_1, \ldots, c_{\ell_d}$, respectively, then all of its entries are zeros. Furthermore, since $H_{0,d}$ is distinctively full rank, we have that $S(r_i) \cap S(r_j) = \varnothing$ for all distinct $i, j \in \{\ell_d + 1, \ldots, 2\ell_d\}$, and $S(c_i) \cap S(c_j) = \varnothing$ for all distinct $i, j \in [\ell_d]$. Thus to show that $\mathrm{drk}(H^{\otimes k}) < 2\ell_d$, it suffices for us to rule out the existence of such sub-stencils H_d in $H^{\otimes k}$.

Consider any collection of rows $R \subseteq ([n] \times [t])^k$ and columns $C \subseteq [n]^k$ in $H^{\otimes k}$ such that $|R| = |C| = \ell_d$ and $S(r_1) \cap S(r_2) = S(c_1) \cap S(c_2) = \varnothing$ for any distinct $r_1, r_2 \in R$ and distinct $c_1, c_2 \in C$. For any pair $(r, c) \in R \times C$, define the collection of pairs $P_{r,c} := \{(p_a(r), p_a(c)) \mid a \in [k]\}$. Let H_d be the $R \times C$ sub-stencil in $H^{\otimes k}$. For any pair $(r, c) \in R \times C$, if $H_d[r, c] = 0$, then by the definition of the tensor power, there must exist an $a \in [k]$ such that $H[p_a(r), p_a(c)] = 0$. Thus by the union bound and the randomness structure of H, we find that

$$\mathbf{Pr}\left[H_d[r, c] = 0\right] \leqslant \sum_{a=1}^{k} \mathbf{Pr}\left[H[p_a(r), p_a(c)] = 0\right] \leqslant \sum_{i=1}^{k} \frac{1}{t} = \frac{k}{t}\,. \tag{1}$$

Now, by the definition of R and C, observe that the sets in $\{P_{r,c} \mid (r, c) \in R \times C\}$ are pairwise disjoint from each other. Additionally, the sets in $\{S_{i,j} \mid i, j \in [n]\}$ are also pairwise disjoint from each other by definition. Since each set $P_{r,c}$ has size most k and each set $S_{i,j}$ has size exactly t, then by a simple greedy process, we can find a collection of pairs $T \subseteq R \times C$ such that $|T| \geqslant \ell_d^2/(kt)$ and for each $i, j \in [n]$, the set $S_{i,j}$ has a nonempty intersection with at most one set in $\{P_{r,c} \mid (r, c) \in T\}$. By the randomness structure of H, this implies that the events $\{H_d[r, c] = 0 \mid (r, c) \in T\}$ are jointly independent from each other. Thus by using Inequality 1 and the inequalities $k \leqslant t - 1$ and $1 - z \leqslant e^{-z}$ for $z \in \mathbb{R}$, we find that

$$\begin{aligned}
\mathbf{Pr}\left[H_d \text{ is all-zeroes}\right] &= \mathbf{Pr}\left[H_d[r, c] = 0 \text{ for all } (r, c) \in R \times C\right] \\
&\leqslant \mathbf{Pr}\left[H_d[r, c] = 0 \text{ for all } (r, c) \in T\right] \\
&= \prod_{(r,c)\in T} \mathbf{Pr}\left[H_d[r, c] = 0\right] \\
&= \prod_{(r,c)\in T} \left(\frac{k}{t}\right) \leqslant \left(\frac{k}{t}\right)^{\ell_d^2/(kt)} \leqslant \left(1 - \frac{1}{t}\right)^{\ell_d^2/(kt)} \leqslant \exp\left(-\frac{\ell_d^2}{kt^2}\right)\,.
\end{aligned}$$

Thus by applying a union bound over all possible choices of rows $R \subseteq ([n] \times [t])^k$ and columns $C \subseteq [n]^k$, we find that

$$\mathbf{Pr}\left[\mathrm{drk}(H^{\otimes k}) \geqslant 2\ell_d\right] \leqslant \binom{(tn)^k}{\ell_d}\binom{n^k}{\ell_d} \exp\left(-\frac{\ell_d^2}{kt^2}\right) \leqslant \left(t^k n^{2k} e^{-\frac{\ell_d}{kt^2}}\right)^{\ell_d}\,.$$

Setting $\ell_d = 4k^2t^2 \log n$, we find that $\mathbf{Pr}\left[\mathrm{drk}(H^{\otimes k}) \geqslant 2\ell_d\right] \leqslant \exp(-\Omega(k^3t^2(\log n)^2))$. Thus for n sufficiently large, the event that $\mathrm{drk}(H^{\otimes k}) < 8k^2t^2 \log n$ for all $k = 1, \ldots, t-1$ occurs with high probability. This finishes the proof. $\square$

8 Further Directions and Discussion

Stencils provide an initial framework for capturing combinatorial methods in lower-bounding matrix ranks. However, visible rank still suffers from some limitations in 2-DRGP stencils as well as small tensor powers of q-LCC stencils. We leave the reader with questions that remain open about the current framework.

1. While we may have shown that the k'th tensor power of a q-LCC does not yield better lower bounds for $k \leqslant n^{o(1)}$, this does not rule out the possibility that the visible capacity might yield better lower bounds. In fact, we do not know if there are any stencils for which the visible capacity does not match the lowest possible rank for the stencil. In other words, does there exist a stencil H such that $\mathrm{rk}_{\mathbb{F}}(H) > \Upsilon(H)$ for every field $\mathbb{F}$? A good starting point for this question would be finding examples where the inequality in Corollary 14 is strict.

2. In this paper, we showed a polynomial in n gap between $\mathrm{rk}_{\mathbb{F}}(H)$ and $\mathrm{vrk}(H)$ by proving that there are 2-DRGP stencils H with $\mathrm{vrk}(H) = O(\log n)$ and $\mathrm{rk}_{\mathbb{F}}(H) = \Omega(\sqrt{n})$. On the other hand, the ranks of q-LCC stencils are in the opposite regime of 2-DRGP stencils, and thus it raises the possibility of having a similar gap in this regime. More concretely, can there also be a similar polynomial gap between the quantities $n - \mathrm{rk}_{\mathbb{F}}(H)$ and $n - \mathrm{vrk}(H)$? From Proposition 20, we have seen that the visible ranks of the k'th tensor power for $k \leqslant n^{o(1)}$ would not suffice to show this polynomial gap. Nonetheless, it still leaves the possibility of using the visible capacity to show this polynomial gap, but we do not know of any methods that can lower bound the visible capacity other than the visible ranks of finite tensor powers. Note that this question is the symmetric spanoid version of Question 2 posed by [DGGW20].

3. One avenue in which progress can be made on lower bounds is through imposing restrictions on what algebraic witnesses are admissible. One natural restriction is to consider *k-colorable stencils*, a stencil wherein any algebraic witness of it can only use at most k distinct nonzero field elements. Can one consider better combinatorial proxies for the algebraic rank of k-colorable stencils (particularly constant-colorable stencils)?

   Constant-colorable stencils can attain asymptotically better lower bounds on their algebraic rank than usual stencils. Indeed, consider a k-colorable stencil $H \in \{0, \star\}^{n \times n}$ where $H[i,i] = 0$ and $H[i,j] = \star$ for $i \neq j$. It is easy to see that $\mathrm{vrk}(H) = 2$, but since H is a k-colorable stencil, the algebraic witness used in Proposition 9 would not apply here anymore to show that the algebraic rank is at most 2. In fact, one can show by a polynomial argument (which we leave for the reader to try) that the rank of any algebraic witness of H is at least $\Omega_k(n^{1/k})$.

4. One component that stencils do not capture is the distance of a code, but that is partly due to its dependence on the field size. To remedy this issue, we can instead consider colorable stencils. In this restricted model, are there combinatorial proxies that upper and lower bound the distance of any algebraic witness of the colorable stencil?

5. Can one explicitly construct an $n \times n$ stencil H whose visible rank is robust to small entry corruptions? Any algebraic witness of H would automatically be a rigid matrix.

Finally, our aim in introducing the visible rank is to study combinatorial proxies for the rank of a matrix. Are there other combinatorial quantities associated with a matrix that could yield more effective lower bounds on its rank?

References

[ABG+20] Noga Alon, Igor Balla, Lior Gishboliner, Adva Mond, and Frank Mousset. The minrank of random graphs over arbitrary fields. *Israel Journal of Mathematics*, 235(1):63–77, 2020.

[AG21] Omar Alrabiah and Venkatesan Guruswami. Visible rank and codes with locality. *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, 2021.

[AGKM23] Omar Alrabiah, Venkatesan Guruswami, Pravesh K. Kothari, and Peter Manohar. A near-cubic lower bound for 3-query locally decodable codes from semirandom csp refutation, 2023.

[AL06] Noga Alon and Eyal Lubetzky. The Shannon capacity of a graph and the independence numbers of its powers. *IEEE Transactions on Information Theory*, 52(5):2172–2176, 2006.

[BBJK11] Ziv Bar-Yossef, Yitzhak Birk, T. S. Jayram, and Tomer Kol. Index coding with side information. *IEEE Trans. Inf. Theory*, 57(3):1479–1494, 2011.

[BDYW11] Boaz Barak, Zeev Dvir, Amir Yehudayoff, and Avi Wigderson. Rank bounds for design matrices with applications to combinatorial geometry and locally correctable codes. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 519–528, 2011.

[BGT17] Arnab Bhattacharyya, Sivakanth Gopi, and Avishay Tal. Lower bounds for 2-query lccs over large alphabet. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2017)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.

[D+12] Zeev Dvir et al. Incidence theorems and their applications. *Foundations and Trends® in Theoretical Computer Science*, 6(4):257–393, 2012.

[DGGW20] Zeev Dvir, Sivakanth Gopi, Yuzhou Gu, and Avi Wigderson. Spanoids - an abstraction of spanning structures, and a barrier for LCCs. *SIAM J. Comput.*, 49(3):465–496, 2020.

[DS05] Zeev Dvir and Amir Shpilka. Locally decodable codes with 2 queries and polynomial identity testing for depth 3 circuits. In Harold N. Gabow and Ronald Fagin, editors, *Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005*, pages 592–601. ACM, 2005.

[DSW14] Zeev Dvir, Shubhangi Saraf, and Avi Wigderson. Improved rank bounds for design matrices and a new proof of kelly's theorem. In *Forum of Mathematics, Sigma*, volume 2. Cambridge University Press, 2014.

[DSW17] Zeev Dvir, Shubhangi Saraf, and Avi Wigderson. Superquadratic lower bound for 3-query locally correctable codes over the reals. *Theory of Computing*, 13(11):1–36, 2017.

[FGW17] S. Luna Frank-Fischer, Venkatesan Guruswami, and Mary Wootters. Locality via partially lifted codes. In *Approximation, Randomization, and Combinatorial Optimization.*

Algorithms and Techniques, APPROX/RANDOM, volume 81 of *LIPIcs*, pages 43:1–43:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017.

[FJJ18] Mathew C Francis, Dalu Jacob, and Satyabrata Jana. Uniquely restricted matchings in interval graphs. *SIAM Journal on Discrete Mathematics*, 32(1):148–172, 2018.

[FVY15] Arman Fazeli, Alexander Vardy, and Eitan Yaakobi. Codes for distributed PIR with low storage overhead. In *IEEE International Symposium on Information Theory*, pages 2852–2856, 2015.

[GHL01] Martin Charles Golumbic, Tirza Hirst, and Moshe Lewenstein. Uniquely restricted matchings. *Algorithmica*, 31(2):139–154, 2001.

[GHSY12] Parikshit Gopalan, Cheng Huang, Huseyin Simitci, and Sergey Yekhanin. On the locality of codeword symbols. *IEEE Transactions on Information theory*, 58(11):6925–6934, 2012.

[GKS13] Alan Guo, Swastik Kopparty, and Madhu Sudan. New affine-invariant codes from lifting. In *Proceedings of the Innovations in Theoretical Computer Science Conference*, pages 529–540, 2013.

[GKST06] Oded Goldreich, Howard Karloff, Leonard J Schulman, and Luca Trevisan. Lower bounds for linear locally decodable codes and private information retrieval. *Computational Complexity*, 15(3):263–296, 2006.

[GKT01] Harold N Gabow, Haim Kaplan, and Robert E Tarjan. Unique maximum matching algorithms. *Journal of Algorithms*, 40(2):159–183, 2001.

[Gop18] Sivakanth Gopi. *Locality in coding theory*. PhD thesis, Princeton University, 2018.

[GRW18] Alexander Golovnev, Oded Regev, and Omri Weinstein. The minrank of random graphs. *IEEE Trans. Inf. Theory*, 64(11):6990–6995, 2018.

[Hae79] Willem H. Haemers. On some problems of Lovász concerning the Shannon capacity of a graph. *IEEE Trans. Inf. Theory*, 25(2):231–232, 1979.

[HMT06] Thanh Minh Hoang, Meena Mahajan, and Thomas Thierauf. On the bipartite unique perfect matching problem. In *International Colloquium on Automata, Languages, and Programming*, pages 453–464. Springer, 2006.

[HSX+12] Cheng Huang, Huseyin Simitci, Yikang Xu, Aaron Ogus, Brad Calder, Parikshit Gopalan, Jin Li, and Sergey Yekhanin. Erasure coding in windows azure storage. In *USENIX Annual Technical Conference*, 2012.

[IS20] Eran Iceland and Alex Samorodnitsky. On coset leader graphs of structured linear codes. *Discrete and Computational Geometry*, 63:560–576, 2020.

[KT00] Jonathan Katz and Luca Trevisan. On the efficiency of local decoding procedures for error-correcting codes. In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing*, pages 80–86, 2000.

[KW04] Iordanis Kerenidis and Ronald de Wolf. Exponential lower bound for 2-query locally decodable codes via a quantum argument. *Journal of Computer and System Sciences*, 69(3):395–420, 2004.

[Lov89] László Lovász. Communication complexity: A survey. Technical report, Princeton University TR-204-89, February 1989. University of Zurich, Department of Informatics.

[LW21] Ray Li and Mary Wootters. Lifted multiplicity codes and the disjoint repair group property. *IEEE Trans. Inf. Theory*, 67(2):716–725, 2021.

[LW22] Ray Li and Mary Wootters. Improved batch code lower bounds. In *2022 IEEE International Symposium on Information Theory (ISIT)*, pages 1596–1599. IEEE, 2022.

[Mis11] Sounaka Mishra. On the maximum uniquely restricted matching for bipartite graphs. *Electronic Notes in Discrete Mathematics*, 37:345–350, 2011.

[Ram20] C. Ramya. Recent progress on matrix rigidity–a survey. *arXiv preprint arXiv:2009.09460*, 2020.

[RV16] Sankeerth Rao and Alexander Vardy. Lower bound on the redundancy of PIR codes. *arXiv preprint arXiv:1605.01869*, 2016.

[SAP+13] Maheswaran Sathiamoorthy, Megasthenis Asteris, Dimitris Papailiopoulos, Alexandros Dimakis, Ramkumar Vadali, Scott Chen, and Dhruba Borthakur. XORing elephants: Novel erasure codes for big data. *Proc. VLDB Endow.*, 6:325–336, 2013.

[SS96] Michael Sipser and Daniel A Spielman. Expander codes. *IEEE transactions on Information Theory*, 42(6):1710–1722, 1996.

[TD13] Maguy Trefois and Jean-Charles Delvenne. Zero forcing sets, constrained matchings and minimum rank. *Linear and Multilinear Algebra*, 2013.

[Tré16] Maguy Tréfois. *Topics in combinatorial matrix theory*. PhD thesis, PhD thesis, UCL, 2016.

[Val77] Leslie G Valiant. Graph-theoretic arguments in low-level complexity. In *International Symposium on Mathematical Foundations of Computer Science*, pages 162–176. Springer, 1977.

[Woo07] David Woodruff. New lower bounds for general locally decodable codes. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 14, 2007.

[Woo12] David P. Woodruff. A quadratic lower bound for three-query linear locally decodable codes over any field. *J. Comput. Sci. Technol.*, 27(4):678–686, 2012.

[Woo16] Mary Wootters. Linear codes with disjoint repair groups. *unpublished manuscript*, 2016. URL: `https://web.stanford.edu/~marykw/files/disjoint_repair_groups.pdf`.

[Yek12] Sergey Yekhanin. Locally decodable codes. *Foundations and Trends in Theoretical Computer Science*, 6(3):139–255, 2012.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il