Explicit Rank Extractors and Subspace Designs via Function Fields, with Applications to Strong Blocking Sets

Zeyu Guo* Roshan Raj* Chong Shangguan† Zihan Zhang*

Abstract

We give new explicit constructions of several fundamental objects in linear-algebraic pseudorandomness and combinatorics, including lossless rank extractors, weak subspace designs, and strong s-blocking sets over finite fields.

Our focus is on the small-field regime, where the field size depends only on a secondary parameter (such as the rank or codimension) and is independent of the ambient dimension. This regime is central to several applications, yet remains poorly understood from the perspective of explicit constructions.

In this setting, we obtain the first explicit constructions of lossless rank extractors and weak subspace designs for $r \ll k$, where r denotes the rank (or codimension), over finite fields $\mathbb{F}_q$ with $q \geq \mathrm{poly}(r)$ and q non-prime, with near-optimal parameters. For other finite fields, including prime fields and small fields, we obtain weaker but still improved bounds.

As a consequence, we construct explicit strong s-blocking sets in $\mathrm{PG}(k-1,q)$ of size $O(s(k-s)q^s)$ for all sufficiently large non-prime fields $q \geq \mathrm{poly}(s)$, matching the best known non-explicit bounds up to constant factors. This significantly improves the previous best bound $2^{O(s^2 \log s)} q^s k$ of Bishnoi and Tomon (Combinatorica, 2026), which requires $q \geq 2^{\Omega(s)}$.

Our approach is primarily algebraic, combining techniques from function fields and polynomial identity testing. In addition, we develop a complementary Fourier-analytic framework based on ε-biased sets, which yields improved explicit constructions of strong s-blocking sets over small fields.

1 Introduction

A central theme in theoretical computer science is the explicit construction of combinatorial and algebraic objects that match the guarantees of the probabilistic method. In recent years, a rich theory of *linear-algebraic pseudorandomness* has emerged, focusing on structured collections of linear maps that behave like random ones with respect to rank and dimension. Prominent examples include lossless rank extractors, subspace designs, and dimension expanders, which play a key role in polynomial identity testing, coding theory, and derandomization.

A recurring challenge in this area is to obtain explicit constructions over *small finite fields*. While the probabilistic method often yields near-optimal parameters over any field, known explicit constructions typically require the field size to grow with the ambient dimension. Bridging this gap, especially in the regime where the field size depends only on a small parameter, is a central open problem with numerous applications.

In this work, we study this problem for several natural objects in linear-algebraic pseudorandomness and combinatorics, including lossless rank extractors, weak subspace designs, and strong

[1]The Ohio State University, zguotcs@gmail.com, amritanshus128@gmail.com, zhang.13691@osu.edu
[2]Shandong University, theoreming@163.com

s-blocking sets. Our main algebraic approach uses function fields and polynomial identity testing, and for strong s-blocking sets we also develop a complementary Fourier-analytic approach via ε-biased sets.

Lossless rank condensers and extractors. A central object in the theory of "linear-algebraic pseudorandomness" [FG14] is the notion of (seeded) lossless rank condensers and extractors [GR08, FS12, FSS14, FG14].

Definition 1.1 (Lossless rank condensers and extractors)**.** Let $r \le t \le k$ be positive integers. A finite collection $\mathcal{E} = \{E_i\}_{i=1}^{n} \subseteq \mathbb{F}^{t\times k}$ of matrices over a field $\mathbb{F}$ is called a (k, r, t, L) lossless rank condenser over $\mathbb{F}$ if for every matrix $M \in \mathbb{F}^{k\times r}$ of rank r, the number of indices $i \in [n]$ for which $\operatorname{rank}(E_i M) < r$ is at most L. We call n the size of $\mathcal{E}$. When $r = t$, which is optimal, such an object is called a (k, r, L) lossless rank extractor.

Explicit lossless rank extractors of small size were first constructed by Gabizon and Raz [GR08], who used them to obtain explicit affine extractors over large fields. In particular, they gave explicit (k, r, kr^2) lossless rank extractors of any size $n \le q$. Forbes and Shpilka [FS12] gave a different construction over fields $\mathbb{F}$ with $|\mathbb{F}| > k$ using Vandermonde matrices, improving the parameter L to kr. Forbes, Saptharishi, and Shpilka [FSS14] refined the analysis of this construction and achieved $L = r(k - r)$, which is optimal over algebraically closed fields (see [FS16, Guo24, BCDZ25b]). It is also known that explicit lossless rank condensers can be constructed from classical Wronskians [GK16, FSS14].

Beyond their role in constructing affine extractors [GR08], lossless rank condensers and extractors play a central role in polynomial identity testing for various models, including depth-3 arithmetic circuits [KS11, KS09, SS13, SS12] and read-once oblivious arithmetic branching programs [FS13, FSS14, AGKS15]. For further applications and a comprehensive discussion, see [FG14].

Nonlinear variants of these objects have also been studied. Dvir, Gabizon, and Wigderson [DGW09] introduced deterministic rank extractors for polynomial sources and gave explicit constructions. Guo, Volk, Jalan, and Zuckerman [GVJZ23] extended this line of work by defining deterministic rank condensers and extractors for varieties as dimension-preserving maps, yielding explicit constructions for algebraic sources that generalize [DGW09, Dvi12]. They also observed that the objects in Definition 1.1 can be viewed as linear seeded rank condensers and extractors for varieties. Finally, Guo [Guo24] introduced variety evasive subspace families as a nonlinear generalization, with further applications to polynomial identity testing.

In this paper, we focus on explicit lossless rank *extractors* over a finite field $\mathbb{F}_q$, i.e., the case $r = t$ and $\mathbb{F} = \mathbb{F}_q$ in Definition 1.1. This setting is crucial for many of the applications discussed above.

From a non-explicit perspective, the probabilistic method shows that (k, r, L) lossless rank extractors of small size exist over any finite field $\mathbb{F}_q$, with $L = \operatorname{poly}(k)$ (and in fact $L = O(r(k-r))$; see Theorem A.1). However, known explicit constructions such as [GR08, FS12] achieve such parameters only when the field size is sufficiently large.

To avoid trivialities, we assume $L < n$, since otherwise one could allow $\operatorname{rank}(E_i M) < r$ for all $i \in [n]$. Under this assumption, the constructions in [GR08, FS12] require $q \ge n$, and hence $q > L$. More concretely, this gives $q > kr^2$ in [GR08] and $q > r(k - r)$ in [FS12]. In addition, the construction in [FS12] requires $q > k$, as it relies on the existence of an element $\gamma \in \mathbb{F}_q^{\times}$ whose multiplicative order is at least k.

A natural approach to constructing lossless rank extractors over small fields is via concatenation or field-reduction techniques. Indeed, [FG14, Proposition 8.5] shows that a (k, r, t, L) lossless rank

condenser over $\mathbb{F}_{q^d}$ can be converted into a (k, r, dt, L) lossless rank condenser over $\mathbb{F}_q$. However, unless $d = 1$, this transformation does not preserve the lossless rank *extractor* property, as it necessarily increases the output dimension to $dt > r$.

This motivates the problem of constructing explicit lossless rank extractors over small fields. In particular, we ask:

Question A. *Suppose $r \ll k$. Do there exist explicit (k, r, L) lossless rank extractors of size $n > L$ over $\mathbb{F}_q$ such that $L = \mathrm{poly}(k, r)$ and q depends only on r?*

As we will see, we answer Question A in the affirmative.

Subspace designs. Informally, a *subspace design* is a collection of linear subspaces arranged so that every subspace of a given dimension intersects the family only in a limited way — either by intersecting only few members (weak designs) or by having small total intersection dimension (strong designs).

Definition 1.2 (Subspace designs)**.** A collection of subspaces $\{V_i\}_{i\in[n]} \subseteq \mathbb{F}_q^k$ is a (t, A) subspace design if for any subspace $W \subseteq \mathbb{F}_q^k$ of dimension t, we have

$$\sum_{i\in[n]} \mathbf{1}[V_i \cap W \neq \{0\}] \leq A \quad \text{(weak design)} \qquad \text{or} \qquad \sum_{i\in[n]} \dim(V_i \cap W) \leq A \quad \text{(strong design)},$$

where $\mathbf{1}[V_i \cap W \neq \{0\}]$ is the indicator function of the event $V_i \cap W \neq \{0\}$.

Introduced by Guruswami and Xing [GX13], subspace designs have since become a central combinatorial primitive in coding theory and beyond. While the probabilistic method yields subspace designs with near-optimal size and dimension [GX13, GK16], obtaining explicit constructions is substantially more challenging. Guruswami and Kopparty [GK16] gave the first explicit construction of even strong subspace designs with near-optimal parameters over large fields, based on folded Reed–Solomon and univariate multiplicity codes, thereby fully derandomizing the code constructions of [GX13]. Subsequently, Guruswami, Xing, and Yuan [GXY18] extended this approach to significantly smaller fields via algebraic function field techniques, albeit with a noticeable loss in parameters.

Subspace designs play a fundamental role in the construction of list-decodable codes. Many of the best-known constructions—ranging from Reed–Solomon and folded Reed–Solomon codes to algebraic geometry, multiplicity, and rank-metric codes—arise by combining classical algebraic codes with subspace design-based pruning [GX13, GK16, GX22]. More recently, this paradigm has been substantially strengthened: the work of Chen and Zhang [CZ25] shows that explicit folded Reed–Solomon and multiplicity codes themselves achieve optimal list-decodability, and in fact identifies subspace designs as the underlying mechanism governing this phenomenon. In particular, any code admitting sufficiently strong subspace designs—termed *subspace designable codes* [CZ25] or *subspace design codes* [GG25]—automatically enjoys near-optimal list-decodability.

Beyond coding theory, strong subspace designs have emerged as a central tool in pseudorandomness and linear-algebraic expansion. In particular, they enable explicit constructions of constant-degree lossless dimension expanders [GRX21], which can be viewed as a linear-algebraic analogue of expander graphs. More recently, influenced by the work of Chen and Zhang [CZ25], a sequence of works [BCDZ25b, BCDZ25a, GG25, JLR26] has uncovered a significantly broader and more structural role for subspace designs. These works position strong subspace designs as a unifying bridge between random and explicit constructions, leading to systematic derandomization frameworks

and new connections to local properties, matroid theory, and proximity gaps—a central notion in modern proof systems.

While most of the aforementioned applications primarily rely on strong subspace designs, explicit constructions of weak subspace designs are equally important. In particular, as shown in [FS14, FS16], weak subspace designs already suffice to yield constructions of strong blocking sets, a prominent object in finite geometry that we discuss in the next paragraph. This connection highlights that weak designs capture a fundamentally combinatorial aspect of the theory and may admit constructions and techniques distinct from those required for strong designs, which are the focus of much of the existing literature.

More specifically, recent work [BCDZ25b] (see also [FS16, Guo24]) establishes lower bounds on the parameters of weak subspace designs. In particular, it shows that for any collection of subspaces $\{V_i\}_{i\in[n]} \subseteq \overline{\mathbb{F}}^k$ of dimension s (where $\overline{\mathbb{F}}$ is an algebraically closed field), if it forms a $(k-s, A)$ subspace design, then necessarily $A \geq s(k-s)$. For sufficiently large fields, explicit constructions due to Guruswami and Kopparty [GK16] match this lower bound even for strong subspace designs. In contrast, obtaining comparable explicit constructions even for weak subspace designs over small fields remains open, as discussed below, and is of particular importance for applications such as constructing strong blocking sets.

Question B. *Can one explicitly construct a collection of subspaces $\{V_i\}_{i\in[n]} \subseteq \mathbb{F}^k$ of dimension s (resp. dimension $k-s$) over fields of size $|\mathbb{F}| = O_s(1)$ that forms a $(k-s, A)$ (resp. (s, A)) weak subspace design with $A = O(s(k-s))$, or with A close to this lower bound?*

Prior to our work, the best known result toward this question was implicitly due to Bishnoi and Tomon [BT26], who obtained a construction over fields of size $q = 2^{\Omega(s)}$ with parameter $A = 2^{O(s^2 \log s)}n$ via expander graphs and algebraic geometry codes. As we will see, our results give a strong positive answer to Question B, achieving $q = \mathrm{poly}(s)$ and $A = O(s(k-s))$ in the regimes covered by our constructions.

Taken together, all these developments demonstrate that both strong and weak subspace designs are not merely a technical ingredient in code constructions, but rather a fundamental combinatorial primitive with far-reaching applications across coding theory, pseudorandomness, and algebraic combinatorics. For a more fine-grained historical overview, we refer the reader to the recent survey by Santonastaso and Zullo [SZ23].

Blocking sets. Blocking sets, introduced by Richardson [Ric56], are classical objects in finite geometry that capture the minimal structure required to intersect all subspaces of a given codimension. For $1 \leq s \leq k$, an *affine s-blocking set* is a subset $B \subseteq \mathbb{F}_q^k$ that intersects every affine subspace of codimension s. In the projective setting, for $1 \leq s \leq k-1$, a set $B \subseteq \mathrm{PG}(k-1, q)$ is a *strong s-blocking set* if for every codimension-s projective subspace Σ of $\mathrm{PG}(k-1, q)$, the intersection $B \cap \Sigma$ spans Σ, i.e., the smallest projective subspace containing $B \cap \Sigma$ equals Σ. Let $b_q(k, s)$ and $b_q^*(k, s)$ denote the minimum sizes of affine and strong blocking sets, respectively.

Blocking sets are related to several well-studied notions in discrete mathematics, theoretical computer science, and coding theory, including vertex covers [Für88], subspace designs [FS16, GK16], trifferent codes [BDGP24], and intersecting and minimal codes [CL85, ABN22, TQLZ21, XKH25].

The affine and projective variants are tightly related: it was shown in [BDGP24] that a set $B \subseteq \mathrm{PG}(k-1, q)$ is a strong s-blocking set if and only if the union of the corresponding lines through the origin forms an affine $(s+1)$-blocking set in $\mathbb{F}_q^k$. In particular,

$$(q-1)b_q^*(k, s) + 1 \geq b_q(k, s+1). \tag{1}$$

The main problem is to determine the values of $b_q(k,s)$ and $b_q^*(k,s)$. For $s=1$, a classical result of Jamison [Jam77] and Brouwer and Schrijver [BS78] shows that

$$b_q(k,1) = (q-1)k+1.$$

For $s \geq 2$, however, the problem becomes significantly more difficult and is related to hard problems in additive combinatorics such as the density Hales–Jewett theorem and the cap set problem (see [BDGP24]).

The best known general lower bound is

$$b_q(k,s) \geq (q^s-1)(k-s+1)+1, \tag{2}$$

which follows from a geometric argument [Bal11].

On the other hand, viewing blocking sets as vertex covers in a natural hypergraph, the Lovász–Stein theorem [Juk11, Theorem 2.16] yields the upper bound

$$b_q(k,s) \leq q^s \left(1 + \ln \binom{k}{s}_q\right),$$

where $\binom{k}{s}_q$ denotes the Gaussian binomial coefficient, i.e., the number of s-dimensional subspaces of $\mathbb{F}_q^k$. A recent result of Bishnoi et al. [BDGP24] improves this to an upper bound that is roughly within a factor $O(s)$ of (2) via a probabilistic argument.

For strong s-blocking sets, combining (1) and (2) yields

$$b_q^*(k,s) \geq \frac{(q^{s+1}-1)(k-s)}{q-1}. \tag{3}$$

For $s=1$, Bishnoi et al. [BDGP24] obtained a stronger asymptotic lower bound

$$b_q^*(k,1) \geq (c_q - o(1))(q+1)(k-1)$$

for some constant $c_q > 1$, using linear programming bounds from coding theory. The same work also gives a probabilistic upper bound for $b_q^*(k,s)$ which is rougly within a factor $O(s)$ of (3) (see [BDGP24, Remark 3.4]).

The main challenge is to obtain explicit constructions of strong[1] s-blocking sets with comparable parameters. When the field size q is sufficiently large as a function of k, constructions based on higgledy-piggledy subspaces [FS14, FS16] yield strong s-blocking sets of size $O_{k,s}(q^s)$. In the regime where q, s are fixed and k grows, the best known explicit constructions of strong s-blocking sets have size $O_s(q^s k)$ [ABDN24, BT26].

For $s=1$, a breakthrough result of Alon, Bishnoi, Das, and Neri [ABDN24] gave the first explicit construction of size $O(qk)$ using expander graphs and coding-theoretic tools. This approach was later extended by Bishnoi and Tomon [BT26] to all $s \geq 2$, yielding constructions of size at most $C_s q^s k$, where $C_s = 2^{O(s^2 \log s)}$. However, the dependence on s in C_s is large, and their construction requires the field size q to be at least exponential in s. For smaller fields, an explicit construction of size $q^{O(s^2)}k$ was given in [BT26].

Our main contribution is explicit constructions of strong s-blocking sets of size $O(s(k-s)q^s)$ over fields $\mathbb{F}_q$ with $q \geq \mathrm{poly}(s)$, together with extensions to smaller fields. This improves the s-dependence in the leading constant of the construction in [BT26] from exponential to polynomial. By (1), these constructions also yield explicit affine $(s+1)$-blocking sets with comparable parameters.

[1] This automatically yields explicit constructions for affine $(s+1)$-blocking sets.

1.1 Main Results

In this paper, we obtain new explicit constructions of lossless rank extractors, weak subspace designs, and strong s-blocking sets over small fields, achieving parameters that were previously unknown.

Lossless rank extractors. Over non-prime finite fields, we obtain the following result on explicit lossless rank extractors.

Theorem 1.3 (Informal version of Theorem 4.5)**.** *For every $r \geq 1$ and every sufficiently large non-prime prime power $q \geq \text{poly}(r)$, there exist infinitely many $k \geq r$ such that one can explicitly construct a (k, r, L) lossless rank extractor of size n over $\mathbb{F}_q$ with $L = O(r(k-r))$ and $L/n \leq q^{-1/4}$.*

Note that the field size q is only required to be polynomial in r, and may be independent of k. The upper bound on the number of bad matrices is $L = O(r(k-r))$, which matches the upper bound $r(k-r)$ in [FSS14, For14] up to a constant factor.

Remark 1.4. *Theorem 1.3 and the subsequent theorems hold only for infinitely many values of k, for the same reason that families of algebraic geometry (AG) codes constructed from function field towers are known to be asymptotically good only for infinitely many lengths, rather than for all lengths. This "infinitely many k" phenomenon also appears in the strong 1-blocking set construction of Alon, Bishnoi, Das, and Neri [ABDN24], which uses either Justesen codes or AG codes, and hence yields asymptotically good codes only for infinitely many lengths. Similarly, it also appears in the explicit constructions of Bishnoi and Tomon [BT26] over small fields, which are also based on AG codes.*

One can extend Theorem 1.3 and the subsequent theorems to all k, and indeed we prove such versions for all k, but then the bound $L = O(r(k-r))$ must increase by an additional factor of at most $O(q)$. See Theorem 4.5, for example.

Over prime finite fields, we obtain the following weaker result.

Theorem 1.5 (Informal version of Theorem 5.10)**.** *For every $r \geq 1$ and $\delta \in (0, 1)$, and every sufficiently large prime $q \geq \text{poly}(r)$, there exist infinitely many $k \geq r$ such that an explicit (k, r, L) lossless rank extractor of size n over $\mathbb{F}_q$ can be constructed with $L \leq (2r/\delta)^{O(\log r)} r(k-r)$ and $L/n \leq \delta$.*

Over fields of absolute constant size, such as $\mathbb{F}_2$, we do not know how to explicitly construct lossless rank extractors for which, for every full-rank matrix M, most $E_i \in \mathcal{E}$ satisfy $\text{rank}(E_i M) = r$. However, we can explicitly construct a weaker object in which *some* E_i satisfies $\text{rank}(E_i M) = r$. Indeed, Forbes [For14] defined rank condensers in this weaker sense.

We restrict to the case where the rank equals the output dimension and call these objects *rank dispersers*, drawing on the intuition that dispersers are one-sided weakenings of extractors.

Definition 1.6 (Lossless rank dispersers)**.** Let $1 \leq r \leq k$ be integers. A finite collection $\mathcal{E} = \{E_i\}_{i=1}^n \subseteq \mathbb{F}^{r \times k}$ of matrices over a field $\mathbb{F}$ is called a (k, r) lossless rank disperser over $\mathbb{F}$ if for every matrix $M \in \mathbb{F}^{k \times r}$ of rank r, there exists $i \in [n]$ such that $\text{rank}(E_i M) = r$. We call n the size of $\mathcal{E}$.

In other words, $\mathcal{E}$ is a (k, r) lossless rank disperser of size n if and only if it is a $(k, r, n-1)$ lossless rank extractor of size n.

The following theorem gives explicit lossless rank dispersers over *any* finite field, including $\mathbb{F}_2$.

Theorem 1.7 (Informal version of Theorem 5.8)**.** *For every $r \geq 1$ and every prime power $q > 1$, there exist infinitely many $k \geq r$ such that one can explicitly construct a (k, r) lossless rank disperser of size n over $\mathbb{F}_q$, where $n = O\left(\max\left\{\frac{c \log r}{\log q}, 2\right\}^r r(k-r)\right)$ and $c > 0$ is some absolute constant.*

Compared with Theorem 1.3 and Theorem 1.5, Theorem 1.7 only yields lossless rank dispersers. However, as we will see, this is still sufficient for constructing strong s-blocking sets.

We also remark that the constant 2 in $\max\left\{\frac{c \log r}{\log q}, 2\right\}$ in Theorem 1.7 (and similarly in Theorem 1.10 and Theorem 1.14) can be improved to 1 when q is non-prime; see the proof of Theorem 5.8 and Remark 5.9. However, this improvement is relevant only for sufficiently large non-prime fields $q \geq \text{poly}(r)$, in which case one can instead apply Theorem 1.5, which yields stronger bounds.

Weak subspace designs. On non-prime, prime, and all finite fields, we obtain explicit weak subspace designs with various parameters.

Theorem 1.8 (Informal version of Theorem 5.11 (1))**.** *For every $r \geq 1$ and every sufficiently large non-prime prime power $q \geq \text{poly}(r)$, there exist infinitely many $k \geq r$ such that one can explicitly construct an (r, A) (resp. $(k-r, A)$) weak subspace design of size n over $\mathbb{F}_q$, consisting of $(k-r)$-dimensional (resp. r-dimensional) subspaces of $\mathbb{F}_q^k$, with $A = O(r(k-r))$ and $n \geq (q^{1/4}/2) \cdot A$.*

Theorem 1.9 (Informal version of Theorem 5.11 (2))**.** *For every $r \geq 1$, $\delta \in (0, 1)$, and every sufficiently large prime $q \geq \text{poly}(r)$, there exist infinitely many $k \geq r$ such that one can explicitly construct an (r, A) (resp. $(k-r, A)$) weak subspace design of size n over $\mathbb{F}_q$, consisting of $(k-r)$-dimensional (resp. r-dimensional) subspaces of $\mathbb{F}_q^k$, with $A \leq (2r/\delta)^{O(\log r)} r(k-r)$ and $n \geq A/\delta$.*

Theorem 1.10 (Informal version of Theorem 5.11 (3))**.** *For every $r \geq 1$ and every prime power $q > 1$, there exist infinitely many $k \geq r$ such that one can explicitly construct an (r, A) (resp. $(k-r, A)$) weak subspace design of size n over $\mathbb{F}_q$, consisting of $(k-r)$-dimensional (resp. r-dimensional) subspaces of $\mathbb{F}_q^k$, with $A = n - 1$ and $n = O\left(\max\left\{\frac{c \log r}{\log q}, 2\right\}^r r(k-r)\right)$ for some absolute constant $c > 0$.*

Remark 1.11. *While explicit weak subspace designs over small finite fields are known in the literature, no explicit constructions achieving Theorems 1.8 to 1.10 were known prior to this work.*

One possible approach is to first construct an explicit (r, A) weak subspace design in $\mathbb{F}_{q^d}^k$ over a large finite field $\mathbb{F}_{q^d}$ with $q^d > k$, consisting of codimension-r subspaces, and then use [GK16, Lemma 8] to obtain an (r, A) weak subspace design in $\mathbb{F}_q^{dk}$.

However, this transformation increases the codimension of the subspaces to $dr > r$. Moreover, when r is constant and $q = O_r(1) = O(1)$, we have $d = O(\log_q k) = O(\log k)$, implying that the codimension of the subspaces becomes logarithmic in k rather than constant.

Strong s-blocking sets. Finally, we obtain explicit strong s-blocking sets of improved size over small fields. By (1), these also yield explicit affine $(s+1)$-blocking sets over small fields. See Table 1 for a summary of our results and prior work.

The following theorem shows that for every sufficiently large non-prime $q \geq \text{poly}(s)$, there exist explicit strong s-blocking sets of size $O(s(k-s)q^s)$, matching the best-known non-explicit upper bound (see [BDGP24, Remark 3.4]) up to an absolute constant factor.

Theorem 1.12 (Informal version of Corollary 6.2 (1))**.** *For every $s \geq 1$ and every sufficiently large non-prime prime power $q \geq \text{poly}(s)$, there exist infinitely many $k > s+1$ such that one can explicitly construct a strong s-blocking set $B \subseteq \text{PG}(k-1, q)$ of size $O(s(k-s)q^s)$.*

Reference	Condition on q	Upper bound on the size
[BT26, Theorem 16]	q is square and $\geq s^{\Omega(s)}$	$2^{O(s^2 \log s)} kq^s$
[BT26, Theorem 17]	–	$kq^{O(s^2)}$
Theorem 1.12	q is non-prime and $\geq \mathrm{poly}(s)$	$O(s(k-s)q^s)$
Theorem 1.13	q is prime and $\geq \mathrm{poly}(s)$	$(2s)^{O(\log s)}(k-s)q^s$
Theorem 1.14	–	$O\left(\max\left\{\frac{c\log(2s)}{\log q}, 2\right\}^{s+1} s(k-s)q^s\right)$
Theorem 1.15	–	$kq^{(2+o_{qs}(1))s}$

Table 1: A comparison of explicit constructions of strong s-blocking sets

When $q \geq \mathrm{poly}(s)$ is prime, we obtain a weaker size bound with a quasi-polynomial factor in s.

Theorem 1.13 (Informal version of Corollary 6.2 (2))**.** *For every $s \geq 1$ and every sufficiently large prime $q \geq \mathrm{poly}(s)$, there exist infinitely many $k > s$ such that one can explicitly construct a strong s-blocking set $B \subseteq \mathrm{PG}(k-1, q)$ of size at most $(2s)^{O(\log s)}(k-s)q^s$.*

For arbitrary q, we obtain the following result.

Theorem 1.14 (Informal version of Corollary 6.2 (3))**.** *For every $s \geq 1$ and every prime power $q > 1$, there exist infinitely many $k > s$ such that one can explicitly construct a strong s-blocking set $B \subseteq \mathrm{PG}(k-1, q)$ of size $O\left(\max\left\{\frac{c\log(2s)}{\log q}, 2\right\}^{s+1} s(k-s)q^s\right)$, where $c > 0$ is an absolute constant. In particular, when $q = 2$, the size of B is at most $2^{O(s \log\log s)}(k-s)q^s$.*

Setting $s = 1$, Theorem 1.14 recovers the main result of Alon, Bishnoi, Das, and Neri that there exist explicit strong 1-blocking sets of size $O(kq)$. Notably, our proof (and construction) does not use expander graphs as in [ABDN24], and can be viewed as a purely algebraic approach to the same result.

Theorems 1.12 to 1.14 are proved using algebraic techniques, including tools from function fields and polynomial identity testing.

We also develop a Fourier-analytic approach, reducing the construction of explicit strong s-blocking sets in $\mathrm{PG}(k-1, q)$ to that of explicit ε-biased sets in $\mathbb{F}_q^k$. By instantiating this reduction with Ta-Shma's breakthrough construction [TS17] (and its extension to general finite fields $\mathbb{F}_q$ by Jalan and Moshkovitz [JM21]), we obtain the following theorem.

Theorem 1.15 (Informal version of Proposition 6.8)**.** *For every $1 \leq s < k$ and every prime power $q > 1$, there exists an explicit strong s-blocking set $B \subseteq \mathrm{PG}(k-1, q)$ of size at most $kq^{(2+o_{qs}(1))s}$.*

Previously, [BT26, Theorem 17] constructed explicit strong s-blocking sets of size $kq^{O(s^2)}$ for all q. Theorem 1.15 improves this to a size bound of $kq^{(2+o_{qs}(1))s}$.

While this bound is weaker than that of Theorem 1.14 for large enough q, e.g., for $q \geq \mathrm{poly}(s)$, it is the best currently known bound when $q = O(1)$, namely $2^{O(s)}kq^s$. In contrast, Theorem 1.14 yields $2^{O(s \log\log s)}kq^s$ in this regime. Moreover, Theorem 1.15 applies to all k, rather than only infinitely many k.

1.2 Technical Overview

We primarily focus on Question A, namely the problem of explicitly constructing (k, r, L) lossless rank extractors over finite fields whose size depends only on r (and in fact $\mathrm{poly}(r)$), in the regime where $r \ll k$. We also explain why some existing constructions do not directly address this question.

For a field $\mathbb{F}$ with $|\mathbb{F}| > k$, Forbes and Shpilka [FS12] constructed an explicit (k, r, L) lossless rank extractor $\mathcal{E}$ over $\mathbb{F}$ with $L = rk - \binom{r+1}{2}$ (later improved to $r(k-r)$ in [FSS14, For14]). Their proof uses the polynomial method and proceeds as follows. First, they construct the symbolic matrix $E(x) = (E_{ij}(x))_{i\in[r],j\in[k]}$ with entries

$$E_{ij}(x) = (\gamma^{i-1}x)^{j-1},$$

where $\gamma \in \mathbb{F}^\times$ is chosen such that $1, \gamma, \gamma^2, \ldots, \gamma^{k-1}$ are distinct. They then show that for any full-rank matrix $M \in \mathbb{F}^{k\times r}$, we have $\det(E(x)M) \neq 0$ symbolically. It is also straightforward to bound the degree of the polynomial $\det(E(x)M) \in \mathbb{F}[x]$ by $rk - \binom{r+1}{2}$. Thus, by the polynomial method, we have $\det(E(a)M) \neq 0$, or equivalently $\mathrm{rank}(E(a)M) = r$, for all but at most $rk - \binom{r+1}{2}$ elements $a \in \mathbb{F}$. It therefore suffices to define $\mathcal{E} = \{E(a) : a \in S\}$ for a finite set $S \subseteq \mathbb{F}$ of size greater than L.

One reason this construction requires a large field size is that the analysis uses Vandermonde determinants associated with elements in the set $\{1, \gamma, \gamma^2, \ldots, \gamma^{k-1}\} \subseteq \mathbb{F}^\times$, and these elements must be distinct for the Vandermonde determinants to be nonzero. In particular, this forces $|\mathbb{F}| > k$.

Our first observation is that, since we are already working with the polynomial ring $\mathbb{F}[x]$, it is not necessary to use distinct values from the field $\mathbb{F}$; instead, we can use elements of $\mathbb{F}[x]$. Indeed, an earlier construction of Gabizon and Raz [GR08] also uses the polynomial method, but defines the symbolic matrix $E = (E_{ij})$ by $E_{ij} = x^{ij}$. This construction uses no constants from $\mathbb{F}$ other than 1, and instead relies on monomials of different degrees. Roughly speaking, this suggests that we can trade distinct constants for elements of different degrees.

However, another issue remains. If we directly adopt the construction of Gabizon and Raz over a small field $\mathbb{F}_q$, then the resulting polynomial $\det(E(x)M)$ has degree much larger than q, and may vanish at every evaluation point in $\mathbb{F}_q$.

Function fields. To address this issue, we draw inspiration from algebraic geometry (AG) codes, whose analysis can be viewed as a function-field generalization of the polynomial method. Geometrically, the idea is to replace a line by an *algebraic curve*, which, even over a small field $\mathbb{F}_q$, can have many evaluation points $N \gg q$. Algebraically, this corresponds to working in a function field. We adopt this approach in this paper.

Specifically, we show that a function-field analog of the Gabizon–Raz construction [GR08] can be carried out over finite fields of size (at least) $\mathrm{poly}(r)$. This already implies a weaker version of Theorem 1.3, although the upper bound L on the number of bad matrices is $O(r^2k)$, as in [GR08], rather than $O(rk)$. This construction and its analysis are presented in Section 3.

We then revisit the Forbes–Shpilka construction [FS12] and show that a function-field analog can also be carried out over finite fields of size (at least) $\mathrm{poly}(r)$, leading to a full proof of Theorem 1.3. The idea is that when $q \ll k$, although we cannot find $\gamma \in \mathbb{F}_q^\times$ such that $1, \gamma, \ldots, \gamma^{k-1}$ are all distinct, we can instead use elements of the form

$$g_1, \gamma g_1, \ldots, \gamma^{q-2}g_1,\ g_2, \gamma g_2, \ldots, \gamma^{q-2}g_2,\ \cdots,\ g_h, \gamma g_h, \ldots, \gamma^{q-2}g_h,$$

for suitable transcendental elements $g_1, \ldots, g_h$ in the function field, where $h \approx \frac{k}{q-1}$. The use of these transcendental elements increases the degree of $\det(E(x)M)$. However, for sufficiently large

$q \geq \mathrm{poly}(r)$ and carefully chosen $g_1, \ldots, g_h$, this increase is relatively small and dominated by other terms. This construction and its analysis are presented in Section 4.

As a side remark, while it is well known that algebraic curves and function fields underlie AG codes, and prior works such as [ABDN24, BT26] also use AG codes, our proof does not use AG codes as a black box. Instead, we directly carry out analogs of the analyses in [GR08, FS12] using the function-field generalization of the polynomial method.

Field reduction. Next, we use a field-reduction technique to obtain explicit lossless rank dispersers over very small fields such as $\mathbb{F}_2$, thereby proving Theorem 1.7. To do so, we first choose an extension field $\mathbb{F}_Q$ of $\mathbb{F}_q$ of size $\mathrm{poly}(r)$, and apply Theorem 1.3 to obtain an explicit lossless rank extractor (and hence disperser) $\mathcal{E}$ over $\mathbb{F}_Q$.

We then transform each matrix $E_i \in \mathbb{F}_Q^{r \times k}$ into a collection $S(E_i)$ of matrices over $\mathbb{F}_q$. Fix a basis $e_1, \ldots, e_d$ of $\mathbb{F}_Q$ over $\mathbb{F}_q$. Each row u of E_i can be written as a linear combination of d row vectors $u_1, \ldots, u_d$ over $\mathbb{F}_q$, and we replace u by one of the u_j for $j \in [d]$. Since there are d choices for each row, this yields $d^r = 2^{\widetilde{O}(r)}$ matrices over $\mathbb{F}_q$, which form the collection $S(E_i)$. We then show that $\bigcup_{E_i \in \mathcal{E}} S(E_i)$ is a lossless rank disperser over $\mathbb{F}_q$, yielding Theorem 1.7.

Addressing prime fields via PIT. Note that Theorem 1.3 requires q to be non-prime. This is due to the need to choose a suitable function field. Specifically, while a function field F provides many evaluation points $N(F) \gg q$, it also introduces an error term depending on its genus $g(F)$. One therefore wants $N(F)/g(F)$ to be large, so that the benefit outweighs the cost.

When q is non-prime, explicit function fields with $N(F)/g(F) \geq q^c$ for some constant $c > 0$ are known. In particular, when q is square, we use the Garcia–Stichtenoth tower [GS96], which attains the optimal Drinfeld–Vlăduţ bound $\sqrt{q} - 1$ [VD83]. When $q = p^d$ with d odd, we instead use the Bassa–Beelen–Garcia–Stichtenoth tower [BBGS15], which similarly guarantees $N(F)/g(F) \geq q^c$.

This leaves the case where $q \geq \mathrm{poly}(r)$ is prime. In this case, Serre [Ser20] used class field towers to show the existence of infinite families of function fields F over $\mathbb{F}_q$ with $N(F)/g(F) \geq c \log q$ for some constant $c > 0$. This result was suggested in [BT26] as a basis for constructing strong s-blocking sets. However, class field towers are generally not regarded as explicit objects; see, e.g., [Sti01].

We develop a new approach, based on polynomial identity testing (PIT), that reduces the prime case to the non-prime prime power case, and thereby yields explicit lossless rank extractors over $\mathbb{F}_q$ for prime $q \geq \mathrm{poly}(r)$. We begin by taking $Q = q^2$ and constructing an explicit lossless rank extractor $\mathcal{E}$ over $\mathbb{F}_Q$ via Theorem 1.3. Starting from $\mathcal{E}$, one can apply the field-reduction procedure described above. In this procedure, each row $u \in \mathbb{F}_Q^k$ is written as $u = u_1 e_1 + u_2 e_2$ with respect to a basis $\{e_1, e_2\}$ of $\mathbb{F}_Q/\mathbb{F}_q$, and one replaces u by either u_1 or u_2. As the choices are made independently for each of the r rows, a direct application of field reduction increases the size by a factor $[\mathbb{F}_Q : \mathbb{F}_q]^r = 2^r$.

To avoid this exponential blow-up, let $u^{(i)} \in \mathbb{F}_Q^k$ denote the i-th row. Write

$$u^{(i)} = u_1^{(i)} e_1 + u_2^{(i)} e_2,$$

where $u_1^{(i)}, u_2^{(i)} \in \mathbb{F}_q^k$. Instead of choosing between $u_1^{(i)}$ and $u_2^{(i)}$, we replace $u^{(i)}$ by a linear combination $a^{(i)} u_1^{(i)} + b^{(i)} u_2^{(i)}$, where $a^{(i)}, b^{(i)} \in \mathbb{F}_q$. For $q \geq \mathrm{poly}(r)$, a random choice of these coefficients is likely to preserve the nonzeroness of the determinant. The problem is therefore to construct a small explicit set of such choices that contains many good ones.

We show that this problem reduces to deterministic black-box PIT for a class of polynomials of the form $\det(A(x))$, where $A(x) = \sum_i A_i x_i$ and each A_i has rank at most one. Gurjar and

Thierauf [GT20] showed that this class admits quasi-polynomial-size hitting sets. We adapt their construction to ensure that evaluation points lie in $\mathbb{F}_q^k$ for sufficiently large $q \geq \text{poly}(r)$, yielding Theorem 1.5. Our method can be viewed as a PIT-based field-reduction technique that reduces the exponential blow-up to quasi-polynomial.

Weak subspace designs and strong s-blocking sets. Our results on weak subspace designs (Theorems 1.8 to 1.10) follow from Theorems 1.3, 1.5 and 1.7 via the equivalence between lossless rank extractors and weak subspace designs (Lemma 2.3; see also [FG14, Proposition 6.1]).

Our results on strong s-blocking sets (Theorems 1.12 to 1.14) use a similar connection, via a reduction to the weak subspace design/lossless rank disperser property ([FS16, Proposition 10]; see also Lemma 6.1).

Finally, Theorem 1.15 provides a different construction based on ε-biased sets and Fourier analysis. The key point is that ε-biased sets are pseudorandom against functions whose Fourier transform has small L_1 norm, while the strong s-blocking set property can be expressed using indicator functions of subspaces and their differences, which have small L_1 norms. This gives a reduction from constructing explicit strong s-blocking sets to constructing explicit ε-biased sets. Plugging in the explicit constructions of ε-biased sets from [TS17, JM21] then yields explicit strong s-blocking sets.

2 Preliminaries and Notation

Define $[n] := \{1, \ldots, n\}$. Denote by $\text{Sym}(n)$ the symmetric group on $[n]$, i.e., the group of all permutations of $[n]$. Denote by $R^{r\times k}$ the set of $r \times k$ matrices over a commutative ring R. For a subspace $V \subseteq \mathbb{F}^k$ over a field $\mathbb{F}$, denote by $V^\perp$ its orthogonal complement with respect to the standard inner product. For a matrix M, let $\text{rowspan}(M)$ and $\text{colspan}(M)$ denote its row space and column space, respectively. The projective space of dimension n over a finite field $\mathbb{F}_q$, denoted by $\text{PG}(n, q)$, is the set of all one-dimensional subspaces of $\mathbb{F}_q^{n+1}$.

Let $r \leq k$. For an $r \times k$ matrix A and a subset $I \subseteq [k]$ of size r, let A_I denote the $r \times r$ submatrix of A consisting of the columns indexed by I. Similarly, for a $k \times r$ matrix B and a subset $J \subseteq [k]$ of size r, let B^J denote the $r \times r$ submatrix of B consisting of the rows indexed by J.

We will use the following classical *Cauchy–Binet formula*, which expresses the determinant of a product of rectangular matrices as a sum over products of minors. For a proof, see, e.g., [For14, Appendix B.3].

Lemma 2.1 (Cauchy–Binet formula)**.** *Let A be an $r \times k$ matrix and B be a $k \times r$ matrix over a commutative ring, with $r \leq k$. Then*

$$\det(AB) = \sum_{I\subseteq[k],|I|=r} \det(A_I)\,\det(B^I).$$

Following [GR08, FS12], we will also need a statement asserting the unique optimality of a maximal minor of a full-rank matrix.

Lemma 2.2. *Let $M \in \mathbb{F}^{r\times k}$ be a full-rank matrix over a field $\mathbb{F}$, with $r \leq k$. Let $I \subseteq [k]$ be the subset produced by the following procedure. Initialize $I = \emptyset$. For $i = k, k-1, \ldots, 1$, add i to I if the i-th column of M does not lie in the $\mathbb{F}$-linear span of the columns indexed by the current set I.*

Then $|I| = r$ and M_I is nonsingular. Moreover, for every $J \subseteq [k]$ of size r such that M_J is nonsingular, the following hold:

1. *$i_\ell \geq j_\ell$ for all $\ell \in [r]$, where i_ℓ and j_ℓ denote the ℓ-th smallest elements of I and J, respectively.*

2. *Let $w_1 > w_2 > \cdots > w_k$ be integers. Then*

$$\sum_{i \in I} w_i \leq \sum_{i \in J} w_i,$$

with equality if and only if $I = J$.

Proof. Let c_i denote the i-th column of M. By construction, each selected column lies outside the span of the previously selected ones, so the columns indexed by I are linearly independent. If they did not span $\mathbb{F}^r$, then some column of M would lie outside their span, and would have been added when considered, a contradiction. Thus $|I| = r$ and M_I is nonsingular.

Let $I = \{i_1 < \cdots < i_r\}$ and $J = \{j_1 < \cdots < j_r\}$, where M_J is nonsingular. We prove that $i_\ell \geq j_\ell$ for all ℓ.

Suppose not, and let ℓ be the largest index such that $i_\ell < j_\ell$. Then for all $t > \ell$, we have $i_t \geq j_t > j_\ell$. Hence when the algorithm considers j_ℓ, the indices already chosen are exactly $i_{\ell+1}, \ldots, i_r$.

Assume that j_ℓ is not selected. Then $c_{j_\ell} \in \operatorname{span}(c_{i_{\ell+1}}, \ldots, c_{i_r})$. For each $t > \ell$, the index j_t is processed before j_ℓ. If j_t is selected, then $j_t \in \{i_{\ell+1}, \ldots, i_r\}$; otherwise, c_{j_t} lies in the span of $c_{i_{\ell+1}}, \ldots, c_{i_r}$. In either case,

$$c_{j_t} \in \operatorname{span}(c_{i_{\ell+1}}, \ldots, c_{i_r}).$$

Thus

$$\operatorname{span}(c_{j_{\ell+1}}, \ldots, c_{j_r}) \subseteq \operatorname{span}(c_{i_{\ell+1}}, \ldots, c_{i_r}).$$

Both spaces have dimension $r - \ell$, since the corresponding sets of columns are linearly independent. Hence the two spans are equal, and therefore

$$c_{j_\ell} \in \operatorname{span}(c_{j_{\ell+1}}, \ldots, c_{j_r}),$$

contradicting the linear independence of the columns indexed by J. Thus j_ℓ must be selected, contradicting $i_\ell < j_\ell$. This proves $i_\ell \geq j_\ell$ for all ℓ.

Finally, if $w_1 > \cdots > w_k$, then $i_\ell \geq j_\ell$ implies $w_{i_\ell} \leq w_{j_\ell}$ for each $\ell \in [r]$, and hence

$$\sum_{i \in I} w_i \leq \sum_{i \in J} w_i.$$

Equality holds only if $w_{i_\ell} = w_{j_\ell}$ for all $\ell \in [r]$, which (since $w_1 > \cdots > w_k$) implies $i_\ell = j_\ell$ for all $\ell \in [r]$, i.e., $I = J$. □

Recall the definition of lossless rank extractors and that of weak subspace designs. The following lemmas show a connection between the two notions. It is similar to [FG14, Proposition 6.1] except that we consider both primal and dual forms and require the matrices to have full rank.

Lemma 2.3. *Let $r \leq k$. Suppose $\mathcal{E} \subseteq \mathbb{F}^{r \times k}$ is a finite collection of matrices, each of rank r. Define*

$$\mathcal{V}(\mathcal{E}) = \{\operatorname{rowspan}(E) : E \in \mathcal{E}\} \quad \text{and} \quad \mathcal{V}^\perp(\mathcal{E}) = \{\operatorname{rowspan}(E)^\perp : E \in \mathcal{E}\}.$$

Then every subspace in $\mathcal{V}(\mathcal{E})$ has dimension r, and every subspace in $\mathcal{V}^\perp(\mathcal{E})$ has dimension $k - r$.

Moreover, $\mathcal{E}$ is a (k, r, L) lossless rank extractor if and only if $\mathcal{V}(\mathcal{E})$ is a $(k - r, L)$ weak subspace design, which holds if and only if $\mathcal{V}^\perp(\mathcal{E})$ is an (r, L) weak subspace design.

Proof. Since each $E \in \mathcal{E}$ has rank r, every $V = \mathrm{rowspan}(E)$ has dimension r, and thus every $V^\perp$ has dimension $k - r$.

For the main claim, fix a full-rank $M \in \mathbb{F}^{k\times r}$ and let

$$W := \mathrm{colspan}(M)^\perp, \quad \dim W = k - r.$$

Then

$$\mathrm{rank}(EM) < r \iff \mathrm{colspan}(M) \cap \ker(E) \neq \{0\} \iff W^\perp \cap V^\perp \neq \{0\} \iff V \cap W \neq \{0\},$$

where $V = \mathrm{rowspan}(E)$, using $\ker(E) = V^\perp$ and dimension counting.

Thus, for each M, the matrices E with $\mathrm{rank}(EM) < r$ are exactly those $V \in \mathcal{V}(\mathcal{E})$ with $V \cap W \neq \{0\}$. Since every $(k-r)$-dimensional W arises this way, $\mathcal{E}$ is a (k, r, L) lossless rank extractor iff $\mathcal{V}(\mathcal{E})$ is a $(k - r, L)$ weak subspace design.

Finally, for any r-dimensional subspace U and any $V \in \mathcal{V}(\mathcal{E})$,

$$V^\perp \cap U \neq \{0\} \iff V \cap U^\perp \neq \{0\},$$

since $\dim V = r$ and $\dim U^\perp = k - r$. Thus, the number of $V^\perp \in \mathcal{V}^\perp(\mathcal{E})$ intersecting U nontrivially equals the number of $V \in \mathcal{V}(\mathcal{E})$ intersecting $U^\perp$ nontrivially. As $U^\perp$ ranges over all $(k-r)$-dimensional subspaces, the claim follows. □

In general, the matrices in a lossless rank extractor need not have full rank. However, by discarding the rank-deficient matrices, we can ensure full rank at the cost of losing at most L elements.

Lemma 2.4. *Let $r \leq L$. Suppose $\mathcal{E}_0 \subseteq \mathbb{F}^{r\times k}$ is a (k, r, L) lossless rank extractor of size $n \geq L$. Let*

$$\mathcal{E} = \{E \in \mathcal{E}_0 : \mathrm{rank}(E) = r\}.$$

Then $\mathcal{E}$ is a (k, r, L) lossless rank extractor of size at least $n - L$ consisting only of rank-r matrices. Consequently, $\mathcal{V}(\mathcal{E})$ and $\mathcal{V}^\perp(\mathcal{E})$ as defined in Lemma 2.3 form a $(k-r, L)$ weak subspace design and an (r, L) weak subspace design, respectively, both of size at least $n - L$.

Proof. Fix any full-rank $M \in \mathbb{F}^{k\times r}$. If $\mathrm{rank}(E) < r$, then $\mathrm{rank}(EM) \leq \mathrm{rank}(E) < r$. Since $\mathcal{E}_0$ is a (k, r, L) lossless rank extractor, there are at most L such matrices E in $\mathcal{E}_0$. Hence $|\mathcal{E}| \geq n - L$.

Removing these matrices does not increase the number of bad matrices for any M, so $\mathcal{E}$ remains a (k, r, L) lossless rank extractor. The rest follows from Lemma 2.3. □

2.1 Preliminaries on Function Fields

We recall basic facts about finitely generated function fields of transcendence degree one over a finite field $\mathbb{F}_q$. The theory of such function fields is equivalent to that of nonsingular irreducible projective curves over $\mathbb{F}_q$, via a well-known correspondence that is functorial up to isomorphism [Har77, Section 1.6]. In this paper, we adopt the algebraic language of function fields. For a comprehensive treatment, see [Sti09].

Throughout this subsection, let F be a finitely generated function field of transcendence degree one over $\mathbb{F}_q$.

Discrete valuations. A *discrete valuation* on F (trivial on $\mathbb{F}_q$) is a map $v : F \to \mathbb{Z} \cup \{\infty\}$ such that for all $x, y \in F$:

1. $v(xy) = v(x) + v(y)$;
2. $v(x + y) \geq \min\{v(x), v(y)\}$;
3. $v(x) = \infty \iff x = 0$;
4. $v(x) = 0$ for all $x \in \mathbb{F}_q^\times$.[2]

We say v is *normalized* if $v(F) = \mathbb{Z}$.

When $v(x) \neq v(y)$, the second condition is in fact an equality:

Lemma 2.5 (Strong triangle inequality)**.** *If $v(x) \neq v(y)$, then $v(x + y) = \min\{v(x), v(y)\}$.*

Proof. By symmetry, we may assume $v(x) < v(y)$. We have $v(x + y) \geq \min\{v(x), v(y)\} = v(x)$. Also,

$$v(x) = v((x + y) - y) \geq \min\{v(x + y), v(-y)\} = \min\{v(x + y), v(y)\}. \tag{4}$$

Since $v(x) < v(y)$, the minimum on the right-hand side of (4) cannot be $v(y)$; hence it must be $v(x+y)$. So (4) becomes $v(x) \geq v(x+y)$. Combining with $v(x+y) \geq v(x)$ yields $v(x+y) = v(x)$. □

Valuation rings. Given a discrete valuation v on F, the associated *valuation ring* is

$$\mathcal{O}_v = \{x \in F : v(x) \geq 0\}.$$

It is a local ring, i.e., a ring with a unique maximal ideal. This maximal ideal is

$$\mathfrak{m}_v = \{x \in F : v(x) > 0\}.$$

An element $x \in \mathcal{O}_v$ is a unit if and only if $v(x) = 0$. The *residue field* of v is defined as $\kappa_v = \mathcal{O}_v/\mathfrak{m}_v$.

Places. A *place* P of F consists of a normalized discrete valuation $v_P : F \to \mathbb{Z} \cup \{\infty\}$ that is trivial on $\mathbb{F}_q$. We denote by $\mathcal{P}_F$ the set of all places of F. For a place $P \in \mathcal{P}_F$, we write $\mathcal{O}_P$ for its valuation ring, $\mathfrak{m}_P$ for its maximal ideal, and $\kappa_P = \mathcal{O}_P/\mathfrak{m}_P$ for the corresponding residue field. The degree of P is defined as $\deg(P) = [\kappa_P : \mathbb{F}_q]$. A place P is called *rational* if $\deg(P) = 1$, or equivalently if $\kappa_P = \mathbb{F}_q$. For such a place P, the *evaluation* of a function $f \in \mathcal{O}_P$ at P, denoted by $f(P) \in \mathbb{F}_q$, is defined as the image of f in the residue field κ_P.

Divisors. A *divisor* on F is a formal integer linear combination of places,

$$D = \sum_{P \in \mathcal{P}_F} n_P P,$$

where all but finitely many coefficients n_P are zero. The degree of D is $\deg(D) = \sum_{P \in \mathcal{P}_F} n_P \deg(P)$. We call D *effective*, and write $D \geq 0$, if $n_P \geq 0$ for all $P \in \mathcal{P}_F$.

The set of divisors on F forms a free abelian group $\mathrm{Div}(F)$ generated by the places of F. The degree map $\deg(\cdot) : \mathrm{Div}(F) \to \mathbb{Z}$ is a group homomorphism.

[2] This condition is redundant: it follows from $v(xy) = v(x)+v(y)$ and the finiteness of $\mathbb{F}_q^\times$, but we state it explicitly for clarity.

For a nonzero function $f \in F^\times$, the associated *principal divisor* is

$$(f) = \sum_{P \in \mathcal{P}_F} v_P(f)\, P,$$

where v_P is the normalized discrete valuation corresponding to the place P. The degree of a principal divisor (f) is always zero. Intuitively, this means that the sum of the orders of the zeros of f equals the sum of the orders of its poles.

Riemann–Roch spaces. For a divisor D, we define the associated Riemann–Roch space

$$\mathcal{L}(D) = \{f \in F^\times \cup \{0\} : (f) + D \geq 0\}.$$

It is a finite-dimensional $\mathbb{F}_q$-vector space, and we denote its dimension by $\ell(D) = \dim_{\mathbb{F}_q} \mathcal{L}(D)$.

We need the following lemma, which is a function-field generalization of the fact that a nonzero polynomial of degree d cannot have more than d zeros.

Lemma 2.6. *$\ell(D) = 0$ if $\deg(D) < 0$.*

Proof. Assume to the contrary that $\ell(D) > 0$, i.e., $\mathcal{L}(D)$ contains a nonzero function $f \in F^\times$. Then $(f) + D \geq 0$. So $\deg((f) + D) \geq 0$. On the other hand, $\deg((f) + D) = \deg((f)) + \deg(D) = \deg(D)$ since the degree of the principal divisor (f) is zero. But this is impossible since $\deg(D) < 0$. □

The following important theorem gives an estimate on the dimension of a Riemann–Roch space.

Theorem 2.7 (Riemann–Roch theorem)**.** *There exist an integer $g \geq 0$ and a divisor K on F such that for every divisor D on F,*

$$\ell(D) = \deg(D) + 1 - g + \ell(K - D).$$

The integer g is called the genus *of F, and K is called a* canonical divisor*. Moreover,* $\deg(K) = 2g - 2$.

As $\ell(K - D) = \dim_{\mathbb{F}_q} \mathcal{L}(K - D) \geq 0$, we have the following corollary.

Corollary 2.8 (Riemann's inequality)**.** *$\ell(D) \geq \deg(D) + 1 - g$.*

This inequality is in fact an equality when the degree of D is sufficiently large:

Corollary 2.9. *Suppose $\deg(D) \geq 2g - 1$. Then $\ell(D) = \deg(D) + 1 - g$.*

Proof. We have $\deg(K - D) = \deg(K) - \deg(D) = (2g - 2) - \deg(D) < 0$. So $\ell(K - D) = 0$ by Lemma 2.6. The claim then follows from the Riemann–Roch theorem. □

Corollary 2.9 further yields the following corollary, which guarantees the existence of elements in a Riemann–Roch space with a prescribed order of zero or pole at a given place.

Corollary 2.10. *Let P be a place of F and D be a divisor on F. Let n be the coefficient of P in D. Suppose $\deg(D - P) \geq 2g - 1$. Then there exists $f \in \mathcal{L}(D)$ such that $v_P(f) = -n$.*

Proof. Applying Corollary 2.9 to D and $D - P$ shows that $\ell(D) = \deg(D) + 1 - g$ and $\ell(D - P) = \deg(D) + 1 - g - \deg(P) = \ell(D) - \deg(P) < \ell(D)$. So there exists $f \in \mathcal{L}(D)$ that is not in $\mathcal{L}(D - P)$. By definition, we have $v_P(f) + n \geq 0$ but $v_P(f) + (n - 1) < 0$. So $v_P(f) = -n$. □

Finally, we record the following corollary of Riemann's inequality on the existence of functions with distinct valuations at a given place.

Corollary 2.11. *Let P be a rational place of F. Let n be a nonnegative integer and $d = n-1+g$. Then there exist $f_1, f_2, \ldots, f_n \in \mathcal{L}(dP)$ such that*

$$0 \geq v_P(f_1) > v_P(f_2) > \cdots > v_P(f_n) \geq -d.$$

Proof. For $t \geq 0$, let $V_t = \mathcal{L}((d-t)P)$. Then

$$V_0 \supseteq V_1 \supseteq \cdots \supseteq V_{d+1} = \{0\}, \tag{5}$$

where the last equality holds by Lemma 2.6. By Riemann's inequality (Corollary 2.8), we have $\dim_{\mathbb{F}_q} V_0 = \ell(dP) \geq d+1-g = n$.

Moreover, $\dim_{\mathbb{F}_q} V_t/V_{t+1} \leq 1$ for every t. Indeed, let $u \in F^\times$ satisfy $v_P(u) = 1$. Then the map

$$V_t \to \kappa_P, \qquad f \mapsto (u^{d-t} f)(P),$$

is well-defined, and its kernel is exactly V_{t+1}. Hence V_t/V_{t+1} embeds into $\kappa_P \cong \mathbb{F}_q$, so it has dimension at most one over $\mathbb{F}_q$.

As $\dim_{\mathbb{F}_q} V_t$ drops from at least n to 0 along the chain (5), and each step decreases the dimension by at most 1, there are at least n indices $0 \leq t_1 < \cdots < t_n \leq d$ such that $\dim V_{t_i}/V_{t_i+1} = 1$.

For each $i \in [n]$, choose $f_i \in V_{t_i} \setminus V_{t_i+1}$. Then $f_i \in \mathcal{L}(dP)$ and $v_P(f_i) = -t_i$. Since $0 \leq t_1 < \cdots < t_n \leq d$, we obtain

$$0 \geq v_P(f_1) > v_P(f_2) > \cdots > v_P(f_n) \geq -d.$$

This proves the claim. □

2.2 Explicit Function Field Towers

Our constructions use explicit function field towers $F_1 \subseteq F_2 \subseteq \cdots$ that are *asymptotically good*, meaning that the ratio N_i/g_i between the number N_i of rational places of F_i and its genus g_i is bounded away from zero by a positive constant. Moreover, as q varies, the towers we use satisfy $N_i/g_i = \Omega(q^c)$ for some constant $c > 0$.

The Garcia–Stichtenoth tower. The Garcia–Stichtenoth tower constructed in [GS96] is a well-known sequence of function fields over finite fields with particularly favorable asymptotic properties. An earlier asymptotically optimal tower was constructed in [GS95]. We briefly recall the definition and basic properties of the tower from [GS96], which will play a central role in our constructions.

Definition 2.12 (Garcia–Stichtenoth tower [GS96])**.** Let q be a square, and write $q = \ell^2$ for some prime power ℓ. The Garcia–Stichtenoth tower over $\mathbb{F}_q$ is the sequence of function fields

$$F_1 \subseteq F_2 \subseteq \cdots$$

defined recursively as follows. Let $F_1 = \mathbb{F}_q(x_1)$ be the rational function field, and for $i \geq 1$, define

$$F_{i+1} = F_i(x_{i+1}), \qquad x_{i+1}^\ell + x_{i+1} = \frac{x_i^\ell}{x_i^{\ell-1}+1}.$$

The following lemma gives estimates for the number of rational places and the genus of the fields in the Garcia–Stichtenoth tower.

Lemma 2.13 ([GS96])**.** *For $i \geq 1$, let N_i and g_i denote the number of rational places and the genus of the i-th field F_i in the Garcia–Stichtenoth tower over $\mathbb{F}_q$, where $q = \ell^2$. Then*

$$[F_i : F_1] = \ell^{i-1},$$
$$N_i \geq \ell^i(\ell - 1) + 1,$$

and

$$g_i = \begin{cases} (\ell^{i/2} - 1)^2, & \text{if } i \text{ is even,} \\ (\ell^{(i+1)/2} - 1)(\ell^{(i-1)/2} - 1), & \text{if } i \text{ is odd.} \end{cases}$$

In particular, $g_i \leq \ell^i$.

In particular, we have

$$\limsup_{i \to \infty} \frac{N_i}{g_i} \geq \ell - 1 = \sqrt{q} - 1,$$

which matches the largest possible asymptotic ratio between the number of rational places and the genus of a function field over $\mathbb{F}_q$, known as the Drinfeld–Vlăduţ bound [VD83].

The Bassa–Beelen–Garcia–Stichtenoth tower. The Garcia–Stichtenoth tower is defined over $\mathbb{F}_q$ only when q is a square. When $q = p^{2m+1}$ with $m \geq 1$, we instead use the Bassa–Beelen–Garcia–Stichtenoth tower [BBGS15]. We recall below a special case of this construction.

Definition 2.14 (Bassa–Beelen–Garcia–Stichtenoth tower [BBGS15])**.** Let $q = p^{2m+1}$ be a prime power with $m \geq 1$, where p is a prime. Define

$$j = \begin{cases} m & \text{if } p \nmid m, \\ m+1 & \text{otherwise,} \end{cases} \qquad k = 2m + 1 - j \in \{m, m+1\}.$$

For an integer $a \geq 1$, define

$$\mathrm{Tr}_a(T) := T + T^p + \cdots + T^{p^{a-1}} \in \mathbb{F}_q[T].$$

The Bassa–Beelen–Garcia–Stichtenoth tower over $\mathbb{F}_q$ is the sequence of function fields

$$E_1 \subseteq E_2 \subseteq \cdots$$

defined recursively as follows. Let $E_1 = \mathbb{F}_q(x_1)$ be the rational function field, and for $i \geq 1$, define

$$E_{i+1} = E_i(x_{i+1}), \qquad \mathrm{Tr}_j\left(\frac{x_{i+1}}{x_i^{p^k}}\right) + \mathrm{Tr}_k\left(\frac{x_{i+1}^{p^j}}{x_i}\right) = 1.$$

The following lemma gives estimates for the number of rational places and the genus of the fields in the Bassa–Beelen–Garcia–Stichtenoth tower.

Lemma 2.15 ([BBGS15])**.** *Let $q = p^{2m+1}$ be a prime power with $m \geq 1$, where p is a prime. For $i \geq 1$, let N_i and g_i denote the number of rational places and the genus of the i-th field E_i in the Bassa–Beelen–Garcia–Stichtenoth tower over $\mathbb{F}_q$. Then*

$$[E_i : E_1] = p^{2m(i-1)},$$
$$N_i \geq [E_i : E_1]\,(q - 1) = p^{2m(i-1)}(q - 1),$$

and

$$g_i \leq \frac{[E_i : E_1]}{2}\left(\frac{q-1}{p^m - 1} + \frac{q-1}{p^{m+1} - 1}\right) = \frac{p^{2m(i-1)}}{2}\left(\frac{q-1}{p^m - 1} + \frac{q-1}{p^{m+1} - 1}\right).$$

In particular, we have

$$\limsup_{i\to\infty} \frac{N_i}{g_i} \geq 2\left(\frac{1}{p^m-1} + \frac{1}{p^{m+1}-1}\right)^{-1}.$$

This does not match the Drinfeld–Vlăduţ bound [VD83], which in this case is $\sqrt{q}-1 = p^{m+1/2}-1$, but it is polynomially related to it.

Remark 2.16. *Our constructions based on function fields require the ability to compute bases of Riemann–Roch spaces $\mathcal{L}(dP_\infty)$ associated with a distinguished rational place P_∞, as well as to evaluate such functions at rational places.*

For the function field towers we use, namely the Garcia–Stichtenoth and Bassa–Beelen–Garcia–Stichtenoth towers, these operations can be carried out in polynomial time using standard algorithms from computational algebraic function field theory; see, e.g., [Hes02, SAK+02].

3 A Function-Field Analog of the Gabizon–Raz Construction

In this section, we present a construction that can be viewed as a function-field analog of the construction of Gabizon and Raz [GR08].

Construction. Let $1 \leq r \leq k$ be integers. Let F be a finitely generated function field of transcendence degree one over a finite field $\mathbb{F}_q$, and let g be its genus. Let P_∞ be the distinguished rational place of F, and let v_{P_∞} be its normalized valuation. Let S be the set of rational places of F other than P_∞.

For each $i \in [r]$ and $j \in [k]$, define

$$d_{ij} = i \cdot (j-1) + 2g \geq 2g,$$

and choose $f_{ij} \in \mathcal{L}(d_{ij}P_\infty)$ such that $v_{P_\infty}(f_{ij}) = -d_{ij}$. The existence of such a function follows from Corollary 2.10 applied to $D = d_{ij}P_\infty$, since $\deg(d_{ij}P_\infty - P_\infty) = d_{ij} - 1 \geq 2g-1$.

Construct the $r \times k$ matrix E over F by

$$E = (f_{ij})_{i\in[r],\, j\in[k]}.$$

For each rational place $P \in S$, define the $r \times k$ matrix $E(P)$ over $\mathbb{F}_q$ by

$$E(P) = (f_{ij}(P))_{i\in[r],\, j\in[k]}.$$

Analysis. The following lemma establishes a nonvanishing property of certain determinants.

Lemma 3.1. *Let $M \in \mathbb{F}_q^{k\times r}$ be a full-rank matrix. Then $\det(EM) \in F$ is nonzero, and*

$$v_{P_\infty}(\det(EM)) \geq -r\left(\frac{(r+1)(3k-r-2)}{6} + 2g\right).$$

Proof. By the Cauchy–Binet formula (Lemma 2.1),

$$\det(EM) = \sum_{I\subseteq[k],\,|I|=r} \det(E_I)\det(M^I).$$

Fix $I = \{j_1 < \cdots < j_r\} \subseteq [k]$. Expanding $\det(E_I)$, we write

$$\det(E_I) = \sum_{\sigma \in \mathrm{Sym}(r)} T_\sigma, \quad T_\sigma := \mathrm{sgn}(\sigma) \prod_{i=1}^{r} f_{ij_{\sigma(i)}}.$$

We claim that the identity permutation uniquely minimizes $v_{P_\infty}(T_\sigma)$. Indeed,

$$v_{P_\infty}(T_\sigma) = \sum_{i=1}^{r} v_{P_\infty}(f_{ij_{\sigma(i)}}) = -\sum_{i=1}^{r} d_{ij_{\sigma(i)}} = -\sum_{i=1}^{r} \big(i \cdot (j_{\sigma(i)} - 1) + 2g\big).$$

If σ has an inversion (i, i'), i.e., $i < i'$ but $\sigma(i) > \sigma(i')$, then $j_{\sigma(i)} > j_{\sigma(i')}$. Let σ' be obtained from σ by swapping $\sigma(i)$ and $\sigma(i')$. Then

$$v_{P_\infty}(T_{\sigma'}) - v_{P_\infty}(T_\sigma) = -\big(i\, j_{\sigma(i')} + i' j_{\sigma(i)}\big) + \big(i\, j_{\sigma(i)} + i' j_{\sigma(i')}\big) = (i' - i)\big(j_{\sigma(i')} - j_{\sigma(i)}\big) < 0,$$

since $i' < i$ is false and $j_{\sigma(i')} - j_{\sigma(i)} < 0$. Thus swapping strictly decreases the valuation. Hence the identity is the unique minimizer.

By the strong triangle inequality, it follows that

$$v_{P_\infty}(\det(E_I)) = -\sum_{i=1}^{r} \big(i \cdot (j_i - 1) + 2g\big) = -\sum_{i=1}^{r} i\, j_i - \sum_{i=1}^{r} (2g - i).$$

Since $\det(M^I) \in \mathbb{F}_q$, we have $v_{P_\infty}(\det(M^I)) \in \{0, \infty\}$, and at least one I satisfies $\det(M^I) \neq 0$ because M has full rank.

Thus the minimum of $v_{P_\infty}(\det(E_I)\det(M^I))$ is attained uniquely by the subset I^* minimizing $-\sum_{i=1}^{r} i\, j_i$ subject to $\det(M^I) \neq 0$. By Lemma 2.2, such I^* is unique.

Let $I^* = \{j_1^* < \cdots < j_r^*\}$. Since the minimum valuation in the above expansion is attained uniquely at I^*, the strong triangle inequality implies that

$$v_{P_\infty}(\det(EM)) = -\sum_{i=1}^{r} i\, j_i^* - \sum_{i=1}^{r} (2g - i) < \infty.$$

In particular, $\det(EM) \neq 0$.

Moreover,

$$v_{P_\infty}(\det(EM)) \geq -\sum_{i=1}^{r} i(k - r + i) - \sum_{i=1}^{r} (2g - i),$$

which simplifies to

$$v_{P_\infty}(\det(EM)) \geq -r\left(\frac{(r+1)(3k - r - 2)}{6} + 2g\right).$$

□

Now we prove the main result of this section.

Theorem 3.2. *Let $M \in \mathbb{F}_q^{k \times r}$ be a full-rank matrix. The number of $P \in S$ such that $E(P)M \in \mathbb{F}_q^{r \times r}$ does not have full rank is at most $r\left(\frac{(r+1)(3k-r-2)}{6} + 2g\right)$.*

Proof. Let

$$d = r\left(\frac{(r+1)(3k-r-2)}{6} + 2g\right).$$

By Lemma 3.1, we have $\det(EM) \neq 0$ and $\det(EM) \in \mathcal{L}(dP_\infty)$. Let $\mathcal{B} \subseteq S$ be the set of places P such that $E(P)M$ does not have full rank. Then for every $P \in \mathcal{B}$, we have $(\det(EM))(P) = \det(E(P)M) = 0$, and hence $\det(EM) \in \mathcal{L}(dP_\infty - \sum_{P\in\mathcal{B}} P)$. Since $\det(EM) \neq 0$, Lemma 2.6 implies that this divisor has nonnegative degree, i.e., $d - |\mathcal{B}| \geq 0$. Thus $|\mathcal{B}| \leq d$, as claimed. □

One can instantiate the function field and the parameters to obtain explicit lossless rank extractors over fields of size at least $\mathrm{poly}(r)$. We defer this to the next section, where we present an alternative construction that improves the $O(r^2k)$ term in Theorem 3.2 to $O(rk)$.

4 A Function-Field Analog of the Forbes–Shpilka Construction

We now present an alternative construction, which can be viewed as a function-field analog of the construction of Forbes and Shpilka [FS12].

Construction. Let $1 \leq r \leq k$ be integers. Let F be a finitely generated function field of transcendence degree one over a finite field $\mathbb{F}_q$, and let g be its genus. Let P_∞ be the distinguished rational place of F, and let v_{P_∞} be its normalized valuation. Let S be the set of rational places of F other than P_∞.

Let $h = \left\lceil \frac{k}{q-1} \right\rceil$. Let $d_1 = k-1+g$ and $d_2 = h-1+g$. By Riemann's inequality (Corollary 2.8), we have $\ell(d_1P_\infty) \geq k$ and $\ell(d_2P_\infty) \geq h$.

By Corollary 2.11, there exist $f_1, \ldots, f_k \in \mathcal{L}(d_1P_\infty)$ and $g_1, \ldots, g_h \in \mathcal{L}(d_2P_\infty)$ such that

$$0 \geq v_{P_\infty}(f_1) > \cdots > v_{P_\infty}(f_k) \geq -d_1 \tag{6}$$

and

$$0 \geq v_{P_\infty}(g_1) > \cdots > v_{P_\infty}(g_h) \geq -d_2. \tag{7}$$

For $j \in [k]$, define $\alpha(j) = \lfloor \frac{j-1}{q-1} \rfloor + 1 \in \{1, \ldots, h\}$ and $\beta(j) = (j-1) \bmod (q-1) \in \{0, \ldots, q-2\}$. Let γ be a generator of $\mathbb{F}_q^\times$.

Construct the $r \times k$ matrix E over F by

$$E_{ij} = (\gamma^{\beta(j)} g_{\alpha(j)})^{i-1} f_j.$$

For each $P \in S$, define

$$E(P) = (E_{ij}(P))_{i\in[r], j\in[k]}.$$

Remark 4.1. *Instead of using Corollary 2.11, which follows from Riemann's inequality (Corollary 2.8), one can use Corollary 2.10, as in Section 3, which relies on the full Riemann–Roch theorem rather than its weaker consequence. This yields functions f_i and g_i with prescribed valuations $v_{P_\infty}(f_i)$ and $v_{P_\infty}(g_i)$, rather than only the bounds in (6) and (7). This increases the g term in the analysis to $2g$, but does not affect the asymptotics up to constant factors.*

Analysis. Next, we prove the following analog of Lemma 3.1 concerning the nonvanishing of determinants.

Lemma 4.2. *Let $M \in \mathbb{F}_q^{k\times r}$ be full rank. Then $\det(EM) \in F$ is nonzero, and*

$$v_{P_\infty}(\det(EM)) \geq -r\left(k - r + \frac{r-1}{2}\left\lceil \frac{k}{q-1}\right\rceil + \frac{r+1}{2}g\right).$$

Proof. By the Cauchy–Binet formula,

$$\det(EM) = \sum_{I\subseteq[k],\,|I|=r} \det(E_I)\det(M^I).$$

Fix $I = \{j_1 < \cdots < j_r\} \subseteq [k]$. The matrix E_I has the form

$$E_I = \left((\gamma^{\beta(j_t)} g_{\alpha(j_t)})^{i-1} f_{j_t}\right)_{i,t\in[r]}.$$

Factoring out f_{j_t} from each column and using multilinearity of the determinant, we obtain

$$\det(E_I) = \left(\prod_{u=1}^{r} f_{j_u}\right) \det\left((\gamma^{\beta(j_t)} g_{\alpha(j_t)})^{i-1}\right)_{i,t\in[r]}.$$

The remaining determinant is a Vandermonde determinant, and hence

$$\det(E_I) = \left(\prod_{u=1}^{r} f_{j_u}\right) \prod_{1\leq t_1<t_2\leq r} \left(\gamma^{\beta(j_{t_2})} g_{\alpha(j_{t_2})} - \gamma^{\beta(j_{t_1})} g_{\alpha(j_{t_1})}\right). \tag{8}$$

We now analyze the valuation of each difference term. For $t_1 < t_2$, consider

$$\gamma^{\beta(j_{t_2})} g_{\alpha(j_{t_2})} - \gamma^{\beta(j_{t_1})} g_{\alpha(j_{t_1})}.$$

There are two cases.

Case 1: $\alpha(j_{t_1}) = \alpha(j_{t_2})$. Then $\beta(j_{t_1}) < \beta(j_{t_2})$, and

$$v_{P_\infty}\left(\gamma^{\beta(j_{t_2})} g_{\alpha(j_{t_2})} - \gamma^{\beta(j_{t_1})} g_{\alpha(j_{t_1})}\right) = v_{P_\infty}\left((\gamma^{\beta(j_{t_2})} - \gamma^{\beta(j_{t_1})}) g_{\alpha(j_{t_2})}\right) = v_{P_\infty}(g_{\alpha(j_{t_2})}).$$

Case 2: $\alpha(j_{t_1}) < \alpha(j_{t_2})$. Then by (7),

$$v_{P_\infty}(g_{\alpha(j_{t_2})}) < v_{P_\infty}(g_{\alpha(j_{t_1})}),$$

and hence, by the strong triangle inequality, $v_{P_\infty}\left(\gamma^{\beta(j_{t_2})} g_{\alpha(j_{t_2})} - \gamma^{\beta(j_{t_1})} g_{\alpha(j_{t_1})}\right) = v_{P_\infty}(g_{\alpha(j_{t_2})})$.

Thus in both cases,

$$v_{P_\infty}\left(\gamma^{\beta(j_{t_2})} g_{\alpha(j_{t_2})} - \gamma^{\beta(j_{t_1})} g_{\alpha(j_{t_1})}\right) = v_{P_\infty}(g_{\alpha(j_{t_2})}). \tag{9}$$

For each $I = \{j_1 < \cdots < j_r\}$, define

$$T_I = \det(E_I)\det(M^I).$$

From (8) and the definition of T_I, we obtain

$$v_{P_\infty}(T_I) = \sum_{u=1}^{r} v_{P_\infty}(f_{j_u}) + \sum_{1\le t_1<t_2\le r} v_{P_\infty}\left(\gamma^{\beta(j_{t_2})} g_{\alpha(j_{t_2})} - \gamma^{\beta(j_{t_1})} g_{\alpha(j_{t_1})}\right) + v_{P_\infty}(\det(M^I)).$$

By (9), each term in the second sum equals $v_{P_\infty}(g_{\alpha(j_{t_2})})$, and hence

$$\begin{aligned} v_{P_\infty}(T_I) &= \sum_{u=1}^{r} v_{P_\infty}(f_{j_u}) + \sum_{t=1}^{r}(t-1) v_{P_\infty}(g_{\alpha(j_t)}) + v_{P_\infty}(\det(M^I)). \\ &= \begin{cases} \sum_{u=1}^{r} v_{P_\infty}(f_{j_u}) + \sum_{t=1}^{r}(t-1) v_{P_\infty}(g_{\alpha(j_t)}) & \text{if } \det(M^I)\neq 0, \\ \infty & \text{otherwise.} \end{cases} \end{aligned} \tag{10}$$

Let $I^* \subseteq [k]$ be the subset produced by the following greedy procedure: initialize $I^* = \emptyset$, and for $i = k, k-1, \dots, 1$, add i to I^* if the i-th column of M does not lie in the $\mathbb{F}$-linear span of the columns indexed by the current set I^*. By Lemma 2.2, we have $|I^*| = r$ and $\det(M^{I^*}) \neq 0$.

First, we show that I^* uniquely minimizes the first sum $\sum_{u=1}^{r} v_{P_\infty}(f_{j_u})$ in (10) among all subsets $I \subseteq [k]$ of size r satisfying $\det(M^I) \neq 0$. Indeed, by (6), the sequence

$$v_{P_\infty}(f_1) > v_{P_\infty}(f_2) > \cdots > v_{P_\infty}(f_k)$$

is strictly decreasing. Applying the second item of Lemma 2.2 with weights $w_i = v_{P_\infty}(f_i)$, we conclude that

$$\sum_{i\in I^*} v_{P_\infty}(f_i) \le \sum_{i\in J} v_{P_\infty}(f_i)$$

for every subset $J \subseteq [k]$ of size r such that $\det(M^J) \neq 0$, with equality if and only if $J = I^*$. This proves the claim.

Next, we show that I^* also minimizes the second sum $\sum_{t=1}^{r}(t-1) v_{P_\infty}(g_{\alpha(j_t)})$ among all subsets $I = \{j_1 < \cdots < j_r\}$ satisfying $\det(M^I) \neq 0$.

Let $I = \{j_1 < \cdots < j_r\}$ be any such subset, and write $I^* = \{j_1^* < \cdots < j_r^*\}$. By the first item of Lemma 2.2, we have $j_t^* \ge j_t$ for all $t \in [r]$. Since $\alpha(\cdot)$ is nondecreasing and the sequence $v_{P_\infty}(g_1) > \cdots > v_{P_\infty}(g_h)$ is strictly decreasing by (7), it follows that

$$v_{P_\infty}(g_{\alpha(j_t^*)}) \le v_{P_\infty}(g_{\alpha(j_t)}) \qquad \text{for all } t \in [r].$$

Multiplying by $(t-1) \ge 0$ and summing over t, we obtain

$$\sum_{t=1}^{r}(t-1) v_{P_\infty}(g_{\alpha(j_t^*)}) \le \sum_{t=1}^{r}(t-1) v_{P_\infty}(g_{\alpha(j_t)}),$$

as claimed.

Combining the two parts above, we conclude that I^* uniquely minimizes $v_{P_\infty}(T_I)$ among all subsets $I \subseteq [k]$ of size r.

Returning to the Cauchy–Binet expansion

$$\det(EM) = \sum_{I\subseteq[k],\,|I|=r} T_I,$$

we have $\det(M^{I^*}) \neq 0$, and hence (10) implies that $v_{P_\infty}(T_{I^*}) < \infty$. Since T_{I^*} is the unique term attaining the minimum valuation, the strong triangle inequality yields

$$v_{P_\infty}(\det(EM)) = v_{P_\infty}(T_{I^*}) < \infty,$$

and thus $\det(EM) \neq 0$.

Moreover, writing $I^* = \{j_1^* < \cdots < j_r^*\}$ and using (10), we have

$$\begin{aligned} v_{P_\infty}(\det(EM)) &= \sum_{u=1}^{r} v_{P_\infty}(f_{j_u^*}) + \sum_{t=1}^{r}(t-1)v_{P_\infty}(g_{\alpha(j_t^*)}) \\ &\overset{(6),(7)}{\geq} -\sum_{u=1}^{r}(d_1 - r + u) - \sum_{t=1}^{r}(t-1)d_2 \\ &= -\sum_{u=1}^{r}(k-1+g-r+u) - \sum_{t=1}^{r}(t-1)\left(\left\lceil \frac{k}{q-1} \right\rceil - 1 + g\right), \end{aligned}$$

which simplifies to

$$v_{P_\infty}(\det(EM)) \geq -r\left(k - r + \frac{r-1}{2}\left\lceil \frac{k}{q-1} \right\rceil + \frac{r+1}{2}g\right).$$

□

The proof of The following theorem is identical to that of Theorem 3.2, except that we use Lemma 4.2 in place of Lemma 3.1, improving the $O(r^2k)$ term in Theorem 3.2 to $O(rk)$.

Theorem 4.3. *Let $M \in \mathbb{F}_q^{k\times r}$ be a full-rank matrix. The number of $P \in S$ such that $E(P)M \in \mathbb{F}_q^{r\times r}$ does not have full rank is at most*

$$r\left(k - r + \frac{r-1}{2}\left\lceil \frac{k}{q-1} \right\rceil + \frac{r+1}{2}g\right).$$

Remark 4.4. *If we take $F = \mathbb{F}_q(X)$ with $g = 0$ and $q > k$, then the above construction reduces to (a generalization of) that of Forbes and Shpilka [FS12]. In this case, the bound in Theorem 4.3 becomes*

$$r\left(k - \frac{r+1}{2}\right) = rk - \binom{r+1}{2},$$

matching the bound in [FS12].

For the construction of Forbes and Shpilka [FS12], this bound was later improved to $r(k-r)$ in [FSS14, For14], using additional properties of the monomials X^{i-1} for $i \in [k]$. It is unclear whether a similar improvement can be obtained for general function fields, or for those defined over small fields that we consider here. In any case, such an improvement would be minor in our setting where $r \ll k$, and would be dominated by the error terms.

Instantiating the function field and parameters. We now choose F to be a function field from either the Garcia–Stichtenoth tower or the Bassa–Beelen–Garcia–Stichtenoth tower, in order to obtain explicit rank extractors over non-prime fields. This construction also yields explicit weak subspace designs; we defer the corresponding discussion to Section 5.3.

Theorem 4.5 (Formal version of Theorem 1.3)**.** *Let $1 \leq r \leq k$ be integers. Suppose q is a prime power that is not a prime and satisfies $q \geq (2r)^{c_0}$ for a sufficiently large absolute constant c_0. Then there exists an explicit (k, r, L) lossless rank extractor over $\mathbb{F}_q$ of size n, where $L = O(r(k-r)q)$ and $L/n \leq q^{-1/4}$. Moreover, for each fixed r and q, we have $L \leq c_1 r(k-r)$ for some absolute constant $c_1 > 0$ and infinitely many k.*

Proof. If $r > k/2$, then $q \geq (2r)^{c_0}$, which is greater than k for suffciently large c_0. In this case, we can use the Forbes–Shpilka construction [FS12] with $\mathbb{F}_q^\times$ as the set of evaluation points, so that $L = r(k-r)$ and $n = q-1$ [FSS14, For14]. See, e.g., [For14, Theorem 5.4.3]. Then $L/n \leq r(k-r)/(q-1) \leq r^2/(q-1) \leq q^{-1/4}$ as $q \geq (2r)^{c_0}$ and c_0 is large enough.

So assume $r \leq k/2$, which implies that $k = O(k-r)$.

If $q = \ell^2$ is a square, let i be the smallest positive integer such that $\ell^i \geq k/r$, and let $F = F_i$ be the i-th field in the Garcia–Stichtenoth tower over $\mathbb{F}_q$. Otherwise, $q = p^{2m+1}$ with $m \geq 1$, and we let i be the smallest positive integer such that

$$\frac{p^{2m(i-1)}}{2}\left(\frac{q-1}{p^m-1} + \frac{q-1}{p^{m+1}-1}\right) \geq k/r,$$

and let $F = F_i$ be the i-th field in the Bassa–Beelen–Garcia–Stichtenoth tower over $\mathbb{F}_q$. In either case, this choice of i ensures that $g \leq (k/r)q$ by Lemmas 2.13 and 2.15.[3] (In fact, this extra factor of q can be improved to ℓ or p^{2m}, but we use this crude bound for simplicity.)

Let

$$L = r\left(k - r + \frac{r-1}{2}\left\lceil\frac{k}{q-1}\right\rceil + \frac{r+1}{2}g\right).$$

By Theorem 4.3, the collection $\mathcal{E} = \{E(P) : P \in S\}$ is an explicit (k, r, L) lossless rank extractor over $\mathbb{F}_q$.

Since $g \leq (k/r)q$ and $q \geq (2r)^{c_0}$, we obtain $L = O(rkq) = O(r(k-r)q)$.

Moreover, for infinitely many k, we have $g = O(k/r)$. For instance, in the case $q = \ell^2$, we may fix i and take $k = \ell^i r$. In the case $q = p^{2m+1}$, we may take i sufficiently large and set

$$k = \left\lfloor r \cdot \frac{p^{2m(i-1)}}{2}\left(\frac{q-1}{p^m-1} + \frac{q-1}{p^{m+1}-1}\right)\right\rfloor.$$

For such k, we obtain $L = O(rk) = O(r(k-r))$.

Finally, we bound L/n. If $q = \ell^2$, then $n = |S| \geq \ell^i(\ell-1)$ by Lemma 2.13. By the choice of i, we have $L = O(r^2\ell^i)$, so

$$L/n = O(r^2\ell^i/n) = O(r^2 q^{-1/2}),$$

which is at most $q^{-1/4}$ for sufficiently large c_0.

If $q = p^{2m+1}$, then $n = |S| \geq p^{2m(i-1)}(q-1) - 1$ by Lemma 2.15. By the choice of i, we have

$$L = O\left(r^2 p^{2m(i-1)}\frac{q-1}{p^m}\right),$$

and hence

$$L/n = O(r^2/p^m) = O\left(r^2 q^{-m/(2m+1)}\right),$$

which is again at most $q^{-1/4}$ for sufficiently large c_0. □

[3] We elaborate on the case $q = p^{2m+1}$; the case $q = \ell^2$ is similar. If $i = 1$, then $F = E_1 = \mathbb{F}_q(x_1)$ has genus $g = 0$, so $g \leq (k/r)q$ trivially. Suppose now that $i \geq 2$. By the minimality of i, we have $\frac{p^{2m(i-2)}}{2}\left(\frac{q-1}{p^m-1} + \frac{q-1}{p^{m+1}-1}\right) < k/r$. So $g \leq \frac{p^{2m(i-1)}}{2}\left(\frac{q-1}{p^m-1} + \frac{q-1}{p^{m+1}-1}\right) \leq q \cdot \frac{p^{2m(i-2)}}{2}\left(\frac{q-1}{p^m-1} + \frac{q-1}{p^{m+1}-1}\right) \leq (k/r)q$.

5 Field Reduction

As mentioned in the introduction, for a field extension $\mathbb{F}_Q/\mathbb{F}_q$, lossless rank condensers over $\mathbb{F}_Q$ can be converted into ones over $\mathbb{F}_q$ via [FG14, Proposition 8.5], but this transformation does not preserve the lossless rank extractor property. In Section 5.1, we develop a different field-reduction technique based on the multilinearity of the determinant, and show that it transforms lossless rank extractors over $\mathbb{F}_Q$ into ones over $\mathbb{F}_q$, albeit with an exponential blow-up in size in r.

In Section 5.2, we further refine this approach by combining it with polynomial identity testing. This is particularly useful over sufficiently large prime fields $\mathbb{F}_q$, where techniques such as the Garcia–Stichtenoth tower and the Bassa–Beelen–Garcia–Stichtenoth tower do not directly apply.

Finally, in Section 5.3, we apply these techniques to prove our main theorems.

5.1 Field Reduction via Multilinearity of the Determinant

Let $r \le k$ be positive integers. Let $\mathbb{F}_Q/\mathbb{F}_q$ be a finite field extension of degree $d = [\mathbb{F}_Q : \mathbb{F}_q]$.

Fix a basis $e_1, \dots, e_d \in \mathbb{F}_Q$ of $\mathbb{F}_Q$ over $\mathbb{F}_q$. For $a \in \mathbb{F}_Q$, let $a^{(1)}, \dots, a^{(d)} \in \mathbb{F}_q$ denote the coordinates of a with respect to this basis, so that $a = \sum_{t=1}^{d} a^{(t)} e_t$.

For an $r \times k$ matrix $E = (E_{ij}) \in \mathbb{F}_Q^{r\times k}$, let E_i denote its i-th row for $i \in [r]$. For $t \in [d]$, define

$$E_i^{(t)} := (E_{i1}^{(t)}, \dots, E_{ik}^{(t)}) \in \mathbb{F}_q^k,$$

the vector obtained by taking the t-th coordinate of each entry of E_i.

For a tuple $\sigma = (\sigma_1, \dots, \sigma_r) \in [d]^r$, define the matrix $E^{(\sigma)} \in \mathbb{F}_q^{r\times k}$ by

$$E^{(\sigma)} := (E_i^{(\sigma_i)})_{i\in[r]}.$$

Finally, for a finite collection $\mathcal{E} \subseteq \mathbb{F}_Q^{r\times k}$, define the collection

$$\mathcal{E}_{\mathbb{F}_Q\to\mathbb{F}_q} := \{E^{(\sigma)} : E \in \mathcal{E},\ \sigma \in [d]^r\} \subseteq \mathbb{F}_q^{r\times k}.$$

Then

$$|\mathcal{E}_{\mathbb{F}_Q\to\mathbb{F}_q}| = d^r|\mathcal{E}| = (\log_q Q)^r|\mathcal{E}|.$$

Recall the definition of a lossless rank disperser (Definition 1.6). We now show that if $\mathcal{E}$ is a lossless rank disperser over $\mathbb{F}_Q$, then $\mathcal{E}_{\mathbb{F}_Q\to\mathbb{F}_q}$ is a lossless rank disperser over $\mathbb{F}_q$.

Theorem 5.1. *Let $\mathcal{E} \subseteq \mathbb{F}_Q^{r\times k}$ be a finite collection. Suppose $M \in \mathbb{F}_Q^{k\times r}$ is a full-rank matrix such that*

$$|\{E \in \mathcal{E} : \operatorname{rank}(EM) = r\}| > 0.$$

Then

$$|\{E \in \mathcal{E}_{\mathbb{F}_Q\to\mathbb{F}_q} : \operatorname{rank}(EM) = r\}| > 0.$$

In particular, this holds for every $M \in \mathbb{F}_q^{k\times r}$ of rank r.

Corollary 5.2. *If $\mathcal{E}$ is a (k, r) lossless rank disperser over $\mathbb{F}_Q$, then $\mathcal{E}_{\mathbb{F}_Q\to\mathbb{F}_q}$ is a (k, r) lossless rank disperser over $\mathbb{F}_q$.*

Proof of Theorem 5.1. By assumption, there exists $E \in \mathcal{E}$ such that $\operatorname{rank}(EM) = r$, or equivalently, $\det(EM) \neq 0$. Fix such an $E = (E_{ij})$.

For each $i \in [r]$ and $j \in [k]$, we have

$$E_i = \sum_{t \in [d]} e_t E_i^{(t)},$$

and hence

$$E_i M = \sum_{t \in [d]} e_t \left(E_i^{(t)} M \right). \tag{11}$$

For $\sigma = (\sigma_1, \ldots, \sigma_r) \in [d]^r$, let A_σ be the $r \times r$ matrix whose i-th row is $E_i^{(\sigma_i)} M$ for $i \in [r]$. Then $A_\sigma = E^{(\sigma)} M$ since the i-th row of $E^{(\sigma)} M$ is also $E_i^{(\sigma_i)} M$ for $i \in [r]$.

By (11) and multilinearity of the determinant with respect to rows, we have

$$\det(EM) = \sum_{\sigma \in [d]^r} \left(\prod_{i=1}^{r} e_{\sigma_i} \right) \det(A_\sigma) = \sum_{\sigma \in [d]^r} \left(\prod_{i=1}^{r} e_{\sigma_i} \right) \det(E^{(\sigma)} M).$$

Since $\det(EM) \neq 0$, there exists $\sigma \in [d]^r$ such that $\det(E^{(\sigma)} M) \neq 0$. As $E^{(\sigma)} \in \mathcal{E}_{\mathbb{F}_Q \to \mathbb{F}_q}$, the claim follows. □

5.2 Field Reduction via PIT for Symbolic Determinants with Rank-One Summands

Corollary 5.2 allows us to construct lossless rank dispersers over an arbitrary finite field $\mathbb{F}_q$ from those over an extension field $\mathbb{F}_Q$. However, it increases the size by a factor exponential in r.

We now show that, when $\mathbb{F}_q$ is sufficiently large, this drawback can be mitigated using techniques from polynomial identity testing. In particular, using the current best known result of [GT20], the size increases by only a quasi-polynomial factor in r, rather than an exponential one. Moreover, this approach also yields the stronger lossless rank extractor property.

We begin by defining a relevant class of polynomials.

Definition 5.3 (Symbolic determinants with rank-one summands)**.** Let r and N be positive integers, let $x = \{x_1, \ldots, x_N\}$ be a set of variables, and let $\mathbb{F}$ be a field. Define $\mathrm{VBP}^1_{r,N,\mathbb{F}} \subseteq \mathbb{F}[x]$ to be the class of polynomials $f \in \mathbb{F}[x]$ that can be written as

$$f(x) = \det\left(\sum_{i=1}^{N} x_i A_i \right),$$

where each $A_i \in \mathbb{F}^{r \times r}$ satisfies $\mathrm{rank}(A_i) \leq 1$. We call such polynomials (r, N) *symbolic determinants with rank-one summands.*

We also need the notion of *δ-hitting sets.*

Definition 5.4 (δ-hitting sets)**.** Let $\mathcal{C} \subseteq \mathbb{F}[x_1, \ldots, x_N]$ be a class of polynomials, and let $\delta \geq 0$. A set $\mathcal{H} \subseteq \mathbb{F}^N$ is called a δ-hitting set for $\mathcal{C}$ if for every nonzero polynomial $f \in \mathcal{C}$,

$$\left| \{ a \in \mathcal{H} : f(a) \neq 0 \} \right| \geq (1 - \delta) |\mathcal{H}|.$$

The following lemma shows that an explicit lossless rank extractor over $\mathbb{F}_Q$ can be transformed into one over $\mathbb{F}_q$, assuming the existence of an explicit δ-hitting set $\mathcal{H} \subseteq \mathbb{F}_q^{rd}$ for the class $\mathrm{VBP}^1_{r,rd,\mathbb{F}_q}$, where $d = [\mathbb{F}_Q : \mathbb{F}_q]$.

Lemma 5.5. *Let $r \le k$ be positive integers, and let $\mathbb{F}_Q/\mathbb{F}_q$ be a finite field extension of degree $d = [\mathbb{F}_Q : \mathbb{F}_q]$. Let $\mathcal{E}_Q \subseteq \mathbb{F}_Q^{r\times k}$ be an explicit (k, r, L) lossless rank extractor over $\mathbb{F}_Q$ of size n. Let $\mathcal{H} \subseteq \mathbb{F}_q^{dr}$ be an explicit δ-hitting set for $\mathrm{VBP}^1_{r,dr,\mathbb{F}_q}$.*

Then one can construct an explicit $(k, r, |\mathcal{H}|(L + \delta(n - L)))$ lossless rank extractor $\mathcal{E}_q$ over $\mathbb{F}_q$ of size $n|\mathcal{H}|$.

Proof. Fix a basis $e_1, \dots, e_d \in \mathbb{F}_Q$ of $\mathbb{F}_Q$ over $\mathbb{F}_q$. For $a \in \mathbb{F}_Q$, let $a^{(1)}, \dots, a^{(d)} \in \mathbb{F}_q$ denote its coordinates with respect to this basis, as defined earlier.

For $E \in \mathcal{E}_Q$, write $E = (E_{ij})$. For each $i \in [r]$ and $j \in [d]$, define $A_{i,j}(E) \in \mathbb{F}_q^{r\times k}$ to be the $r \times k$ matrix whose i-th row is $(E_{i1}^{(j)}, \dots, E_{ik}^{(j)})$ and whose other rows are zero. Then $\mathrm{rank}(A_{i,j}(E)) \le 1$.

Let $x = \{x_{i,j} : i \in [r],\ j \in [d]\}$ be variables, and define

$$E^*(x) := \sum_{i\in[r],\, j\in[d]} x_{i,j} A_{i,j}(E) \in \mathbb{F}_q[x]^{r\times k}.$$

Define the collection of matrices

$$\mathcal{E}_q := \{E^*(a) : E \in \mathcal{E}_Q,\ a \in \mathcal{H}\} \subseteq \mathbb{F}_q^{r\times k}.$$

Then $|\mathcal{E}_q| = n|\mathcal{H}|$.

We claim that $\mathcal{E}_q$ is an $(k, r, |\mathcal{H}|(L + \delta(n - L)))$ lossless rank extractor over $\mathbb{F}_q$. Let $M \in \mathbb{F}_q^{k\times r}$ be of full column rank.

By multilinearity of the determinant (cf. the proof of Theorem 5.1), we have

$$\det(EM) = \sum_{\sigma\in[d]^r} \left(\prod_{i=1}^{r} e_{\sigma_i}\right) \det\big(E^{(\sigma)}M\big).$$

Thus, if $\det(EM) \neq 0$, then there exists $\sigma \in [d]^r$ such that $\det(E^{(\sigma)}M) \neq 0$. In this case, since

$$\det(E^*(x)M) = \sum_{\sigma\in[d]^r} \left(\prod_{i=1}^{r} x_{i,\sigma_i}\right) \det(E^{(\sigma)}M),$$

it follows that $\det(E^*(x)M)$ is a nonzero polynomial.

Moreover,

$$E^*(x)M = \sum_{i\in[r],\, j\in[d]} x_{i,j}(A_{i,j}(E)M),$$

and since each $A_{i,j}(E)$ has rank at most one, the same holds for $A_{i,j}(E)M$. Hence $\det(E^*(x)M) \in \mathrm{VBP}^1_{r,dr,\mathbb{F}_q}$.

Since $\mathcal{E}_Q$ is an (k, r, L) lossless rank extractor over $\mathbb{F}_Q$, there are at least $n - L$ matrices $E \in \mathcal{E}_Q$ such that $\det(EM) \neq 0$. For each such E, the polynomial $\det(E^*(x)M)$ is nonzero, and since $\mathcal{H}$ is a δ-hitting set for $\mathrm{VBP}^1_{r,dr,\mathbb{F}_q}$, there are at least $(1-\delta)|\mathcal{H}|$ points $a \in \mathcal{H}$ such that $\det(E^*(a)M) \neq 0$.

Therefore, among the matrices in $\mathcal{E}_q$, the number of matrices $E^*(a)$ such that $\mathrm{rank}(E^*(a)M) < r$ is at most

$$|\mathcal{H}|L + \delta|\mathcal{H}|(n - L) = |\mathcal{H}|(L + \delta(n - L)).$$

This proves the claim. □

Gurjar and Thierauf [GT20] constructed explicit hitting sets for symbolic determinants with rank-one summands. We adapt their construction to prove the following lemma. The proof is deferred to Appendix B.

Lemma 5.6. *Let N be a positive integer, $\delta \in (0,1)$, and $\mathbb{F}$ a field such that $|\mathbb{F}| \geq (N/\delta)^c$ for some large enough absolute constant $c > 0$. Then, one can construct an explicit δ-hitting set $\mathcal{H} \subseteq \mathbb{F}^N$ for $\mathrm{VBP}^1_{r,N,\mathbb{F}}$ of size polynomial in $(N/\delta)^{\log N}$.*

By combining Lemma 5.5 and Lemma 5.6, we obtain the following theorem, which is the main result of this subsection.

Theorem 5.7. *Let $\mathbb{F}_Q/\mathbb{F}_q$ be a finite field extension of degree $d = [\mathbb{F}_Q : \mathbb{F}_q]$. Let $r \leq k$ be positive integers, let $\delta \in (0,1)$, and let q be a prime power satisfying $q \geq (dr/\delta)^c$ for a sufficiently large absolute constant $c > 0$. Let $\mathcal{E}_Q \subseteq \mathbb{F}_Q^{r \times k}$ be an explicit (k, r, L) lossless rank extractor over $\mathbb{F}_Q$ of size n. Then there exists an explicit (k, r, L') lossless rank extractor $\mathcal{E}_q$ over $\mathbb{F}_q$ of size n', where*

$$L' = (L + \delta(n - L))h, \qquad n' = nh, \qquad h \leq (dr/\delta)^{O(\log(dr))}.$$

5.3 Putting It Together

We begin by proving the following result over arbitrary finite fields via field reduction:

Theorem 5.8 (Formal version of Theorem 1.7)**.** *Let $1 \leq r \leq k$ be integers, and let $q > 1$ be a prime power. Then there exists an explicit (k, r) lossless rank disperser over $\mathbb{F}_q$ of size n, where*

$$n = O\left(\max\left\{\frac{c\log(2r)}{\log q}, 2\right\}^r r(k-r)q\right),$$

for some absolute constant $c > 0$. Moreover, for each fixed r and q, we have

$$n \leq c_1 \max\left\{\frac{c\log(2r)}{\log q}, 2\right\}^r r(k-r)$$

for some absolute constant $c_1 > 0$ and infinitely many k.

Proof. Choose the smallest integer $d \geq 2$ such that $Q = q^d$ satisfies $Q \geq (2r)^{c_0}$, where c_0 is the constant from Theorem 4.5. Apply Theorem 4.5 to obtain an explicit (k, r, L) lossless rank extractor $\mathcal{E}_0$ over $\mathbb{F}_Q$ of size $n_0 > L$, where $L + 1 = O(r(k-r)q)$. Moreover, for each fixed r and q, we have $L + 1 \leq c_1 r(k-r)$ for some absolute constant $c_1 > 0$ and infinitely many k.

Remove $n_0 - (L+1)$ elements from $\mathcal{E}_0$. The resulting collection $\mathcal{E}$ is still a (k, r) lossless rank disperser over $\mathbb{F}_Q$, and its size is $L + 1$.

By Theorem 5.1, the collection $\mathcal{E}_{\mathbb{F}_Q \to \mathbb{F}_q}$ defined in Section 5.1 is a (k, r) lossless rank disperser over $\mathbb{F}_q$ of size

$$d^r(L+1) \leq \max\left\{\frac{c\log(2r)}{\log q}, 2\right\}^r (L+1)$$

for some absolute constant $c > 0$. Since $L + 1 = O(r(k-r)q)$ for all k and $L + 1 = O(r(k-r))$ for infinitely many k, the claim follows. □

Remark 5.9. *It follows from the proof above that the number 2 in $\max\left\{\frac{c\log(2r)}{\log q}, 2\right\}$ can be improved to 1 if q is non-prime. The reason we choose 2 in general is that we need Q to be non-prime so that Theorem 4.5 applies.*

Next, we prove the following result over prime fields:

Theorem 5.10 (Formal version of Theorem 1.5)**.** *Let $1 \le r \le k$ be integers, and let $\delta \in (0,1)$. Suppose $q > 1$ is a prime satisfying $q \ge (2r/\delta)^{c_0}$ for a sufficiently large absolute constant c_0. Then there exists an explicit (k, r, L) lossless rank extractor over $\mathbb{F}_q$ of size n, where*

$$L \le (2r/\delta)^{c \log(2r)} r(k-r)q$$

for some absolute constant $c > 0$, and $L/n \le \delta$. Moreover, for each fixed r and q, we have

$$L \le (2r/\delta)^{c \log(2r)} r(k-r)$$

for infinitely many k.

Proof. We assume $r < k$, since otherwise the $k \times k$ identity matrix forms an $(k, r, 0)$ lossless rank extractor.

Let $Q = q^2$. Apply Theorem 4.5 to obtain an explicit (k, r, L_0) lossless rank extractor $\mathcal{E}_0$ over $\mathbb{F}_Q$ of size $n_0 > L_0$, where $L_0 = O(r(k-r)q)$ and $L_0/n_0 \le Q^{-1/4} = q^{-1/2}$. Moreover, for each fixed r and q, we have $L_0 \le c_1 r(k-r)$ for some absolute constant $c_1 > 0$ and infinitely many k.

Let $\delta' = \delta/2$. Since $q \ge (2r/\delta)^{c_0}$, we may assume $q^{-1/2} \le \delta'$. By removing elements from $\mathcal{E}_0$, we obtain an explicit (k, r, L_0) lossless rank extractor $\mathcal{E}_Q$ over $\mathbb{F}_Q$ of size $n_1 = \lceil L_0/\delta' \rceil$.

By Theorem 5.7, there exists an explicit (k, r, L) lossless rank extractor $\mathcal{E}$ over $\mathbb{F}_q$ of size n, where $L = (L_0 + \delta'(n_1 - L_0))h$, $n = n_1 h$, and $h \le (2r/\delta')^{O(\log(2r))}$. Then

$$L/n \le (L_0/n_1) + \delta' \le 2\delta' = \delta.$$

Moreover, since $L_0 = O(r(k-r)q)$,

$$L = (L_0 + \delta'(n_1 - L_0))h \le (2L_0 + \delta')h = (2L_0 + \delta/2)h \le (2r/\delta)^{c \log(2r)} r(k-r)q,$$

where $c > 0$ is a sufficiently large absolute constant. Finally, since $L_0 \le c_1 r(k-r)$ for infinitely many k, by choosing c large enough, we obtain

$$L \le (2r/\delta)^{c \log(2r)} r(k-r)$$

for infinitely many k. □

Next, we prove our main results on explicit weak subspace designs.

Theorem 5.11 (Formal version of Theorems 1.8 to 1.10)**.** *Let $1 \le r < k$ be integers, and let $r' = \min\{r, k-r\}$. Let $\delta \in (0,1)$. Suppose $q > 1$ is a prime power. Then there exists an explicit $(k-r, A)$ weak subspace design $\mathcal{V}$ of size n consisting of r-dimensional subspaces of $\mathbb{F}_q^k$, and an explicit (r, A) weak subspace design $\mathcal{V}^\perp$ of size n consisting of $(k-r)$-dimensional subspaces of $\mathbb{F}_q^k$, such that $\mathcal{V} = \{V^\perp : V \in \mathcal{V}^\perp\}$ and $\mathcal{V}^\perp = \{V^\perp : V \in \mathcal{V}\}$. Moreover,*

1. *Suppose q is a non-prime prime power satisfying $q \ge (2r')^{c_0}$ for some sufficiently large absolute constant $c_0 > 0$. Then $A = O(r(k-r)q)$ and $n \ge q^{1/4}A/2$. Moreover, for each fixed r and q, we have $A \le c_1 r(k-r)$ for some absolute constant $c_1 > 0$ and infinitely many k.*

2. *Suppose q is a prime satisfying $q \ge (2r')^{c_0}$ for some sufficiently large absolute constant $c_0 > 0$. Then*

   $$A \le (2r'/\delta)^{c \log(2r')} r(k-r)q$$

   for some absolute constant $c > 0$, and $n \ge A/\delta$. Moreover, for each fixed r and q, we have

   $$A \le (2r'/\delta)^{c \log(2r')} r(k-r)$$

   for infinitely many k.

3. *In general,*

$$A = O\left(\max\left\{\frac{c\log(2r')}{\log q}, 2\right\}^{r'} r(k-r)q\right)$$

for some absolute constant $c > 0$, and $n > A$. Moreover, for each fixed r and q, we have

$$A \leq c_1 \max\left\{\frac{c\log(2r')}{\log q}, 2\right\}^{r'} r(k-r)$$

for some absolute constant $c_1 > 0$ and infinitely many k.

Proof. Note that if the theorem holds for $r \leq k/2$, then the case $r > k/2$ follows by replacing r with $k-r$ and swapping $\mathcal{V}$ and $\mathcal{V}^\perp$. Thus, we may assume $r \leq k/2$, and hence $r' = r$.

To prove Item 1, we apply Theorem 4.5 to obtain an explicit (k, r, L) lossless rank extractor $\mathcal{E}_0$. In particular, $L/|\mathcal{E}_0| \leq q^{-1/4} \leq 1/2$. We then apply Lemma 2.4 to obtain two weak subspace designs $\mathcal{V}$ and $\mathcal{V}^\perp$ of size $|\mathcal{E}_0| - L \geq \frac{1}{2}|\mathcal{E}_0| \geq q^{1/4}L/2$. Let $A = L$, and Item 1 follows. The proof of Item 2 is similar, using Theorem 5.10 in place of Theorem 4.5.

Finally, to prove Item 3, we apply Theorem 5.8 to obtain an explicit (k, r) lossless rank disperser $\mathcal{E}_0$, and then remove all matrices of rank less than r. This does not affect the lossless rank disperser property: if $E \in \mathbb{F}_q^{r\times k}$ satisfies $\operatorname{rank}(EM) = r$ for some $M \in \mathbb{F}_q^{k\times r}$, then necessarily $\operatorname{rank}(E) = r$, and hence such matrices are not removed. Let $\mathcal{E}_1$ denote the resulting collection. We then apply Lemma 2.3 to obtain $\mathcal{V} = \mathcal{V}(\mathcal{E}_1)$ and $\mathcal{V}^\perp = \mathcal{V}^\perp(\mathcal{E}_1)$. Let $A = |\mathcal{V}| - 1 = |\mathcal{V}^\perp| - 1$, and Item 3 follows. □

Remark 5.12. *Using the same idea of exchanging r with $k-r$ via duality, one may relax the condition $q \geq \operatorname{poly}(r)$ in Theorem 1.3 and Theorem 1.5 to $q \geq \operatorname{poly}(r')$, where $r' = \min\{r, k-r\}$. We choose not to adopt this formulation for simplicity.*

6 Explicit Constructions of Strong s-Blocking Sets

Recall that a set $B \subseteq \mathrm{PG}(k-1, q)$ is a *strong s-blocking set* if for every codimension-s projective subspace Σ of $\mathrm{PG}(k-1, q)$, the intersection $B \cap \Sigma$ spans Σ. In this section, we present two types explicit constructions of strong s-blocking sets in $\mathrm{PG}(k-1, q)$. The first arises from lossless rank extractors over $\mathbb{F}_q$, while the second is obtained from ε-biased sets in $\mathbb{F}_q^k$ with sufficiently small bias.

For a subset $S \subseteq \mathbb{F}_q^k$, let $\operatorname{span}(S)$ denote its $\mathbb{F}_q$-linear span, and let $\widetilde{S} \subseteq \mathrm{PG}(k-1, q)$ denote its projectivization, i.e.,

$$\widetilde{S} = \{\langle u\rangle : u \in S \setminus \{0\}\},$$

where $\langle u\rangle$ denotes the one-dimensional subspace spanned by u.

6.1 Strong s-Blocking Sets via Lossless Rank Dispersers

The following lemma shows that every $(k, s+1)$ lossless rank disperser over $\mathbb{F}_q$ (see Definition 1.6) gives rise to a strong s-blocking set in $\mathrm{PG}(k-1, q)$. This was essentially already observed in [FS16, Proposition 10].[4]

[4] A minor technical difference is that [FS16, Proposition 10] assumes the matrices all have full rank, i.e., their row spaces have dimension $s+1$. This is unimportant, as rank-deficient matrices can be safely ignored.

Lemma 6.1. *Let $s < k$ be positive integers. Suppose that $\mathcal{E} = \{E_i\}_{i=1}^n \subseteq \mathbb{F}_q^{(s+1)\times k}$ is a $(k, s+1)$ lossless rank disperser over $\mathbb{F}_q$. For $i \in [n]$, let $V_i = \operatorname{rowspan}(E_i) \subseteq \mathbb{F}_q^k$. Then $B := \bigcup_{i=1}^n \widetilde{V}_i$ is a strong s-blocking set in $\mathrm{PG}(k-1, q)$. Moreover, $|B| \le n\frac{q^{s+1}-1}{q-1} \le 2nq^s$.*

Proof. Suppose for contradiction that B is not a strong s-blocking set. Then there exists a codimension-s subspace $L \subseteq \mathbb{F}_q^k$ and a codimension-$(s+1)$ subspace $H \subseteq L$ such that for all $i \in [n]$,

$$V_i \cap L \subseteq H. \tag{12}$$

Choose $M \in \mathbb{F}_q^{k\times(s+1)}$ of rank $s+1$ such that $\ker(M^T) = H$. Since $\mathcal{E}$ is a $(k, s+1)$ lossless rank disperser, there exists $E_i \in \mathcal{E}$ such that $\operatorname{rank}(E_iM) = s+1$. In particular, $\dim V_i = s+1$.

Since $\dim(V_i) + \dim(L) = k+1$, we have $V_i \cap L \neq \{0\}$. By (12), this implies $V_i \cap H \neq \{0\}$.

On the other hand, $\operatorname{rank}(E_iM) = s+1$ implies that M^T is injective on V_i, hence

$$V_i \cap \ker(M^T) = \{0\},$$

i.e., $V_i \cap H = \{0\}$, a contradiction.

Thus B is a strong s-blocking set. The size bound is immediate. □

Combining Lemma 6.1 with Theorem 4.5, Theorem 5.10, and Theorem 5.8, respectively, we obtain the following results on explicit strong s-blocking sets.

Corollary 6.2 (Formal version of Theorems 1.12 to 1.14)**.** *Let $1 \le s < k$ be integers, and let $q > 1$ be a prime power. Then the following hold:*

1. *Suppose q is a non-prime prime power satisfying $q \ge (2s)^{c_0}$ for some sufficiently large absolute constant $c_0 > 0$. Then there exists an explicit strong s-blocking set $B \subseteq \mathrm{PG}(k-1, q)$ of size $|B| = O(s(k-s)q^{s+1})$. Moreover, for each fixed s and q, we have $|B| = O(s(k-s)q^s)$ for infinitely many k.*

2. *Suppose q is prime and satisfies $q \ge (2s)^{c_0}$ for some sufficiently large absolute constant $c_0 > 0$. Then there exists an explicit strong s-blocking set $B \subseteq \mathrm{PG}(k-1, q)$ of size at most $(2s)^{c\log(2s)}(k-s)q^{s+1}$ for some absolute constant $c > 0$. Moreover, for each fixed s and q, we have $|B| \le (2s)^{c\log(2s)}(k-s)q^s$ for infinitely many k.*

3. *In general, there exists an explicit strong s-blocking set $B \subseteq \mathrm{PG}(k-1, q)$ of size*

$$O\left(\max\left\{\frac{c\log(2s)}{\log q}, 2\right\}^{s+1} s(k-s)q^{s+1}\right)$$

for some absolute constant $c > 0$. Moreover, for each fixed s and q, we have

$$|B| \le c_1 \max\left\{\frac{c\log(2s)}{\log q}, 2\right\}^{s+1} s(k-s)q^s$$

for some absolute constant $c_1 > 0$ and infinitely many k.

6.2 Strong s-Blocking Sets via ε-Biased Sets

In this subsection, we first show that every ε-biased set $S \subseteq \mathbb{F}_q^k$ with sufficiently small bias forms a strong s-blocking set. We then obtain explicit strong s-blocking sets by applying known constructions of ε-biased sets from [TS17, JM21].

We begin with basic terminologies of discrete Fourier analysis over $\mathbb{F}_q^n$; for more details, see, e.g., [TV06, O'D14]. Let $G = \mathbb{F}_q^k$, and let $\widehat{G}$ be its character group consisting of the characters $\chi : G \to \mathbb{C}^\times$. The identity of $\widehat{G}$ is called the *trivial* character, while the others are *nontrivial* characters.

Let $L(G, \mathbb{C})$ be the space of functions $f : G \to \mathbb{C}$. Define the inner product $\langle \cdot, \cdot \rangle$ on $L(G, \mathbb{C})$ via $\langle f, g \rangle = \mathbb{E}_{x \in G} \left[f(x)\overline{g(x)} \right]$, where $\overline{g(x)}$ denotes the complex conjugate of $g(x)$. The character group $\widehat{G}$ is an orthonormal basis of $L(G, \mathbb{C})$.

The orthogonality relation of characters states that for two characters $\chi, \psi \in \widehat{G}$, we have $\langle \chi, \psi \rangle = \mathbb{E}_{x \in G} \left[\chi(x)\overline{\psi(x)} \right] = \mathbf{1}[\chi = \psi]$. This implies that $\mathbb{E}_{x \in G} [\chi(x)] = 0$ for every nontrivial character χ.

For $f \in L(G, \mathbb{C})$ and $\chi \in \widehat{G}$, define the Fourier coefficient $\widehat{f}(\chi)$ of f to be

$$\widehat{f}(\chi) = \langle f, \chi \rangle = \mathop{\mathbb{E}}_{x \in G} \left[f(x)\overline{\chi(x)} \right].$$

This defines a function $\widehat{f} : \widehat{G} \to \mathbb{C}$ sending χ to $\widehat{f}(\chi)$, called the *Fourier transform* of f. Its L_1-norm is defined to be $\|\widehat{f}\|_1 = \sum_{\chi \in \widehat{G}} |\widehat{f}(\chi)|$.

As $\widehat{G}$ is an orthonormal basis of $L(G, \mathbb{C})$, any $f \in L(G, \mathbb{C})$ can be expanded as $f = \sum_{\chi \in G} \widehat{f}(\chi)\chi$.

Definition 6.3 (ε-biased set)**.** Let $\varepsilon > 0$. A nonempty set $S \subseteq G$ is called an *ε-biased set* if

$$\left| \mathop{\mathbb{E}}_{x \in S} [\chi(x)] \right| \leq \varepsilon$$

for all nontrivial characters $\chi \in \widehat{G}$.

The following lemma is well-known.

Lemma 6.4. *Let $f : G \to \mathbb{C}$ be a function, and let $S \subseteq G$ be an ε-biased set. Then*

$$\left| \mathop{\mathbb{E}}_{x \in S} [f(x)] - \mathop{\mathbb{E}}_{x \in G} [f(x)] \right| \leq \|\widehat{f}\|_1 \cdot \varepsilon.$$

Proof. Expand f as $f = \sum_{\chi \in \widehat{G}} \widehat{f}(\chi)\, \chi$. The claim holds for each character $\chi \in \widehat{G}$ by the definition of ε-biased sets. The result then follows by linearity of expectation. □

For a set $S \subseteq G$, let $1_S : G \to \mathbb{C}$ denote its indicator function, i.e., $1_S(x) = 1$ if $x \in S$ and $1_S(x) = 0$ otherwise. We will use the following lemma.

Lemma 6.5. *Let $V \subseteq G$ be an affine subspace over $\mathbb{F}_q$. Then $\|\widehat{1_V}\|_1 = 1$.*

Proof. We have $V = u + L$ for some vector $u \in G$ and some linear subspace $L \subseteq G$. Let $H \subseteq \widehat{G}$ be the subgroup of characters vanishing on L, i.e., $H = \{\chi \in \widehat{G} : \chi|_L = 1\}$. Note that H is the kernel of the surjective map $\widehat{G} \to \widehat{L}$ that dualizes the inclusion $L \subseteq G$. So $|H| = |\widehat{G}|/|\widehat{L}| = |G|/|L|$.

For $\chi \in \widehat{G}$, we have

$$\widehat{1_V}(\chi) = \mathop{\mathbb{E}}_{x\in G}\left[1_V(x)\overline{\chi(x)}\right] = \frac{1}{|G|}\sum_{x\in V}\overline{\chi(x)} = \frac{1}{|G|}\sum_{y\in L}\overline{\chi(u+y)}$$
$$= \frac{1}{|G|}\overline{\chi(u)}\sum_{y\in L}\overline{\chi(y)} = \frac{|L|}{|G|}\overline{\chi(u)}\mathop{\mathbb{E}}_{y\in L}\left[\overline{\chi(y)}\right].$$

Since $\chi|_L$ is either the trivial character or nontrivial, we have

$$\mathop{\mathbb{E}}_{y\in L}\left[\overline{\chi(y)}\right] = \begin{cases} 1, & \text{if } \chi|_L = 1, \text{ or equivalently, } \chi \in H, \\ 0, & \text{otherwise.} \end{cases}$$

It follows that

$$\|\widehat{1_V}\|_1 = \sum_{\chi\in\widehat{G}}|\widehat{1_V}(\chi)| = \sum_{\chi\in H}|\widehat{1_V}(\chi)| = \frac{|H||L|}{|G|}|\overline{\chi(u)}| = \frac{|H||L|}{|G|} = 1. \qquad \square$$

Recall that $B \subseteq \mathbb{F}_q^k$ is an affine s-blocking set if it intersects every affine subspace of codimension s. The next theorem connects ε-biased sets with affine and strong s-blocking sets.

Theorem 6.6. *Let $B \subseteq G$ be an ε-biased set. Then the following statements hold:*

1. *If $\varepsilon < q^{-s}$, then B is an affine s-blocking set.*
2. *If $\varepsilon < \frac{q-1}{2q^{s+1}}$, then $\widetilde{B}$ is a strong s-blocking set.*

Proof. For Item 1, assume $\varepsilon < q^{-s}$. Let L be an affine subspace of codimension s. By Lemmas 6.4 and 6.5,

$$\left|\frac{|B\cap L|}{|B|} - q^{-s}\right| = \left|\mathop{\mathbb{E}}_{x\in B}[1_L(x)] - \mathop{\mathbb{E}}_{x\in G}[1_L(x)]\right| \le \varepsilon\|\widehat{1_L}\|_1 = \varepsilon.$$

Hence

$$\frac{|B\cap L|}{|B|} \ge q^{-s} - \varepsilon > 0,$$

so $B \cap L \neq \emptyset$. Since this holds for every affine subspace L of codimension s, it follows that B is an affine s-blocking set.

For Item 2, assume $\varepsilon < \frac{q-1}{2q^{s+1}}$. It suffices to show that for every codimension-s linear subspace $L \subseteq \mathbb{F}_q^k$ and every hyperplane H of L, we have

$$\widetilde{B} \cap (\widetilde{L} \setminus \widetilde{H}) \neq \emptyset.$$

Writing $1_{L\setminus H} = 1_L - 1_H$, Lemma 6.5 and the triangle inequality give

$$\|\widehat{1_{L\setminus H}}\|_1 \le \|\widehat{1_L}\|_1 + \|\widehat{1_H}\|_1 = 2.$$

Applying Lemma 6.4 with $f = 1_{L\setminus H}$, we obtain

$$\left|\frac{|B\cap(L\setminus H)|}{|B|} - \frac{q^{k-s}-q^{k-s-1}}{q^k}\right| = \left|\mathop{\mathbb{E}}_{x\in B}[1_{L\setminus H}(x)] - \mathop{\mathbb{E}}_{x\in G}[1_{L\setminus H}(x)]\right| \le 2\varepsilon.$$

Therefore,

$$\frac{|B\cap(L\setminus H)|}{|B|} \ge q^{-s} - q^{-s-1} - 2\varepsilon > 0,$$

and hence $B \cap (L \setminus H) \neq \emptyset$. Since $0 \in H$, this implies $(B \setminus \{0\}) \cap (L \setminus H) \neq \emptyset$, and therefore $\widetilde{B} \cap (\widetilde{L} \setminus \widetilde{H}) \neq \emptyset$. This proves Item 2. $\square$

Theorem 6.6 shows that explicit constructions of ε-biased sets with sufficiently small bias yield explicit constructions of affine and strong s-blocking sets. We now briefly recall the explicit constructions of ε-biased sets over $\mathbb{F}_q^k$ that are relevant for our purposes. For $q = 2$, a breakthrough result of Ta-Shma [TS17] shows that there exist explicit ε-biased sets $S \subseteq \mathbb{F}_2^k$ with $|S| = O\left(\frac{k}{\varepsilon^{2+o(1)}}\right)$. Jalan and Moshkovitz [JM21] extended Ta-Shma's construction to every prime power q as follows.

Theorem 6.7 ([JM21]). *There exists an explicit ε-biased set $S \subseteq \mathbb{F}_q^k$ with*

$$|S| \leq \frac{C_1 k (\log q)^{C_2}}{\varepsilon^{2+\alpha}},$$

where

$$\alpha \leq C_3 \left(\frac{\log\log(1/\varepsilon)}{\log(1/\varepsilon)}\right)^{1/3} = o_{1/\varepsilon}(1),$$

and $C_1, C_2, C_3 > 0$ are absolute constants.

Combining Theorem 6.6 and Theorem 6.7, we obtain the main result of this section.

Proposition 6.8 (Formal version of Theorem 1.15). *For every pair of positive integers $s < k$ and every prime power $q > 1$, there exist an explicit affine s-blocking set in $\mathbb{F}_q^k$ and an explicit strong s-blocking set in $\mathrm{PG}(k-1, q)$, both of size at most*

$$C_1\, k\, (\log q)^{C_2}\, q^{2s + C_3 s \left(\frac{\log(s \log q)}{s \log q}\right)^{1/3}},$$

where $C_1, C_2, C_3 > 0$ are absolute constants.

Proof. Choose $\varepsilon = \Theta(q^{-s})$ sufficiently small so that $\varepsilon < q^{-s}$ and $\varepsilon < \frac{q-1}{2q^{s+1}}$. Let $B \subseteq \mathbb{F}_q^k$ be an explicit ε-biased set of size at most $C_1\, k\, (\log q)^{C_2}\, q^{2s + C_3 s \left(\frac{\log(s \log q)}{s \log q}\right)^{1/3}}$, whose existence is guaranteed by Theorem 6.7. Note that $|\widetilde{B}| \leq |B|$. Applying Theorem 6.6, we conclude that B is an affine s-blocking set and that $\widetilde{B}$ is a strong s-blocking set. □

7 Concluding Remarks and Open Problems

We conclude by highlighting several open problems that arise naturally from our work.

1. One of the most interesting open problems is whether there exist, for $r \ll k$, explicit *strong* (r, A) subspace designs in $\mathbb{F}_q^k$ with $A = O(r(k-r))$, where the field size q depends only on r (e.g., $q = \mathrm{poly}(r)$) and is independent of k. Theorem 1.8 achieves this for weak subspace designs over non-prime finite fields. It would be very interesting to know whether a similar result can be obtained for strong subspace designs. We note that Guruswami, Xing, and Yuan [GXY18] constructed nontrivial explicit strong subspace designs using function fields (specifically, cyclotomic function fields), but their parameters fall short of the above regime.

2. The gap between the best-known upper and lower bounds on the sizes of strong and affine s-blocking sets remains roughly a factor of $O(s)$. Closing this gap would be of significant interest.

3. For every sufficiently large non-prime prime power q, we constructed explicit strong s-blocking sets of size $O(skq^s)$, matching the best-known existential bound up to an absolute constant factor. In contrast, when q is prime or a small prime power, our explicit constructions have a worse dependence on s in the leading coefficient. It would be very interesting to obtain improved explicit constructions in these cases as well.

4. In particular, when $q = O(1)$, do there exist explicit strong s-blocking sets of size $O(C(s)kq^s)$ with $C(s)$ subexponential in s? Note that Theorem 1.15 achieves $C(s) = \exp(O(s))$ in this regime.

Acknowledgments

We thank discussions with Abhranil Chatterjee, Sumanta Ghosh, Rohit Gurjar, Ran Tao, Ben Lee Volk, and Chaoping Xing. An early discussion of this project took place at the 2025 Oberwolfach workshop "New Mathematical Directions in Coding Theory," and we thank the institute for its hospitality and the organizers for the opportunity to participate.

References

[ABDN24] Noga Alon, Anurag Bishnoi, Shagnik Das, and Alessandro Neri. Strong blocking sets and minimal codes from expander graphs. *Transactions of the American Mathematical Society*, 377(08):5389–5410, 2024. doi:10.1090/tran/9205.

[ABN22] Gianira N. Alfarano, Martino Borello, and Alessandro Neri. A geometric characterization of minimal codes and their asymptotic performance. *Advances in Mathematics of Communications*, 16(1):115–133, 2022. doi:10.3934/amc.2020104.

[AGKS15] Manindra Agrawal, Rohit Gurjar, Arpita Korwar, and Nitin Saxena. Hitting-sets for ROABP and sum of set-multilinear circuits. *SIAM Journal on Computing*, 44(3):669–697, 2015. doi:10.1137/140975103.

[Bal11] Simeon Ball. The polynomial method in Galois geometries. *Current research topics in Galois geometry*, pages 105–130, 2011.

[BBGS15] Alp Bassa, Peter Beelen, Arnaldo Garcia, and Henning Stichtenoth. Towers of function fields over non-prime finite fields. *Moscow Mathematical Journal*, 15(1):1–29, 2015. doi:10.17323/1609-4514-2015-15-1-1-29.

[BCDZ25a] Joshua Brakensiek, Yeyuan Chen, Manik Dhar, and Zihan Zhang. Combinatorial bounds for list recovery via discrete Brascamp–Lieb inequalities, 2025. arXiv:2510.13775.

[BCDZ25b] Joshua Brakensiek, Yeyuan Chen, Manik Dhar, and Zihan Zhang. From random to explicit via subspace designs with applications to local properties and matroids, 2025. arXiv:2510.13777.

[BDGP24] Anurag Bishnoi, Jozefien D'haeseleer, Dion Gijswijt, and Aditya Potukuchi. Blocking sets, minimal codes and trifferent codes. *Journal of the London Mathematical Society*, 109(6):e12938, 2024. doi:10.1112/jlms.12938.

[BS78] Andries E. Brouwer and Alexander Schrijver. The blocking number of an affine space. *Journal of Combinatorial Theory, Series A*, 24(2):251–253, 1978. doi:10.1016/0097-3165(78)90013-4.

[BT26] Anurag Bishnoi and István Tomon. Explicit constructions of optimal blocking sets and minimal codes. *Combinatorica*, 46(2):13, 2026. doi:10.1007/s00493-026-00202-5.

[CL85] Gérard Cohen and Abraham Lempel. Linear intersecting codes. *Discrete Mathematics*, 56(1):35–43, 1985. doi:10.1016/0012-365X(85)90190-6.

[CZ25] Yeyuan Chen and Zihan Zhang. Explicit folded Reed-Solomon and multiplicity codes achieve relaxed generalized Singleton bounds. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pages 1–12, 2025. doi:10.1145/3717823.3718114.

[DGW09] Zeev Dvir, Ariel Gabizon, and Avi Wigderson. Extractors and rank extractors for polynomial sources. *computational complexity*, 18(1):1–58, 2009. doi:10.1007/s00037-009-0258-4.

[Dvi12] Zeev Dvir. Extractors for varieties. *Computational complexity*, 21(4):515–572, 2012. doi:10.1007/s00037-011-0023-3.

[FG14] Michael A. Forbes and Venkatesan Guruswami. Dimension expanders via rank condensers, 2014. Conference version in RANDOM 2015. arXiv:1411.7455.

[For14] Michael Andrew Forbes. *Polynomial identity testing of read-once oblivious algebraic branching programs*. PhD thesis, Massachusetts Institute of Technology, 2014.

[FS12] Michael A. Forbes and Amir Shpilka. On identity testing of tensors, low-rank recovery and compressed sensing. In *Proceedings of the 44th Annual ACM Symposium on Theory of Computing*, pages 163–172, 2012. doi:10.1145/2213977.2213995.

[FS13] Michael A. Forbes and Amir Shpilka. Quasipolynomial-time identity testing of non-commutative and read-once oblivious algebraic branching programs. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 243–252, 2013. doi:10.1109/FOCS.2013.34.

[FS14] Szabolcs L. Fancsali and Péter Sziklai. Lines in higgledy-piggledy arrangement. *The Electronic Journal of Combinatorics*, 21(2):P2.56, 2014. doi:10.37236/4149.

[FS16] Szabolcs L. Fancsali and Péter Sziklai. Higgledy-piggledy subspaces and uniform subspace designs. *Designs, Codes and Cryptography*, 79(3):625–645, 2016. doi:10.1007/s10623-016-0189-4.

[FSS14] Michael A. Forbes, Ramprasad Saptharishi, and Amir Shpilka. Hitting sets for multilinear read-once algebraic branching programs, in any order. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 867–875, 2014. doi:10.1145/2591796.2591816.

[Für88] Zoltán Füredi. Matchings and covers in hypergraphs. *Graphs and Combinatorics*, 4(1):115–206, 1988. doi:10.1007/BF01864160.

[GG25] Rohan Goyal and Venkatesan Guruswami. Optimal proximity gaps for subspace-design codes and (random) Reed-Solomon codes. Cryptology ePrint Archive, Paper 2025/2054, 2025. URL: https://eprint.iacr.org/2025/2054.

[GK16] Venkatesan Guruswami and Swastik Kopparty. Explicit subspace designs. *Combinatorica*, 36(2):161–185, 2016. doi:10.1007/s00493-014-3169-1.

[GR08] Ariel Gabizon and Ran Raz. Deterministic extractors for affine sources over large fields. *Combinatorica*, 28(4):415–440, 2008. doi:10.1007/s00493-008-2259-3.

[GRX21] Venkatesan Guruswami, Nicolas Resch, and Chaoping Xing. Lossless dimension expanders via linearized polynomials and subspace designs. *Combinatorica*, 41(4):545–579, 2021. doi:10.1007/s00493-020-4360-1.

[GS95] Arnaldo Garcia and Henning Stichtenoth. A tower of Artin-Schreier extensions of function fields attaining the Drinfeld-Vladut bound. *Inventiones mathematicae*, 121(1):211–222, 1995. doi:10.1007/BF01884295.

[GS96] Arnaldo Garcia and Henning Stichtenoth. On the asymptotic behaviour of some towers of function fields over finite fields. *Journal of Number Theory*, 61(2):248–273, 1996. doi:10.1006/jnth.1996.0147.

[GT20] Rohit Gurjar and Thomas Thierauf. Linear matroid intersection is in quasi-NC. *computational complexity*, 29(2):9, 2020. doi:10.1007/s00037-020-00200-z.

[Guo24] Zeyu Guo. Variety evasive subspace families. *computational complexity*, 33(2):10, 2024. doi:10.1007/s00037-024-00256-1.

[GVJZ23] Zeyu Guo, Ben Lee Volk, Akhil Jalan, and David Zuckerman. Extractors for images of varieties. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 46–59, 2023. doi:10.1145/3564246.3585109.

[GX13] Venkatesan Guruswami and Chaoping Xing. List decoding Reed-Solomon, algebraic-geometric, and Gabidulin subcodes up to the Singleton bound. In *Proceedings of the 45th Annual ACM Symposium on Theory of Computing*, pages 843–852, 2013. doi:10.1145/2488608.2488715.

[GX22] Venkatesan Guruswami and Chaoping Xing. Optimal rate list decoding over bounded alphabets using algebraic-geometric codes. *Journal of the ACM*, 69(2):1–48, 2022. doi:10.1145/3506668.

[GXY18] Venkatesan Guruswami, Chaoping Xing, and Chen Yuan. Subspace designs based on algebraic function fields. *Transactions of the American Mathematical Society*, 370(12):8757–8775, 2018. doi:10.1090/tran/7369.

[Har77] Robin Hartshorne. *Algebraic Geometry*, volume 52 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1977. doi:10.1007/978-1-4757-3849-0.

[Hes02] Florian Hess. Computing Riemann–Roch spaces in algebraic function fields and related topics. *Journal of Symbolic Computation*, 33(4):425–445, 2002. doi:10.1006/jsco.2001.0513.

[Jam77] Robert E. Jamison. Covering finite fields with cosets of subspaces. *Journal of Combinatorial Theory, Series A*, 22(3):253–266, 1977. doi:10.1016/0097-3165(77)90001-2.

[JLR26] Fernando Granha Jeronimo, Lenny Liu, and Pranav Rajpal. Optimal proximity gap for folded Reed–Solomon codes via subspace designs, 2026. arXiv:2601.10047.

[JM21] Akhil Jalan and Dana Moshkovitz. Near-optimal Cayley expanders for abelian groups. In *41st IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2021)*, pages 24:1–24:23, 2021. doi:10.4230/LIPIcs.FSTTCS.2021.24.

[Juk11] Stasys Jukna. *Extremal Combinatorics: With Applications in Computer Science*, volume 571. Springer, 2011. doi:10.1007/978-3-642-17364-6.

[KS09] Neeraj Kayal and Shubhangi Saraf. Blackbox polynomial identity testing for depth 3 circuits. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 198–207, 2009. doi:10.1109/FOCS.2009.67.

[KS11] Zohar S. Karnin and Amir Shpilka. Black box polynomial identity testing of generalized depth-3 arithmetic circuits with bounded top fan-in. *Combinatorica*, 31(3):333–364, 2011. doi:10.1007/s00493-011-2537-3.

[O'D14] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014. doi:10.1017/CBO9781139814782.

[Ric56] Moses Richardson. On finite projective games. *Proceedings of the American Mathematical Society*, 7(3):458–465, 1956. doi:10.2307/2032754.

[SAK+02] Kenneth W. Shum, Ilia Aleshnikov, P. Vijay Kumar, Henning Stichtenoth, and Vinay Deolalikar. A low-complexity algorithm for the construction of algebraic-geometric codes better than the Gilbert-Varshamov bound. *IEEE Transactions on Information Theory*, 47(6):2225–2241, 2002. doi:10.1109/18.945244.

[Ser20] Jean-Pierre Serre. *Rational Points on Curves over Finite Fields*. Société mathématique de France, 2020.

[SS12] Nitin Saxena and Comandur Seshadhri. Blackbox identity testing for bounded top-fanin depth-3 circuits: The field doesn't matter. *SIAM Journal on Computing*, 41(5):1285, 2012. doi:10.1137/10848232.

[SS13] Nitin Saxena and Comandur Seshadhri. From Sylvester-Gallai configurations to rank bounds: Improved blackbox identity test for depth-3 circuits. *Journal of the ACM*, 60(5):1–33, 2013. doi:10.1145/2528403.

[Sti01] Henning Stichtenoth. Explicit constructions of towers of function fields with many rational places. In *European Congress of Mathematics*, pages 219–224. Birkhäuser Basel, 2001. doi:10.1007/978-3-0348-8266-8_17.

[Sti09] Henning Stichtenoth. *Algebraic Function Fields and Codes*. Springer, 2009. doi:10.1007/978-3-540-76878-4.

[SZ23] Paolo Santonastaso and Ferdinando Zullo. On subspace designs. *EMS Surveys in Mathematical Sciences*, 11(1):1–62, 2023. doi:10.4171/EMSS/77.

[TQLZ21] Chunming Tang, Yan Qiu, Qunying Liao, and Zhengchun Zhou. Full characterization of minimal linear codes as cutting blocking sets. *IEEE Transactions on Information Theory*, 67(6):3690–3700, 2021. doi:10.1109/TIT.2021.3070377.

[TS17] Amnon Ta-Shma. Explicit, almost optimal, epsilon-balanced codes. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 238–251, 2017. doi:10.1145/3055399.3055408.

[TV06] Terence Tao and Van H. Vu. *Additive Combinatorics*. Cambridge University Press, 2006. doi:10.1017/CBO9780511755149.

[VD83] S. G. Vlăduţ and V. G. Drinfeld. Number of points of an algebraic curve. *Functional Analysis and Its Applications*, 17(1):53–54, 1983. doi:10.1007/BF01083182.

[XKH25] Yang Xu, Haibin Kan, and Guangyue Han. r-minimal codes with respect to rank metric. *IEEE Transactions on Information Theory*, 2025. doi:10.1109/TIT.2025.3585604.

A Existence of Non-Explicit Lossless Rank Extractors

The main result of this section is the following theorem, establishing the existence of non-explicit lossless rank extractors with good parameters over any finite field $\mathbb{F}_q$.

Theorem A.1. *There exists a function $\delta^* : \{2, 3, 4, \dots\} \to (0, 1)$ such that $\lim_{q\to\infty} \delta^*(q) = 0$ and the following holds. For every prime power q and all integers $k \geq r > 0$, there exist (k, r, L) lossless rank extractors over $\mathbb{F}_q$ of size n such that*

$$L = O(r(k - r)), \qquad L/n \leq \delta^*(q).$$

The proof is based on the probabilistic method. A similar result was obtained in [For14, Section 5.3], but for a weaker notion of lossless rank extractors, which coincides with our notion of lossless rank dispersers (Definition 1.6).

We will use the following bound.

Lemma A.2. *Let $q \geq 2$. Let $r > 0$ be an integer. Then $\prod_{i=1}^{r}(1 - q^{-i}) \geq e^{-2/(q-1)}$.*

Proof. For $x \in [0, 1/2]$, we claim that $\log(1 - x) \geq -2x$. Indeed, let $f(x) := \log(1 - x) + 2x$. Then $f'(x) = 2 - \frac{1}{1-x} = \frac{1-2x}{1-x} \geq 0$ for all $x \in [0, 1/2]$. Since $f(0) = 0$, it follows that $f(x) \geq 0$ for $x \in [0, 1/2]$, proving the claim.

Applying the claim to $x = q^{-i}$ gives $\log(1 - q^{-i}) \geq -2q^{-i}$ for $i \geq 1$. Summing over $i \in [r]$, we obtain

$$\sum_{i=1}^{r} \log(1 - q^{-i}) \geq -2 \sum_{i=1}^{r} q^{-i} \geq -2 \sum_{i=1}^{\infty} q^{-i} = -2/(q - 1).$$

Exponentiating both sides proves the lemma. □

We now prove Theorem A.1.

Proof of Theorem A.1. Let $M \in \mathbb{F}_q^{k\times r}$ be a full-rank matrix. Let $E \in \mathbb{F}_q^{r\times k}$ be uniformly random, and let $u_1, \dots, u_r \in \mathbb{F}_q^k$ denote its rows. Then the vectors $u_1, \dots, u_r$ are independent and uniformly

distributed over $\mathbb{F}_q^k$. Since M has full rank, each $u_i M$ is uniformly distributed over $\mathbb{F}_q^r$, and these vectors remain independent. Hence EM is uniformly distributed over $\mathbb{F}_q^{r\times r}$. Therefore,

$$\Pr[\operatorname{rank}(EM) < r] = 1 - \frac{\prod_{j=0}^{r-1}(q^r - q^j)}{q^{r^2}} = 1 - \prod_{i=1}^{r}(1 - q^{-i}) \le 1 - e^{-2/(q-1)}, \tag{13}$$

where the last inequality follows from Lemma A.2.

Let S be the set of full-rank matrices in $\mathbb{F}_q^{k\times r}$. The group $\mathrm{GL}(r,q)$ acts on S by right multiplication, and each orbit has size $\prod_{j=0}^{r-1}(q^r - q^j)$. Let $M_1, \dots, M_k \in S$ be a complete set of orbit representatives. Then

$$k = \frac{|S|}{\prod_{j=0}^{r-1}(q^r - q^j)} = \prod_{j=0}^{r-1}\frac{q^k - q^j}{q^r - q^j} \le q^{r(k-r)} \prod_{i=1}^{r}\frac{1}{1 - q^{-i}} \le q^{r(k-r)} e^{2/(q-1)}, \tag{14}$$

again by Lemma A.2.

Now let $E_1, \dots, E_n$ be independent uniformly random matrices in $\mathbb{F}_q^{r\times k}$. Fix $M \in \{M_1, \dots, M_k\}$. By (13) and a union bound,

$$\Pr\big[\#\{i \in [n] : \operatorname{rank}(E_i M) < r\} > L\big] \le \binom{n}{L+1}\big(1 - e^{-2/(q-1)}\big)^{L+1}. \tag{15}$$

Assume that

$$q^{r(k-r)} e^{2/(q-1)} \cdot \binom{n}{L+1}\big(1 - e^{-2/(q-1)}\big)^{L+1} < 1. \tag{16}$$

Then (14), (15), and a union bound over $M_1, \dots, M_k$ imply that there exist matrices $E_1, \dots, E_n \in \mathbb{F}_q^{r\times k}$ such that $\#\{i \in [n] : \operatorname{rank}(E_i M) < r\} \le L$ for every $M \in \{M_1, \dots, M_k\}$. Fix such a collection.

We claim that $\{E_1, \dots, E_n\}$ is a (k, r, L) lossless rank extractor. Indeed, let $M \in \mathbb{F}_q^{k\times r}$ be any full-rank matrix. Since M lies in the same $\mathrm{GL}(r,q)$-orbit as some representative M_j, we can write $M = M_j G$ for some $j \in [k]$ and some $G \in \mathrm{GL}(r,q)$. Then for every $i \in [n]$, $\operatorname{rank}(E_i M) = \operatorname{rank}(E_i M_j G) = \operatorname{rank}(E_i M_j)$, and hence

$$\#\{i \in [n] : \operatorname{rank}(E_i M) < r\} = \#\{i \in [n] : \operatorname{rank}(E_i M_j) < r\} \le L.$$

This proves the claim.

It remains to choose $\delta^*(\cdot)$, L, and n so that (16) holds and the stated bounds are satisfied. Taking logarithms, (16) is equivalent to

$$r(k-r)\ln q + \frac{2}{q-1} + \ln\binom{n}{L+1} + (L+1)\ln\big(1 - e^{-2/(q-1)}\big) < 0. \tag{17}$$

Let

$$C_\delta := (\delta^{-1} - 1)\ln\left(\frac{\delta^{-1} e}{\delta^{-1} - 1}\right), \qquad \delta \in (0,1).$$

Define $\delta^*(q)$ to be the smallest $\delta \in (0,1)$ such that

$$C_\delta \le -\frac{1}{2}\ln\big(1 - e^{-2/(q-1)}\big).$$

This is well defined by continuity, since $C_\delta \to +\infty$ as $\delta \to 0^+$ and $C_\delta \to 0$ as $\delta \to 1^-$. Moreover,

$$\lim_{q\to\infty} \delta^*(q) = 0,$$

since $-\ln\bigl(1 - e^{-2/(q-1)}\bigr) = \Theta(\log q)$ as $q \to \infty$.

Write δ^* for $\delta^*(q)$ and set $n = \lceil L/\delta^* \rceil$. Then $L \le \delta^* n$ and also $n \le (\delta^*)^{-1}(L+1)$. Hence

$$\ln \binom{n}{L+1} \le \ln \binom{(\delta^*)^{-1}(L+1)}{L+1} = \ln \binom{(\delta^*)^{-1}(L+1)}{(\delta^*)^{-1}(L+1) - (L+1)} \le C_{\delta^*}(L+1). \tag{18}$$

By definition, $C_{\delta^*} \le -\frac{1}{2}\ln\bigl(1 - e^{-2/(q-1)}\bigr)$. Combining this with (18), we obtain

$$\ln \binom{n}{L+1} + (L+1)\ln\bigl(1 - e^{-2/(q-1)}\bigr) \le \frac{1}{2}(L+1)\ln\bigl(1 - e^{-2/(q-1)}\bigr). \tag{19}$$

Substituting into (17), it suffices to choose L such that

$$r(k-r)\ln q + \frac{2}{q-1} + \frac{1}{2}(L+1)\ln\bigl(1 - e^{-2/(q-1)}\bigr) < 0.$$

Since $-\ln(1 - e^{-2/(q-1)}) = \Theta(\log q)$, this holds whenever $L = c\,r(k-r)$ for a sufficiently large constant c. Thus we may choose $L = O(r(k-r))$. □

Remark A.3. *Using sharper estimates in the above argument, one can strengthen the conclusion to $L = (1 + o_q(1))\, r(k-r)$. We omit the details.*

B δ-Hitting Sets for Symbolic Determinants with Rank-One Summands

In this section, we prove Lemma 5.6, which constructs explicit δ-hitting sets for symbolic determinants with rank-one summands in $\mathbb{F}_q^N$ over a sufficiently large field $\mathbb{F}_q$. For convenience, we first restate the lemma.

Lemma 5.6. *Let N be a positive integer, $\delta \in (0,1)$, and $\mathbb{F}$ a field such that $|\mathbb{F}| \ge (N/\delta)^c$ for some large enough absolute constant $c > 0$. Then, one can construct an explicit δ-hitting set $\mathcal{H} \subseteq \mathbb{F}^N$ for $\mathrm{VBP}^1_{r,N,\mathbb{F}}$ of size polynomial in $(N/\delta)^{\log N}$.*

Let $N \ge 1$ be an integer, let $x = \{x_1, \ldots, x_N\}$ be variables, and let $\mathbb{F}$ be a field. Consider symbolic matrices of the form

$$A(x) = \sum_{i\in[N]} x_i A_i \in \mathbb{F}[x_1, \ldots, x_N]^{r\times r},$$

where $A_i \in \mathbb{F}^{r\times r}$ and $\mathrm{rank}(A_i) \le 1$ for all $i \in [N]$. Since each A_i has rank at most one, it can be written as $u_i v_i^T$ for some vectors $u_i, v_i \in \mathbb{F}^r$. Let $U, V \in \mathbb{F}^{r\times N}$ such that for $i \in [N]$, the i-th column of U and that of V are u_i and v_i, respectively. Let X be the $N \times N$ diagonal matrix whose i-th diagonal entry is the variable x_i for $i \in [N]$. It can be easily shown that $A = UXV^T$. By the Cauchy–Binet formula,

$$\det(A) = \det(UXV^T) = \sum_{S\subseteq[N],|S|=r} \det((UX)_S)\det(V_S) = \sum_{S\subseteq[N],|S|=r} \det(U_S)\det(V_S)\prod_{i\in S} x_i.$$

For a subset $S \subseteq [N]$, let x_S denote $\prod_{i \in S} x_i$. Note that the coefficient of x_S in A is nonzero if and only if $\det(U_S)$ and $\det(V_S)$ are nonzero. In other words, the coefficient of x_S in A is nonzero if and only if S is a common base of the linear matroids represented by the matrices U and V. We denote these two matroids by $\mathcal{M}_1^A$ and $\mathcal{M}_2^A$. For A as above, define the support

$$\mathrm{supp}(A) := \{S \subseteq [N] \mid \text{the coefficient of } x_S \text{ in } \det(A) \text{ is nonzero}\}.$$

From above discussion, $\mathrm{supp}(A)$ is the set of common bases of the linear matroids represented by U and V.

For $u, v \in \mathbb{Z}^t$, write $u > v$ if u is lexicographically larger than v.

We also need the notion of *isolating weight assignments*.

Definition B.1 (Isolating weight assignments)**.** Let $t \leq N$, and let $A = \sum_{i \in [N]} x_i A_i$ with $A_i \in \mathbb{F}^{r \times r}$ and $\mathrm{rank}(A_i) \leq 1$ for $i \in [N]$. Let $w = (w_1, \ldots, w_N) \in (\mathbb{Z}^t_{\geq 0})^N$. For $S \subseteq [N]$, define $w(S) := \sum_{i \in S} w_i \in \mathbb{Z}^t_{\geq 0}$. We say that w is *isolating* for A if there exists a unique set $S \in \mathrm{supp}(A)$ maximizing $w(S)$ lexicographically. A set $\mathcal{W} \subseteq (\mathbb{Z}^t_{\geq 0})^N$ is isolating for A if it contains such a vector w.

The importance of isolating weight assignments is captured by the following claim.

Claim B.2. *Let $A(x) = \sum_{i \in [N]} x_i A_i$ with $\mathrm{rank}(A_i) \leq 1$ and $\det(A) \neq 0$. Let $w \in (\mathbb{Z}^t_{\geq 0})^N$ be isolating for A. Let $y = \{y_1, \ldots, y_t\}$ be variables, and define $A_w(y) \in \mathbb{F}[y_1, \ldots, y_t]^{r \times r}$ by substituting*

$$x_i \mapsto \prod_{j \in [t]} y_j^{w_i[j]},$$

where $w_i[j]$ denotes the j-th coordinate of w_i. Then $\det(A_w(y)) \neq 0$.

Proof. Write $\det(A) = \sum_{S \in \mathrm{supp}(A)} c_S x_S$ with $c_S \neq 0$. Under the substitution, the monomial x_S maps to $\prod_{j \in [t]} y_j^{w(S)[j]}$. Since w is isolating for A, there is a unique subset $S \subseteq [N]$ maximizing $w(S)$ lexicographically, so the corresponding monomial $\prod_{j \in [t]} y_j^{w(S)[j]}$ is not canceled by the others. □

We follow the construction of Gurjar and Thierauf [GT20], which yields isolating weight assignments w with $t = 1$, i.e., $\det(A_w(y))$ is univariate. However, the resulting weights can be quasi-polynomial in N, which would force us to work over a finite field $\mathbb{F}_q$ of size quasi-polynomial in N. To avoid this, we adapt their method and set $t = \Theta(\log N)$, which allows us to use polynomially bounded weights.

As mentioned earlier, $\mathrm{supp}(A)$ is a set of common bases of two linear matroids. We take a slight detour and discuss the result of [GT20] on the construction of weight isolation assignments for matroid intersection.

Matroid Intersection Polytope. Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two matroids of rank r on common ground set E with rank function r_1 and r_2. Then, the convex hull of characteristic vectors of common bases in $\mathbb{R}^{|E|}$, denoted by $P(\mathcal{M}_1, \mathcal{M}_2)$, is described by the following linear program:

$$x_e \geq 0 \text{ for each } e \in E, \quad x(S) = \sum_{e \in S} x_e \leq r_i(S) \text{ for each } S \subseteq E,\ i \in \{1, 2\}, \quad x(E) = r.$$

The following lemma [GT20, Lemma 3.5] gives a characterization for the tight inequalities in the LP for faces of the matroid intersection polytope.

Lemma B.3. *Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two matroids on a common ground set E with rank functions r_1 and r_2, respectively. Let F be a face of $P(\mathcal{M}_1, \mathcal{M}_2)$. Then, there exists two partitions $\mathcal{P}_1, \mathcal{P}_2$ of E such that for any $S_1 \in \mathcal{P}_1$ and $S_2 \in \mathcal{P}_2$, there exists $\nu_{S_1}, \mu_{S_2} \in \mathbb{N}$ such that for each $x \in F$, $x(S_1) = \nu_{S_1}$ and $x(S_2) = \mu_{S_2}$.*

Moreover,

1. *If for some $T \subseteq E$, $x(T) = r_1(T)$ for all $x \in F$, or $x(T) = r_2(T)$ for all $x \in F$, then T is the union of sets from $\mathcal{P}_1$, respectively $\mathcal{P}_2$.*
2. *If for some $e \in E$, $x_e = 0$ for all $e \in F$, then there is a $S_1 \in \mathcal{P}_1$ and $S_2 \in \mathcal{P}_2$ such that $S_1 = S_2 = \{e\}$ and $\nu_{S_1} = \mu_{S_2} = 0$.*

Definition B.4. Let F be a face of the polytope $P(\mathcal{M}_1, \mathcal{M}_2)$ with the partitions $\mathcal{P}_1, \mathcal{P}_2$ from Lemma B.3. A sequence $C = (e_1, e_2, \ldots, e_{2\ell})$ of distinct elements of E is called a *cycle* with respect to face F, if consecutive pairs are alternately in a set from $\mathcal{P}_1$ and a set from $\mathcal{P}_2$. That is, for $i \in [\ell]$,

$$e_{2i-1}, e_{2i} \in S_{2i-1} \text{ for some } S_{2i-1} \in \mathcal{P}_1,$$
$$e_{2i}, e_{2i+1} \in S_{2i} \text{ for some } S_{2i} \in \mathcal{P}_2.$$

We denote the set of all cycles with respect to a face F as $\mathcal{C}_F$. For a vector $u \in \mathbb{Z}^E$, let F be the face of $P(\mathcal{M}_1, \mathcal{M}_2)$ that maximizes $u^T \cdot y$ over all the points y in the polytope. Then, we denote $\mathcal{C}_F$ by $\mathcal{C}_u$. For a bipartite graph G with edge set E and a vector $u \in \mathbb{Z}^E$, the *circulation* of a cycle $C = (e_1, e_2, \ldots, e_{2\ell})$ with respect to u, denoted by $u(C)$, is defined by

$$u(C) := \left| \sum_{i=1}^{2\ell} (-1)^i u[e_i] \right|.$$

The following lemma summarizes key properties of $\mathcal{C}_u$. Additionally, it establishes a sufficient condition for a weight assignment on the ground set to guarantee the existence of a unique maximum-weight common base. Such a weight assignment is said to be *isolating* for the two matroids.

Lemma B.5 ([GT20, Corollary 3.10 and Lemma 3.14])**.** *Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two matroids defined on a common ground set E of size N, and let $u \in \mathbb{Z}^E$.*

- *If $\mathcal{C}_u$ contains no cycles of length at most c, then the number of cycles in $\mathcal{C}_u$ of length at most $2c$ is at most N^4.*
- *If $\mathcal{C}_u$ is empty, then u is isolating for $\mathcal{M}_1$ and $\mathcal{M}_2$.*

The following lemma combines Claim 3.18 and Lemma 3.19 of [GT20].

Lemma B.6. *Let $\mathcal{M}_1, \mathcal{M}_2$ be two matroids on a common ground set E of size N. Let $t = \lceil \log N \rceil$, and let H be a positive integer. Let $W_0 \in \mathbb{Z}^N$ be the all-ones vector.*

For each $i \in [t]$, let $u_i \in \mathbb{Z}^N_{\geq 0}$ satisfy $H > N \max_{j \in [N]} u_i[j]$ and

$$u_i(C) \neq 0 \quad \text{for every cycle } C \text{ in } \mathcal{C}_{W_{i-1}} \text{ of length at most } 2^i.$$

Define $W_i = \sum_{j=1}^{i} H^{i-j} u_j$. Then $\mathcal{C}_{W_i}$ contains no cycles of length at most 2^i for every $i \in [t]$. Moreover, W_t is an isolating weight assignment for the set of common bases of $\mathcal{M}_1$ and $\mathcal{M}_2$.

For a symbolic matrix A with rank-one summands, if we apply the above lemma for matroids $\mathcal{M}_1^A$ and $\mathcal{M}_2^A$, we get an N-tuple of vectors in one dimension that is isolating for A. The following corollary directly follows from above lemma and gives an N-tuple of vectors in $t = \lceil \log N \rceil$ dimension that is isolating for A.

Corollary B.7. *Let $A(x) = \sum_{i\in[N]} x_i A_i$ with $\operatorname{rank}(A_i) \le 1$ for all $i \in [N]$, and suppose $\det(A) \neq 0$. Let $t = \lceil \log N \rceil$, and let H be a positive integer. Let $W_0 \in \mathbb{Z}^N$ be the all-ones vector. Let $\mathcal{M}_1^A$ and $\mathcal{M}_2^A$ be matroids such that $\operatorname{supp}(A)$ is the set of their common bases.*

For each $i \in [t]$, let $u_i \in \mathbb{Z}^N_{\ge 0}$ satisfy $H > N \max_{j\in[N]} u_i[j]$ and

$$u_i(C) \neq 0 \quad \text{for every cycle } C \text{ in } \mathcal{C}_{W_{i-1}} \text{ of length at most } 2^i.$$

Define $W_i = \sum_{j=1}^{i} H^{i-j} u_j$. Then $\mathcal{C}_{W_i}$ contains no cycles of length at most 2^i for all $i \in [t]$.

Moreover, for each $i \in [N]$, define $w_i \in \mathbb{Z}^t_{\ge 0}$ by $w_i[j] = u_j[i]$ for all $j \in [t]$. Then $w = (w_1, \dots, w_N) \in (\mathbb{Z}^t_{\ge 0})^N$ is isolating for A.

Proof. By Lemma B.6, the weight vector $W_t = \sum_{j=1}^{t} H^{t-j} u_j$ admits a unique maximum-weight common base for $\mathcal{M}_1^A$ and $\mathcal{M}_2^A$.

For a subset $S \subseteq [N]$, write $w(S) = \sum_{i\in S} w_i \in \mathbb{Z}^t_{\ge 0}$, and compare such vectors in lexicographic order. By definition,

$$W_t(S) = \sum_{j=1}^{t} H^{t-j} \sum_{i\in S} u_j[i].$$

Since $H > N \max_{i,j} u_j[i]$, there is no carry between the coefficients, and therefore for any two subsets $S_1, S_2 \subseteq [N]$,

$$W_t(S_1) > W_t(S_2) \quad \Longleftrightarrow \quad w(S_1) > w(S_2)$$

in lexicographic order.

Thus, the unique maximum-weight common base under W_t is also the unique maximum under w, and hence w is isolating for A. □

Lemma B.8. *Let N be a positive integer, $t = \lceil \log N \rceil$, and $0 < \varepsilon < 1$. Let $\mathcal{U} = \{u_1, u_2, \dots, u_D\} \subseteq \mathbb{Z}^N_{\ge 0}$ be a collection such that for any graph G with at most N edges and any set $\mathcal{C}$ of at most N^4 cycles of G,*

$$\Pr_{u\sim\mathcal{U}}[u(C) \neq 0 \ \forall C \in \mathcal{C}] \ge 1 - \varepsilon.$$

Define $\mathcal{W} \subseteq (\mathbb{Z}^t_{\ge 0})^N$ as follows: for each $T = (a_1, \dots, a_t) \in [D]^t$, let $w^T = (w_1^T, \dots, w_N^T)$ with

$$w_i^T[j] = u_{a_j}[i] \quad \text{for all } i \in [N],\ j \in [t].$$

Then $\mathcal{W} = \{w^T : T \in [D]^t\}$.

Then, for any $A(x) = \sum_{i\in[N]} x_i A_i$ with $\operatorname{rank}(A_i) \le 1$ and $\det(A) \neq 0$, at least a $(1-\varepsilon)^t$ fraction of $w \in \mathcal{W}$ are isolating for A.

Proof. Let $T = (a_1, a_2, \dots, a_t) \sim [D]^t$ be chosen uniformly and independently. Let H be an integer greater than $N \max_{i\in[t],\, j\in[N]} u_{a_i}[j]$. Let $W_0 \in \mathbb{Z}^N$ be the all-ones vector, and for each $i \in [t]$, define $W_i = \sum_{j=1}^{i} H^{i-j} u_{a_j}$.

Let $\mathcal{M}_1^A$ and $\mathcal{M}_2^A$ be the matroids such that $\operatorname{supp}(A)$ is the set of their common bases. By assumption on $\mathcal{U}$, for any set $\mathcal{C}$ of cycles of size at most N^4,

$$\Pr_{u\sim\mathcal{U}}[u(C) \neq 0 \ \forall C \in \mathcal{C}] \ge 1 - \varepsilon.$$

From Lemma B.5, if $\mathcal{C}_{W_{i-1}}$ has no cycles of length at most 2^{i-1}, then it contains at most N^4 cycles of length at most 2^i. Hence, for each $i \in [t]$,

$$\Pr_{a_i \sim [D]}\left[u_{a_i}(C) \neq 0 \ \forall C \in \mathcal{C}_{W_{i-1}},\ |C| \leq 2^i \ \middle|\ \mathcal{C}_{W_{i-1}} \text{ has no cycles of length } \leq 2^{i-1}\right] \geq 1 - \varepsilon.$$

We now prove by induction that for every $i \in \{0, 1, \ldots, t\}$,

$$\Pr_{(a_1,\ldots,a_i)\sim[D]^i}\left[\mathcal{C}_{W_i} \text{ has no cycles of length } \leq 2^i\right] \geq (1-\varepsilon)^i.$$

The base case $i = 0$ is trivial. For the inductive step,

$$\begin{aligned}
&\Pr_{(a_1,\ldots,a_i)}\left[\mathcal{C}_{W_i} \text{ has no cycles of length } \leq 2^i\right] \\
&\geq \Pr\left[u_{a_i}(C) \neq 0 \ \forall C \in \mathcal{C}_{W_{i-1}},\ |C| \leq 2^i \text{ and } \mathcal{C}_{W_{i-1}} \text{ has no cycles of length } \leq 2^{i-1}\right] \\
&= \Pr\left[u_{a_i}(C) \neq 0 \ \forall C \in \mathcal{C}_{W_{i-1}},\ |C| \leq 2^i \mid \mathcal{C}_{W_{i-1}} \text{ has no cycles of length } \leq 2^{i-1}\right] \\
&\quad \times \Pr\left[\mathcal{C}_{W_{i-1}} \text{ has no cycles of length } \leq 2^{i-1}\right] \\
&\geq (1-\varepsilon)\cdot(1-\varepsilon)^{i-1} = (1-\varepsilon)^i,
\end{aligned}$$

where the first inequality holds by Lemma B.6, and the last inequality follows from the conditional bound above and the induction hypothesis.

Thus, with probability at least $(1-\varepsilon)^t$, the graph $\mathcal{C}_{W_t}$ has no cycles of length at most $2^t \geq N$, and hence is empty. By Lemma B.5, W_t is isolating for $\mathcal{M}_1^A$ and $\mathcal{M}_2^A$. By Corollary B.7, this implies that w^T is isolating for A. □

Now we construct the collection $\mathcal{U}$ that will be used together with Lemma B.8 to obtain an isolating set.

Claim B.9. *Let N, s be positive integers, and let $0 < \varepsilon < 1$. Then there exists a collection $\mathcal{U} \subseteq \mathbb{Z}_{\geq 0}^N$ such that for any bipartite graph G with edge set $[N]$ and any set $\mathcal{C}$ of s cycles in G, at least a $(1-\varepsilon)$ fraction of the vectors $u \in \mathcal{U}$ satisfy $u(C) \neq 0$ for every $C \in \mathcal{C}$. Moreover,*

$$|\mathcal{U}|,\ \max_{u\in\mathcal{U},\, j\in[N]} u[j] \leq \frac{2sN}{\varepsilon}.$$

Proof. Let p be a prime satisfying $\frac{sN}{\varepsilon} \leq p < \frac{2sN}{\varepsilon}$, which exists by Bertrand's postulate. For each $a \in \{0, 1, \ldots, p-1\}$, define a vector $u_a \in \{0, 1, \ldots, p-1\}^N$ by

$$u_a[i] = a^i \bmod p \qquad \text{for all } i \in [N].$$

Let $\mathcal{U} = \{u_a : 0 \leq a < p\} \subseteq \mathbb{Z}_{\geq 0}^N$. Then $|\mathcal{U}| = p < 2sN/\varepsilon$, and for every $u \in \mathcal{U}$ and $j \in [N]$, we have $u[j] < p < \frac{2sN}{\varepsilon}$, as desired.

Let $C = (e_1, e_2, \ldots, e_{2\ell})$ be a cycle in G, where each $e_i \in [N]$. Define

$$f_C(y) = \sum_{i=1}^{\ell} y^{e_{2i-1}} - \sum_{i=1}^{\ell} y^{e_{2i}} \in \mathbb{Z}[y],$$

which is nonzero since the edges of C are distinct.

For any $a \in \{0, 1, \ldots, p-1\}$, we have

$$f_C(a) \equiv \sum_{i=1}^{\ell} u_a[e_{2i-1}] - \sum_{i=1}^{\ell} u_a[e_{2i}] \pmod{p}.$$

In particular, if $f_C(a) \not\equiv 0 \pmod{p}$, then $u_a(C) \neq 0$.

Now let $f_{\mathcal{C}} = \prod_{C \in \mathcal{C}} f_C$. Since each f_C is nonzero, we have $f_{\mathcal{C}} \neq 0$. Its degree is at most sN. Hence $f_{\mathcal{C}}$ has at most sN roots in $\mathbb{F}_p$. Therefore, there are at least $p - sN$ values $a \in \{0, 1, \ldots, p-1\}$ such that $f_{\mathcal{C}}(a) \not\equiv 0 \pmod{p}$. For each such a, we have $f_C(a) \not\equiv 0 \pmod{p}$ for all $C \in \mathcal{C}$, and hence $u_a(C) \neq 0$ for all $C \in \mathcal{C}$. Since $p \geq sN/\varepsilon$, we have $\frac{p-sN}{p} \geq 1 - \varepsilon$. Therefore, at least a $(1-\varepsilon)$ fraction of the vectors in $\mathcal{U}$ satisfy $u(C) \neq 0$ for every $C \in \mathcal{C}$. □

Now we complete the proof of Lemma 5.6.

Proof of Lemma 5.6. Let $t = \lceil \log N \rceil$. Let $\mathcal{U} \subseteq \mathbb{Z}_{\geq 0}^N$ be given by Claim B.9 with parameters $s = N^4$ and $\varepsilon = \delta/(2t)$, and let $\mathcal{W} \subseteq (\mathbb{Z}_{\geq 0}^t)^N$ be the corresponding collection from Lemma B.8. Let $S \subseteq \mathbb{F}$ be any set of size at least $\frac{8N^6t^2}{\delta^2}$. Such a set exists by the assumption on $|\mathbb{F}|$.

Define

$$\mathcal{H} = \left\{ a_{w,v} \in \mathbb{F}^N \;\middle|\; w = (w_1, \ldots, w_N) \in \mathcal{W},\ v \in S^t,\ a_{w,v}[i] = \prod_{j \in [t]} v_j^{w_i[j]} \right\}.$$

We claim that $\mathcal{H}$ is a δ-hitting set for $\mathrm{VBP}^1_{r,N,\mathbb{F}}$.

Let

$$A(x) = \sum_{i \in [N]} x_i A_i$$

with $A_i \in \mathbb{F}^{r \times r}$, $\operatorname{rank}(A_i) \leq 1$ for all $i \in [N]$, and $\det(A) \neq 0$. By Lemma B.8, at least a $\left(1 - \frac{\delta}{2t}\right)^t \geq 1 - \frac{\delta}{2}$ fraction of the elements of $\mathcal{W}$ are isolating for A. Fix such a weight assignment $w = (w_1, \ldots, w_N)$, where $w_i \in \mathbb{Z}_{\geq 0}^t$. Let $A_w(y)$ be the matrix obtained from A by replacing each x_i with $\prod_{j \in [t]} y_j^{w_i[j]}$. By Claim B.2, we have $\det(A_w(y)) \neq 0$.

By the construction of $\mathcal{W}$ and the properties of $\mathcal{U}$ from Claim B.9, each entry of $A_w(y)$ is a polynomial in $y_1, \ldots, y_t$ of total degree at most $t \max_{u \in \mathcal{U},\, j \in [N]} u[j] \leq \frac{4N^5t^2}{\delta}$. Since $r \leq N$, it follows that

$$\deg(\det(A_w(y))) \leq \frac{4N^6t^2}{\delta}.$$

By construction of $\mathcal{H}$, for every $v \in S^t$ we have $\det(A)(a_{w,v}) = \det(A_w(y))(v)$. Therefore, by the Schwartz–Zippel lemma, for at least a $1 - \frac{\deg(\det(A_w(y)))}{|S|} \geq 1 - \frac{\delta}{2}$ fraction of points $v \in S^t$, we have $\det(A_w(y))(v) \neq 0$.

Combining the two bounds, for at least a $\left(1 - \frac{\delta}{2}\right)^2 \geq 1 - \delta$ fraction of points in $\mathcal{H}$, the polynomial $\det(A)$ evaluates to a nonzero value. Hence $\mathcal{H}$ is a δ-hitting set for $\mathrm{VBP}^1_{r,N,\mathbb{F}}$.

Finally, by construction,

$$|\mathcal{H}| = |\mathcal{W}| \cdot |S|^t = |\mathcal{U}|^t \cdot |S|^t = \left(\frac{N}{\delta}\right)^{O(\log N)}.$$

□