On Low Field Size Constructions of Access-Optimal Convertible Codes

Saransh Chopra, Francisco Maturana, and K. V. Rashmi
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA, USA
Email: {saranshc,fmaturan,rvinayak}@andrew.cmu.edu

2024 IEEE International Symposium on Information Theory (ISIT) | 979-8-3503-8284-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISIT57864.2024.10619440

***Abstract*—Most large-scale storage systems employ erasure coding to provide resilience against disk failures. Recent work has shown that tuning this redundancy to changes in disk failure rates leads to substantial storage savings. This process requires *code conversion*, wherein data encoded using an $[n^I, k^I]$ initial code has to be transformed into data encoded using an $[n^F, k^F]$ final code, a resource-intensive operation. *Convertible codes* are a class of codes that enable efficient code conversion while maintaining other desirable properties. In this paper, we focus on the *access cost* of conversion (total number of code symbols accessed in the conversion process) and on an important subclass of conversions known as the merge regime (combining multiple initial codewords into a single final codeword).**

In this setting, explicit constructions are known for systematic access-optimal Maximum Distance Separable (MDS) convertible codes for all parameters in the merge regime. However, the existing construction for a key subset of these parameters, which makes use of Vandermonde parity matrices, requires a large field size making it unsuitable for practical applications. In this paper, we provide (1) sharper bounds on the minimum field size requirement for such codes, and (2) explicit constructions for low field sizes for several parameter ranges. In doing so, we provide a proof of super-regularity of specially designed classes of Vandermonde matrices that could be of independent interest.

I. Introduction

Erasure codes are used widely in modern large scale distributed storage systems as a means to mitigate data loss in the event of disk failures. In this context, erasure coding involves dividing data into groups of k chunks that are each encoded into stripes of n chunks using an $[n, k]$ erasure code. These encoded chunks are then stored across n distinct storage nodes in the system. The code parameters n and k determine the amount of redundancy added to the system and the degree of durability guaranteed.

There are various classes of codes that are commonly used in real-world systems. For example, *systematic* codes are those in which the original message symbols are embedded among the code symbols. This is highly desirable in practice as in the event that there are no observed disk failures, there is no decoding process needed to recover the original data. Systematic codes with *Vandermonde parity matrices* (see §II-A) are even more advantageous as there are known efficient algorithms utilizing Fast Fourier Transform (FFT) for computing the product between vectors and Vandermonde matrices [1], [2], speeding up the encoding process. This attribute is becoming increasingly important given the recent trend to use wider (high k) and longer (high n) erasure codes [3], [4]. Additionally, *Maximum Distance Separable (MDS)* codes are a subset of erasure codes that require the least amount of additional storage in order to meet a specific failure tolerance goal. An $[n, k]$ MDS code can tolerate loss of any $n - k$ out of the n code symbols. In this paper, our interest is on systematic MDS codes with Vandermonde parity matrices.

Recent findings by Kadekodi et al. [5] reveal the dynamic variability in disk failure rates over time. Their research highlights the potential for meaningful savings in storage and associated operational expenses through tuning code parameters to observed failure rates. However, the resource overhead associated with the *default approach* of re-encoding all of the data in order to modify n and k is prohibitively expensive [6].

The *code conversion* problem introduced in [6] formalizes the problem of efficiently transforming data that has been encoded under an $[n^I, k^I]$ initial code $\mathcal{C}^I$ to its new representation under an $[n^F, k^F]$ final code $\mathcal{C}^F$. One of the key measures of the cost of conversion is the *access cost*, which represents the total number of code symbols accessed (read/written) during conversion. *Convertible codes* [6] are a class of codes that enable efficient conversion while maintaining other desirable properties such as being MDS and systematic (more details in §II-B).

Among various types of conversions, the *merge regime*, where $k^F = \lambda k^I$ for any integer $\lambda \geq 2$ (i.e., combining multiple initial codewords into a single final codeword), is the most important one. First, the merge regime requires the least resource utilization [7] among all types of conversions and hence are a highly favorable choice for practical systems. Second, constructions for the merge regime are key building blocks for the constructions for codes in the *general regime* which allows for any set of initial parameters and any set of final parameters [7]. In this paper, our focus is on systematic MDS convertible codes in the merge regime.

In [6], the authors established lower bounds on the access cost of conversion and provided constructions of *access-optimal* convertible codes for all parameters in the merge regime. Let us denote $r^I := n^I - k^I$ and $r^F := n^F - k^F$, (which correspond to the number of parity symbols in the initial and final codes

This work was funded in part by an NSF CAREER award (CAREER-1943409), a Sloan Fellowship and a VMware Systems Research Award.

if the codes are systematic). For cases where $r^I > r^F$ (i.e., when the initial configuration has more parities than the final configuration), the authors provide explicit constructions of systematic MDS access-optimal convertible codes over fields of size linear in n^F. For cases where $r^I < r^F$ (i.e., when more parities are needed in the final configuration than in the initial), it has been shown [6] that the access cost of conversion for MDS erasure codes is lower bounded by that of the default approach to decode and re-encode all of the data. As a consequence, it is not possible to realize any savings with specialized code constructions.

However, in the case where $r^I = r^F$, the best-known construction requires a minimum field size of p^D for any prime p and some $D \in \Theta((n^F)^3)$ [6]. This field size is far too high for efficient practical implementations. Most current instruction-set architectures are optimized to operate on bytes of data at a time. Utilizing erasure codes defined over larger field sizes can hamper the encoding/decoding speed. Hence most (if not all) practical implementations of storage codes use $\mathbb{F}_{256}$ (which translates each field symbol to a one-byte representation). Thus, the problem of constructing low-field-size access-optimal convertible codes remains open for the case $r^I = r^F$.

In this paper, we study the setting of systematic MDS access-optimal convertible codes in the merge regime in the case where $r^I = r^F$. The best known construction of convertible codes in this setting is a systematic code with a very specific choice of super-regular Vandermonde parity matrix with a singular degree of freedom [6] (as will be detailed in §II-A). In §III, we improve on this construction by allowing more freedom in the choice of *scalars* of the Vandermonde matrix. We then study the minimum field size $q^*(k, r)$ required for existence of the underlying $k \times r$ super-regular Vandermonde parity matrices of such codes. We provide two lower bounds on the minimum field size required, one applicable for codes over general prime power fields (Theorem 1) and one for codes over fields of characteristic 2 where $k > r$ (Theorem 2). For fields of characteristic 2, the bound takes the form $q^*(k, r) \geq \Omega(2^r)$. Additionally, we establish an upper bound $q^*(k, r) \leq O(k^r)$ (Theorem 3), which in turn results in an improved upper bound $q \leq O((k^F)^{r^F})$ on the field size required for the existence systematic MDS access-optimal convertible codes in the merge regime in the case where $r^I = r^F$.

Furthermore, in §IV, we provide the first explicit low-field-size constructions of convertible codes in this setting for several parameter ranges via constructing their corresponding super-regular Vandermonde parity matrices. The proposed construction makes use of field automorphisms in designing the Vandermonde matrices. For any general prime power field $\mathbb{F}_q$ where $q = p^w$, we find explicit constructions of $k \times 3$ super-regular Vandermonde matrices for all k such that $k < w$ (Theorem 4). This, in turn, gives us a construction of systematic MDS access-optimal convertible codes for all parameters in the merge regime such that $r^F = r^I \leq 3$ and $k^F < w$. For any finite field $\mathbb{F}_q$ where $q = 2^w$ (that is, characteristic 2), we present a stronger result covering a larger range of k by showing that the same proposed construction is super-regular for all k such that $k < q$ (Theorem 5).

These results are also of independent interest beyond the setting considered in this paper as systematic MDS codes with Vandermonde parity matrices serve as the base codes for *bandwidth-optimal* convertible codes [8], [9] and have also been studied in various other settings [2], [10], [11].

II. Background and Related Work

Let us begin with an overview of important concepts and notation referred to throughout this paper, along with a literature review of previous related work.

A. Systematic MDS codes and Vandermonde matrices

An $[n, k]$ linear erasure code $\mathcal{C}$ with generator matrix $\mathbf{G} \in \mathcal{M}(\mathbb{F})_{k\times n}$ over a finite field $\mathbb{F}$ is said to be systematic, or in standard form, if $\mathbf{G} = [\mathbf{I_k} \mid \mathbf{P}]$ where $\mathbf{I_k}$ is the $k \times k$ identity matrix and $\mathbf{P}$ is a $k \times (n-k)$ matrix also known as the parity matrix. Let $\mathbf{m}$ be a message and $\mathbf{c}$ be its corresponding codeword under $\mathcal{C}$, where $\mathbf{m} = (m_i)_{i=1}^k$ and $\mathbf{c} = (c_i)_{i=1}^n$ are vectors of message and code symbols, respectively. As $\mathbf{m}$ is encoded under $\mathcal{C}$ via the multiplication $\mathbf{c} = \mathbf{m}^T\mathbf{G}$, it follows that $c_i = m_i$ for all $i \leq k$ if $\mathcal{C}$ is systematic.

An $[n, k]$ linear erasure code $\mathcal{C}$ is Maximum Distance Separable (MDS) if and only if every k columns of its generator matrix $\mathbf{G}$ are linearly independent; in other words, every $k \times k$ submatrix of $\mathbf{G}$ is non-singular [12]. As a result, data encoded by an $[n, k]$ MDS code can withstand any erasure pattern of $n-k$ out symbols in any codeword and still successfully recover the original data. If $\mathcal{C}$ is also systematic with parity matrix $\mathbf{P}$, this is equivalent to the property that every square submatrix of $\mathbf{P}$ is non-singular [12]. Such a matrix is also referred to as *super-regular*. It is useful to note that any submatrix of a super-regular matrix is also super-regular.

A systematic code with a Vandermonde parity matrix $\mathbf{P} \in \mathcal{M}(\mathbb{F}_{k\times r})$ is one where $\mathbf{P}$ is of the form

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ \xi_1 & \xi_2 & \dots & \xi_r \\ \xi_1^2 & \xi_2^2 & \dots & \xi_r^2 \\ \vdots & \vdots & \ddots & \vdots \\ \xi_1^{k-1} & \xi_2^{k-1} & \dots & \xi_r^{k-1} \end{bmatrix} \tag{1}$$

for some *scalars* $\xi = (\xi_i)_{i=1}^r \in \mathbb{F}^r$. Let us denote the above $k \times r$ Vandermonde matrix as $V_k(\xi)$. Such a matrix is not always guaranteed to be super-regular [12] and thus careful selection of the scalars is required to ensure the resulting systematic code is MDS.

B. Convertible Codes [6]

Recall that a *code conversion* is a procedure that converts data from its initial representation under an $[n^I, k^I]$ code $\mathcal{C}^I$ to its final representation under an $[n^F, k^F]$ code $\mathcal{C}^F$. In order to capture the potential change in dimension, let $M := \mathrm{lcm}(k^I, k^F)$ and consider any message $\mathbf{m} \in \mathbb{F}_q^M$. This is equivalent to $\lambda^I := \frac{M}{k^I}$ codewords in the initial configuration

and $\lambda^F := \frac{M}{k^F}$ codewords in the final configuration. Let $[i] := \{1, 2, \ldots, i\}$ and let $|S|$ denote the size of a set S. Let $\mathbf{m}[S]$ be the vector formed by projecting $\mathbf{m}$ onto the coordinates in the set S, and let $\mathcal{C}(\mathbf{m})$ stand for the encoding of $\mathbf{m}$ under the code $\mathcal{C}$. Let $r^I := n^I - k^I$ and $r^F := n^F - k^F$.

Definition 1 (Convertible Code [6]): An $(n^I, k^I; n^F, k^F)$ convertible code over $\mathbb{F}_q$ is defined by: (1) a pair of codes $(\mathcal{C}^I, \mathcal{C}^F)$ over $\mathbb{F}_q$ such that $\mathcal{C}^I$ is an $[n^I, k^I]$ code and $\mathcal{C}^F$ is an $[n^F, k^F]$ code; (2) a pair of partitions $\mathcal{P}^I := \{P_i^I \mid i \in [\lambda^I]\}$ and $\mathcal{P}^F := \{P_j^F \mid j \in [\lambda^F]\}$ of $[M = \mathrm{lcm}(k^I, k^F)]$ such that $|P_i^I| = k^I$ for all $P_i^I \in \mathcal{P}^I$ and $|P_j^F| = k^F$ for all $P_j^F \in \mathcal{P}^F$; and (3) a conversion procedure which, for any $m \in \mathbb{F}_q^M$, maps the initial set of codewords $\{\mathcal{C}^I(\mathbf{m}[P_i^I]) \mid P_i^I \in \mathcal{P}^I\}$ to the corresponding set of codewords $\{\mathcal{C}^F(\mathbf{m}[P_j^F]) \mid P_j^F \in \mathcal{P}^F\}$ over the final code.

Recall that access cost during code conversion refers to the number of code symbols that are read or written during conversion. Access-optimal convertible codes are those which meet the lower bounds on access cost established in [6] that are known to be tight. It is known that *any* $(n^I, k^I; n^F, k^F)$ convertible code for the merge regime where $r^I = r^F$ formed by a pair of systematic codes with Vandermonde parity matrices $\mathbf{P^I} = V_{k^I}(\xi)$ and $\mathbf{P^F} = V_{k^F}(\xi)$ over the same scalars is access-optimal [6]. This is due to $\mathbf{P^F}$ being "*r^F-column block-constructible*" from $\mathbf{P^I}$; that is, each new parity of a merged codeword can directly be computed as a linear combination of the parities of the original codewords. If the parity matrices are super-regular, then the resulting convertible code is guaranteed to be MDS as well. The best known construction [6] of a systematic MDS access-optimal convertible code for the merge regime where $r^I = r^F$ is formed by a pair of systematic codes with Vandermonde parity matrices over the scalars $\xi = (\theta^{i-1})_{i=1}^{r^I}$, for any primitive element $\theta \in \mathbb{F}$. This construction requires a field size $q \geq p^D$ where p is any prime and $D \in \Theta((n^F)^3)$ [6].

C. Additional Notation and Preliminaries

This section presents notation and terminology used in this paper that follows and expands on the notation introduced in [6], and reviews some preliminaries from Galois theory that will be used in the rest of the paper.

For any two sets I, J, let $I \triangle J$ denote the symmetric difference of I and J. For any two integers a, b, let $a \perp b$ denote that a and b are coprime. Let $\mathbf{x}$ denote the vector $(x_i)_{i=1}^r$ for some r. Let $\mathbf{M}_{i,j}$ denote the entry in the ith row and jth column of the matrix $\mathbf{M}$, with both indices 1-indexed. Let $\mathbf{M}_{I \times J}$ denote the submatrix of $\mathbf{M}$ formed by the intersection of the rows indexed by I and the columns indexed by J, with all indices 1-indexed. Let $\mathrm{row}_i(\mathbf{M})$ stand for the ith row vector of the matrix $\mathbf{M}$. Let χ_P be the indicator function for whether the proposition P is true.

Let $\mathbb{F}_p$ denote the prime field of size p, and let us reserve $\mathbb{F}_q$ for prime power fields of size $q = p^w$ for some prime p and $w > 1$. Let $\mathbb{F}^\times$ denote the multiplicative group of the field, or $\mathbb{F} \setminus \{0\}$. Let $\mathrm{ord}(a)$ denote the order of an element $a \in \mathbb{F}^\times$. Let $\mathbb{F}[x_1, \ldots, x_r]$ denote the ring of polynomials in $x_1, \ldots, x_r$ over the field $\mathbb{F}$. Let $\mathrm{Aut}(\mathbb{F})$ denote the group of automorphisms over the field $\mathbb{F}$. Let S_n denote the group of permutations of $[n]$.

Recall that a field automorphism is a bijective map $\sigma : \mathbb{F} \to \mathbb{F}$ such that for all $x, y \in \mathbb{F}$, $\sigma(x + y) = \sigma(x) + \sigma(y)$ and $\sigma(xy) = \sigma(x)\sigma(y)$; in essence, the map preserves the structure of the field. Note also by definition, it must be the case that $\sigma(0) = 0$ and $\sigma(1) = 1$, which also gives us that $\sigma(-a) = -\sigma(a)$, $\sigma(a^{-1}) = \sigma(a)^{-1}$, and $\mathrm{ord}(a) = \mathrm{ord}(\sigma(a))$ for all $a \in \mathbb{F}^\times$. It is easy to verify that the set of fixed points of an automorphism form a sub-field of $\mathbb{F}$, termed the *fixed field* of the automorphism. It is also a consequence of Galois theory that the fixed field of an automorphism over the field $\mathbb{F}_q$ where $q = p^w$ is always an extension of the base prime field $\mathbb{F}_p$ [13].

D. Related Work

The most directly related works on access-optimal convertible codes [6], [7] were already discussed in §I. In this section, we will discuss other closely related works. In addition to the access cost, previous works on convertible codes have also studied other costs of conversion such as bandwidth cost [8] and locality of repair [14], [15]. In this paper, while we focus on the access cost of conversion, the proposed new constructions do enable better constructions of bandwidth-optimal convertible codes as well. This is because access-optimal convertible codes serve as the base codes of the Piggybacking framework [8] when constructing convertible codes efficient in bandwidth cost.

There also have been previous efforts to study the fundamental limits of existence of super-regular Vandermonde matrices. Shparlinski [10] provided an upper bound on the total number of singular square submatrices of a Vandermonde matrix by showing that any $(q-1) \times m$ Vandermonde matrix $V_{q-1}(\xi_1, \ldots, \xi_m)$ over the field $\mathbb{F}_q$ has at most $3(m-1)(q-1)^m T^{\frac{-1}{m-1}}$ singular $m \times m$ square submatrices where $T := \min_{i \neq j \in [m]} \mathrm{ord}(\frac{\xi_i}{\xi_j})$; however, this bound has been shown to be not tight upon closer investigation [2]. Additionally, Intel's Intelligent Storage Acceleration Library (ISA-L), commonly used to implement erasure coding in practice, has published bounds on the range of parameters $[n, k]$ over $\mathbb{F}_{256}$ for which its code supports generation of super-regular Vandermonde parity matrices, based on a very specific construction [16]. There is no proof provided alongside these bounds; they were likely determined by running a code script to test each submatrix for invertibility.

In addition, there has been independent work studying systematic linear MDS codes with various other constructions of super-regular parity matrices. For example, it is known that a Cauchy matrix $\mathbf{C}$, that is, one of the form $C_{i,j} = (a_i + b_j)^{-1}$ for all $i, j \in [n]$ given two vectors $(a_i)_{i=1}^n$ and $(b_j)_{j=1}^n$, is super-regular so long as the a_i's and b_j's are all distinct from each other [11], [17], [18]. Additionally, Lacan and Fimes introduced a construction of super-regular matrices formed by taking the product of two Vandermonde matrices [2]. To add on, there has been considerable progress in constructing super-regular

Toeplitz matrices in the development of convolutional codes [19]–[21]. Nonetheless, none of these alternatives are suitable for the construction of access-optimal convertible codes.

To our knowledge, in this paper we establish the best known bounds on the field size required for the existence of systematic MDS access-optimal convertible codes for the merge regime where $r^F = r^I$. This paper is also the first to provide, with proof, explicit constructions of systematic MDS access-optimal convertible codes for the merge regime where $r^F = r^I$ over practically usable field sizes.

III. Fundamental Limits on Field Size

In this section, we study a new construction of systematic MDS access-optimal convertible codes for the merge regime where $r^I = r^F$ that generalizes the construction introduced in [6]. The new construction is still based on systematic codes with super-regular Vandermonde parity matrices, but we allow the scalars to take on any distinct nonzero values, rather than being restricted to consecutive powers of a primitive element in the field. By virtue of the parity matrices being Vandermonde matrices, as detailed in §II-B, the new construction of convertible codes remains access-optimal. Thus, a proof of the existence of any $k \times r$ super-regular Vandermonde matrix yields $(n^I, k^I; n^F, k^F = \lambda k^I)$ systematic MDS access-optimal convertible codes for any $\lambda \geq 2$, $k^F \leq k$, and $r^I = r^F \leq r$. We will establish several bounds (Theorems 1, 2 and 3) on the field sizes for which there exist $[n, k]$ systematic MDS codes with Vandermonde parity matrices. This is done by studying their underlying super-regular Vandermonde matrices.

We start with a result which provides a requirement on the field sizes over which such matrices exist. This result draws upon intuition that an optimal choice of scalars for the Vandermonde matrix would avoid selecting elements with smaller order to avoid repetition along the corresponding columns.

Theorem 1: Over the field $\mathbb{F}_q$, a $k \times r$ super-regular Vandermonde matrix can only exist if the following condition holds: for every divisor m of $q-1$ where $m < k$, $q \geq rm + 1$.

Proof: Omitted for brevity. Refer to full version [22]. ■

For the field $\mathbb{F}_{256}$, for example, this result tells us that $[n = 90, k = 86]$ and $[n = 58, k = 52]$ systematic MDS codes with Vandermonde parity matrices do not exist.

The next lemma is a simple consequence of viewing finite prime power fields as vector spaces over their base prime fields.

Lemma 1: Over the field $\mathbb{F}_q$, where $q = 2^w$, for any $r > w$, for any $S = \{\xi_i\}_{i=1}^r \subseteq \mathbb{F}_q$, there must exist some nonempty subset $I \subseteq [r]$ such that $\sum_{i \in I} \xi_i = 0$.

Proof: Omitted for brevity. Refer to full version [22]. ■

This lemma stems from the fact that any collection of field elements larger than the field's dimension must be linearly dependent. Over fields of characteristic 2, this simply corresponds to a nonempty subset of elements that add to 0. This will be used later to identify a singular submatrix in a proposed Vandermonde matrix. This in turn, yields a lower bound on the minimum field size required for the existence of super-regular Vandermonde matrices specific to fields of characteristic 2.

Theorem 2: Over the field $\mathbb{F}_q$, where $q = 2^w$, for any r, k such that $k > r$, a $k \times r$ super-regular Vandermonde matrix with distinct, nonzero scalars can only exist if $q \geq 2^r$.

Proof: Omitted for brevity. Refer to full version [22]. ■

For example, again considering $\mathbb{F}_{256}$, this bound informs us that $[n = 19, k = 10]$ systematic MDS codes with Vandermonde parity matrices do not exist.

The first result for general fields (Theorem 1) is a tighter bound for regimes where $k \gg r$ and $\exists m \approx k$ such that $m < k$ and m divides $q-1$ for a proposed field size q; in this case, we get the bound $q^*(k, r) \geq \Omega(kr)$. On the other hand, the lower bound specific to fields of characteristic 2 (Theorem 2) is more relevant in settings such as storage in unreliable environments which demand narrow codes with higher storage overhead, or when when $k \approx r$.

We will next prove the existence of $k \times r$ super-regular Vandermonde matrices over all fields of size greater than a threshold in terms of k and r. We first start with a lemma that narrows down the set of square submatrices of a Vandermonde matrix that need to be tested for singularity to establish super-regularity. More specifically, we show that it is sufficient to only consider submatrices formed by a set of rows that includes the first row.

Lemma 2: Over the field $\mathbb{F}_q$, for any r, k, ℓ such that $\ell \leq \min(r, k)$, for any $k \times r$ Vandermonde matrix $V_k(\xi)$ with $(\xi_i)_{i=1}^r \in (\mathbb{F}_q^\times)^r$, the submatrix $\mathbf{H} := V_k(\xi)_{I \times J}$ defined by $I := \{\alpha_1, \ldots, \alpha_\ell\} \subseteq [k]$ and $J := \{\beta_1, \ldots, \beta_\ell\} \subseteq [r]$, where $\alpha_i < \alpha_j$ for all $i < j$, is non-singular if and only if the submatrix $\mathbf{H}' := V_k(\xi)_{I' \times J}$ defined by $I' := \{1, \alpha_2 - (\alpha_1 - 1), \ldots, \alpha_\ell - (\alpha_1 - 1)\} \subseteq [k]$ and J is non-singular.

Proof: Omitted for brevity. Refer to full version [22]. ■

We now utilize the Schwartz–Zippel lemma [23], [24] in a probabilistic argument for the existence of a super-regular Vandermonde matrix given a sufficiently large field size. This, in effect, establishes an *upper bound* on the *minimum* field size required for the existence of super-regular Vandermonde matrices.

Theorem 3: Over the field $\mathbb{F}_q$, for any r, k, if $q > 1 + \binom{k}{2} \sum_{\ell=2}^r \binom{r}{\ell} \binom{k-2}{\ell-2} \in O(k^r)$, then there must exist scalars $(\xi_i)_{i=1}^r \in (\mathbb{F}_q^\times)^r$ such that the $k \times r$ Vandermonde matrix $V_k(\xi)$ is super-regular.

Proof: Omitted for brevity. Refer to full version [22]. ■

Recall that the previously known upper bound [6] on the minimum field size q required for the existence of systematic MDS access-optimal convertible codes for the merge regime where $r^I = r^F$ was $\log q \leq \Theta((n^F)^3)$. Theorem 3 establishes the improved upper bound of $\log q \leq O(r^F \log k^F)$, an order of magnitude smaller.

IV. Low Field Size Constructions

In this section, we present several explicit constructions of systematic MDS access-optimal convertible codes in the merge regime (that is, for $(n^I, k^I; n^F, k^F = \lambda k^I)$ convertible codes where $\lambda \geq 2$), with field sizes smaller than existing

constructions. Specifically, for *general prime power fields* $\mathbb{F}_q$ where $q = p^w$, we provide explicit constructions of convertible codes in the merge regime for all parameters such that $r^F = r^I \leq 3$ and $w > k^F$. For fields $\mathbb{F}_q$ of characteristic 2, we present explicit constructions of convertible codes in the merge regime for all parameters such that $r^F = r^I \leq 3$ and $q > k^F$. We do this by providing constructions of $k \times 3$ super-regular Vandermonde matrices for field sizes: $q > p^k$ for general prime power fields (Theorem 4) and $q > k$ for finite fields of characteristic 2 (Theorem 5). These matrices serve as the parity matrices for the systematic MDS codes that underlie the aforementioned convertible codes. As every submatrix of a super-regular matrix is also super-regular, a valid parity matrix for three parities gives us one for any fewer than three parities as well.

We start with a lemma that builds on the intuition to choose primitive elements of the finite field for the scalars of the super-regular Vandermonde parity matrix.

Lemma 3: Over the field $\mathbb{F}_q$, for all $k < q$, given any primitive element $\theta \in \mathbb{F}_q$, given $2 \leq e \leq q-1$ such that $e, e-1 \perp q-1$, the $k \times 3$ Vandermonde matrix $V_k(1, \theta, \theta^e)$ has no singular 2×2 square submatrices.

Proof: Omitted for brevity. Refer to full version [22]. ∎

Next, we introduce the idea of field automorphisms into our construction and choice of scalars, in particular as automorphisms are order preserving maps. Recall some key properties of field automorphisms from §II-C.

Lemma 4: Over the field $\mathbb{F}_q$ where $q = p^w$, for all $k < q$, given any primitive element $\theta \in \mathbb{F}_q$ and nontrivial automorphism $\sigma \in \mathrm{Aut}(\mathbb{F}_q)$ with fixed field $\mathbb{F}_p$, the $k \times 3$ Vandermonde matrix $V_k(1, \theta, \sigma(\theta))$ has no 2×2 singular square submatrices.

Proof: First, recall that $\mathrm{Aut}(\mathbb{F}_q)$ is a group generated by the Frobenius automorphism, or the map $\sigma : x \to x^p$, and thus any nontrivial element $\sigma \in \mathrm{Aut}(\mathbb{F}_q)$ is of the form $\sigma(x) = x^{p^e}$ for some $1 \leq e < w$. It follows that $p \leq p^e < p^w = q$, and because $q \equiv 0 \mod p$, $q - 1 \not\equiv 0 \mod p$ and clearly $p^e \perp q-1$. Next, see that if σ has fixed field $\mathbb{F}_p$, this can only occur if the polynomial $p_1(x) = x^{p^e} - x$, and consequently the polynomial $p_2(x) = x^{p^e-1} - 1$, have no roots in $\mathbb{F}_q$ outside of $\mathbb{F}_p$. This implies that $p^e - 1 \perp q - 1$, and thus we can apply Lemma 3 to get that this matrix has no 2×2 singular submatrices. ∎

For the same construction of Vandermonde matrices as in Lemma 4, we next consider its 3×3 square submatrices and establish the necessary and sufficient conditions under which they are singular. We are able to show a significantly tighter end result for fields of characteristic 2 in particular, but a lot of the arguments used apply to all finite fields as well. Thus, we start with an intermediate result using the shared ideas.

Lemma 5: Over the field $\mathbb{F}_q$ where $q = p^w$, for all $k < q$, given any primitive element $\theta \in \mathbb{F}_q$ and nontrivial automorphism $\sigma \in \mathrm{Aut}(\mathbb{F}_q)$ with fixed field $\mathbb{F}_p$, the $k \times 3$ Vandermonde matrix $V_k(1, \theta, \sigma(\theta))$ has a 3×3 singular square submatrix if and only if $\exists e_1, e_2 \in [k-1]$ and $c_1, c_2 \in \mathbb{F}_p^\times$ such that $e_1 < e_2$ and $\{1, \theta, \sigma(\theta)\}$ are all roots of the polynomial $f(x) = c_1 + c_2 x^{e_1} + x^{e_2}$.

Proof: Omitted for brevity. Refer to full version [22]. ∎

We now arrive at the first of our major results in this section, on explicit constructions of super-regular Vandermonde matrices over arbitrary prime power fields.

Theorem 4: Over the field $\mathbb{F}_q$ where $q = p^w$, for all $k \leq w$, given any primitive element $\theta \in \mathbb{F}_q$ and a non-trivial automorphism $\sigma \in \mathrm{Aut}(\mathbb{F}_q)$ with fixed field $\mathbb{F}_p$, the $k \times 3$ Vandermonde matrix $V_k(1, \theta, \sigma(\theta))$ is super-regular.

Proof: First, note that every 1×1 submatrix of $V_k(1, \theta, \sigma(\theta))$ is non-singular as every element is a power of a nonzero element of $\mathbb{F}_q$. Next, by Lemma 4, every 2×2 submatrix of $V_k(1, \theta, \sigma(\theta))$ is also non-singular. Finally, assume for sake of contradiction that $V_k(1, \theta, \sigma(\theta))$ has a singular 3×3 square submatrix. Then by Lemma 5, $\exists e_1, e_2 \in [k-1]$ and $c_1, c_2 \in \mathbb{F}_p^\times$ such that $e_1 < e_2$ and $\{1, \theta, \sigma(\theta)\}$ are all roots of the polynomial $f(x) = c_1 + c_2 x^{e_1} + x^{e_2}$. However, as $f \in \mathbb{F}_p[x]$, it must be a multiple of the minimum polynomial of θ in $\mathbb{F}_p[x]$, which we know is of degree $w \geq k > e_2 = \deg(f)$ as θ is a generator of $\mathbb{F}_q^\times$, resulting in a contradiction. Thus, every 3×3 square submatrix is also non-singular and $V_k(1, \theta, \sigma(\theta))$ is super-regular, as desired. ∎

Finally, we show an analogous but stronger result for fields of characteristic 2. This is of particular interest as finite fields of characteristic 2 are the most efficient choice for the representation of data in compute nodes and on storage devices.

Theorem 5: Over the field $\mathbb{F}_q$ where $q = 2^w$, for all $k < q$, given any primitive element $\theta \in \mathbb{F}_q$ and a non-trivial automorphism $\sigma \in \mathrm{Aut}(\mathbb{F}_q)$ with fixed field $\mathbb{F}_2$, the $k \times 3$ Vandermonde matrix $V_k(1, \theta, \sigma(\theta))$ is super-regular.

Proof: First, note that every 1×1 submatrix of $V_k(1, \theta, \sigma(\theta))$ is non-singular as every element is a power of a nonzero element of $\mathbb{F}_q$. Next, by Lemma 4, every 2×2 submatrix of $V_k(1, \theta, \sigma(\theta))$ is also non-singular. Finally, assume for sake of contradiction that $V_k(1, \theta, \sigma(\theta))$ has a singular 3×3 square submatrix. Then by Lemma 5, $\exists e_1, e_2 \in [k-1]$ and $c_1, c_2 \in \mathbb{F}_2^\times$ such that $\{1, \theta, \sigma(\theta)\}$ are all roots of the polynomial $f(x) = c_1 + c_2 x^{e_1} + x^{e_2}$. However, this implies $c_1 = c_2 = 1$, but then $f(1) = 1 + 1 + 1 = 1$, contradicting the fact that 1 is a root of f. Therefore, every 3×3 square submatrix is also non-singular and $V_k(1, \theta, \sigma(\theta))$ is super-regular, as desired. ∎

Using this result and the Frobenius automorphism, which is known to have fixed field $\mathbb{F}_p$ over any finite extension $K/\mathbb{F}_p$ [13], we show a family of constructions of super-regular Vandermonde matrices for fields of characteristic 2. We also give results specific to the field $\mathbb{F}_{256}$, which is the most commonly used finite field in practice.

Corollary 1: Over the field $\mathbb{F}_q$ where $q = 2^w$, for all $k < q$, given any primitive element $\theta \in \mathbb{F}_q$, the $k \times 3$ Vandermonde matrix $V_k(1, \theta, \theta^2)$ is super-regular. ∎

Corollary 2: Over the field $\mathbb{F}_{256}$, for all $k < 256$, given any primitive element $\theta \in \mathbb{F}_{256}$, the $k \times 3$ Vandermonde matrices $V_k(1, \theta, \theta^2)$, $V_k(1, \theta, \theta^8)$, $V_k(1, \theta, \theta^{32})$, and $V_k(1, \theta, \theta^{128})$ are super-regular. ∎

REFERENCES

[1] I. Gohberg and V. Olshevsky, "Fast Algorithms with Preprocessing for Matrix-Vector Multiplication Problems," *Journal of Complexity*, vol. 10, no. 4, pp. 411–427, 1994.

[2] J. Lacan and J. Fimes, "Systematic MDS erasure codes based on Vandermonde matrices," *IEEE Communications Letters*, vol. 8, no. 9, pp. 570–572, 2004.

[3] Y. Hu, L. Cheng, Q. Yao, P. P. C. Lee, W. Wang, and W. Chen, "Exploiting Combined Locality for Wide-Stripe Erasure Coding in Distributed Storage," in *19th USENIX Conference on File and Storage Technologies (FAST 21)*. USENIX Association, Feb. 2021, pp. 233–248.

[4] S. Kadekodi, S. Silas, D. Clausen, and A. Merchant, "Practical Design Considerations for Wide Locally Recoverable Codes (LRCs)," *Association for Computing Machinery Transactions on Storage*, vol. 19, no. 4, nov 2023.

[5] S. Kadekodi, K. V. Rashmi, and G. R. Ganger, "Cluster storage systems gotta have HeART: improving storage efficiency by exploiting disk-reliability heterogeneity," in *17th USENIX Conference on File and Storage Technologies, FAST 2019, Boston, MA, February 25-28, 2019*, A. Merchant and H. Weatherspoon, Eds. USENIX Association, 2019, pp. 345–358.

[6] F. Maturana and K. V. Rashmi, "Convertible codes: enabling efficient conversion of coded data in distributed storage," *IEEE Transactions on Information Theory*, vol. 68, pp. 4392–4407, 2022.

[7] F. Maturana, V. S. C. Mukka, and K. V. Rashmi, "Access-optimal linear MDS convertible codes for all parameters," in *IEEE International Symposium on Information Theory, ISIT 2020, Los Angeles, California, USA, June 21-26, 2020*, 2020.

[8] F. Maturana and K. V. Rashmi, "Bandwidth Cost of Code Conversions in the Split Regime," in *2022 IEEE International Symposium on Information Theory (ISIT)*, 2022, pp. 3262–3267.

[9] ——, "Bandwidth Cost of Code Conversions in Distributed Storage: Fundamental Limits and Optimal Constructions," *IEEE Transactions on Information Theory*, vol. 69, no. 8, pp. 4993–5008, 2023.

[10] I. E. Shparlinski, "On the singularity of generalised Vandermonde matrices over finite fields," *Finite Fields and Their Applications*, vol. 11, no. 2, pp. 193–199, 2005.

[11] R. Roth and G. Seroussi, "On generator matrices of MDS codes (Corresp.)," *IEEE Transactions on Information Theory*, vol. 31, no. 6, pp. 826–830, 1985.

[12] F. J. MacWilliams and N. J. A. Sloane, *The theory of error correcting codes*. Elsevier, 1977, vol. 16.

[13] D. Dummit and R. Foote, *Abstract Algebra*. Wiley, 2003.

[14] F. Maturana and K. V. Rashmi, "Locally Repairable Convertible Codes: Erasure Codes for Efficient Repair and Conversion," in *2023 IEEE International Symposium on Information Theory (ISIT)*, 2023, pp. 2033–2038.

[15] X. Kong, "Locally repairable convertible codes with optimal access costs," *ArXiv*, vol. abs/2308.06802, 2023.

[16] Intel, "Corrupted fragment on decode · issue #10 · Intel/ISA-L." [Online]. Available: https://github.com/intel/isa-l/issues/10

[17] R. Roth and A. Lempel, "On MDS codes via Cauchy matrices," *IEEE Transactions on Information Theory*, vol. 35, no. 6, pp. 1314–1319, 1989.

[18] J.-J. Climent, D. Napp, C. Perea, and R. Pinto, "A construction of MDS 2D convolutional codes of rate $1/n$ based on superregular matrices," *Linear Algebra and its Applications*, vol. 437, no. 3, pp. 766–780, 2012.

[19] P. J. F. Almeida and D. N. Avelli, "Superregular matrices over small finite fields," *ArXiv*, vol. abs/2008.00215, 2020.

[20] R. Hutchinson, R. Smarandache, and J. Trumpf, "On superregular matrices and MDP convolutional codes," *Linear Algebra and its Applications*, vol. 428, no. 11, pp. 2585–2596, 2008.

[21] P. Almeida, D. Napp, and R. Pinto, "A new class of superregular matrices and MDP convolutional codes," *Linear Algebra and its Applications*, vol. 439, no. 7, pp. 2145–2157, 2013.

[22] S. Chopra, F. Maturana, and K. V. Rashmi, "On Low Field Size Constructions of Access-Optimal Convertible Codes." [Online]. Available: https://saranshchopra.github.io/research/ISIT_2024

[23] J. T. Schwartz, "Fast Probabilistic Algorithms for Verification of Polynomial Identities," *Journal of the Association for Computing Machinery*, vol. 27, no. 4, p. 701–717, oct 1980.

[24] R. Zippel, "Probabilistic algorithms for sparse polynomials," in *Symbolic and Algebraic Computation*, E. W. Ng, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 1979, pp. 216–226.