Low Field Size Constructions of Access-Optimal Convertible Codes

Saransh Chopra

CMU-CS-24-144

August 2024

Computer Science Department
School of Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Rashmi Vinayak, Chair
Ryan O'Donnell

Submitted in partial fulfillment of the requirements for the degree of Master of Science.

This work was funded in part by an NSF CAREER award (CAREER-1943409), a Sloan Fellowship and a VMware Systems Research Award. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of any sponsoring institution, the U.S. government or any other entity.

Keywords: Coding theory, distributed storage systems, redundancy tuning, code conversion, super-regularity

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 thesis, 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 thesis, 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.

Acknowledgments

I would like to start by thanking my advisor Professor Rashmi Vinayak. I will always be incredibly grateful for her openness to taking on a student who didn't have much previous research experience in an academic setting. I have already learned and grown so much working with her just over the past year. I would also like to thank my other committee member, Professor Ryan O'Donnell, for taking the time to attend my presentation and review this thesis. His observations and contributions were critical in ensuring the correctness of this work; I am fortunate to have had the opportunity to share my work with him. Next, I would like to thank my mentor Dr. Francisco Maturana. I am so lucky to have started my research journey under his guidance. His passion helped me fall in love with this problem and coding theory as a whole. I would next like to thank my former managers at Amazon Web Services, Dr. Sarvesh Bhardwaj and Dr. Samit Chaudhuri, for both instilling leadership principles that I rely on every day as well as supporting and encouraging my efforts to apply for graduate school to chase my true passions. Finally, I would be remiss if I did not thank my parents, family, and friends who all supported me and believed in me throughout this journey. It truly takes a village.

Contents

1 Introduction **1**

2 Background and Related Work **5**
2.1 Systematic MDS codes and Vandermonde matrices 5
2.2 Code Conversion 6
2.3 Convertible Codes 7
2.4 Additional Notation and Preliminaries 10
2.5 Other Related Works 10

3 Fundamental Limits on Field Size **13**

4 Low Field Size Constructions **19**

5 Conclusion **23**

Bibliography **25**

List of Figures

2.1 Conversion from an $[n^I, k^I]$ initial code to an $[n^F, k^F]$ final code. Each box denotes a codeword symbol; empty boxes denote *unchanged symbols*, dotted boxes denote *retired symbols*, and cross-hatched boxes denote *new symbols*. The c node denotes the location where new symbols are computed from the symbols read during conversion. 6

2.2 A *merge conversion* of two codewords encoded under an initial $[6, 4]$ code into a final codeword encoded under a $[10, 8]$ code. Dashed arrows indicate unchanged symbols, and solid arrows indicate reads and writes. 8

List of Tables

2.1 Field size requirements of various access-optimal convertible code constructions . 9

Chapter 1

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 §2.1) are even more advantageous as there are known efficient algorithms utilizing Fast Fourier Transform (FFT) for computing the product between vectors and Vandermonde matrices [5, 12], 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 [6, 10]. 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 thesis, the focus is on systematic MDS codes with Vandermonde parity matrices.

Recent findings by Kadekodi et al. [9] 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 [15].

The *code conversion* problem introduced in [15] 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* [15] are a class of codes that enable efficient conversion while maintaining

other desirable properties such as being MDS and systematic (more details in §2.3).

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 [18] 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 [18]. This thesis focuses on systematic MDS convertible codes in the merge regime.

In [15], the authors established lower bounds on the access cost of conversion between pairs of linear MDS codes and provided constructions of *access-optimal* convertible codes for all parameters in the merge regime, which meet the established lower bounds. 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 if the codes are systematic). For several 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 [15] 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 $D \in \Theta((n^F)^3)$ [15]. 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 $r^I = r^F$.

This thesis studies 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 [15] (as will be detailed in §2.1). In Chapter 3, this construction is improved upon by allowing more freedom in the choice of *scalars* of the Vandermonde matrix.

We start with an existence condition for the underlying $k \times r$ super-regular Vandermonde parity matrices over a candidate field size q (Theorem 3.1). We then establish a lower bound on the minimum field size $q^*(k, r)$ required to guarantee the existence of such matrices, when $k > r$ (Theorem 3.3). The bound takes the form $q^*(k, r) \geq 2^r$ for fields of characteristic 2. Additionally, we establish an upper bound $q^*(k, r) \leq O(k^r)$ (Theorem 3.6), which in turn results in an improved upper bound $q \leq O((k^F)^{r^F})$ on the field size required for the existence of systematic MDS access-optimal convertible codes in the merge regime in the case where $r^I = r^F$.

Furthermore, in Chapter 4, 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 4.4). 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 4.6).

These results are also of independent interest beyond the setting considered in this thesis as systematic MDS codes with Vandermonde parity matrices serve as the base codes for *bandwidth-optimal* convertible codes [14, 17] and have also been studied in various other settings [12, 20, 22].

Chapter 2

Background and Related Work

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

2.1 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 [13]. 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 [13]. 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

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ \xi_1 & \xi_2 & \dots & \xi_r \\ \xi_1^2 & \xi_2^2 & \dots & \xi_r^2 \\ \vdots & \vdots & \ddots & \vdots \\ \xi_1^{k-1} & \xi_2^{k-1} & \dots & \xi_r^{k-1} \end{bmatrix} \tag{2.1}$$

for some *scalars* $\xi = (\xi_i)_{i=1}^{r} \in \mathbb{F}^r$. Let $V_k(\xi)$ denote the $k \times r$ Vandermonde matrix over scalars ξ as depicted in Eq. (2.1). Such a matrix is not always guaranteed to be super-regular [13] and

thus careful selection of the scalars is required to ensure the resulting systematic code is MDS.

2.2 Code Conversion

Code conversion [15] refers to the theoretical problem of efficiently converting 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 $\mathcal{C}^F$, both over the same finite field $\mathbb{F}_q$. Generally, conversion maps multiple codewords in the initial configuration to (potentially) multiple codewords in the final configuration. As depicted in Fig. 2.1, r^I typically denotes the value $n^I - k^I$, or the number of parities per codeword if the initial code $\mathcal{C}^I$ is systematic, and likewise r^F denotes the value $n^F - k^F$ for the final code $\mathcal{C}^F$.

Additionally, in order to capture the potential change in dimension, the conversion is defined over every set of $M := \mathrm{lcm}(k^I, k^F)$ message symbols, the *smallest instance* of the problem. 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. As a consequence, code conversion also formally requires an initial partition $\mathcal{P}^I$ and a final partition $\mathcal{P}^F$ of the set $\{1, 2, \ldots, M\}$ mapping the M message symbols to their initial and final codewords, respectively.

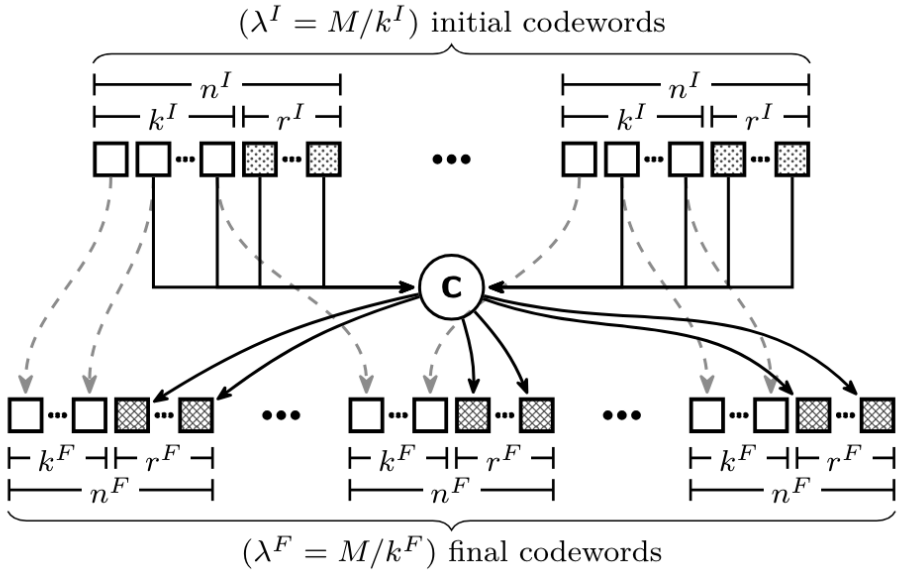

Figure 2.1: Conversion from an $[n^I, k^I]$ initial code to an $[n^F, k^F]$ final code. Each box denotes a codeword symbol; empty boxes denote *unchanged symbols*, dotted boxes denote *retired symbols*, and cross-hatched boxes denote *new symbols*. The c node denotes the location where new symbols are computed from the symbols read during conversion.

2.3 Convertible Codes

Convertible codes [15] are a class of pairs of codes that enable efficient conversion while maintaining other desired decodability constraints, such as being MDS and systematic. Let $[i] := \{1, 2, \ldots, i\}$ and let $|S|$ denote the size of a set S. Let $\mathbf{m}$ denote the vector $(m_i)_{i=1}^{\ell}$ for some ℓ. Let $\mathbf{m}[S]$ denote the vector formed by projecting $\mathbf{m}$ onto the coordinates in the set S, and let $\mathcal{C}(\mathbf{m})$ denote the encoding of $\mathbf{m}$ under the code $\mathcal{C}$. Formally, convertible codes are defined as follows:

Definition 2.1 (Convertible Code [15])**:** *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 $P_j^F \in \mathcal{P}^F$ indicating which message symbols correspond to each codeword; and (3) a conversion procedure which, for any $\mathbf{m} \in \mathbb{F}_q^M$, maps the set of codewords $\{\mathcal{C}^I(\mathbf{m}[P_i^I]) \mid P_i^I \in \mathcal{P}^I\}$ over the initial code 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.*

The cost of a conversion is defined as a function over its parameters $(n^I, k^I; n^F, k^F)$. The general motivation for designing convertible codes is to minimize the conversion cost for any given set of parameters. Thus, for any construction of a convertible code, typically only the optimal conversion procedure between its initial and final codes is considered. There are various cost metrics affecting cluster storage systems, such as compute, bandwidth, and disk IO. This thesis will focus on *access cost* of conversion, defined as follows:

Definition 2.2 (Access Cost [15])**:** *The read access cost of a conversion procedure is defined as the total number of symbols read during the procedure. Similarly, the write access cost of a conversion procedure is the total number of new symbols written during the procedure. The access cost of a conversion procedure is the sum of its read and write access costs. The access cost of a convertible code is the access cost of its conversion procedure.*

Additionally, in this work, we focus on an important subclass of conversions known as the *merge regime*, the set of all conversions in which multiple initial codewords are combined into a single final codeword. Specifically, this is the case where $M = k^F = \lambda k^I$ for some integer $\lambda \geq 2$. Other regimes of interest include the *split regime*, or when $M = k^I = \lambda k^F$ for some integer $\lambda \geq 2$, and the *general regime*, when there are no such restrictions on initial and final code dimensions. However, it has been shown [18] that there is much greater potential for savings in access cost over the default approach in the merge regime than in other types of conversions; moreover, improved constructions of convertible codes in the merge regime directly result in improved constructions in the general regime [18].

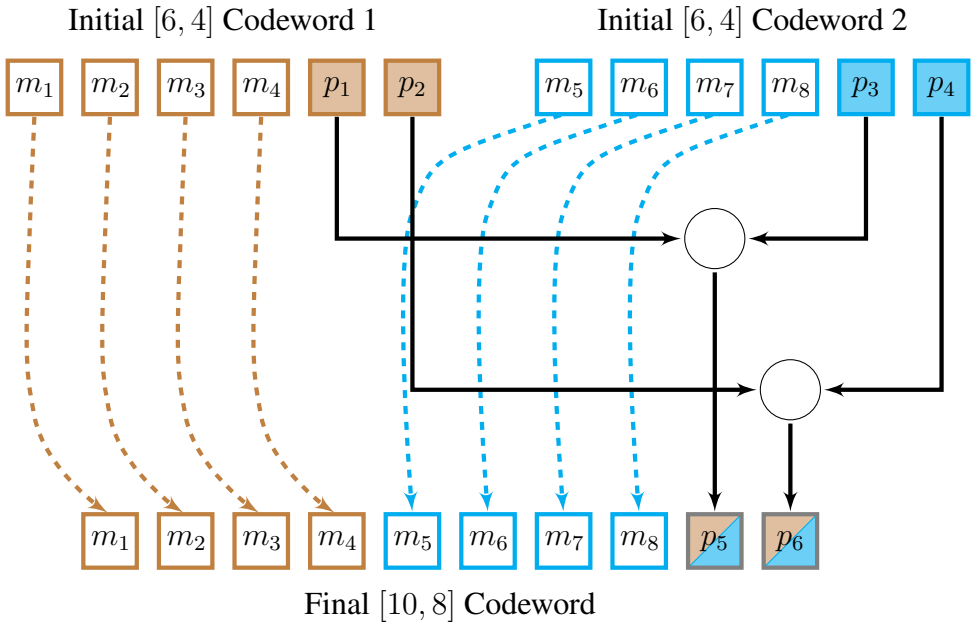

Figure 2.2: A *merge conversion* of two codewords encoded under an initial $[6, 4]$ code into a final codeword encoded under a $[10, 8]$ code. Dashed arrows indicate unchanged symbols, and solid arrows indicate reads and writes.

A convertible code is said to be *access-optimal* over a class of codes if it meets established lower bounds on the access cost of conversion between any pair of codes in that class. It was previously established for conversions between linear MDS codes, within the merge regime, if $r^F \leq r^I$ and $r^F < k^I$, then access cost is lower bounded by $\lambda r^F + r^F$; otherwise access cost is guaranteed to be at least $\lambda k^I + r^F$ [15]. This second bound is attainable by the default approach for any pair of linear MDS codes. Therefore, constructing a linear MDS access-optimal code is only nontrivial when we assume $r^F \leq r^I$ and $r^F < k^I$.

In the same work, the authors showed that within the merge regime, when $r^F \leq r^I$, a systematic convertible code is necessarily access-optimal if the $\lambda k^I \times r^F$ final parity matrix $\mathbf{P^F}$ is r^F*-column block-constructible* from the $k^I \times r^I$ initial parity matrix $\mathbf{P^I}$; that is, the columns of each $k^I \times r^F$ block of $\mathbf{P^F}$ are spanned by at most r^F columns of $\mathbf{P^I}$. Let $\mathbf{M}_{I \times J}$ denote the submatrix of a matrix $\mathbf{M}$ formed by the intersection of the rows indexed by I and the columns indexed by J, with all indices 1-indexed. Block-constructibility is formally defined as follows:

Definition 2.3 (t-column block-constructible [15])**:** *A $\lambda n \times m_1$ matrix $\mathbf{A}$ is t-column block-constructible from an $n \times m_2$ matrix $\mathbf{B}$ if and only if for every $i \in [\lambda]$, there exists $S_i \subseteq [m_2]$ of size t such that the columns of $\mathbf{A}_{R_i \times [m_1]}$ where $R_i = \{(i-1)n + j \mid j \in [n]\}$ are contained in the span of the columns of $\mathbf{B}_{[n] \times S_i}$.*

It follows that given *any* scalars $\xi = (\xi_i)_{i=1}^{r^I} \in \mathbb{F}^{r^I}$, any subset $S := \{\alpha_i \mid i \in [r^F]\} \subseteq [r^I]$, the $\lambda k^I \times r^F$ Vandermonde matrix $\mathbf{P^F} := V_{\lambda k^I}(\xi[S])$ is r^F-column block-constructible from the

$k^I \times r^I$ Vandermonde matrix $\mathbf{P^I} := V_{k^I}(\xi)$. To see this, observe that for every $i \in [\lambda]$, $\mathbf{P^F}_{R_i \times [r^F]}$ where $R_i = \{(i-1)k^I + j \mid j \in [k^I]\}$ is of the form

$$\begin{bmatrix} \xi_{\alpha_1}^{(i-1)k^I} & \xi_{\alpha_2}^{(i-1)k^I} & \cdots & \xi_{\alpha_{r^F}}^{(i-1)k^I} \\ \xi_{\alpha_1}^{(i-1)k^I+1} & \xi_{\alpha_2}^{(i-1)k^I+1} & \cdots & \xi_{\alpha_{r^F}}^{(i-1)k^I+1} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{\alpha_1}^{(i-1)k^I+k^I-1} & \xi_{\alpha_2}^{(i-1)k^I+k^I-1} & \cdots & \xi_{\alpha_{r^F}}^{(i-1)k^I+k^I-1} \end{bmatrix} \tag{2.2}$$

It is clear that the columns of $\mathbf{P^F}_{R_i \times [r^F]}$ are spanned by the columns of $\mathbf{P^I}_{[k^I] \times S}$, which is of the form

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \xi_{\alpha_1} & \xi_{\alpha_2} & \cdots & \xi_{\alpha_{r^F}} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{\alpha_1}^{k^I-1} & \xi_{\alpha_2}^{k^I-1} & \cdots & \xi_{\alpha_{r^F}}^{k^I-1} \end{bmatrix} \tag{2.3}$$

Given this condition guaranteeing access-optimality, the authors provided the following explicit construction of parity matrices that yield systematic MDS access-optimal $(n^I, k^I; n^F, k^F = \lambda k^I)$ convertible codes, given sufficiently large field size:

Theorem 2.4 ([15])**:** *Let $\mathbb{F}_q$ be a finite field of size $q = p^D$, where p is any prime and D is a specific exponent required by the construction that is $\mathcal{O}((\max^3(n^I, n^F)))$. Let $\theta \in \mathbb{F}_q$ be any primitive element. For any $(n^I, k^I; n^F, k^F = \lambda k^I)$ such that $r^F \leq r^I$, the pair of initial and final systematic codes formed by parity matrices $\mathbf{P^I} := V_{k^I}((\theta^i)_{i=1}^{r^I})$ and $\mathbf{P^F} := V_{\lambda k^I}((\theta^i)_{i=1}^{r^F})$ over $\mathbb{F}_q$ yield a systematic MDS access-optimal $(n^I, k^I; n^F, k^F)$ convertible code.*

An extremely large field size is required to ensure super-regularity of the parity matrices in this construction. To address this limitation, the authors also introduced low field size constructions of systematic MDS access-optimal convertible codes for several parameter ranges in the merge regime. The constructions utilized submatrices of super-regular Hankel arrays [20] for parity matrices of the initial and final codes. A summary of the field sizes required to guarantee existence of the best-known constructions of systematic MDS access-optimal $(n^I, k^I; n^F, k^F = \lambda k^I)$ convertible codes in the merge regime is provided in Table 2.1:

Regime	Optimal Construction	Minimum Field Size
$r^F > r^I$ or $r^F \geq k^I$	All systematic MDS convertible codes access-optimal	
$r^F \leq \lfloor r^I/\lambda \rfloor$	Hankel-I [15]	$q \geq \max\{n^I, n^F\} - 1$
$r^F \leq r^I - \lambda + 1$	Hankel-II [15]	$q \geq k^I r^I$
All remaining cases	Vandermonde [15]	$\log q \in \mathcal{O}((\max^3(n^I, n^F)))$

Table 2.1: Field size requirements of various access-optimal convertible code constructions

It is important to note that for the critical regime where $r^I = r^F$, or when the amount of failure tolerance remains constant, there does not exist any known low field size construction. We will focus on this regime for the remainder of this work.

2.4 Additional Notation and Preliminaries

This section presents notation and terminology used in this thesis that follows and expands on the notation introduced in [15], and reviews some preliminaries from field theory that will be used in the rest of the thesis.

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 $\mathbf{M}_{i,j}$ denote the entry in the ith row and jth column of the matrix $\mathbf{M}$, with both 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 field 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 $\mathbb{F}_p$ [4].

2.5 Other Related Works

The most directly related works on the code conversion problem [15] and access-optimal convertible codes [15, 18] were already discussed in §§2.2 and 2.3. 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 [14] and locality of repair [11, 16]. In this thesis, 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 [14] 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 [22] 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 [12]. 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 [8]. There is no proof provided alongside the ISA-L 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 [3, 19, 20]. Additionally, Lacan and Fimes introduced a construction of super-regular matrices formed by taking the product of two Vandermonde matrices [12]. To add on, there has been considerable progress in constructing super-regular Toeplitz matrices in the development of convolutional codes [1, 2, 7]. Nonetheless, none of these alternatives are suitable for the construction of access-optimal convertible codes.

Chapter 3

Fundamental Limits on Field Size

Recall from §2.3 that the best-known construction of systematic MDS access-optimal convertible codes for the merge regime where $r^I = r^F$, introduced in [15], leads to a very high field size requirement. In this chapter, we consider a generalization of this previously best-known construction. The new construction is still based on codes with 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 initial and final parity matrices being Vandermonde matrices over the same set of scalars, as detailed in §2.3, the new construction of convertible codes remains access-optimal. It follows that existence of *any* $k \times r$ super-regular Vandermonde matrix over the field $\mathbb{F}_q$ yields $(n^I, k^I; n^F, k^F = \lambda k^I)$ systematic MDS access-optimal convertible codes over $\mathbb{F}_q$ for any $\lambda \geq 2$, $k^F \leq k$, and $r^I = r^F \leq r$. Thus, in this chapter, we will study the fundamental limits on the field sizes that ensure the existence of $k \times r$ super-regular Vandermonde matrices. We will establish an existence condition (Theorem 3.1), a lower bound for fields of characteristic 2 (Theorem 3.3), and a general upper bound on the minimum field size that guarantees the existence of $k \times r$ super-regular Vandermonde matrices (Theorem 3.6).

We start with a result which provides a requirement on the field sizes over which super-regular Vandermonde 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 3.1: *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. Consider the $k \times r$ Vandermonde matrix $V_k(\xi)$ for any scalars $(\xi_i)_{i=1}^r \in \mathbb{F}_q^r$. Given any divisor m of $q-1$, as $k > m$, the ith entry in the $(m+1)$th row of $V_k(\xi)$ is of the form ξ_i^m. Note that raising this entry to the $(q-1)/m$ power would result in $\xi_i^{q-1} = 1$, so it follows that all of the entries in this row of $V_k(\xi)$ are roots of the polynomial $x^{(q-1)/m} - 1$ over $\mathbb{F}_q$. This polynomial has exactly $(q-1)/m$ distinct roots in $\mathbb{F}_q$, so if $r > (q-1)/m$, then $\exists i, j \in [r]$ such that $i \neq j$

and $\xi_i^m = \xi_j^m$. It follows that $V_k(\xi)_{I\times J}$ where $I = \{1, m+1\}$ and $J = \{i, j\}$ is of the form

$$\begin{bmatrix} 1 & 1 \\ \xi_i^m & \xi_j^m \end{bmatrix} \tag{3.1}$$

and is singular. Therefore, in order for the matrix to be super-regular, we must have $r \leq (q-1)/m \Rightarrow q \geq rm + 1$. □

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 3.2: *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. As there are 2^r distinct subsets of $[r]$, and $q < 2^r$, then $\exists I, J \subseteq [r]$ such that $I \neq J$ and $\sum_{i\in I} \xi_i = \sum_{i\in J} \xi_i$. As every element is its own additive inverse in fields of characteristic 2, it follows that $0 = \sum_{i\in I} \xi_i + \sum_{i\in J} \xi_i = \sum_{i\in I\setminus J} \xi_i + \sum_{i\in J\setminus I} \xi_i + \sum_{i\in I\cap J} \xi_i + \sum_{i\in I\cap J} \xi_i = \sum_{i\in I\triangle J} \xi_i$. As $I \neq J$, $I \triangle J$ must be nonempty, as desired. □

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 3.3: *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. Let $q < 2^r$, and consider the $k \times r$ Vandermonde matrix $V_k(\xi)$ for any distinct scalars $(\xi_i)_{i=1}^r \in (\mathbb{F}_q^\times)^r$ and r, k such that $k > r$. Then, it follows by Lemma 3.2, that $\exists I \subseteq [r]$ nonempty such that $\sum_{i\in I} \xi_i = 0$, and we must have $|I| > 2$ as the ξ_i's are nonzero and distinct. Let us define $\ell := |I|$ and $(c_i)_{i=1}^{\ell+1} \in \mathbb{F}_q^{\ell+1}$ to be the coefficient vector of the polynomial $f(x) := \prod_{i\in I}(x - \xi_i)$ such that $f(x) = \sum_{i=1}^{\ell+1} c_i x^{i-1}$, and note by construction $c_\ell = \sum_{i\in I} \xi_i = 0$. Now consider the square submatrix $\mathbf{H} := V_k(\xi)_{J\times I}$ where $J = [\ell + 1] \setminus \{\ell\}$. If we take the linear combination $\mathbf{y} = c_{\ell+1}\mathrm{row}_\ell(\mathbf{H}) + \sum_{i=1}^{\ell-1} c_i \mathrm{row}_i(\mathbf{H})$, it follows that $\mathbf{y} = (f(\xi_i))_{i\in I} = \mathbf{0}$. As $c_{\ell+1} = 1$, this is a nontrivial linear combination of the rows of $\mathbf{H}$, and thus $\mathbf{H}$ is singular. Therefore, in order for the matrix to be super-regular, we must have $q \geq 2^r$. □

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 3.1) is tighter 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. On the other hand, the lower bound specific to fields of characteristic 2 (Theorem 3.3) 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 3.4: *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. Observe that $\mathbf{H}$ is of the form

$$\begin{bmatrix} \xi_{\beta_1}^{\alpha_1-1} & \xi_{\beta_2}^{\alpha_1-1} & \cdots & \xi_{\beta_\ell}^{\alpha_1-1} \\ \xi_{\beta_1}^{\alpha_2-1} & \xi_{\beta_2}^{\alpha_2-1} & \cdots & \xi_{\beta_\ell}^{\alpha_2-1} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{\beta_1}^{\alpha_\ell-1} & \xi_{\beta_2}^{\alpha_\ell-1} & \cdots & \xi_{\beta_\ell}^{\alpha_\ell-1} \end{bmatrix} \tag{3.2}$$

while $\mathbf{H}'$ is of the form

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \xi_{\beta_1}^{\alpha_2-\alpha_1} & \xi_{\beta_2}^{\alpha_2-\alpha_1} & \cdots & \xi_{\beta_\ell}^{\alpha_2-\alpha_1} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{\beta_1}^{\alpha_\ell-\alpha_1} & \xi_{\beta_2}^{\alpha_\ell-\alpha_1} & \cdots & \xi_{\beta_\ell}^{\alpha_\ell-\alpha_1} \end{bmatrix} \tag{3.3}$$

As the ξ_i's are all non-zero, it can be seen that we can get from $\mathbf{H}'$ to $\mathbf{H}$ by multiplying through the ith column by $\xi_{\beta_i}^{\alpha_1-1}$ for all $i \in [\ell]$. Therefore, $\det(\mathbf{H}) = \det(\mathbf{H}') \prod_{i=1}^{\ell} \xi_{\beta_i}^{\alpha_1-1}$, so either $\det(\mathbf{H}) = \det(\mathbf{H}') = 0$ or both matrices are non-singular, as desired. □

We now utilize the Schwartz–Zippel lemma [21, 23] 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. We have restated the lemma for convenience.

Lemma 3.5 (Schwartz–Zippel)**:** *Let $f \in \mathbb{F}[x_1, \ldots, x_r]$ be a non-zero polynomial over a field $\mathbb{F}$ and let d be the total degree of f. Let S be a finite subset of $\mathbb{F}$. If we independently and uniformly at random select values from S to assign to each of $x_1, \ldots, x_r$, then $\mathbf{Pr}[f(x_1, \ldots, x_r) = 0] \leq \frac{d}{|S|}$.*

Theorem 3.6: *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. Let us start by considering an arbitrary square submatrix of our proposed $k \times r$ Vandermonde matrix $V_k(\xi)$- that is, let $I := \{\alpha_1, \ldots, \alpha_\ell\} \subseteq [k]$ and $J := \{\beta_1, \ldots, \beta_\ell\} \subseteq [r]$ for some $\ell \leq \min(k, r)$ and let us define $\mathbf{H} := V_k(\xi)_{I\times J}$ so that $\mathbf{H}$ is an $\ell \times \ell$ submatrix of $V_k(\xi)$. Observe that

$$\begin{aligned}\det(\mathbf{H}) &= \sum_{\sigma\in S_\ell}\left(sgn(\sigma)\prod_{i=1}^{\ell}\mathbf{H}_{i,\sigma(i)}\right)\\ &= \sum_{\sigma\in S_\ell}\left(sgn(\sigma)\prod_{i=1}^{\ell}\xi_{\beta_{\sigma(i)}}^{\alpha_i-1}\right)\end{aligned}$$

where S_ℓ denotes the group of permutations of $[\ell]$. See that we can treat the scalars as variables $(x_i)_{i=1}^r$ and the overall determinant as a multivariate polynomial

$$f_{\mathbf{H}}(\mathbf{x}) = \sum_{\sigma\in S_\ell}\left(sgn(\sigma)\prod_{i=1}^{\ell}x_{\beta_{\sigma(i)}}^{\alpha_i-1}\right) \in \mathbb{F}_q[x_1, \ldots, x_r]$$

We deduce that the degree of the term in this summation corresponding to any arbitrary $\sigma \in S_\ell$ is $\sum_{i=1}^{\ell}(\alpha_i - 1)$, and thus this is the total degree of $f_{\mathbf{H}}$ as well. Also, note that because every term in this summation corresponds to a unique permutation of $[\ell]$ and the α_i's are distinct, the resulting monomial terms are also all unique, so no terms cancel out and $f_{\mathbf{H}}$ is not identically 0 so long as $q > \deg(f_{\mathbf{H}})$. From here, see that for any family $\mathcal{H}$ of square submatrices of $V_k(\xi)$, if we define $f_{\mathcal{H}} := \prod_{\mathbf{H}\in\mathcal{H}} f_{\mathbf{H}}$, then $f_{\mathcal{H}}$ is also not identically 0 so long as $q > \deg(f_{\mathcal{H}})$. Note also that $f_{\mathcal{H}}$ evaluates to 0 if and only if one of the square submatrices in $\mathcal{H}$ has determinant 0 and is singular. Moreover, as $\deg(f_{\mathcal{H}}) = \sum_{(I,J)|V_k(\xi)_{I\times J}\in\mathcal{H}}\sum_{\alpha_i\in I}(\alpha_i - 1)$, we can then apply Schwartz–Zippel to get that the probability that a uniformly randomly drawn vector from $(\mathbb{F}_q^\times)^r$ is a root of $f_{\mathcal{H}}$ is at most

$$\begin{aligned}\Pr_{\mathbf{x}}\left[f_{\mathcal{H}}(\mathbf{x}) = 0\right] &\leq \frac{\sum_{(I,J)|V_k(\xi)_{I\times J}\in\mathcal{H}}\sum_{\alpha_i\in I}(\alpha_i - 1)}{q-1}\\ &= \frac{\sum_{i\in[k]}(i-1)\sum_{(I,J)|V_k(\xi)_{I\times J}\in\mathcal{H}}\chi_{i\in I}}{q-1}\\ &= \frac{\sum_{i\in[k]}(i-1)\left|\{V_k(\xi)_{I\times J}\in\mathcal{H} \mid i\in I\}\right|}{q-1}\end{aligned}$$

Now see that by Lemma 3.4, it is sufficient to test for super-regularity by only considering

$\mathcal{H} := \{V_k(\xi)_{I\times J} \mid 1 \in I\}$. Therefore, it follows that

$$\begin{aligned}\Pr_{\mathbf{x}}\left[f_{\mathcal{H}}(\mathbf{x}) = 0\right] &\leq \frac{\sum_{i\in[k]}(i-1)\left|\{V_k(\xi)_{I\times J} \mid 1, i \in I\}\right|}{q-1} \\ &= \frac{\sum_{i\in[k]}(i-1)\sum_{\ell=2}^{r}\binom{r}{\ell}\binom{k-2}{\ell-2}}{q-1} \\ &= \frac{\binom{k}{2}\sum_{\ell=2}^{r}\binom{r}{\ell}\binom{k-2}{\ell-2}}{q-1} < 1\end{aligned}$$

if $q > 1 + \binom{k}{2}\sum_{\ell=2}^{r}\binom{r}{\ell}\binom{k-2}{\ell-2}$. If there is a nonzero probability that a uniformly randomly drawn vector $\mathbf{x}$ from $(\mathbb{F}_q^{\times})^r$ is not a root of any of the determinant polynomials, then there must exist some assignment of scalars $(\xi_i)_{i=1}^r$ such that the $k \times r$ Vandermonde matrix $V_k(\xi)$ is super-regular, as desired. □

Recall that the previously known upper bound [15] 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 3.6 establishes the improved upper bound of $\log q \leq O(r^F \log k^F)$, an order of magnitude smaller.

Chapter 4

Low Field Size Constructions

In this chapter, 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 4.4) and $q > k$ for finite fields of characteristic 2 (Theorem 4.6). The super-regular Vandermonde 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 4.1: *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. First, note by Lemma 3.4, we can assume any candidate submatrix contains the first row of the original Vandermonde matrix. Next, see that because $e \perp q-1$, we must have that θ^e is in fact another primitive element of $\mathbb{F}_q$. Thus, we can handle the cases of 2×2 submatrices formed by the first and second columns or the first and third columns identically. In both of these cases, the matrix is of the form

$$\begin{bmatrix} 1 & 1 \\ 1 & \theta^j \end{bmatrix} \tag{4.1}$$

where $q-1 > k-1 \geq j$. It follows that the determinant of this matrix is equal to $\theta^j - 1$ and thus the matrix is singular if and only if $\theta^j = 1 \iff q-1 \mid j$, a contradiction. Similarly, for the case

that the 2×2 submatrix is formed by the second and third columns, the matrix is of the form

$$\begin{bmatrix} 1 & 1 \\ \theta^j & (\theta^e)^j \end{bmatrix} \tag{4.2}$$

See that as $(\theta^j)^{-1} = (\theta^{-1})^j$ exists and is nonzero, we can multiply through the second row by this constant and it would not affect the singularity of the matrix. As a result, it is sufficient to consider the matrix

$$\begin{bmatrix} 1 & 1 \\ 1 & (\theta^{e-1})^j \end{bmatrix} \tag{4.3}$$

and note that as $e-1 \perp q-1$, θ^{e-1} is again a primitive element, and this matrix is thus non-singular by the same proof as in the previous case. □

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 §2.4.

Lemma 4.2: *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 4.1 to get that this matrix has no 2×2 singular submatrices. □

For the same construction of Vandermonde matrices as in Lemma 4.2, 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 4.3: *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. First, let us consider an arbitrary 3×3 square submatrix of the Vandermonde matrix; by Lemma 3.4, we can assume it to be of the form

$$\mathbf{H} := \begin{bmatrix} 1 & 1 & 1 \\ 1 & \theta^{e_1} & (\sigma(\theta))^{e_1} \\ 1 & \theta^{e_2} & (\sigma(\theta))^{e_2} \end{bmatrix} \tag{4.4}$$

where $q-1 > k-1 \geq e_2 > e_1 > 0$. Next, see that $\mathbf{H}$ is singular if and only if there exists nontrivial $(c_i)_{i=1}^3 \in \mathbb{F}_q^3$ such that $\sum_{i=1}^3 c_i \text{row}_i(\mathbf{H}) = \mathbf{0}$; in other words, $\{1, \theta, \sigma(\theta)\}$ are all roots of the polynomial $f(x) = c_1 + c_2 x^{e_1} + c_3 x^{e_2}$. See also that if $c_i = 0$ for any $i \in [3]$, then if we let $J = [3] \setminus \{i\}$, it follows that $\sum_{j \in J} c_j \text{row}_j(\mathbf{H}_{[3]\times[2]}) = \mathbf{0}$. This would equate to a singular 2×2 square submatrix of $V_k(1, \theta, \sigma(\theta))$, a direct contradiction of Lemma 4.2. Thus, we can assume the c_i's are all nonzero. From here, see that $\{1, \theta, \sigma(\theta)\}$ are all roots of the polynomial $f(x) = c_1 + c_2 x^{e_1} + c_3 x^{e_2}$ if and only if they are also roots of the polynomial $g(x) = c_3^{-1} f(x)$, so we can assume without loss of generality that $c_3 = 1$. In summary, $\mathbf{H}$ is singular if and only if $\exists c_1, c_2 \in \mathbb{F}_q^\times$ such that $\{1, \theta, \sigma(\theta)\}$ are all roots of the polynomial $f(x) = c_1 + c_2 x^{e_1} + x^{e_2}$.

Plugging in our three roots into the polynomial, we get the following:

$$c_1 + c_2 + 1 = 0 \tag{4.5}$$
$$c_1 + c_2 \theta^{e_1} + \theta^{e_2} = 0 \tag{4.6}$$
$$c_1 + c_2 (\sigma(\theta))^{e_1} + (\sigma(\theta))^{e_2} = 0 \tag{4.7}$$

Furthermore, we can plug in both sides of Eq. (4.6) into σ, giving us the additional equation

$$0 = \sigma(c_1) + \sigma(c_2)(\sigma(\theta))^{e_1} + (\sigma(\theta))^{e_2} \tag{4.8}$$

We can combine Eq. (4.7) and Eq. (4.8) to get

$$c_1 - \sigma(c_1) = (\sigma(c_2) - c_2)(\sigma(\theta))^{e_1}$$

We can then substitute $c_2 = -c_1 - 1$ from manipulating Eq. (4.5) to get

$$\begin{aligned} c_1 - \sigma(c_1) &= (c_1 + 1 + \sigma(-c_1 - 1))(\sigma(\theta))^{e_1} \\ &= (c_1 - \sigma(c_1))(\sigma(\theta))^{e_1} \end{aligned}$$

Assume for sake of contradiction that $c_1 \notin \mathbb{F}_p$. Then c_1 is not fixed by σ and thus $(c_1 - \sigma(c_1)) \neq 0$. We can then multiply through by $(c_1 - \sigma(c_1))^{-1}$ to get $1 = (\sigma(\theta))^{e_1}$. Note that as $\sigma(\theta)$ is a primitive element of $\mathbb{F}_q$, we must have $q - 1 \mid e_1$, contradicting the assumption that $e_1 < q - 1$. Therefore, we must have $c_1 \in \mathbb{F}_p^\times$, and as $c_2 = -(c_1 + 1)$ and fields are closed under addition and inverses, $c_2 \in \mathbb{F}_p^\times$ as well, as desired. □

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

Theorem 4.4: *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 $\sigma \in \text{Aut}(\mathbb{F}_q)$ 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 4.2, 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 4.3, $\exists e_1, e_2 \in [k-1]$ 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. □

Using this result and the Frobenius automorphism, which is known to have fixed field $\mathbb{F}_p$ over any finite extension $K/\mathbb{F}_p$ [4], we show a family of constructions of super-regular Vandermonde matrices for arbitrary prime power fields.

Corollary 4.5: *Over the field $\mathbb{F}_q$ where $q = p^w$, for all $k \leq w$, given any primitive element $\theta \in \mathbb{F}_q$, the $k \times 3$ Vandermonde matrix $V_k(1, \theta, \theta^p)$ is super-regular.*

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 4.6: *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 4.2, 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 4.3, $\exists e_1, e_2 \in [k-1]$ 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. □

Again using the Frobenius automorphism, 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 4.7: *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 4.8: *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.*

Chapter 5

Conclusion

Code conversion provides a theoretical framework to model the problem of redundancy adaptation, a significant challenge to most large-scale cluster storage systems. Convertible codes are a class of specially designed codes that enable efficient conversion while maintaining desired decodability constraints. The access cost of conversion represents the total number of symbols read or written during the conversion process, which corresponds to the number of disks accessed in the system for the conversion process. Further, the merge regime is an important subclass of conversions which involve merging multiple codewords under an $[n^I, k^I]$ initial code $\mathcal{C}^I$ into a final codeword under an $[n^F, k^F]$ final code $\mathcal{C}^F$.

In this thesis, we studied the setting of systematic MDS access-optimal convertible codes for all parameters $(n^I, k^I; n^F, k^F)$ in the merge regime such that $r^I = r^F$. The previously best-known constructions of codes in this setting required an extremely large field size. In this work, we presented the best-known upper bounds on the field size required to guarantee existence of systematic MDS access-optimal convertible codes in the merge regime. We did so by considering constructions of codes based on super-regular Vandermonde parity matrices. First, we presented an existence condition, a lower bound, and an upper bound on the minimum field size required to guarantee existence of super-regular Vandermonde matrices. In doing so, we improved upon the previously best-known upper bounds on the field size requirement of access-optimal convertible codes by orders of magnitude. Additionally, we provided, to our knowledge, the first explicit constructions of systematic MDS access-optimal convertible codes in the merge regime when $r^I = r^F$ over practically usable field sizes. Specifically, for arbitrary prime power fields $\mathbb{F}_q$, we showed an explicit construction of the super-regular Vandermonde parity matrices of $(n^I, k^I; n^F, k^F = \lambda k^I)$ systematic MDS access-optimal convertible codes for all parameters such that $k^F < \log q$ and $r^I = r^F \leq 3$. We made use of field automorphisms in the choice of scalars of these Vandermonde matrix constructions. With the additional assumption that the fields are of characteristic 2, we extend this same construction to all parameters such that $k^F < q$ and $r^I = r^F \leq 3$.

This work leaves many open questions and potential directions for future work. First, it is of great interest to find a way to extend the explicit construction of access-optimal convertible codes beyond three parities per codeword. A second direction of work could be closing the gap between

fields of characteristic 2 and arbitrary prime power fields. This is purely of theoretical interest, as in practice we almost exclusively use fields of characteristic 2 for the representation of data. Third, it would be beneficial to computationally search for explicit super-regular Vandermonde constructions over various field sizes, as not only would these codes become immediately usable in practice, but they would also illuminate how far theoretical knowledge is from the ground truth. Finally, it would be of interest to find better candidates than Vandermonde matrices for parity matrices of constructions of systematic MDS access-optimal convertible codes. This would help overcome the shortcomings of Vandermonde matrices, especially as in this work, we have identified several parameter ranges for which super-regular Vandermonde matrices do not exist over practical field sizes.

Bibliography

[1] P. Almeida, D. Napp, and R. Pinto. A new class of superregular matrices and MDP convolutional codes. *Linear Algebra and its Applications*, 439(7):2145–2157, 2013. ISSN 0024-3795. doi: https://doi.org/10.1016/j.laa.2013.06.013. 2.5

[2] Paulo José Fernandes Almeida and Diego Napp Avelli. Superregular matrices over small finite fields. *ArXiv*, abs/2008.00215, 2020. 2.5

[3] Joan-Josep Climent, Diego Napp, Carmen Perea, and Raquel Pinto. A construction of MDS 2D convolutional codes of rate $1/n$ based on superregular matrices. *Linear Algebra and its Applications*, 437(3):766–780, 2012. ISSN 0024-3795. doi: https://doi.org/10.1016/j.laa.2012.02.032. 2.5

[4] D.S. Dummit and R.M. Foote. *Abstract Algebra*. Wiley, 2003. ISBN 9780471433347. 2.4, 4

[5] I. Gohberg and V. Olshevsky. Fast Algorithms with Preprocessing for Matrix-Vector Multiplication Problems. *Journal of Complexity*, 10(4):411–427, 1994. ISSN 0885-064X. 1

[6] Yuchong Hu, Liangfeng Cheng, Qiaori Yao, Patrick P. C. Lee, Weichun Wang, and Wei Chen. Exploiting Combined Locality for Wide-Stripe Erasure Coding in Distributed Storage. In *19th USENIX Conference on File and Storage Technologies (FAST 21)*, pages 233–248. USENIX Association, February 2021. ISBN 978-1-939133-20-5. 1

[7] Ryan Hutchinson, Roxana Smarandache, and Jochen Trumpf. On superregular matrices and MDP convolutional codes. *Linear Algebra and its Applications*, 428(11):2585–2596, 2008. ISSN 0024-3795. doi: https://doi.org/10.1016/j.laa.2008.02.011. 2.5

[8] Intel. Corrupted fragment on decode · issue #10 · Intel/ISA-L, 2024. URL `https://github.com/intel/isa-l/issues/10`. 2.5

[9] Saurabh Kadekodi, K. V. Rashmi, and Gregory R. Ganger. Cluster storage systems gotta have HeART: improving storage efficiency by exploiting disk-reliability heterogeneity. In Arif Merchant and Hakim Weatherspoon, editors, *17th USENIX Conference on File and Storage Technologies, FAST 2019, Boston, MA, February 25-28, 2019*, pages 345–358. USENIX Association, 2019. 1

[10] Saurabh Kadekodi, Shashwat Silas, David Clausen, and Arif Merchant. Practical Design Considerations for Wide Locally Recoverable Codes (LRCs). *Association for Computing Machinery Transactions on Storage*, 19(4), nov 2023. ISSN 1553-3077. doi: 10.1145/3626198. 1

[11] Xiangliang Kong. Locally repairable convertible codes with optimal access costs. *ArXiv*, abs/2308.06802, 2023. 2.5

[12] J. Lacan and J. Fimes. Systematic MDS erasure codes based on Vandermonde matrices. *IEEE Communications Letters*, 8(9):570–572, 2004. 1, 2.5

[13] Florence Jessie MacWilliams and Neil James Alexander Sloane. *The theory of error correcting codes*, volume 16. Elsevier, 1977. 2.1, 2.1

[14] Francisco Maturana and K. V. Rashmi. Bandwidth Cost of Code Conversions in the Split Regime. In *2022 IEEE International Symposium on Information Theory (ISIT)*, pages 3262–3267, 2022. doi: 10.1109/ISIT50566.2022.9834604. 1, 2.5

[15] Francisco Maturana and K. V. Rashmi. Convertible codes: enabling efficient conversion of coded data in distributed storage. *IEEE Transactions on Information Theory*, 68:4392–4407, 2022. ISSN 1557-9654. doi: 10.1109/TIT.2022.3155972. 1, 2.2, 2.3, 2.1, 2.2, 2.3, 2.3, 2.4, 2.3, 2.4, 2.5, 3, 3

[16] Francisco 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)*, pages 2033–2038, 2023. doi: 10.1109/ISIT54713.2023.10206604. 2.5

[17] Francisco Maturana and K. V. Rashmi. Bandwidth Cost of Code Conversions in Distributed Storage: Fundamental Limits and Optimal Constructions. *IEEE Transactions on Information Theory*, 69(8):4993–5008, 2023. doi: 10.1109/TIT.2023.3265512. 1

[18] Francisco Maturana, V. S. Chaitanya 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. 1, 2.3, 2.5

[19] R.M. Roth and A. Lempel. On MDS codes via Cauchy matrices. *IEEE Transactions on Information Theory*, 35(6):1314–1319, 1989. doi: 10.1109/18.45291. 2.5

[20] R.M. Roth and G. Seroussi. On generator matrices of MDS codes (Corresp.). *IEEE Transactions on Information Theory*, 31(6):826–830, 1985. doi: 10.1109/TIT.1985.1057113. 1, 2.3, 2.5

[21] J. T. Schwartz. Fast Probabilistic Algorithms for Verification of Polynomial Identities. *Journal of the Association for Computing Machinery*, 27(4):701–717, oct 1980. ISSN 0004-5411. doi: 10.1145/322217.322225. 3

[22] Igor E. Shparlinski. On the singularity of generalised Vandermonde matrices over finite fields. *Finite Fields and Their Applications*, 11(2):193–199, 2005. ISSN 1071-5797. doi: https://doi.org/10.1016/j.ffa.2004.11.001. 1, 2.5

[23] Richard Zippel. Probabilistic algorithms for sparse polynomials. In Edward W. Ng, editor, *Symbolic and Algebraic Computation*, pages 216–226, Berlin, Heidelberg, 1979. Springer Berlin Heidelberg. ISBN 978-3-540-35128-3. 3