Rank-Metric Codes and Their Applications
Other titles in Foundations and Trends® in Communications and Information Theory

Asymptotic Frame Theory for Analog Coding
Marina Haikin, Matan Gavish, Dustin G. Mixon and Ram Zamir
ISBN: 978-1-68083-908-1

Modeling and Optimization of Latency in Erasure-coded Storage Systems
Vaneet Aggarwal and Tian Lan
ISBN: 978-1-68083-842-8

An Algebraic and Probabilistic Framework for Network Information Theory
S. Sandeep Pradhan, Arun Padakandla and Farhad Shirani
ISBN: 978-1-68083-766-7

Theoretical Foundations of Adversarial Binary Detection
Mauro Barni and Benedetta Tondi

Polynomial Methods in Statistical Inference
Yihong Wu and Pengkun Yang
ISBN: 978-1-68083-730-8

Information-Theoretic Foundations of Mismatched Decoding
Jonathan Scarlett, Albert Guillen i Fabregas, Anelia Somekh-Baruch and Alfonso Martinez
ISBN: 978-1-68083-712-4
Rank-Metric Codes and Their Applications

Hannes Bartz
German Aerospace Center (DLR)
hannes.bartz@dlr.de

Lukas Holzbaur
Technical University of Munich
lukas.holzbaur@tum.de

Hedongliang Liu
Technical University of Munich
lia.liu@tum.de

Sven Puchinger
Hensoldt Sensors GmbH
mail@svenpuchinger.de

Julian Renner
Technical University of Munich
julian.renner@tum.de

Antonia Wachter-Zeh
Technical University of Munich
antonia.wachter-zeh@tum.de
Editorial Scope

Topics

Foundations and Trends® in Communications and Information Theory publishes survey and tutorial articles in the following topics:

- Coded modulation
- Coding theory and practice
- Communication complexity
- Communication system design
- Cryptology and data security
- Data compression
- Data networks
- Demodulation and Equalization
- Denoising
- Detection and estimation
- Information theory and statistics
- Information theory and computer science
- Joint source/channel coding
- Modulation and signal design
- Multiuser detection
- Multiuser information theory
- Optical communication channels
- Pattern recognition and learning
- Quantization
- Quantum information processing
- Rate-distortion theory
- Shannon theory
- Signal processing for communications
- Source coding
- Storage and recording codes
- Speech and Image Compression
- Wireless Communications

Information for Librarians

Foundations and Trends® in Communications and Information Theory, 2022, Volume 19, 4 issues. ISSN paper version 1567-2190. ISSN online version 1567-2328. Also available as a combined paper and online subscription.
## Contents

**1** Introduction .................................................. 2

**2** Basics on Rank-Metric Codes .................................. 7
   2.1 Notation ...................................................... 8
   2.2 Linearized Polynomials ...................................... 9
   2.3 Rank-Metric Codes .......................................... 13
   2.4 Weight Distribution of MRD Codes .......................... 16
   2.5 Constant-Rank Codes ....................................... 18
   2.6 Covering Property .......................................... 20
   2.7 Gabidulin Codes ............................................. 21
   2.8 Decoding of Gabidulin Codes ............................... 23
   2.9 Considerations on List Decoding Gabidulin Codes .......... 34
   2.10 Interleaved Gabidulin Codes ............................... 36
   2.11 Folded Gabidulin Codes .................................... 37
   2.12 Decoding of Symmetric Errors ............................. 40
   2.13 Further Classes of MRD Codes ............................. 41

**3** Applications to Code-Based Cryptosystems .................. 45
   3.1 The Hardness of Problems in the Rank Metric ............. 47
   3.2 McEliece-like Systems ...................................... 55
   3.3 Systems based on the Hardness of List Decoding .......... 76
   3.4 A System based on Rank Quasi-Cyclic Codes ............... 80
3.5 Parameters of Public-Key Encryption Schemes ........ 84
3.6 Signature Schemes .................................. 84

4 Applications to Storage 86
4.1 Locality in Distributed Data Storage ..................... 86
4.2 Coded Caching Scheme with MRD Codes .................. 95

5 Applications to Network Coding 99
5.1 Introduction ............................................. 99
5.2 Solutions of Generalized Combination Networks .......... 102
5.3 Error Control in (Random) Linear Network Coding ....... 107
5.4 Subspace Codes .......................................... 112
5.5 Upper Bounds on Subspace Codes ......................... 119

6 Conclusion 132

Acknowledgements 136
Rank-Metric Codes and Their Applications

Hannes Bartz¹, Lukas Holzbaur², Hedongliang Liu², Sven Puchinger³, Julian Renner² and Antonia Wachter-Zeh²

¹German Aerospace Center (DLR); hannes.bartz@dlr.de
²Technical University of Munich; lukas.holzbaur@tum.de, lia.liu@tum.de, julian.renner@tum.de, antonia.wachter-zeh@tum.de
³Hensoldt Sensors GmbH; mail@svenpuchinger.de

ABSTRACT
The rank metric measures the distance between two matrices by the rank of their difference. Codes designed for the rank metric have attracted considerable attention in recent years, reinforced by network coding and further motivated by a variety of applications. In code-based cryptography, the hardness of the corresponding generic decoding problem can lead to systems with reduced public-key size. In distributed data storage, codes in the rank metric have been used repeatedly to construct codes with locality, and in coded caching, they have been employed for the placement of coded symbols. This survey gives a general introduction to rank-metric codes, explains their most important applications, and highlights their relevance to these areas of research.

©2022 H. Bartz et al.
Codes composed of matrices are a natural generalization of codes composed of vectors. Codes in the rank metric of length $n \leq m$ can be considered as a set of $m \times n$ matrices over a finite field $\mathbb{F}_q$ or equivalently as a set of vectors of length $n$ over the extension field $\mathbb{F}_{q^m}$. The rank weight of each codeword vector is the rank of its matrix representation and the rank distance between two matrices is the rank of their difference. These definitions rely on the fact that the rank distance is indeed a metric. Several code constructions and basic properties of the rank metric show strong similarities to codes in the Hamming metric. However, there are also notable differences, e.g., in the list decoding properties.

Error-correcting codes in the rank metric were first considered by Delsarte \[64\], who proved a Singleton-like upper bound on the cardinality of rank-metric codes and constructed a class of codes achieving this bound\(^1\). This class of codes was reintroduced by Gabidulin \[78\] in his fundamental paper \textit{“Theory of Codes with Maximum Rank Distance”}. Further, in his paper several properties of codes in the rank metric and

\(^1\)In analogy to MDS codes, such codes are called Maximum Rank Distance (MRD) codes.
an efficient decoding algorithm based on an equivalent of the Euclidean algorithm were shown. Since Gabidulin’s publication contributed significantly to the development of error-correcting codes in the rank metric, the most famous class of codes in the rank metric — the equivalents of Reed–Solomon codes — are nowadays called Gabidulin codes. These codes can be defined by evaluating non-commutative linearized polynomials, proposed by Ore [201], [202]. Independently of the previous work, Roth [241] discovered in 1991 codes in the rank metric and applied them for correcting crisscross error patterns.

The goal of this survey is to provide an overview of the known properties of rank-metric codes and their application to problems in different areas of coding theory and cryptography.

Section 2 provides a brief introduction to rank-metric codes, their properties and their decoding. After providing basic notations for finite fields and linearized polynomials, we consider codes in the rank metric. We first define the rank metric and give basic properties and bounds on the cardinality of codes in the rank metric (namely, equivalents of the Singleton, sphere-packing, and the Gilbert–Varshamov bounds). Then, we define Gabidulin codes, show that they attain the Singleton-like upper bound on the cardinality and give their generator and parity-check matrices. We describe their decoding up to half the minimum rank-distance by syndrome-based decoding. A summary of how to accomplish this efficiently is given and the problem of error-erasure correction is considered. We also give an overview on list decoding of Gabidulin codes and consider interleaved and folded Gabidulin codes. Finally, further classes of rank-metric codes such as twisted Gabidulin codes are briefly discussed.

Rank-metric codes have several applications in communications and security, including public-key code-based cryptography. In 1978, Rivest, Shamir and Adleman (RSA) [239] proposed the first public-key cryptosystem in order to guarantee secure communication in an asymmetric manner. Since then, public-key cryptography is essential to protect data via encryption, to enable secure key exchange for symmetric encryption, and to protect the authenticity and integrity of data via digital signature schemes. Only one year after the RSA cryptosystem was introduced, whose security relies on the hardness of the integer
factorization problem, McEliece [186] proposed the first public key cryptosystem based on error-correcting codes. In his pioneering work McEliece showed that hard problems in coding theory can be used to derive public-key cryptosystems. A crucial drawback of the McEliece cryptosystem compared to other public-key cryptosystems, such as RSA or elliptic curve cryptosystems (ECC), is its large public-key size. The recent developments in quantum computing rendered all of the currently used public-key cryptosystems whose security relies on the integer factorization or the discrete logarithm problem insecure. In particular, Shor’s algorithm [250] allows to solve both the integer factorization problem and the discrete logarithm problem in polynomial time, which in turn allows to break the corresponding public-key cryptosystems completely, given a sufficiently large quantum computer. Since code-based public-key cryptosystems are resilient against all known attacks on quantum computers, including Shor’s algorithm, they are considered to be quantum-resistant (or post-quantum secure) cryptosystems.

Quantum-resistant cryptography is an important research area to ensure the long-term security of transmitted and stored data. Therefore, the National Institute of Standards and Technology (NIST) opened a standardization call, which meanwhile has reached its final round [200]. In order to reduce the public-key size, many new McEliece variants based on several codes were proposed, both before and independent of the NIST competition and also as submissions to the NIST competition. This includes a long history of variants based on codes in the rank metric. The first McEliece variant in the rank metric was proposed by Gabidulin, Paramonov, and Tretjakov [85] and is therefore known as the GPT cryptosystem. Although no rank-metric based schemes are among the finalists, rank-metric based schemes are considered as potential candidates for future standards [6].

Section 3 gives an overview of rank-metric code-based quantum-resistant encryption and authentication schemes. First, hard problems which can be used to design rank-metric code-based cryptosystems are considered. Then, a general framework to define most GPT variants is given, and the particular variants are described. Finally, an overview on non-GPT-like cryptosystems, including the NIST submission Rank Quasi Cyclic (RQC), and rank-metric code-based signature schemes is given.
Rank-metric codes find applications not only in the cryptographic protection of data, but also in ensuring its integrity. The increase in the amount of data that is stored by distributed storage systems has motivated a transition from replication of the data to the use of more involved storage codes, most commonly Maximum Distance Separable (MDS) codes. By storing one symbol of a codeword on each node, a node failure then corresponds to a symbol erasure and the Hamming distance of the storage code provides a guarantee on the number of failures the system can tolerate before data loss occurs. However, as the number of nodes in these systems grows, not only the number of tolerable node failures, but also the efficiency of the node repair process becomes a concern. Codes with locality [53], [111], [133] address this issue by reducing the number of nodes required for repair in the more likely event of a single or small number of node failures. While these codes are designed for the Hamming metric, codes for the rank metric, in particular, Gabidulin codes have repeatedly been used to construct these codes, especially for the stronger notion of maximally recoverable (MR) locally repairable/recoverable codes (LRCs)\(^2\). Further, rank-metric codes have also been used in another area related to distributed storage, referred to as coded caching. Caching is a commonly used strategy to reduce the traffic rate during the peak hours. The communication procedure consists of two phases: placement and delivery. The seminal work by Maddah-Ali and Niesen [178] has shown that applying coding merely in the delivery phase can reduce the traffic rate. As a further improved scheme [277] has been shown to be order-optimal under uncoded placement, schemes with coded placement [54], [108] become of interest in order to further reduce the traffic rate during the delivery phase. Rank-metric codes have been utilized in the scheme with coded placement by Tian and Chen [262], which has been shown to outperform the optimal scheme with uncoded placement [277] in the regime of small cache size.

In Section 4, the application of rank-metric codes to distributed data storage is outlined. First, we explore the connection between codes with locality and rank-metric codes by providing a high-level description of

\(^2\)MR LRCs are also referred to as partial MDS (PMDS) codes.
the property exploited by many constructions of (MR) LRCs. Second, we present the application of Maximum Rank Distance (MRD) codes in the coded caching scheme by Tian and Chen [262].

Network coding has been attracting attention since the fundamental works by Ahlswede et al. [5] and Li, Yeung, and Cai [161] showed that the capacity of multicast networks can be achieved by performing linear combinations of packets instead of just forwarding them. Rank-metric codes have been used in network coding solutions [74] and error correction in coherent networks [256]. For random networks, rank-metric codes are used to correct errors by the lifting construction [258]. In addition, the subspace metric [150] was introduced for error control, as this metric perfectly captures the type of errors that occur in (random) linear network coding. Due to the close relation between the rank metric and the subspace metric, rank-metric codes are a natural choice to construct subspace codes for error control in random network coding.

Section 5 introduces constructions of network codes based on MRD codes, constructions of subspace codes by lifting rank-metric codes, bounds on the cardinality, and the list decoding capability of subspace codes. We first present constructions based on MRD codes for a class of deterministic multicast networks, which guarantee that all the receivers decode all the messages. Two error models commonly considered in networks are described. We introduce subspace codes, with a focus on constructions based on lifting rank-metric codes and provide upper bounds on the size of subspace codes. Further, an analysis of list decoding subspace codes is provided.

Finally, Section 6 concludes this survey and shortly mentions further applications of rank-metric codes.
References


References


P. Loidreau, “A new rank metric codes based encryption scheme,” in *8th Int. Conf. on Post-Quantum Cryptography (PQCrypto)*, 2017.


References


References


