

# Low Density MDS Codes and Factors of Complete Graphs \*

Lihao Xu Vasken Bohossian Jehoshua Bruck  
 California Institute of Technology  
 Mail Code 136-93  
 Pasadena CA 91125  
 Email: {lihao,vincent,bruck}@paradise.caltech.edu

David G. Wagner  
 University of Waterloo  
 Department of Combinatorics and Optimization  
 Waterloo, Ontario, Canada N2L 3G1  
 dgwagner@math.uwaterloo.ca

*Abstract* — We reveal an equivalence relation between the construction of a new class of low density MDS array codes, that we call B-Code, and a combinatorial problem known as *perfect one-factorization* of complete graphs. We use known perfect one-factors of complete graphs to create constructions and decoding algorithms for both B-Code and its *dual code*. B-Code and its dual are optimal in the sense that (i) they are MDS, (ii) they have an *optimal encoding* property, i.e., the number of the parity bits that are affected by change of a single information bit is *minimal* and (iii) they have *optimal length*. The existence of perfect one-factorizations for *every* complete graph with an even number of nodes is a 35 years long conjecture in graph theory. The construction of B-codes of arbitrary odd length will provide an affirmative answer to the conjecture.

## I. INTRODUCTION

A common property of MDS array codes is that the encoding and decoding procedures use only simple *XOR* (*exclusive OR*) and thus are more computationally efficient [1]. One important parameter of array codes is the *update complexity*, the average number of parity bits affected by a change of a single information bit in the codes. This parameter is particularly crucial when the codes are used in storage applications that need frequent updates of information.

In this paper, we present the *B-Code*, a new MDS array code of size  $n \times l$  bits, where  $l = 2n$  or  $2n + 1$ . B-Code is of column distance 3 and its dual code is of distance  $l - 1$ . It achieves the *lower* bound of the update complexity. B-Code also achieves the *maximum* length that MDS codes with optimal update property can have, which is called as *optimal length* here. In addition, parity bits of B-Code are evenly distributed over all columns, and each parity bit requires the same amount of *XOR* operations, and thus the computation complexity for computing parity bits is *balanced*.

Figure 1 shows,  $\hat{B}_6$ , the *dual B-Code* of length 6. It is represented as an array of information ( $a_1$  through  $a_6$ ) and parity bits as well as by a labeled graph in which every vertex corresponds to an information bit and each edge represents a parity bit summing the two information bits that are its vertices. The edges and vertices of the graph are labeled by the index of the column to which the corresponding information and parity bits belong.  $\hat{B}_6$  has distance 5 and can tolerate the erasure of 4 columns. Figure 1 (b) shows a decoding path for the erasure of columns 3 through 6. Use  $a_2$  (from column 2) together with parity  $a_2 + a_3$  (from column 1) to recover  $a_3$ . Use the latter along with parity  $a_3 + a_4$  (from column 2) to

recover  $a_4$ , etc. For any 4-column erasures, such a decoding path exists.

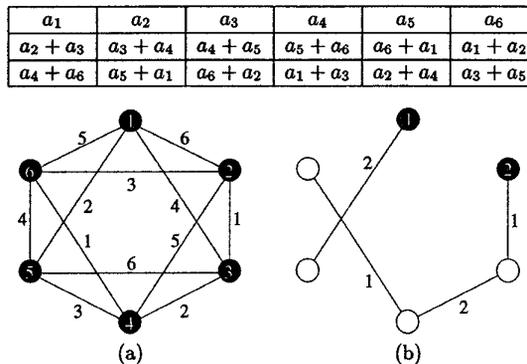


Fig. 1:  $\hat{B}_6$ , the dual *B-Code* of length 6, is a  $3 \times 6$  MDS array code. (a) a graph representation of  $\hat{B}_6$ , (b) a decoding path for the erasure of columns 3, 4, 5 and 6.

The novelty of this paper is to use the above new graph description of array codes and prove the equivalence of B-code constructions to the problem of *perfect one-factorizations* of complete graphs [2]. The problem is to partition all the edges of a complete graph into *disjoint* sets such that 1) each set covers all vertices of the graph, and 2) union of any pair of the sets forms a *Hamilton cycle*. Results on perfect one-factorizations can be used to construct two infinite families of B-Codes. The existence of perfect one-factorizations for *every* complete graph with an even number of nodes is a 35-year-long conjecture in graph theory. A positive answer to it will provide B-code constructions of arbitrary length. On the other hand, the construction of B-codes of arbitrary odd length will provide an affirmative answer to the conjecture.

The full version of this paper can be found at:  
<http://paradise.caltech.edu/ETR.html>

## REFERENCES

- [1] M. Blaum, P. G. Farrell and H. C. A. van Tilborg, "Chapter on Array Codes", Handbook of Coding Theory, edited by V. S. Pless and W. C. Huffman, to appear.
- [2] W. D. Wallis, *One-Factorizations*, Kluwer Academic Publisher, 1997.

\*Supported in part by the NSF Young Investigator Award CCR-9457811, by an IBM Partnership Award, by the Sloan Research Fellowship, and by DARPA through an agreement with NASA/OSAT.