This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Li, Jie; Tang, Xiaohu

Systematic Construction of MDS Codes with Small Sub-packetization Level and Near Optimal Repair Bandwidth

Published in:
2019 IEEE International Symposium on Information Theory, ISIT 2019 - Proceedings

DOI:
[10.1109/ISIT.2019.8849209](https://doi.org/10.1109/ISIT.2019.8849209)

Published: 01/07/2019

Document Version
Peer reviewed version

Please cite the original version:
Li, J., & Tang, X. (2019). Systematic Construction of MDS Codes with Small Sub-packetization Level and Near Optimal Repair Bandwidth. In *2019 IEEE International Symposium on Information Theory, ISIT 2019 - Proceedings* (pp. 1067-1071). [8849209] (IEEE International Symposium on Information Theory - Proceedings; Vol. 2019-July). IEEE. <https://doi.org/10.1109/ISIT.2019.8849209>

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

Systematic Construction of MDS Codes with Small Sub-packetization Level and Near Optimal Repair Bandwidth

Jie Li^{*†} and Xiaohu Tang[‡]

^{*}Department of Mathematics and Systems Analysis, Aalto University, Espoo, Finland, jie.0.li@aalto.fi, jieli873@gmail.com

[†]Hubei Key Laboratory of Applied Mathematics, Faculty of Mathematics and Statistics, Hubei University, Wuhan, China

[‡]Information Security and National Computing Grid Laboratory, Southwest Jiaotong University, Chengdu, 610031, China, xhutang@swjtu.edu.cn

Abstract—In the literature, all the known high-rate MDS codes with the optimal repair bandwidth possess a significantly large sub-packetization level, which may prevent the codes to be implemented in practical systems. To build MDS codes with small sub-packetization level, existing constructions and theoretical bounds imply that one may sacrifice the optimality of the repair bandwidth. Partly motivated by the work of Tamo *et al.* (IEEE Trans. Inform. Theory, 59(3), 1597-1616, 2013), in this paper, we present a powerful transformation that can greatly reduce the sub-packetization level of any MDS codes with respect to the same code length n . As applications of the transformation, four high-rate MDS codes having both small sub-packetization level and near optimal repair bandwidth can be obtained, where two of them are also explicit and the required field sizes are comparable to the code length n .

I. INTRODUCTION

Maximum distance separable (MDS) codes, which can provide the optimal tradeoff between fault tolerance and storage overhead, have been used extensively in distributed storage systems. By distributing the codeword across distinct storage nodes, it can be ensured that in the case of node failure, the missing data can be recovered from the data at some surviving nodes, named helper nodes as well. In storage scenarios, one of the most important parameters is the *repair bandwidth*, which is defined as the amount of data downloaded from the helper nodes to repair the failed node. Particularly, Dimakis *et al.* [1] derived a lower bound on the repair bandwidth of MDS codes, which motivated abundant recent research in coding for distributed storage [2]–[13].

In the literature, most existing MDS codes with the repair bandwidth achieving the lower bound in [1] are known as *array codes* [14]. A codeword of an (n, k) array code is an $N \times n$ matrix, where the parameter N is called the *sub-packetization level* and n is called the *code length*. When deploying an array code to a distributed storage system, a code symbol (i.e., a column) corresponds to a storage node. Then, an array code is said to have the *MDS property* if every k out of the n columns of the matrix can recover the rest $n-k$ columns. It was proved in [1] that the repair bandwidth $\gamma(d)$ of an (n, k) MDS array code with sub-packetization level N should satisfy $\gamma(d) \geq \gamma^*(d) \triangleq \frac{d}{d-k+1}N$, where d ($k \leq d \leq n-1$) is the number of helper nodes. An MDS array code is said to have the optimal repair bandwidth if $\gamma(d) = \gamma^*(d)$. In the particular case, when $d = n-1$, $\gamma^*(d)$ can be reduced to the minimal value $\frac{n-1}{n-k}N$. Therefore, $d = n-1$ is the main concern in most

known works [3]–[12]. In this paper, we also follow the same setting and abbreviate $\gamma^*(n-1)$ to γ^* . Besides, we focus on MDS array codes, which will be abbreviated as MDS codes.

In the literature, all the known high-rate MDS code constructions with the optimal repair bandwidth possess a significantly large sub-packetization level N , usually $N \geq r^{\frac{n}{r+1}}$ [12]. Further in [15], it was shown that for an MDS code with the optimal repair bandwidth, a sub-packetization level N with N being exponential in n is necessary. An MDS code with larger sub-packetization level can lead to a reduced design space in terms of various system parameters and makes management of meta-data difficult. Furthermore, it is not easy to be implemented in practical systems [16].

Existing constructions and theoretical bounds imply that one may construct high-rate MDS codes with small sub-packetization level by sacrificing the optimality of the repair bandwidth. In [16], two high-rate (n, k) MDS codes with small sub-packetization level were presented, the first one can have a sub-packetization level as small as $N = r^\tau$ where τ is any positive integer and $r = n - k$, while the repair bandwidth is $(1 + \frac{1}{\tau})\gamma^*$. Nevertheless, the code is constructed over a significantly large finite field $\mathbf{F}_q$ with $q > n^{(r-1)N+1}$, which may prevent it to be deployed in practical systems. While the second one in [16] is obtained by combining an MDS code with the optimal repair bandwidth and another scalar linear code operating on the GV bound [20]. In [3], an $(n = sk' + 2, k = sk')$ MDS code with sub-packetization level $2^{k'-1}$ and near optimal repair bandwidth only for systematic nodes was proposed, which is termed duplication-zigzag code in this paper.

In this paper, we aim to construct high-rate MDS codes that have both small sub-packetization level and near optimal repair bandwidth for general parameters n and k . We notice that there exist abundant high-rate MDS codes with the optimal repair bandwidth but require a large sub-packetization level in the literature [3]–[10], [19], which intrigue us to think whether can we reduce the sub-packetization level of those codes by slightly sacrificing the optimality of the repair bandwidth. Partly motivated by the work in [3], we present a powerful transformation that can convert any MDS code into another MDS code with much longer code length, such that the repair bandwidth of the new MDS code is slightly larger than the optimal value but the sub-packetization level is the same as that of the original MDS code, or equivalently the generic

transformation can reduce the sub-packetization level N of the original code with respect to the same code length n . By directly applying the generic transformation to several known high-rate MDS codes with the optimal repair bandwidth, we get four high-rate (n, k) MDS codes that have both small sub-packetization level N and near optimal repair bandwidth, with two of them are explicit and the required field sizes are comparable to the code length n , which outperform the first MDS code construction in [16] in terms of the field size and outperform the first code in both [6] and [19], the second MDS code construction in [16] in terms of the sub-packetization level.

The remainder of the paper is organized as follows. Section II presents some necessary preliminaries. Section III gives the generic transformation and the asserted properties. Section IV presents direct applications of the generic transformation, and a concrete coefficient assignment to two of them. Section V gives comparisons of some key parameters among the MDS codes proposed in this paper and some existing notable MDS codes. Finally, Section VI provides some concluding remarks.

II. PRELIMINARIES

A. (n, k) MDS codes

Denote by q a prime power and $\mathbf{F}_q$ the finite field with q elements. Let $\mathbf{f}_0, \mathbf{f}_1, \dots, \mathbf{f}_{n-1}$ be the data stored across a distributed storage system consisting of n nodes based on an (n, k) MDS code, where $\mathbf{f}_i$ is a column vector of length N over $\mathbf{F}_q$. Throughout this paper, we consider the (n, k) MDS codes that defined in the following parity-check form:

$$\underbrace{\begin{pmatrix} A_{0,0} & A_{0,1} & \cdots & A_{0,n-1} \\ A_{1,0} & A_{1,1} & \cdots & A_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{r-1,0} & A_{r-1,1} & \cdots & A_{r-1,n-1} \end{pmatrix}}_A \begin{pmatrix} \mathbf{f}_0 \\ \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_{n-1} \end{pmatrix} = \mathbf{0}, \quad (1)$$

where $r = n - k$, $A_{t,i}$, $t \in [0, r)$, $i \in [0, n)$ are nonsingular matrices of order N over $\mathbf{F}_q$, A is called the *parity-check matrix* of the code, where A can be written as

$$A = (A_{t,i})_{t \in [0, r), i \in [0, n-1]} \quad (2)$$

to indicate the block entries.

For every $t \in [0, r)$, by (1), we have

$$A_{t,0}\mathbf{f}_0 + A_{t,1}\mathbf{f}_1 + \cdots + A_{t,n-1}\mathbf{f}_{n-1} = \mathbf{0}, \quad (3)$$

which contains N equations, for convenience, we say that $\sum_{i=0}^{N-1} A_{t,i}\mathbf{f}_i = \mathbf{0}$ is the t -th *parity-check group*.

B. The MDS property

An (n, k) MDS code defined by (1) possesses the MDS property that the source file can be reconstructed by connecting any k out of the n nodes. That is, any $r \times r$ sub-block matrix of $A = (A_{t,i})_{t \in [0, r), i \in [0, n-1]}$ is nonsingular [6].

Particularly, if

$$A_{t,i} = A_i^{t-1}, \quad t \in [0, r), \quad i \in [0, n) \quad (4)$$

for some matrices A_i , then we have the following result.

Lemma 1 ([6]). *An (n, k) MDS code defined by (1) and (4) is said to have the MDS property if $A_i A_j = A_j A_i$ and $A_i - A_j$ is nonsingular for all $i, j \in [0, n)$ with $i \neq j$.*

C. Repair

When repairing a failed node i of an (n, k) MDS code, the data downloaded from helper node j can be represented by $R_{i,j}\mathbf{f}_j$, where $R_{i,j}$ is a $\beta_{i,j} \times N$ matrix of full rank, and is called *repair matrix* of node i .

Clearly, a failed node can be repaired if there are N linearly independent equations with respect to the N unknowns of $\mathbf{f}_i$, where the N equations should be chosen elaborately so that the interference in these equations can be cancelled by downloading a small amount of data from the helper nodes. In this paper, similarly to that in [19], for convenience, we only consider the symmetric situation that exactly N/r linearly independent equations are acquired from each of the r parity-check groups, where the N/r equations are linear combinations of the corresponding N parity-check equations. Precisely, the N/r linearly independent equations from the t -th parity-check group can be obtained by multiplying (3) with an $N/r \times N$ matrix $S_{i,t}$ of full rank, where $S_{i,t}$ is called the *select matrix* in [19], which is used to generate appropriate N/r equations from the t -th parity-check group. As a consequence, the following linear equations are available.

$$\underbrace{\begin{pmatrix} S_{i,0}A_{0,i} \\ S_{i,1}A_{1,i} \\ \vdots \\ S_{i,r-1}A_{r-1,i} \end{pmatrix}}_{\text{useful data}} \mathbf{f}_i + \sum_{j=0, j \neq i}^{n-1} \underbrace{\begin{pmatrix} S_{i,0}A_{0,j} \\ S_{i,1}A_{1,j} \\ \vdots \\ S_{i,r-1}A_{r-1,j} \end{pmatrix}}_{\text{interference by } \mathbf{f}_j} \mathbf{f}_j = \mathbf{0},$$

thus regenerating node i ($i \in [0, n)$) requires that

- (i) the coefficient matrix of the useful data is of full rank, i.e.,

$$\text{rank}\left(\begin{pmatrix} S_{i,0}A_{0,i} \\ S_{i,1}A_{1,i} \\ \vdots \\ S_{i,r-1}A_{r-1,i} \end{pmatrix}\right) = N, \quad (5)$$

- (ii) the interference caused by $\mathbf{f}_j$ can be determined by the data $R_{i,j}\mathbf{f}_j$ downloaded from node j , i.e.,

$$\text{rank}\left(\begin{pmatrix} R_{i,j} \\ S_{i,0}A_{0,j} \\ S_{i,1}A_{1,j} \\ \vdots \\ S_{i,r-1}A_{r-1,j} \end{pmatrix}\right) = \text{rank}(R_{i,j}) \quad (6)$$

for all $j \in [0, n) \setminus \{i\}$, which means that

$$\text{rank}\left(\begin{pmatrix} R_{i,j} \\ S_{i,t}A_{t,j} \end{pmatrix}\right) = \text{rank}(R_{i,j}) \quad (7)$$

for all $j \in [0, n) \setminus \{i\}$ and $t \in [0, r)$.

The repair bandwidth of node i is then

$$\gamma_i = \sum_{j=0, j \neq i}^{n-1} \text{rank}(R_{i,j}) = \sum_{j=0, j \neq i}^{n-1} \beta_{i,j}. \quad (8)$$

Obviously, if $\gamma_i = (n-1)\frac{N}{r}$, then node i is said to have the optimal repair bandwidth, which can be accomplished if $\beta_{i,j} = N/r$ for all $j \in [0, n) \setminus \{i\}$.

In addition to the (near) optimal repair bandwidth, an (n, k) MDS code is also preferred to have the *repair-by-transfer* property, i.e., repairing of a failed node involves mere transfer of data and without need for any arithmetic operations either at the helper nodes or at the replacement node [21]. Therefore, an MDS code has the repair-by-transfer property if each row of its repair matrix has only one nonzero entry.

III. A GENERIC TRANSFORMATION

In this section, we present a generic transformation that can convert any MDS code with the optimal repair bandwidth defined in the form of (1) to a new MDS code with longer code length and near optimal repair bandwidth. The transformation can be performed through the following two steps.

Step 1: Choosing an (n', k') MDS code with the optimal repair bandwidth as the base code

Assuming the base code is defined over a finite field contain at least q' elements. Let N denote its sub-packetization level, $r = n' - k'$, and let $(A'_{t,i})_{t \in [0, r), i \in [0, n')}$ denote its parity-check matrix while the $N/r \times N$ matrices $R'_{i,j}$ and $S'_{i,t}$, $i, j \in [0, n')$ with $j \neq i$, $t \in [0, r)$, respectively denote the repair matrices and select matrices.

Step 2: The transition from the base code to the new MDS code

Through the generic transformation, we intend to design a new $(n = k + r, k)$ MDS code over certain finite field $\mathbf{F}_q$ having arbitrary code length n while maintaining the same sub-packetization level N . By convenience, we assume that n is a multiple of n' . Note that through puncturing, we can obtain (n, k) MDS codes where n is not a multiple of n' [8], [20]. In the following, let $n = sn'$, where $s \geq 2$.

The transition from the base code to the new MDS code is done by designing the parity-check matrix, the repair matrices, and the select matrices of the new MDS code from those of the base code as follows.

$$A_{t, ln'+p} = x_{t, ln'+p} A'_{t,p}, \quad (9)$$

$$R_{i,j} = \begin{cases} R'_{i \% n', j \% n'}, & \text{if } j \not\equiv i \pmod{n'}, \\ I, & \text{otherwise,} \end{cases} \quad (10)$$

and

$$S_{i,t} = S'_{i \% n', t} \quad (11)$$

where $x_{t, ln'+p} \in \mathbf{F}_q \setminus \{0\}$, $t \in [0, r)$, $l \in [0, s)$, $p \in [0, n')$, $i, j \in [0, n)$ with $j \neq i$, and $\%$ denotes the modulo operation.

Similarly to Theorem 7 in [12] and Lemma 7 in [13], by using the Combinatorial Nullstellensatz in [17], we have the following result to ensure the MDS property of the new code.

Theorem 1. *The new (n, k) code over $\mathbf{F}_q$ obtained by the generic transformation can possess the MDS property if $q > N \binom{n-1}{r-1}$.*

Theorem 2. *Every failed node of the new $(n = sn', k)$ code obtained by the generic transformation can be regenerated by the repair matrices defined in (10), where the repair bandwidth is $(1 + \frac{(s-1)(r-1)}{n-1})\gamma^*$.*

Proof. Since the (n', k') base code possesses the optimal repair bandwidth, then (5) and (7) hold for the base code.

Firstly, we verify (5) for the new code. For any $i \in [0, n)$, write i as $i = jn' + i'$, where $j \in [0, s)$ and $i' \in [0, n')$. Then, by (9) and (11),

$$\text{rank}\left(\begin{pmatrix} S_{i,0}A_{0,i} \\ S_{i,1}A_{1,i} \\ \vdots \\ S_{i,r-1}A_{r-1,i} \end{pmatrix}\right) = \text{rank}\left(\begin{pmatrix} S'_{i',0}A'_{0,i'} \\ S'_{i',1}A'_{1,i'} \\ \vdots \\ S'_{i',r-1}A'_{r-1,i'} \end{pmatrix}\right) = N,$$

where the last equality holds since (5) holds for the base code.

Next, we check (7) for the new code. For $i, j \in [0, n)$ with $j \neq i$, we rewrite i and j as $i = un' + i'$ and $j = vn' + j'$, where $u, v \in [0, s)$ and $i', j' \in [0, n')$. When $i' \neq j'$,

$$\text{rank}\left(\begin{pmatrix} R_{i,j} \\ S_{i,t}A_{t,j} \end{pmatrix}\right) = \text{rank}\left(\begin{pmatrix} R'_{i',j'} \\ S'_{i',t}A'_{t,j'} \end{pmatrix}\right) = \frac{N}{r}, \quad t \in [0, r),$$

where the last equality holds since (7) holds for the base code. When $i' = j'$,

$$\text{rank}\left(\begin{pmatrix} R_{i,j} \\ S_{i,t}A_{t,j} \end{pmatrix}\right) = \text{rank}\left(\begin{pmatrix} I \\ S_{i,t}A_{t,j} \end{pmatrix}\right) = N, \quad t \in [0, r).$$

Therefore, according to (8), the repair bandwidth of node i is

$$\gamma_i = \sum_{j=0, j \neq i}^{n-1} \text{rank}(R_{i,j}) = (1 + \frac{(s-1)(r-1)}{n-1})\gamma^*.$$

□

IV. APPLICATIONS OF THE GENERIC TRANSFORMATION

In this section, by directly applying the generic transformation in Section III respectively to the two (n', k') code construction in [6], which are respectively termed Ye-Barg codes 1 and 2 in this paper, the first (n', k') MDS code construction in [19], which is termed the improved Ye-Barg code 2 (since it is an improvement of the Ye-Barg code 2 in [6] w.r.t. the field size), and the (n', k') optimal sub-packetization code in [8], [9], we get four MDS codes with small sub-packetization level. Note that the sub-packetization level of these four MDS codes are respectively $r^{n'}$, $r^{n'-1}$, $r^{n'-1}$, and $r^{\frac{n'}{r}}$, where $r = n' - k'$.

Theorem 3. *Respectively choosing the (n', k') Ye-Barg codes 1 and 2, the (n', k') improved Ye-Barg code 2 in [19], and the (n', k') optimal sub-packetization code in [8], [9] as the base code for the generic transformation in Section III, we can get four $(n = sn', k)$ MDS codes C_1 - C_4 over $\mathbf{F}_q$ with $q > N \binom{n-1}{r-1}$, where $r = n - k = n' - k'$, and the repair*

bandwidth is $(1 + \frac{(s-1)(r-1)}{n-1})\gamma^*$. Besides, the MDS code $\mathcal{C}_1$ has the optimal update property while $\mathcal{C}_2$, $\mathcal{C}_3$, and $\mathcal{C}_4$ have the repair-by-transfer property.

Note that for the resultant codes obtained by the generic transformation, the required field size is relatively large and the constructions are inexplicit. In the following, we provide a solution to dramatically reduce the field size of the MDS codes $\mathcal{C}_2$ and $\mathcal{C}_3$ by providing a concrete assignment of the coefficients $x_{t,j}$, $t \in [0, r)$ and $j \in [0, n)$ in (9).

Theorem 4. *The field size q of the $(n = sn', k)$ MDS codes $\mathcal{C}_2$ and $\mathcal{C}_3$ can be respectively reduced to $q > \max\{sr, n'\}$ and $q > sr$ by setting $x_{t,j} = x_j^t = c^{\lfloor \frac{j}{n'} \rfloor t}$ in (9) for $t \in [0, r)$ and $j \in [0, n)$, where c is a primitive element of $\mathbf{F}_q$.*

Hereafter, we only prove Theorem 4 for $\mathcal{C}_2$ to save space, which is similar to the proof of Theorem 15 of [6].

Proof of Theorem 4 for code $\mathcal{C}_2$:

For consistency, we borrow the notations in [6] to introduce the parity-check matrix of the Ye-Barg code 2. Let $N = r^{n'-1}$ where $r = n' - k'$. For any $a \in [0, N)$ with $(a_{n'-2}, a_{n'-1}, \dots, a_0)$ being its r -ary expansion, define

$$a(i, u) = (a_{n'-2}, \dots, a_{i+1}, u, a_{i-1}, \dots, a_0)$$

where $i \in [0, n')$ and $u \in [0, r)$.

Let $\{e_a : a = 0, 1, \dots, N-1\}$ denote the standard basis of $\mathbf{F}_{q'}^N$ over $\mathbf{F}_{q'}$ where $q' > n'$. The parity-check matrix $(A'_{t,i})_{t \in [0, r), i \in [0, n')}$ of the (n', k') Ye-Barg code 2 in [6] is defined by

$$A'_{t,i} = (A'_i)^t, \quad (12)$$

$$A'_i = \sum_{a=0}^{N-1} \lambda_{i,a_i} e_a^\top e_{a(i, a_i \oplus 1)}, \quad i \in [0, n'-1), A'_{n'-1} = I, \quad (13)$$

where $\oplus$ denotes addition modulo r , $\lambda_{i,a_i} = c^i$ if $a_i = 0$ and 1 otherwise with c being a primitive element of $\mathbf{F}_{q'}$.

Note that by (9), (12), and (13), we have that the parity-check matrix of code $\mathcal{C}_2$ is defined in the form of (4) where

$$A_{un'+n'-1} = c^u A'_{n'-1} = c^u I, \quad (14)$$

$$A_{un'+i'} = c^u A'_{i'} = \sum_{a=0}^{N-1} c^u \lambda_{i',a_{i'}} e_a^\top e_{a(i', a_{i'}+1)} \quad (15)$$

for $i' \in [0, n'-1)$, $u \in [0, s)$.

It is obvious that $A_i A_j = A_j A_i$ holds for any $i, j \in [0, n)$ with $i \neq j$. Then according to Lemma 1, it suffice to show that $A_i - A_j$ is nonsingular, which is equivalent to say that for any $X = \sum_{a=0}^{N-1} x_a e_a^\top$, $(A_i - A_j)X = \mathbf{0}$ implies $X = \mathbf{0}$. Let us rewrite $i = un' + i'$ and $j = vn' + j'$ for some $u, v \in [0, s)$ and $i', j' \in [0, n')$, where $(u, i') \neq (v, j')$.

Case 1: If $i \equiv j \pmod{n'}$, i.e., $i' = j'$ and $u \neq v$, then by (14) and (15), we have

$$A_i - A_j = (c^u - c^v) A'_{i'},$$

which is nonsingular.

Case 2: If $i \not\equiv j \pmod{n'}$, $i' \neq n' - 1$, and $j' \neq n' - 1$, then we have that

$$\begin{aligned} & (A_i - A_j)X \\ &= (c^u A'_{i'} - c^v A'_{j'})X \\ &= \sum_{a=0}^{N-1} (c^u \lambda_{i',a_{i'}} x_{a(i', a_{i'}+1)} e_a^\top - c^v \lambda_{j',a_{j'}} x_{a(j', a_{j'}+1)} e_a^\top) \\ &= \mathbf{0} \end{aligned}$$

if and only if

$$c^u \lambda_{i',a_{i'}} x_{a(i', a_{i'}+1)} - c^v \lambda_{j',a_{j'}} x_{a(j', a_{j'}+1)} = 0, \quad a \in [0, N),$$

which is equivalent to

$$x_a = \frac{\prod_{t=0}^{r-1} c^v \lambda_{j',t}}{\prod_{t=0}^{r-1} c^u \lambda_{i',t}} x_a = c^{t(v-u)} x_a, \quad a \in [0, N), \quad (16)$$

which implies $X = \mathbf{0}$ since $0 < t(v-u) < sr$.

Case 3: If $i \not\equiv j \pmod{n'}$, similarly to Case 2, we can get that $(A_i - A_j)X = \mathbf{0}$ implies $X = \mathbf{0}$.

This finishes the proof.

V. COMPARISONS

In this section, we do comparisons of some key parameters among the MDS codes proposed in this paper and some existing notable MDS codes, where Table I illustrates the details.

Under the same parameters n and k , it is seen that the new MDS codes $\mathcal{C}_1$ - $\mathcal{C}_4$ have the following advantages and disadvantages compared with existing notable MDS codes.

- The new MDS codes $\mathcal{C}_1$ - $\mathcal{C}_4$ have the same repair bandwidth as that of the first MDS code in [16].
- Compared with the first MDS code in [16], the new MDS codes $\mathcal{C}_1$ - $\mathcal{C}_4$ especially $\mathcal{C}_2$ and $\mathcal{C}_3$ are built on a much smaller finite field while maintain the same repair bandwidth, the new MDS code $\mathcal{C}_4$ also has the same sub-packetization level as that of the first MDS code in [16].
- The MDS code $\mathcal{C}_1$ has the optimal update property while the other new codes have the repair-by-transfer property.
- If setting $n' = r\tau$ in the new MDS codes such as in code $\mathcal{C}_3$, then it has a smaller sub-packetization level, which is around $\frac{1}{\log n}$ times that of the second MDS code in [16] while under the same field size level and the same repair bandwidth level.
- The new MDS codes $\mathcal{C}_1$ - $\mathcal{C}_4$ can support any number of parity nodes while the punctured duplication-zigzag code¹ in [3] can only support two parity nodes.

¹Note that the code length of the duplication-zigzag code in [3] is in the form of $uk' + 2$ with $uk' \gg 2$, in order to do a fair comparison under the same code length, we delete two nodes of the duplication-zigzag code in [3] and term the resultant code as punctured duplication-zigzag code.

TABLE I
A COMPARISON OF SOME KEY PARAMETERS AMONG THE (n, k) MDS CODES PROPOSED IN THIS PAPER AND SOME EXISTING NOTABLE (n, k) MDS CODES, WHERE $n = sn'$, $r = n - k$

	Sub-packetization level N	Field size	The ratio of repair bandwidth to the optimal value	Property	Memo
The new MDS code $\mathcal{C}_1$	$r^{n'}$	$q > N^{\binom{n-1}{r-1}}$	$= 1 + \frac{(s-1)(r-1)}{n-1} < 1 + \frac{r}{n'}$	optimal update	Thm 3
The second MDS code in [16]	$O(r^{r^r} \log n)$	$O(n)$	$\leq 1 + \frac{1}{r}$		
Ye-Barg code 1 in [6]	r^n	$q \geq rn$	1 (optimal)	optimal update	
The new MDS code $\mathcal{C}_2$	$r^{n'-1}$	$q > \max\{sr, n'\}$	$= 1 + \frac{(s-1)(r-1)}{n-1} < 1 + \frac{r}{n'}$	repair-by-transfer	Thms 3, 4
The new MDS code $\mathcal{C}_3$	$r^{n'-1}$	$q > sr$	$= 1 + \frac{(s-1)(r-1)}{n-1} < 1 + \frac{r}{n'}$	repair-by-transfer	Thms 3, 4
The improved Ye-Barg code 2 in [19]	r^{n-1}	$q > r$	1 (optimal)	repair-by-transfer	
Punctured duplication-zigzag code [3]	$r^{n'-1}$	$q > s$	$= 1 + \frac{(s-1)(r-1)}{n-1} < 1 + \frac{r}{n'}$	repair-by-transfer	
The new MDS code $\mathcal{C}_4$	$r^{\frac{n}{r}}$	$q > N^{\binom{n-1}{r-1}}$	$= 1 + \frac{(s-1)(r-1)}{n-1} < 1 + \frac{r}{n'}$	repair-by-transfer	Thm 3
The first MDS code in [16]	$r^{\frac{n}{r}}$	$q > n^{(r-1)N+1}$	$= 1 + \frac{(s-1)(r-1)}{n-1} < 1 + \frac{r}{n'}$	repair-by-transfer	
Optimal sub-packetization code [8], [9]	$r^{\frac{n}{r}}$	$q > n$	1 (optimal)	optimal sub-packetization w.r.t. the bound in [18]	

VI. CONCLUDING REMARKS

In this paper, we provided a powerful transformation that can greatly reduce the sub-packetization level N of the original codes with respect to the same code length n at the cost of slightly increase the repair bandwidth. Four applications of the transformation were also given with two of them are explicit and over a small finite field. Comparisons show that the obtained MDS codes outperform the first MDS code construction in [16] in terms of the field size and outperform the first code in both [6] and [19] in terms of the sub-packetization level.

ACKNOWLEDGEMENT

The authors would like to thank the anonymous reviewers and TPC members for their valuable suggestions and comments, which have greatly improved the presentation and quality of this paper. This work was supported in part by the National Science Foundation of China under Grants 61801176 and 61871331.

REFERENCES

- [1] A. G. Dimakis, P. Godfrey, Y. Wu, M. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," *IEEE Trans. Inform. Theory*, vol. 56, no. 9, pp. 4539-4551, Sep. 2010.
- [2] K.V. Rashmi, N.B. Shah, and P.V. Kumar, "Optimal exact-regenerating codes for distributed storage at the MSR and MBR points via a product-matrix construction," *IEEE Trans. Inform. Theory*, vol. 57, no. 8, pp. 5227-5239, Aug. 2011.
- [3] T. Tamo, Z. Wang, and J. Bruck, "Zigzag codes: MDS array codes with optimal rebuilding," *IEEE Trans. Inform. Theory*, vol. 59, no. 3, pp. 1597-1616, Mar. 2013.
- [4] D.S. Papailiopoulos, A.G. Dimakis, and V.R. Cadambe, "Repair optimal erasure codes through hadamard designs," *IEEE Trans. Inform. Theory*, vol. 59, no. 5, pp. 3021-3037, May 2013.
- [5] J. Li and X. Tang, "Optimal exact repair strategy for the parity nodes of the $(k+2, k)$ Zigzag code," *IEEE Trans. Inform. Theory*, vol. 62, no. 9, pp. 4848-4856, Sep. 2016.
- [6] M. Ye and A. Barg, "Explicit constructions of high-rate MDS array codes with optimal repair bandwidth," *IEEE Trans. Inform. Theory*, vol. 63, no. 4, pp. 2001-2014, Apr. 2017.
- [7] J. Li, X. Tang, and C. Tian, "A generic transformation for optimal repair bandwidth and rebuilding access in MDS codes", in *Proc. IEEE Int. Symp. Inform. Theory*, Aachen, Germany, Jun. 2017, pp. 1623-1627.
- [8] B. Sasidharan, M. Vajha, and P.V. Kumar, "An explicit, coupled-layer construction of a high-rate MSR code with low sub-packetization level, small field size and all-node repair," *arXiv: 1607.07335 [cs.IT]*
- [9] M. Ye and A. Barg, "Explicit constructions of optimal-access MDS codes with nearly optimal sub-packetization," *IEEE Trans. Inform. Theory*, vol. 63, no. 10, pp. 6307-6317, Oct. 2017.
- [10] J. Li, X. Tang, and C. Tian, "A generic transformation to enable optimal repair in MDS codes for distributed storage systems", *IEEE Trans. Inform. Theory*, vol. 64, no. 9, pp. 6257-6267, Sept. 2018.
- [11] J. Li, X. Tang, and U. Parampalli, "A framework of constructions of minimal storage regenerating codes with the optimal access/update property," *IEEE Trans. Inform. Theory*, vol. 61, no. 4, pp. 1920-1932, Apr. 2015.
- [12] Z. Wang, T. Tamo, and J. Bruck, "Explicit minimum storage regenerating codes," *IEEE Trans. Inform. Theory*, vol. 62, no. 8, pp. 4466-4480, Aug. 2016.
- [13] S. Goparaju, A. Fazeli, and A. Vardy, "Minimum storage regenerating codes for all parameters," *IEEE Trans. Inform. Theory*, vol. 63, no. 10, pp. 6318-6328, Oct. 2017.
- [14] M. Blaum, P.G. Farrell, and H. van Tilborg, "Array codes," in *Handbook of Coding Theory*, V. Pless and W. C. Huffman, Eds. Elsevier Science, 1998, vol. II, ch. 22, pp. 1855-1909.
- [15] S. Goparaju, I. Tamo, and R. Calderbank, "An Improved Sub-Packetization Bound for Minimum Storage Regenerating Codes," *IEEE Trans. Inform. Theory*, vol. 60, no. 5, pp. 2770-2779, May 2014.
- [16] A.S. Rawat, I. Tamo, V. Guruswami, and K. Efremenko, "MDS code constructions with small sub-packetization and near-optimal repair bandwidth," *IEEE Trans. Inform. Theory*, vol. 64, no. 10, pp. 6506-6525, Oct. 2018.
- [17] N. Alon, "Combinatorial nullstellensatz," *Combinat. Probab. Comput.*, vol. 8, no. 1-2, pp. 7-29, Jan. 1999.
- [18] S.B. Balaji and P.V. Kumar, "A tight lower bound on the sub-packetization level of optimal-access MSR and MDS codes," [Online]. Available at: arXiv: 1710.05876v1 [cs.IT]
- [19] Y. Liu, J. Li, and X. Tang, "Explicit constructions of high-rate MSR codes with optimal access property over small finite fields," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4405-4413, Oct. 2018.
- [20] F.J. MacWilliams and N.J. Sloane, "The theory of error-correcting codes," *Elsevier*, 1977.
- [21] N.B. Shah, K.V. Rashmi, P.V. Kumar, and K. Ramchandran, "Distributed storage codes with repair-by-transfer and nonachievability of interior points on the storage-bandwidth tradeoff," *IEEE Trans. Inform. Theory*, vol. 58, no. 3, pp. 1837-1852, Mar. 2012.
- [22] T. Tamo, Z. Wang, and J. Bruck, "Access versus bandwidth in codes for storage," *IEEE Trans. Inform. Theory*, vol. 60, no. 4, pp. 2028-2037, Apr. 2014.