Bounds for Cooperative Locality Using Generalized Hamming Weights

Abdel-Ghaffar, Khaled A.S.; Weber, Jos H.

DOI
10.1109/ISIT.2017.8006618

Publication date
2017

Document Version
Accepted author manuscript

Published in
Proceedings of the 2017 IEEE International Symposium on Information Theory

Citation (APA)
Abdel-Ghaffar, K. A. S., & Weber, J. H. (2017). Bounds for Cooperative Locality Using Generalized Hamming Weights. In *Proceedings of the 2017 IEEE International Symposium on Information Theory* (pp. 699-703). IEEE. https://doi.org/10.1109/ISIT.2017.8006618

Bounds for Cooperative Locality Using Generalized Hamming Weights

Khaled A. S. Abdel-Ghaffar
University of California
Davis, CA 95616, USA
Email: ghaffar@ece.ucdavis.edu

Jos H. Weber
Delft University of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands
Email: j.h.weber@tudelft.nl

***Abstract*—The Cadambe-Mazumdar bound gives a necessary condition for a code to have a certain locality in case of a single erasure in terms of length, dimension, and Hamming distance of the code and of certain shortened codes. The bound has been generalized by Rawat, Mazumdar, and Vishwanath to recover multiple erasures in a cooperative repair scenario. In this paper, the generalized Hamming weights of the code and its shortened codes, which include the Hamming distance as one component, are incorporated to obtain bounds on locality to recover a single erasure or multiple erasures cooperatively. The new bounds give sharper necessary conditions than existing bounds.**

I. Introduction

A. Background

In cloud storage, data is stored on multiple nodes at geographically different locations. The effects of localized disruption of service can be effectively mitigated if the system has the ability to recover the data stored at a failed node by accessing other nodes that form a *repair set* for the failed node. In order to achieve this, coding is employed where data stored at a failed node is considered as an erasure in coding-theoretic terminology. The ability of a code to recover from node failures is measured by the well-known concept of Hamming distance. Inevitably coding introduces storage overhead to store redundant data and transmission overhead to exchange information between nodes in order to recover the lost data. Storage overhead is measured in terms of redundancy, a classical coding-theoretic concept. As a measure of transmission overhead, Gopalan *et al.* [2] introduced the new concept of *locality*, which is the number of nodes that need to be accessed in the repair process. In particular, a code has r-locality if the data stored at any given node can be recovered by accessing at most r other nodes, i.e., each node has a repair set of size at most r.

The above concept of locality assumes that only one node fails. This guarantees that all nodes in a repair set of a failed node are reliable. However, based on practical considerations, it is natural to address the case of more than one failed node. One approach proposes having multiple disjoint repair sets, each of size at most r. In particular, if each node has e disjoint repair sets and if the total number of failed nodes in the system is at most e, then each failed node has at least one repair set that does not contain any failed nodes. In this case, the code is said to have availability [9]. Another approach proposed in [7] associates to each node a set of at most $r+e-1$ other nodes such that if the node fails and up to $e-1$ nodes in the set also fail, the remaining nodes in the set form a repair set for the node associated with the set. In the above two approaches, repairing e failed nodes may require accessing data from er nodes as each failed node may require accessing r nodes to repair it.

To keep the number of accessed nodes from growing linearly with e, Rawat, Mazumdar, and Vishwanath [8] proposed a third approach in which each set of e nodes is assigned a disjoint repair set of at most r nodes, called *cooperative repair set*. If up to e nodes fail, then a repair set associated with the failing nodes can be used to recover the data stored at the e failed nodes. In this approach, the e failed nodes are not repaired independently as in the previous two approaches where r nodes are involved in the repair process of each failing node, but rather collectively as r nodes are involved in the repair process of all the failed nodes. Therefore, a code achieving this requirement is said to have (r,e)-*cooperative locality*. For each one of the three approaches, considerable literature is devoted to both the study of bounds on the code's length, redundancy, Hamming distance, locality, and the number of failed nodes allowed, as well as the construction of codes that are optimal in the sense of achieving these bounds. In case $e=1$, these three approaches reduce to the concept of r-locality as proposed in [2].

In this paper, we consider bounds on linear codes with (r,e)-cooperative locality as proposed in [8], including the special case in which $e=1$. Our approach is based on the concept of generalized Hamming weights, proposed by Wei [10] as a generalization of Hamming distance of a linear code. This allows us to generalize many known bounds on such codes leading to bounds which, for some code parameters, improve upon the tightest bounds known in the literature. The concept of generalized Hamming weights has been used already in [7] to bound locality but not in the cooperative scenario. Furthermore, even in the common case of $e=1$, our approach and results are different from [7].

B. Known Bounds

A fundamental inequality relating the parameters of any linear code over $\mathbb{F}_q$ of length n, dimension k, and Hamming distance d, i.e., an $[n,k,d]_q$ linear code, with r-locality in case of a single erasure was derived in [2] and states that

$$d \leq n-k-\left\lceil \frac{k}{r} \right\rceil + 2. \tag{1}$$

Notice that this implies the Singleton bound

$$d \leq n-k+1 \tag{2}$$

and suggests that there is a price to be paid in terms of Hamming distance for being able to correct a single erasure with small locality. Also, (1) implies that

$$k \leq \frac{r}{r+1}\, n. \tag{3}$$

A much stronger bound on k, due to Cadambe and Mazumdar [1], states that

$$k \leq \min_{1\leq t\leq \lfloor \frac{k-1}{r}\rfloor} \left\{tr + k^{\text{opt}}[n-t(1+r), d]_q\right\}, \tag{4}$$

where $k^{\text{opt}}[n,d]_q$ denotes the maximum dimension, k, of an $[n,k,d]_q$ linear code*. By setting t to $\lceil k/r\rceil - 1$ in the Cadambe-Mazumdar (CM)-bound (4) and bounding $k^{\text{opt}}[n-t(r+1),d]_q$ by $n-t(r+1)-d+1$ based on (2), the bound (1) follows.

For (r,e)-cooperative repair, Rawat, Mazumdar, and Vishwanath [8] generalized the bounds (1), (3), and (4) as

$$d \leq n-k-e\left\lceil\frac{k}{r}\right\rceil + e + 1, \tag{5}$$

$$k \leq \frac{r}{r+e}\, n, \tag{6}$$

and

$$k \leq \min_{1\leq t\leq \lfloor (k-1)/r\rfloor} \{tr + k^{\text{opt}}[n-t(e+r), d)]_q\}. \tag{7}$$

C. Our Contributions

Our main result is Theorem 2 which bounds the parameters of any linear code over $\mathbb{F}_q$ of length n, dimension k, and κ^{th}-generalized Hamming weight d_κ that has (r,e)-cooperative locality. In particular, we show that for $1\leq\kappa\leq k-r$,

$$k \leq \min_{1\leq t\leq \lfloor (k-\kappa)/r\rfloor} \{tr + k_\kappa^{\text{opt}}[n-t(e+r), d_\kappa]_q\}.$$

Here $k_\kappa^{\text{opt}}[n,d_\kappa]_q$ denotes the maximum dimension, k, of a linear code of length n over $\mathbb{F}_q$ with κ^{th}-generalized Hamming weight equal to d_κ. Notice that by setting $\kappa=1$, we obtain (7) as d_1 equals the Hamming distance, d, of the code. We also show that

$$d_\kappa \leq n-k-e\left\lceil\frac{k-\kappa+1}{r}\right\rceil + e + \kappa.$$

Again by setting $\kappa=1$, we obtain (5). We also prove that

$$d \leq \frac{q^\kappa - q^{\kappa-1}}{q^\kappa - 1}\left(n-k-e\left\lceil\frac{k-\kappa+1}{r}\right\rceil + e + \kappa\right)$$

for $1\leq\kappa\leq k-r$. This gives a new condition derived from the generalized Hamming weights that does not involve any of them except for the Hamming distance.

The rest of this paper is organized as follows. Section II derives basic results on cooperative repair sets leading to the new bounds presented in Section III. The paper is concluded in Section IV.

*The Cadambe-Mazumdar bound holds also for nonlinear codes. However, here we restrict it to linear codes.

II. Cooperative Repair sets

We are interested in the use of an $[n,k,d]_q$ linear code, $\mathcal{C}$, for repairing erasures, i.e., retrieving symbols erased during transmission. Suppose that a codeword $\mathbf{c}=(c_1,c_2,\ldots,c_n)$ is transmitted and, due to failures, symbols with indices in a set $\mathcal{E}\subseteq\{1,2,\ldots,n\}$ are erased. Then, the codeword $\mathbf{c}$ becomes the word $\mathbf{v}=(v_1,v_2,\ldots,v_n)$, where $v_i = ?$ for $i\in\mathcal{E}$ and $v_i=c_i$ for $i\notin\mathcal{E}$. Here ? denotes an erasure and $\mathcal{E}$ is called an *erasure set*. From $\mathbf{v}$, we would like to retrieve the codeword $\mathbf{c}$, i.e., repair all the erased symbols. This is possible if and only if $\mathbf{c}$ is the only codeword in the code that agrees with $\mathbf{v}$ in all its unerased symbols, i.e., symbols with indices in $\bar{\mathcal{E}}=\{1,2,\ldots,n\}\backslash\mathcal{E}$. In this case, it is possible to retrieve the erased symbols with indices in $\mathcal{E}$ by examining the symbols with indices in $\bar{\mathcal{E}}$. However, it may be sufficient to examine only symbols with indices in a subset of $\bar{\mathcal{E}}$ to retrieve all the erased symbols with indices in $\mathcal{E}$. Let $\mathcal{R}\subseteq\bar{\mathcal{E}}$ be a set of indices such that the erased symbols v_i, for all $i\in\mathcal{E}$, can be repaired using the unerased symbols v_i, $i\in\mathcal{R}$. Then, we say that $\mathcal{R}$ is a *cooperative repair set* for the set $\mathcal{E}$. This is the case if and only if all codewords that agree on symbols indexed by $\mathcal{R}$ also agree on symbols indexed by $\mathcal{E}$. By linearity of $\mathcal{C}$, this is the same as saying that every codeword in $\mathcal{C}$ which is zero on $\mathcal{R}$ is also zero on $\mathcal{E}$. Since we only consider linear codes in this paper, it is convenient to take this criterion as a definition of cooperative repair sets.

Definition 1. *Let $\mathcal{C}$ be an $[n,k,d]_q$ linear code. The set $\mathcal{R}\subseteq\{1,2,\ldots,n\}$ is a cooperative repair set for a set $\mathcal{E}$, disjoint from $\mathcal{R}$, if every codeword in $\mathcal{C}$ which is zero on $\mathcal{R}$ is also zero on $\mathcal{E}$.*

In general, a set may have more than one cooperative repair set. In practice, for a given set, it is desirable to specify a cooperative repair set of smallest size. From Definition 1, it follows that if $\mathcal{R}$ is a cooperative repair set for $\mathcal{E}$, then any superset of $\mathcal{R}$ disjoint from $\mathcal{E}$ is also a cooperative repair set for $\mathcal{E}$. Also, if $\mathcal{R}$ is a cooperative repair set for the set $\mathcal{E}$, then it is a cooperative repair set for any subset $\mathcal{E}'\subseteq\mathcal{E}$.

Recall that the support of a vector $(c_1,c_2,\ldots,c_n)$ is the set $\{i : 1\leq i\leq n, c_i\neq 0\}$. Then, from the definition, it follows that a necessary and sufficient condition for a nonempty set $\mathcal{E}$ to have a cooperative repair set is that it does not contain the support of a nonzero codeword in $\mathcal{C}$. Based on this condition, we conclude that for the $[n,k,d]_q$ linear code, $\mathcal{C}$, every nonempty subset of $\{1,2,\ldots,n\}$ of size less than d has a cooperative repair set, there is at least one subset of size d that has no repair set, and every subset of size greater than $n-k$ has no repair set.

The following lemma gives a necessary and sufficient condition for a set $\mathcal{R}$ to be a cooperative repair set for a set $\mathcal{E}$ in terms of a generator matrix of a code.

Lemma 1. *Let $\mathcal{C}$ be an $[n,k,d]_q$ linear code with generator matrix $\mathbf{G}$. The set $\mathcal{R}\subseteq\{1,2,\ldots,n\}$ is a cooperative repair set for the nonempty set $\mathcal{E}\subseteq\{1,2,\ldots,n\}$ disjoint from $\mathcal{R}$ if and only if every column in $\mathbf{G}$ indexed by $\mathcal{E}$ is in the space* *spanned by the columns in* $\mathbf{G}$ *indexed by* $\mathcal{R}$.

Proof. From the definition, $\mathcal{R}$ is a cooperative repair set for $\mathcal{E}$ if and only if for every vector $\mathbf{u} = (u_1, u_2, \ldots, u_k)$ over $\mathbb{F}_q$ for which the components of $\mathbf{uG}$ indexed by $\mathcal{R}$ are zeros, the components of $\mathbf{uG}$ indexed by $\mathcal{E}$ are also zeros. This is the same as saying that every vector $\mathbf{u}$ in the null space of the columns in $\mathbf{G}$ indexed by $\mathcal{R}$ is orthogonal to every column in $\mathbf{G}$ indexed by an element in $\mathcal{E}$. This is the case if and only if the space spanned by the columns indexed by $\mathcal{R}$ contains every column of $\mathbf{G}$ indexed by an element in $\mathcal{E}$. $\square$

Since $\mathbf{G}$ has at most k linearly independent columns, if a nonempty set $\mathcal{E} \subseteq \{1, 2, \ldots, n\}$ has a cooperative repair set $\mathcal{R}$ of size greater than k, then there is a subset of $\mathcal{R}$ of size not greater than k which is a cooperative repair set for $\mathcal{E}$.

Lemma 1 gives a characterization of cooperative repair sets in terms of generator matrices. The next lemma gives a characterization in terms of parity-check matrices. Let $\mathcal{C}^\perp$ be the dual code of $\mathcal{C}$, i.e., the vector space composed of all vectors over $\mathbb{F}_q$ orthogonal to every codeword in $\mathcal{C}$.

Lemma 2. *Let* $\mathcal{C}$ *be an* $[n, k, d]_q$ *linear code. The set* $\mathcal{R} \subseteq \{1, 2, \ldots, n\}$ *is a cooperative repair set for the nonempty set* $\mathcal{E} \subseteq \{1, 2, \ldots, n\}$ *disjoint from* $\mathcal{R}$ *if and only if there are* $|\mathcal{E}|$ *vectors* $\mathbf{x}_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,n})$*,* $i \in \mathcal{E}$*, in* $\mathcal{C}^\perp$ *such that for each* $i \in \mathcal{E}$*,* $x_{i,i} = 1$ *and* $x_{i,j} = 0$ *for* $j \notin \mathcal{R} \cup \{i\}$*. Hence,* $\mathcal{R} \subseteq \{1, 2, \ldots, n\}$ *is a cooperative repair set for the set* $\mathcal{E} \subseteq \{1, 2, \ldots, n\}$ *disjoint from* $\mathcal{R}$ *if and only if* $\mathcal{C}$ *has a parity-check matrix that includes* $\mathbf{x}_i$*,* $i \in \mathcal{E}$*, as rows.*

Proof. From Lemma 1, $\mathcal{R}$ is a cooperative repair set for $\mathcal{E}$ if and only if every column in $\mathbf{G}$ indexed by $i \in \mathcal{E}$ is a linear combination of the columns indexed by $\mathcal{R}$. This is the case if and only if for every index $i \in \mathcal{E}$, there is such a vector $\mathbf{x}_i$ in $\mathcal{C}^\perp$. $\square$

Next, we define the cooperative locality of a code.

Definition 2. *An* $[n, k, d]_q$ *linear code* $\mathcal{C}$ *has* (r, e)*-cooperative locality, where* $1 \le e < d$*, if every set* $\mathcal{E}$ *of size* e *has a cooperative repair set of size* r *or less.*

Example 1. We consider the $[2^m - 1, 2^m - m - 1, 3]_2$ Hamming code. Any parity-check matrix of the code is of size $m \times (2^m - 1)$, the columns of which are the $2^m - 1$ nonzero vectors of length m. It follows that each row has weight 2^{m-1}. From Lemma 2, a set $\mathcal{R} \subseteq \{1, 2, \ldots, 2^m - 1\} \backslash \{i\}$ is a cooperative repair set for $\{i\}$, where $1 \le i < 2^m$, if and only if there is such a row, $\mathbf{x}_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,2^m-1})$, for which $x_{i,i} = 1$ and $x_{i,j} = 0$ for all $j \notin \mathcal{R} \cup \{i\}$. Hence, the smallest value of r such that the code has $(r, 1)$-cooperative locality is $r = 2^{m-1} - 1$. (Notice that this result is derived in [5] for cyclic Hamming codes.) Furthermore, from Lemma 2, a set $\mathcal{R} \subseteq \{1, 2, \ldots, 2^m - 1\} \backslash \{i_1, i_2\}$ is a cooperative repair set for $\{i_1, i_2\}$, where $1 \le i_1 < i_2 < 2^m$, if and only if there are two rows $\mathbf{x}_{i_1} = (x_{i_1,1}, x_{i_1,2}, \ldots, x_{i_1,2^m-1})$ and $\mathbf{x}_{i_2} = (x_{i_2,1}, x_{i_2,2}, \ldots, x_{i_2,2^m-1})$ in a parity-check matrix of the code for which $x_{i_1,i_1} = x_{i_2,i_2} = 1$, $x_{i_1,i_2} = x_{i_2,i_1} = 0$, and $x_{i_1,j} = x_{i_2,j} = 0$ for all $j \notin \mathcal{R} \cup \{i_1, i_2\}$. Since the sum of any two rows of weight 2^{m-1} in the parity-check matrix is a vector of the same weight, it follows that $x_{i_1,j} = x_{i_2,j} = 0$ for exactly $2^{m-2} - 1$ values of j, $1 \le j < 2^m$. Hence, the smallest value of r such that the code has $(r, 2)$-cooperative locality is $r = (2^m - 1) - 2 - (2^{m-2} - 1) = 3 \times 2^{m-2} - 2$. $\square$

In general, finding for each e the smallest r for which a given code has (r, e)-cooperative locality can be difficult. In the next section we give lower bounds on such r.

III. Bounds Using Generalized Hamming Weights

In the following, we give a generalization and a strengthening of the bounds (1), (4), (5), and (7) using generalized Hamming weights.

Recall that the *support,* $\chi(\mathcal{C})$*, of a code* $\mathcal{C}$ is the set of not-always-zero symbol positions, i.e.,

$$\chi(\mathcal{C}) = \{i : \exists\, (c_1, c_2, \ldots, c_n) \in \mathcal{C}, c_i \ne 0\}.$$

For $1 \le \kappa \le k$, the κ^{th} generalized Hamming weight, $d_\kappa(\mathcal{C})$, of an $[n, k, d]_q$ linear code $\mathcal{C}$ is the size of a smallest support of a κ-dimensional linear subspace of $\mathcal{C}$, i.e.,

$$d_\kappa(\mathcal{C}) \;=\; \min\{|\chi(\mathcal{D})| : \mathcal{D} \text{ is a linear subcode of } \mathcal{C} \text{ of dimension } \kappa\}.$$

In particular, $d_1(\mathcal{C})$ equals the Hamming distance, d, of the linear code. In [10], Wei studied the generalized Hamming weights of linear codes and has shown that they obey a generalized Singleton bound [10, Corollary 1] given by

$$d_\kappa(\mathcal{C}) \le n - k + \kappa. \tag{8}$$

We start with the following result.

Theorem 1. *For an* $[n, k, d]_q$ *linear code* $\mathcal{C}$ *with* (r, e)*-cooperative locality, where* $1 \le e < d$*, we have*

$$r \ge d_e(\mathcal{C}^\perp) - e.$$

Proof. Let $\mathcal{E} \subseteq \{1, 2, \ldots, n\}$ be a set of size e with a repair set $\mathcal{R}$ of size at most r. From Lemma 2, the space spanned by the vectors $\mathbf{x}_i$, $i \in \mathcal{E}$, which are linearly independent, is a subcode of $\mathcal{C}^\perp$ of dimension e with support in $\mathcal{E} \cup \mathcal{R}$ of size at most $e + r$. According to the definition of the generalized Hamming weight, the size of this support gives an upper bound on the e^{th} generalized Hamming weight, $d_e(\mathcal{C}^\perp)$, of $\mathcal{C}^\perp$. $\square$

Example 2. It follows from [10, Corollary 3] that the dual code, $\mathcal{C}^\perp$, of the $[2^m - 1, 2^m - m - 1, 3]_2$ Hamming code $\mathcal{C}$ has generalized Hamming weights given by $d_\kappa(\mathcal{C}^\perp) = \sum_{i=1}^{\kappa} 2^{m-i}$ for $1 \le \kappa \le m$. Hence, for $e = 1$, we have $r \ge 2^{m-1} - 1$ and for $e = 2$, we have $r \ge 3 \times 2^{m-2} - 2$. From Example 1, we notice that these bounds are tight. $\square$

Although Theorem 1 gives the smallest value of locality in Example 2, it is not always tight. For example, consider an $[n, k, d]_q$ linear code $\mathcal{C}$ for which the dual code, $\mathcal{C}^\perp$, has a codeword of weight two. Then, $d_1(\mathcal{C}^\perp) \le 2$ and Theorem 1 does not eliminate the possibility that $\mathcal{C}$ has $(1, 1)$-cooperative locality regardless of its Hamming distance or rate. However, this possibility is eliminated by the simple bound (3) if the rate is greater than 1/2.

Next, we proceed to give a generalization of the bounds (4) and (7) that involves generalized Hamming weights. We start with two lemmas.

Lemma 3. *Let $\mathcal{C}$ be an $[n,k,d]_q$ linear code. If $\mathcal{E}$ is a nonempty set of size e that has a cooperative repair set $\mathcal{R}$ of size $r<k$, then there exists a linear code $\mathcal{C}_r$ over $\mathbb{F}_q$ of length at most $n-(e+r)$, dimension $k-r$, and generalized Hamming weight $d_\kappa(\mathcal{C}_r)\geq d_\kappa(\mathcal{C})$ for $1\leq\kappa\leq k-r$.*

Proof. As the space spanned by the columns of $\mathbf{G}$ is of dimension k and the space spanned by the columns of $\mathbf{G}$ indexed by $\mathcal{R}$ is of dimension at most r, there is a set $\mathcal{S}\subseteq\{1,2,\ldots,n\}$ of size $k-r$ such that the columns in $\mathbf{G}$ indexed by $\mathcal{S}$ are linearly independent and none of them is a linear combination of the r columns in $\mathbf{G}$ indexed by $\mathcal{R}$. As $\mathcal{R}$ is a repair set for $\mathcal{E}$, it follows, from Lemma 1, that $\mathcal{S}$ and $\mathcal{E}$ are disjoint. For each $j\in\mathcal{S}$, there is a nonzero vector, $\mathbf{u}^{(j)}$, of length k which is orthogonal to every column of $\mathbf{G}$ indexed by $\mathcal{R}\cup\mathcal{S}\backslash\{j\}$ but not orthogonal to the column indexed by j. Hence, $\mathbf{u}^{(j)}\mathbf{G}$ is a nonzero codeword $\mathbf{c}^{(j)}=(c_1^{(j)},c_2^{(j)},\ldots,c_n^{(j)})$ in $\mathcal{C}$ such that $c_j^{(j)}\neq 0$ and $c_{j'}^{(j)}=0$ for all $j'\in\mathcal{R}\cup\mathcal{S}\backslash\{j\}$. Since $\mathbf{c}^{(j)}$ is zero on $\mathcal{R}$, which is a repair set for $\mathcal{E}$, then $\mathbf{c}^{(j)}$ is also zero on $\mathcal{E}$. The collection of the $k-r$ vectors $\mathbf{c}^{(j)}$, $j\in\mathcal{S}$, spans a $(k-r)$-dimensional subcode of $\mathcal{C}$, the support of which does not intersect with $\mathcal{E}\cup\mathcal{R}$ of size $e+r$. In particular, this support is of size at most $n-(e+r)$. Deleting the symbols with indices not in this support gives the code $\mathcal{C}_r$. Clearly, $d_\kappa(\mathcal{C}_r)\geq d_\kappa(\mathcal{C})$ for $1\leq\kappa\leq k-r$. $\square$

Lemma 4. *Let $\mathcal{C}$ be an $[n,k,d]_q$ linear code with (r,e)-cooperative locality, where $1\leq e<d$, and let $1\leq\kappa\leq k-r$. Then, for $1\leq t\leq\lfloor(k-\kappa)/r\rfloor$, there exists a linear code $\mathcal{C}_{tr}$ over $\mathbb{F}_q$ of length at most $n-t(e+r)$, dimension $k-tr$, and generalized Hamming weight $d_\kappa(\mathcal{C}_{tr})\geq d_\kappa(\mathcal{C})$.*

Proof. We iteratively construct a subset of $\{1,2,\ldots,n\}$ of size te that has a cooperative repair set of size at most tr. Pick a subset $\mathcal{E}_1\subseteq\{1,2,\ldots,n\}$ of size e and let $\mathcal{R}_1$ be a cooperative repair set of size at most r for $\mathcal{E}_1$. Let $\mathcal{E}_1'=\mathcal{E}_1$ and $\mathcal{R}_1'=\mathcal{R}_1$. Next, if $|\mathcal{E}_1'\cup\mathcal{R}_1'|\leq n-e$, pick a subset $\mathcal{E}_2\subseteq\{1,2,\ldots,n\}$ of size e disjoint from $\mathcal{E}_1'\cup\mathcal{R}_1'$. Let $\mathcal{R}_2$ be a cooperative repair set of size at most r for $\mathcal{E}_2$ and

$$\mathcal{R}_2'=(\mathcal{R}_1'\cup\mathcal{R}_2)\backslash\mathcal{E}_1'. \tag{9}$$

We will argue, using Definition 1, that $\mathcal{R}_2'$ is a cooperative repair set for $\mathcal{E}_2'=\mathcal{E}_1'\cup\mathcal{E}_2$. Consider an arbitrary codeword in $\mathcal{C}$ which is zero on $\mathcal{R}_2'$. Since $\mathcal{R}_1'\subseteq\mathcal{R}_2'$ as $\mathcal{E}_1'$ and $\mathcal{R}_1'$ are disjoint, then such a codeword is zero on $\mathcal{E}_1'$. From (9), it follows that the codeword is zero on $\mathcal{R}_2$ and, hence, is zero on $\mathcal{E}_2$ as well. This proves that $\mathcal{R}_2'$ is a cooperative repair set for $\mathcal{E}_2'$. Notice that $\mathcal{E}_2'$ is of size $2e$ and $\mathcal{R}_2'$ is of size at most $2r$. This procedure can be repeated to form a set $\mathcal{E}_i'$ of size ie with a cooperative repair set $\mathcal{R}_i'$ of size at most ir. Indeed, suppose that we have a set $\mathcal{R}_{i-1}'$ of size at most $(i-1)r$ which is a cooperative repair set for $\mathcal{E}_{i-1}'=\mathcal{E}_1\cup\mathcal{E}_2\cup\cdots\cup\mathcal{E}_{i-1}$. If $|\mathcal{E}_{i-1}'\cup\mathcal{R}_{i-1}'|\leq n-e$, pick a subset $\mathcal{E}_i\subseteq\{1,2,\ldots,n\}$ of size e disjoint from $\mathcal{E}_{i-1}'\cup\mathcal{R}_{i-1}'$. Let $\mathcal{R}_i$ be a cooperative repair set of size at most r for $\mathcal{E}_i$ and

$$\mathcal{R}_i'=(\mathcal{R}_{i-1}'\cup\mathcal{R}_i)\backslash\mathcal{E}_{i-1}'.$$

Using the same argument stated above for $i=2$, it follows that $\mathcal{R}_i'$ is a cooperative repair set for $\mathcal{E}_i'=\mathcal{E}_{i-1}'\cup\mathcal{E}_i$. Notice that $\mathcal{E}_i'$ is of size ie and $\mathcal{R}_i'$ is of size at most ir. Since $t\leq\lfloor(k-\kappa)/r\rfloor$, then from (6), we have $t\leq(n-\kappa)/(e+r)$, $(t-1)(e+r)\leq n-e$, and the procedure can continue until $i=t$. Hence, we can indeed construct a set $\mathcal{E}=\mathcal{E}_t'$ of size te with a cooperative repair set $\mathcal{R}$ of size tr where $\mathcal{R}=\mathcal{R}_t'$ if $|\mathcal{R}_t'|=tr$ or a superset of $\mathcal{R}_t'$ obtained by adding $tr-|\mathcal{R}_t'|$ indices not in $\mathcal{E}\cup\mathcal{R}_t'$ to $\mathcal{R}_t'$ if $|\mathcal{R}_t'|<tr$. The result then follows from Lemma 3. $\square$

Let $d_\kappa^{\mathrm{opt}}[n,k]_q$ denote the κ^{th} generalized Hamming weight, d_κ, maximized over all linear codes over $\mathbb{F}_q$ of length n and dimension k. Let $k_\kappa^{\mathrm{opt}}[n,d_\kappa]_q$ denote the dimension, k, maximized over all linear codes over $\mathbb{F}_q$ of length n and κ^{th} generalized Hamming weight equal to d_κ. Let $n_\kappa^{\mathrm{opt}}[k,d_\kappa]_q$ denote the length, n, minimized over all linear codes over $\mathbb{F}_q$ of dimension k and κ^{th} generalized Hamming weight equal to d_κ. In case $\kappa=1$, we drop the subscript κ. The following result is a direct consequence of Lemma 4.

Theorem 2. *Let $\mathcal{C}$ be an $[n,k,d]_q$ linear code with (r,e)-cooperative locality, where $1\leq e<d$. Let d_κ be the κ^{th} generalized Hamming weight of the code, where $1\leq\kappa\leq k-r$. Then,*

$$\begin{aligned}
d_\kappa &\leq \min_{1\leq t\leq\lfloor(k-\kappa)/r\rfloor}\{d_\kappa^{\mathrm{opt}}[n-t(e+r),k-tr]_q\},\\
k &\leq \min_{1\leq t\leq\lfloor(k-\kappa)/r\rfloor}\{tr+k_\kappa^{\mathrm{opt}}[n-t(e+r),d_\kappa)]_q\},\\
n &\geq \max_{1\leq t\leq\lfloor(k-\kappa)/r\rfloor}\{t(e+r)+n_\kappa^{\mathrm{opt}}[k-tr,d_\kappa]_q\}.
\end{aligned}$$

From Theorem 2, we have the following explicit bound on the parameters of an $[n,k,d]_q$ linear code with (r,e)-cooperative locality.

Theorem 3. *Let $\mathcal{C}$ be an $[n,k,d]_q$ linear code with (r,e)-cooperative locality, where $1\leq e<d$. Then, for $1\leq\kappa\leq k-r$,*

$$d_\kappa(\mathcal{C})\leq n-k-e\left\lceil\frac{k-\kappa+1}{r}\right\rceil+e+\kappa.$$

Proof. Bounding $d_\kappa^{\mathrm{opt}}[n-t(e+r),k-tr]_q$ in Theorem 2 using the generalized Singleton bound (8), we get

$$\begin{aligned}
d_\kappa(\mathcal{C}) &\leq \min_{1\leq t\leq\lfloor(k-\kappa)/r\rfloor}\{n-t(e+r)-(k-tr)+\kappa\}\\
&= \min_{1\leq t\leq\lfloor(k-\kappa)/r\rfloor}\{n-k-te+\kappa\}.
\end{aligned}$$

Setting $t=\lfloor(k-\kappa)/r\rfloor=\lceil(k-\kappa+1)/r\rceil-1$, we get the stated result. $\square$

Example 3. It follows from [10, Corollary 4] that the $[15,11,3]_2$ Hamming code $\mathcal{C}$ has

$$d_\kappa(\mathcal{C})=3,5,6,7,9,10,11,12,13,14,15$$

TABLE I
LOWER BOUNDS ON r FOR (r, e)-COOPERATIVE LOCALITY OF THE $[15, 11, 3]_2$ HAMMING CODE FOR $e = 1$ AND $e = 2$ BASED ON THEOREM 3.

$e\backslash\kappa$	1	2	3	4	5	6	7	8	9	10	11
1	4	5	5	4	7	6	5	4	3	2	1
2	6	10	9	8	7	6	5	4	3	2	1

for $\kappa = 1, 2, \ldots, 11$, respectively. Applying Theorem 3 in the cases $e = 1$ and $e = 2$, we obtain the lower bounds on r given in Table I. From this table, it is clear that the sharpest bounds on r for $e = 1$ and $e = 2$, which are 7 and 10, are attained for $\kappa = 5$ and $\kappa = 2$, respectively, and both significantly improve upon the corresponding bounds for $\kappa = 1$. Actually, from Example 1, we have $r = 7$ and $r = 10$ for $e = 1$ and $e = 2$, respectively. In general, for the $[2^m-1, 2^m-m-1, 3]_2$ Hamming code $\mathcal{C}$, we show that Theorem 3 gives $r \geq 2^{m-1}-1$ and $r \geq 3\times 2^{m-2}-2$ for $e = 1$ and $e = 2$, respectively. From [10, Corollary 4], it can be deduced that $d_{2^s-s}(\mathcal{C}) = 2^s + 1$ for $1 \leq s \leq m-1$. For $e = 1$, setting $s = m-1$ and $\kappa = 2^{m-1} - (m-1)$ in Theorem 3 gives $r \geq 2^{m-1} - 1$. For $e = 2$, setting $s = m-2$ and $\kappa = 2^{m-2} - (m-2)$ in Theorem 3 gives $r \geq 3\times 2^{m-2}-2$. Again, these bounds agree with the smallest values of r for $e = 1$ and 2 as deduced in Example 1. □

Example 4. The binary second order Reed-Muller code, RM$(2, 5)$, has length $n = 32$, dimension $k = 16$, and Hamming distance $d = 8$. It is shown in [5] to have 7-locality to correct $e = 1$ erasure. We show that the number 7 is the minimum locality for this code, i.e., it does not have 6-locality. It is reported in [10] that $d_{10} = 26$. With $\kappa = 10$, Theorem 3 eliminates the possibility that $r = 6$. On the other hand, using the tables in [3], we notice that the CM bound (4) does not eliminate the possibility that $r = 3$. □

In applying Theorems 2 and 3, the generalized Hamming weights need to be known. The reader may refer to [10] where the generalized Hamming weights are determined for Hamming codes, Reed-Muller codes, binary Golay code, and Reed-Solomon codes, and to [6, Chapter 1, Section 3] for references for other codes. However, in general it is not easy to determine the generalized Hamming weights for an arbitrary code. We can weaken Theorem 3 to obtain a bound on locality that does not involve any of the generalized Hamming weights except for the Hamming distance. Consider an $[n, k, d]_q$ linear code, $\mathcal{C}$, with κ^{th} generalized Hamming weight equal to d_κ, where $1 \leq \kappa \leq k-1$. Let $\mathcal{D}$ be a linear subcode of $\mathcal{C}$ of dimension κ that has support of size d_κ. Deleting all the symbols with indices not in the support of $\mathcal{D}$ from its codewords gives a code of length d_κ, dimension κ, and Hamming distance at least d. Applying the Griesmer bound [6, Chapter 1, Theorem 3.12] to this code yields $d_\kappa(\mathcal{C}) \geq \sum_{i=0}^{\kappa-1} \lceil d/q^i \rceil$. (See also [4, Corollary 2] where this inequality is first stated for $q = 2$.) Bounding $d_\kappa(\mathcal{C})$ in Theorem 3 using this inequality, we have the following result.

Theorem 4. *Let $\mathcal{C}$ be an $[n, k, d]_q$ linear code with (r, e)-cooperative locality, where $1 \leq e < d$. Then, for $1 \leq \kappa \leq k - r$,*

$$\sum_{i=0}^{\kappa-1} \left\lceil \frac{d}{q^i} \right\rceil \leq n - k - e \left\lceil \frac{k-\kappa+1}{r} \right\rceil + e + \kappa,$$

which implies that

$$d \leq \frac{q^\kappa - q^{\kappa-1}}{q^\kappa - 1} \left(n - k - e \left\lceil \frac{k-\kappa+1}{r} \right\rceil + e + \kappa \right).$$

In simplicity, the bounds in Theorem 4 are comparable to the bounds (1) and (5) as they give an explicit necessary condition for an $[n, k, d]_q$ linear code to have a given r-locality or a given (r, e)-cooperative locality. Actually, setting $\kappa = 1$ in Theorem 4 gives the bound in (5) which reduces to that in (1) for $e = 1$. However, tighter bounds may be obtained by setting $\kappa > 1$.

Example 5. For the $[15, 11, 3]_2$ Hamming code, (1) and (5) eliminate the possibilities that the code has $(3, 1)$ and $(5, 2)$-cooperative localities but not the possibilities that the code has $(4, 1)$ and $(6, 2)$-cooperative localities. On the other hand, with $\kappa = 2$, the second, and weaker, inequality in Theorem 4 eliminates the last two possibilities. Actually, this inequality shows that the code does not have $(9, 2)$-cooperative locality. This is sharp as it is shown in Example 1 that the code has $(10, 2)$-cooperative locality. □

IV. CONCLUSION

By incorporating the generalized Hamming weights, new bounds on cooperative localities are derived. Through examples, it is shown that these bounds improve upon other bounds available in the literature.

REFERENCES

[1] V. Cadambe and A. Mazumdar, "Bounds on the size of locally recoverable codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 5787–5794, Nov. 2015.

[2] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, "On the locality of codeword symbols," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, ipp. 6925–6934, Nov. 2012.

[3] M. Grassl. (2008, Jun. 10). *Bounds on the minimum distance of linear codes* [Online]. Available: http://www.codetables.de/

[4] T. Helleseth, T. Kløve, and Ø. Ytrehus, "Generalizations of the Griesmer bound," in Error Control, Cryptology, and Speech Compression. New York, NY, USA: Springer-Verlag, 1994, vol. LNCS 829, pp. 41–52.

[5] P. Huang, E. Yaakobi, H. Uchikawa, and P. H. Siegel, "Binary linear locally repairable codes," *IEEE Trans. Inf. Theory*, vol. 62, no. 11, pp. 6268–6283, Nov. 2016.

[6] V. S. Pless and W. C. Hoffman, Eds. *Handbook of Coding Theory*. Amsterdam, The Netherlands: Elsevier, 1998.

[7] N. Prakash, G. M. Kamath, V. Lalitha, and P. V. Kumar, "Optimal linear codes with a local-error-correction property," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Cambridge, MA, Jul. 2012, pp. 2776–2780.

[8] A. S. Rawat, A. Mazumdar, and S. Vishwanath, "Cooperative local repair in distributed storage," *EURASIP J. Adv. Signal Process.*, pp. 1–17, Dec. 2015.

[9] A. S. Rawat, D. S. Papailiopoulos, A. G. Dimakis, and S. Vishwanath, "Locality and availability in distributed storage," *IEEE Trans. Inf. Theory*, vol. 62, no. 8, pp. 4481–4493, Aug. 2016.

[10] V. K. Wei, "Generalized Hamming weights for linear codes," *IEEE Trans. Inf. Theory*, vol. 37, no. 5, pp. 1412–1418, Sep. 1991.