Beyond Cut-Set Bounds - The Approximate Capacity of D2D Networks

Avik Sengupta
Hume Center for National Security and Technology
Department of Electrical and Computer Engineering
Virginia Tech, Blacksburg, VA 24060, USA
Email: aviksg@vt.edu

Ravi Tandon
Discovery Analytics Center
Department of Computer Science
Virginia Tech, Blacksburg, VA 24060, USA
Email: tandonr@vt.edu

Abstract—Device-to-Device (D2D) communication is emerging as a viable solution for alleviating the severe capacity crunch in content-centric wireless networks. D2D encourages backhaul-free communication directly between devices with similar content requirements grouped into clusters. In this work, a self-sustaining D2D network is considered, where a set of commonly requested files are completely stored within the collective devices memories in a cluster and file requests from devices are serviced by local inter-device multicast transmissions. For such a network, new information theoretic converse results are developed, in the form of a lower bound on the minimum D2D multicast rate as a function of the storage per device. The proposed converse is then used to characterize the approximate tradeoff between the device storage and D2D multicast rate to within a constant multiplicative gap of 8.

I. INTRODUCTION

In recent times, the dynamics of traffic over wireless networks has undergone a paradigm shift to become increasingly content centric with multimedia content distribution holding precedence. This necessitates careful design of content storage and over-the-air distribution schemes within the networks to efficently use scarce spectral resources to handle heavy traffic loads as evidenced by a wealth of recent research [1]–[12].

Device-to-device (D2D) communication systems have been proposed as one of the solutions that could help in alleviating the severe capacity crunch in wireless networks [13]–[16]. D2D uses low-power direct communications within devices grouped into clusters to facilitate large scale frequency reuse thereby freeing up precious spectral resources. We consider a D2D network with K devices, and N files (denoted by $F_1, \ldots, F_N$, each of size B units), where the storage of each device is MB units. The base-station populates the devices with functions of files, depending on the number of users and available device storage. The D2D system is self-sustaining i.e., all N files must be recoverable from the collective storage memory of the devices, which is equivalent to the condition $KM \geq N$. Once the devices reveal their file requests, all the devices within the cluster transmit multicast messages over the shared link to service the requests. Figure 1 illustrates the system model. The storage/multicast mechanisms must be designed such that any device must be able to decode its requested file using its local storage and the multicast transmissions from other devices.

Recently, Maddah-Ali and Niesen have investigated a related caching problem [1]–[4] for which it has been shown that, by jointly designing the content storage and delivery, order-wise improvement in transmission rates can be achieved. The main difference between the caching problem studied in [1,2] and the D2D problem is the *distributed nature of multicast transmissions*. The caching problem allows for centralized delivery, in which the multicast can be any arbitrary function of all N files. In the D2D problem however, the outgoing multicast from each device can only depend on the content stored at that device. The focus of this work is on the fundamental tradeoff between the device storage and the total multicast transmission rate (referred to as the (M, R) tradeoff) for the D2D problem. In [15,16], Ji et.al. present new storage/delivery mechanisms for the D2D problem. The results in [16] show that for a self-sustaining D2D network, when the devices can use local inter-device multicast transmissions to satisfy the requests of others within the cluster, order-wise rate improvements similar to [1,2] can be achieved. The authors present an achievable scheme as well as a cut-set based argument to obtain upper and lower bounds on the (M, R) tradeoff for the D2D problem.

In this work, we develop a new converse (lower bound) on the (M, R) tradeoff for the D2D problem. The new lower bound addresses constraints that are unique to the D2D network, i.e., the local dependence of device multicast on its stored content. We show that the new converse is tighter than the existing one presented in [16, Theorem 2] (which are primarily based on the relaxation of the D2D problem to the centralized caching problem) for all values of problem parameters. Our result shows that the achievable scheme presented in [16] is indeed optimal for the lowest allowable value of storage memory for a self-sustaining D2D network i.e., at $M = N/K$. Using our new converse result, together with the achievable scheme of [16], we characterize the (M, R) tradeoff for the D2D network to within a maximum constant multiplicative gap of 8 for all values of problem parameters.

The paper is organized as follows: Section II outlines the system model. Section III details the analysis of the new lower bound on the storage vs. rate tradeoff for the D2D network, while Section IV illustrates the intuition behind the new bound through an example. Finally, Section V concludes the paper.
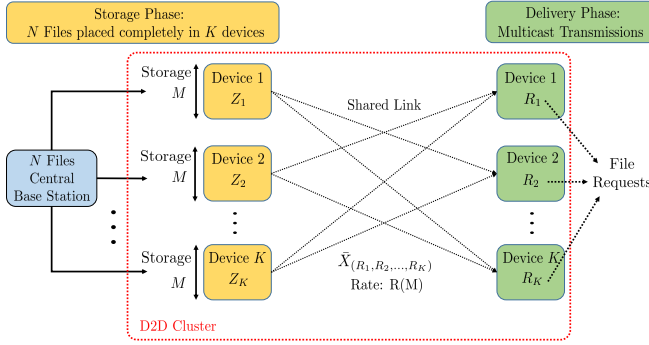
Fig. 1. System Model for device-to-device (D2D) Network.

Notation: Let Y_i be a random variable. $Y_{[a:b]}$, where $a < b$ denotes the set of random variables $\{Y_i : i = a, a+1, \ldots, b-1, b\}$. Also, $Y_{[a,b]}$ denotes the set $\{Y_i : i = a, b\}$. $Y_{[n]}$ denotes a set of any n *arbitrary* random variables Y_i such that $|Y_{[n]}| = n$. $\mathbb{N}^+$ denotes the set of natural numbers; the function $(x)^+ = \max(0, x)$; $\lceil x \rceil$ denotes the ceil function; $\lfloor x \rfloor$ denotes the floor function.

II. SYSTEM MODEL

The D2D network has K users, and a total of N files, denoted by $F_1, .., F_N$, where each file is assumed to be of size B, for some $B \in \mathbb{N}^+$. Formally, the files F_n are i.i.d and distributed as

$$: F_n \sim \text{Unif}\{1, 2, \ldots, 2^B\}. \tag{1}$$

We next define the key components of the D2D problem:

Definition 1 (**Storage**). *The content storage phase consists of K storage functions, which map the files $(F_1, \ldots, F_N)$ into the device storage content*

$$Z_k \triangleq \phi_k(F_1, \ldots, F_N), \tag{2}$$

for each user $k \in \{1, 2, \ldots, K\}$. The maximum allowable size of the contents of each device storage, Z_k, is MB units. An additional constraint of self-sustainability is enforced i.e., $KM \geq N$, which ensures all the files in the system are completely stored within the D2D cluster.

Definition 2 (**Delivery**). *In the delivery phase each device $k \in \{1, 2, \ldots, K\}$ uses one of N^K encoding functions to map its storage contents, Z_k, to a multicast transmission:*

$$X^k_{(R_1, \ldots, R_K)} \triangleq \psi_{(R_1, \ldots, R_K)}(Z_k) \quad \forall k \in \{1, 2, \ldots, K\}. \tag{3}$$

over the shared link in response to the requests $(R_1, \ldots, R_K) \in \{1, 2, \ldots, N^K\}$. Each such transmission has rate RB/K units. The total multicast rate. when all K devices transmit, is RB units.

Definition 3 (**File Decoding**). *Once the multicast transmission is received, KN^K decoding functions map the received signal over the shared link $X^1_{(R_1, \ldots, R_K)}, \ldots, X^K_{(R_1, \ldots, R_K)}$ and the storage content Z_k to the estimate*

$$\hat{F}_{R_k} \triangleq \mu_{(R_1, \ldots, R_K), k}\left(X^1_{(R_1, \ldots, R_K)}, \ldots, X^K_{(R_1, \ldots, R_K)}, Z_k\right), \tag{4}$$

of the requested file F_{R_k} for user $k \in \{1, 2, \ldots, K\}$.

For the (M, R) D2D scheme, the probability of error is defined as:

$$P_e \triangleq \max_{(R_1, \ldots, R_K) \in [N]^K} \max_{k \in [K]} \mathbb{P}(\hat{F}_{R_k} \neq F_{R_k}). \tag{5}$$

Definition 4 (**Storage vs. Rate Tradeoff**). *The storage-rate pair (M, R) is achievable if, for any $\epsilon > 0$, there exists an (M, R) D2D scheme for which $P_e \leq \epsilon$. The storage vs. rate tradeoff is defined as:*

$$R(M) \triangleq \inf \{R : (M, R) \text{ is achievable}\}. \tag{6}$$

III. MAIN RESULTS AND DISCUSSION

We next present our main result which gives a new information theoretic lower bound on the storage vs. rate tradeoff for the D2D problem.

Theorem 1. *In a D2D system with K devices and N files, with each device having storage size M, and $KM \geq N$, the optimal content delivery rate $R^*(M)$ is lower bounded by:*

$$R^*(M) \geq R_{LB}(M) \triangleq$$
$$\max_{\substack{s \in \{1, \ldots, K\}, \\ \ell \in \{1, \ldots, \lceil \frac{N}{s} \rceil\}}} \left\{ \frac{N - sM - \left(\frac{\mu}{s+\mu}\right)(N - \ell s)^+}{\ell \left(\frac{K-s}{K}\right)} \right\}, \tag{7}$$

where $\mu = \min\left(\lceil \frac{N - \ell s}{\ell} \rceil, K - s\right) \ \forall s, \ell$.

Proof. The proof of Theorem 1 is given in Appendix A $\qquad \square$

The expression in Theorem 1 has two parameters - (1) s, which is related to the number of devices and as well as (2) a new parameter ℓ. Compared to [16, Theorem 2], the additional parameter ℓ adds further flexibility to the lower bound expression and better models the file decoding through the interaction of the device storage contents and the multicast transmissions. Considering the cut-set based bound presented in [16, Theorem 2], the first term inside the $\max(\cdot)$ function is simply the expression for the centralized caching problem presented in [1, Theorem 2]. Furthermore, setting $\ell = \lceil N/s \rceil$ in Theorem 1 yields

$$R^*(M) \geq \max_{s \in \{1, \ldots, K\}} \left\{ \frac{N - sM}{\lceil \frac{N}{s} \rceil \left(\frac{K-s}{K}\right)} \right\}, \tag{8}$$

which is greater than the cut-set expression in the first term inside the $\max(\cdot)$ function owing to the factor $\frac{K-s}{K} \leq 1$ in the denominator. The second term inside the $\max(\cdot)$ function can be obtained by setting $s = 1, \ell = N$ in Theorem 1. Thus, the proposed bound is generally tighter than the bound in [16, Theorem 2]. It is to be noted that the bound in [16] is tight only for large values of device storage size M. As shown in the sequel, the new bound in Theorem 1 is tighter for smaller values of M and yields the existing bound as a special case for large values of M.

Next, we present our second main result, which establishes the optimal storage vs. rate tradeoff for the D2D network to within a constant multiplicative factor. In [16], the authors propose a D2D content distribution scheme which achieves a rate given by [16, Theorem 1]. Using the result in (7), we show a constant multiplicative gap exists between the new lower bound and the achievable rate in [16, Theorem 1]:

Theorem 2. *Let $R_{UB}(M)$ be the achievable rate of the D2D scheme given in [16, Theorem 1] and let $R_{LB}(M)$ be the lower bound on the optimal rate given in (7). For any $K \in \mathbb{N}^+$ devices, $N \in \mathbb{N}^+$ files, and device storage in the range $\frac{N}{K} \leq M \leq N$, we have:*

$$Gap = \frac{R_{UB}(M)}{R_{LB}(M)} \begin{cases} = 1 & M = \frac{N}{K} \\ \leq 3 & \frac{N}{K} < M \leq \frac{2}{3} \\ \leq 6 & \frac{2}{3} < M \leq 1 \\ \leq 8 & 1 \leq M \leq N \end{cases}. \quad (9)$$

Proof. The proof of Theorem 2 is omitted for brevity. $\square$

IV. INTUITION BEHIND PROOF OF THEOREM 1

We next present a detailed example to illustrate the intuition behind the proposed converse.

Example 1. Consider a D2D system with $K = 3$ devices, each with a storage memory of M, and $N = 3$ files A, B, C each of unit size. For this problem, the lower bounds [16, Theorem 2] are given by:

$$R^* + 3M \geq 3 \quad (10)$$
$$2R^* + M \geq 3. \quad (11)$$

In contrast, the proposed bound in Theorem 1 gives following tighter bounds for different choices of s and ℓ:

$$R^* + 6M \geq 8, \quad s = 2, \ \ell = 1 \quad (12)$$
$$8R^* + 6M \geq 15, \quad s = 1, \ \ell = 2 \quad (13)$$
$$2R^* + M \geq 3, \quad s = 1, \ \ell = 3, \quad (14)$$

where (14) is the bound in (11). Next, we detail the derivation of the new lower bounds (12) and (13). First, we consider the request vectors $(R_1, R_2, R_3) = (A, B, C)$ and $(R_1, R_2, R_3) = (B, C, A)$. The first $s = 2$ device storage contents Z_1, Z_2 along with two transmissions $\bar{X}_{ABC} = \{X^3_{ABC}\}, \bar{X}_{BCA} = \{X^3_{BCA}\}$ from the third device corresponding to the two request vectors are able to decode all 3 files. Here each transmission has the rate of $\frac{1}{3}R^*$. We upper bound the entropy of $\ell = 1$ transmission with this rate, while using the other transmission's decoding capability (in conjunction with the device storage contents) to derive a tighter bound. We have:

$$3 \leq H(Z_{[1,2]}, \bar{X}_{ABC}, \bar{X}_{BCA}) \quad (15)$$
$$\leq H(Z_{[1,2]}) + H(\bar{X}_{ABC}, \bar{X}_{BCA}|Z_{[1,2]}) \quad (16)$$
$$\leq 2M + H(\bar{X}_{ABC}) + H(\bar{X}_{BCA}|Z_{[1,2]}, \bar{X}_{ABC}) \quad (17)$$
$$\leq 2M + \frac{R^*}{3} + H(\bar{X}_{BCA}|Z_{[1,2]}, \bar{X}_{ABC}, A, B) \quad (18)$$
$$\leq 2M + \frac{R^*}{3} + H(\bar{X}_{BCA}, Z_3|Z_{[1,2]}, \bar{X}_{ABC}, A, B) \quad (19)$$
$$\leq 2M + \frac{R^*}{3} + H(Z_3|Z_{[1,2]}, \bar{X}_{ABC}, A, B)$$
$$\quad + H(\bar{X}_{BCA}|Z_{[1:3]}, \bar{X}_{ABC}, A, B) \quad (20)$$
$$\leq 2M + \frac{R^*}{3} + H(Z_3|Z_{[1,2]}, A, B)$$
$$\quad + H(\bar{X}_{BCA}|Z_{[1:3]}, \bar{X}_{ABC}, A, B, C) \quad (21)$$
$$\leq 2M + \frac{R^*}{3} + H(Z_3|Z_{[1,2]}, A, B), \quad (22)$$

where (22) follows from the fact that $H(\bar{X}_{BCA}|Z_{[1:3]}, X_{ABC}, A, B, C) = 0$ since X_{BCA} is a function of the files A, B, C which are present inside the conditioning. Considering the term $H(Z_3|Z_{[1,2]}, A, B)$, we have:

$$H(Z_3|Z_{[1,2]}, A, B) = H(Z_{[1:3]}|A, B) - H(Z_{[1,2]}|A, B). \quad (23)$$

Using (23) in (22), we have:

$$3 \leq 2M + \frac{R^*}{3} + H(Z_{[1:3]}|A, B) - H(Z_{[1,2]}|A, B). \quad (24)$$

Now considering all possible subsets of $Z_{[1:3]}$ having cardinality 2, in the RHS of (24), we have the following:

$$3 \leq 2M + \frac{R^*}{3} + H(Z_{[1:3]}|A, B) - H(Z_{[2,3]}|A, B). \quad (25)$$

$$3 \leq 2M + \frac{R^*}{3} + H(Z_{[1:3]}|A, B) - H(Z_{[1,3]}|A, B). \quad (26)$$

Symmetrizing over the inequalities in (24), (25) and (26), we have:

$$3 \leq 2M + \frac{R^*}{3} + H(Z_{[1:3]}|A, B) - \sum_{i,j=1, i \neq j}^{3} \frac{H(Z_{[i,j]}|A, B)}{3}. \quad (27)$$

Han's Inequality: We next state Han's Inequality [17, Theorem 17.6.1] on subsets of random variables, which we use throughout the paper for deriving the new lower bound. Let $\{X_1, X_2, \ldots, X_n\}$ denote a set of random variables. Further, let $\mathsf{X}_{[s]} \subseteq \{X_1, X_2, \ldots, X_n\}$ denote a subset of cardinality s. Then given two subsets $\mathsf{X}_{[r]}, \mathsf{X}_{[m]}$ where $r \geq m$, the statement of Han's Inequality can be stated as:

$$\frac{1}{\binom{n}{r}} \sum_{\mathsf{X}_{[r]}:|\mathsf{X}_{[r]}|=r} \frac{H(\mathsf{X}_{[r]})}{r} \leq \frac{1}{\binom{n}{m}} \sum_{\mathsf{X}_{[m]}:|\mathsf{X}_{[m]}|=m} \frac{H(\mathsf{X}_{[m]})}{m}, \quad (28)$$

where the sums are over all possible subsets of size r, m respectively.

Next, consider $Z_{[1:3]}$ as the set of random variables and the subset $\mathsf{Z}_{[i,j]} \subseteq Z_{[1:3]} : i \neq j, \forall i, j = 1, 2, 3$ of cardinality 2. Applying Han's Inequality using the conditional entropy of the sets, we have:

$$\frac{1}{\binom{3}{3}} \sum_{i=1}^{\binom{3}{3}} \frac{H(Z_{[1:3]}|A, B)}{3} \leq \frac{1}{\binom{3}{2}} \sum_{\substack{i,j=1 \\ i \neq j}}^{\binom{3}{2}} \frac{H(Z_{[i,j]}|A, B)}{2} \quad (29)$$

$$\Rightarrow \frac{2}{3} H(Z_{[1:3]}|A, B) \leq \frac{1}{3} \sum_{i,j=1, i \neq j}^{3} H(Z_{[i,j]}|A, B). \quad (30)$$

Substituting (30) into (24), we have:

$$3 \leq 2M + \frac{R^*}{3} + H(Z_{[1:3]}|A, B) - \frac{2H(Z_{[1:3]}|A, B)}{3} \quad (31)$$
$$\leq 2M + \frac{R^*}{3} + \frac{H(Z_{[1:3]}|A, B)}{3} \quad (32)$$
$$\leq 2M + \frac{R^*}{3} + \frac{H(Z_{[1:3]}, C|A, B)}{3} \quad (33)$$
$$\leq 2M + \frac{R^*}{3} + \frac{H(C|A, B) + H(Z_{[1:3]}|A, B, C)}{3} \quad (34)$$
$$\leq 2M + \frac{R^*}{3} + \frac{1}{3} \quad \Rightarrow R^* + 6M \geq 8. \quad (35)$$
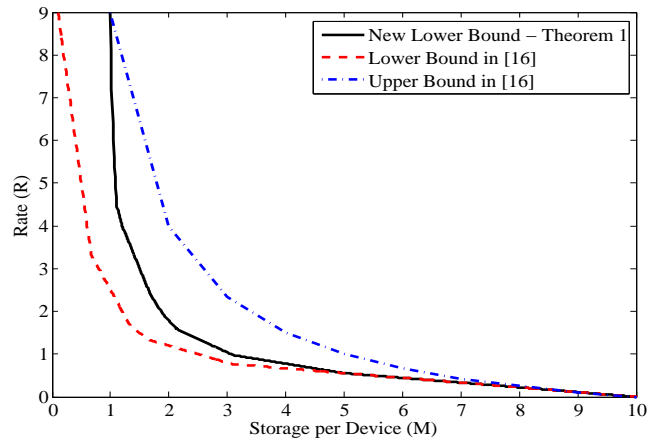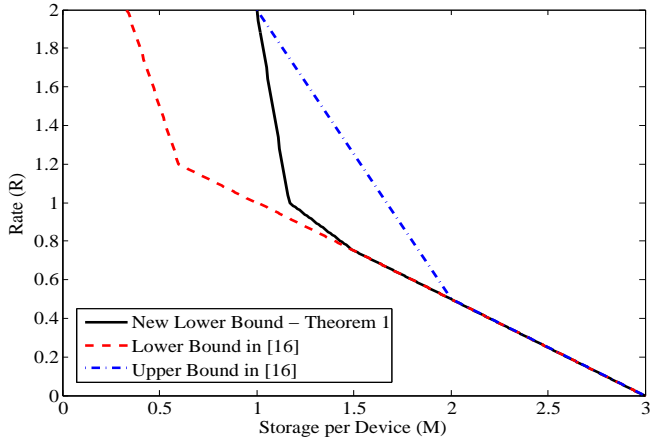
Fig. 2. M vs. R tradeoff for a D2D cluster with (a) $N = K = 3$ and (b) $N = K = 10$.

Next, we consider $s = 1$ device storage content, Z_1, and three request vectors $(R_1, R_2, R_3) = (A, B, C)$, $(R_1, R_2, R_3) = (B, C, A)$ and $(R_1, R_2, R_3) = (C, A, B)$ along with transmissions $\bar{X}_{ABC} = \{X_{ABC}^2, X_{ABC}^3\}, \bar{X}_{BCA} = \{X_{BCA}^2, X_{BCA}^3\}, \bar{X}_{CAB} = \{X_{CAB}^2, X_{CAB}^3\}$ which are capable of decoding all 3 files. Each composite transmission is of rate $\frac{2}{3}R^*$. We upper bound the entropy of $\ell = 2$ transmissions with their rate. Thus we have:

$$3 \leq H(Z_1, \bar{X}_{ABC}, \bar{X}_{BCA}, \bar{X}_{CAB}) \tag{36}$$

$$\leq H(Z_1) + H(\bar{X}_{ABC}, \bar{X}_{BCA}, \bar{X}_{CAB}|Z_1) \tag{37}$$

$$\leq M + H(\bar{X}_{ABC}, \bar{X}_{BCA}|Z_1)$$
$$\quad + H(\bar{X}_{CAB}|Z_1, \bar{X}_{ABC}, \bar{X}_{BCA}) \tag{38}$$

$$\leq M + 2 \cdot \frac{2R^*}{3} + H(\bar{X}_{CAB}, Z_2|Z_1, \bar{X}_{ABC}, \bar{X}_{BCA}) \tag{39}$$

$$\leq M + \frac{4R^*}{3} + H(Z_2|Z_1, \bar{X}_{ABC}, \bar{X}_{BCA}, A, B)$$
$$\quad + H(\bar{X}_{CAB}|Z_{[1,2]}, \bar{X}_{ABC}, \bar{X}_{BCA}, A, B, C) \tag{40}$$

$$\leq M + \frac{4R^*}{3} + H(Z_2|Z_1, A, B) \tag{41}$$

$$\leq M + \frac{4R^*}{3} + H(Z_{[1,2]}|A, B) - H(Z_1|A, B), \tag{42}$$

where, (40) follows from the fact that $H(\bar{X}_{CAB}|Z_{[1,2]}, \bar{X}_{ABC}, \bar{X}_{BCA}, A, B, C) = 0$ since $\bar{X}_{CAB}$ is a function of the files A, B, C. Considering all possible subsets of $Z_{[1,2]}$ of cardinality 1 and using a similar symmetrization argument and Han's Inequality, we get:

$$3 \leq M + \frac{4R^*}{3} + H(Z_{[1,2]}|A, B) - \frac{H(Z_{[1,2]}|A, B)}{2} \tag{43}$$

$$\leq M + \frac{4R^*}{3} + \frac{H(C|A, B) + H(Z_{[1,2]}|A, B, C)}{2} \tag{44}$$

$$\leq M + \frac{4R^*}{3} + \frac{1}{2} \tag{45}$$

$$\Rightarrow 8R^* + 6M \geq 15. \tag{46}$$

Finally, considering again, $s = 1$ device storage content, Z_1, and three request vectors $(R_1, R_2, R_3) = (A, B, C)$, $(R_1, R_2, R_3) = (B, C, A)$ and $(R_1, R_2, R_3) = (C, A, B)$ along with three transmissions $\bar{X}_{ABC} =$

$\{X_{ABC}^2, X_{ABC}^3\}, \bar{X}_{BCA} = \{X_{BCA}^2, X_{BCA}^3\}, \bar{X}_{CAB} = \{X_{CAB}^2, X_{CAB}^3\}$ which are capable of decoding all 3 files. Each transmission has rate $2R^*/3$. We upper bound the entropy of $\ell = 3$ transmissions by their rates. We have:

$$3 \leq H(Z_1, \bar{X}_{ABC}, \bar{X}_{BCA}, \bar{X}_{CAB}) \tag{47}$$

$$\leq H(Z_1) + H(\bar{X}_{ABC}) + H(\bar{X}_{BCA}) + H(\bar{X}_{CAB}) \tag{48}$$

$$\leq M + 3\frac{2}{3}R^* \tag{49}$$

$$\Rightarrow 2R^* + M \geq 3. \tag{50}$$

Finally, combining (12), (13) and (14), gives the proposed lower bound on the optimal transmission rate as shown in Figure 2(a). The last bound on the rate given in (50) is tight only in the case when the transmissions are almost independent of each other. The proposed converse is tight at the point $M = \frac{N}{K}$ where we show that the achievable point proposed in [16, Theorem 1] is optimal. For increasing N, K we can see from Figure 2(b) that the proposed converse significantly improves on the existing bounds from [16, Theorem 2].

V. CONCLUSION

In this paper, we presented a new information theoretic lower bound for a self-sustaining D2D network. We have leveraged Han's Inequality to better model the interaction of shared content stored in the device memories and file decoding capability of multicast transmissions to derive new lower bounds which are generally tighter than the existing cut-set based bounds. Using the new lower bound, we characterized the device storage vs. transmission rate tradeoff of the D2D network to within a maximum constant multiplicative factor of 8 for all possible values of problem parameters.

APPENDIX A
PROOF OF THEOREM 1

Let there be a library of $N \in \mathbb{N}^+$ files $F_{[1:N]}$ each of unit size (without loss of generality), and $K \in \mathbb{N}^+$ devices in the D2D cluster, with storage contents $Z_{[1:K]}$. The self sustainability condition of the D2D cluster, we have $KM \geq N$. Let s be an integer such that $s \in \{1, 2, \ldots, K\}$. Consider

the first s device storage contents, $Z_{[1:s]}$, and a request vector $(R_1, R_2, \ldots, R_s, R_{s+1}, \ldots, R_K) = (1, 2, \ldots, s, \phi, \ldots, \phi)$ i.e., the first s requests are unique files while the remaining $K - s$ can be for arbitrary files. To service this request consider the composite transmission vector:

$$\bar{X}_1 = \begin{bmatrix} X^{s+1}_{(1,2,\ldots,s,\phi,\ldots,\phi)} \\ X^{s+2}_{(1,2,\ldots,s,\phi,\ldots,\phi)} \\ \vdots \\ X^{s+K}_{(1,2,\ldots,s,\phi,\ldots,\phi)} \end{bmatrix}, \tag{51}$$

where $X^{s+1}_{(1,2,\ldots,s,\phi,\ldots,\phi)} = \psi(F_{[1:s]}, \mathsf{F}_{[K-s]})$ denotes the multicast transmission from the $(s+1)$−th device to service the request of the s devices. Note that the files in the set $\mathsf{F}_{[K-s]}$, denoting the requests of the other $K - s$ devices, are considered arbitrary. For achieving optimal transmission rate $R^*(M)$, each device in the D2D cluster multicasts with a rate of $R^*(M)/K$. Thus the rate of each composite transmission vector is given by $\frac{K-s}{K}R^*(M)$. Next, we note that the composite transmission vector $\bar{X}_1$, composed of $(K - s)$ sub-multicast transmissions, along with the contents of s device memories is capable of decoding s files i.e., $F_{[1:s]}$. Similarly, consider a request vector $(R_1, R_2, \ldots, R_s, R_{s+1}, \ldots, R_K) = (s+1, s+2, \ldots, 2s, \phi, \ldots, \phi)$ where the first s requests are again for unique files and rest are arbitrary. A composite transmission vector

$$\bar{X}_2 = \begin{bmatrix} X^{s+1}_{(s+1,s+2,\ldots,2s,\phi,\ldots,\phi)} \\ X^{s+2}_{(s+1,s+2,\ldots,2s,\phi,\ldots,\phi)} \\ \vdots \\ X^{s+K}_{(s+1,s+2,\ldots,2s,\phi,\ldots,\phi)} \end{bmatrix}, \tag{52}$$

along with device storage contents $Z_{[1:s]}$, can decode another s files $F_{[s+1:2s]}$. Considering $\lceil N/s \rceil$ composite transmission vectors $\bar{X}_{[1:\lceil N/s \rceil]}$ and s device storage contents $Z_{[1:s]}$, the library of files $F_{[1:N]}$ can be completely decoded. Thus, we have:

$$N \leq H\left(Z_{[1:s]}, \bar{X}_{[1:\lfloor N/s \rfloor]}\right) \tag{53}$$

$$\leq H\left(Z_{[1:s]}\right) + H\left(\bar{X}_{[1:\lfloor N/s \rfloor]}|Z_{[1:s]}\right) \tag{54}$$

$$\leq sM + H\left(\bar{X}_{[1:\lfloor N/s \rfloor]}|Z_{[1:s]}\right) \tag{55}$$

$$\leq sM + H\left(\bar{X}_{[1:\ell]}|Z_{[1:s]}\right) + H\left(\bar{X}_{[\ell+1:\lceil N/s \rceil]}|Z_{[1:s]}, \bar{X}_{[1:\ell]}\right) \tag{56}$$

$$\leq sM + \frac{\ell(K-s)R^*(M)}{K} + H\left(\bar{X}_{[\ell+1:\lceil N/s \rceil]}|Z_{[1:s]}, \bar{X}_{[1:\ell]}, F_{[1:\ell s]}\right) \tag{57}$$

$$\leq sM + \frac{\ell(K-s)R^*(M)}{K} + H\left(\bar{X}_{[\ell+1:\lceil N/s \rceil]}, Z_{[s+1:s+\mu]}|Z_{[1:s]}, \bar{X}_{[1:\ell]}, F_{[1:\ell s]}\right) \tag{58}$$

$$\leq sM + \frac{\ell(K-s)R^*(M)}{K} + \underbrace{H\left(Z_{[s+1:s+\mu]}|Z_{[1:s]}, \bar{X}_{[1:\ell]}, F_{[1:\ell s]}\right)}_{\triangleq \delta} + \underbrace{H\left(\bar{X}_{[\ell+1:\lceil N/s \rceil]}|Z_{[1:s+\mu]}, \bar{X}_{[1:\ell]}, F_{[1:\ell\mu+\ell s]}\right)}_{\triangleq \lambda}, \tag{59}$$

where (57) follows from the fact that each composite transmission vector has rate $\frac{K-s}{K}R^*(M)$ and that the device storage contents, $Z_{[1:s]}$, along with the composite transmission vectors $\bar{X}_{[1:\ell]}$ are capable of decoding the files $F_{[1:\ell s]}$. In (58), $\mu = \min\left\{\left\lceil \frac{N-\ell s}{\ell}\right\rceil, K - s\right\}$ is the number of additional device memories which, along with the transmissions $\bar{X}_{[1:\ell]}$ can decode all N files. For $s = K$, we have:

$$H\left(\bar{X}_{[1:\lceil N/s \rceil]}|Z_{[1:s]}\right) = 0, \tag{60}$$

and (55) yields the self-sustainability condition $KM \geq N$.

Upper Bound on δ: We consider the factor δ, from (59):

$$\delta = H\left(Z_{[s+1:s+\mu]}|Z_{[1:s]}, \bar{X}_{[1:\ell]}, F_{[1:\ell s]}\right) \tag{61}$$

$$\leq H\left(Z_{[s+1:s+\mu]}|Z_{[1:s]}, F_{[1:\ell s]}\right) \tag{62}$$

$$= H\left(Z_{[1:s+\mu]}|F_{[1:\ell s]}\right) - H\left(Z_{[1:s]}|F_{[1:\ell s]}\right). \tag{63}$$

Considering all possible subsets of $Z_{[1:s+\mu]}$ having cardinality s, i.e., all possible combination of s device storage contents in (53), and all possible combinations of files in the set $\mathsf{F}_{[K-s]}$ for each transmission $\bar{X}_{[1:\ell]}$ in (57), we can obtain $\binom{s+\mu}{s}$ different inequalities of the form of (63). Symmetrizing over all the inequalities, we have:

$$\delta \leq H\left(Z_{[1:s+\mu]}|F_{[1:\ell s]}\right) - \frac{1}{\binom{s+\mu}{s}} \sum_{i=1}^{\binom{s+\mu}{s}} H\left(\mathsf{Z}^i_{[s]}|F_{[1:\ell s]}\right), \tag{64}$$

where, $\mathsf{Z}^i_{[s]}$ is the i-th size-s subset of $Z_{[1:s+\mu]}$. Next, consider $Z_{[1:s+\mu]}$ as a set of random variables and subsets $\mathsf{Z}^i_{[s]} \subseteq Z_{[1:s+\mu]}\ \forall i = 1, \ldots, \binom{s+\mu}{s}$. Applying Han's Inequality from (28), we have:

$$\frac{1}{\binom{s+\mu}{s+\mu}} \sum_{i=1}^{\binom{s+\mu}{s+\mu}} \frac{H\left(Z_{[1:s+\mu]}|F_{[1:\ell s]}\right)}{s+\mu}$$

$$\leq \frac{1}{\binom{s+\mu}{s}} \sum_{i=1}^{\binom{s+\mu}{s}} \frac{H\left(\mathsf{Z}^i_{[s]}|F_{[1:\ell s]}\right)}{s} \tag{65}$$

$$\Rightarrow \frac{s}{s+\mu}H\left(Z_{[1:s+\mu]}|F_{[1:\ell s]}\right) \leq \frac{1}{\binom{s+\mu}{s}} \sum_{i=1}^{s+\mu} H\left(\mathsf{Z}^i_{[s]}|F_{[1:\ell s]}\right). \tag{66}$$

Substituting (66) into (64), we have:

$$\delta \leq H\left(Z_{[1:s+\mu]}|F_{[1:\ell s]}\right) - \frac{s}{s+\mu}H\left(Z_{[1:s+\mu]}|F_{[1:\ell s]}\right) \tag{67}$$

$$= \frac{\mu}{s+\mu}H\left(Z_{[1:s+\mu]}|F_{[1:\ell s]}\right) \tag{68}$$

$$\leq \frac{\mu}{s+\mu}H\left(Z_{[1:s+\mu]}, F_{[\ell s+1:N]}|F_{[1:\ell s]}\right) \tag{69}$$

$$= \frac{\mu}{s+\mu}H\left(F_{[\ell s+1:N]}|F_{[1:\ell s]}\right) + \underbrace{H\left(Z_{[1:s+\mu]}|F_{[1:N]}\right)}_{=0} \tag{70}$$

$$\leq \frac{\mu}{s+\mu}(N - \ell s)^+, \tag{71}$$

where (70) follows from the fact that the device storage contents are functions of all the N files in the library.

Upper Bound on λ: We consider the factor λ, from (59). We have two cases:

Case 1: $N \leq \ell s + \ell\mu$: We consider the case that all N files

can be decoded with $\mu \leq K - s$ additional device storage contents and transmissions $\bar{X}_{[1:\ell]}$, within the conditioning in the factor λ in (59)i.e., $\ell s + \ell\mu \leq N$. We have:

$$\begin{aligned}
\lambda &= H\left(\bar{X}_{[\ell+1:\lceil N/s\rceil]}|Z_{[1:s+\mu]}, \bar{X}_{[1:\ell]}, F_{[1:N]}\right) \\
&\leq H\left(\bar{X}_{[\ell+1:\lceil N/s\rceil]}|Z_{[1:s+\mu]}, F_{[1:N]}\right) \\
&= 0,
\end{aligned} \tag{72}$$

which follows from the fact that the transmissions are functions of all N files.

Case 2: $N > \ell s + \ell\mu$**:** We consider the case where $\mu = K - s$ additional device storage contents along with the transmissions $\bar{X}_{[1:\ell]}$, cannot decode all N files. We have:

$$\begin{aligned}
\lambda &= H\left(\bar{X}_{[\ell+1:\lceil N/s\rceil]}|Z_{[1:K]}, \bar{X}_{[1:\ell]}, F_{[1:K\ell]}\right) \\
&\leq H\left(\bar{X}_{[\ell+1:\lceil N/s\rceil]}|Z_{[1:K]}\right) = 0,
\end{aligned} \tag{73}$$

which follows from the fact that $KM \geq N$ i.e., all files are stored within the collective device memories in the D2D cluster and hence all transmissions are functions of all the device storage contents. Thus combining (72) and (73) we have:

$$\lambda = 0. \tag{74}$$

Substituting (71) and (74) into (59), we have:

$$N \leq sM + \frac{\ell(K-s)}{K} \cdot R^*(M) + \frac{\mu}{s+\mu}(N-\ell s)^+$$

$$\Rightarrow R^*(M) \geq \left\{ \frac{N - sM - \left(\frac{\mu}{s+\mu}\right)(N-\ell s)^+}{\ell\left(\frac{K-s}{K}\right)} \right\}. \tag{75}$$

Optimizing over all parameter values of s, ℓ, μ, we have:

$$R^*(M) \geq R_{LB}(M) \triangleq$$

$$\max_{\substack{s \in \{1,\dots,K\}, \\ \ell \in \{1,\dots,\lceil\frac{N}{s}\rceil\}}} \left\{ \frac{N - sM - \left(\frac{\mu}{s+\mu}\right)(N-\ell s)^+}{\ell\left(\frac{K-s}{K}\right)} \right\}, \tag{76}$$

which completes the proof of Theorem 1. $\quad\square$

REFERENCES

[1] M. A. Maddah-Ali and U. Niesen, "Fundamental Limits of Caching," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2856–2867, May 2014.

[2] M. Maddah-Ali and U. Niesen, "Decentralized Coded Caching Attains Order-Optimal Memory-Rate Tradeoff," *IEEE/ACM Transactions on Networking*, April 2014.

[3] U. Niesen and M. A. Maddah-Ali, "Coded Caching with Nonuniform Demands," in *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, April 2014, pp. 221–226.

[4] R. Pedarsani, M. Maddah-Ali, and U. Niesen, "Online Coded Caching," in *IEEE International Conference on Communications (ICC)*, June 2014, pp. 1878–1883.

[5] A. Sengupta, R. Tandon, and T. C. Clancy, "Fundamental Limits of Caching with Secure Delivery," *Wireless Physical Layer Security Workshop - IEEE International Conference on Communications (ICC)*, pp. 771–776, June 2014.

[6] ——, "Decentralized Caching with Secure Delivery," *IEEE International Symposium on Information Theory (ISIT)*, pp. 41–45, July 2014.

[7] ——, "Secure Caching with Non-Uniform Demands," *IEEE Global Wireless Summit (GWS)*, pp. 1–5, May 2014.

[8] ——, "Fundamental Limits of Caching with Secure Delivery," *IEEE Transactions on Information Forensics and Security*, vol. 10, pp. 355–370, Feb 2015.

[9] A. Sengupta, S. Amuru, R. Tandon, R. M. Buehrer, and T. C. Clancy, "Learning Distributed Caching Strategies in Small Cell Networks," *The Eleventh International Symposium on Wireless Communication Systems (ISWCS)*, pp. 917–921, Aug 2014.

[10] M. Ji, G. Caire, and A. F. Molisch, "Optimal Throughput-Outage Tradeoff in Wireless One-Hop Caching Networks," in *IEEE International Symposium on Information Theory (ISIT)*, July 2013, pp. 1461–1465.

[11] M. Ji, A. M. Tulino, J. Llorca, and G. Caire, "Order Optimal Coded Caching-Aided Multicast under Zipf Demand Distributions," in *The Eleventh International Symposium on Wireless Communication Systems (ISWCS)*, 2014.

[12] N. Golrezaei, K. Shanmugam, A. Dimakis, A. Molisch, and G. Caire, "FemtoCaching: Wireless Video Content Delivery through Distributed Caching Helpers," *IEEE Transactions on Information Theory*, vol. 59(12), pp. 8402–8413, Dec. 2013.

[13] M. Ji, G. Caire, and A. F. Molisch, "Wireless Device-to-Device Caching Networks: Basic Principles and System Performance," *arXiv : 1305.5216*, May 2013.

[14] N. Golrezaei, A. G. Dimakis, and A. F. Molisch, "Wireless Device-to-Device Communications with Distributed Caching," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, July 2012, pp. 2781–2785.

[15] M. Ji, G. Caire, and A. F. Molisch, "Fundamental Limits of Distributed Caching in D2D Wireless Networks," in *IEEE Information Theory Workshop (ITW) 2013*, Sept 2013, pp. 1–5.

[16] ——, "Fundamental limits of caching in wireless D2D networks," *arXiv:1405.5336*, 2014. [Online]. Available: http://arxiv.org/abs/1405.5336

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd Edition*. Hoboken, NJ, USA: Wiley-Interscience, John Wiley and Sons.