Bounds on the size and asymptotic rate of subblock-constrained codes

Tandon, Anshoo; Kiah, Han Mao; Motani, Mehul

2018

Tandon, A., Kiah, H. M., & Motani, M. (2018). Bounds on the Size and Asymptotic Rate of Subblock-Constrained Codes. *IEEE Transactions on Information Theory*, 64(10), 6604–6619. doi:10.1109/tit.2018.2864137

<https://hdl.handle.net/10356/144707>

<https://doi.org/10.1109/TIT.2018.2864137>

© 2018 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at: <https://doi.org/10.1109/TIT.2018.2864137>.

Downloaded on 26 Apr 2023 21:47:58 SGT

Bounds on the Size and Asymptotic Rate of Subblock-Constrained Codes

Anshoo Tandon, *Member, IEEE*, Han Mao Kiah, and Mehul Motani, *Fellow, IEEE*

Abstract—The study of subblock-constrained codes has recently gained attention due to their application in diverse fields. We present bounds on the size and asymptotic rate for two classes of subblock-constrained codes. The first class is binary *constant subblock-composition codes* (CSCCs), where each codeword is partitioned into equal sized subblocks, and every subblock has the same fixed weight. The second class is binary *subblock energy-constrained codes* (SECCs), where the weight of every subblock exceeds a given threshold. We present novel upper and lower bounds on the code sizes and asymptotic rates for binary CSCCs and SECCs. For a fixed subblock length and small relative distance, we show that the asymptotic rate for CSCCs (resp. SECCs) is strictly lower than the corresponding rate for constant weight codes (CWCs) (resp. heavy weight codes (HWCs)). Further, for codes with high weight and low relative distance, we show that the asymptotic rate for CSCCs is strictly lower than that of SECCs, which contrasts with the fact that the asymptotic rate for CWCs is equal to that of HWCs. We also provide a correction to an earlier result by Chee et al. (2014) on the asymptotic CCCC rate. Additionally, we present several numerical examples comparing the rates for CSCCs and SECCs with those for constant weight codes and heavy weight codes.

Index Terms—Subblock-constrained codes, constant subblock-composition codes, subblock energy-constrained codes, optimal code size, asymptotic rate.

I. INTRODUCTION

The study of subblock-constrained codes has recently gained attention as they are suitable candidates for varied applications such as simultaneous energy and information transfer [3], powerline communications [4], and design of low-cost authentication methods [5]. A special class of subblock-constrained codes are codes where each codeword is partitioned into equal sized subblocks, and every subblock has the same fixed composition. Such codes were called *constant subblock-composition codes* (CSCCs) in [3], and were labeled as multiply constant-weight codes (MCWC) in [5].

Subblock energy-constrained codes (SECCs) were proposed in [3], [6] for providing real-time energy and information transfer from a powered transmitter to an energy harvesting receiver. For binary alphabet, SECCs are characterized by the

property that weight of every subblock exceeds a given threshold. The CCCC and SECC capacities, and computable bounds, were presented in [3] for discrete memoryless channels.

In this paper, we study bounds on the size and asymptotic rate for binary CSCCs and SECCs with given error correction capability, i.e., minimum distance of the code. By studying the asymptotic ball sizes of the respective subblock constrained codes, we explore the performance gap between these two classes of codes. Using standard Gilbert-Varshamov (GV) and sphere packing bounds, we carry out a careful analysis to find useful upper and lower bounds on the asymptotic ball sizes that lead to estimates of the performance gap.

A. Notation

The input alphabet is denoted by $\mathcal{X}$ which comprises q symbols. An n -length, q -ary code $\mathcal{C}$ over $\mathcal{X}$ is a subset of $\mathcal{X}^n$. The elements of $\mathcal{C}$ are called *codewords* and $\mathcal{C}$ is said to have *minimum distance* d if the smallest *Hamming distance* between any two distinct codewords is equal to d . A q -ary code of length n and minimum distance d is called an $(n, d)_q$ -code, and the largest size of an $(n, d)_q$ -code is denoted by $A_q(n, d)$. For binary alphabet ($q = 2$), an $(n, d)_2$ -code is just called an (n, d) -code, and the largest size for this code is simply denoted by $A(n, d)$.

A *constant weight code* (CWC) with parameter w is a binary code where each codeword has weight exactly w . We denote a CWC with weight parameter w , blocklength n , and minimum distance d by (n, d, w) -CWC, and denote its maximum possible size by $A(n, d, w)$. A *heavy weight code* (HWC) [7] with parameter w is a binary code where each codeword has weight *at least* w . We denote a HWC with weight parameter w , blocklength n , and minimum distance d by (n, d, w) -HWC, and denote its maximum possible size by $H(n, d, w)$. Since an (n, d, w) -CWC is an (n, d, w) -HWC, we have that $A(n, d, w) \leq H(n, d, w)$.

A subblock-constrained code is a code where each codeword is divided into subblocks of equal length, and each subblock satisfies a fixed set of constraints. For a subblock-constrained code, we denote the codeword length by n , the subblock length by L , and the number of subblocks in a codeword by m . For the binary alphabet $\mathcal{X} = \{0, 1\}$, a CCCC is characterized by the property that each subblock in every codeword has the same *weight*, i.e. each subblock has the same number of ones. A binary CCCC with codeword length $n = mL$, subblock length L , minimum distance d , and weight w_s per subblock is called an (m, L, d, w_s) -CCCC. We denote the maximum possible size of (m, L, d, w_s) -CCCC by $C(m, L, d, w_s)$. Since an (m, L, d, w_s) -CCCC is an (mL, d, mw_s) -CWC, we have that $C(m, L, d, w_s) \leq A(mL, d, mw_s)$.

This paper was presented in part at the 2017 IEEE International Symposium on Information Theory [1], [2].

A. Tandon and M. Motani are with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117583 (email: anshoo.tandon@gmail.com, motani@nus.edu.sg).

H. M. Kiah is with the School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371 (email: hmkiah@ntu.edu.sg). H. M. Kiah was also supported in part by the Singapore Ministry of Education under Research Grants MOE2016-T1-001-156 and MOE2015-T2-2-086.

Copyright (c) 2018 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

For providing regular energy content in a codeword for the application of simultaneous energy and information transfer from a powered transmitter to an energy harvesting receiver, the use of CSCCs was proposed in [3]. When *on-off keying* is employed, with bit-1 (bit-0) represented by the presence (absence) of a high energy signal, regular energy content in a CCCC codeword can be ensured by appropriately choosing the weight w_s per subblock. A natural extension of binary CSCCs are binary SECCs, which allow the weight of each subblock to exceed w_s , thereby ensuring that the energy content within every subblock duration is sufficient [3]. A binary SECC with codeword length $n = mL$, subblock length L , minimum distance d , and weight at least w_s per subblock is called an (m, L, d, w_s) -SECC. We denote the maximum possible size of an (m, L, d, w_s) -SECC by $S(m, L, d, w_s)$. Since an (m, L, d, w_s) -SECC is an (mL, d, mw_s) -HWC, we have that $S(m, L, d, w_s) \leq H(mL, d, mw_s)$. Also, since an (m, L, d, w_s) -CSCC is an (m, L, d, w_s) -SECC, we have that $C(m, L, d, w_s) \leq S(m, L, d, w_s)$.

The relation among code sizes is summarized below. For all m , and $1 \leq d, w_s \leq L$, we have

$$\begin{aligned} S(m, L, d, w_s) &\geq C(m, L, d, w_s) \\ \wedge \quad H(mL, d, mw_s) &\geq A(mL, d, mw_s) \end{aligned} \quad (1)$$

We also analyze bounds on the rate in the asymptotic setting where the number of subblocks m tends to infinity, d scales linearly with m , but L and w_s are fixed. In the following, the base for log is assumed to be 2. Formally, for fixed $0 \leq \delta \leq 1$, the asymptotic rates for CSCCs and SECCs with fixed subblock length L , subblock weight parameter w_s , number of subblocks in a codeword $m \rightarrow \infty$, and minimum distance d scaling as $d = \lfloor mL\delta \rfloor$ are defined as

$$\gamma(L, \delta, w_s/L) \triangleq \limsup_{m \rightarrow \infty} \frac{\log C(m, L, \lfloor mL\delta \rfloor, w_s)}{mL}, \quad (2)$$

$$\sigma(L, \delta, w_s/L) \triangleq \limsup_{m \rightarrow \infty} \frac{\log S(m, L, \lfloor mL\delta \rfloor, w_s)}{mL}. \quad (3)$$

These rate can be compared with related exponents:

$$\alpha(\delta) \triangleq \limsup_{n \rightarrow \infty} \frac{\log A(n, \lfloor n\delta \rfloor)}{n}, \quad (4)$$

$$\alpha(\delta, w_s/L) \triangleq \limsup_{n \rightarrow \infty} \frac{\log A(n, \lfloor n\delta \rfloor, \lfloor nw_s/L \rfloor)}{n}, \quad (5)$$

$$\eta(\delta, w_s/L) \triangleq \limsup_{n \rightarrow \infty} \frac{\log H(n, \lfloor n\delta \rfloor, \lfloor nw_s/L \rfloor)}{n}. \quad (6)$$

The relation between asymptotic rates can be obtained by using the relation among code sizes in (1), and the above rate definitions. For all $1 \leq w_s \leq L$ and $0 \leq \delta \leq 1$, we have

$$\begin{aligned} \sigma(L, \delta, w_s/L) &\stackrel{(b)}{\geq} \gamma(L, \delta, w_s/L) \\ (c) \wedge & \quad (d) \wedge \\ \eta(\delta, w_s/L) &\stackrel{(a)}{\geq} \alpha(\delta, w_s/L) \end{aligned} \quad (7)$$

Now, observe that when the subblock length L is fixed and we let the number of blocks m go to infinity, the rates of the space of all CCCC and SECC words approach $\frac{1}{L} \log \binom{L}{w_s}$ and $\frac{1}{L} \log \sum_{i=w_s}^L \binom{L}{i}$, respectively. In contrast,

when we assume $w_s \geq L/2$, the rates of the space of all constant weight and heavy weight words are the same and are equal to $h(w_s/L)$ when word length approach infinity. Since $\frac{1}{L} \log \binom{L}{w_s} < \frac{1}{L} \log \sum_{i=w_s}^L \binom{L}{i} < h(w_s/L)$ for $w_s \geq L/2$, this implies that inequalities (b), (c), and (d) are strict for sufficiently small δ . We state and prove this formally in Section IV.

We summarize our notation on code size and asymptotic rates for CWCs, HWCs, CSCCs and SECCs in Table I. Further, we use the notation $[z]^+$ to denote $\max(z, 0)$. Note that wherever possible, the notation for the optimal asymptotic rate for a given constrained code is chosen as the Greek counterpart of the corresponding Latin letter denoting the optimal code size. For example, for SECCs, the optimal code size is denoted $S(m, L, d, w_s)$ whereas the optimal asymptotic rate is denoted $\sigma(L, \delta, w_s/L)$.

B. Previous Work

Among the codes discussed above, although CWCs have been widely studied, the exact characterization of $\alpha(\delta, \omega)$, for $0 < \omega < 1$, has remained elusive. A good upper bound for $\alpha(\delta, \omega)$ was given in [8], by using a linear programming bound for the CWC code size. When $\omega = 0.5$, it is known that $\alpha(\delta, 1/2) = \alpha(\delta)$. The class of HWCs was introduced by Cohen *et al.* [7], motivated by certain asynchronous communication problems. The asymptotic rates for HWCs was later established by Bachoc *et al.* [9].

Theorem 1 (Bachoc *et al.* [9]). *Let $0 \leq \delta, \omega \leq 1$. Then*

$$\eta(\delta, \omega) = \begin{cases} \alpha(\delta), & \text{when } 0 \leq \omega \leq 1/2, \\ \alpha(\delta, \omega), & \text{when } 1/2 \leq \omega \leq 1. \end{cases} \quad (8)$$

In view of the above theorem, the inequality (a) in (7) is in fact an equality for $L/2 \leq w_s \leq L$.

Chee *et al.* [5] introduced the class of CSCCs and provided rudimentary bounds for $C(m, L, d, w_s)$. Later, constructions of CSCCs were proposed by various authors [10], [11]. The asymptotic rate for CSCCs was also studied in [5]. However, an inconsistent asymptotic rate definition in [5] led to an erroneous claim regarding the CCCC rate (see [5, Prop. 6.1]). In this paper, we also provide a correct statement for the CCCC rate in the scenario where the subblock length tends to infinity via Proposition 9 in Section III.

SECCs were proposed in [6], owing to their natural application in real-time simultaneous energy and information transfer. As shown in Section II-B, the SECC space, comprising words where each subblock has weight exceeding a given threshold, has an interesting property that different balls of same radius may have different sizes. The lower bound on the code size for such spaces, where balls of same radius may have different sizes, was studied in [12], where a *generalized Gilbert-Varshamov bound* was presented. The *generalized sphere-packing bound*, providing an upper bound on the code size in such spaces, has been recently presented in [13], [14], using graph-based techniques.

Notation	Description
L	Subblock length
m	Number of subblocks in a codeword
n	Codeword length ($n = mL$)
d	Minimum distance of the code
δ	Relative distance of the code ($\delta = d/n$)
w_s	Weight per subblock
w	Weight per codeword ($w = mw_s$)
ω	Fraction of ones in a codeword ($\omega = w/n$)
$h(\cdot)$	Binary entropy function
$(n, d)_q$ -code	General q -ary code
$A_q(n, d)$	Maximum size of $(n, d)_q$ -code
(n, d) -code	General <i>binary</i> code
$A(n, d)$	Maximum size of (n, d) -code
$\alpha(\delta)$	Asymptotic rate of (n, d) -code
(n, d, w) -CWC	Constant weight code (each codeword has weight w)
$A(n, d, w)$	Maximum size of (n, d, w) -CWC
$\alpha(\delta, \omega)$	Asymptotic rate of (n, d, w) -CWC
$\delta^*(\omega)$	Minimum δ for which $\alpha(\delta, \omega) = 0$. It is known that $\delta^*(\omega) = 2\omega(1 - \omega)$
$\alpha_{GV}(\delta, \omega)$	Lower bound on $\alpha(\delta, \omega)$ using Gilbert Varshamov bound for (n, d, w) -CWC
$\alpha_{SP}(\delta, \omega)$	Upper bound on $\alpha(\delta, \omega)$ using sphere packing bound for (n, d, w) -CWC
(n, d, w) -HWC	Heavy weight code (each codeword has weight <i>at least</i> w)
$H(n, d, w)$	Maximum size of (n, d, w) -HWC
$\eta(\delta, \omega)$	Asymptotic rate of (n, d, w) -HWC
(m, L, d, w_s) -CSCC	Binary constant subblock-composition code (each subblock has weight w_s)
$C(m, L, d, w_s)$	Maximum size of (m, L, d, w_s) -CSCC
$\mathcal{C}(m, L, w_s)$	Space of all CSCC words
$\gamma(L, \delta, w_s/L)$	Asymptotic rate of (m, L, d, w_s) -CSCC
$\delta_\gamma^*(L, w_s/L)$	Minimum δ for which $\gamma(L, \delta, w_s/L) = 0$
$\gamma_{GV}(L, \delta, w_s/L)$	Lower bound on $\gamma(L, \delta, w_s/L)$ using Gilbert Varshamov bound for (m, L, d, w_s) -CSCC
$\gamma_{SP}(L, \delta, w_s/L)$	Upper bound on $\gamma(L, \delta, w_s/L)$ using sphere-packing bound for (m, L, d, w_s) -CSCC
(m, L, d, w_s) -SECC	Binary subblock energy-constrained code (each subblock has weight <i>at least</i> w_s)
$S(m, L, d, w_s)$	Maximum size of (m, L, d, w_s) -SECC
$\mathcal{S}(m, L, w_s)$	Space of all SECC words
$\sigma(L, \delta, w_s/L)$	Asymptotic rate of (m, L, d, w_s) -SECC
$\delta_\sigma^*(L, w_s/L)$	Minimum δ for which $\sigma(L, \delta, w_s/L) = 0$
$\sigma_{GV}(L, \delta, w_s/L)$	Lower bound on $\sigma(L, \delta, w_s/L)$ using Gilbert Varshamov bound for (m, L, d, w_s) -SECC
$\sigma_{SP}(L, \delta, w_s/L)$	Upper bound on $\sigma(L, \delta, w_s/L)$ using sphere-packing bound for (m, L, d, w_s) -SECC
$G_{\alpha-\gamma}(L, \delta, w_s/L)$	Asymptotic rate gap between CWC and CSCC, $\alpha(\delta, w_s/L) - \gamma(L, \delta, w_s/L)$
$G_{\alpha-\gamma}^{LB}(L, \delta, w_s/L)$	Lower bound on the asymptotic rate gap between CWC and CSCC
$G_{\eta-\sigma}(L, \delta, w_s/L)$	Asymptotic rate gap between HWC and SECC, $\eta(\delta, w_s/L) - \sigma(L, \delta, w_s/L)$
$G_{\eta-\sigma}^{LB}(L, \delta, w_s/L)$	Lower bound on the asymptotic rate gap between HWC and SECC
$G_{\sigma-\gamma}(L, \delta, w_s/L)$	Asymptotic rate gap between SECC and CSCC, $\sigma(\delta, w_s/L) - \gamma(L, \delta, w_s/L)$
$G_{\sigma-\gamma}^{LB}(L, \delta, w_s/L)$	Lower bound on the asymptotic rate gap between SECC and CSCC

TABLE I: Table of Notation

C. Our Contributions

The contributions of this paper are as follows:

- 1) By studying the space of CSCC and SECC codewords, we compute both upper and lower bounds for the optimal CSCC code size $C(m, L, d, w_s)$ and the optimal SECC code size $S(m, L, d, w_s)$ in Section II. Interestingly, using these bounds, we are able to compute new *exact* values of $S(1, L, d, w) = H(n, d, w)$ (see (26) and (27)).
- 2) We analyze the limiting behavior of ball sizes for these spaces in high dimensions, to derive both upper and lower bounds on the asymptotic rates for CSCC and SECC in Section III.
- 3) For fixed L and w_s , we estimate the value of δ_L such that inequalities (b), (c), and (d) in (7) are *strict* for all $\delta < \delta_L$ (see Section IV). This implies that: (i) Relative to codeword-based constraints for CWCs (resp. HWCs), the stricter subblock-based weight constraints for CSCCs (resp. SECCs), leads to a rate penalty. (ii) SECCs provide higher rates than CSCCs due to greater flexibility in choosing bits within each subblock (in contrast to Theorem 1).

- 4) For fixed values of δ , we quantify the rate penalty due to subblock-based constraints in Section V, by numerically evaluating the corresponding rate bounds.
- 5) We also provide a correction to a result by Chee *et al.* [5], on the asymptotic CSCC rate in the scenario where the subblock length tends to infinity (see Proposition 9 in Section III).

II. BOUNDS ON OPTIMAL CODE SIZE

We derive novel bounds for $C(m, L, d, w_s)$ and $S(m, L, d, w_s)$. While bounds for the former were also discussed in [5], those results are insufficient to provide good bounds on the asymptotic rates $\gamma(L, \delta, w_s/L)$. Among other bounds, we derive the Gilbert-Varshamov (GV) bound and the sphere-packing bound for both CSCCs and SECCs in this section, and their asymptotic versions in Section III.

A. CCCC code size

For an (m, L, d, w_s) -CCCC, it is easy to see from symmetry, via complementing bits in codewords, that we have the relation

$$C(m, L, d, w_s) = C(m, L, d, L - w_s).$$

Let $\mathcal{C}(m, L, w_s)$ denote the space of all binary words comprising m subblocks, each subblock having length L , with weight w_s per subblock. For $\mathbf{x} \in \mathcal{C}(m, L, w_s)$, we define a ball centered at $\mathbf{x}$ and having radius t as

$$\mathcal{B}_C(\mathbf{x}, t; m, L, w_s) \triangleq \{\mathbf{y} \in \mathcal{C}(m, L, w_s) : d(\mathbf{x}, \mathbf{y}) \leq t\}$$

The following lemma shows that the size of the CCCC ball, $|\mathcal{B}_C(\mathbf{x}, t; m, L, w_s)|$, is independent of choice of $\mathbf{x} \in \mathcal{C}(m, L, w_s)$. We will see later in Sec. II-B that this is not true for the space of SECC words.

Lemma 1. *If $\mathbf{x}$ and $\tilde{\mathbf{x}}$ are two words in $\mathcal{C}(m, L, w_s)$, then $|\mathcal{B}_C(\tilde{\mathbf{x}}, t; m, L, w_s)| = |\mathcal{B}_C(\mathbf{x}, t; m, L, w_s)|$.*

Proof: See Appendix A. ■

In view of the above lemma, the size of CCCC ball is independent of the center word. The following lemma quantifies the size of CCCC ball of radius t .

Lemma 2. *If $v \triangleq \min\{w_s, L - w_s\}$, then for any $\mathbf{x}$ in $\mathcal{C}(m, L, w_s)$, we have*

$$|\mathcal{B}_C(\tilde{\mathbf{x}}, t; m, L, w_s)| = \sum_{\substack{2(u_1+u_2+\dots+u_m) \leq t, \\ 0 \leq u_i \leq v}} \prod_{i=1}^m \binom{w_s}{u_i} \binom{L-w_s}{u_i}. \quad (9)$$

Proof: See Appendix B. ■

The following GV bound for $C(m, L, d, w_s)$ is obtained using the above lemmas.

Proposition 1. *If $v \triangleq \min\{w_s, L - w_s\}$, then*

$$C(m, L, d, w_s) \geq \frac{\binom{L}{w_s}^m}{\sum_{\substack{2(u_1+u_2+\dots+u_m) \leq d-1, \\ 0 \leq u_i \leq v}} \prod_{i=1}^m \binom{w_s}{u_i} \binom{L-w_s}{u_i}}. \quad (10)$$

Proof: Using standard Gilbert construction in the space $\mathcal{C}(m, L, w_s)$, we have the lower bound

$$C(m, L, d, w_s) \geq \frac{|\mathcal{C}(m, L, w_s)|}{|\mathcal{B}_C(\mathbf{x}, d-1; m, L, w_s)|}, \quad (11)$$

where $\mathbf{x}$ is any word in $\mathcal{C}(m, L, w_s)$, and $|\mathcal{C}(m, L, w_s)| = \binom{L}{w_s}^m$. From Lemma 1, we note that $|\mathcal{B}_C(\mathbf{x}, d-1; m, L, w_s)|$ is independent of the choice of $\mathbf{x}$, and using Lemma 2 we observe that this ball size of radius $d-1$ is given by the denominator in (10). ■

The following proposition provides the sphere-packing bound for CCCCCs.

Proposition 2. *If $v \triangleq \min\{w_s, L - w_s\}$ and $t \triangleq \lfloor (d-1)/2 \rfloor$, then*

$$C(m, L, d, w_s) \leq \frac{\binom{L}{w_s}^m}{\sum_{\substack{2(u_1+u_2+\dots+u_m) \leq t, \\ 0 \leq u_i \leq v}} \prod_{i=1}^m \binom{w_s}{u_i} \binom{L-w_s}{u_i}}. \quad (12)$$

Proof: The claim follows from the standard sphere-packing argument that for any (m, L, d, w_s) -CCCC, the balls of radius $t = \lfloor (d-1)/2 \rfloor$ around codewords should be non-intersecting, and the fact that the denominator in (12) is equal to $|\mathcal{B}_C(\mathbf{x}, t; m, L, w_s)|$. ■

To obtain Proposition 2, the sphere-packing argument considered balls centered on the codewords in $\mathcal{C}(m, L, w_s)$ in the same space. As pointed out by a reviewer, we may follow Freiman's and Berger's methods [15], [16] and center the balls on codewords in $\mathcal{C}(m, L, w_s)$ but consider words in a different space, say $\mathcal{C}(m, L, w_s - j)$. Formally, we define $\hat{\mathcal{B}}_C(\mathbf{x}, t; m, L, w_s, j) \triangleq \{\mathbf{y} \in \mathcal{C}(m, L, w_s - j) : d(\mathbf{x}, \mathbf{y}) \leq t\}$ for $w_s - L \leq j \leq w_s$ and following a similar argument to the proof of Prop. 2, we may obtain the following upper bound. Set $t \triangleq \lfloor (d-1)/2 \rfloor$. Then

$$C(m, L, d, w_s) \leq \min_{w_s - L \leq j \leq w_s} \frac{\binom{L}{w_s - j}^m}{|\hat{\mathcal{B}}_C(\mathbf{x}, t; m, L, w_s, j)|} \quad (13)$$

where the ball size $|\hat{\mathcal{B}}_C(\mathbf{x}, t; m, L, w_s, j)|$ is given by

$$\sum_{\substack{2(u_1+u_2+\dots+u_m) - mj \leq t, \\ [j]^+ \leq u_i \leq \min\{w_s, L - w_s + j\}}} \prod_{i=1}^m \binom{w_s}{u_i} \binom{L - w_s}{u_i - j}.$$

For certain values of j , the set $\hat{\mathcal{B}}_C(\mathbf{x}, t; m, L, w_s, j)$ may be empty, and for these cases we ignore the corresponding j values in the minimization in (13). Observe that evaluating the objective value of (13) at $j = 0$ yields (12). In other words, (13) is a possibly tighter upper bound. However, it is unclear if (13) yields a stronger upper bound on the *asymptotic* rate.

B. SECC code size

1) *Lower bounds on SECC code size:* Let $\mathcal{S}(m, L, w_s)$ denote the space of all binary words comprising m subblocks, each subblock having length L , with weight per subblock at least w_s . For $\mathbf{x} \in \mathcal{S}(m, L, w_s)$, we define a ball centered at $\mathbf{x}$ and having radius t as

$$\mathcal{B}_S(\mathbf{x}, t; m, L, w_s) \triangleq \{\mathbf{y} \in \mathcal{S}(m, L, w_s) : d(\mathbf{x}, \mathbf{y}) \leq t\}$$

Unfortunately, in contrast to CCCCCs, the size of $\mathcal{B}_S(\mathbf{x}, t; m, L, w_s)$ depends on $\mathbf{x}$. Take for example, $m = 1, L = 4, w_s = 2$ and $t = 1$. We have that $\mathcal{B}_S(0111, t; m, L, w_s) = \{0111, 1111, 0011, 0101, 0110\}$, while $\mathcal{B}_S(1001, t; m, L, w_s) = \{1001, 1101, 1011\}$.

We denote the smallest and the average ball size in the SECC space as follows:

$$|\mathcal{B}_S^{\min}(t; m, L, w_s)| \triangleq \min_{\mathbf{x} \in \mathcal{S}(m, L, w_s)} |\mathcal{B}_S(\mathbf{x}, t; m, L, w_s)|, \quad (14)$$

$$|\mathcal{B}_S^{\text{avg}}(t; m, L, w_s)| \triangleq \sum_{\mathbf{x} \in \mathcal{S}(m, L, w_s)} \frac{|\mathcal{B}_S(\mathbf{x}, t; m, L, w_s)|}{|\mathcal{S}(m, L, w_s)|}. \quad (15)$$

The total number of words in the SECC space $\mathcal{S}(m, L, w_s)$ are $\left(\sum_{i=w_s}^L \binom{L}{i}\right)^m$. Thus, the *generalized Gilbert-Varshamov bound* [12, Thm. 4], [17] (for spaces where balls with fixed radius and different centers may have different sizes) when

applied to the SECC space $\mathcal{S}(m, L, w_s)$ gives us the following lower bound on $S(m, L, d, w_s)$.

Proposition 3. *We have*

$$\begin{aligned} S(m, L, d, w_s) &\geq \sum_{\mathbf{x} \in \mathcal{S}(m, L, w_s)} \frac{1}{|\mathcal{B}_S(\mathbf{x}, t; m, L, w_s)|} \\ &\geq \frac{\left(\sum_{i=w_s}^L \binom{L}{i}\right)^m}{|\mathcal{B}_S^{\text{avg}}(d-1; m, L, w_s)|}. \end{aligned} \quad (16)$$

The first inequality in (16) follows from Turan's theorem [18], [19], while the second inequality follows from the arithmetic-harmonic mean inequality. The next proposition demonstrates how to construct SECCs from CSCCs.

Proposition 4.

(i) *We have*

$$S(m, L, d, w_s) \geq C(m, L, d, j), \text{ for } j \geq w_s. \quad (17)$$

(ii) *When $m \geq d$, we have*

$$S(m, L, d, w_s) \geq \sum_{j=w_s}^L C(m, L, d, j). \quad (18)$$

(iii) *When m is an integer multiple of d , with $m = kd$,*

$$S(m, L, d, w_s) \geq S(d, L, d, w_s)^k \geq \left(\sum_{j=w_s}^L C(d, L, d, j)\right)^k. \quad (19)$$

Proof: See Appendix C. ■

The next proposition extends the concatenation approach [20] for SECCs.

Proposition 5. *If $q \leq H(L, d_1, w_s)$, then*

$$S(m, L, d_1 d_2, w_s) \geq A_q(m, d_2).$$

Proof: Adapt the concatenated code construction scheme in [5, Prop. 4.1] by replacing the constant weight inner code by a heavy weight inner code. ■

The following bound on the optimal SECC code size is analogous to the Elias-Bassalygo bound (see for example, [8, eq. 2.7]).

Proposition 6. *We have*

$$S(m, L, d, w_s) \geq \frac{\left(\sum_{i=w_s}^L \binom{L}{i}\right)^m}{2^{mL}} A(mL, d). \quad (20)$$

Proof: See Appendix D. ■

2) *Upper bounds on SECC code size:* We next provide several upper bounds on the SECC code size, including the SECC sphere-packing bound (Prop. 7) and the SECC Johnson type bound (Prop. 8). Observing that the balls of radius $t = \lfloor (d-1)/2 \rfloor$ around codewords should be non-intersecting in an (m, L, d, w_s) -SECC, we having the following sphere-packing upper bound on $S(m, L, d, w_s)$.

Proposition 7. *Let $t \triangleq \lfloor (d-1)/2 \rfloor$. Then, we have*

$$S(m, L, d, w_s) \leq \frac{\left(\sum_{i=w_s}^L \binom{L}{i}\right)^m}{|\mathcal{B}_S^{\text{min}}(t; m, L, w_s)|}. \quad (21)$$

As discussed earlier, for a given radius t , different SECC balls may have different sizes, depending on the center word. In view of this, note that the SECC sphere-packing upper bound (21) is obtained by considering the smallest ball size of radius t . The *generalized sphere-packing bound*, for spaces where different balls of same radius have different sizes, was investigated in [13], [14]. As pointed out by a reviewer, the points with minimum size spheres are *not typical* in the space $\mathcal{S}(m, L, w_s)$ and a finer analysis akin to the techniques in [13], [14] may yield tighter *asymptotic* upper bound than that given in the next section via Theorem 4. However, this is beyond the scope of the paper and we defer this analysis to future work.

Furthermore, we point out that the average sphere-packing value is *not* an upper bound for the code size of SECCs. Specifically, for a t -error-correcting code, the *average sphere-packing value* was defined in [13] to be the ratio of size of the space, to the average ball size of radius t . It was observed that for many spaces, this average sphere-packing value is an upper bound for the optimal code size.

However, we now show that there exist SECC spaces where the average sphere-packing value is *not* an upper bound on the optimal code size. Towards this, consider the SECC space, $\mathcal{S}(m, L, w_s)$, corresponding to $m = 1$, $L = 3$, and $w_s = 1$. Here, the size of space, $|\mathcal{S}(m, L, w_s)|$, is 7 while the average ball size, $|\mathcal{B}_S^{\text{avg}}(t; m, L, w_s)|$, corresponding to $t = 1$ is equal to 25/7. In this case, the *average sphere-packing value*, for a single error correcting code, is 49/25. But this value is readily seen to be strictly less than the size of the SECC code $\mathcal{C} = \{100, 011\}$.

Other upper bounds on the optimal SECC code size are discussed next. The following inequality is immediate from the definition of SECC.

$$S(m, L, d, w_1) \leq S(m, L, d, w_2), \text{ if } w_1 > w_2.$$

We now present an upper bound on $S(m, L, d, w_s)$ which is analogous to the Johnson bound for constant weight codes [21], [22]. Towards this, we consider a generalization of SECC where different subblocks in a codeword may have different length and weight constraints. Let $T(m, [L_1, \dots, L_m], d, [w_1, \dots, w_m])$ denote the largest size of a binary code where each codeword has m subblocks, the i th subblock has length L_i and weight at least w_i , and the minimum distance of the code is d . Here, the length of each codeword is $n = \sum_{i=1}^m L_i$.

Now, let $\mathcal{C}$ be such a generalized code of size $T(m, [L_1, \dots, L_m], d, [w_1, \dots, w_m])$. Consider a matrix with n columns, whose rows comprise of the $T(m, [L_1, \dots, L_m], d, [w_1, \dots, w_m])$ codewords of $\mathcal{C}$. By focusing on the i th subblock of each codeword, we observe that there exists a column having at least $T(m, [L_1, \dots, L_m], d, [w_1, \dots, w_m]) \cdot (w_i/L_i)$ ones, and denote such a column as l . Pick a subcode of $\mathcal{C}$ where each codeword has a 1 in the l -th position. Delete the l -th

component in the subcode to obtain

$$\begin{aligned} & T(m, [L_1, \dots, L_i, \dots, L_m], d, [w_1, \dots, w_i, \dots, w_m]) \\ & \leq \frac{L_i}{w_i} T(m, [L_1, \dots, L_i - 1, \dots, L_m], d, [w_1, \dots, w_i - 1, \dots, w_m]). \end{aligned} \quad (22)$$

By varying i from 1 to m and recursively applying (22),

$$\begin{aligned} & T(m, [L_1, \dots, L_m], d, [w_1, \dots, w_m]) \\ & \leq \left(\prod_{i=1}^m \frac{L_i}{w_i} \right) T(m, [L_1 - 1, \dots, L_m - 1], d, [w_1 - 1, \dots, w_m - 1]). \end{aligned} \quad (23)$$

Specializing (23) to the case when each $L_i = L$ and $w_i = w_s$, we obtain the following upper bound for SECCs.

Proposition 8. *We have*

$$S(m, L, d, w_s) \leq \frac{L^m}{w_s^m} S(m, L - 1, d, w_s - 1). \quad (24)$$

We remark that the above proposition generalizes the upper bound on the optimal size of heavy weight codes in [9, Prop. 1], which can be recovered by setting $m = 1$ in (24). Further, it can be shown that the bound in Prop. 8 is tight for certain parameters. In fact, the following example shows that (24) can be applied to improve upon some of the entries for the optimal heavy weight code size in [9, Table I].

Example: Consider the case where $m = 1$, $L = 7$, $d = 4$, and $w_s = 4$. Then using Prop. 8 we have

$$\begin{aligned} S(1, 7, 4, 4) & \stackrel{(a)}{\leq} \frac{7}{4} S(1, 6, 4, 3) \\ & \stackrel{(b)}{\leq} \frac{7}{4} A(6, 4) \\ & \stackrel{(c)}{=} 7, \end{aligned} \quad (25)$$

where (a) follows from (24), inequality (b) follows because the size of a constrained code of length 6 and minimum distance 4 is upper bounded by the size of the corresponding unconstrained code, and (c) follows because $A(6, 4) = 4$ [23]. Now, SECC with $m = 1$ corresponds to a heavy weight code, and hence using [9, Table I] we get $S(1, 7, 4, 4) \geq 7$. Combining this with (25), we obtain the following results

$$S(1, 7, 4, 4) = 7, \quad (26)$$

$$S(1, 6, 4, 3) = 4. \quad (27)$$

From (26) it follows that the bound in Prop. 8 is tight for the parameters chosen in this example. More importantly, it implies that the optimal size of a heavy weight code with blocklength 7, minimum weight 4, and minimum distance 4 is *exactly* equal to 7, thereby improving upon the corresponding bound in [9, Table I]. Further, using (27) it follows the optimal size of a heavy weight code with blocklength 6, minimum weight 4, and minimum distance 3 is *exactly* 4, thereby improving upon the corresponding bound in [9, Table I].

In summary, we have provided an example which not only shows that the upper bound in Prop. 8 is tight for certain parameters, but also helps in improving the bound on the optimal size of heavy weight codes for two different parameters in [9].

We next present bounds on the asymptotic rate for CSCCs and SECCs.

III. ASYMPTOTIC BOUNDS ON RATE

The asymptotic rate for subblock constrained codes may be studied in scenarios where the number of subblocks m , or the subblock length L , or both, tend to infinity. The following proposition states that the asymptotic rate of CSCC is equal to that of CWC when the subblock length L tends to infinity, which is not surprising as the subblock constraint fades asymptotically.

Proposition 9. *For any positive integer m and $0 \leq \delta, \omega \leq 1$,*

$$\lim_{L \rightarrow \infty} \frac{\log C(m, L, \lfloor \delta m L \rfloor, \lfloor \omega L \rfloor)}{mL} = \alpha(\delta, \omega). \quad (28)$$

Proof: We have the inequality

$$\lim_{L \rightarrow \infty} \frac{\log C(m, L, \lfloor \delta m L \rfloor, \lfloor \omega L \rfloor)}{mL} \leq \alpha(\delta, \omega), \quad (29)$$

as $C(m, L, \lfloor \delta m L \rfloor, \lfloor \omega L \rfloor) \leq A(mL, \lfloor \delta m L \rfloor, m \lfloor \omega L \rfloor)$. On the other hand, from [5, Lemma 6.1], we have that

$$C(m, L, \lfloor \delta m L \rfloor, \lfloor \omega L \rfloor) \geq \left(\frac{\lfloor \omega L \rfloor}{\lfloor \delta m L \rfloor} \right)^m A(mL, \lfloor \delta m L \rfloor, m \lfloor \omega L \rfloor). \quad (30)$$

If $h(\cdot)$ denotes the binary entropy function, then

$$\lim_{L \rightarrow \infty} \frac{1}{mL} \log \left(\frac{\lfloor \omega L \rfloor}{\lfloor \delta m L \rfloor} \right)^m = h(\omega) - h(\omega) = 0,$$

and hence using (30) we have

$$\lim_{L \rightarrow \infty} \frac{\log C(m, L, \lfloor \delta m L \rfloor, \lfloor \omega L \rfloor)}{mL} \geq \alpha(\delta, \omega). \quad (31)$$

The proof is complete by combining (29) and (31). ■

Note that (28) also holds when $m \rightarrow \infty$. Asymptotic rate results were also presented in [5]. However, there were some inconsistencies in the definition of the asymptotic CSCC rate and the resulting claim in [5, Prop. 6.1] was incorrect. Proposition 9 above provides a correction. The inconsistency in the rate definition in [5] also renders [5, Thm. 6.3] incorrect, whose proof also contained some anomalies.

By combining Thm. 1 and Prop. 9, we obtain the following proposition on the SECC asymptotic rate in the scenario where the subblock length L tends to infinity.

Proposition 10. *For any positive integer m , and $0 \leq \delta \leq 1$, $1/2 \leq \omega \leq 1$, we have*

$$\lim_{L \rightarrow \infty} \frac{\log S(m, L, \lfloor \delta m L \rfloor, \lfloor \omega L \rfloor)}{mL} = \eta(\delta, \omega). \quad (32)$$

In the remainder of the paper, we fix the relative distance δ , the subblock length L , and the parameter w_s , and provide estimates of the asymptotic rates for CSCCs and SECCs as the number of blocks m tends to infinity. The motivation for fixing L to relatively small values comes from the application of CSCCs and SECCs to *simultaneous energy and information transfer* [3]. Here, it can be shown that if the weight of each subblock is sufficiently high, then a receiver with limited energy storage will not suffer from energy outage when the subblock length is less than a certain threshold [3], [6].

A. CCCC Rate

Recall the definitions of $\gamma(L, \delta, w_s/L)$ and $\alpha(\delta, w_s/L)$ given by (2) and (5). Furthermore, these quantities are related via the following inequality

$$\gamma(L, \delta, w_s/L) \leq \alpha(\delta, w_s/L). \quad (33)$$

The following proposition shows that for the case when $L = 2$ and $w_s = 1$, the CCCC rate $\gamma(L, \delta, w_s/L)$ is strictly less than $\alpha(\delta, w_s/L)$ when $0 < \delta < 1/2$.

Proposition 11. *We have*

$$\gamma(2, \delta, 1/2) = \frac{1}{2}\alpha(\delta, 1/2) = \frac{1}{2}\alpha(\delta). \quad (34)$$

Proof: It was shown in [5, Cor. 4.2] that

$$C(m, 2, 2d, 1) = A(m, d).$$

Then (34) follows immediately from the definitions of asymptotic rates. ■

Since $\alpha(\delta, 1/2) = \alpha(\delta)$ [8], the relation in (34) can alternately be expressed as $\gamma(2, \delta, 1/2) = (1/2)\alpha(\delta)$. Now, from the GV bound for general binary codes [8], we know that $\alpha(\delta) > 0$ for $0 < \delta < 0.5$, while from the asymptotic Plotkin bound [24] for binary codes, we have $\alpha(\delta) = 0$ for $\delta \geq 0.5$. Thus, from (34), it follows that the inequality in (33) is strict for the case when $L = 2$, $w_s = 1$, and $0 < \delta < 0.5$.

We now define

$$\delta^*(w_s/L) \triangleq 2 \left(\frac{w_s}{L} \right) \left(1 - \frac{w_s}{L} \right), \quad (35)$$

where $\delta^*(w_s/L)$ denotes the expected relative distance between two randomly chosen vectors in the space $\mathcal{C}(m, L, w_s)$. From the MRRW upper bound and the Gilbert lower bound for constant weight codes [8], we have

$$\alpha(\delta, w_s/L) = 0, \quad \text{if } \delta \geq \delta^*(w_s/L) \quad (36)$$

$$\alpha(\delta, w_s/L) > 0, \quad \text{if } \delta < \delta^*(w_s/L). \quad (37)$$

As $\alpha(\delta) = \alpha(\delta, 0.5)$, we have that $\alpha(\delta) > 0$ if $\delta < \delta^*(0.5) = 0.5$, while $\alpha(\delta) = 0$ if $\delta \geq 0.5$. This result on the rate of unconstrained code can also be obtained using the Gilbert-Varshamov lower bound and the Plotkin upper bound [24]. From (33) and (36), it follows that

$$\gamma(L, \delta, w_s/L) = 0, \quad \text{if } \delta \geq \delta^*(w_s/L). \quad (38)$$

Therefore, we are interested in determining $\gamma(L, \delta, w_s/L)$ for $\delta < \delta^*(w_s/L)$. In fact, we will show that the inequality (33) is strict for relatively small values of L and δ . To this end, Theorem 2 presents a lower bound for $\gamma(L, \delta, w_s/L)$ using the GV bound for $C(m, L, d, w_s)$ when $\delta < \delta^*(w_s/L)$. The following lemmas will be used towards proving this theorem.

Lemma 3. *For fixed positive integers m , n and z , let k_i , with $1 \leq i \leq m$, be integers which satisfy $0 \leq k_i \leq n$, $\sum_{i=1}^m k_i = z$. Then we have the inequality*

$$\prod_{i=1}^m \binom{n}{k_i} \leq \binom{n}{\lfloor z/m \rfloor}^{m_1} \binom{n}{\lceil z/m \rceil}^{m-m_1}, \quad (39)$$

where $m_1 = m \lceil z/m \rceil - z$.

Proof: Follows from log-concavity of the binomial coefficients [25]. ■

Lemma 4. *For $0 < k \leq w_s(L - w_s)/L$, we have the inequality*

$$\binom{w_s}{k} \binom{L - w_s}{k} > \binom{w_s}{k-1} \binom{L - w_s}{k-1}. \quad (40)$$

Proof: See Appendix E. ■

Theorem 2 (Asymptotic GV bound for CCCCs). *For $0 < \delta < \delta^*(w_s/L)$, we have*

$$\gamma(L, \delta, w_s/L) \geq \gamma_{GV}(L, \delta, w_s/L), \quad (41)$$

where $\gamma_{GV}(L, \delta, w_s/L)$ is defined as follows

a) For $L = 2$,

$$\gamma_{GV}(2, \delta, 1/2) \triangleq \frac{1}{2}(1 - h(\delta)). \quad (42)$$

b) For $L > 2$,

$$\begin{aligned} \gamma_{GV}(L, \delta, w_s/L) \triangleq & \left[\frac{1}{L} \log \binom{L}{w_s} \right. \\ & - \left(\frac{\lceil u \rceil - u}{L} \right) \log \binom{w_s}{\lfloor u \rfloor} \\ & - \left(\frac{1 + u - \lceil u \rceil}{L} \right) \log \binom{w_s}{\lceil u \rceil} \\ & - \left(\frac{\lceil u \rceil - u}{L} \right) \log \binom{L - w_s}{\lfloor u \rfloor} \\ & - \left(\frac{1 + u - \lceil u \rceil}{L} \right) \log \binom{L - w_s}{\lceil u \rceil} \\ & \left. - \min\{\theta(L, w_s), \phi(L, \delta)\} \right]^+, \quad (43) \end{aligned}$$

where $u \triangleq \delta L/2$, and

$$\begin{aligned} \theta(L, w_s) &\triangleq \frac{1}{L} \log (\min\{w_s, L - w_s\} + 1) \\ \phi(L, \delta) &\triangleq \left(\frac{1}{L} + \frac{\delta}{2} \right) h \left(\frac{1}{1 + \delta L/2} \right). \end{aligned}$$

Proof: See Appendix F. ■

Remark: In general, we have $\phi(L, \delta) < \theta(L, w_s)$ when w_s/L does not deviate significantly from $1/2$. On the other hand, $\theta(L, w_s)$ may be less than $\phi(L, \delta)$ when w_s is extremely close to L (or 0). In particular, $\theta(L, L - 1) < \phi(L, \delta)$ if $\delta \geq \frac{2}{3L}$.

Note that for $L = 2$, we have $\gamma_{GV}(2, \delta, 1/2) > 0$ for $0 < \delta < \delta^*(1/2) = 0.5$. However, in general, $\gamma_{GV}(L, \delta, w_s/L)$ is not necessarily positive when δ lies in the interval $0 < \delta < \delta^*(w_s/L)$. In particular, when $L = 3$ and $w_s = 2$, we have $\delta^*(2/3) = 4/9$, but $\gamma_{GV}(3, 0.4, 2/3) = 0$.

Further, let $\delta_\gamma^*(L, w_s/L)$ be defined as the smallest value of δ for which $\gamma(L, \delta, w_s/L) = 0$. Then, using (34), (36), and (37), we have $\delta_\gamma^*(2, 1/2) = 0.5$. As an $(m, 2, d, 1)$ -CCCC is also an $(m, 2w_s, d, w_s)$ -CCCC, when L is even, we have that $\delta_\gamma^*(L, 1/2) \geq \delta_\gamma^*(2, 1/2) = 0.5$. Further, as CCCCs are also constant weight codes, we have $\delta_\gamma^*(L, 1/2) \leq \delta^*(1/2) = 0.5$. Combining these results, we obtain that for even L ,

$$\delta_\gamma^*(L, 1/2) = 0.5. \quad (44)$$

However, the value of $\delta_\gamma^*(L, w_s)$ for general values of L and w_s is not known.

The following theorem presents the asymptotic sphere-packing upper bound on $\gamma(L, \delta, w_s/L)$.

Theorem 3 (Asymptotic sphere-packing bound for CSCCs). *The CSCC rate can be upper bounded as*

$$\gamma(L, \delta, w_s/L) \leq \gamma_{SP}(L, \delta, w_s/L), \quad (45)$$

where $\gamma_{SP}(L, \delta, w_s/L)$ is defined as

$$\begin{aligned} \gamma_{SP}(L, \delta, w_s/L) \triangleq & \left[\frac{1}{L} \log \left(\frac{L}{w_s} \right) \right. \\ & - \left(\frac{[\tilde{u}] - \tilde{u}}{L} \right) \log \left(\frac{w_s}{[\tilde{u}]} \right) \\ & - \left(\frac{1 + \tilde{u} - [\tilde{u}]}{L} \right) \log \left(\frac{w_s}{[\tilde{u}]} \right) \\ & - \left(\frac{[\tilde{u}] - \tilde{u}}{L} \right) \log \left(\frac{L - w_s}{[\tilde{u}]} \right) \\ & - \left(\frac{1 + \tilde{u} - [\tilde{u}]}{L} \right) \log \left(\frac{L - w_s}{[\tilde{u}]} \right) \\ & \left. - \frac{1}{L} h([\tilde{u}] - \tilde{u}) \right]^+, \end{aligned} \quad (46)$$

where $\tilde{u} \triangleq \delta L/4$.

Proof: See Appendix G. ■

Remark: For $L = 2, w_s = 1$, we have $\gamma_{SP}(2, \delta, 0.5) = 0.5(1 - h(\delta/2))$, which also follows from $\gamma(2, \delta, 1/2) = (1/2)\alpha(\delta)$ and then applying the standard sphere-packing bound for unconstrained binary codes.

For $0 < \delta < \min\{\delta^*(w_s/L), \frac{4}{L}\}$, (46) simplifies to

$$\begin{aligned} \gamma_{SP}(L, \delta, w_s/L) = & \frac{1}{L} \log \left(\frac{L}{w_s} \right) - \frac{\delta}{4} \log(w_s(L - w_s)) \\ & - \frac{1}{L} h\left(\frac{\delta L}{4}\right). \end{aligned} \quad (47)$$

B. SECC Rate

Recall the definitions of $\sigma(L, \delta, w_s/L)$ and $\eta(\delta, w_s/L)$ given by (3) and (6). We have the following inequality

$$\sigma(L, \delta, w_s/L) \leq \eta(\delta, w_s/L). \quad (48)$$

The gap $\eta(\delta, w_s/L) - \sigma(L, \delta, w_s/L)$ denotes the rate penalty on HWC due to the additional constraint on sufficient weight within every subblock duration. The asymptotic rates of HWCs were studied in [9] where Theorem 1 was established. Therefore, it follows that for $w_s \geq L/2$ we have

$$\sigma(L, \delta, w_s/L) = 0, \text{ if } \delta \geq \delta^*(w_s/L).$$

Let $\delta_\sigma^*(L, w_s/L)$ be defined as the smallest value of δ for which $\sigma(L, \delta, w_s/L) = 0$. As, an (m, L, d, w_s) -CSCC is also an (m, L, d, w_s) -SECC, using (44) it follows that for even L we have $\delta_\sigma^*(L, 1/2) \geq 0.5$. Combining this with the fact that $\alpha(\delta) = 0$ for $\delta \geq 0.5$, we obtain

$$\delta_\sigma^*(L, 1/2) = 0.5. \quad (49)$$

Also, when L is even and $w_s \leq L/2$, then an $(m, L, d, L/2)$ -SECC is also an (m, L, d, w_s) -SECC, and hence (49) can be generalized to obtain

$$\delta_\sigma^*(L, w_s/L) = 0.5, \text{ when } L \text{ is even and } w_s \leq L/2. \quad (50)$$

On the other hand, when $w_s > L/2$, then we have $\sigma(L, \delta, w_s/L) \leq \eta(\delta, w_s/L) = \alpha(\delta, w_s/L)$ and hence

$$\delta_\sigma^*(L, w_s/L) \leq \delta^*(w_s/L), \text{ when } w_s > L/2. \quad (51)$$

In the following, we present the asymptotic GV bound and the sphere-packing bound on $\sigma(L, \delta, w_s/L)$.

Proposition 12 (Asymptotic GV bound for SECCs). *We have $\sigma(L, \delta, w_s/L) \geq \sigma_{GV}(L, \delta, w_s/L)$ where*

$$\sigma_{GV}(L, \delta, w_s/L) \triangleq \left[\frac{1}{L} \log \left(\sum_{j=w_s}^L \binom{L}{j} \right) - h(\delta) \right]^+. \quad (52)$$

Proof: A simple upper bound on the average SECC ball size of radius $d - 1$ is given by

$$|\mathcal{B}_S^{\text{avg}}(d - 1; m, L, w_s)| \leq \sum_{i=1}^{d-1} \binom{mL}{i}. \quad (53)$$

Using Proposition 3 and (53), we get

$$S(m, L, d, w_s) \geq \frac{\left(\sum_{j=w_s}^L \binom{L}{j} \right)^m}{\sum_{i=1}^{d-1} \binom{mL}{i}}. \quad (54)$$

The proposition now follows by combining (3) and (54). ■

The above proposition presents a lower bound on $\sigma(L, \delta, w_s/L)$. Next, in Theorem 4 we present the sphere-packing upper bound on $\sigma(L, \delta, w_s/L)$ for relatively small values of δ . We will use the following lemma towards proving this theorem.

Lemma 5. *For $0 < \delta < 2/L$, we have*

$$\lim_{m \rightarrow \infty} \frac{1}{mL} \log |\mathcal{B}_S^{\text{min}}(\lfloor (mL\delta - 1)/2 \rfloor; m, L, w_s)| \geq \nu(L, \delta, w_s), \quad (55)$$

where $\nu(L, \delta, w_s)$ is defined as

$$\begin{aligned} \max \left\{ \frac{1}{L} h\left(\frac{L\delta}{4}\right) + \frac{\delta}{4} \log((L - w_s)(w_s + 1)), \right. \\ \left. \frac{1}{L} h\left(\frac{L\delta}{2}\right) + \frac{\delta}{2} \log(L - w_s) \right\}. \end{aligned} \quad (56)$$

Proof: See Appendix H. ■

Theorem 4 (Asymptotic sphere-packing bound for SECCs). *For $0 < \delta < 2/L$, we have $\sigma(L, \delta, w_s/L) \leq \sigma_{SP}(L, \delta, w_s/L)$ where*

$$\sigma_{SP}(L, \delta, w_s/L) \triangleq \frac{1}{L} \log \left(\sum_{j=w_s}^L \binom{L}{j} \right) - \nu(L, \delta, w_s), \quad (57)$$

where $\nu(L, \delta, w_s)$ is defined in (56).

Proof: Follows by combining (3), Prop. 7 and Lem. 5. ■

For $L = 2, w_s = 1$, and $\delta < 0.4$, the asymptotic sphere-packing bound for SECCs reduces to

$$\sigma_{SP}(2, \delta, 0.5) = \frac{1}{2} (\log_2(3) - h(\delta)). \quad (58)$$

IV. RATE PENALTY DUE TO SUBBLOCK CONSTRAINTS

In this section, we quantify the penalty in rate due to imposition of subblock constraints, relative to the application of corresponding constraints per codeword. For each pair of codeword / subblock constraints, we do the following:

- i) Provide a value of δ_L such that rate penalty is strictly positive for $0 \leq \delta \leq \delta_L$. To do so, we make use of the asymptotic upper and lower bounds in the previous section to compute a lower bound for this penalty. Using continuity arguments, we then show that this penalty is nonzero in the given interval.
- ii) Provide a value of δ^* such that the rate penalty is zero for $\delta^* \leq \delta \leq 1$.
- iii) Demonstrate that the estimate of the rate gap is tight when δ approaches zero.

A. CWC versus CSCC

The rate penalty due to constant weight per *subblock*, relative to the constraint requiring constant weight per *codeword*, is quantified by $G_{\alpha-\gamma}(L, \delta, w_s/L)$, defined as

$$G_{\alpha-\gamma}(L, \delta, w_s/L) \triangleq \alpha(\delta, w_s/L) - \gamma(L, \delta, w_s/L).$$

A lower bound to this rate gap is given by

$$G_{\alpha-\gamma}^{LB}(L, \delta, w_s/L) \triangleq [\alpha_{GV}(\delta, w_s/L) - \gamma_{SP}(L, \delta, w_s/L)]^+, \quad (59)$$

where $\gamma_{SP}(L, \delta, w_s/L)$ is defined in (46) and

$$\alpha_{GV}(\delta, \omega) \triangleq h(\omega) - \omega h\left(\frac{\delta}{2\omega}\right) - (1-\omega)h\left(\frac{\delta}{2(1-\omega)}\right), \quad (60)$$

with $\alpha_{GV}(\delta, w_s/L)$ denoting the asymptotic GV lower bound for CWCs [8], [26]. The sphere-packing upper bound on the asymptotic rate for CWCs is given by $\alpha_{SP}(\delta, \omega)$, defined as

$$\alpha_{SP}(\delta, \omega) \triangleq h(\omega) - \omega h\left(\frac{\delta}{4\omega}\right) - (1-\omega)h\left(\frac{\delta}{4(1-\omega)}\right). \quad (61)$$

If $L = 2$ and $w_s = 1$, then using (34) we have the strict inequality $\alpha(\delta, 0.5) > \gamma(2, \delta, 0.5)$ for $0 < \delta < 0.5$. For relatively large values of the subblock length, L , the following proposition quantifies a lower bound on δ for which the rate penalty is strictly positive.

Proposition 13. *For even L with $L \geq 4$, we have the strict inequality $G_{\alpha-\gamma}^{LB}(L, \delta, 0.5) > 0$ for $0 < \delta < \tilde{\delta}_L$, where $\tilde{\delta}_L$ is the smallest positive root of $\tilde{f}_L(\delta)$ defined as*

$$\tilde{f}_L(\delta) \triangleq 1 - h(\delta) - \frac{1}{L} \log\left(\frac{L}{L/2}\right) + \frac{\delta}{2} \log \frac{L}{2} + \frac{1}{L} h\left(\frac{\delta L}{4}\right). \quad (62)$$

Proof: See Appendix I. ■

The following proposition addresses the converse question on identifying an interval for δ when the rate gap between CWCs and CSCCs is provably zero.

Proposition 14. *The rate gap between CWCs and CSCCs, $G_{\alpha-\gamma}(L, \delta, w_s/L)$, is identically zero for $\delta^*(w_s/L) \leq \delta \leq 1$.*

Proof: Follows from (36) and (38). ■

In [3], the gap between CWC capacity and CCCC capacity on noisy binary input channels was upper bounded by the rate penalty term, $r(L, \omega)$, defined as

$$r(L, \omega) \triangleq h(\omega) - (1/L) \log\left(\frac{L}{L\omega}\right), \quad (63)$$

where $\omega = w_s/L$. Further, it was shown in [3] that the actual capacity gap is equal to $r(L, \omega)$ for a noiseless channel. The following proposition shows that $G_{\alpha-\gamma}^{LB}(L, \delta, w_s/L)$ tends to $r(L, w_s/L)$ as δ tends to 0.

Proposition 15. *For $0 < w_s < L$, we have*

$$\lim_{\delta \rightarrow 0} G_{\alpha-\gamma}^{LB}(L, \delta, w_s/L) = r(L, w_s/L) > 0. \quad (64)$$

Proof: From (60) we have $\lim_{\delta \rightarrow 0} \alpha_{GV}(\delta, w_s/L) = h(w_s/L)$, while using (47) we obtain the limit $\lim_{\delta \rightarrow 0} \gamma_{SP}(L, \delta, w_s/L) = (1/L) \log\left(\frac{L}{w_s}\right)$, and hence the claim follows from definitions (59) and (63). ■

Proposition 16. *The lower bound on the rate gap between CWCs and CSCCs, $G_{\alpha-\gamma}^{LB}(L, \delta, w_s/L)$, is tight when $\delta \rightarrow 0$.*

Proof: An upper bound on $G_{\alpha-\gamma}(L, \delta, w_s/L)$ is given by $\alpha_{SP}(\delta, w_s/L) - \gamma_{GV}(\delta, \omega)$. Using (42), (43), and (61), we observe that this upper bound on the rate gap also tends to $r(L, w_s/L)$ as δ tends to 0. The proof is complete by combining this observation with Proposition 15. ■

B. HWC versus SECC

In SECCs, the fraction of ones in every subblock is at least w_s/L , and hence the fraction of ones in the entire codeword is also at least w_s/L . Relative to the constraint requiring at least w_s/L fraction of bits to be 1 for all *codewords*, the rate penalty due to the constraint requiring minimum weight w_s per *subblock* is quantified by $G_{\eta-\sigma}(L, \delta, w_s/L)$, defined as

$$G_{\eta-\sigma}(L, \delta, w_s/L) \triangleq \eta(\delta, w_s/L) - \sigma(L, \delta, w_s/L). \quad (65)$$

For $w_s \geq L/2$, using Theorem 1, we note that $G_{\eta-\sigma}(L, \delta, w_s/L)$ can equivalently be expressed as $\alpha(\delta, w_s/L) - \sigma(L, \delta, w_s/L)$. Thus, a lower bound for $G_{\eta-\sigma}(L, \delta, w_s/L)$, for $w_s \geq L/2$, is given by

$$G_{\eta-\sigma}^{LB}(L, \delta, w_s/L) \triangleq [\alpha_{GV}(\delta, w_s/L) - \sigma_{SP}(L, \delta, w_s/L)]^+, \quad (66)$$

where $\alpha_{GV}(\delta, w_s/L)$ and $\sigma_{SP}(L, \delta, w_s/L)$ are defined in (60) and (57), respectively. When $w_s \leq L/2$, we have $\eta(\delta, w_s/L) = \alpha(\delta, 0.5)$, and in this case, the corresponding rate gap lower bound is defined as

$$G_{\eta-\sigma}^{LB}(L, \delta, w_s/L) \triangleq [\alpha_{GV}(\delta, 0.5) - \sigma_{SP}(L, \delta, w_s/L)]^+. \quad (67)$$

The following proposition quantifies a lower bound on δ for which the rate gap between HWCs and SECCs is strictly positive.

Proposition 17. For even L with $L \geq 4$, we have the strict inequality $G_{\eta-\sigma}^{LB}(L, \delta, 0.5) > 0$ for $0 < \delta < \hat{\delta}_L$, where $\hat{\delta}_L$ is the smallest positive root of $\hat{f}_L(\delta)$ defined as

$$\hat{f}_L(\delta) \triangleq 1 - h(\delta) - \frac{1}{L} \log \left(\sum_{j=L/2}^L \binom{L}{j} \right) + \frac{1}{L} h \left(\frac{L\delta}{2} \right) + \frac{\delta}{2} \log \left(\frac{L}{2} \right). \quad (68)$$

Proof: See Appendix J. ■

When $L = 2$ and $w_s = 1$, it can be verified using (58) that $G_{\eta-\sigma}^{LB}(2, \delta, 0.5) > 0$ for $0 \leq \delta < 0.083$.

Prop. 17 considers the case where $w_s = L/2$. In general for $0 < w_s < L$, using a similar technique, we can obtain lower bounds on δ for which the rate gap between HWCs and SECCs is strictly positive. The following proposition addresses the converse question on identifying an interval for δ when this gap is provably zero.

Proposition 18. For $w_s \leq L/2$, the rate gap between HWCs and SECCs, $G_{\eta-\sigma}(L, \delta, w_s/L)$ is identically zero when $1/2 \leq \delta \leq 1$, while for $w_s \geq L/2$, this gap is zero when $\delta^*(w_s/L) \leq \delta \leq 1$.

Proof: Follows from (8) and (36). ■

The following proposition shows that the bound on rate gap is tight when $\delta \rightarrow 0$.

Proposition 19. The lower bound on the rate gap between HWCs and SECCs, $G_{\eta-\sigma}^{LB}(L, \delta, w_s/L)$, is tight when $\delta \rightarrow 0$.

Proof: See Appendix K. ■

C. SECC versus CSCC

The SECCs, relative to CSCCs, provide the flexibility of allowing different subblocks to have different weights. In this subsection, we show that this flexibility leads to an improvement in asymptotic rate when the relative distance of the code is sufficiently small. The gap between SECC rate and CSCC rate is quantified by $G_{\sigma-\gamma}(L, \delta, w_s/L)$, defined as

$$G_{\sigma-\gamma}(L, \delta, w_s/L) \triangleq \sigma(L, \delta, w_s/L) - \gamma(L, \delta, w_s/L).$$

A lower bounded to this rate gap is given by

$$G_{\sigma-\gamma}^{LB}(L, \delta, w_s/L) \triangleq [\sigma_{GV}(L, \delta, w_s/L) - \gamma_{SP}(L, \delta, w_s/L)]^+, \quad (69)$$

where $\sigma_{GV}(L, \delta, w_s/L)$ and $\gamma_{SP}(L, \delta, w_s/L)$ are given by (52) and (46), respectively. The following proposition quantifies the range of δ over which $G_{\sigma-\gamma}^{LB}(L, \delta, w_s/L)$ is strictly positive.

Proposition 20. For even L with $L \geq 4$, we have the strict inequality $G_{\sigma-\gamma}^{LB}(L, \delta, 0.5) > 0$ for $0 < \delta < \hat{\delta}_L$, where $\hat{\delta}_L$ is the smallest positive root of $\hat{f}_L(\delta)$ defined as

$$\hat{f}_L(\delta) \triangleq \frac{1}{L} \log \left(\sum_{j=L/2}^L \binom{L}{j} \right) - h(\delta) - \frac{1}{L} \log \left(\frac{L}{L/2} \right) + \frac{\delta}{2} \log \frac{L}{2} + \frac{1}{L} h \left(\frac{\delta L}{4} \right). \quad (70)$$

Proof: See Appendix L. ■

For the case when $L = 2$ and $w_s = 1$, we have

$$G_{\sigma-\gamma}^{LB}(2, \delta, 0.5) = \left[\frac{1}{2} \log(3/2) - h(\delta) + \frac{1}{2} h(\delta/2) \right]^+, \quad (71)$$

and $G_{\sigma-\gamma}^{LB}(2, \delta, 0.5)$ is strictly positive for $0 \leq \delta < 0.084$.

From Proposition 12 and Theorem 4, note that for $0 < \delta < \min\{\delta^*(w_s/L), 4/L\}$, we have $\sigma_{GV}(L, \delta, w_s/L) \leq \sigma_{SP}(L, \delta, w_s/L)$, and hence it follows from definitions (59), (66), and (69) that

$$G_{\alpha-\gamma}^{LB}(L, \delta, w_s/L) \geq G_{\alpha-\sigma}^{LB}(L, \delta, w_s/L) + G_{\sigma-\gamma}^{LB}(L, \delta, w_s/L). \quad (72)$$

Although Prop. 20 only considers the case $w_s = L/2$, a similar argument can be applied to show that the rate gap between SECCs and CSCCs is strictly positive in a general setting where $0 < w_s < L$, provided δ is sufficiently small. The following converse, providing an interval for δ which results in zero rate gap, is obtained by using an argument similar to that in Proposition 18.

Proposition 21. For $w_s \leq L/2$, the rate gap between SECCs and CSCCs, $G_{\sigma-\gamma}(L, \delta, w_s/L)$, is identically zero when $1/2 \leq \delta \leq 1$, while for $w_s \geq L/2$, this gap is zero when $\delta^*(w_s/L) \leq \delta \leq 1$.

The following proposition establishes the tightness of $G_{\sigma-\gamma}^{LB}(L, \delta, w_s/L)$ when δ tends to 0.

Proposition 22. The lower bound on the rate gap between SECCs and CSCCs, $G_{\sigma-\gamma}^{LB}(L, \delta, w_s/L)$, is tight when $\delta \rightarrow 0$.

Proof: From (46) and (52), we have that

$$G_{\sigma-\gamma}^{LB}(L, 0, w_s/L) = \frac{1}{L} \log \left(\sum_{j=w_s}^L \binom{L}{j} \right) - \frac{1}{L} \log \binom{L}{w_s}. \quad (73)$$

An upper bound on $G_{\sigma-\gamma}(L, \delta, w_s/L)$ is given by $\sigma_{SP}(L, \delta, w_s/L) - \gamma_{GV}(L, \delta, w_s/L)$. From (42), (43), and (57), we note that this upper bound tends to the right hand side of (73) as $\delta \rightarrow 0$. ■

V. NUMERICAL RESULTS

In this section, we provide numerical bounds on rate penalties due to weight constraint per subblock, relative to imposing similar constraint per codeword.

Fig. 1 plots the lower bound on the rate gap between CWCs and CSCCs, $G_{\alpha-\gamma}^{LB}(L, \delta, 0.5)$, as a function of the subblock length. The upper bound on the gap between CWC capacity and CSCC capacity on noisy binary input channels for $w_s = L/2$, given by $r(L, 0.5)$ (see (63)), is also plotted in red. As suggested by Proposition 15, the figure shows that $G_{\alpha-\gamma}^{LB}(L, \delta, 0.5)$ tends to $r(L, 0.5)$ as δ gets close to zero. For a fixed value of w_s/L , note that $\alpha_{GV}(L, \delta, w_s/L)$ is independent of L . Thus, for a given δ , the decrease in $G_{\alpha-\gamma}^{LB}(L, \delta, 0.5)$ with increasing L is due to an increase in CSCC rate. This is intuitively expected, because an increase in L allows for greater flexibility in the choice of bits within every subblock. Further, from Proposition 9, it follows that $G_{\alpha-\gamma}^{LB}(L, \delta, 0.5) \rightarrow 0$ as $L \rightarrow \infty$.

Fig. 1: $G_{\alpha-\gamma}^{LB}(L, \delta, 0.5)$ versus subblock length, L .

Fig. 2: $G_{\alpha-\gamma}^{LB}(16, \delta, w_s/16)$ versus subblock weight, w_s .

Fig. 2 plots $G_{\alpha-\gamma}^{LB}(L, \delta, w_s/L)$ when the subblock length is fixed at $L = 16$, and w_s varies from $L/2 = 8$ to $L - 1 = 15$. Note that $\alpha_{GV}(L, \delta, (L - w_s)/L) = \alpha_{GV}(L, \delta, w_s/L)$ and $\gamma_{SP}(L, \delta, (L - w_s)/L) = \gamma_{SP}(L, \delta, w_s/L)$ (see (46) and (60), respectively), and thus $G_{\alpha-\gamma}^{LB}(L, \delta, (L - w_s)/L) = G_{\alpha-\gamma}^{LB}(L, \delta, w_s/L)$. Note that Figs. 1 and 2 illustrate that $G_{\alpha-\gamma}^{LB}(L, \delta, w_s/L)$ decreases with δ .

Fig. 3 depicts the region where the gap between CWC rate and CSCC rate is provably strictly positive. Note that $\tilde{\delta}_L$ is the smallest value of δ for which the lower bound $G_{\alpha-\gamma}^{LB}(L, \delta, w_s/L)$ is zero, when L is fixed, and $w_s = L/2$ (see Prop. 13). The figure shows that $\tilde{\delta}_L$ decreases with L , and from Proposition 9 it follows that $\tilde{\delta}_L \rightarrow 0$ when $L \rightarrow \infty$. Moreover, using Proposition 14, it is seen that the actual rate gap $G_{\alpha-\gamma}(L, \delta, 0.5)$ is provably zero for $\delta \geq 0.5$.

Fig. 4 plots $G_{\eta-\sigma}^{LB}(L, \delta, w_s/L)$, lower bound for the rate gap between HWCs and SECCs, as a function of L , with $w_s = L/2$. For a given δ , it is seen from the figure that $G_{\eta-\sigma}^{LB}(L, \delta, 0.5)$ decreases with L . Note that for $w_s \geq L/2$, using Prop. 10, we have $G_{\eta-\sigma}^{LB}(L, \delta, w_s/L) \rightarrow 0$ as $L \rightarrow \infty$.

Fig. 3: Area with $G_{\alpha-\gamma}^{LB}(L, \delta, 0.5) > 0$. Note that $G_{\alpha-\gamma}(L, \delta, 0.5) = 0$ for $\delta \geq \delta^*(0.5) = 0.5$.

Fig. 4: $G_{\eta-\sigma}^{LB}(L, \delta, 0.5)$ as a function of subblock length, L .

Fig. 5 plots $G_{\eta-\sigma}^{LB}(L, \delta, w_s/L)$ versus w_s , for fixed $L = 16$.

The shaded area in Fig. 6 depicts the region where the rate gap between HWC and SECC is provably strictly positive. Here, $\hat{\delta}_L$ is the smallest value of δ for which the lower bound $G_{\eta-\sigma}^{LB}(L, \delta, w_s/L)$ is zero, when L is fixed, and $w_s = L/2$ (see Prop. 17). The figure shows that $\hat{\delta}_L$ decreases with L , and from Proposition 10 it follows that $\hat{\delta}_L \rightarrow 0$ when $L \rightarrow \infty$. Moreover, using Proposition 18, it is seen that the actual rate gap $G_{\eta-\sigma}(L, \delta, 0.5)$ is provably zero for $\delta \geq 0.5$.

Relative to CSCCs, the SECCs allow for greater flexibility in choice of bits within each subblock, by allowing the subblock weight to vary, provided it exceeds a certain threshold. This flexibility results in higher rate for SECCs and Fig. 7 plots $G_{\sigma-\gamma}^{LB}(L, \delta, 0.5)$, lower bound on the rate gap between SECCs and CSCCs. The figure shows that for a given δ , the rate gap bound decreases with L , and we have $G_{\sigma-\gamma}^{LB}(L, \delta, 0.5) \rightarrow 0$ as $L \rightarrow \infty$. The last assertion follows by combining Theorem 1, Proposition 9, and the fact that $\gamma(L, \delta, w_s/L) \leq \sigma(L, \delta, w_s/L) \leq \eta(L, \delta, w_s/L)$. Additionally,

Fig. 5: $G_{\eta-\sigma}^{LB}(16, \delta, w_s/16)$ as a function of w_s .

Fig. 7: $G_{\sigma-\gamma}^{LB}(L, \delta, 0.5)$ versus subblock length, L .

Fig. 6: Area with $G_{\eta-\sigma}^{LB}(L, \delta, 0.5) > 0$. Note that $G_{\eta-\sigma}(L, \delta, 0.5) = 0$ for $\delta \geq \delta^*(0.5) = 0.5$.

Fig. 8: $G_{\sigma-\gamma}^{LB}(16, \delta, w_s/16)$ as a function of w_s .

comparing Figs. 1, 4, and 7, we observe that the inequality in (72) is satisfied.

Fig. 8 plots $G_{\sigma-\gamma}^{LB}(L, \delta, w_s/L)$ versus w_s , for fixed $L = 16$, and $\delta \in \{0.001, 0.01\}$. On comparing Figs. 2, 5, and 8, it is observed that lower bounds on respective rate gaps satisfy (72). Fig. 9 depicts the region where the rate gap between SECC and CSCC is provably strictly positive. Here, $\hat{\delta}_L$ is the smallest value of δ for which the lower bound $G_{\sigma-\gamma}^{LB}(L, \delta, w_s/L)$ is zero, when L is fixed, and $w_s = L/2$ (see Prop. 20). Fig. 9 shows that $\hat{\delta}_L$ decreases with L , and from Thm. 1, Prop. 9, and Prop. 10 it follows that $\hat{\delta}_L \rightarrow 0$ when $L \rightarrow \infty$. Moreover, using Proposition 21, we see that the true gap $G_{\sigma-\gamma}(L, \delta, 0.5)$ is provably zero for $\delta \geq 0.5$.

VI. CONCLUDING REMARKS AND REFLECTIONS

We now reflect on the results presented in this paper and point out avenues for further work.

1) We derived upper and lower bounds for the sizes of CSCCs and SECCs. Interestingly, using these bounds, we are able

to compute new *exact* values of $S(1, L, d, w) = H(n, d, w)$ (see (26) and (27)). As the focus of the current paper is to quantify the rate gaps between the classes of codes, we defer the exploration of these optimal sizes to future work.

2) For a fixed subblock length L and weight parameter w_s , via Propositions 13, 17, and 20, we estimate the values of $\tilde{\delta}_L$, $\hat{\delta}_L$, and $\check{\delta}_L$ such that the gaps

$$\begin{aligned} G_{\alpha-\gamma}(L, \delta, w_s/L) &> 0 && \text{for } 0 \leq \delta < \tilde{\delta}_L, \\ G_{\eta-\sigma}(L, \delta, w_s/L) &> 0 && \text{for } 0 \leq \delta < \hat{\delta}_L, \text{ and,} \\ G_{\sigma-\gamma}(L, \delta, w_s/L) &> 0 && \text{for } 0 \leq \delta < \check{\delta}_L. \end{aligned}$$

These gaps then reflect the rate penalties due to imposition of subblock constraints, relative to the application of corresponding constraints per codeword. A natural question is whether these computed values are tight. In our future work, we plan to refine our analysis using techniques in [13], [14].

3) The converse problem, on identifying an interval for δ where the respective rate penalties are provably zero, is addressed via Propositions 14, 18, and 21. An interesting

Fig. 9: Area with $G_{\sigma-\gamma}^{LB}(L, \delta, 0.5) > 0$. Note that $G_{\sigma-\gamma}(L, \delta, 0.5) = 0$ for $\delta \geq \delta^*(0.5) = 0.5$.

but unsolved problem in this regard is to characterize the smallest δ beyond which the respective rate penalties are zero. We can get some insight from the numerical computations in [3], which indicate that there is a nonzero gap between CCCC and CWC capacities and a nonzero gap between CCCC and SECC capacities. This suggests that, for a fixed subblock length L , the rate penalties are zero if and only if the respective asymptotic rates themselves are zero. However, this remains an open problem.

- 4) Recall that $\delta_\sigma^*(L, w_s/L)$ refers to the smallest value of δ for which the SECC rate satisfies $\sigma(L, \delta, w_s/L) = 0$ and we showed that $\delta_\sigma^*(L, w_s/L) = \delta^*(1/2) = 0.5$ when $w_s/L \leq 0.5$. For the case $w_s/L > 0.5$, we showed that $\delta_\sigma^*(L, w_s/L) \leq \delta^*(w_s/L)$, and further conjecture that in this case $\delta_\sigma^*(L, w_s/L) = \delta^*(w_s/L)$. Hence, combined with the above remark, we conjecture that for a fixed subblock length L , and $w_s \geq L/2$, the rate gaps are nonzero if and only if $\delta \geq \delta^*(w_s/L)$.

ACKNOWLEDGMENT

The authors thank the editor and the reviewers for their constructive feedback and helpful suggestions. In particular, the results given by Eq. (13), (55) and (56) are based on suggestions by the reviewers.

APPENDIX A PROOF OF LEMMA 1

For $1 \leq i \leq m$, let $\mathbf{x}_{[i]}$ (resp. $\tilde{\mathbf{x}}_{[i]}$) denote the i th subblock of $\mathbf{x}$ (resp. $\tilde{\mathbf{x}}$). As $\mathbf{x}_{[i]}$ and $\tilde{\mathbf{x}}_{[i]}$ have constant weight w_s , there exists a permutation π_i on L letters such that $\tilde{\mathbf{x}}_{[i]} = \pi_i(\mathbf{x}_{[i]})$. Now, if π denotes the permutation on mL letter, induced by $\pi_i, 1 \leq i \leq m$, defined as

$$\pi(\mathbf{x}) \triangleq [\pi_1(\mathbf{x}_{[1]}) \cdots \pi_m(\mathbf{x}_{[m]})],$$

then, $\tilde{\mathbf{x}} = \pi(\mathbf{x})$. The proof is complete by observing that

$$\mathcal{B}_C(\tilde{\mathbf{x}}, t; m, L, w_s) = \{\pi(\mathbf{y}) : \mathbf{y} \in \mathcal{B}_C(\mathbf{x}, t; m, L, w_s)\}.$$

APPENDIX B PROOF OF LEMMA 2

From Lemma 1 we note that $|\mathcal{B}_C(\mathbf{x}, t; m, L, w_s)|$ is independent of the choice of $\mathbf{x}$. Let $\mathbf{x}_{[i]}$ be the i th subblock of $\mathbf{x}$. Then the distance of $\mathbf{x}_{[i]}$ from any length L binary vector of weight w_s is always even, and the number of length L binary vectors of weight w_s at a distance $2u_i$ from $\mathbf{x}_{[i]}$ is $\binom{w_s}{u_i} \binom{L-w_s}{u_i}$ when $0 \leq u_i \leq v$ (and 0 otherwise). Now, if $\mathbf{y} \in \mathcal{C}(m, L, w_s)$, and distance between i th subblocks of $\mathbf{x}$ and $\mathbf{y}$ is $2u_i$, then $\mathbf{y} \in \mathcal{B}_C(\mathbf{x}, t; m, L, w_s)$ if and only if $2 \sum_{i=1}^m u_i \leq t$. Hence, the size of CCCC ball of radius t is given by (9). ■

APPENDIX C PROOF OF PROPOSITION 4

(i) A CCCC with fixed subblock weight $j \geq w_s$ is also an SECC having subblock weight at least w_s .

(ii) If $\mathbf{s}$ and $\tilde{\mathbf{s}}$ are two CCCC sequences with m subblocks, constant weight per subblock j and $j+1$, respectively, then the Hamming distance between $\mathbf{s}$ and $\tilde{\mathbf{s}}$ is at least m .

(iii) Let $\mathcal{C}$ denote an (d, L, d, w_s) -SECC having size $S(d, L, d, w_s)$. Now construct a SECC code having $m = kd$ subblocks, where each block comprising d consecutive subblocks is chosen from $\mathcal{C}$. This construction provides the first inequality in (19), while the second inequality follows from (18). ■

APPENDIX D PROOF OF PROPOSITION 6

Let $\mathcal{C}$ be a (mL, d) -code with $A(mL, d)$ codewords. Let $\mathbb{F}_2^{mL}$ denote the space of binary vectors of length mL , and $\mathbf{x} \in \mathbb{F}_2^{mL}$ be chosen so that $|\mathcal{S}(m, L, w_s) \cap (\mathbf{x} + \mathcal{C})|$ is maximal. Then

$$\begin{aligned} S(m, L, d, w_s) &\geq |\mathcal{S}(m, L, w_s) \cap (\mathbf{x} + \mathcal{C})| \\ &\geq \frac{1}{2^{mL}} \sum_{\mathbf{y} \in \mathbb{F}_2^{mL}} |\mathcal{S}(m, L, w_s) \cap (\mathbf{y} + \mathcal{C})| \\ &= \frac{1}{2^{mL}} \sum_{\mathbf{y} \in \mathbb{F}_2^{mL}} \sum_{\mathbf{b} \in \mathcal{S}(m, L, w_s)} \sum_{\mathbf{c} \in \mathcal{C}} |\{\mathbf{b}\} \cap \{\mathbf{y} + \mathbf{c}\}| \\ &= \frac{1}{2^{mL}} \sum_{\mathbf{b} \in \mathcal{S}(m, L, w_s)} \sum_{\mathbf{c} \in \mathcal{C}} 1 \\ &= \frac{|\mathcal{S}(m, L, w_s)| |\mathcal{C}|}{2^{mL}}. \end{aligned}$$

APPENDIX E PROOF OF LEMMA 4

We have

$$\begin{aligned} \frac{\binom{w_s}{k} \binom{L-w_s}{k}}{\binom{w_s}{k-1} \binom{L-w_s}{k-1}} &= \frac{(w_s - (k-1)) ((L-w_s) - (k-1))}{k^2} \\ &\stackrel{(a)}{\geq} \frac{Lk - L(k-1) + (k-1)^2}{k^2} \\ &= \frac{(L-2k) + (k^2+1)}{k^2} \stackrel{(b)}{>} 1, \end{aligned}$$

where (a) follows because $w_s(L-w_s) \geq Lk$, and (b) follows from the fact that $k \leq \min\{w_s, L-w_s\} \leq L/2$. ■

APPENDIX F
PROOF OF THEOREM 2

The claim for $L = 2$ follows from (34) and the GV bound for general binary codes.

For establishing the result for $L > 2$, we use Proposition 1. The challenge here is to provide an appropriate upper bound on the CSCC ball size of radius $d-1$, $|\mathcal{B}_C(\mathbf{x}, d-1; m, L, w_s)|$ (the denominator in (10)). From (35) it follows that conditions $\delta < \delta^*(w_s/L)$ and $d = \lfloor mL\delta \rfloor$ imply

$$d-1 < 2m w_s(L-w_s)/L. \quad (74)$$

Further, if $v \triangleq \min\{w_s, L-w_s\}$ and $t \triangleq \lfloor (d-1)/2 \rfloor$, we note that $|\mathcal{B}_C(\mathbf{x}, d-1; m, L, w_s)|$ is equal to

$$\begin{aligned} & 1 + \sum_{\tau=1}^t \sum_{\substack{u_1+u_2+\dots+u_m=\tau, \\ 0 \leq u_i \leq v}} \left[\prod_{i=1}^m \binom{w_s}{u_i} \right] \left[\prod_{i=1}^m \binom{L-w_s}{u_i} \right] \\ & \stackrel{(i)}{\leq} 1 + \sum_{\tau=1}^t \sum_{\substack{u_1+u_2+\dots+u_m=\tau, \\ 0 \leq u_i \leq v}} Q_\tau, \end{aligned} \quad (75)$$

where (i) follows from Lemma 3 and the definition

$$Q_\tau \triangleq \left[\binom{w_s}{\lfloor \tau/m \rfloor} \binom{L-w_s}{\lfloor \tau/m \rfloor} \right]^{m_\tau} \left[\binom{w_s}{\lceil \tau/m \rceil} \binom{L-w_s}{\lceil \tau/m \rceil} \right]^{m-m_\tau} \quad (76)$$

$$m_\tau \triangleq m \lceil \tau/m \rceil - \tau.$$

From (74) and relations $1 \leq \tau \leq t$, with $t = \lfloor (d-1)/2 \rfloor$, it follows that $\tau/m < w_s(L-w_s)/L$. Now, applying Lemma 4 we note that Q_τ is a non-decreasing function of τ . Hence, using (75) we get

$$|\mathcal{B}_C(\mathbf{x}, d-1; m, L, w_s)| < Q_t \sum_{\substack{u_1+u_2+\dots+u_m \leq t, \\ 0 \leq u_i \leq v}} 1, \quad (77)$$

where the summation term simply denotes the number of tuples $(u_1, u_2, \dots, u_m)$ under two constraints: $u_1 + \dots + u_m \leq t$ and $0 \leq u_i \leq v$. This summation term can hence be upper bounded as

$$\sum_{\substack{u_1+u_2+\dots+u_m \leq t, \\ 0 \leq u_i \leq v}} 1 \leq \min \left\{ (v+1)^m, \binom{t+m}{m} \right\}, \quad (78)$$

where $(v+1)^m$ is obtained by counting the tuples $(u_1, u_2, \dots, u_m)$ which only satisfy the constraint $0 \leq u_i \leq v$, while $\binom{t+m}{m}$ is obtained using the concept of *weak compositions* [27, Chap. 2] by counting the tuples which satisfy the constraints $u_1 + \dots + u_m \leq t$ and $u_i \geq 0$ (with no upper bound on u_i for $1 \leq i \leq m$). Combining (77) and (78),

$$|\mathcal{B}_C(\mathbf{x}, d-1; m, L, w_s)| < Q_t \min \left\{ (v+1)^m, \binom{t+m}{m} \right\}, \quad (79)$$

where Q_t is obtained by substituting $\tau = t$ in (76). Finally, using (11), (2), and (79), we observe that $\gamma(L, \delta, w_s/L)$ is lower bounded by the following term

$$\begin{aligned} & \frac{1}{L} \log \left(\frac{L}{w_s} \right) - \lim_{m \rightarrow \infty} \frac{1}{mL} \log Q_t \\ & - \min \left\{ \frac{1}{L} \log(v+1), \lim_{m \rightarrow \infty} \frac{1}{mL} \log \binom{t+m}{m} \right\}. \end{aligned} \quad (80)$$

It can be verified that

$$\lim_{m \rightarrow \infty} \frac{1}{mL} \log \binom{t+m}{m} = \left(\frac{1}{L} + \frac{\delta}{2} \right) h \left(\frac{1}{1 + \delta L/2} \right). \quad (81)$$

Now, the t/m term in the expression for Q_t (76) is equal to $\delta L/2 =: u$, and it follows using (81) that the lower bound on $\gamma(L, \delta, w_s/L)$ given by (80) simplifies to the expression on the right hand side in (43). ■

APPENDIX G
PROOF OF THEOREM 3

For proving the claim, we apply Proposition 2 while providing an appropriate lower bound on the CSCC ball size of radius t , $|\mathcal{B}_C(\mathbf{x}, t; m, L, w_s)|$ (the denominator in (12)), where $t = \lfloor (d-1)/2 \rfloor$ and minimum distance d scales as $d = \lfloor mL\delta \rfloor$. If we define $v \triangleq \min\{w_s, L-w_s\}$, $\tilde{t} \triangleq \lfloor t/2 \rfloor$, and $\tilde{v} \triangleq \tilde{t}/m$, then we note that $\tilde{v}$ satisfies the following inequality

$$0 < \tilde{v} < \frac{d}{4m} \stackrel{(i)}{\leq} \frac{L\delta}{4} < \frac{v}{2}, \quad (82)$$

where (i) follows from the inequality $\delta < \delta^*(w_s/L)$. Further, the ball size $|\mathcal{B}_C(\mathbf{x}, t; m, L, w_s)|$ is equal to

$$\begin{aligned} & 1 + \sum_{\tau=1}^{\tilde{t}} \sum_{\substack{u_1+u_2+\dots+u_m=\tau, \\ 0 \leq u_i \leq v}} \prod_{i=1}^m \binom{w_s}{u_i} \binom{L-w_s}{u_i} \\ & \stackrel{(ii)}{>} \sum_{\substack{u_1+u_2+\dots+u_m=\tilde{t}, \\ \lfloor \tilde{v} \rfloor \leq u_i \leq \lceil \tilde{v} \rceil}} \prod_{i=1}^m \binom{w_s}{u_i} \binom{L-w_s}{u_i} \\ & = \binom{m}{\tilde{m}} \left[\binom{w_s}{\lfloor \tilde{v} \rfloor} \binom{L-w_s}{\lfloor \tilde{v} \rfloor} \right]^{\tilde{m}} \left[\binom{w_s}{\lceil \tilde{v} \rceil} \binom{L-w_s}{\lceil \tilde{v} \rceil} \right]^{m-\tilde{m}}, \end{aligned} \quad (83)$$

where (ii) follows using (82) (as the constraint $\lfloor \tilde{v} \rfloor \leq u_i \leq \lceil \tilde{v} \rceil$ is stricter than the constraint $0 \leq u_i \leq v$), and $\tilde{m}$ in (83) is defined as $\tilde{m} \triangleq m \lceil \tilde{t}/m \rceil - \tilde{t}$. Note that asymptotically we get the following limits

$$\lim_{m \rightarrow \infty} \tilde{v} = \lim_{m \rightarrow \infty} \frac{1}{m} \left\lfloor \frac{1}{2} \left\lfloor \frac{d-1}{2} \right\rfloor \right\rfloor = \frac{\delta L}{4} = \tilde{u}, \quad (84)$$

$$\lim_{m \rightarrow \infty} \frac{\tilde{m}}{m} = \lim_{m \rightarrow \infty} (\lceil \tilde{v} \rceil - \tilde{v}) = \lceil \tilde{u} \rceil - \tilde{u}. \quad (85)$$

Using (83), (84), and (85), we see that $\lim_{m \rightarrow \infty} \frac{1}{mL} \log |\mathcal{B}_C(\mathbf{x}, t; m, L, w_s)|$ is greater than

$$\begin{aligned} & \frac{1}{L} h(\lceil \tilde{u} \rceil - \tilde{u}) + \left(\frac{\lceil \tilde{u} \rceil - \tilde{u}}{L} \right) \log \left(\frac{w_s}{\lfloor \tilde{u} \rfloor} \right) \\ & + \left(\frac{1 + \tilde{u} - \lceil \tilde{u} \rceil}{L} \right) \log \left(\frac{w_s}{\lceil \tilde{u} \rceil} \right) + \left(\frac{\lceil \tilde{u} \rceil - \tilde{u}}{L} \right) \log \left(\frac{L-w_s}{\lfloor \tilde{u} \rfloor} \right) \\ & + \left(\frac{1 + \tilde{u} - \lceil \tilde{u} \rceil}{L} \right) \log \left(\frac{L-w_s}{\lceil \tilde{u} \rceil} \right). \end{aligned} \quad (86)$$

Now, the theorem is proved by combining Proposition 2, (2), and (86). ■

APPENDIX H PROOF OF LEMMA 5

Let $\mathbf{x}_i$ be a binary vector of length L whose weight $\tilde{w}_s$ satisfies $\tilde{w}_s \geq w_s$. Then we will show that the number of binary vectors with length L , weight at least w_s , which are at a distance of either 1 or 2 from $\mathbf{x}_i$ is lower bounded by $(L - w_s)(w_s + 1)$.

Let N_1 (resp. N_2) be the number of L length vectors of weight at least w_s which are at a distance 1 (resp. 2) from $\mathbf{x}_i$. We consider the following cases:

- 1) $\tilde{w}_s = w_s$: In this case $N_1 = L - w_s$. If $(L - w_s) \geq 2$, then $N_2 = (L - w_s)w_s + \binom{L - w_s}{2}$, else $N_2 = (L - w_s)w_s$.
- 2) $\tilde{w}_s = w_s + 1$: In this case $N_1 = L$. If $(L - w_s) \geq 2$, then $N_2 = (L - w_s)w_s + \binom{L - w_s}{2}$, else $N_2 = (L - w_s)w_s$.
- 3) $\tilde{w}_s \geq w_s + 2$: In this scenario, $N_1 = L$ and $N_2 = \binom{L}{2}$.

For all the above three cases, it can easily be verified that $N_1 + N_2 \geq (L - w_s)(w_s + 1)$. Now, let $t = \lfloor (d - 1)/2 \rfloor$ with minimum distance scaling as $d = \lfloor mL\delta \rfloor$. Let $\tilde{m} \triangleq \lfloor t/2 \rfloor$ and note that the constraint $\delta < 2/L$ implies that $\tilde{m} < m$. For a given $\mathbf{x} \in \mathcal{S}(m, L, w_s)$, let $\mathbf{x}_{[j]}$ denote the j -th subblock of $\mathbf{x}$, i.e. $\mathbf{x} = (\mathbf{x}_{[1]} \mathbf{x}_{[2]} \cdots \mathbf{x}_{[m]})$. Let $\Lambda_{\mathbf{x}} \subset \mathcal{S}(m, L, w_s)$ be the set of vectors which satisfy the following conditions:

- a) For every $\mathbf{y} \in \Lambda_{\mathbf{x}}$, exactly $\tilde{m}$ subblocks of $\mathbf{y}$ differ from corresponding subblocks of $\mathbf{x}$.
- b) If $\mathbf{y}_{[j]} \neq \mathbf{x}_{[j]}$, then $d(\mathbf{x}_{[j]}, \mathbf{y}_{[j]}) \in \{1, 2\}$.

From the above conditions, it follows that if $\mathbf{y} \in \Lambda_{\mathbf{x}}$, then $d(\mathbf{x}, \mathbf{y}) \leq 2\tilde{m} \leq t$, and hence $\Lambda_{\mathbf{x}} \subseteq \mathcal{B}_{\mathcal{S}}(\mathbf{x}, t; m, L, w_s)$ with

$$|\mathcal{B}_{\mathcal{S}}(\mathbf{x}, t; m, L, w_s)| \geq |\Lambda_{\mathbf{x}}| \stackrel{(i)}{\geq} \binom{m}{\tilde{m}} [(L - w_s)(w_s + 1)]^{\tilde{m}}, \quad (87)$$

where (i) follows from the observation at start of the proof. Because above inequality holds for all $\mathbf{x} \in \mathcal{S}(m, L, w_s)$,

$$|\mathcal{B}_{\mathcal{S}}^{\min}(t; m, L, w_s)| \geq \binom{m}{\tilde{m}} [(L - w_s)(w_s + 1)]^{\tilde{m}}. \quad (88)$$

Now,

$$\frac{\tilde{m}}{m} = \frac{1}{m} \left\lfloor \frac{1}{2} \left\lfloor \frac{\lfloor mL\delta \rfloor - 1}{2} \right\rfloor \right\rfloor \implies \lim_{m \rightarrow \infty} \frac{\tilde{m}}{m} = \frac{L\delta}{4}, \quad (89)$$

and hence $\lim_{m \rightarrow \infty} \frac{1}{mL} \log |\mathcal{B}_{\mathcal{S}}^{\min}(\lfloor (mL\delta - 1)/2 \rfloor; m, L, w_s)|$ is lower bounded by the term

$$\frac{1}{L} h\left(\frac{L\delta}{4}\right) + \frac{\delta}{4} \log((L - w_s)(w_s + 1)). \quad (90)$$

An alternate lower bound on $|\mathcal{B}_{\mathcal{S}}^{\min}(t; m, L, w_s)|$ can be obtained by observing that for $0 < \delta < 2/L$, we have $t = \lfloor (d - 1)/2 \rfloor < m$, and hence an SECC ball of radius t contains vectors such that a constituent subblock differs from the corresponding subblock of the center word of the ball in at most one bit. Thus, we have

$$|\mathcal{B}_{\mathcal{S}}^{\min}(t; m, L, w_s)| \geq \binom{m}{t} (L - w_s)^t, \quad (91)$$

and $\lim_{m \rightarrow \infty} \frac{1}{mL} \log |\mathcal{B}_{\mathcal{S}}^{\min}(\lfloor (mL\delta - 1)/2 \rfloor; m, L, w_s)|$ is lower bounded by the term

$$\frac{1}{L} h\left(\frac{L\delta}{2}\right) + \frac{\delta}{2} \log(L - w_s). \quad (92)$$

The proof is complete by combining (90) and (92). ■

APPENDIX I PROOF OF PROPOSITION 13

Using (47), (59), and (60), we have $G_{\alpha-\gamma}^{LB}(L, \delta, 0.5) = \tilde{f}_L(\delta)$ when $\delta < 2/L$. We observe from (62) that $\tilde{f}_L(\delta)$ is a continuous function of δ with

$$\tilde{f}_L(0) = 1 - \frac{1}{L} \log\left(\frac{L}{L/2}\right) > 0. \quad (93)$$

Further, when $\delta = 1/L$, we have

$$\begin{aligned} \tilde{f}_L\left(\frac{1}{L}\right) &< 1 - \frac{1}{L} \log\left(\frac{L}{L/2}\right) + \frac{1}{2L} \log \frac{L}{2} + \frac{1}{L} - h\left(\frac{1}{L}\right) \\ &\stackrel{(i)}{\leq} \frac{1}{2L} \log(2L) + \frac{1}{2L} \log \frac{L}{2} + \frac{1}{L} - h\left(\frac{1}{L}\right) \\ &= \frac{1}{L} \log(2L) - h\left(\frac{1}{L}\right) \stackrel{(ii)}{\leq} 0, \end{aligned} \quad (94)$$

where (i) and (ii) follow from [28, Ex. 5.8]. Now using (93), (94), and the intermediate value theorem [29], it follows that the equation $\tilde{f}_L(\delta) = 0$ has a solution in the interval $(0, 1/L)$. The proof is complete by denoting the smallest positive root of $\tilde{f}_L(\delta)$ by $\tilde{\delta}_L$. ■

APPENDIX J PROOF OF PROPOSITION 17

Using (66), (60), and (57), we have $G_{\eta-\sigma}^{LB}(L, \delta, 0.5) = \hat{f}_L(\delta)$ for $0 < \delta \leq 1/L$ when $L \geq 4$. We observe from (68) that $\hat{f}_L(\delta)$ is a continuous function of δ with

$$\hat{f}_L(0) = 1 - \frac{1}{L} \log\left(\sum_{j=L/2}^L \binom{L}{j}\right) > 0. \quad (95)$$

As $\sum_{j=L/2}^L \binom{L}{j} > 2^{L-1}$, for $L \geq 4$ we have

$$\hat{f}_L\left(\frac{1}{L}\right) < \frac{2}{L} + \frac{1}{2L} \log\left(\frac{L}{2}\right) - h\left(\frac{1}{L}\right) \stackrel{(i)}{<} 0, \quad (96)$$

where (i) follow using [28, Ex. 5.8]. Now from (95) and (96), it follows that the equation $\hat{f}_L(\delta) = 0$ has a solution in the interval $(0, 1/L)$. The proposition now follows by denoting the smallest positive root of $\hat{f}_L(\delta)$ by $\hat{\delta}_L$. ■

APPENDIX K PROOF OF PROPOSITION 19

For $w_s \leq L/2$, from (67) we have that

$$G_{\eta-\sigma}^{LB}(L, 0, w_s/L) = 1 - \frac{1}{L} \log\left(\sum_{j=w_s}^L \binom{L}{j}\right). \quad (97)$$

Now, from (8) and the relation $\alpha(\delta, 0.5) = \alpha(\delta)$, an upper bound on $G_{\eta-\sigma}(L, \delta, w_s/L)$ is given by $\alpha_{SP}(\delta, 0.5) - \sigma_{GV}(L, \delta, w_s/L)$. For $w_s \leq L/2$, from (52) and (61), we note that this upper bound tends to the right hand side of (97) as $\delta \rightarrow 0$. This proves the claim for $w_s \leq L/2$.

For $w_s \geq L/2$, from (66) we have that

$$G_{\eta-\sigma}^{LB}(L, 0, w_s/L) = h(w_s/L) - \frac{1}{L} \log\left(\sum_{j=w_s}^L \binom{L}{j}\right). \quad (98)$$

For $w_s \geq L/2$, an upper bound on the rate gap $G_{\eta-\sigma}(L, \delta, w_s/L)$ is given by $\alpha_{SP}(\delta, w_s/L) - \sigma_{GV}(L, \delta, w_s/L)$ (using (8)), and this upper bound tends to the right hand side of (98) as $\delta \rightarrow 0$. ■

APPENDIX L

PROOF OF PROPOSITION 20

Using (47), (52), and (69), we have $G_{\sigma-\gamma}^{LB}(L, \delta, 0.5) = \hat{f}_L(\delta)$ for $\delta < 2/L$. From (70) we note that $\hat{f}_L(\delta)$ is a continuous function of δ with

$$\hat{f}_L(0) = \frac{1}{L} \log \left(\sum_{j=L/2}^L \binom{L}{j} \right) - \frac{1}{L} \log \left(\frac{L}{L/2} \right) > 0. \quad (99)$$

Further, comparing (62) and (70), we observe that $\hat{f}_L(\delta) < \tilde{f}_L(\delta)$. In particular, for $\delta = 1/L$ we have

$$\hat{f}_L \left(\frac{1}{L} \right) < \tilde{f}_L \left(\frac{1}{L} \right) < 0, \quad (100)$$

where the last inequality follows from (94). From (99) and (100) it follows that the equation $\hat{f}_L(\delta) = 0$ has a solution in the interval $(0, 1/L)$. The proof is complete by denoting the smallest positive root of $\hat{f}_L(\delta)$ by δ_L . ■

REFERENCES

- [1] A. Tandon, H. M. Kiah, and M. Motani, "Bounds on the asymptotic rate of binary constant subblock-composition codes," in *Proc. 2017 IEEE Int. Symp. Inf. Theory*, Jun. 2017, pp. 704–708.
- [2] —, "Binary subblock energy-constrained codes: Bounds on code size and asymptotic rate," in *Proc. 2017 IEEE Int. Symp. Inf. Theory*, Jun. 2017, pp. 1480–1484.
- [3] A. Tandon, M. Motani, and L. R. Varshney, "Subblock-constrained codes for real-time simultaneous energy and information transfer," *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 4212–4227, Jul. 2016.
- [4] Y. M. Chee, H. M. Kiah, and P. Purkayastha, "Matrix codes and multitone frequency shift keying for power line communications," in *Proc. 2013 IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 2870–2874.
- [5] Y. M. Chee *et al.*, "Multiply constant-weight codes and the reliability of loop physically unclonable functions," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7026–7034, Nov. 2014.
- [6] A. Tandon, M. Motani, and L. R. Varshney, "Subblock energy-constrained codes for simultaneous energy and information transfer," in *Proc. 2016 IEEE Int. Symp. Inf. Theory*, Jul. 2016, pp. 1969–1973.
- [7] G. Cohen, P. Solé, and A. Tchamkerten, "Heavy weight codes," in *Proc. 2010 IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 1120–1124.
- [8] R. McEliece *et al.*, "New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities," *IEEE Trans. Inf. Theory*, vol. 23, no. 2, pp. 157–166, Mar. 1977.
- [9] C. Bachoc *et al.*, "On bounded weight codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6780–6787, Oct. 2011.
- [10] Y. M. Chee *et al.*, "Constructions of optimal and near-optimal multiply constant-weight codes," Nov. 2014, arXiv:1411.2513 [cs.IT].
- [11] X. Wang *et al.*, "New bounds and constructions for multiply constant-weight codes," *IEEE Trans. Inf. Theory*, vol. 62, no. 11, pp. 6315–6327, Sep. 2016.
- [12] J. Gu and T. Fuja, "A generalized Gilbert-Varshamov bound derived via analysis of a code-search algorithm," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 1089–1093, May 1993.
- [13] A. Fazeli, A. Vardy, and E. Yaakobi, "Generalized sphere packing bound," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2313–2334, May 2015.
- [14] D. Cullina and N. Kiyavash, "Generalized sphere-packing bounds on the size of codes for combinatorial channels," *IEEE Trans. Inf. Theory*, vol. 62, no. 8, pp. 4454–4465, Aug. 2016.
- [15] C. Freiman, "Upper bounds for fixed-weight codes of specified minimum distance (Corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-10, no. 3, pp. 246–248, 1964.
- [16] E. Berger, "Some additional upper bounds for fixed-weight codes of specified minimum distance," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 2, pp. 307–308, 1967.
- [17] L. M. G. M. Tolhuizen, "The generalized Gilbert-Varshamov bound is implied by Turan's theorem," *IEEE Trans. Inf. Theory*, vol. 43, no. 5, pp. 1605–1606, Sep. 1997.
- [18] Y. Caro, "New results on the independence number," Tel Aviv University, Tech. Rep., 1979.
- [19] V. K. Wei, "A lower bound on the stability number of a simple graph," Bell Laboratories Technical Memorandum 81-11217-9, 1981.
- [20] G. D. Forney, Jr., *Concatenated Codes*. Cambridge, MA, USA: MIT Press, 1966.
- [21] S. Johnson, "A new upper bound for error-correcting codes," *IRE Trans. Inf. Theory*, vol. 8, no. 3, pp. 203–207, April 1962.
- [22] —, "Upper bounds for constant weight error correcting codes," *Discrete Math.*, vol. 3, no. 1-3, pp. 109–124, 1972.
- [23] M. Best *et al.*, "Bounds for binary codes of length less than 25," *IEEE Trans. Inf. Theory*, vol. 24, no. 1, pp. 81–93, Jan. 1978.
- [24] M. Plotkin, "Binary codes with specified minimum distance," *IRE Trans. Inform. Theory*, vol. 6, no. 4, pp. 445–450, Sep. 1960.
- [25] R. P. Stanley, "Log-concave and unimodal sequences in algebra, combinatorics, and geometry," *Ann. N.Y. Acad. Sci.*, vol. 576, no. 1, pp. 500–535, 1989.
- [26] R. Graham and N. Sloane, "Lower bounds for constant weight codes," *IEEE Trans. Inf. Theory*, vol. 26, no. 1, pp. 37–43, Jan. 1980.
- [27] M. Bona, *Combinatorics of Permutations*. Boca Raton, FL, USA: CRC Press, Inc., 2004.
- [28] R. G. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley and Sons, Inc., 1968.
- [29] W. Rudin, *Principles of mathematical analysis*, ser. International series in pure and applied mathematics. McGraw-Hill, 1964.

Anshoo Tandon (S'13–M'17) received the B.E. degree in Computer Science and Engineering from Kumaun University, Nainital, India, in 1998, the M.E. degree in Signal Processing from the Indian Institute of Science, Bangalore, India, in 2000, and the Ph.D. degree from the National University of Singapore (NUS), Singapore, in 2016. Between 2000 and 2011, he worked in different capacities in the industry towards developing efficient cellular and wireless connectivity solutions.

Dr. Tandon is currently a Research Fellow in the department of Electrical and Computer Engineering at NUS, Singapore. His research interests include information theory, queueing theory, and algebra.

Han Mao Kiah received his Ph.D. degree in mathematics from Nanyang Technological University (NTU), Singapore, in 2014. From 2014 to 2015 he was a Postdoctoral Research Associate at the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign. Currently he is a Lecturer at the School of Physical and Mathematical Sciences, NTU, Singapore. His research interests include combinatorial design theory, coding theory, and enumerative combinatorics.

Mehul Motani (S'92–M'01–SM'15–F'18) received the B.E. degree from Cooper Union, New York, NY, the M.S. degree from Syracuse University, Syracuse, NY, and the Ph.D. degree from Cornell University, Ithaca, NY, all in Electrical and Computer Engineering.

Dr. Motani is currently an Associate Professor in the Electrical and Computer Engineering Department at the National University of Singapore (NUS) and a Visiting Research Collaborator at Princeton University. Previously, he was a Visiting Fellow at Princeton University. He was also a Research Scientist at the Institute for Infocomm Research in Singapore, for three years, and a Systems Engineer at Lockheed Martin in Syracuse, NY for over four years. His research interests include information theory, machine learning, and wireless & sensor networks.

Dr. Motani was the recipient of the Intel Foundation Fellowship for his Ph.D. research, the NUS Annual Teaching Excellence Award, the NUS Faculty of Engineering Innovative Teaching Award, and the NUS Faculty of Engineering Teaching Honours List Award. He is a Fellow of the IEEE and has served as the Secretary of the IEEE Information Theory Society Board of Governors. He has served as an Associate Editor for both the IEEE Transactions on Information Theory and the IEEE Transactions on Communications. He has also served on the Organizing and Technical Program Committees of numerous IEEE and ACM conferences.