Distance bounds for an ensemble of LDPC convolutional codes

Sridharan, Arvind; Truhachev, Dmitri; Lentmaier, Michael; Costello Jr., Daniel J.; Zigangirov, Kamil

Published in:
IEEE Transactions on Information Theory

DOI:
10.1109/TIT.2007.909113

2007

Distance Bounds for an Ensemble of LDPC Convolutional Codes

Arvind Sridharan, *Member, IEEE,* Dmitri Truhachev, *Member, IEEE,* Michael Lentmaier *Member, IEEE,* Daniel J. Costello, Jr., *Fellow, IEEE,* and Kamil Sh. Zigangirov, *Fellow, IEEE*

***Abstract*— An ensemble of (J, K)-regular LDPC convolutional codes is introduced and existence-type lower bounds on the minimum distance $d_{\rm L}$ of code segments of finite length L and on the free distance $d_{\rm free}$ are derived. For sufficiently large constraint lengths ν, the distances are shown to grow linearly with ν and the ratio $d_{\rm L}/\nu$ approaches the ratio $d_{\rm free}/\nu$ for large L. Moreover, the ratio of free distance to constraint length is several times larger than the ratio of minimum distance to block length for Gallager's ensemble of (J, K)-regular LDPC block codes.**

***Index Terms*— low-density parity check (LDPC) codes, LDPC Convolutional Codes, free distance lower bounds, minimum distance lower bounds**

I. Introduction

LDPC block codes were first introduced by Gallager in [1]. Specifically, Gallager considered block codes described by binary parity-check matrices having J ones in each column and K ones in each row. We refer to LDPC block codes with this property as (J, K)-regular LDPC block codes. The convolutional counterpart of LDPC block codes, LDPC convolutional codes, was first proposed by Tanner in a 1981 patent application [2] and specific constructions were independently described in [3]. Other constructions for LDPC convolutional codes have been presented in [4]-[5]. Both variants of LDPC codes, block and convolutional, are defined by sparse parity-check matrices and can be decoded iteratively with computational complexity per bit per iteration independent of block/constraint length.

LDPC convolutional codes have some advantages in comparison with LDPC block codes, especially for transmitting streaming data [6]. Another desirable feature (for example in Ethernet applications [7]) of LDPC convolutional codes is that the same encoder can be used to obtain a sequence of codes of varying frame lengths with very good performance. Implementation aspects of LDPC convolutional codes, including termination, are discussed in [7]-[8]. It has also been proved, using the same ensemble described here, that (J, K)-regular LDPC convolutional codes have better iterative decoding convergence thresholds than comparable (J, K)-regular LDPC block codes [9][10].

We consider a class of (J, K)-regular LDPC convolutional codes with parity-check matrices (or, equivalently, syndrome formers) composed of blocks of $M \times M$ permutation matrices. These codes are the convolutional counterparts of the (J, K)-regular LDPC block codes introduced in Appendix B of [1] and in [11][12].

Encoding and decoding are carried out on blocks of symbols (the number of symbols in a block depends on M). The code structure makes an analysis of distance properties, similar to that carried out in [12] for (J, K)-regular LDPC block codes, possible.

One way of characterizing the strength of a block code is its minimum distance $d_{\rm min}$. The well known (asymptotic) Gilbert-Varshamov (GV) bound [13][14] guarantees, for sufficiently large block lengths N, the existence of linear block codes of rate $R, 0 < R < 1$, whose minimum distance is lower bounded by a linear function of N, i.e., $d_{\rm min} \geq \alpha_{\rm GV}(R)N$, where $\alpha_{\rm GV}(R)$ is the GV coefficient. Analogously, Gallager proved the existence of (J, K)-regular LDPC block codes $(J > 2)$ satisfying the inequality $d_{\rm min} \geq \alpha_{\rm G}(J, K)N$ for sufficiently large block lengths N [1]. The coefficient $\alpha_{\rm G}(J, K)$ can be calculated numerically. For practically interesting J and K, $\alpha_{\rm G}(J, K)$ is several times smaller than the corresponding GV coefficient $\alpha_{\rm GV}(1 - \frac{J}{K})$. (Note that (J, K)-regular LDPC codes typically have rate $R \approx 1 - J/K$.)

The convolutional counterpart of minimum distance is free distance and the corresponding analog of the GV bound is the Costello bound [15]. Costello proved the existence of convolutional codes of rate $R, 0 \leq R < 1$, with free distance increasing linearly with constraint length, i.e., $d_{\rm free} \geq \alpha_{\rm C}(R)\nu$, for sufficiently large constraint lengths ν. For rate $R = \frac{1}{2}$ codes, the coefficient $\alpha_{\rm C}(R)$ is about three and a half times larger than $\alpha_{\rm GV}(R)$.

In [16], the distance spectrum of a special ensemble of (J, K)-regular LDPC convolutional codes based on Markov permutors with $K = 2J$ was analyzed, and a technique to numerically calculate the distance spectrum of the codes in the ensemble as a function of constraint length was described. The results obtained in [16] suggest the existence of LDPC

This work was supported in part by NSF Grants CCR-02-05310 and CCF05-15012, NASA Grant NAG5-12792, and the Alberta Ingenuity Fund. Some of the material in this paper was previously presented at ISIT 2004 (Chicago, USA).

A. Sridharan was with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556. He is now with Seagate Technologies, 389 Disc Drive, Longmont, CO 80503.

D. Truhachev is with the Electrical and Computer Engineering Research Faculty (ECERF), University of Alberta, Edmonton, Alberta, Canada T6G 2V4

M. Lentmaier was with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556. He is now with the German Aerospace Center (DLR), Institute of Communications and Navigation, Oberpfaffenhofen, P.O. Box 1116 D-82234, Wessling, Germany

Daniel J. Costello, Jr., is with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556, USA.

K. Sh. Zigangirov is with the Institute for Problems of Information Transmission, Moscow, Russia, and the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556, USA.

convolutional codes with free distance increasing linearly with constraint length.

In this paper existence-type lower bounds on the minimum weight of code segments of length L, the Lth order segment distance d_{L}, of (J,K)-regular LDPC convolutional codes $(J>2)$ in a permutation-matrix-based ensemble are derived. Moreover, for the same code ensemble, we derive an existence-type lower bound on the free distance d_{free}. In particular, we prove that the Lth order segment distance is lower bounded by the inequality $d_{\mathrm{L}} \geq (\rho^*_{\mathrm{L}}/6)\nu$ for sufficiently large constraint lengths ν. The values ρ^*_{L} are decreasing with L and for any L are lower bounded by $\rho^* = 0.5$. We then prove that the free distance of the codes satisfies $d_{\mathrm{free}} \geq (\rho^*/6)\nu \stackrel{\mathrm{def}}{=} \alpha_{\mathrm{LDPCC}}(J,K)\nu$[1]. Numerical results indicate that, for practically interesting J and K, the coefficient $\alpha_{\mathrm{LDPCC}}(J,K)$ is several times smaller than the corresponding Costello coefficient $\alpha_{\mathrm{C}}(1-\frac{J}{K})$, where the convolutional code rate $R=1-J/K$. This parallels the result of Gallager for (J,K)-regular LDPC block codes relative to the GV coefficient $\alpha_{\mathrm{GV}}(R)$.

Consider, for example, the $(3,6)$ case. Based on numerical evaluation of the bound on free distance, we find that $\alpha_{\mathrm{LDPCC}}(3,6) = 0.083$. This is weaker than the Costello coefficient $\alpha_{\mathrm{C}}(1/2) = 0.39$ for rate $R=1/2$ codes. However, $\alpha_{\mathrm{LDPCC}}(3,6)$ is about three and a half times larger than the Gallager coefficient $\alpha_{\mathrm{G}}(3,6) = 0.023$ for $(3,6)$-regular LDPC block codes. This essentially mimics the relationship between the Costello bound (for convolutional codes) and the GV bound (for block codes) noted above.

The analysis and bounding techniques used here are significantly different from the traditional techniques for lower bounding the free distance of conventional convolutional codes [15][17]. The traditional techniques rely on the fact that the weight at the beginning and the end of a code sequence increases with constraint length. However, the ensemble of (J,K)-regular LDPC convolutional codes we investigate has code sequences with negligible weight at either end. In fact, most code sequences have their weight concentrated in the middle. This fact significantly complicates the analysis of these codes.

The paper is organized as follows. We start with the code ensemble description in Section II. Section III presents the main results, formulated in terms of two theorems: a segment distance bound and a free distance bound. The theorems are proved in Sections IV and V, respectively. A discussion of the results is given in Section VI, and Section VII offers some concluding remarks.

II. An LDPC Convolutional Code Ensemble

A rate $R=b/c$ binary convolutional code can be defined as the set of sequences $\boldsymbol{v} = (\ldots, \boldsymbol{v}_{-1}, \boldsymbol{v}_0, \boldsymbol{v}_1, \ldots),\ \boldsymbol{v}_t \in \mathbb{F}_2^c$, satisfying the equality $\boldsymbol{v}\boldsymbol{H}^{\mathrm{T}} = \boldsymbol{0}$, where the infinite syndrome former $\boldsymbol{H}^{\mathrm{T}}$ is given by

$$\boldsymbol{H}^{\mathrm{T}} = \begin{pmatrix} \ddots & & & & & \\ & \boldsymbol{H}^{\mathrm{T}}_{m_{\mathrm{s}}}(-1) & & & & \\ \ddots & \vdots & \boldsymbol{H}^{\mathrm{T}}_{m_{\mathrm{s}}}(0) & & & \\ \ddots & \boldsymbol{H}^{\mathrm{T}}_{1}(-1) & \vdots & \boldsymbol{H}^{\mathrm{T}}_{m_{\mathrm{s}}}(1) & & \\ & \boldsymbol{H}^{\mathrm{T}}_{0}(-1) & \boldsymbol{H}^{\mathrm{T}}_{1}(0) & \vdots & \ddots & \\ & & \boldsymbol{H}^{\mathrm{T}}_{0}(0) & \boldsymbol{H}^{\mathrm{T}}_{1}(1) & & \\ & & & \boldsymbol{H}^{\mathrm{T}}_{0}(1) & \ddots & \\ & & & & \ddots & \end{pmatrix}, \quad (1)$$

and each $\boldsymbol{H}^{\mathrm{T}}_i(t)$ is a $c\times(c-b)$ binary matrix, $i=0,1,\ldots,m_{\mathrm{s}}$, $t\in\mathbb{Z}$. If $\boldsymbol{H}^{\mathrm{T}}$ defines a rate $R=b/c$ convolutional code, the matrix $\boldsymbol{H}^{\mathrm{T}}_0(t)$ must have full rank for all time instants t. In this case, by suitable row permutations, we can ensure that the last $(c-b)$ rows are linearly independent. Then the first b symbols at each time instant are information symbols and the last $(c-b)$ symbols are the corresponding parity symbols. The largest i such that $\boldsymbol{H}^{\mathrm{T}}_i(t)$ is a non-zero matrix for some t is called the syndrome former memory m_s.

LDPC convolutional codes have sparse syndrome formers. A (J,K)-regular LDPC convolutional code is defined by a syndrome former that contains exactly J ones in each row and K ones in each column.

Let $a = \gcd(J,K)$ denote the greatest common divisor of J and K. Then there exist positive integers J' and K' such that $J=aJ'$ and $K=aK'$ and $\gcd(J',K')=1$. We consider (J,K)-regular LDPC convolutional codes defined by syndrome formers $\boldsymbol{H}^{\mathrm{T}}$ with syndrome former memory $a-1$. For $i=0,1,\ldots,a-1$, the sub-matrices $\boldsymbol{H}^{\mathrm{T}}_i(t)$ of the syndrome former are

$$\boldsymbol{H}^{\mathrm{T}}_i(t) = \begin{pmatrix} \boldsymbol{P}_i^{(0,0)}(t) & \boldsymbol{P}_i^{(0,1)}(t) & \ldots & \boldsymbol{P}_i^{(0,J'-1)}(t) \\ \boldsymbol{P}_i^{(1,0)}(t) & \boldsymbol{P}_i^{(1,1)}(t) & \ldots & \boldsymbol{P}_i^{(1,J'-1)}(t) \\ \ldots & \ldots & \ldots & \ldots \\ \boldsymbol{P}_i^{(K'-1,0)}(t) & \boldsymbol{P}_i^{(K'-1,1)}(t) & \ldots & \boldsymbol{P}_i^{(K'-1,J'-1)}(t) \end{pmatrix},$$

where each $\boldsymbol{P}_i^{(k,j)}(t)$, $k=0,1,\ldots,K'-1$, $j=0,1,\ldots,J'-1$, is an $M\times M$ permutation matrix. All other entries of the syndrome former are zero matrices. Equivalently, each $\boldsymbol{H}^{\mathrm{T}}_i(t)$, $i=0,1,\ldots,a-1$, is a $c\times(c-b)$ binary matrix where $c=K'M$ and $b=(K'-J')M$. By construction, it follows that each row of the syndrome former $\boldsymbol{H}^{\mathrm{T}}$ has J ones and each column K ones. Let $\mathcal{C}_P(J,K,M)$ denote the ensemble of (J,K)-regular LDPC convolutional codes obtained by choosing each $M\times M$ permutation matrix in $\boldsymbol{H}^{\mathrm{T}}$ independently and such that each of the $M!$ possible permutation matrices is equally likely. (Analogous to the codes introduced in [3], the codes in the ensemble $\mathcal{C}_P(J,K,M)$ are time-varying, but in contrast to [3], they are, generally speaking, non-periodic.) Fig. 1 shows the syndrome former of a $(3,6)$-regular LDPC convolutional code in $\mathcal{C}_P(3,6,M)$.

The syndrome formers in the ensemble $\mathcal{C}_P(J,K,M)$ have syndrome former memory $m_{\mathrm{s}} = a-1$ independent of M

[1] In addition to providing the intuitively pleasing result that its limit for large L approaches the free distance bound, the segment distance bound is interesting in its own right as an indicator of the asymptotic performance of LDPC convolutional codes that are decoded over a window of finite length (see, for example, the pipeline decoder described in [3]).

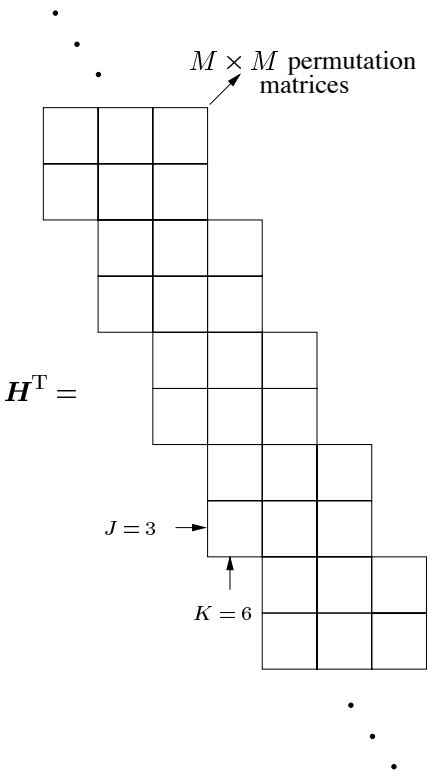

Fig. 1. Syndrome former for a code in $C_P(3, 6, M)$.

while b and c depend on M. This is different from the LDPC convolutional codes considered in [3]–[5], where the codes have varying syndrome former memories m_{s}, while b and c are fixed. For the ensemble $\mathcal{C}_P(J, K, M)$, as M increases, i.e., as b and c increase, the syndrome formers become increasingly sparse.

By virtue of their sparse parity-check matrices, the codes in $\mathcal{C}_P(J, K, M)$ can be iteratively decoded using message passing algorithms (e.g., belief propagation), and decoding can be scheduled so as to obtain a continuous time pipeline decoder [3]. At each time instant a block of $c = K'M$ received symbols is input to the decoder and $b = RK'M$ information symbols are decoded and output from the decoder, where R is the code rate.

For the ensemble $\mathcal{C}_P(3, 6, M)$, the matrices $\boldsymbol{H}_i^{\mathrm{T}}(t)$ consist of two $M \times M$ permutation matrices, denoted $\boldsymbol{P}_i^{(0)}(t)$ and $\boldsymbol{P}_i^{(1)}(t)$, and hence have rank equal to M, i.e., the code rate is $R = M/2M$. In this case, by permuting rows of the syndrome former an equivalent rate $R = 1/2$ ($b = 1, c = 2$) code with syndrome former memory at most $3M - 1$ can be obtained (see Fig. 2). Since distance properties are unaffected by row permutations, the distance bounds obtained for codes in the ensemble $\mathcal{C}_P(3, 6, M)$ are also valid for the equivalent $b = 1, c = 2$ codes.

In general, however, there are at least $J' - 1$ dependent columns in $\boldsymbol{H}_0^{\mathrm{T}}(t)$ for any code in $\mathcal{C}_P(J, K, M)$. Hence, $\boldsymbol{H}^{\mathrm{T}}$ defines a rate $R \geq 1 - \frac{J'M-(J'-1)}{K'M}$ code. The constraint length[2] of codes of $\mathcal{C}_P(J, K, M)$ is defined as

$$\nu = (m_{\mathrm{s}} + 1)c = aK'M = KM \ .$$

Thus the codes in the ensemble $\mathcal{C}_P(3, 6, M)$ have constraint length $6M$.

The syndrome formers of the (J, K)-regular LDPC convolutional code ensemble described above have a structure similar to that of the permutation-matrix-based (J, K)-regular LDPC block code ensemble described in [12]. The parity-check matrices of the codes in the ensemble considered in [12] are composed of $J \times K$ permutation matrices, where each permutation matrix is of size $M \times M$. Thus the parity-check matrices are of size $JM \times KM$ and have exactly K ones in each row and J ones in each column. This ensemble is a vanishingly small sub-ensemble of Gallager's original ensemble [1].

If $a = \gcd(J, K) = 1$, i.e., J and K are relatively prime, then the LDPC convolutional codes in the ensemble $\mathcal{C}_P(J, K, M)$ have syndrome former memory $m_{\mathrm{s}} = a - 1 = 0$. The ensemble of memory zero (J, K)-regular LDPC convolutional codes so obtained is equivalent to the block code ensemble considered in [12]. In [12], we show that asymptotically, i.e., as the block length $N = KM \to \infty$, almost all codes in the ensemble have minimum distance satisfying Gallager's bound[3], i.e., $d_{\min} \geq \alpha_{\mathrm{G}}(J, K)N$.

A probability distribution is defined on the ensemble $\mathcal{C}_P(J, K, M)$ as follows. Assume that all of the permutation matrices comprising the syndrome former of a code in $\mathcal{C}_P(J, K, M)$ are chosen independently and such that each of the $M!$ possible permutation matrices is equally likely.

III. Lower Bounds on Segment Distance and Free Distance

We seek a lower bound on the minimum weight of code sequences having a non-zero segment of length at most L, i.e., we lower bound the Lth order segment distance d_{L}. To calculate or lower bound the Lth order segment distance for the class of periodically time-varying codes, it is sufficient to consider code sequences with starting positions within one period (see [17]). However, in the most general case of non-periodically time-varying codes, such as codes from the ensemble $\mathcal{C}_P(J, K, M)$, all possible starting positions must be considered. This complicates the analysis.

To avoid cumbersome notation, we henceforth focus on the $(3, 6)$ case, i.e., the ensemble $\mathcal{C}_P(3, 6, M)$, though the same technique can also be used more generally.

Consider sequences

$$\boldsymbol{v}_{[i+1,i+L]} = (\ldots, \boldsymbol{0}, \boldsymbol{v}_{i+1}^{(0)}, \boldsymbol{v}_{i+1}^{(1)}, \ldots, \boldsymbol{v}_{i+L}^{(0)}, \boldsymbol{v}_{i+L}^{(1)}, \boldsymbol{0}, \ldots) \ ,$$

[2]From (1), we see that $(m_{\mathrm{s}} + 1)c$ is the total number of code symbols involved in the parity-check constraints at any time instant t. This also corresponds to the total number of encoder output symbols that directly depend on a given block of b information symbols (see, e. g., [8]). Finally, the pipeline decoder described in [3] requires a processor that can exchange messages among $(m_{\mathrm{s}} + 1)c$ code symbols. Thus it makes sense to define the constraint length of an LDPC convolutional code as $\nu = (m_{\mathrm{s}} + 1)c$, and we note that constraint length for LDPC convolutional codes plays a role analogous to block length for LDPC block codes.

[3]This holds for small values of J and K.

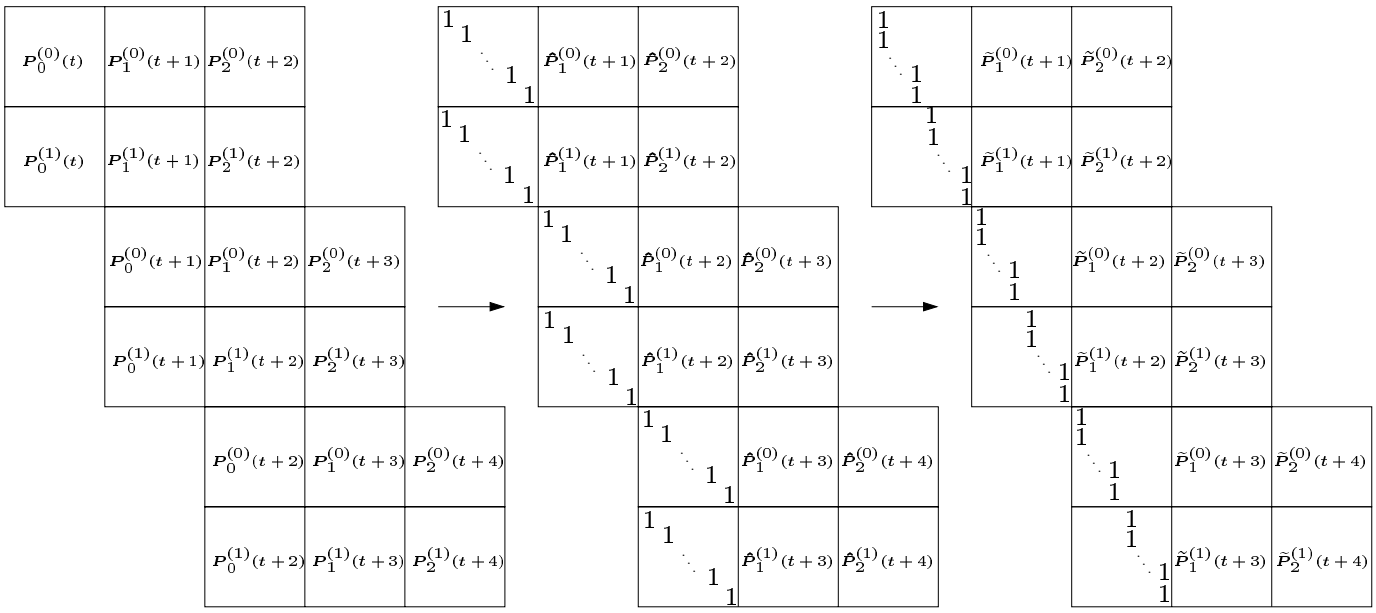

Fig. 2. Permuting rows of a syndrome former $\boldsymbol{H}^{\mathrm{T}}$ in $\mathcal{C}_P(3,6,M)$.

$\boldsymbol{v}_j^{(h)} \in \mathbb{F}_2^M$, for $j = i+1, i+2, \ldots, i+L$, $h = 0,1$, with $\boldsymbol{v}_{[i+1,i+L]} \neq \boldsymbol{0}$, i.e., sequences with a non-zero segment of length at most L. The sequence $\boldsymbol{v}_{[i+1,i+L]}$ consists of at most $2LM$ non-zero binary symbols. If $\boldsymbol{v}_{[i+1,i+L]}$ is a code segment, then at each time instant t, $i < t < (i+L+1)$, $\boldsymbol{v}_{[i+1,i+L]}$ contains an information block $\boldsymbol{v}_t^{(0)}$ of length M and the corresponding length M parity block $\boldsymbol{v}_t^{(1)}$.

For convenience we first define a segment distance measure related to the starting position. The "local" Lth order segment distance at time i, $d_{\mathrm{L}}(i)$, of a code in $\mathcal{C}_P(3,6,M)$ is defined as the minimum weight over code segments of the form $\boldsymbol{v}_{[i+1,i+L]}$, $\boldsymbol{v}_{i+1}^{(0)} \neq \boldsymbol{0}$. Observe that $\boldsymbol{v}_{i+1}^{(0)} \neq \boldsymbol{0}$ implies $\boldsymbol{v}_{i+1}^{(1)} \neq \boldsymbol{0}$. The Lth order segment distance d_{L} of a code in $\mathcal{C}_P(3,6,M)$ is defined as $\min_i d_{\mathrm{L}}(i)$. Note that the segment distance is a non-increasing function of L.

This definition of the Lth order segment distance associated with a syndrome former is analogous to the traditional definition of the Lth order row distance associated with an encoding matrix [17]. In particular, for a time-invariant or periodically time-varying convolutional encoder, the Lth order row distance is defined as the minimum weight of code sequences having a non-zero segment of length at most $L+m+1$, where m is the encoder memory [17]. The symbols in the last m time instants of the encoder input sequence are determined so as to force the encoder to the zero state.

We note that row distance is an encoder property, whereas our definition of segment distance is a code property.[4] However, as with the definition of row distance, our definition of segment distance also looks at weight properties of finite length sequences. Further, the sequences used to determine segment distance correspond to a row-truncated syndrome former. This is similar to the traditional case, where the row distance is calculated by considering sequences obtained from a row-truncated generator matrix.

We are now ready to state and prove the main results of the paper, given by the following two theorems.[5]

Theorem 1: For any L and any starting position i, there exists an M_L such that for any $M > M_L$, there exists a code in $\mathcal{C}_P(3,6,M)$ with local Lth order segment distance $d_{\mathrm{L}}(i)$ lower bounded by

$$d_{\mathrm{L}}(i) \geq (\rho_{\mathrm{L}}^*/6)\nu, \tag{2}$$

where $\nu = 6M$ is the constraint length of the code and ρ_L^* is given in (26). Further, for any L, there exists an M_L such that for any $M > M_L$, there exists a code in $\mathcal{C}_P(3,6,M)$ with Lth order segment distance d_L lower bounded by

$$d_{\mathrm{L}} \geq (\rho_{\mathrm{L}}^*/6)\nu. \tag{3}$$

□

Numerical techniques are needed to solve the max-min problem in (26) to evaluate ρ_{L}^* for a given L. We were able to obtain ρ_{L}^* for values of L up to $L = 16$. On the other hand, we prove in Appendix IV that for any L the value $\rho_{\mathrm{L}}^* \geq \rho^* = 0.5$. For $L = 16$, the coefficient $\rho_L^*/6 = 0.085$, which is only slightly higher than $\rho^*/6 = 0.083$.

The main idea of the proof is outlined as follows. First we prove (2) and show that the fraction of codes in the ensemble $\mathcal{C}_P(3,6,M)$ with $d_{\mathrm{L}}(i) < (\rho_{\mathrm{L}}^*/6)\nu$ tends to zero with increasing M. The key step is to obtain the probability that a segment $\boldsymbol{v}_{[i+1,i+L]}$ is a valid code segment. The parity-check matrices of codes in $\mathcal{C}_P(3,6,M)$ are comprised of blocks of independently chosen $M \times M$ permutation matrices. Hence this probability can be calculated using a technique similar to the one described in [12]. As we shall see, the probability depends not only on the overall weight of the sequence $\boldsymbol{v}_{[i+1,i+L]}$ but also on the weight of the individual

[4]We would like to thank an anonymous referee for pointing this out and motivating the comparison between our definition of segment distance and the traditional definition of row distance.

[5]For simplicity we state the theorems only for the $(3,6)$ case.

components $\boldsymbol{v}_t^{(h)},\ t = i+1,\ldots,i+L, h = 0,1$. This is a key difference compared to the traditional lower bounding techniques for convolutional codes ([15][17]) and significantly complicates the proof. Section IV presents the proof of (2), the lower bound on local segment distance.

In order to extend the bound of (2) for the local Lth order segment distance to obtain the general Lth order segment distance bound of (3), we use a special *expurgation procedure*. We expurgate code sequences leading to a local Lth order segment distance less than $(\rho_{\mathrm{L}}^*/6)\nu$ by fixing some information symbols to be zero. This lowers the code rate but the loss in rate tends to zero as M tends to infinity. The expurgation procedure is explained in more detail at the end of Section IV.

The second theorem provides a lower bound on the free distance of codes in $\mathcal{C}_P(3,6,M)$. The proof is not based on bounding the segment distance and considers instead different sets of low weight sequences. Evaluating the segment distance bounds requires considering sequences of a fixed length, possibly merging and diverging from the all-zero state several times. However, for the free distance bound it is sufficient to look at "detours", i.e., sequences diverging from the all-zero state exactly once. For long sequences, the latter number is significantly smaller than the former. Note also that it is the overall weight of the sequence, regardless of its length, that is of interest for evaluating the free distance. As in Theorem 1, we start by bounding the "local" free distance and use the same expurgation technique mentioned above to show that the result is valid globally.

Theorem 2: There exists an M_f such that for any $M > M_f$, there exists a code in the ensemble $\mathcal{C}_P(3,6,M)$ with free distance d_{free} lower bounded by

$$d_{\text{free}} \geq (\rho^*/6)\nu \stackrel{\text{def}}{=} \alpha_{\text{LDPCC}}(3,6)\nu\ . \tag{4}$$

□

As noted above, we prove in Appendix IV that $\rho_{\mathrm{L}}^* \geq \rho^* = 0.5$ for any L, and we observe from numerical calculations that the ratio ρ_{L}^*/ν approaches the ratio ρ^*/ν as L goes to infinity. It also follows from the definition of segment distance that $\lim_{L\to\infty} d_{\mathrm{L}} = d_{\text{free}}$ for a given code. However, these facts along with (3) do not imply (4), since for every L, (3) is valid only for M greater than some M_L, but there is no fixed M_L, and thus no fixed constraint length ν, for which (3) holds for all $L = 1,2,\ldots$. Thus a separate proof is required for Theorem 2. Section V presents the proof of (4), the lower bound on free distance.

IV. Proof of Theorem 1

We first seek a lower bound on $d_{\mathrm{L}}(i)$ for a particular starting position i. Without loss of generality, we can investigate code sequences starting at time $t = 1$, i.e., of the form $\boldsymbol{v}_{[1,L]}$ and obtain a lower bound on $d_{\mathrm{L}}(0)$. For a segment $\boldsymbol{v}_{[1,L]}$, let $\boldsymbol{d}_t = (d_t^{(0)}, d_t^{(1)}),\ t = 1,\ldots,L$, where $d_t^{(h)}$ is the Hamming weight of $\boldsymbol{v}_t^{(h)},\ t = 1,\ldots,L,\ h = 0,1$. Define the $2LM \times (L+2)M$ matrix $\boldsymbol{H}_{[1,L]}^{\mathrm{T}}$ as in (5). If the segment $\boldsymbol{v}_{[1,L]}$ is part of a code sequence, then it must satisfy the $(L+2)M$ constraints imposed by the matrix $\boldsymbol{H}_{[1,L]}^{\mathrm{T}}$. Fig. 3 illustrates

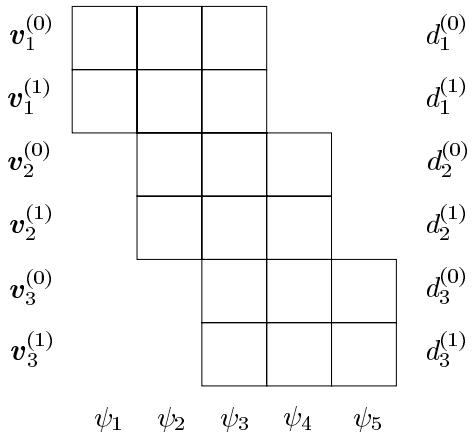

Fig. 3. Illustration of length $L = 3$ segment $\boldsymbol{v}_{[1,3]}$ and $\boldsymbol{H}_{[1,3]}^{\mathrm{T}}$.

segment $\boldsymbol{v}_{[1,3]}$ and the matrix $\boldsymbol{H}_{[1,3]}^{\mathrm{T}}$ for a randomly chosen code in $\mathcal{C}_P(3,6,M)$.

For any t, $1 \leq t \leq L+2$, let ψ_t be the probability in the ensemble $\mathcal{C}_P(3,6,M)$ that $\boldsymbol{v}_{[1,L]}$ satisfies the constraints (M equations) imposed by the sub-matrices $\boldsymbol{H}_i^{\mathrm{T}}(t),\ i = 0,1,2$, of $\boldsymbol{H}_{[1,L]}^{\mathrm{T}}$, i.e., the tth column of $\boldsymbol{H}_{[1,L]}^{\mathrm{T}}$. Let $\psi_{[1,L]}$ be the probability that $\boldsymbol{v}_{[1,L]}$ is a valid code segment. Then

$$\psi_{[1,L]} = \prod_{t=1}^{L+2} \psi_t\ . \tag{6}$$

The sub-matrices $\boldsymbol{H}_i^{\mathrm{T}}(t)$ of the syndrome former of a code in $\mathcal{C}_P(3,6,M)$ are of size $2M \times M$ and are composed of two $M \times M$ permutation matrices. Hence each of the terms ψ_t, $t = 1,2,\ldots,L+2$, can be calculated and bounded by following the technique introduced in [12]. The details of this technique can be found in Appendix I.

Let $\boldsymbol{\rho}_t = (\rho_t^{(0)}, \rho_t^{(1)}),\ t = 1,\ldots,L$, where $\rho_t^{(h)} = d_t^{(h)}/M$ is the normalized Hamming weight of $\boldsymbol{v}_t^{(h)},\ t = 1,\ldots,L$, $h = 0,1$, and let $\boldsymbol{\lambda}_i(t) = (\lambda_i^{(0)}(t), \lambda_i^{(1)}(t))$, $t = 2,\ldots,L+1$, $i = 0,1,2$, where $\lambda_i^{(0)}(t)$ and $\lambda_i^{(1)}(t)$ are arbitrary constants.

Further, let $\boldsymbol{\lambda}_{[1,I]} = (\lambda_1, \lambda_2, \ldots, \lambda_I)$ and $\boldsymbol{d}_{[1,I]} = (d_1, d_2, \ldots, d_I)$, and $\boldsymbol{\rho}_{[1,I]} = (\rho_1, \rho_2, \ldots, \rho_I)$, where λ_i is an arbitrary constant, d_i is a Hamming weight, $\rho_i = d_i/M$ is a normalized Hamming weight, and $i = 1,\ldots,I$ is a time index, and define the function $\boldsymbol{G}_I(\boldsymbol{\lambda}_I, \boldsymbol{\rho}_I)$ as

$$G_I(\boldsymbol{\lambda}_{[1,I]}, \boldsymbol{\rho}_{[1,I]}) = g_I(\lambda_1, \lambda_2, \ldots, \lambda_I) - \sum_{i=1}^{I} \lambda_i \rho_i, \tag{7}$$

where

$$g_I(\lambda_1, \lambda_2, \ldots, \lambda_I) \stackrel{\triangle}{=} \ln \frac{\prod_{i=1}^{I}(1+e^{\lambda_i}) + \prod_{i=1}^{I}(1-e^{\lambda_i})}{2}. \tag{8}$$

In Appendix I we derive the probability $\gamma_I(\boldsymbol{d}_{[1,I]})$ that an IM dimensional vector, $I \geq 2$, with a given weight composition $\boldsymbol{d}_{[1,I]}$, satisfies a set of M parity check constraints imposed by I permutation matrices. For $I = 2$, this probability is given by (72), and for $I > 2$ it is upper bounded by (76). It follows from the definition of ψ_t that $\psi_1 = \gamma_2(\boldsymbol{\rho}_1 M)$,

$$\boldsymbol{H}_{[1,L]}^{\mathrm{T}} = \begin{pmatrix} \boldsymbol{H}_0^{\mathrm{T}}(1) & \boldsymbol{H}_1^{\mathrm{T}}(2) & \boldsymbol{H}_2^{\mathrm{T}}(3) & & & & \\ & \boldsymbol{H}_0^{\mathrm{T}}(2) & \boldsymbol{H}_1^{\mathrm{T}}(3) & \boldsymbol{H}_2^{\mathrm{T}}(4) & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \boldsymbol{H}_0^{\mathrm{T}}(L-1) & \boldsymbol{H}_1^{\mathrm{T}}(L) & \boldsymbol{H}_2^{\mathrm{T}}(L+1) & \\ & & & & \boldsymbol{H}_0^{\mathrm{T}}(L) & \boldsymbol{H}_1^{\mathrm{T}}(L+1) & \boldsymbol{H}_2^{\mathrm{T}}(L+2) \end{pmatrix} \tag{5}$$

$\psi_2 = \gamma_4(\boldsymbol{\rho}_2 M, \boldsymbol{\rho}_1 M)$, $\psi_t = \gamma_6(\boldsymbol{\rho}_t M, \boldsymbol{\rho}_{t-1} M, \boldsymbol{\rho}_{t-2} M)$ for $t = 3, \ldots, L$, $\psi_{L+1} = \gamma_4(\boldsymbol{\rho}_L M, \boldsymbol{\rho}_{L-1} M)$, and $\psi_{L+2} = \gamma_2(\boldsymbol{\rho}_L M)$. Hence, (72) and (76) imply that ψ_t satisfies

$$\psi_1 = \frac{1}{\binom{M}{M\rho_1^{(0)}}} \tag{9}$$

$$\psi_{L+2} = \frac{1}{\binom{M}{M\rho_L^{(0)}}} \tag{10}$$

$$\psi_2 \le \frac{\exp\left[M G_4(\boldsymbol{\lambda}_0(2), \boldsymbol{\lambda}_1(2), \boldsymbol{\rho}_2, \boldsymbol{\rho}_1)\right]}{\prod_{t=1}^{2}\prod_{h=0}^{1}\binom{M}{M\rho_t^{(h)}}} \tag{11}$$

$$\psi_{L+1} \le \frac{\exp\left[M G_4(\boldsymbol{\lambda}_1(L+1), \boldsymbol{\lambda}_2(L+1), \boldsymbol{\rho}_L, \boldsymbol{\rho}_{L-1})\right]}{\prod_{t=L-1}^{L}\prod_{h=0}^{1}\binom{M}{M\rho_t^{(h)}}} \tag{12}$$

$$\psi_t \le \frac{\exp\left[M G_6(\boldsymbol{\lambda}_0(t), \boldsymbol{\lambda}_1(t), \boldsymbol{\lambda}_2(t), \boldsymbol{\rho}_t, \boldsymbol{\rho}_{t-1}, \boldsymbol{\rho}_{t-2})\right]}{\prod_{i=t-2}^{t}\prod_{h=0}^{1}\binom{M}{M\rho_i^{(h)}}}, \quad t \ne 1, 2, L+1, L+2. \tag{13}$$

For any code sequence $\boldsymbol{v}_{[1,L]}$ we must have $(\boldsymbol{v}_1^{(0)}, \boldsymbol{v}_1^{(1)})\boldsymbol{H}_0^{\mathrm{T}}(1) = 0$ and $(\boldsymbol{v}_L^{(0)}, \boldsymbol{v}_L^{(1)})\boldsymbol{H}_2^{\mathrm{T}}(L+2) = 0$. Hence it follows that $\rho_1^{(0)} = \rho_1^{(1)}$ and $\rho_L^{(0)} = \rho_L^{(1)}$, i.e., the first two and last two blocks of the code segment $\boldsymbol{v}_{[1,L]}$ have the same weight.

Now let

$$\boldsymbol{\lambda}_{[1,L]} = (\boldsymbol{\lambda}_0(2), \boldsymbol{\lambda}_1(2), \boldsymbol{\lambda}_0(3), \boldsymbol{\lambda}_1(3), \boldsymbol{\lambda}_2(3), \ldots, \boldsymbol{\lambda}_0(L), \boldsymbol{\lambda}_1(L), \boldsymbol{\lambda}_2(L), \boldsymbol{\lambda}_1(L+1), \boldsymbol{\lambda}_2(L+1))$$

be a $(6L-4)$-dimensional vector of arbitrary constants and $\boldsymbol{\rho}_{[1,L]} = (\boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_L)$ be a $2L$-dimensional vector of normalized Hamming weights[6]. Then from (6) and (9)-(13) it follows that $\psi_{[1,L]}$ can be upper bounded as

$$\psi_{[1,L]} \le \frac{\exp\left[M\widetilde{F}(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})\right]}{\left[\prod_{t=1}^{L}\prod_{h=0}^{1}\binom{M}{M\rho_t^{(h)}}\right]^3}\binom{M}{M\rho_1^{(1)}}\binom{M}{M\rho_L^{(1)}}, \tag{14}$$

where

$$\begin{aligned}\widetilde{F}(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]}) &= G_4(\boldsymbol{\lambda}_0(2), \boldsymbol{\lambda}_1(2), \boldsymbol{\rho}_2, \boldsymbol{\rho}_1) \\ &\quad + \sum_{t=3}^{L} G_6(\boldsymbol{\lambda}_0(t), \boldsymbol{\lambda}_1(t), \boldsymbol{\lambda}_2(t), \boldsymbol{\rho}_t, \boldsymbol{\rho}_{t-1}, \boldsymbol{\rho}_{t-2}) \\ &\quad + G_4(\boldsymbol{\lambda}_1(L+1), \boldsymbol{\lambda}_2(L+1), \boldsymbol{\rho}_L, \boldsymbol{\rho}_{L-1}).\end{aligned} \tag{15}$$

[6]We note that the entries of $\boldsymbol{\lambda}_{[1,L]}$ and $\boldsymbol{\rho}_{[1,L]}$ are now 2-dimensional vectors. Also, in a slight abuse of notation, $\boldsymbol{\lambda}_{[1,L]}$ contains entries for time units 2 through $L+1$, rather than 1 through L.

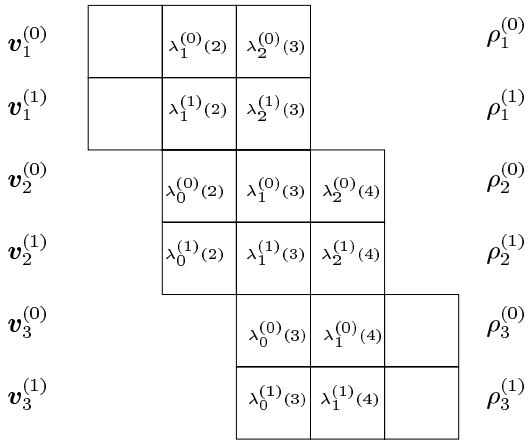

Fig. 4. Set of parameters $\boldsymbol{\lambda}_{[1,3]}$ and $\boldsymbol{\rho}_{[1,3]}$ for calculating the function $\widetilde{F}(\boldsymbol{\lambda}_{[1,3]}, \boldsymbol{\rho}_{[1,3]})$

Fig. 4 shows the parameters which are needed to calculate $\widetilde{F}(\boldsymbol{\lambda}_{[1,3]}, \boldsymbol{\rho}_{[1,3]})$, where it can be seen that $\lambda_i^{(0)}(t)$ and $\lambda_i^{(1)}(t)$ are associated with the permutation matrices $\boldsymbol{P}_i^{(0)}(t)$ and $\boldsymbol{P}_i^{(1)}(t)$ that comprise the matrix $\boldsymbol{H}_i(t)$.

The expected number $E(\boldsymbol{\rho}_{[1,L]})$ of code segments $\boldsymbol{v}_{[1,L]}$ having normalized weight composition $\boldsymbol{\rho}_{[1,L]}$ in a code from the ensemble $\mathcal{C}_P(3,6,M)$ is given by

$$E(\boldsymbol{\rho}_{[1,L]}) = \psi_{[1,L]}\prod_{t=1}^{L}\prod_{h=0}^{1}\binom{M}{M\rho_t^{(h)}}. \tag{16}$$

Now using (14) we can obtain an upper bound on $E(\boldsymbol{\rho}_{[1,L]})$.

Lemma 1: For normalized weight compositions $\boldsymbol{\rho}_{[1,L]}$ with total normalized weight

$$\rho_{[1,L]} \stackrel{\text{def}}{=} w(\boldsymbol{\rho}_{[1,L]}) = \sum_{t=1}^{L}\sum_{h=0}^{1}\rho_t^{(h)} = \frac{d}{M}\ , \tag{17}$$

$E(\boldsymbol{\rho}_{[1,L]})$ is upper bounded by

a)

$$E(\boldsymbol{\rho}_{[1,L]}) \le \exp\left\{M\left(F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]}) + \sum_{\rho_t^{(h)} \ne 0}\frac{\ln(2\pi e \rho_t^{(h)} M)}{M}\right)\right\}, \tag{18}$$

b)

$$E(\boldsymbol{\rho}_{[1,L]}) \le (2\pi e M)^{2L}\exp\left\{M F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})\right\}, \text{ or} \tag{19}$$

c)

$$E(\boldsymbol{\rho}_{[1,L]}) \le \exp\Big\{ M\Big(F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]}) + \rho_{[1,L]} \ln(2\pi e) \Big) \Big\} , \quad (20)$$

where the functions $F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})$ and $H(\rho)$ are given by

$$F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]}) = \tilde{F}(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]}) - 3H(\rho_1^{(0)}) - 2\sum_{t=2}^{L-1}\sum_{h=0}^{1} H(\rho_t^{(h)}) - 3H(\rho_L^{(0)}) , \quad (21)$$

$$H(\rho) = -\rho \ln \rho - (1-\rho)\ln(1-\rho) ,$$

and $\boldsymbol{\lambda}_{[1,L]}$ is a vector of arbitrary constants.

Proof: See Appendix II. ∎

We see that the function $F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})$ is the key component in the upper bounds (18)-(20). If $F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]}) < 0$ for some $\boldsymbol{\lambda}_{[1,L]}$, then the expected number of codes having codewords with normalized weight composition $\boldsymbol{\rho}_{[1,L]}$ goes to zero, i.e., there are codes without the composition $\boldsymbol{\rho}_{[1,L]}$. Let us define

$$U(\boldsymbol{\rho}_{[1,L]}) \stackrel{\text{def}}{=} \min_{\boldsymbol{\lambda}_{[1,L]}} F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]}) \quad (22)$$

$$\theta(\rho_{[1,L]}) \stackrel{\text{def}}{=} \max_{\boldsymbol{\rho}_{[1,L]}: w(\boldsymbol{\rho}_{[1,L]})=\rho_{[1,L]}} U(\boldsymbol{\rho}_{[1,L]}) \quad (23)$$

and

$$\rho_L^* \stackrel{\text{def}}{=} \max\{\rho_{[1,L]} : \theta(\widehat{\rho}_{[1,L]}) < 0, \text{ if } 0 < \widehat{\rho}_{[1,L]} < \rho_{[1,L]}\} \quad (24)$$

$$= \max\{\rho_{[1,L]} : \max_{\substack{\boldsymbol{\rho}_{[1,L]}: \\ w(\boldsymbol{\rho}_{[1,L]})=\widehat{\rho}_{[1,L]}}} U(\boldsymbol{\rho}_{[1,L]}) < 0, \text{ if } 0 < \widehat{\rho}_{[1,L]} < \rho_{[1,L]}\} \quad (25)$$

$$= \max\{\rho_{[1,L]} : \max_{\substack{\boldsymbol{\rho}_{[1,L]}: \\ w(\boldsymbol{\rho}_{[1,L]})=\widehat{\rho}_{[1,L]}}} \min_{\boldsymbol{\lambda}_{[1,L]}} F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]}) < 0, \text{ if } 0 < \widehat{\rho}_{[1,L]} < \rho_{[1,L]}\}. \quad (26)$$

Finally, we note that $\theta(\cdot)$ is a continuous function and it therefore follows from (24) that ρ_L^* is the smallest positive root of $\theta(\rho_{[1,L]})$.

Now we can begin the proof of Theorem 1. We start by proving

$$d_L(0) \ge \rho_L^* M = (\rho_L^*/6)\nu . \quad (27)$$

Let

$$\mathcal{A}_L(\rho_{[1,L]}) = \{\boldsymbol{\rho}_{[1,L]} : w(\boldsymbol{\rho}_{[1,L]}) = \rho_{[1,L]}\} ,$$

i.e., $\mathcal{A}_L(\rho_{[1,L]})$ is the set of normalized weight compositions $\boldsymbol{\rho}_{[1,L]}$ with total normalized Hamming weight $\rho_{[1,L]}$. We then show that, for an arbitrarily small $\delta > 0$,

$$\sum_{d=2}^{\lfloor M(\rho_L^*-\delta)\rfloor} \sum_{\boldsymbol{\rho}_{[1,L]} \in \mathcal{A}_L\left(\frac{d}{M}\right)} E(\boldsymbol{\rho}_{[1,L]}) \to 0 \quad (28)$$

as $M \to \infty$[7]. This implies the existence of a code without any codewords of normalized weight smaller than ρ_L^*, i.e., $d_L(0) \ge \rho_L^* M$. (At the end of this section, using an expurgation procedure and allowing for a negligible rate loss, we then show that the Lth order segment distance $d_L \ge \rho_L^* M$.)

First we calculate the cardinality of $\mathcal{A}_L\left(\frac{d}{M}\right)$, which is the number of ways of representing an integer d as a sum of $2L$ nonnegative integers, i.e.,

$$\left|\mathcal{A}_L\left(\frac{d}{M}\right)\right| = \binom{2L-1+d}{2L-1} = \frac{2L-1+d}{2L-1}\cdot\frac{2L-2+d}{2L-2}\cdots\frac{d+1}{1} \le d^{2L-2}(d+1) < d^{2L} \quad (29)$$

for $d \ge 2$.

Now consider one term of the sum in (28). Using Lemma 1 b), (22), and (23), this can be upper bounded as

$$\sum_{\boldsymbol{\rho}_{[1,L]} \in \mathcal{A}_L\left(\frac{d}{M}\right)} E(\boldsymbol{\rho}_{[1,L]}) \le \sum_{\boldsymbol{\rho}_{[1,L]} \in \mathcal{A}_L\left(\frac{d}{M}\right)} (2\pi e M)^{2L} e^{MU(\boldsymbol{\rho}_{[1,L]})} \le \left|\mathcal{A}_L\left(\frac{d}{M}\right)\right| (2\pi e M)^{2L} e^{M\theta\left(\frac{d}{M}\right)} \quad (30)$$

by replacing each term in the sum with the largest term. Now upper bounding the right hand side of (30) using (29), we obtain

$$\sum_{\boldsymbol{\rho}_{[1,L]} \in \mathcal{A}_L\left(\frac{d}{M}\right)} E(\boldsymbol{\rho}_{[1,L]}) \le d^{2L}(2\pi e M)^{2L} e^{M\theta\left(\frac{d}{M}\right)} \quad (31)$$

$$\le \exp\left\{ M\left[\frac{2L}{M}\ln d + \frac{2L}{M}\ln(2\pi e M) + \theta\left(\frac{d}{M}\right)\right]\right\} . \quad (32)$$

Since we must consider only values of d such that $d \le \lfloor(\rho_L^* - \delta)M\rfloor$, the first two terms within the square brackets in (32) go to zero as $M \to \infty$. With respect to the third term, we note that the function $\theta(\cdot)$ is negative between its two roots 0 and ρ_L^*. Therefore, for the case when $\frac{d}{M}$ is bounded away from 0, i.e., $\delta' \le \frac{d}{M} \le \rho_L^* - \delta$ for some δ', as $M \to \infty$, there exists a positive constant $\epsilon > 0$ such that $\theta\left(\frac{d}{M}\right) < -\epsilon$. In this case (32) is upper bounded by

$$\exp\left\{ M\left[\frac{2L}{M}\ln d + \frac{2L}{M}\ln(2\pi e M) - \epsilon\right]\right\} \to 0 \quad (33)$$

as $M \to \infty$, where the two first terms inside the square brackets go to 0 as $M \to \infty$ and $-\epsilon$ stays constant independent of M.

For the case when $\frac{d}{M}$ goes to zero as $M \to \infty$, we need the following lemma.

Lemma 2:

$$\theta(\rho_{[1,L]}) \le 3 f_6(\rho_{[1,L]}) , \quad (34)$$

[7]The equal weight condition for the first two and last two blocks of a code segment $\boldsymbol{v}_{[1,L]}$ implies that all non-zero code segments of length L must have weight at least 2.

where

$$f_6(\rho_{[1,L]}) \stackrel{\text{def}}{=} \frac{1}{6}\rho_{[1,L]}(3\ln 5 - 1 + \ln \rho_{[1,L]}) \ . \quad (35)$$

Proof: See Appendix III. ∎

Now using Lemma 1 c), (22), (23), (29), and Lemma 2, we obtain

$$\begin{aligned}
&\sum_{\boldsymbol{\rho}_{[1,L]} \in \mathcal{A}_L\left(\frac{d}{M}\right)} E(\boldsymbol{\rho}_{[1,L]}) \\
&\leq \sum_{\boldsymbol{\rho}_{[1,L]} \in \mathcal{A}_L\left(\frac{d}{M}\right)} \exp\left\{M\left(U(\boldsymbol{\rho}_{[1,L]}) + \frac{d}{M}\ln(2\pi e)\right)\right\} \\
&\leq \left|\mathcal{A}_L\left(\frac{d}{M}\right)\right| \exp\left\{M\left(\theta\left(\frac{d}{M}\right) + \frac{d}{M}\ln(2\pi e)\right)\right\} \quad (36) \\
&\leq d^{2L} \exp\left\{M\left(3f_6\left(\frac{d}{M}\right) + \frac{d}{M}\ln(2\pi e)\right)\right\} \quad (37) \\
&\leq d^{2L} \exp\Bigg\{M\Bigg(\frac{1}{2}(3\ln 5 - 1)\frac{d}{M} \\
&\quad + \frac{1}{2}\frac{d}{M}\ln\left(\frac{d}{M}\right) + \frac{d}{M}\ln(2\pi e)\Bigg)\Bigg\} \quad (38) \\
&\leq \exp\Bigg\{M\Bigg(\frac{1}{2}(3\ln 5 - 1)\frac{d}{M} + \frac{1}{2}\frac{d}{M}\ln\left(\frac{d}{M}\right) \\
&\quad + \frac{d}{M}\ln(2\pi e) + 2L\frac{\ln d}{M}\Bigg)\Bigg\} \quad (39) \\
&\leq \exp\Bigg\{M\Bigg(\frac{1}{2}(3\ln 5 - 1)\frac{d}{M} + \frac{1}{2}\frac{d}{M}\ln\left(\frac{d}{M}\right) \\
&\quad + \frac{d}{M}\ln(2\pi e) + 2L\frac{d}{M}\Bigg)\Bigg\} \quad (40) \\
&\leq \exp\left\{M\left(a_L\frac{d}{M} + \frac{1}{2}\frac{d}{M}\ln\left(\frac{d}{M}\right)\right)\right\} \\
&= \left(e^{a_L}\sqrt{\frac{d}{M}}\right)^d , \quad (41)
\end{aligned}$$

where

$$a_L \stackrel{\text{def}}{=} \frac{1}{2}(3\ln 5 - 1) + \ln(2\pi e) + 2L \ . \quad (42)$$

Observe that for $\frac{d}{M} < e^{-2a_L}$,

$$a_L\frac{d}{M} + \frac{1}{2}\frac{d}{M}\ln\left(\frac{d}{M}\right) < 0 \ , \quad (43)$$

and (41) goes to zero as $M \to \infty$. (Note, however, that the convergence is not exponential.)

Now by choosing $\delta' = e^{-2a_L}$, we can consider separately the cases $d < \lfloor \delta' M \rfloor$ and $\lfloor \delta' M \rfloor \leq d \leq \lfloor M(\rho_L^* - \delta) \rfloor$. We begin by splitting (28) into two terms as follows:

$$\begin{aligned}
\sum_{d=2}^{\lfloor M(\rho_L^*-\delta)\rfloor} \sum_{\boldsymbol{\rho}_{[1,L]} \in \mathcal{A}_L\left(\frac{d}{M}\right)} E(\boldsymbol{\rho}_{[1,L]}) = \\
\sum_{d=2}^{\lfloor M\delta' \rfloor - 1} \sum_{\boldsymbol{\rho}_{[1,L]} \in \mathcal{A}_L\left(\frac{d}{M}\right)} E(\boldsymbol{\rho}_{[1,L]}) + \\
\sum_{d=\lfloor M\delta' \rfloor}^{\lfloor M(\rho_L^*-\delta)\rfloor} \sum_{\boldsymbol{\rho}_{[1,L]} \in \mathcal{A}_L\left(\frac{d}{M}\right)} E(\boldsymbol{\rho}_{[1,L]}) \ . \quad (44)
\end{aligned}$$

Using (41), the first term in (44) can be upper bounded as

$$\begin{aligned}
&\sum_{d=2}^{\lfloor M\delta' \rfloor - 1} \sum_{\boldsymbol{\rho}_{[1,L]} \in \mathcal{A}_L\left(\frac{d}{M}\right)} E(\boldsymbol{\rho}_{[1,L]}) \\
&\leq \sum_{d=2}^{\lfloor Me^{-2a_L} \rfloor - 1} \left(e^{a_L}\sqrt{\frac{d}{M}}\right)^d \\
&\leq \frac{2e^{2a_L}}{M} + \frac{3\sqrt{3}e^{3a_L}}{M^{3/2}}(Me^{-2a_L} - 3) \to 0 \quad (45)
\end{aligned}$$

as $M \to \infty$, where the second inequality in (45) is obtained by overbounding the terms for $d = 4, \ldots, \lfloor Me^{-2a_L} \rfloor - 1$ by the $d = 3$ term. Also, using (33), the second term in (44) can be upper bounded as

$$\begin{aligned}
&\sum_{d=\lfloor Me^{-2a_L} \rfloor}^{\lfloor M(\rho_L^*-\delta)\rfloor} \sum_{\boldsymbol{\rho}_{[1,L]} \in \mathcal{A}_L\left(\frac{d}{M}\right)} E(\boldsymbol{\rho}_{[1,L]}) \\
&\qquad \leq \sum_{d=\lfloor Me^{-2a_L} \rfloor}^{\lfloor M(\rho_L^*-\delta)\rfloor} \exp\Bigg\{M\Bigg(-\epsilon + \frac{2L}{M}\ln M \\
&\qquad\quad + \frac{2L}{M}\ln(2\pi e M)\Bigg)\Bigg\} \to 0 \quad (46)
\end{aligned}$$

as $M \to \infty$, and (28) is proved.

Hence, for a fixed L, we have shown that as $M \to \infty$ the expected number of code segments starting at $t = 1$ with normalized Hamming weight less than ρ_L^* in the ensemble $\mathcal{C}_P(3, 6, M)$ tends to zero. Therefore, there exists a code in the ensemble $\mathcal{C}_P(3, 6, M)$ such that $d_{\text{L}}(0) \geq (\rho_L^*/6)\nu$.

For $L = 1, 2, \ldots, 16$, the values ρ_L^* can be calculated numerically and are given in Table I. We observe that ρ_L^* rapidly decreases for small L. On the other hand, for $L \geq 10$ the calculated values of ρ_L^* stabilize at 0.51. In Appendix IV, we prove that $\rho_L^* \geq \rho^* = 0.5$ for any L, and the numerical calculations confirm that this lower bound is closely approached for $L \geq 10$.

L	1	2	3	4	5	10	16
ρ_L^*	2	1.65	0.85	0.66	0.56	0.51	0.51

TABLE I
NUMERICALLY CALCULATED VALUES OF ρ_L^*.

Using the same argument as above, it follows that for all possible starting positions i, there exists a code in the ensemble $\mathcal{C}_P(3, 6, M)$ such that $d_{\text{L}}(i) \geq (\rho_L^*/6)\nu$. In order to prove that the Lth order segment distance $d_{\text{L}} \geq (\rho_L^*/6)\nu$, we still need to show the existence of one single code such that $d_{\text{L}}(i) \geq (\rho_L^*/6)\nu$ for all possible starting positions i. We prove this using a special *expurgation procedure*.

Let i_j, $j = 1, 2, \ldots$, be the starting positions of all the code segments of length L with weight less than $(\rho_{\text{L}}^*/6)\nu$ for some code in the ensemble $\mathcal{C}_P(3, 6, M)$, i.e., the local segment distances $d_{\text{L}}(i_j)$, $j = 1, 2, \ldots$, for this code are upper bounded by $d_{\text{L}}(i_j) < (\rho_L^*/6)\nu$. The above analysis shows that, for a given starting position i, the fraction of codes with

$d_{\mathrm{L}}(i) < (\rho_L^*/6)\nu$ tends to zero with increasing M. It follows that there exists at least one code in the ensemble $\mathcal{C}_P(3,6,M)$ such that the fraction of starting positions i_j among all starting positions with local segment distance $d_{\mathrm{L}}(i_j) < (\rho_L^*/6)\nu$ tends to zero with increasing M. Recall that any code segment $\boldsymbol{v}_{[i_j+1,i_j+L]}$ has at least one non-zero information bit in the initial block of M information symbols. Each of the low weight code segments can then be expurgated by fixing at most one of the non-zero information bits in the initial information block to zero. (Note that the total number of information symbols that are fixed to zero is no more than the total number of low weight code segments.)

Fixing information symbols leads to a loss in rate. Since both the fraction of starting positions and the number of low weight code sequences tends to zero with increasing M, the fraction of fixed information symbols also tends to zero, and the corresponding rate loss is negligible.

Hence, inequality (3) of Theorem 1, i.e.,

$$d_{\mathrm{L}} \geq (\rho_L^*/6)\nu \ , \tag{47}$$

follows, and Theorem 1 is proved.

V. Proof of Theorem 2

For each particular code $C \in \mathcal{C}_P(3,6,M)$, the local free distance at position i is given by

$$d_{\text{free}}(i) = \min_{\boldsymbol{v}\in\mathcal{V}(C,i)} w_{\mathrm{H}}(\boldsymbol{v}) \ , \tag{48}$$

where $\mathcal{V}(C,i)$ is the set of code sequences (detours)

$$\boldsymbol{v} = (\ldots, \mathbf{0}, \boldsymbol{v}_{i+1}^{(0)}, \boldsymbol{v}_{i+1}^{(1)}, \ldots, \boldsymbol{v}_{i+L}^{(0)}, \boldsymbol{v}_{i+L}^{(1)}, \mathbf{0}, \ldots)$$

such that the partial syndrome former encoder (see [8]) starts in the zero state at time i, ends in the zero state at time $i+L$ (for some L), and does not pass through the zero state in between. We note that any sequence from $\mathcal{V}(C,i)$ cannot have more than five consecutive all-zero blocks; otherwise, the zero state would be reached before the end of the sequence. Moreover, the first two and last two blocks of a non-zero code segment $\boldsymbol{v}_{[i+1,i+L]}$ have the same weight. Finally, the global free distance is given by

$$d_{\text{free}} = \min_i d_{\text{free}}(i) \ . \tag{49}$$

We now consider the set $\mathcal{V}(i)$ of all possible code sequences $\boldsymbol{v}$ in the ensemble $\mathcal{C}_P(3,6,M)$ having at most five consecutive all-zero blocks and such that the weights of the first two and last two blocks of a non-zero code segment are equal. Thus $\mathcal{V}(C,i) \subset \mathcal{V}(i)$ for any code $C \in \mathcal{C}_P(3,6,M)$. By $\mathcal{V}(d,i)$ we denote the subset of $\mathcal{V}(i)$ with sequences having weight d, and by $\mathcal{R}\left(\frac{d}{M}, i\right)$ we denote the set of corresponding normalized weight sequences.

Without loss of generality, we can consider $d_{\text{free}}(0)$. Therefore, we will omit the index 0 in the notation for the normalized sequence set and write $\mathcal{R}\left(\frac{d}{M}\right)$ instead of $\mathcal{R}\left(\frac{d}{M}, 0\right)$. The number of code sequences leading to $d_{\text{free}}(0) < \rho^* M$ in a random code is upper bounded by

$$\sum_{d=2}^{\lfloor \rho^* M \rfloor} \sum_{\boldsymbol{\rho}\in\mathcal{R}\left(\frac{d}{M}\right)} E(\boldsymbol{\rho}) \ , \tag{50}$$

where $\boldsymbol{\rho}$ is the normalized weight sequence corresponding to a code sequence $\boldsymbol{v}$. We will show that this upper bound goes to zero as M tends to infinity.

First, we estimate the cardinality of $\mathcal{R}\left(\frac{d}{M}\right)$[8]. Consider a subset $\mathcal{R}_N\left(\frac{d}{M}\right)$ consisting of sequences from $\mathcal{R}(\frac{d}{M})$ having exactly N nonzero elements, i.e. there are N non-zero normalized weights $\rho_t^{(h)}$. The number of ways of distributing a weight of d among N nonzero terms is $\binom{d-1}{N-1}$. There can be from 0 to 5 zeros between any two adjacent nonzero elements. Thus

$$\left|\mathcal{R}_N\left(\frac{d}{M}\right)\right| = \binom{d-1}{N-1} 6^{N-1} \ , \tag{51}$$

and the cardinality of $\mathcal{R}\left(\frac{d}{M}\right)$ is

$$\left|\mathcal{R}\left(\frac{d}{M}\right)\right| \leq \sum_{N=1}^{d} \binom{d-1}{N-1} 6^{N-1} = 7^{d-1} \ . \tag{52}$$

Analogous to the proof of Theorem 1, we now separate the terms in (50) corresponding to small $\frac{d}{M}$ from the terms corresponding to large $\frac{d}{M}$. Choosing

$$\delta'' = \exp(-2\ln 7 - 3\ln 5 + 1 - 2\ln(2\pi e)) \tag{53}$$

and splitting (50) into two sums, we obtain

$$\begin{aligned}\sum_{d=2}^{\lfloor \rho^* M \rfloor} \sum_{\boldsymbol{\rho}\in\mathcal{R}\left(\frac{d}{M}\right)} E(\boldsymbol{\rho}) = &\sum_{d=2}^{\lfloor \delta'' M \rfloor} \sum_{\boldsymbol{\rho}\in\mathcal{R}\left(\frac{d}{M}\right)} E(\boldsymbol{\rho}) \\ &+ \sum_{d=\lfloor \delta'' M \rfloor + 1}^{\lfloor \rho^* M \rfloor} \sum_{\boldsymbol{\rho}\in\mathcal{R}\left(\frac{d}{M}\right)} E(\boldsymbol{\rho}) \ . \end{aligned} \tag{54}$$

Starting with the first term in (54), we use Lemma 1 c), (22), (23), Lemma 2, and (53) to obtain

$$\begin{aligned}&\sum_{d=2}^{\lfloor \delta'' M \rfloor} \sum_{\boldsymbol{\rho}\in\mathcal{R}\left(\frac{d}{M}\right)} E(\boldsymbol{\rho}) \\ &\leq \sum_{d=2}^{\lfloor \delta'' M \rfloor} \sum_{\boldsymbol{\rho}\in\mathcal{R}\left(\frac{d}{M}\right)} \exp\left\{M(U(\boldsymbol{\rho}) + \frac{d}{M}\ln(2\pi e))\right\} \\ &\leq \sum_{d=2}^{\lfloor \delta'' M \rfloor} \left|\mathcal{R}\left(\frac{d}{M}\right)\right| \exp\left\{M\left(\theta\left(\frac{d}{M}\right) + \frac{d}{M}\ln(2\pi e)\right)\right\} \\ &\leq \sum_{d=2}^{\lfloor \delta'' M \rfloor} 7^{d-1} \exp\left\{M\left(3f_6\left(\frac{d}{M}\right) + \frac{d}{M}\ln(2\pi e)\right)\right\}\end{aligned}$$

[8]Note that, for a given code segment length L, the set $\mathcal{R}\left(\frac{d}{M}\right)$, which only allows detours, is much smaller than the corresponding set $\mathcal{A}_L\left(\frac{d}{M}\right)$ (see (29)) which allows remergers.

$$\leq \sum_{d=2}^{\lfloor \delta'' M \rfloor} 7^{d-1} \exp\Bigg\{ M \Bigg(\frac{1}{2}(3\ln 5 - 1 + 2\ln(2\pi e)) \frac{d}{M} + \frac{1}{2}\frac{d}{M}\ln\frac{d}{M} \Bigg) \Bigg\} \quad (55)$$

$$\leq \sum_{d=2}^{\lfloor \delta'' M \rfloor} \exp\left\{ M\left(-\frac{1}{2}\ln\delta'' \frac{d}{M} + \frac{1}{2}\frac{d}{M}\ln\frac{d}{M} \right) \right\} \quad (56)$$

$$= \sum_{d=2}^{\lfloor \delta'' M \rfloor} \exp\left\{ d\left(-\frac{1}{2}\ln\delta'' + \frac{1}{2}\ln\frac{d}{M} \right) \right\} \quad (57)$$

$$= \sum_{d=2}^{\lfloor \delta'' M \rfloor} \left(\sqrt{\frac{d}{\delta'' M}} \right)^d . \quad (58)$$

Now observing that $\sqrt{\frac{d}{\delta'' M}} \leq 1$ for $d = 2, \ldots, \lfloor \delta'' M \rfloor$, (58) can be written as

$$\sum_{d=2}^{\lfloor \delta'' M \rfloor} \sum_{\boldsymbol{\rho} \in \mathcal{R}(\frac{d}{M})} E(\boldsymbol{\rho}) \leq \sum_{d=2}^{\lfloor \delta'' M \rfloor} \left(\sqrt{\frac{d}{\delta'' M}} \right)^d \leq \frac{2}{\delta'' M} + \frac{3\sqrt{3}}{(\delta'' M)^{\frac{3}{2}}} + (\delta'' M - 3)\frac{16}{\delta''^2 M^2} \to 0 \quad (59)$$

as $M \to \infty$.

To bound the second term in (54) we need the following Lemma.

Lemma 3: Consider a normalized weight sequence $\boldsymbol{\rho} = \boldsymbol{\rho}_{[1,L]}$ with N nonzero elements such that

$$\delta'' < w(\boldsymbol{\rho}_{[1,L]}) \leq \rho^* .$$

Then there exists an M' such that for any $M > M'$

$$U(\boldsymbol{\rho}_{[1,L]}) + \frac{1}{M} \sum_{\rho_t^{(h)} \neq 0} \left(\ln(2\pi e) + \ln \rho_t^{(h)} M \right) \leq -\epsilon + \frac{N}{2}\left(\frac{3\alpha}{M} - \frac{\ln M}{2M} \right) - N'\left(\frac{\alpha}{M} - \frac{\ln M}{6M} \right) , \quad (60)$$

where α, ϵ, and N' are positive constants.

Proof: See Appendix V. ∎

Using Lemmas 1 a) and 3, the second term in (54) can be written as

$$\sum_{d=\lfloor \delta'' M \rfloor + 1}^{\lfloor \rho^* M \rfloor} \sum_{\boldsymbol{\rho} \in \mathcal{R}(\frac{d}{M})} E(\boldsymbol{\rho}) \leq \sum_{d=\lfloor \delta'' M \rfloor + 1}^{\lfloor \rho^* M \rfloor} \sum_{N=1}^{d} \binom{d-1}{N-1} 6^{N-1} \cdot \exp\left\{ M\left(U(\boldsymbol{\rho}) + \sum_{\rho_t^{(h)} \neq 0} \frac{\ln(2\pi e \rho_t^{(h)} M)}{M} \right) \right\} \quad (61)$$

$$\leq \sum_{d=\lfloor \delta'' M \rfloor + 1}^{\lfloor \rho^* M \rfloor} \sum_{N=1}^{d} \binom{d-1}{N-1} 6^{N-1} \cdot \exp\left\{ M\left(-\epsilon + \frac{N}{2}\left(\frac{3\alpha}{M} - \frac{\ln M}{2M} \right) - N'\left(\frac{\alpha}{M} - \frac{\ln M}{6M} \right) \right) \right\} \quad (62)$$

$$\leq M^{N'/6} e^{-N'\alpha} e^{-\epsilon M} \sum_{d=\lfloor \delta'' M \rfloor + 1}^{\lfloor \rho^* M \rfloor} \sum_{N=0}^{d} \binom{d}{N} 6^N e^{N(\frac{3\alpha}{2} - \frac{\ln M}{4})} \leq M^{N'/6} e^{-N'\alpha} e^{-\epsilon M} \sum_{d=\lfloor \delta'' M \rfloor + 1}^{\lfloor \rho^* M \rfloor} \left(1 + 6e^{\frac{3\alpha}{2} - \frac{\ln M}{4}} \right)^d \leq M^{N'/6} e^{-N'\alpha} e^{-\epsilon M} M \left(1 + \frac{6e^{\frac{3\alpha}{2}}}{M^{\frac{1}{4}}} \right)^{\rho^* M} , \quad (63)$$

where the last inequality is obtained by replacing each term in the sum by the largest term and upperbounding the number of terms by M. Now we continue upperbounding (63) and obtain

$$\begin{aligned} & M^{N'/6} e^{-N'\alpha} e^{-\epsilon M} M \left(1 + \frac{6e^{\frac{3\alpha}{2}}}{M^{\frac{1}{4}}} \right)^{\rho^* M} \\ &= \exp\Bigg\{ (N'/6+1)\ln M - N'\alpha - \epsilon M + \rho^* M \ln\left(1 + \frac{6e^{\frac{3\alpha}{2}}}{M^{\frac{1}{4}}} \right) \Bigg\} \\ &\leq \exp\Bigg\{ (N'/6+1)\ln M - N'\alpha - \epsilon M + \rho^* M \frac{6e^{\frac{3\alpha}{2}}}{M^{\frac{1}{4}}} \Bigg\} \\ &= \exp\Bigg\{ (N'/6+1)\ln M - N'\alpha - \epsilon M + \rho^* 6 e^{\frac{3\alpha}{2}} M^{\frac{3}{4}} \Bigg\} \to 0 \end{aligned} \quad (64)$$

as $M \to \infty$, since the term $-\epsilon M$ dominates.

Hence the fraction of codes having local free distance $d_{\text{free}}(0) < (\rho^*/6)\nu$ goes to zero, and this is also valid for any starting position i. Now the expurgation procedure from Theorem 1 can be used to prove the existence of a code with global free distance $d_{\text{free}} \geq (\rho^*/6)\nu$, and Theorem 2 is proved.

VI. Results and Discussion

In order to calculate the segment distance to constraint length ratio in Theorem 1 for small values of L, we use numerical methods to solve the max-min problem of (26) and obtain ρ_L^*. The symmetry condition $\rho_t^{(0)} = \rho_t^{(1)}$, for $t = 2, \ldots, L-1$, can be shown to hold in the calculation of $\max_{\boldsymbol{\rho}_{[1,L]}: w(\boldsymbol{\rho}_{[1,L]}) = \hat{\rho}_{[1,L]}} \min_{\boldsymbol{\lambda}_{[1,L]}} F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})$. This reduces the number of variables by half and simplifies the numerical evaluation of $F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})$.

In Fig. 5 the ratio $\rho_L^*/6$, a lower bound on the segment distance to constraint length ratio $d_{\text{L}}/(6M)$, is plotted as a function of L for $L \leq 16$. Observe that, for $L = 1$, the bound has its maximum value of $\rho_1^*/6 = 1/3$, i.e., the code segment

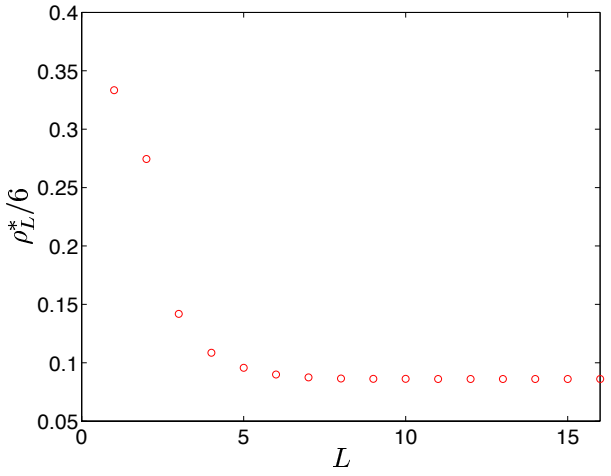

Fig. 5. Lower bound on segment distance to constraint length ratio $\frac{d_{\mathrm{L}}}{6M}$ as a function of L.

has weight $d_1 = 2M$. Also note that, for $L = 16$, the bound reaches its minimum value of $\rho_L^*/6 = 0.085$.

Fig. 5 shows that the lower bound on segment distance to constraint length ratio $\rho_L^*/6$ is virtually unchanged from $L = 9$ to $L = 16$. In Fig. 6 we plot the ratio $\rho_t^{(0)}/6$ for the normalized weight composition $\boldsymbol{\rho}_{[1,L]}$ satisfying $w(\boldsymbol{\rho}_{[1,L]}) = \rho_L^*$ that maximizes $U(\boldsymbol{\rho}_{[1,L]})$ for $L = 9, 11, 13, 15$ (Fig. 6(a)) and $L = 10, 12, 14, 16$ (Fig. 6(b)), where the time index t represents distance from the middle block of the segment. We see that the maximizing weight composition always satisfies $\rho_t^{(0)} = \rho_{L-t}^{(0)}$, i.e., the normalized Hamming weights are symmetrically distributed about the middle block. Also the weight distribution of the maximizing weight composition $\boldsymbol{\rho}_{[1,L]}$ is largest in the middle and both ends have almost zero weight.

The free distance of a convolutional code satisfies

$$d_{\mathrm{free}} = \min_{L} d_{\mathrm{L}}, \ L \geq 1 \ , \tag{65}$$

where d_{L} is the Lth order segment distance, and the numerical calculations of Fig. 5 indicate that the segment distance decreases as a function of L until the free distance is attained, after which it stays unchanged at d_{free}. Further, we have shown that both the Lth order segment distance and the free distance are lower bounded by $(\rho^*/6)\nu = \alpha_{\mathrm{LDPCC}}(3,6)\nu$, where $\alpha_{\mathrm{LDPCC}}(3,6) = 0.083$ for any L and correspondingly large M (or, equivalently, ν). This bound is weaker than the Costello coefficient $\alpha_{\mathrm{C}}(1/2) = 0.39$ for rate $R = 1/2$ convolutional codes. However, $\alpha_{\mathrm{LDPCC}}(3,6)$ is about three and a half times larger than the Gallager coefficient $\alpha_{\mathrm{G}}(3,6) = 0.023$ for $(3,6)$-regular block codes. Interestingly, for the general class of rate $R = 1/2$ codes, $\alpha_{\mathrm{C}}(1/2) = 0.39$ is also about three and a half times larger than the corresponding Gilbert-Varshamov coefficient $\alpha_{\mathrm{GV}}(1/2) = 0.11$.

In Table II we compare the numerically calculated values of the parameter $\alpha_{\mathrm{LDPCC}}(J,K)$ with $\alpha_{\mathrm{G}}(J,K)$ for the $(3,6)$ and $(4,8)$ cases, and we see that the asymptotic distance bound ratio for LDPC convolutional codes is more than three times larger than for the corresponding LDPC block codes in both cases.

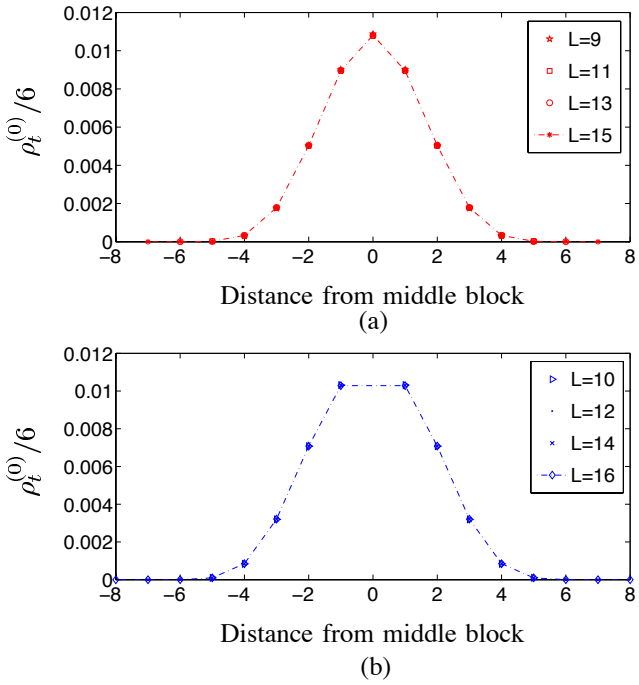

Fig. 6. Weight distribution of the maximizing normalized weight vector $\boldsymbol{\rho}_{[1,L]}$ for different L.

(J,K)	$\alpha_{\mathrm{LDPCC}}(J,K)$	$\alpha_{\mathrm{G}}(J,K)$
(3,6)	0.0833	0.023
(4,8)	0.1908	0.0627

TABLE II

NUMERICALLY CALCULATED DISTANCE BOUNDS FOR LDPC BLOCK AND CONVOLUTIONAL CODES ($\alpha_{\mathrm{LDPCC}}(4,8)$ IS DEFINED ANALOGOUSLY TO $\alpha_{\mathrm{LDPCC}}(3,6)$, AS IN (4)).

In the case when J and K are relatively prime, the convolutional code ensemble $\mathcal{C}_P(J,K,M)$ has syndrome former memory $m_{\mathrm{s}} = 0$ and is identical to the block code ensemble of [12]. Hence the Lth order segment distance and the free distance in this case satisfy the same lower bound[9], i.e., the lower bound derived by Gallager for LDPC block codes. For other values of J and K, obtaining upper bounds on $U(\cdot)$ that permit an analytical evaluation of $\alpha_{\mathrm{LDPCC}}(J,K)$ becomes more complicated, and a numerical solution of the max-min optimization problem is also difficult to obtain.

The convolutional code ensemble $\mathcal{C}_P(3,6,M)$ is composed of permutation matrices, and hence in any code there always exists a code sequence with weight $2M$ of the form $(\ldots, \mathbf{0}, \boldsymbol{v}_t^{(0)}, \boldsymbol{v}_t^{(1)}, \mathbf{0}, \ldots)$ with $\boldsymbol{v}_t^{(0)} = \boldsymbol{v}_t^{(1)} = \mathbf{1}$, where $\mathbf{1}$ is the M-dimensional all-one vector. In fact, for all J and K, such a code sequence with weight $2M$ always exists for any code in the ensemble $\mathcal{C}_P(J,K,M)$. This limits the

[9]When J and K are relatively prime the constraint length ν of the convolutional codes in $\mathcal{C}_P(J,K,M)$ equals the block length N of the codes considered in [12]. Hence, for such values of J and K, we have $d_{\mathrm{min}} = d_{\mathrm{free}} \geq \alpha_{\mathrm{G}}(J,K)\nu$.

asymptotic segment distance ratio and the free distance ratio to $\alpha_{\text{LDPCC}}(J,K) \leq \frac{2}{K}$. This is a severe restriction for codes with large K. However, we can expurgate such code sequences from the ensemble at the expense of a small loss in rate. For example, in the $(J,2J)$ case, we can fix the first information symbol in each block to be a zero, so that the rate is reduced to $R = \frac{M-1}{2M}$. In general, such low weight code sequences can be avoided by fixing one information symbol in the first block to zero.

VII. Conclusions

We have introduced an ensemble of LDPC convolutional codes with syndrome formers comprised of permutation matrices. Such code ensembles lend themselves to an analysis of their distance and threshold[10] properties. In particular, we can derive lower bounds on the Lth order segment distance and the free distance of these codes. We have proved that this ensemble contains codes whose Lth order segment distance and free distance increases linearly with constraint length. Further, for the same (J,K), the numerically evaluated asymptotic free distance to constraint length ratio $\alpha_{\text{LDPCC}}(J,K)$ is several times larger than the asymptotic minimum distance to block length ratio obtained by Gallager for LDPC block codes. For example, in the $(3,6)$ case, we show that $\alpha_{\text{LDPCC}}(3,6) = \frac{d_{\text{free}}}{\nu} \geq 0.083$. This value is about three and a half times larger than the corresponding coefficient $\frac{d_{\min}}{N} = 0.023$ for $(3,6)$ LDPC block codes.

References

[1] R. G. Gallager, *Low-density parity-check codes*. M.I.T. Press, Cambridge, Massachusetts, 1963.

[2] R. M. Tanner, "Error-correcting coding system." Patent No. 4,295,218, Oct. 13 1981.

[3] A. Jiménez Feltström and K. S. Zigangirov, "Time-varying periodic convolutional codes with low-density parity-check matrix," *IEEE Transactions on Information Theory*, vol. IT-45, no. 5, pp. 2181–2190, Sept. 1999.

[4] K. Engdahl and K. S. Zigangirov, "On the theory of low density convolutional codes I," *Problems of Information Transmission (Problemy Peredachi Informatsii)*, vol. 35, no. 4, pp. 295–310, Oct.-Dec. 1999.

[5] R. M. Tanner, D. Sridhara, A. Sridharan, T. E. Fuja, and D. J. Costello, Jr., "LDPC block and convolutional codes based on circulant matrices," *IEEE Transactions on Information Theory*, vol. IT-50, no. 12, pp. 2966–2984, Dec. 2004.

[6] D. J. Costello, Jr., A. E. Pusane, S. Bates, and K. S. Zigangirov, "A comparison between LDPC block and convolutional codes," in *Proceedings of the Information Theory and Application Workshop*, (San Diego, USA), Feb. 2006.

[7] S. Bates, Z. Chen, and X. Dong, "Low-density parity check convolutional codes for ethernet networks," in *Proceedings of the IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, (Victoria, B. C. Canada), Aug. 2005.

[8] A. E. Pusane, A. Jiménez Feltström, A. Sridharan, M. Lentmaier, K. S. Zigangirov, and D. J. Costello, Jr., "Implementation aspects of LDPC convolutional codes." IEEE Transactions on Communications, to appear.

[9] A. Sridharan, M. Lentmaier, D. J. Costello, Jr., and K. S. Zigangirov, "Convergence analysis of a class of LDPC convolutional codes for the erasure channel," in *Proccedings of the 42nd Allerton Conference on Communication, Control, and Computing*, (Monticello, IL, USA), Oct. 2004.

[10] M. Lentmaier, A. Sridharan, D. J. Costello, Jr., and K. S. Zigangirov, "Terminated LDPC convolutional codes with thresholds close to capacity," in *Proceedings of the 2005 IEEE Intl. Symposium on Information Theory*, (Adelaide, AUS), pp. 1372–1376, Sept. 2005.

[11] D. J. C. MacKay and M. C. Davey, "Evaluation of Gallager codes for short block length and high rate applications," in *Codes, Systems and Graphical Models* (B. Marcus and J. Rosenthal, eds.), vol. 123 of *IMA Volumes in Mathematics and its Applications*, ch. 5, pp. 113–130, New York: Springer-Verlag, 2001.

[12] A. Sridharan, M. Lentmaier, D. V. Truhachev, D. J. Costello, Jr., and K. S. Zigangirov, "On the minimum distance of low-density parity-check codes with parity-check matrices constructed from permutation matrices," *Problems of Information Transmission (Problemy Peredachi Informatsii)*, vol. 41, no. 1, pp. 33–44, 2005.

[13] E. N. Gilbert, "A comparison of signalling alphabets," *Bell System Technical Journal*, vol. 31, pp. 504–522, 1952.

[14] R. R. Varsharmov, "Estimates of the number of signals in error correcting codes," *Doklady A.N.S.S.R*, vol. 117, no. 5, pp. 739–741, 1957.

[15] D. J. Costello, Jr., "Free distance bounds for convolutional codes," *IEEE Transactions on Information Theory*, vol. IT-20, no. 3, pp. 356–365, May 1974.

[16] K. Engdahl, M. Lentmaier, and K. S. Zigangirov, "On the theory of low-density convolutional codes," in *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, 13th International Symposium, AAECC-13* (M. P. C. Fossorier, H. Imai, S. Lin, and A. Poli, eds.), vol. 1719 of *Lecture Notes in Computer Science*, pp. 77–86, Springer, Nov. 1999.

[17] R. Johannesson and K. S. Zigangirov, *Fundamentals of Convolutional Coding*. IEEE Press, 1999.

[18] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*. Addison - Wesley, 1989.

[19] R. M. Fano, *Transmission of Information, A Statistical Theory of Communications*. M.I.T. Press and John Wiley and Sons, Inc, 1961.

[10] See [9][10].

Appendix I
Ensemble Analysis

We calculate the probability that an IM dimensional vector, $I \geq 2$, $\boldsymbol{w}_{[1,I]} = (\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_I)$, $\boldsymbol{w}_i \in \mathbb{F}_2^M$, $i = 1, \ldots, I$, satisfies the condition

$$\boldsymbol{w}_{[1,I]}\boldsymbol{\Pi} = \boldsymbol{0}, \tag{66}$$

where the $IM \times M$ matrix $\boldsymbol{\Pi}$ is given by

$$\boldsymbol{\Pi} = [\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_I]^{\mathrm{T}}$$

and the $\boldsymbol{P}_i$, $i = 1, \ldots, I$, are $M \times M$ permutation matrices. For the block code ensemble in [12] only the case $I = K$ was needed. However, for the (J,K) LDPC convolutional code ensemble considered here, the cases $I = 2, 4, \ldots, K$, are required. Thus, for the ensemble $\mathcal{C}_P(3,6,M)$, the cases $I = 2,4,6$, must be considered. Since the permutation matrices comprising the syndrome formers in $\mathcal{C}_P(3,6,M)$ are chosen independently and equally likely, we assume the same for the matrices $\boldsymbol{P}_i$, i.e., they are chosen independently and take on one of the $M!$ possible values with equal probability.

Let d_i, $i = 1, \ldots, I$, be the Hamming weight of the M dimensional vector $\boldsymbol{w}_i$. We say that $\boldsymbol{w}_{[1,I]}$ has weight composition $\boldsymbol{d}_{[1,I]} = (d_1, d_2, \ldots, d_I)$. The Hamming weight d of the vector $\boldsymbol{w}_{[1,I]}$ with weight composition $\boldsymbol{d}_{[1,I]}$ is $d = d_1 + d_2 + \cdots + d_I$.

If $\boldsymbol{w}_{[1,I]}$ satisfies (66), then each of the M constraint equations defined by the columns of $\boldsymbol{\Pi}$ must include an even number of ones, i.e., $0, 2, \ldots, 2\lfloor I/2 \rfloor$, from $\boldsymbol{w}_{[1,I]}$. Since $\boldsymbol{\Pi}$ is composed of blocks of permutation matrices, a constraint equation involves at most a single one from each $\boldsymbol{w}_i$, $i = 1, 2, \ldots, I$.

For the mth constraint equation of $\boldsymbol{\Pi}$, $m = 1, \ldots, M$, we associate an I-dimensional binary vector $\boldsymbol{p}_m$. The ith component of $\boldsymbol{p}_m$, $i = 1, \ldots, I$, is one if a one from $\boldsymbol{w}_i$ is involved in the mth constraint equation and is zero otherwise.

Since any constraint equation involves an even number of ones, $\boldsymbol{p}_m$ can take on one of 2^{I-1} values.

Let ν_0 denote the number of constraint equations involving only zeros of $\boldsymbol{w}_{[1,I]}$. Then, let $\nu_2(i_1,i_2)$ be the number of constraints involving two ones, one from each of the vectors $\boldsymbol{w}_{i_1}$ and $\boldsymbol{w}_{i_2}$, and zeros from the other $I-2$ components of $\boldsymbol{w}_{[1,I]}$. In general, let $\nu_{2q}(i_1,\ldots,i_{2q})$, $q=1,\ldots,\lfloor I/2 \rfloor$, denote the number of constraints involving $2q$ ones, one from each of the vectors $w_{i_1},\ldots,w_{i_{2q}}$, and zeros from the remaining $(I-2q)$ components of $\boldsymbol{w}_{[1,I]}$. Observe that the arguments of $\nu_{2q}(i_1,\ldots,i_{2q})$ are distinct, i.e., $i_1 \neq i_2 \neq \cdots \neq i_{2q}$. Further, $\nu_{2q}(i_1,\ldots,i_{2q})$ is invariant to permutations of the arguments, for example, $\nu_{2q}(i_1,i_2,\ldots,i_{2q}) = \nu_{2q}(i_2,i_1,\ldots,i_{2q})$. In other words, $\nu_{2q}(i_1,\ldots,i_{2q})$ is a function of the set $\{i_1,i_2,\ldots,i_{2q}\}$. To emphasize this fact, we henceforth write $\nu_{2q}(\{i_1,\ldots,i_{2q}\})$ for $\nu_{2q}(i_1,\ldots,i_{2q})$. There exist $\binom{I}{2q}$ different sets $\{i_1,i_2,\ldots,i_{2q}\}$. Therefore the function $\nu_{2q}(\{i_1,i_2,\ldots,i_{2q}\})$ can take on $\binom{I}{2q}$ values, not necessarily distinct.

Now assume that $I=6$. Then the number of ones in each constraint equation is at most six. In this case, since the total number of constraints is M, we have[11]

$$\nu_0 + \sum_{\{i_1,i_2\}} \nu_2(\{i_1,i_2\}) + \sum_{\{i_1,i_2,i_3,i_4\}} \nu_4(\{i_1,i_2,i_3,i_4\}) + \nu_6 = M \ . \quad (67)$$

Further, it follows from the definition of $\nu_{2q}(\{i_1,i_2,\ldots,i_{2q}\})$ that, for any $i_1 \in \{1,2,\ldots,I\}$,

$$\sum_{\{i_2\}} \nu_2(\{i_1,i_2\}) + \sum_{\{i_2,i_3,i_4\}} \nu_4(\{i_1,i_2,i_3,i_4\}) + \nu_6 = d_{i_1}. \quad (68)$$

For any vector $\boldsymbol{w}_{[1,6]}$ with weight composition $\boldsymbol{d}_{[1,6]}$,

$$(\nu_0, \{\nu_2(\{i_1,i_2\})\}, \{\nu_4(\{i_1,i_2,i_3,i_4\})\}, \nu_6)$$

is called a constraint composition of $\boldsymbol{w}_{[1,6]}$ if

$$\nu_0, \{\nu_2(\{i_1,i_2\})\}, \{\nu_4(\{i_1,i_2,i_3,i_4\})\} \ ,$$

and ν_6 satisfy (67) and the six equations implied in (68). To clarify some of the above notation we present a simple example.

Example 1: Let $I=6$ and $M=5$.
In this case, each of the vectors $\boldsymbol{w}_i$, $i=1,\ldots,6$, is a five dimensional vector. Let the vectors $\boldsymbol{w}_i$ be

$$\begin{aligned} \boldsymbol{w}_1 &= (0,1,1,1,0) \\ \boldsymbol{w}_2 &= (1,0,0,1,0) \\ \boldsymbol{w}_3 &= (0,0,1,1,0) \\ \boldsymbol{w}_4 &= (0,1,0,0,1) \\ \boldsymbol{w}_5 &= (1,0,0,1,1) \\ \boldsymbol{w}_6 &= (0,0,1,0,1). \end{aligned}$$

[11]We omit the arguments and write ν_6 since, in this case, we have a one from each of the six components of $\boldsymbol{w}_{[1,6]}$.

Hence, $\boldsymbol{w}_{[1,6]}$ has weight composition $\boldsymbol{d}_{[1,6]} = (3,2,2,2,3,2)$. Suppose now that $\boldsymbol{\Pi} = [\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_6]^{\mathrm{T}}$, with

$$\boldsymbol{P}_1^{\mathrm{T}} = \begin{pmatrix} 0&0&1&0&0 \\ 0&0&0&1&0 \\ 0&1&0&0&0 \\ 0&0&0&0&1 \\ 1&0&0&0&0 \end{pmatrix}, \boldsymbol{P}_2^{\mathrm{T}} = \begin{pmatrix} 1&0&0&0&0 \\ 0&0&1&0&0 \\ 0&0&0&0&1 \\ 0&0&0&1&0 \\ 0&1&0&0&0 \end{pmatrix},$$

$$\boldsymbol{P}_3^{\mathrm{T}} = \begin{pmatrix} 0&1&0&0&0 \\ 0&0&1&0&0 \\ 0&0&0&1&0 \\ 1&0&0&0&0 \\ 0&0&0&0&1 \end{pmatrix}, \boldsymbol{P}_4^{\mathrm{T}} = \begin{pmatrix} 0&0&1&0&0 \\ 1&0&0&0&0 \\ 0&1&0&0&0 \\ 0&0&0&0&1 \\ 0&0&0&1&0 \end{pmatrix},$$

$$\boldsymbol{P}_5^{\mathrm{T}} = \begin{pmatrix} 0&0&0&1&0 \\ 0&0&1&0&0 \\ 1&0&0&0&0 \\ 0&0&0&0&1 \\ 0&1&0&0&0 \end{pmatrix}, \boldsymbol{P}_6^{\mathrm{T}} = \begin{pmatrix} 0&0&1&0&0 \\ 0&1&0&0&0 \\ 1&0&0&0&0 \\ 0&0&0&0&1 \\ 0&0&0&1&0 \end{pmatrix}.$$

Since $I=6$, the vector $\boldsymbol{p}_m$ associated with the mth constraint equation, $m=1,\ldots,5$, is a six dimensional vector. The first constraint equation involves a one from each of the vectors $\boldsymbol{w}_2, \boldsymbol{w}_3, \boldsymbol{w}_4$, and $\boldsymbol{w}_6$. Hence, $\boldsymbol{p}_1 = (0,1,1,1,0,1)$. Similarly, we obtain

$$\begin{aligned} \boldsymbol{p}_2 &= (1,0,0,0,1,0) \\ \boldsymbol{p}_3 &= (0,0,0,0,0,0) \\ \boldsymbol{p}_4 &= (1,1,1,1,1,1) \\ \boldsymbol{p}_5 &= (1,0,0,0,1,0). \end{aligned}$$

The third constraint equation involves no one's from $\boldsymbol{w}_{[1,6]}$, and hence $\nu_0 = 1$. The second and fifth constraint equations involve one's from $\boldsymbol{w}_1$ and $\boldsymbol{w}_5$, so $\nu_2(\{1,5\}) = 2$. In fact, this is the only non-zero term in the set $\{\nu_2(\{i_1,i_2\})\}$. The first constraint equation involves ones from $\boldsymbol{w}_2, \boldsymbol{w}_3, \boldsymbol{w}_4$, and $\boldsymbol{w}_6$. This is the only non-zero term in the set $\{\nu_4(\{i_1,i_2,i_3,i_4\})\}$, and we have $\nu_4(\{2,3,4,6\}) = 1$. The fourth constraint equation involves a one from each of the vectors $\boldsymbol{w}_i$, $i=1,\ldots,6$. Hence $\nu_6 = 1$. Now we have

$$\nu_0 + \nu_2(\{1,5\}) + \nu_4(\{2,3,4,6\}) + \nu_6 = 1+2+1+1 = 5 = M$$

so that (67) is satisfied. Similarly, it can be checked that the six equations implied in (68) are also satisfied. □

Let $\boldsymbol{c} = (\boldsymbol{p}_1,\ldots,\boldsymbol{p}_M)$. If $\boldsymbol{w}_{[1,6]}$ has constraint composition

$$(\nu_0, \{\nu_2(\{i_1,i_2\})\}, \{\nu_4(\{i_1,i_2,i_3,i_4\})\}, \nu_6) \ ,$$

then it follows that

- ν_0 of the vectors $\boldsymbol{p}_m$, $m=1,\ldots,M$, are the all-zero vector
- For each set $\{i_1,i_2\}$, $\nu_2(\{i_1,i_2\})$ of the vectors $\boldsymbol{p}_m$ have ones only in positions i_1 and i_2
- For each set $\{i_1,i_2,i_3,i_4\}$, $\nu_4(\{i_1,i_2,i_3,i_4\})$ of the vectors $\boldsymbol{p}_m$ have ones only in positions i_1, i_2, i_3, and i_4
- ν_6 of the vectors $\boldsymbol{p}_m$ are the all-one vector

For any constraint composition

$$(\nu_0, \{\nu_2(\{i_1,i_2\})\}, \{\nu_4(\{i_1,i_2,i_3,i_4\})\}, \nu_6) \ ,$$

there are in total

$$\frac{M!}{\nu_0! \prod_{\{i_1,i_2\}} \nu_2(\{i_1,i_2\})! \prod_{\{i_1,i_2,i_3,i_4\}} \nu_4(\{i_1,i_2,i_3,i_4\})! \ \nu_6!}$$

possible vectors $\boldsymbol{c}$. Let

$$\boldsymbol{\nu} = \{(\nu_0, \{\nu_2(\{i_1, i_2\})\}, \{\nu_4(\{i_1, i_2, i_3, i_4\})\}, \nu_6)\}$$

represent the set of possible constraint compositions of $\boldsymbol{w}_{[1,6]}$. Note that $\boldsymbol{\nu}$ is only a function of the weight composition $\boldsymbol{d}_{[1,6]}$ of $\boldsymbol{w}_{[1,6]}$, i.e., $\boldsymbol{\nu} = \boldsymbol{\nu}(\boldsymbol{d}_{[1,6]})$. For a given weight composition $\boldsymbol{d}_{[1,6]}$, there are in total

$$\phi_6\left(\boldsymbol{d}_{[1,6]}\right) = \sum_{\boldsymbol{\nu}(\boldsymbol{d}_{[1,6]})} \frac{M!}{\nu_0! \prod_{\{i_1,i_2\}} \nu_2(\{i_1,i_2\})! \prod_{\substack{\{i_1,i_2,\\ i_3,i_4\}}} \nu_4(\{i_1,i_2,i_3,i_4\})! \nu_6!} \tag{69}$$

possible vectors $\boldsymbol{c}$, where the summation in (69) is over the set of possible constraint compositions. The formulas (67)-(69) can be similarly extended to other values of I. In the general case, the probability that $\boldsymbol{w}_{[1,I]}$ satisfies (66) is given by the following lemma.

Lemma 4: The probability $\gamma_I\left(\boldsymbol{d}_{[1,I]}\right)$ that a vector

$$\boldsymbol{w}_{[1,I]} = (\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_I)$$

with weight composition $\boldsymbol{d}_{[1,I]} = \{d_1, d_2, \ldots, d_I\}$ satisfies (66) is

$$\gamma_I\left(\boldsymbol{d}_{[1,I]}\right) = \frac{\phi_I\left(\boldsymbol{d}_{[1,I]}\right)}{\prod_{i=1}^{I} \binom{M}{d_i}}. \tag{70}$$

Proof: Consider an arbitrary vector $\boldsymbol{w}_{[1,I]}$ with weight composition $\boldsymbol{d}_{[1,I]}$. For a fixed $\boldsymbol{w}_{[1,I]}$ and $\boldsymbol{c}$, there are $d_i!(M-d_i)!$ ways of choosing the permutation matrix $\boldsymbol{P}_i,\ i = 1, \ldots, I$. Since the number of possible $\boldsymbol{c}$ is equal to $\phi_I\left(\boldsymbol{d}_{[1,I]}\right)$ and each of the permutation matrices can be chosen in $M!$ ways, we obtain

$$\gamma_I\left(\boldsymbol{d}_{[1,I]}\right) = \frac{\phi_I\left(\boldsymbol{d}_{[1,I]}\right) \prod_{i=1}^{I} d_i!(M-d_i)!}{(M!)^I},$$

which results in (70). □

For the case $I = 2$, we must have $d_1 = d_2$. Thus $\nu_2 = d_1$, $\nu_0 = M - d_1$, and therefore in this case (69) simplifies to

$$\phi_2\left(\boldsymbol{d}_{[1,2]}\right) = \frac{M!}{d_1!(M-d_1)!}. \tag{71}$$

Hence we have

$$\gamma_2\left(\boldsymbol{d}_{[1,2]}\right) = \frac{\phi_2\left(\boldsymbol{d}_{[1,2]}\right)}{\prod_{i=1}^{2} \binom{M}{d_i}} = \frac{\binom{M}{d_1}}{\binom{M}{d_1}\binom{M}{d_2}} = \frac{1}{\binom{M}{d_1}}. \tag{72}$$

In general, however, the function $\gamma_I\left(\boldsymbol{d}_{[1,I]}\right)$ has a complex structure. Therefore we obtain an upper bound on $\gamma_I\left(\boldsymbol{d}_{[1,I]}\right)$ by first upper bounding $\phi_I\left(\boldsymbol{d}_{[1,I]}\right)$.

Lemma 5: The function $\phi_I\left(\boldsymbol{d}_{[1,I]}\right)$ is upper bounded by the inequality

$$\phi_I\left(\boldsymbol{d}_{[1,I]}\right) \le \exp\Big[M\Big(g_I(\lambda_1, \lambda_2, \ldots, \lambda_I) - \sum_{i=1}^{I} \lambda_i \rho_i\Big)\Big], \tag{73}$$

where $g_I(\lambda_1, \lambda_2, \ldots, \lambda_I)$ is as defined in (8), $\rho_i = d_i/M$, is the normalized Hamming weight of the vector $\boldsymbol{w}_i$, and the λ_i are arbitrary constants, $i = 1, \ldots, I$.

Proof: We present the proof for the $I = 6$ case, but generalization to other values of I is straightforward. Multiply each of the terms in (69) by

$$\exp\Big[\sum_{i_1=1}^{6} \lambda_{i_1}\Big(\sum_{\{i_2\}} \nu_2(\{i_1,i_2\}) + \sum_{\{i_2,i_3,i_4\}} \nu_4(\{i_1,i_2,i_3,i_4\}) + \nu_6 - d_{i_1}\Big)\Big], \tag{74}$$

where each λ_{i_1} is an arbitrary constant. Observe that this does not change the sum in (69) by virtue of the constraints of (68). To obtain an upper bound on $\phi_6(\boldsymbol{d}_{[1,6]})$, we sum over all constraint compositions

$$(\nu_0, \{\nu_2(\{i_1,i_2\})\}, \{\nu_4(\{i_1,i_2,i_3,i_4\})\}, \nu_6)$$

satisfying (67) but not necessarily (68). The multinomial theorem [18], (69), and (74) together imply that

$$\phi_6\left(\boldsymbol{d}_{[1,6]}\right) \le \left[\frac{\prod_{i=1}^{6}(1+e^{\lambda_i}) + \prod_{i=1}^{6}(1-e^{\lambda_i})}{2}\right]^M \cdot e^{-\sum_{i=1}^{6} \lambda_i d_i}, \tag{75}$$

and (73) follows from (8) and (75). □

From Lemmas 4 and 5 we obtain the upper bound

$$\gamma_I\left(\boldsymbol{d}_{[1,I]}\right) \le \frac{\exp\left[M G_I(\boldsymbol{\lambda}_{[1,I]}, \boldsymbol{\rho}_{[1,I]})\right]}{\prod_{i=1}^{I} \binom{M}{M\rho_i}}, \tag{76}$$

where $G_I(\boldsymbol{\lambda}_{[1,I]}, \boldsymbol{\rho}_{[1,I]})$ is as defined in (7).

APPENDIX II
PROOF OF LEMMA 1

From (14) and (16) we have

$$E(\boldsymbol{\rho}_{[1,L]}) = \psi_{[1,L]} \prod_{t=1}^{L} \prod_{h=0}^{1} \binom{M}{M\rho_t^{(h)}} \le \frac{\exp\left(M\tilde{F}(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})\right)}{\left[\prod_{t=1}^{L} \prod_{h=0}^{1} \binom{M}{M\rho_t^{(h)}}\right]^2} \binom{M}{M\rho_1^{(1)}} \binom{M}{M\rho_L^{(1)}}, \tag{77}$$

where for $0 < \rho_t^{(h)} < 1$ we can use the inequalities (see [19])

$$\binom{M}{M\rho_t^{(h)}} \le \exp\left(MH(\rho_t^{(h)})\right) \frac{1}{\sqrt{2\pi M \rho_t^{(h)}(1-\rho_t^{(h)})}} \tag{78}$$

and

$$\binom{M}{M\rho_t^{(h)}}^{-1} \le \exp\left(-MH(\rho_t^{(h)})\right) \sqrt{2\pi M \rho_t^{(h)}(1-\rho_t^{(h)})} \cdot \exp \frac{1}{12M\rho_t^{(h)}(1-\rho_t^{(h)})}. \tag{79}$$

Observing that $\rho_k(1-\rho_k) \ge \frac{1}{M}\left(1 - \frac{1}{M}\right)$, we obtain

$$12M\rho_k(1-\rho_k) > 2 \tag{80}$$

and

$$2\pi M\rho_k(1-\rho_k) > 1 \tag{81}$$

for $M \geq 2$. Thus (78) and (79) can be simplified to

$$\binom{M}{M\rho_t^{(h)}} \leq \exp\left(MH(\rho_t^{(h)})\right) \tag{82}$$

and

$$\binom{M}{M\rho_t^{(h)}}^{-1} \leq \exp\left(-MH(\rho_t^{(h)})\right) \cdot \sqrt{2\pi M\rho_t^{(h)}}\exp\left(\frac{1}{2}\right) . \tag{83}$$

For $\rho_t^{(h)} = 1$,

$$\binom{M}{M\rho_t^{(h)}} = 1 = \exp\left(MH(\rho_t^{(h)})\right) , \tag{84}$$

and therefore (82) and (83) are valid in this case as well.

Now consider only t and h such that $\rho_t^{(h)} \neq 0$. Let us enumerate them with one index k. We assume that there are N such normalized weights, $\rho_1, \rho_2, \ldots, \rho_N$. The corresponding weights are $d_1 = \rho_1 M$, $d_2 = \rho_2 M, \ldots$, $d_N = \rho_N M$. Substituting (82) and (83) into (77) and recalling the equal weight condition for the first and last two blocks of $\boldsymbol{v}_{[1,L]}$, it follows from (21) that

$$E(\boldsymbol{\rho}_{[1,L]}) \leq \exp\left\{M \cdot F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})\right\} \prod_{k=1}^{N} (2\pi e d_k) , \tag{85}$$

which implies (18). In order to find the maximum of the product term we consider a simple optimization problem: for N variables $d_k \geq 1$, $\sum_{k=1}^{N} d_k = d$, find the maximum of the product $\prod_{k=1}^{N} d_k$. It is easy to show that all d_k should be equal to maximize $\prod_{k=1}^{N} d_k$ (otherwise, instead of d_1, d_2, one would take $(d_1+d_2)/2$, $(d_1+d_2)/2$), and the maximum is $\prod_{k=1}^{N} d_k = (d/N)^N$. We are left with

$$E(\boldsymbol{\rho}_{[1,L]}) \leq \exp\left\{M \cdot F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})\right\} \left(\frac{2\pi e d}{N}\right)^N . \tag{86}$$

Now recalling that $N \leq 2L$ and $d \leq MN$ (the maximum possible total weight), we see that (86) implies (19). Moreover, noting that $1 \leq d_k$, $k = 1, \ldots, N$, and

$$N \leq \sum_{k=1}^{N} d_k = d = \rho_{[1,L]} M ,$$

we see that the maximum of (86) is reached when $N = d = \rho_{[1,L]} M$, and it follows that

$$\begin{aligned} E(\boldsymbol{\rho}_{[1,L]}) &\leq \exp\left\{M \cdot F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})\right\} (2\pi e)^d \\ &= \exp\left\{M \cdot F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})\right\} (2\pi e)^{\rho_{[1,L]} M} , \end{aligned} \tag{87}$$

which implies (20). ■

APPENDIX III
PROOF OF LEMMA 2

Before proving Lemma 2, we derive some important properties of the function $U(\boldsymbol{\rho}_{[1,L]})$ defined in (22). First, we note from (21) and (15) that the function $F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})$ consists of several terms. Although $F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})$ depends on the entire set $\boldsymbol{\lambda}_{[1,L]}$, only the term at time t in $\widetilde{F}(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})$ depends on the variables $\lambda_i^{(h)}(t)$, $i = 0, 1, 2$, $h = 0, 1$. Thus, the minimization $\min_{\boldsymbol{\lambda}_{[1,L]}} F(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})$ can be carried out individually for each of the terms in $\widetilde{F}(\boldsymbol{\lambda}_{[1,L]}, \boldsymbol{\rho}_{[1,L]})$.

After some reorganization of terms, and recalling the equal weight condition for the first and last two blocks of $\boldsymbol{v}_{[1,L]}$, it follows that we can represent $U(\boldsymbol{\rho}_{[1,L]})$ as

$$\begin{aligned} U(\boldsymbol{\rho}_{[1,L]}) = U_2(\boldsymbol{\rho}_1) + U_4(\boldsymbol{\rho}_2, \boldsymbol{\rho}_1) + \sum_{t=3}^{L} U_6(\boldsymbol{\rho}_t, \boldsymbol{\rho}_{t-1}, \boldsymbol{\rho}_{t-2}) \\ + U_4(\boldsymbol{\rho}_L, \boldsymbol{\rho}_{L-1}) + U_2(\boldsymbol{\rho}_L) , \end{aligned} \tag{88}$$

where the functions U_2, U_4, and U_6 are defined as

$$\begin{aligned} U_6(\rho_1, \rho_2, \rho_3, \rho_4, \rho_5, \rho_6) \stackrel{\text{def}}{=} \min_{\lambda_1, \ldots, \lambda_6} \Big[g_6(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) \\ - \sum_{i=1}^{6} \lambda_i \rho_i - \frac{2}{3} \sum_{i=1}^{6} H(\rho_i) \Big] , \end{aligned} \tag{89}$$

$$\begin{aligned} U_4(\rho_1, \rho_2, \rho_3, \rho_4) \stackrel{\text{def}}{=} \min_{\lambda_1, \ldots, \lambda_4} \Big[g_4(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \\ - \sum_{i=1}^{4} \lambda_i \rho_i - \frac{2}{3} \sum_{i=1}^{4} H(\rho_i) \Big] , \end{aligned} \tag{90}$$

$$U_2(\rho_1, \rho_2) \stackrel{\text{def}}{=} -\frac{1}{6}(H(\rho_1) + H(\rho_2)) , \tag{91}$$

and we see that $U_2(\rho_1, \rho_2) \leq 0$ for any $\rho_1, \rho_2 \in [0, 1]$.

Proposition 1: For any $\rho_1, \ldots, \rho_6$ such that $\rho_i \in [0, 1)$, $i = 1, \ldots, 6$,

$$U_6(\rho_1, \ldots, \rho_6) \leq U_6\left(\frac{1}{6}\sum_{i=1}^{6}\rho_i, \frac{1}{6}\sum_{i=1}^{6}\rho_i, \ldots, \frac{1}{6}\sum_{i=1}^{6}\rho_i\right) , \tag{92}$$

$$U_4(\rho_1, \ldots, \rho_4) \leq U_4\left(\frac{1}{4}\sum_{i=1}^{4}\rho_i, \ldots, \frac{1}{4}\sum_{i=1}^{4}\rho_i\right) . \tag{93}$$

Proof: A proof is given in [12]. ■

Corollary 1: Consider a normalized weight composition $\boldsymbol{\rho}_{[1,L]}$ such that, for each $t \in \{3, 4, \ldots, L\}$, the sum of six consecutive terms $\rho_{t-2+i}^{(h)}$, $i = 0, 1, 2$, $h = 0, 1$, satisfies

$$\sum_{i=0}^{2}\sum_{h=0}^{1} \rho_{t-2+i}^{(h)} < 6 \cdot 0.023 = 0.138 . \tag{94}$$

Then

$$U(\boldsymbol{\rho}_{[1,L]}) < 0 . \tag{95}$$

Proof: We consider (88) and apply the bounds from Proposition 1 to every term on the right hand side. For

$t \in \{3,4,\ldots,L\}$ we obtain

$$U_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2}) \leq U_6\Bigg(\frac{1}{6}\sum_{i=0}^{2}\sum_{h=0}^{1}\rho_{t-2+i}^{(h)}, \frac{1}{6}\sum_{i=0}^{2}\sum_{h=0}^{1}\rho_{t-2+i}^{(h)}, \ldots, \frac{1}{6}\sum_{i=0}^{2}\sum_{h=0}^{1}\rho_{t-2+i}^{(h)}\Bigg) < 0 \ , \quad (96)$$

since (94) implies

$$\frac{1}{6}\sum_{i=0}^{2}\sum_{h=0}^{1}\rho_{t-2+i}^{(h)} < 0.023 \ ,$$

where 0.023 is approximately equal to the smallest positive root of $U_6(\rho,\ldots,\rho)$ (note that 0.023 is the value of $\alpha_{\mathrm{G}}(3,6)$ obtained by Gallager for (3,6) LDPC block codes). The term $U_4(\boldsymbol{\rho}_2,\boldsymbol{\rho}_1)$ can be upperbounded as

$$U_4(\boldsymbol{\rho}_2,\boldsymbol{\rho}_1) \leq U_4\left(\frac{1}{4}\sum_{t=1}^{2}\sum_{h=0}^{1}\rho_t^{(h)},\ldots,\frac{1}{4}\sum_{t=1}^{2}\sum_{h=0}^{1}\rho_t^{(h)}\right) < 0 \ , \quad (97)$$

since (94) implies

$$\frac{1}{4}\sum_{t=1}^{2}\sum_{h=0}^{1}\rho_t^{(h)} < \frac{0.138}{4} < 0.112 \ ,$$

where 0.112 is approximately equal to the smallest positive root of $U_4(\rho,\ldots,\rho)$. Analogously, we can show that $U_4(\boldsymbol{\rho}_L,\boldsymbol{\rho}_{L-1}) < 0$. Finally, since the terms $U_2(\boldsymbol{\rho}_1)$ and $U_2(\boldsymbol{\rho}_L)$ cannot be positive, (95) follows. ∎

Proposition 2: For any $\rho_1,\ldots,\rho_6$ such that $\rho_i \in [0,1)$, $i = 1,\ldots,6$,

$$U_6(\rho_1,\ldots,\rho_6) \leq 6 f_6\left(\frac{1}{6}\sum_{i=1}^{6}\rho_i\right) \ , \quad (98)$$

$$U_4(\rho_1,\ldots,\rho_4) \leq 4 f_6\left(\frac{1}{4}\sum_{i=1}^{4}\rho_i\right) \ , \quad (99)$$

where $f_6(\rho)$ is defined in (35).

Proof: From numerical calculations of $U_6(\rho,\ldots,\rho)$ for $\rho \in [0,1)$, we determine that it is upper bounded by

$$\rho(3\ln 5 - 1 + \ln\rho) = 6 f_6(\rho) \ .$$

Then we lower bound $U_6(\rho,\ldots,\rho)$ using Proposition 1 to obtain (98). Function $U_4(\cdot)$ is upper bounded in the same way. ∎

We now note that the smallest positive root of $f_6(\cdot)$ is slightly larger than 0.021, and we define $\rho_{\mathrm{f}} \stackrel{\mathrm{def}}{=} 0.021$. For $\rho \in (0,\rho_{\mathrm{f}}]$, the function $f_6(\rho) < 0$.

Proposition 3: For any combination $\rho_1,\ldots,\rho_6$ such that $\rho_i \in [0,1)$, $i = 1,\ldots,6$,

$$U_6(\rho_1,\ldots,\rho_6) \leq \sum_{i=1}^{6} f_6(\rho_i) \ , \quad (100)$$

$$U_4(\rho_1,\ldots,\rho_4) \leq \sum_{i=1}^{4} f_6(\rho_i) \ , \quad (101)$$

and

$$U_2(\rho_1,\rho_2) \leq f_6(\rho_1) + f_6(\rho_2) \ . \quad (102)$$

Proof: We can write

$$U_6(\rho_1,\ldots,\rho_6) \leq 6 f_6\left(\frac{1}{6}\sum_{i=1}^{6}\rho_i\right) \leq \sum_{i=1}^{6} f_6(\rho_i) \ , \quad (103)$$

where the first inequality in (103) follows from Proposition 2 and the second from the convexity of the function $f_6(\cdot)$. The same argument holds for $U_4(\cdot)$. For $U_2(\cdot)$, we make use of the fact that $-\frac{1}{6}H(\rho) \leq f_6(\rho)$ for $\rho \in [0,1]$. ∎

Proposition 4: For any positive integer k and $\rho_1,\ldots,\rho_k$ such that $\rho_i \in [0,1)$, $i = 1,\ldots,k$,

$$\sum_{i=1}^{k} f_6(\rho_i) \leq f_6\left(\sum_{i=1}^{k}\rho_i\right) \ . \quad (104)$$

Proof:

$$\begin{aligned}\sum_{i=1}^{k} f_6(\rho_i) &= \sum_{i=1}^{k}\frac{1}{6}\rho_i(3\ln 5 - 1 + \ln\rho_i) \\ &= \frac{1}{6}(3\ln 5 - 1)\sum_{i=1}^{k}\rho_i + \sum_{i=1}^{k}\frac{1}{6}\rho_i\ln\rho_i \\ &\leq \frac{1}{6}(3\ln 5 - 1)\sum_{i=1}^{k}\rho_i + \sum_{i=1}^{k}\frac{1}{6}\rho_i\ln\left(\sum_{j=1}^{k}\rho_j\right) \\ &= f_6\left(\sum_{i=1}^{k}\rho_i\right) \ . \quad (105)\end{aligned}$$

∎

Proof of Lemma 2: Starting with (88) and upper bounding each term on the right hand side using Proposition 3, we obtain

$$U(\boldsymbol{\rho}_{[1,L]}) \leq 3\sum_{t=1}^{L}\sum_{h=0}^{1} f_6(\rho_t^{(h)}) \ . \quad (106)$$

(Note that every normalized weight $\rho_t^{(h)}$ belongs to exactly three terms in (88), and hence $f_6(\rho_t^{(h)})$ appears in the inequality exactly three times for each t and h.) Now Proposition 4 implies that

$$U(\boldsymbol{\rho}_{[1,L]}) \leq 3\sum_{t=1}^{L}\sum_{h=0}^{1} f_6(\rho_t^{(h)}) \leq 3 f_6\left(w(\boldsymbol{\rho}_{[1,L]})\right) \ , \quad (107)$$

which, along with (23), leads directly to (34). ∎

APPENDIX IV
A LOWER BOUND ON ρ_L^*

Proposition 5: For ρ_L^* defined in (25), and for any L,

$$\rho_L^* \geq \rho^* = 0.5 \ . \quad (108)$$

Proof: Consider an arbitrary $\boldsymbol{\rho}_{[1,L]}$ such that

$$w(\boldsymbol{\rho}_{[1,L]}) \leq \rho^* = 0.5$$

and subdivide this sequence into tuples of length six. There can be at most three 6-tuples whose sum

$$\sum_{t'=t-2}^{t}\sum_{h=0}^{1}\rho_{t'}^{(h)} \geq 0.126 = 6\rho_{\mathrm{f}} \ .$$

Now assume that there are exactly three such 6-tuples,

$$\rho_{t_1}^{(1)},\rho_{t_1}^{(0)},\ldots,\rho_{t_1-2}^{(0)},\quad \rho_{t_2}^{(1)},\rho_{t_2}^{(0)},\ldots,\rho_{t_2-2}^{(0)},\text{ and } \rho_{t_3}^{(1)},\rho_{t_3}^{(0)},\ldots,\rho_{t_3-2}^{(0)}$$

and consider the set of time instants

$$T_{\max}=\{t_1-6,t_1-5,\ldots,t_1+6\}\cup\{t_2-6,t_2-5,\ldots,t_2+6\} \cup\{t_3-6,t_3-5,\ldots,t_3+6\}\ . \quad (109)$$

Observe that $|T_{\max}|\le 3\cdot 13=39$. Using numerical optimization we can show that

$$\sum_{t\in T_{\max}} U_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})<-10^{-5}\ . \quad (110)$$

For any other 6-tuple,

$$\sum_{t'=t-2}^{t}\sum_{h=0}^{1}\rho_{t'}^{(h)}<0.126\ ,$$

and it follows from Proposition 2 and the fact that $\rho_{\rm f}=0.021$ is slightly less than the smallest positive root of $f_6(\rho)$ that $U_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})<0$, and the same holds for $U_4(\cdot)$ and $U_2(\cdot)$. Combining these results into (88) we obtain

$$U(\boldsymbol{\rho}_{[1,L]})<0\ . \quad (111)$$

If there are two, one, or zero "high-weight" 6-tuples whose sum

$$\sum_{t'=t-2}^{t}\sum_{h=0}^{1}\rho_{t'}^{(h)}\ge 0.126\ ,$$

we can use the same arguments to show (111). Since (111) holds for any $\boldsymbol{\rho}_{[1,L]}$ such that $w(\boldsymbol{\rho}_{[1,L]})\le\rho^*=0.5$, (108) follows from (25). ∎

APPENDIX V
PROOF OF LEMMA 3

Proposition 6: There exists an M' such that for any $M>M'$ and any $\rho\in\left[\frac{1}{M},\rho_{\rm f}=0.021\right]$,

$$f_6(\rho)+\frac{1}{3M}\ln(2\pi e\rho M)\le\frac{\alpha}{M}-\frac{\ln M}{6M}\ , \quad (112)$$

where

$$\alpha=\frac{1}{2}\left(\ln 5-\frac{1}{3}\right)+\frac{\ln 2\pi e}{3}\ . \quad (113)$$

Proof: To find maximum of the function

$$\varphi(\rho)\stackrel{\rm def}{=}f_6(\rho)+\frac{1}{3M}\ln(2\pi e\rho M) =\frac{1}{2}\left(\ln 5-\frac{1}{3}\right)\rho+\frac{1}{6}\rho\ln\rho+\frac{\ln(2\pi e\rho M)}{3M}\ , \quad (114)$$

we calculate its first and second derivative as follows:

$$\frac{\partial\varphi(\rho)}{\partial\rho}=\frac{1}{2}\ln 5+\frac{1}{6}\ln\rho+\frac{1}{3\rho M} \quad (115)$$

$$\frac{\partial^2\varphi(\rho)}{\partial\rho^2}=\frac{1}{6\rho}-\frac{1}{3\rho^2M}\ . \quad (116)$$

By looking at the sign of $\frac{\partial^2\varphi(\rho)}{\partial\rho^2}$, we observe that $\frac{\partial\varphi(\rho)}{\partial\rho}$ is decreasing in the interval $\left[\frac{1}{M},\frac{2}{M}\right]$ and increasing in the interval $\left[\frac{2}{M},\rho_{\rm f}\right]$. Moreover,

$$\left.\frac{\partial\varphi(\rho)}{\partial\rho}\right|_{\rho=\frac{1}{M}}=\frac{1}{2}\ln 5+\frac{1}{3}-\frac{1}{6}\ln M<0 \quad (117)$$

and

$$\left.\frac{\partial\varphi(\rho)}{\partial\rho}\right|_{\rho=\rho_{\rm f}}=0.1608+\frac{1}{0.063M}>0 \quad (118)$$

for large M. Therefore, $\frac{\partial\varphi(\rho)}{\partial\rho}$ has exactly one root in the interval $\left[\frac{1}{M},\rho_{\rm f}\right]$. This root corresponds to the point of minimum for $\varphi(\rho)$, and the maximum is achieved on the boundary for $\rho=\frac{1}{M}$ or $\rho=\rho_{\rm f}$. For these values of ρ, we obtain

$$\varphi\left(\frac{1}{M}\right)=\frac{1}{M}\left(\frac{1}{2}\left(\ln 5-\frac{1}{3}\right)+\frac{\ln 2\pi e}{3}\right) -\frac{1}{6M}\ln M\ , \quad (119)$$

$$\varphi\left(\rho_{\rm f}\right)=-1.22\times 10^{-4}+\frac{\ln(0.359M)}{3M}\ , \quad (120)$$

and for large M the first expression, which coincides with the right hand side of (112), is greater. ∎

Now we define the function

$$\tilde{U}_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})\stackrel{\rm def}{=}U_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2}) +\sum_{\rho_t^{(h)}\in\mathcal{N}_t}\frac{\ln(2\pi eM\rho_t^{(h)})}{3M}\ , \quad (121)$$

where $\mathcal{N}_t$ is the set of positions corresponding to nonzero values of $\rho_t^{(h)}$ in the 6-tuple $(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})$. We also define $\tilde{U}_2(\boldsymbol{\rho}_t)$ and $\tilde{U}_4(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1})$ analogously, where in these cases $\mathcal{N}_t$ denotes the set of positions corresponding to nonzero values in the 2-tuple $(\boldsymbol{\rho}_t)$ or 4-tuple $(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1})$, respectively.

Corollary 2: Consider a 6-tuple $(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})$ such that $0<\sum_{\rho_t^{(h)}\in\mathcal{N}_t}\rho_t^{(h)}<6\rho_{\rm f}=0.126$. Then there exists an M' such that for any $M>M'$,

$$\tilde{U}_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})\le|\mathcal{N}_t|\left(\frac{\alpha}{M}-\frac{\ln M}{6M}\right)<0\ , \quad (122)$$

$$\tilde{U}_4(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1})\le|\mathcal{N}_t|\left(\frac{\alpha}{M}-\frac{\ln M}{6M}\right)<0\ ,$$

$$\tilde{U}_2(\boldsymbol{\rho}_t)\le|\mathcal{N}_t|\left(\frac{\alpha}{M}-\frac{\ln M}{6M}\right)<0\ .$$

Proof: Consider $\rho=\frac{1}{6}\sum_{\rho_t^{(h)}\in\mathcal{N}_t}\rho_t^{(h)}$. From Proposition 2, Proposition 6, and the convexity of the function $\ln(\cdot)$, it follows that

$$\begin{aligned}&\tilde{U}_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})\\ &=U_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})+\sum_{\rho_t^{(h)}\in\mathcal{N}_t}\frac{\ln(2\pi eM\rho_t^{(h)})}{3M}\\ &\le 6f_6(\rho)+6\frac{\ln(2\pi eM\rho)}{3M}\\ &<6\left(\frac{\alpha}{M}-\frac{\ln M}{6M}\right)\le|\mathcal{N}_t|\left(\frac{\alpha}{M}-\frac{\ln M}{6M}\right)<0\ , \quad (123)\end{aligned}$$

where the last two inequalities hold for large enough M. Similar arguments lead to the bounds on $U_2(\cdot)$ and $U_4(\cdot)$. ■

Additionally, we note that

$$\sum_{\rho_t^{(h)}\in\mathcal{N}_t} \frac{\ln(2\pi eM\rho_t^{(h)})}{3M} \to 0 \tag{124}$$

as $M\to\infty$, which leads to the next corollary.

Corollary 3: Consider a 6-tuple $(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})$ such that $0<\delta<\sum_{\rho_t^{(h)}\in\mathcal{N}_t}\rho_t^{(h)}<6\rho_{\mathrm{f}}=0.126$. Then, for M large enough,

$$\tilde{U}_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})<6\max\{f_6(\delta/6),f_6(\rho_{\mathrm{f}})\}<0\ . \tag{125}$$

Proof: First we use Proposition 2 to obtain

$$\begin{aligned}&\tilde{U}_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})\\ &=U_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})+\sum_{\rho_t^{(h)}\in\mathcal{N}_t}\frac{\ln(2\pi eM\rho_t^{(h)})}{3M}\\ &\le 6f_6\left(\frac{1}{6}\sum_{\rho_t^{(h)}\in\mathcal{N}_t}\rho_t^{(h)}\right)+\sum_{\rho_t^{(h)}\in\mathcal{N}_t}\frac{\ln(2\pi eM\rho_t^{(h)})}{3M}\ .\end{aligned} \tag{126}$$

Then, following the same procedure as in the proof of Proposition 6, we find that $f_6(\cdot)$ achieves its maximum at one of the boundary points, either $\delta/6$ or ρ_{f}. Using this maximum, which is a negative constant, as an upper bound, and making use of (124), we obtain (125). ■

Proof of Lemma 3: It follows from (88) and the definition of $\tilde{U}(\cdot)$ that

$$\begin{aligned}&U(\boldsymbol{\rho}_{[1,L]})+\sum_{\rho_t^{(h)}\neq 0}\frac{1}{M}\left(\ln(2\pi e)+\ln M\rho_t^{(h)}\right)\\ &\quad=\sum_{t=3}^{L}\tilde{U}_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})+\tilde{U}_4(\boldsymbol{\rho}_2,\boldsymbol{\rho}_1)+\tilde{U}_2(\boldsymbol{\rho}_1)\\ &\quad\quad+\tilde{U}_4(\boldsymbol{\rho}_L,\boldsymbol{\rho}_{L-1})+\tilde{U}_2(\boldsymbol{\rho}_L)\ ,\end{aligned} \tag{127}$$

where we note that each time unit is included three times in the sum on the right hand side of (127). Now we consider three different cases.

Case 1: There exists a 6-tuple $(\rho_t^{(1)},\rho_t^{(0)},\ldots,\rho_{t-2}^{(0)})$ such that $\sum_{\rho_t^{(h)}\in\mathcal{N}_t}\rho_t^{(h)}\ge 0.126$.

We follow the same procedure as in the proof of Proposition 5 and consider the set of indices $T_{\max}$ in (109). By $N_{\max}$ we denote the total number of nonzero elements $\rho_t^{(h)}$ in the 6-tuples $(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})$, for all $t\in T_{\max}$. (If the 6-tuples overlap, we count repetitions of overlapping elements.) If all elements in the 6-tuples indexed by $t\in T_{\max}$ are nonzero, then $N_{\max}$ attains its maximum value $6\cdot 39=234$. Using the numerical optimization result of (110), it follows that

$$\begin{aligned}&\sum_{t\in T_{\max}}\tilde{U}_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})\\ &\quad=\sum_{t\in T_{\max}}U_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})\\ &\quad\quad+\sum_{t\in T_{\max}}\sum_{\rho_t^{(h)}\in\mathcal{N}_t}\frac{\ln(2\pi eM\rho_t^{(h)})}{3M}\end{aligned} \tag{128}$$

$$<-10^{-5}+N_{\max}\frac{\ln(2\pi eM)}{3M}<-10^{-5}\stackrel{\text{def}}{=}-\epsilon\ , \tag{129}$$

where the last inequality holds for large enough M and we have used the fact that $\rho_t^{(h)}\le 1$ for all t and h.

Now, for $t\in\{1,\ldots,L\}\setminus T_{\max}$, we use the bound from Corollary 2 to obtain

$$\begin{aligned}\tilde{U}_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},&\boldsymbol{\rho}_{t-2})\\ &\le|\mathcal{N}_t|\left(\frac{\alpha}{M}-\frac{\ln M}{6M}\right),\\ &\qquad t\in\{3,\ldots,L\}\setminus T_{\max}\ ,\\ \tilde{U}_4(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},)&\le|\mathcal{N}_t|\left(\frac{\alpha}{M}-\frac{\ln M}{6M}\right),\ t=\{2,L\}\setminus T_{\max}\ ,\\ \tilde{U}_2(\boldsymbol{\rho}_t)&\le|\mathcal{N}_t|\left(\frac{\alpha}{M}-\frac{\ln M}{6M}\right),\ t=\{1,L\}\setminus T_{\max}\ .\end{aligned} \tag{130}$$

Combining (129) and (130) and substituting into (127) gives

$$\begin{aligned}&U(\boldsymbol{\rho}_{[1,L]})+\sum_{\rho_t^{(h)}\neq 0}\frac{1}{M}\left(\ln(2\pi eM\rho_t^{(h)})\right)\\ &\quad\le-\epsilon+(3N-N_{\max})\left(\frac{\alpha}{M}-\frac{\ln M}{6M}\right)\end{aligned} \tag{131}$$

$$\begin{aligned}&\le-\epsilon+N\left(\frac{3\alpha}{M}-\frac{\ln M}{2M}\right)\\ &\quad-N_{\max}\left(\frac{\alpha}{M}-\frac{\ln M}{6M}\right)\ ,\end{aligned} \tag{132}$$

where again we note that each non-zero element $\rho_t^{(h)}$ will appear three times in the sum on the right hand side of (127) and M is chosen large enough to guarantee that $\frac{\alpha}{M}-\frac{\ln M}{6M}<0$. Since the middle term on the right hand side of (132) is negative, and ϵ,α, and $N_{\max}$ are positive constants, (60) follows.

Case 2: Case 1 is not true, but there exists at least one pair t,h such that $\frac{\delta''}{2}\le\rho_t^{(h)}$.

We consider the 6-tuple containing $\rho_t^{(h)}$ and use Corollary 3 to obtain

$$\begin{aligned}&\tilde{U}_6(\boldsymbol{\rho}_t,\boldsymbol{\rho}_{t-1},\boldsymbol{\rho}_{t-2})\\ &\qquad\le 6\max\left\{f_6\left(\frac{\delta''}{12}\right),f_6(\rho_{\mathrm{f}})\right\}\stackrel{\text{def}}{=}-\epsilon<0\end{aligned} \tag{133}$$

for M large enough. For the other terms we use Corollary 2,

resulting in

$$U(\boldsymbol{\rho}_{[1,L]}) + \sum_{\rho_t^{(h)}\neq 0} \frac{1}{M}\left(\ln(2\pi e M\rho_t^{(h)}\right) \leq -\epsilon + (3N-6)\left(\frac{\alpha}{M} - \frac{\ln M}{6M}\right) \quad (134)$$

$$= -\epsilon + N\left(\frac{3\alpha}{M} - \frac{\ln M}{2M}\right) - 6\left(\frac{\alpha}{M} - \frac{\ln M}{6M}\right) \ , \quad (135)$$

and (60) again follows.

Case 3: $\rho_t^{(h)} < \frac{\delta''}{2}$ for all $t \in \{1,2,\ldots,L\}$, $h = 0,1$.

By $\mathcal{TH}_1$ we denote the set of index pairs (t,h) corresponding to the k largest elements $\rho_t^{(h)}$, where k is chosen such that

$$\frac{\delta''}{2} \leq \sum_{(t,h)\in\mathcal{TH}_1} \rho_t^{(h)} < \delta'' \ , \quad (136)$$

and $k = |\mathcal{TH}_1| \leq \frac{N}{2}$. The remaining index pairs we denote by $\mathcal{TH}_2$. Now using (106) we can write

$$U(\boldsymbol{\rho}_{[1,L]}) + \sum_{\rho_t^{(h)}\neq 0} \frac{1}{M}\left(\ln(2\pi e) + \ln M\rho_t^{(h)}\right) \leq \sum_{\rho_t^{(h)}\neq 0} 3f_6(\rho_t^{(h)}) + \sum_{\rho_t^{(h)}\neq 0} \frac{1}{M}\left(\ln(2\pi e) + \ln M\rho_t^{(h)}\right) \quad (137)$$

$$= 3\sum_{(t,h)\in\mathcal{TH}_1} f_6(\rho_t^{(h)}) + \sum_{(t,h)\in\mathcal{TH}_1} \frac{\ln(2\pi e M\rho_t^{(h)})}{M} + 3\sum_{(t,h)\in\mathcal{TH}_2}\left(f_6(\rho_t^{(h)}) + \frac{\ln(2\pi e M\rho_t^{(h)})}{3M}\right) \ . \quad (138)$$

We use now Proposition 4 to upper bound the first term of (138) and the linear bound $\ln x \leq x$ to upper bound the second term of (138), which gives

$$3\sum_{(t,h)\in\mathcal{TH}_1} f_6(\rho_t^{(h)}) + \sum_{(t,h)\in\mathcal{TH}_1} \frac{\ln(2\pi e M\rho_t^{(h)})}{M}] \leq 3f_6\left(\sum_{(t,h)\in\mathcal{TH}_1}\rho_t^{(h)}\right) + \ln(2\pi e)\sum_{(t,h)\in\mathcal{TH}_1}\rho_t^{(h)} \leq 3f_6\left(\frac{\delta''}{2}\right) + \ln(2\pi e)\frac{\delta''}{2} \stackrel{\text{def}}{=} -\epsilon < 0 \ , \quad (139)$$

where the second inequality follows from (136) and the fact that $f_6'(\rho) < 0$ for all $\rho \in [0,1)$.

Finally, we use Proposition 6 to bound the third term, which gives

$$3\sum_{(t,h)\in\mathcal{TH}_2}\left(f_6(\rho_t^{(h)}) + \frac{\ln(2\pi e M\rho_t^{(h)})}{3M}\right) \leq 3|\mathcal{TH}_2|\left(\frac{\alpha}{M} - \frac{\ln M}{6M}\right) \leq \frac{N}{2}\left(\frac{3\alpha}{M} - \frac{\ln M}{2M}\right) \ , \quad (140)$$

where the last inequality follows from the fact that $|\mathcal{TH}_2| \geq N/2$ and $\frac{\alpha}{M} - \frac{\ln M}{6M} < 0$ for M large enough. Now we combine (138), (139), and (140) to obtain

$$U(\boldsymbol{\rho}_{[1,L]}) + \sum_{\rho_t^{(h)}\neq 0} \frac{1}{M}\left(\ln(2\pi e) + \ln(M\rho_t^{(h)})\right) \leq -\epsilon + \frac{N}{2}\left(\frac{3\alpha}{M} - \frac{\ln M}{2M}\right) \ , \quad (141)$$

and (60) again follows.

Author biographies for IT submission 5-352

Arvind Sridharan was born in Madras, India. He received the B.Tech degree in electrical engineering from the Indian Institute of Technology, Madras, India in 1999 the M.S. and Ph.D degrees in electrical engineering from the University of Notre Dame, Notre Dame, IN, in 2001 and 2005, respectively. Since February 2005 he has been with the Coding and signal procesing group at Seagate Technology in Longmont, Colorado. His research interests include coding theory, iterative algorithms and information theory.

Dmitri Truhachev (S'00-M'04) was born in Saint Petersburg, Russia, in 1978. He received the B.S. degree in applied mathematics from Saint Petersburg State Electrotechnical University, Russia, in 1999 and the Ph.D. degree in electrical engineering in 2004 from Lund University, Sweden. In 2004 he joined High Capacity Digital Communications Laboratory at University of Alberta, Edmonton, Canada as a Postdoctoral Fellow. His major research interests include communications, error control coding, information theory, and ad-hoc wireless networks.

Michael Lentmaier (S'98-M'03) was born in Ellwangen, Germany. He received the Dipl.-Ing. degree in electrical engineering from University of Ulm, Ulm, Germany in 1998, and the Ph.D. degree in telecommunication theory from Lund University, Lund, Sweden in 2003. As a Postdoctoral Research Associate he spent 15 months with the coding research group at University of Notre Dame, Indiana, and four months in the Department of Telecommunications and Applied Information Theory at University of Ulm. Since January 2005, he has been with the Institute of Communications and Navigation at the German Aerospace Center (DLR), Oberpfaffenhofen, Germany. His research interests include coding theory, with emphasis on iterative decoding of block and convolutional codes, and sequential Bayesian estimation techniques with applications to multipath mitigation in navigation receivers.

Daniel J. Costello, Jr. was born in Seattle, WA, on August 9, 1942. He received the B.S.E.E. degree from Seattle University, Seattle, WA, in 1964, and the M.S. and Ph.D. degrees in electrical engineering from the University of Notre Dame, Notre Dame, IN, in 1966 and 1969, respectively.

In 1969 he joined the faculty of the Illinois Institute of Technology, Chicago, IL, as an Assistant Professor of Electrical Engineering. He was promoted to Associate Professor in 1973, and to Full Professor in 1980. In 1985 he became Professor of Electrical Engineering at the University of Notre Dame, Notre Dame, IN, and from 1989 to 1998 served as Chair of the Department of Electrical Engineering. In 1991, he was selected as one of 100 Seattle University alumni to receive the Centennial Alumni Award in recognition of alumni who have displayed outstanding service to others, exceptional leadership, or uncommon achievement. In 1999, he received a Humboldt Research Prize from the Alexander von Humboldt Foundation in Germany. In 2000, he was named the Leonard Bettex Professor of Electrical Engineering at Notre Dame.

Dr. Costello has been a member of IEEE since 1969 and was elected Fellow in 1985. Since 1983, he has been a member of the Information Theory Society Board of Governors, and in 1986 served as President of the BOG. He has also served as Associate Editor for Communication Theory for the IEEE Transactions on Communications, Associate Editor for Coding Techniques for the IEEE Transactions on Information Theory, and Co-Chair of the IEEE International Symposia on Information Theory in Kobe, Japan (1988), Ulm, Germany (1997), and Chicago, IL (2004). In 2000, he was selected by the IEEE Information Theory Society as a recipient of a Third-Millennium Medal.

Dr. Costello's research interests are in the area of digital communications, with special emphasis on error control coding and coded modulation. He has numerous technical publications in his field, and in 1983 co-authored a textbook entitled "Error Control Coding: Fundamentals and Applications", the 2nd edition of which was published in 2004.

Kamil Sh. Zigangirov was born in the U.S.S.R. in 1938. He received the M.S. degree in 1962 from the Moscow Institute for Physics and Technology, Moscow, U.S.S.R., and the Ph.D. degree in 1966 from the Institute of Radio Engineering and Electronics of the U.S.S.R. Academy of Sciences, Moscow, U.S.S.R.

From 1965 to 1991, he held various research positions at the Institute for Problems of Information Transmission of the U.S.S.R. Academy of Sciences, Moscow, first as a Junior Scientist, and later as a Main Scientist. During this period, he visited several universities in the United States, Sweden, Italy, and Switzerland as a Guest Researcher. He organized several symposia on information theory in the U.S.S.R. In 1994, he received the Chair of Telecommunication Theory at Lund University, Lund, Sweden. In 2003 and 2004, he has been a Visiting Professor at the University of Notre Dame, Notre Dame, IN, and at the University of Alberta, Edmonton, AB, Canada. His scientific interests include information theory, coding theory, de- tection theory, and mathematical statistics. In addition to papers in these areas, he published a book on sequential decoding of convolutional codes (in Russian) in 1974. With R. Johannesson, he coauthored the textbook Fundamentals of Convolutional Coding (Piscataway, NJ: IEEE Press, 1999). His book Theory of CDMA Communication was published by IEEE Press in 2004.