Subspace Decomposition of Extreme-Rate Secrecy Codes

David Hunn
Department of Electrical and Computer Engineering
Brigham Young University
Provo, Utah 84602
Email: david.hunn@byu.edu

Willie K. Harrison
Department of Electrical and Computer Engineering
Brigham Young University
Provo, Utah 84602
Email: willie.harrison@byu.edu

2022 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-2159-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISIT50566.2022.9834735

***Abstract*—The performance and structure of wiretap codes are analyzed in the limit of large code size and very low- or very high-rate codes. Under these conditions, code performance may be calculated using the properties of subspaces of the available code space. Using this technique, a code defined by a generator matrix with a uniform distribution of nonzero columns is proven to be locally optimal.**

I. Introduction

Wiretap codes have been explored now for several decades as a mechanism for secure and reliable data transmission over communication channels with eavesdroppers [1], [2], [3], [4], [5]. The performance limits of wiretap codes in the limit of large blocklength are well understood [6], [7], [8]. A recent line of inquiry has focused on optimal finite blocklength wiretap codes. In this regime, the codes operating over the type-II wiretap channel are analyzed by, e.g., generalized Hamming weights [9], [10], while over the binary erasure wiretap channel (BEWC) combinatoric analysis techniques are applied to the problem of designing *best* codes [11], [12], [13], [14], [15]. In this work, we study an intermediate class of codes with large blocklength but extremely high or low rates operating over the BEWC. Under these conditions, simplifying analytical assumptions permit the code's performance to be expressed in terms of subspaces of the code's generator matrix.

The paper is organized as follows. Section II provides background material on coset code construction over binary erasure wiretap channels. The extreme-rate analysis is provided in Section III, and the main results of the paper are given in Section IV.

II. Coset Code Background

The wiretap channel used in this paper is the binary erasure wiretap channel of Wyner. In this channel, a message is encoded and sent by the transmitter. It is then received via a perfect channel by a legitimate receiver and received via a binary erasure channel (BEC) by an eavesdropper. The usual form of secrecy coding for this type of channel is called coset coding and involves expanding the original k-symbol message to an n-symbol codeword. This is done by creating a base $(n, n-k)$ linear block code C defined by generator matrix G, then assigning each possible message m to one of the cosets of C. This may be accomplished in practice by defining an auxiliary generator matrix G' which is linearly independent relative to G. A random $(n-k)$-bit auxiliary message m' is then appended to the message m, and the resulting vector is multiplied by the matrix G^*, which is formed by the vertical concatenation of G and G'. The final codeword x is given by

$$x = \begin{bmatrix} m & m' \end{bmatrix} \begin{bmatrix} G' \\ G \end{bmatrix}. \tag{1}$$

To calculate the eavesdropper's equivocation $H(M \mid Z)$, the equivocation for a given value z of Z is averaged over all possible values of z. An important result arising from [12] is that the equivocation may be calculated based on the set, denoted $r(z)$, of revealed bit positions of the codeword z, as

$$H(M \mid Z = z) = H(M) - |r(z)| + \operatorname{rank}(G_{r(z)}), \tag{2}$$

where $G_{r(z)}$ is the submatrix of G formed by concatenating the columns of G which are indexed by the set $r(z)$. It is assumed throughout this work that the message is uniformly distributed across the 2^k possibilities and, therefore, $H(M) = k$.

The message equivocation of a given coset code depends on the particulars of the code and the erasure probability ϵ of the BEC. Calculating the expected message equivocation involves determining the equivocation and probability of each possible erasure pattern $r(z)$ received by the eavesdropper, as shown in (2). This information is commonly summarized in an equivocation matrix. For a given (n, k) coset code with $(n-k) \times n$ generator G, the equivocation matrix A is a $(k+1) \times (n+1)$ matrix with elements defined by

$$A_{i,j} = |\{r : |r(z)| = j-1, k - H(M \mid r(z)) = i-1\}|. \tag{3}$$

III. Analytical Methods and Definitions

In this work, we consider codes for which the codeword size n is much greater than either the message size k or the random vector size $n-k$. Specifically, we utilize one of the following assumptions:

$$n \gg 2^{n-k} \tag{4}$$

$$n \gg 2^k. \tag{5}$$

This work was funded by the US National Science Foundation: Grant Award Number #1910812.

Because of the relationships between a coset code and its dual [13], the analyses which follow from assumptions (4) and (5), referred to as the "high-rate assumption" and the "low-rate assumption", respectively, are analogous. The analysis in the remainder of this work is performed in terms of the high-rate assumption. In this configuration, the generator matrix G consists of a few rows and many columns.

A. Vector-Fraction Code Specification

When considering such a matrix, the first useful observation is that because of the extreme width of the matrix, it is necessary that the contents of some columns will be repeated. In fact, the bulk of the matrix will necessarily consist of repeated columns. This observation leads to a computational simplification. Rather than specifying the generator matrix element-by-element, it may be specified by listing, for each of the $2^{n-k}-1$ possible nonzero column vectors, the fraction of the total columns equal to that vector. Let $u \in \mathbb{F}_2^{(n-k)}$ denote a particular unique binary vector of size $n-k$, then let $b(u)$ represent the binary integer representation of u and $v(x)$ represent the binary vector expansion of the integer x. Then define the vector q of size $2^{n-k}-1$ where each $q_x, 1 \leq x \leq 2^{n-k}-1$, represents the fraction of the columns of G equal to $v(x)$. Thus, if $n-k=3$, $q_3 = 0.2$ indicates that one fifth of the columns of G are equal to $[0\ 1\ 1]^T$.

B. Equivocation Probability Matrix

This method of specifying the generator matrix naturally suggests a new method of specifying the performance of the corresponding coset code. Instead of reporting the equivocation matrix A, it is convenient to specify an equivocation probability matrix, denoted $\mathcal{A}$. Whereas the equivocation matrix expresses the number of erasure patterns of a given size that result in a given number of bits of message information being leaked, the equivocation probability matrix expresses the probability, conditional on an erasure pattern of a given size, of a given number of message bits being leaked. Symbolically,

$$\mathcal{A}_{i,j} = Pr\left[k - H(M\,|\,r(z)) = i-1 \mid |r(z)| = j-1\right]. \quad (6)$$

Thus, $\mathcal{A}$ is related to A in that the element values of $\mathcal{A}$ equal the fraction of the total of the column elements that are represented by the corresponding element of A.

For a given coset code defined by n, k, and $q = [q_1, q_2, \cdots q_{2^k-1}]$, (4) provides a computational simplification which enables the calculation of values in the equivocation probability matrix. If the number of columns equal to any u is large, the probability of any given erasure occurring in a column equal to u_i may be assumed to be equal to $q_{b(u_i)}$ regardless of the number of erasures occurring in other columns equal to u_i. That is, the number of erasures occurring in columns equal to u_i is assumed to have a Poisson distribution rather then a binomial distribution. Using this assumption, an expression for the i,j element of $\mathcal{A}$ is

$$\mathcal{A}_{i,j} = \sum_{\substack{x_1,x_2,\ldots,x_{j-1}:\\ \mathrm{rank}([v(x_1);\ldots;v(x_{j-1})])=j-i}} \left(\prod_{\iota=1}^{j-1} q_{x_\iota}\right), \quad (7)$$

where $x_\iota \in \{0,1,\ldots,2^{n-k}\}$ is the numeric representation $b(u)$ of one of the possible column vectors u. Thus, the sum is over all possible size-$(j-1)$ collections of column vectors which produce a matrix of rank $j-i$.

Calculating elements of $\mathcal{A}$ using this formula directly is computationally intractable, but further simplification is possible by analysis of the subspace structure of G.

C. Subspace Structure Definitions

Let W be a vector space comprised of all the vectors in $\mathbb{F}_2^{(n-k)}$, where binary vector addition and scalar multiplication are defined in the usual way.

Next, define the function $\Xi(S,d)$ acting on a space S (equal to W or a subspace of W) and a scalar dimension d. This function returns the set of all dimension-d subspaces of S. In this work, braced superscript notation may be used to indicate the dimension of a vector space, so for example $S^{\{d\}}$ indicates a vector space S which has dimension d. This notation is for clarity and may be omitted if the dimension of the vector space is clear. The number of subspaces of dimension d' contained within a space S of dimension d is given by the Gaussian binomial coefficient

$$|\Xi(S^{\{d\}}, d')| = \binom{d}{d'}_2 = \prod_{\iota=0}^{d'-1} \frac{2^d - 2^\iota}{2^{d'} - 2^\iota} = \prod_{\iota=0}^{d'-1} \frac{2^{d-\iota}-1}{2^{d'-\iota}-1}. \quad (8)$$

A function $\phi(S)$ is also defined which acts on a space S. This function expresses the fraction of the columns of G that lie within S and is defined as

$$\phi(S) = \sum_{x:v(x)\in S} q_x. \quad (9)$$

The final definition required for the analysis of the subspace structure of coset codes is a recursively-defined function, denoted $\psi(S,\mu)$, which expresses the probability that a codeword z with $|r(z)| = \mu$ revealed bits will result in a submatrix G_r which exactly spans a space S. (For this analysis, the terminology "exactly spans" indicates that a set of vectors spans a space and does not span any higher-dimensional space.) A set of columns in G will exactly span S if and only if two criteria are satisfied: (1) all of the columns of G_r are contained within S and (2) the columns of G_r do not all lie within any proper subspace of S. The probability that a set of μ revealed bits selected uniformly at random results in G_r with columns that all lie within S is equal to $\phi(S)^\mu$. Using these observations, $\psi(S,\mu)$ may be defined as

$$\psi(S^{\{d\}},\mu) = \phi(S^{\{d\}})^\mu - \sum_{\iota=1}^{d-1} \left(\sum_{T^{\{\iota\}}\in\Xi(S^{\{d\}},\iota)} \psi(T^{\{\iota\}},\mu) \right). \quad (10)$$

Using these definitions, a number of results may be obtained related to the subspace structure of coset codes. These results are presented below.

IV. Results

Lemma 1. *The probability of $G_{r(z)}$, defined by $z : |r(z)| = \mu$ revealed bits, exactly spanning a subspace S is expressible as*

$$\psi(S^{\{d\}}, \mu) = \sum_{\iota=1}^{d} \left(c_\iota(d) \sum_{U^{\{\iota\}} \in \Xi(S,\iota)} \phi(U^{\{\iota\}})^\mu \right), \quad (11)$$

where the $c_\iota(d)$ are a series of constants which depend on d.

Proof. Begin with (10). Because of the recursive nature of this formula, the value $\phi(U)^\mu$ for every subspace $U^{\{d'\}}$ of dimension $1 \le d' \le d$ will appear in the calculation of $\psi(S^{\{d\}}, \mu)$. Because all the terms in (10) are either of the form $\phi(U)^\mu$ or of the form $\psi(U, \mu)$, after the full expansion of every instance of each $\psi(U, \mu)$ term, only $\phi(U)^\mu$ terms will remain, with each U being a subspace of S. Additionally, by symmetry, the term $\phi(U^{\{d'\}})^\mu$ for every such subspace $U^{\{d'\}}$ of dimension d' appears the same number of times. Thus, the final expansion may be expressed by finding the sum of the $\phi(U^{\{d'\}})^\mu$ terms for every U of dimension d', multiplying by a constant that depends only on d and $d' = \iota$, and summing this result for every $d' = 1, ..., d$, as expressed in (11). □

Lemma 2. *The elements $\mathcal{A}_{i,j}$ of the equivocation probability matrix may be expressed as*

$$\mathcal{A}_{i,j} = \sum_{\iota=1}^{n-k} \left(\gamma_\iota(j-i, n-k) \sum_{S \in \Xi(W,\iota)} \phi(S^{\{\iota\}})^{j-1} \right) \quad (12)$$

for $j > 1$, where $\gamma_\iota(d, \kappa)$ are a series of constants which depend on $d = j - i$ and $\kappa = n - k$.

Proof. Combining the definition of the equivocation probability matrix (6) with (2), the elements of the equivocation probability matrix may be expressed as

$$\mathcal{A}_{i,j} = Pr[\mathrm{rank}(G_{r(z)}) = j - i \mid |r(z)| = j - 1]. \quad (13)$$

The submatrix $G_{r(z)}$ has rank $j - i$ if and only if the $|r(z)| = \mu$ columns of $G_{r(z)}$ exactly span one (and only one) dimension-$j - i$ subspace $S^{\{j-i\}}$ of W. Because the probabilities of spanning any $S^{\{j-i\}}$ are disjoint across all the $S^{\{j-i\}} \in \Xi(W, j - i)$, the probability that $G_{r(z)}$ has rank $j - i$ is equal to the probability of $G_{r(z)}$ exactly spanning a subspace $S^{\{j-i\}}$, summed across all the $S^{\{j-i\}} \in \Xi(W, j-i)$. Because $\psi(S^{\{j-i\}}, i-1)$ represents this probability, combining (13) with (11) yields

$$\mathcal{A}_{i,j} = \sum_{S \in \Xi(W,j-i)} \sum_{\iota=1}^{j-i} \left(c_\iota(j-i) \sum_{T^{\{\iota\}} \in \Xi(S,\iota)} \phi(T^{\{\iota\}})^{i-1} \right). \quad (14)$$

As in Lemma 1, the final expansion of (13) consists entirely of multiples of $\phi(T^{\{\iota\}})^{i-1}$ for various $T^{\{\iota\}}$ of various dimensions ι. Also as in Lemma 1, the $T^{\{\iota\}}$ include all the subspaces of W of dimension ι, and the instances of $\phi(T^{\{\iota\}})^{i-1}$ occur in the same frequency for any given ι. Therefore, (13) may be converted to the form of (12) with the $\gamma_\iota(j-i, n-k)$ given by

$$\gamma_\iota(d, \kappa) = \binom{\kappa}{d}_2 c_\iota(d) \binom{d}{\iota}_2 / \binom{\kappa}{\iota}_2. \quad (15)$$

□

Lemma 3. *The expected message equivocation loss $E[H(M) - H(M \mid |r(z)| = \mu)]$, denoted $L(\mu)$ for a given number $\mu > 0$ of revealed codeword bits may be expressed as*

$$L(\mu) = \sum_{\iota=1}^{n-k} \left(C_\iota(\mu, n-k) \sum_{S \in \Xi(W,\iota)} \phi(S^{\{\iota\}})^\mu \right), \quad (16)$$

where the $C_\iota(\mu, \kappa)$ are a series of constants which may depend on μ and $\kappa = n - k$.

Proof. The expected equivocation loss for a given number of revealed bits may be calculated from the equivocation probability matrix as

$$L(\mu) = E[H(M) - H(M| \ |r(z)| = \mu)] = \sum_{i=1}^{\mu+1} (i-1) \cdot \mathcal{A}_{i,\mu+1}$$
$$= \sum_{i=1}^{\mu+1} (i-1) \cdot \sum_{\iota=1}^{n-k} \left(\gamma_\iota(\mu - i + 1, n-k) \sum_{S \in \Xi(W,\iota)} \phi(S^{\{\iota\}})^\mu \right). \quad (17)$$

Changing the order of summation yields the form specified in (16), with the $C_\iota(\mu)$ given by

$$C_\iota(\mu, \kappa) = \sum_{i=1}^{\mu} ((i-1) \cdot \gamma_\iota(\mu - i + 1, \kappa)), \quad (18)$$

or by eliminating the zero-valued case of $i = 1$ and shifting the index of summation by 1,

$$C_\iota(\mu, \kappa) = \sum_{i=1}^{\mu-1} (i \cdot \gamma_\iota(\mu - i, \kappa)). \quad (19)$$

□

Theorem 4. *The expected message equivocation loss $E[H(M) - H(M \mid |r(z)| = \mu)]$ for a given number $\mu > 0$ of revealed codeword bits is equal to*

$$L(\mu) = \mu - (n-k) + \sum_{\delta=1}^{n-k-1} \left(K_\delta \sum_{S \in \Xi(W, n-k-\delta)} \phi(S)^\mu \right). \quad (20)$$

Where the K_δ are a series of constants which do not depend on μ or $n - k$.

Proof. The constants $C_\iota(\mu, \kappa)$ of (16) and (18) depend on the $\gamma_\iota(d, \kappa)$ of (12) and (15), which in turn depend on the $c_\iota(d)$ of (11). To find the value of the coefficients $c_\iota(d)$, a first step is to count the number of times that a dimension-d' subspace is encountered during the summations specified in the expansion of (10). For each instance of a call to $\phi(U^{\{d'\}})$ for a dimension-d' subspace $U^{\{d'\}}$, a "path" θ may be specified as a set of integers $\theta_i \in \theta, d' \le \theta_i \le d$, $\theta_{i+1} < \theta_i$, which indicates the dimension θ_i of each call to $\psi(T^{\{\theta_i\}}, \mu)$, starting with $\theta_1 = d$ and ending with $\theta_{|\theta|} = d'$. Next, let $\Theta_{d,d'}$ represent the set of all valid paths which start with d and end with d'. (Note that if $d = d'$, there is exactly one path

θ, of size one, in $\Theta_{d,d'}$. Also note that if $d = d' + 1$, there is exactly one path θ, of size two, in $\Theta_{d,d'}$.) For each path $\theta \in \Theta_{d,d'}$, the number of times $\phi(U^{\{d'\}})$ is encountered for any subspace $U^{\{d'\}} \in \Xi(S^{\{d\}}, d')$ is equal to the product of $\binom{\theta_{i-1}}{\theta_i}_2$ for each $1 < i \leq |L|$. Because the sign of summation alternates with each successive recursive call to $\psi(T, \mu)$, the sign of any given $\phi(U^{\{d'\}})$ encountered via path θ during the expansion of (10) is positive if $|\theta|$ is even and negative if $|\theta|$ is odd. It is useful to represent the product over such a path set by a function $\eta(d, d')$ defined as

$$\eta(d, d') = \sum_{\theta \in \Theta_{d,d'}} (-1)^{|\theta|+1} \cdot \left(\prod_{j=2}^{|\theta|} \binom{\theta_{j-1}}{\theta_j}_2 \right). \tag{21}$$

Finally, because each $c_\iota(d)$ in (11) is multiplied by the sum of all dimension-ι subspaces $U^{\{\iota\}}$ of S, the number of instances of $\phi(U^{\{\iota\}})$ for all dimension-ι subspaces must be divided by the number $\binom{d}{\iota}_2$ of dimension-ι subspaces of $S^{\{d\}}$.

Using these steps, the $c_\iota(d)$ may be expressed as

$$c_\iota(d) = \eta(d, \iota) / \binom{d}{\iota}_2. \tag{22}$$

Combining this expression with (15) yields

$$\gamma_\iota(d, \kappa) = \binom{\kappa}{d}_2 \eta(d, \iota) / \binom{\kappa}{\iota}_2, \tag{23}$$

and combining this equation with (19) yields

$$C_\iota(\mu, \kappa) = \sum_{i=1}^{\mu-1} \left(i \cdot \binom{\kappa}{\mu - i}_2 \eta(\mu - i, \iota) / \binom{\kappa}{\iota}_2 \right). \tag{24}$$

Several observations can help to simplify the computation of (24). First, from the definition of the Gaussian binomial, $\binom{n-k}{\mu-i}_2 = 0$ if $i < \mu - (n-k)$, so the index of summation only need start at $\mu - (n-k)$. Next, by the definition of $\eta(\cdot)$, if $\mu - i < \iota$, then $\eta(\mu - i < \iota) = 0$, so the index of summation may end at $\mu - \iota$. (24) then becomes

$$C_\iota(\mu, \kappa) = \sum_{i=\mu-\kappa}^{\mu-\iota} \left(i \cdot \binom{\kappa}{\mu - i}_2 \eta(\mu - i, \iota) / \binom{\kappa}{\iota}_2 \right), \tag{25}$$

and making the substitution $i = j + \mu - \kappa$ gives

$$C_\iota(\mu, \kappa) = \frac{1}{\binom{\kappa}{\iota}_2} \sum_{j=0}^{\kappa-\iota} \left((\mu - \kappa + j) \cdot \binom{\kappa}{\kappa - j}_2 \eta(\kappa - j, \iota) \right). \tag{26}$$

Another simplification arises due to the fact that the $\eta(\cdot)$ function has the property that

$$\eta(d, d') = \begin{cases} 1 & \text{if } d = d' \\ -\sum_{i=d'}^{d-1} \binom{d}{i}_2 \eta(i, d'), & \text{if } d > d'. \end{cases} \tag{27}$$

The $C_\iota(\mu, \kappa)$ of (26) may then be divided into two cases. First, in the case that $\iota = \kappa = n - k$, it is easy to verify that

$$C_\kappa(\mu, \kappa) = \mu - \kappa. \tag{28}$$

Second, in the case that $\iota < \kappa = n - k$, the first ($j = 0$) term in the summation of (26) can then be split out and transformed using (27) to yield

$$\begin{aligned} C_{\iota:\iota<\kappa}(\mu, \kappa) = &\sum_{j=1}^{\kappa-\iota} \frac{(\mu - \kappa + j) \cdot \binom{\kappa}{\kappa-j}_2 \eta(\kappa - j, \iota)}{\binom{\kappa}{\iota}_2} \\ &- \sum_{i=\iota}^{\kappa-1} \frac{(\mu - \kappa) \binom{\kappa}{i}_2 \eta(i, \iota)}{\binom{\kappa}{\iota}_2}. \end{aligned} \tag{29}$$

Performing the replacement $j = \kappa - i$ for the index of summation in the second sum (which also swaps the order of the limits) and combining the sums yields

$$C_{\iota:\iota<\kappa}(\mu) = \sum_{j=1}^{\kappa-\iota} \frac{j \cdot \binom{\kappa}{\kappa-j}_2 \eta(\kappa - j, \iota)}{\binom{\kappa}{\iota}_2}. \tag{30}$$

At this point, for any $\iota < (n - k)$, $C_\iota(\mu, \kappa)$ does not depend on μ. Next, define the series of constants K_δ as

$$K_\delta = C_{\kappa-\delta}(\mu, \kappa). \tag{31}$$

Expanding this definition with (30) yields

$$K_\delta = \sum_{j=1}^{\delta} \frac{j \cdot \binom{\kappa}{\kappa-j}_2 \eta(\kappa - j, \kappa - \delta)}{\binom{\kappa}{\kappa-\delta}_2}. \tag{32}$$

This expression may be simplified using the following property of product chains of Gaussian binomials:

$$\frac{\binom{x_1}{x_2}_2 \binom{x_2}{x_3}_2 \cdots \binom{x_{n-1}}{x_n}_2}{\binom{x_1}{x_n}_2} = \binom{x_1 - x_n}{x_2 - x_n}_2 \binom{x_2 - x_n}{x_3 - x_n}_2 \cdots \binom{x_{n-1} - x_n}{0}_2. \tag{33}$$

Because the expansion of $\eta(d, d')$ results in a sum of products of Gaussian binomial terms, each of which consists of a chain starting at d and ending at d', (32) may be simplified using (33) to yield

$$K_\delta = \sum_{j=1}^{\delta} j \cdot \binom{\delta}{\delta - j}_2 \eta(\delta - j, 0). \tag{34}$$

These constants K_δ do not depend on either μ or $\kappa = n - k$. Finally, the summation in (16) may be split into the last term ($\iota = n - k$) and the remainder of the terms ($\iota < n - k$). The last term is given by (28), while the remainder of the terms may be transformed using the definition (31) of K_δ to yield the desired equation (20). □

The first few K_δ are: $1, -1, 3, -21, 315, -9965$. The following pattern is immediately apparent:

$$K_\delta = \prod_{i=1}^{\delta-1} (1 - 2^i). \tag{35}$$

This pattern has been verified for $\delta \leq 16$, and it is conjectured that the pattern holds for all δ.

Theorem 5. *Assuming* (35) *holds for all positive integers δ, the extreme-rate code defined by a uniform proportion of all nonzero generator matrix columns,*

$$q = t : t_i = \frac{1}{2^{n-k} - 1}, \tag{36}$$

is locally optimal in terms of equivocation loss.

Proof. Let the vector $s = \begin{bmatrix} s_1 & s_2 & \cdots & s_{2^{n-k}-1} \end{bmatrix}^T$ be a zero-sum ($\sum_i s_i = 0$) unit vector defining an offset direction from the uniform-proportion vector t of (36). Then let $\alpha = t + sx$ be a vector offset from t by magnitude x in direction s. The derivative of $L(\mu)$ with respect to x is then

$$\frac{\delta}{\delta x} L(\mu)\Big|_{x=0} = \sum_{\delta=1}^{n-k-1} \left(K_\delta \sum_{S \in \Xi(W, n-k-\delta)} \frac{\delta}{\delta x} \phi(S)^\mu \Big|_{x=0} \right). \quad (37)$$

Recalling the definition (9) of $\phi(\cdot)$,

$$\begin{aligned} \frac{\delta}{\delta x} \phi(S^{\{d\}})^\mu \Big|_{x=0} &= \mu \phi(S^{\{d\}})^{\mu-1} \frac{\delta}{\delta x} \phi(S^{\{d\}}) \Big|_{x=0} \\ &= \mu \phi(S^{\{d\}})^{\mu-1} \sum_{x : v(x) \in S^{\{d\}}} s_x = \mu \left(\frac{2^d - 1}{2^{n-k} - 1}\right)^{\mu-1} \sum_{x : v(x) \in S^{\{d\}}} s_x. \end{aligned} \quad (38)$$

Substituting (38) into (37) gives $\frac{\delta}{\delta x} L(\mu)\big|_{x=0}$ as a sum of constant multiples of s_i. Because of the symmetry of the subspaces of W, each $s_i, 1 \le i < 2^{n-k}$ appears the same number of times with the same constant multipliers. However, because $\sum_i s_i = 0$, the derivative of the equivocation loss in any direction s is zero at $q = t$.

The calculation of the second derivative of the equivocation loss begins in a similar manner to that of the first derivative:

$$\frac{\delta^2}{\delta x^2} L(\mu)\Big|_{x=0} = \sum_{\delta=1}^{n-k-1} \left(K_\delta \sum_{S \in \Xi(W, n-k-\delta)} \frac{\delta^2}{\delta x^2} \phi(S)^\mu \Big|_{x=0} \right). \quad (39)$$

$$\frac{\delta^2}{\delta x^2} \phi(S^{\{d\}})^\mu \Big|_{x=0} = \mu(\mu-1) \left(\frac{2^d - 1}{2^{n-k} - 1}\right)^{\mu-2} \left(\sum_{x : v(x) \in S^{\{d\}}} s_x \right)^2. \quad (40)$$

The expansion of the squared sum term in (40) results in $(2^d - 1)$ terms of the form s_i^2 and $(2^d - 1)(2^d - 2)$ cross terms of the form $s_i s_j$. When these terms are evaluated across all subspaces of dimension d, a total of $\binom{n-k}{d}_2 (2^d - 1)$ squared terms and $\binom{n-k}{d}_2 (2^d - 1)(2^d - 2)$ cross terms are accumulated. Again by symmetry, all squared terms occur in equal frequency, and because all pairs of vectors are symmetric, all cross terms also occur in equal frequency. Next, because $(\sum_i s_i)^2$ accumulates $(2^{n-k} - 1)$ squared terms and $(2^{n-k} - 1)(2^{n-k} - 2)$ cross terms and sums to zero, the second sum in (39) becomes

$$\begin{aligned} &\sum_{S \in \Xi(W, n-k-\delta)} \frac{\delta^2}{\delta x^2} \phi(S)^\mu \Big|_{x=0} = \mu(\mu-1) \left(\frac{2^{n-k-\delta} - 1}{2^{n-k} - 1}\right)^{\mu-2} \cdot \\ &\binom{n-k}{\delta}_2 \left(\frac{(2^{n-k-\delta} - 1)}{2^{n-k} - 1} - \frac{(2^{n-k-\delta} - 1)(2^{n-k-\delta} - 2)}{(2^{n-k} - 1)(2^{n-k} - 2)} \right) \\ &= \mu(\mu-1) \left(\frac{2^{n-k-\delta} - 1}{2^{n-k} - 1} \right)^{\mu-2} \binom{n-k-2}{\delta-1}_2 2^{n-k-\delta-1}. \end{aligned} \quad (41)$$

Clearly if $\mu = 0$ or $\mu = 1$, $\frac{\delta^2}{\delta x^2} L(\mu) = 0$. In these cases, because $L(\mu) = 0$ for all x, t represents a local optimum. All that remains to show that $\frac{\delta^2}{\delta x^2} L(\mu) > 0$ for $\mu \ge 2$. Substituting (41) into (39) and removing an always-positive factor of $\frac{\mu(\mu-1)}{(2^{n-k}-1)^{\mu-2}}$ gives

$$\sum_{\delta=1}^{n-k-1} \left(K_\delta (2^{n-k-\delta} - 1)^{\mu-2} \binom{n-k-2}{\delta-1}_2 2^{n-k-\delta-1} \right) > 0, \quad (42)$$

and using the expression (35) for K_δ and the definition (8) of the Gaussian binomial yields

$$\sum_{\delta=1}^{n-k-1} \left((2^{n-k-\delta} - 1)^{\mu-2} 2^{n-k-\delta-1} \prod_{i=0}^{\delta-2} (1 - 2^{n-k-i-2}) \right) > 0. \quad (43)$$

When considering the terms in (43), all of the terms include a (trivial) factor of 1, all but one include a factor of $(1 - 2^{n-k-2})$, all but two include a factor of $(1 - 2^{n-k-3})$ and so on until only the last term includes a factor of $(1 - 2^1)$. Because of this pattern, the additive terms of the sum may be accumulated sequentially and multiplied progressively by the terms in the product. The result is the final term of a sequence $a_i(\mu)$, $1 \le i < (n-k)$ with $a_0(\mu) = 0$ and with the recurrence relation

$$a_i(\mu) = 2^{i-1}(2^i - 1)^{\mu-2} + (1 - 2^{i-1}) \cdot a_{i-1}(\mu). \quad (44)$$

Consider the following expression for the $a_i(\mu)$ to satisfy (44):

$$a_i(\mu) = \sum_{j=0}^{\mu-2} \left(u_{\mu,j} \prod_{l=2}^{j+1} (2^{i+1} - 2^l) \right). \quad (45)$$

If the $u_{\mu,j}$ are all positive, then each $a_i(\mu)$ is positive because the product $\prod_{l=2}^{j+1} (2^{i+1} - 2^l)$ is composed of either all positive terms if $i > j$ or contains a zero term if $i \le j$. The function $D(j, \nu)$ is used to represent the coefficient of $2^{\nu i}$ in the expansion of $\prod_{l=2}^{j+1} (2^{i+1} - 2^l)$ and is given by

$$D(j, \nu) = (-1)^{j-\nu} 2^\nu \binom{j}{\nu}_2 \cdot 2^{\left(\frac{1}{2}(j-\nu)^2 + \frac{3}{2}(j-\nu)\right)}. \quad (46)$$

Using (46) with (44) gives an equation relating the $u_{\mu,j}$ with factors of $2^{\nu i}$. Because this equation must hold for all i, the coefficients of $2^{\nu i}$ must equate for any ν. Using this procedure, the following equation may be obtained expressing $u_{\mu,\nu}$ in terms of the $u_{\mu,\nu+1}$, $u_{\mu,\nu+2}$, …, $u_{\mu,\mu-2}$:

$$u_{\mu,\nu} = (-1)^{\mu-\nu} \binom{\mu-2}{\nu} - \sum_{j=\nu+1}^{\mu-2} \left(u_{\mu,j} (-1)^{j-\nu} \binom{j}{\nu}_2 2^{\frac{1}{2}(j-\nu)(j-\nu+1)} \right). \quad (47)$$

The recursive equation (47) is solved by

$$u_{\mu,\nu} = \sum_{i=\nu}^{\mu-2} \left((-1)^{\mu-i} \binom{\mu-2}{i} 2^i \binom{i}{\nu}_2 \right). \quad (48)$$

The $u_{\mu,\nu}$ defined in (48) also have the property that

$$u_{\mu,\nu} = u_{\mu-1,\nu-1} + (2^{\nu+1} - 1) u_{\mu-1,\nu}, \quad (49)$$

which implies that if $u_{\mu,\nu} > 0$ for all $0 \le \nu \le \mu - 2$, then $u_{\mu+1,\nu} > 0$ for all $0 \le \nu \le \mu - 1$. It is easy to verify that $u_{2,0} = 1 > 0$ produces the sequence $a_i(2) = 1$ which solves (44), so by induction, $u_{\mu,\nu} > 0$ for every $\mu \ge 2, 0 \le \nu \le \mu - 2$, which implies that every $a_i(\mu) > 0$, and the second derivative of $L(\mu)$ is always positive for $\mu \ge 2$.

□

REFERENCES

[1] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[2] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011. [Online]. Available: https://books.google.com/books?id=ov5jYjrrNCIC

[3] M. Bellare, S. Tessaro, and A. Vardy, "Semantic security for the wiretap channel," in *Advances in Cryptology – CRYPTO 2012*, R. Safavi-Naini and R. Canetti, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 294–311.

[4] W. K. Harrison, J. Almeida, M. R. Bloch, S. W. McLaughlin, and J. Barros, "Coding for secrecy: An overview of error-control coding techniques for physical-layer security," *IEEE Signal Processing Magazine*, vol. 30, no. 5, pp. 41–50, Sept. 2013.

[5] M. R. Bloch, M. Hayashi, and A. Thangaraj, "Error-control coding for physical-layer secrecy," *Proceedings of IEEE*, vol. 103, no. 10, pp. 1725–1746, Oct. 2015.

[6] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, "Applications of LDPC codes to the wiretap channels," *IEEE Transactions on Information Theory*, vol. 53, no. 8, pp. 2933–2945, Aug. 2007.

[7] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6428–6443, 2011.

[8] Z. Goldfeld, P. Cuff, and H. H. Permuter, "Semantic-security capacity for wiretap channels of type II," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 3863–3879, 2016.

[9] L. H. Ozarow and A. D. Wyner, "Wire-tap channel II," *AT T Bell Laboratories Technical Journal*, vol. 63, no. 10, pp. 2135–2157, 1984.

[10] V. Wei, "Generalized Hamming weights for linear codes," *IEEE Transactions on Information Theory*, vol. 37, no. 5, pp. 1412–1418, Sept. 1991.

[11] S. Al-Hassan, M. Z. Ahmed, and M. Tomlinson, "Secrecy coding for the wiretap channel using best known linear codes," in *Global Information Infrastructure Symposium - GIIS 2013*, 2013, pp. 1–6.

[12] J. Pfister, M. A. C. Gomes, J. P. Vilela, and W. K. Harrison, "Quantifying equivocation for finite blocklength wiretap codes," in *2017 IEEE International Conference on Communications (ICC)*, 2017, pp. 1–6.

[13] W. K. Harrison and M. R. Bloch, "On dual relationships of secrecy codes," in *2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2018, pp. 366–372.

[14] ——, "Attributes of generators for best finite blocklength coset wiretap codes over erasure channels," in *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 827–831.

[15] W. Yang, R. F. Schaefer, and H. V. Poor, "Wiretap channels: Nonasymptotic fundamental limits," *IEEE Transactions on Information Theory*, vol. 65, no. 7, pp. 4069–4093, 2019.