EG-LDPC Codes for the Erasure Wiretap Channel

V. Rajaraman and Andrew Thangaraj
Department of Electrical Engineering
Indian Institute of Technology, Madras
Chennai 600036, India
Email: rajaraman.v@gmail.com, andrew@iitm.ac.in

Abstract—The wiretap channel model, proposed by Wyner, has been studied by various authors from the perspectives of security, reliability and cryptographic protocols. A basic theme of these discussions has been information theoretically secure communication whose degree of secrecy can be theoretically proved. This paper explains a practical implementation of Euclidean Geometry (EG) - Low Density Parity Check (LDPC) codes for wiretap channels of type I and II. The focus has been on efficient encoding and decoding by taking advantage of the cyclic nature of EG codes. The generalized Hamming weights of EG codes can be bounded using those of primitive BCH codes. This provides security guarantees for EG codes in wiretap channels. The asymptotic threshold analysis from the LDPC viewpoint provides added guarantees at long blocklengths. Overall, EG-LDPC codes are promising candidates for low-complexity coding over wiretap channels.

I. INTRODUCTION

The wiretap channel, introduced by Wyner [1] and later generalized by Csiszar and Korner [2], provides a good model for developing schemes for information-theoretic security. In a wiretap channel, the legitimate users (Alice and Bob) are separated by a channel called the main channel, while an adversary (Eve) listens to all of Alice’s transmissions through another channel called the wiretapper’s channel. Under certain conditions on the main and wiretapper’s channels, Alice can achieve both error-free communication to Bob (reliability objective) and information-theoretic security against Eve (security objective). The maximum rate (bits per channel use) at which both objectives are attainable is called the secrecy capacity of the wiretap channel.

The simplest example of a wiretap channel is the (binary) erasure wiretap channel, in which the main channel is assumed to be noiseless and the wiretapper’s channel is assumed to be a binary erasure channel with erasure probability $\epsilon$. Hence, every bit sent by Alice is received correctly by Bob, and Eve receives bits sent by Alice with probability $1-\epsilon$ (iid for each bit). Such a probabilistic erasure wiretap channel is referred to as a type-I erasure wiretap channel.

In a type-II wiretap channel [3], we assume, once again, that the main channel is noiseless. The wiretapper will receive an $\epsilon$-fraction of the transmitted bits, and the wiretapper can choose the exact bits that will be revealed. We see that the wiretap-II model is significantly different from the wiretap-I model, and different coding ideas might be needed for each of them.

Several recent papers have studied various types of wiretap channels and provided results on secrecy capacity (see [4] and the references therein). However, there has not been a lot of emphasis on coding for wiretap channels. In this work, we propose and study the use of Euclidean Geometry (EG) - Low Density Parity Check (LDPC) codes for the erasure wiretap channel. We show that the properties of EG-LDPC codes are well-suited for low-complexity implementation of encoding and decoding over the erasure wiretap channel. Then, we show that the various properties of EG-LDPC codes are useful for providing security guarantees against adversaries in both types (I and II) of erasure wiretap channels.

The rest of this article is organized as follows. In Section II, we provide a formal definition for erasure wiretap channels, and discuss coding schemes and performance criteria for both type-I and type-II models. In Section III, we define the EG codes that are proposed for erasure wiretap channels and discuss their useful properties. We provide simulation results in Section IV and make concluding remarks in Section V.

II. ERASURE WIRETAP CHANNEL MODEL

The Wiretap channel model, studied by Wyner [1] in the seventies, is still a very relevant model in today’s highly connected world. The particular scenario that will be considered in the current study is the erasure wiretap model shown in Fig 1.

![Fig. 1. The Wiretap Channel Model](image)

Under this model Alice has to transmit a $k$-bit secret message $m$ to a legitimate recipient Bob, and Eve is an unauthorized eavesdropper trying to tap into the communication. We assume that Alice encodes the $k$-bit message $m$ into a $n$-bit codeword $c$. The encoder is assumed to be one-to-many but invertible i.e. no two messages can be encoded into the
same codeword, but each message can be encoded to one of many possible codewords. Such an encoder is important for achieving the security objective as detailed below.

In the erasure wiretap model considered in this work, the main channel is assumed to be noiseless. So, Bob receives the codeword $c$ and can decode $m$ uniquely by the invertibility property assumed for the encoder. Hence, the rate of communication from Alice to Bob is $R_s = k/n$ bits per channel use.

The $n$-length vector seen by Eve is denoted $z$. In the erasure wiretap model of type I, Eve receives $c$ across a binary erasure channel (BEC) with erasure probability $\epsilon$ i.e. if $c = [c_1 \ c_2 \cdots c_n]$ and $z = [z_1 \ z_2 \cdots z_n]$, we have $z_i = c_i$ with probability $1-\epsilon$ or $z_i = ?$ with probability $\epsilon$ for each $i$ independently.

The type-II wiretap model is more stringent on the designer: Eve can now access $\mu$ bits of her choice from the $n$-bit code word i.e. the locations that are received correctly are under Eve’s control, but she can receive only $\mu$ bits; the remaining bits have to be erased.

The wiretapped’s channel for both the type-I and type-II models are shown in Fig. 1. For type II, the wiretapper’s channel is called as a bit selector parametrized by the number of revealed bits $\mu$. In both these models, it is assumed that Eve has full a priori knowledge of the decoding scheme and its parameters, but only the random process by which the $k$ random bits are chosen remains a run time secret.

In this work, we use the weak secrecy criterion. We require that the rate of information leaked to the eavesdropper should tend to zero asymptotically. More formally, we need that $\frac{1}{k}I(m; m^b) \rightarrow 0$ as $k \rightarrow \infty$, where $m^b$ denotes Eve’s decoded message. Note that we use the same notation for a random variable and the value taken by it inside the mutual information function. The secrecy capacity is the largest $R_s$ for which the security criterion is achievable. For the erasure wiretap channel of type-I, it has been shown in [1] that the secrecy capacity is $1 - (1 - \epsilon) = \epsilon$. In [3], the secrecy capacity for the erasure wiretap channel of type-II has been shown to be $1 - \mu/n$.

### A. Coset Encoding Scheme

The coset encoding process for a wiretap channel, introduced and studied in [1] and [3], works as follows. We start with a $(n, n-k)$ code $C$ as the base code. Let $G$ be the generator matrix of $C$, and let $G^*$ be the generator matrix for the code $C'$ such that $C \oplus C' = \{0,1\}^n$ i.e. the rows of $G$ and $G^*$ form a basis for $\{0,1\}^n$. A secret $k$-bit message $m$ is mapped to a codeword $c$ using the transformation

$$c = [m \ u] \begin{bmatrix} G^* \ G \end{bmatrix},$$

where $u$ is a uniformly random $(n-k)$-bit random vector. To view this otherwise, the code word to be transmitted is an $n$-bit vector chosen randomly from the coset indexed by the secret message $m$. Therefore, there exists a parity check matrix $H$ for $C$ such that $Hc^T = m$.

The decoding process is straight forward, since all we need is the syndrome of $c$ with respect to the parity check matrix $H$. At the legitimate decoder, we simply set

$$m(\text{received}) = He^T.$$

Suppose Eve receives $s$ bits at positions $I \subseteq \{1,2,\cdots,n\}$ through the wiretapper’s channel. The secrecy of the transmission will be ensured if the cosets of $C$ have all the $2^s$ possible $s$-tuples in the revealed positions $I$ [3], or equivalently, if the $s$ columns of $G$ corresponding to the set $I$ have full column rank. In general, if $G_I$ denotes the submatrix of $G$ formed by the $s$ columns indexed by $I$, the number of bits of the message leaked to the eavesdropper, denoted $m(I)$, is given by [3]

$$m(I) = s - \text{rank}(G_I) = |I| - \text{rank}(G_I).$$

Let us further define a quantity $s_e(C)$ as the minimum number of codeword bits needed by an eavesdropper to find at least $e$ bits of the message. More precisely,

$$s_e(C) = \min_{I \subseteq \{1,2,\cdots,n\} : |I| \geq e} \min_{I : |I| - \text{rank}(G_I) \geq e} |I|.$$

### B. Generalized Hamming Weights

The support $\chi(C')$ of a sub code $C'$ of a linear code $C$ is the set of bit positions in which the codewords of $C'$ are not always zero. Formally,

$$\chi(C') = \{i : (x_1, x_2, \cdots x_n) \in C, \ x_i \neq 0\}.$$  

The $r$-th generalized Hamming weight of a code $C$, denoted $d_r(C)$, is the size of the smallest support of an $r$-dimensional subcode of $C$. In other words,

$$d_r(C) = \min_{C' \subseteq C, \ \text{subcode}} \{|\chi(C')| : \text{dimension}(C') = r\}.$$  

As shown in [5, Theorem 2], the $r$-th generalized Hamming weight can be written as

$$d_r(C) = \min_{I \subseteq \{1,2,\cdots,n\}, |I| - \text{rank}(H_I) \geq r} |I|,$$

where $H$ is a parity-check matrix for the code $C$. Comparing with (2), when Alice and Bob are using an $(n, n-k)$ code $C$ in the erasure wiretap channel, we have

$$s_e(C) = d_e(C^\perp).$$

Therefore, the generalized Hamming weights of the dual code $C^\perp$ control the number of leaked bits to the eavesdropper. For instance, if the number of bits leaked to Eve is less than the minimum distance of $C^\perp$ (also equal to $d_1(C^\perp)$), then the message is fully secure.

### C. LDPC coding

As shown in [6], a low density parity check (LDPC) code can be used over erasure wiretap channels of type I. The idea is to use the LDPC matrix as the generator matrix $G$ of the $(n, n-k)$ code $C$ used for coset encoding. Note that $C$ is the dual of the LDPC code. Suppose the threshold of the LDPC code over the BEC is $\alpha$ [7]. Let the wiretapper’s channel be BEC($\epsilon$). Then, for $1 - \epsilon < \alpha$ or $\epsilon > 1 - \alpha$, the submatrix
of $G$ formed by the revealed bits is of full rank with high probability for large $n$.

In summary, we see that the coset encoding scheme can be used in erasure wiretap channels of types I and II. In type I channels, we might prefer to use LDPC codes, and the BEC threshold of LDPC codes controls the security. In type II channels, we might prefer to use structured codes, since security is controlled by generalized Hamming weights. Therefore, EG-LDPC codes, which can be seen as both structured cyclic codes and LDPC codes, are very useful in erasure wiretap channels.

### III. Euclidean Geometry Codes

Finite Geometry codes are based on the lines and points of Euclidean or Projective Geometries defined over finite fields. Such codes not only have good minimum distance properties, but are also cyclic or quasi-cyclic in nature. This makes them amenable to simple linear time encoding using shift registers.

In this paper, we focus mainly on 2-dimensional EG codes with blocklength $n = 2^{2s} - 1$ and dimension $k = 2^{2s} - 3^s$ denoted as $EG(s)$. The minimum distance of the code is $d = 2^s + 1$. Parameters of some EG codes are given in Table I.

<table>
<thead>
<tr>
<th>$s$</th>
<th>$n$</th>
<th>$k$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>15</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>63</td>
<td>37</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>255</td>
<td>175</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>1023</td>
<td>781</td>
<td>33</td>
</tr>
<tr>
<td>6</td>
<td>4095</td>
<td>3367</td>
<td>65</td>
</tr>
</tbody>
</table>

**TABLE I**  
Some $EG(s)$ codes.

The EG code $EG(s)$ has an $n \times n$ sparse parity check matrix with constant row and column weight equal to $w = 2^s$. With this parity-check matrix, the code $EG(s)$ can be treated as an LDPC code [8].

#### A. EG codes in erasure wiretap channels

We propose to use the coset encoding scheme of Section II-A with the code $C$ set to be $EG(s)^\perp$, which is the dual of the EG LDPC code $EG(s)$.

The generator matrix for the $(n, n-k)$ code $C$ will be a parity check matrix for $EG(s)$. However, it is much simpler to implement the encoder in the coset encoding scheme, since the code $C$ is a cyclic code. Let $g(x)$ be the generator polynomial of $C$ with degree equal to $k$. The secret message is taken to be a polynomial $m(x)$ with degree not greater than $k-1$. We generate a $(n-k)$-bit random vector and use it as coefficients of the polynomial $u(x)$, whose degree will be at most $n-k-1$.

Then, we follow the systematic encoding method for cyclic codes with a minor modification to perform coset encoding as follows. We divide $x^{n-k}u(x)$ by $g(x)$ to obtain a quotient $q(x)$ and a remainder $r(x)$ (degree $\leq k - 1$). So, $x^{n-k}u(x) = q(x)g(x) + r(x)$. We let the transmitted codeword polynomial to be

$$c(x) = x^{n-k}u(x) + r(x) + m(x).$$

We see that $c(x)$ belongs to the coset with syndrome $m(x)$, since $c(x) = q(x)g(x) + m(x)$ will give $m(x)$ as remainder, when divided by $g(x)$. The legitimate receiver will simply deconvolve the received vector with $g(x)$ to retrieve the secret message $m(x)$.

Compared to LDPC codes in erasure wiretap channels, the encoding and decoding of EG-LDPC codes is significantly less complex. Encoding and decoding can be implemented using shift register circuitry at the encoder and the decoder.

#### B. Type-I erasure wiretap channel

Over the type-I channel, we use the low density parity check matrix with constant row and column weight $2^s$ for analysis purposes. For $s = 6$ and above, the block length of the EG-LDPC code is large. Hence, we can hope that the assumptions in the asymptotic analysis of the LDPC codes hold. The BEC thresholds of the EG-LDPC codes are given in Table II along with the threshold for security, which is the threshold subtracted from 1.

<table>
<thead>
<tr>
<th>$s$</th>
<th>$n$</th>
<th>BEC threshold</th>
<th>Security threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>255</td>
<td>0.3058</td>
<td>0.6942</td>
</tr>
<tr>
<td>5</td>
<td>1023</td>
<td>0.1832</td>
<td>0.8168</td>
</tr>
<tr>
<td>6</td>
<td>4095</td>
<td>0.1062</td>
<td>0.8928</td>
</tr>
<tr>
<td>7</td>
<td>16383</td>
<td>0.06</td>
<td>0.94</td>
</tr>
</tbody>
</table>

**TABLE II**  
BEC thresholds of EG-LDPC codes.

Since the EG codes are regular, their thresholds will be weaker when compared to irregular LDPC codes. Also, the asymptotic analysis will only be roughly valid for $s = 4$ and $s = 5$. However, the benefits of easy encoding and structure might be useful.

#### C. Type-II erasure wiretap channel

Over the type-II channel, the cyclic structure of EG-LDPC codes can be exploited to determine security from the generalized Hamming weights of $C^{\perp} = EG(s)$ using (6). However, determining the generalized Hamming weights of a code is likely to be a hard problem [9]. Some results are known for the first $l - m + 2$ generalized Hamming weights of primitive BCH codes of length $2^l - 1$ with designed distance of the form $2^m - 1$ [10].

The EG code $EG(s)$ is a cyclic code of length $n = 2^{2s} - 1$ with zeros in GF($2^{2s}$). The zero set of the code contains the set $\{1, 2, \cdots, 2^s\}$ as a subset, and the BCH bound results in a minimum distance at least $2^s + 1$ [11]. So, $EG(s)$ is a subcode of the primitive BCH code of length $n$ and designed distance $2^s - 1$. Denoting the primitive BCH code of length $n$ and designed distance $d$ as BCH($n, d$), we see

$$d_r(EG(s)) \geq d_r(BCH(n, d)),$$

since $EG(s) \subseteq BCH(n, d)$. Combining with the results from [10], we obtain lower bounds on the first $2s - s + 2 = s + 2$ generalized Hamming weights of $EG(s)$. The BCH bound on the minimum distance of $EG(s)$ gives $d_1(EG(s)) \geq 2^s + 1$. 


These bounds are presented in Table III for some values of \( s \).

<table>
<thead>
<tr>
<th>( s )</th>
<th>( r )</th>
<th>BCH bound on ( d_r(\text{EG}(s)) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>(1,2,3,4,5)</td>
<td>(9,11,13,14,15)</td>
</tr>
<tr>
<td>4</td>
<td>(1,2,3,4,5,6)</td>
<td>(17,23,27,29,30,31)</td>
</tr>
<tr>
<td>5</td>
<td>(1,2,3,4,5,6,7)</td>
<td>(33,47,55,59,61,62,63)</td>
</tr>
<tr>
<td>6</td>
<td>(1,2,3,4,5,6,7,8)</td>
<td>(65,95,111,119,123,125,126,127)</td>
</tr>
</tbody>
</table>

**TABLE III**

**BCH BOUNDS ON GENERALIZED HAMMING WEIGHTS OF EG-LDPC CODES.**

From the table and the BCH bound on the minimum distance, the security properties of EG codes on type-II erasure wiretap channels can be estimated. For instance, for \( s = 6 \), no message bit is leaked, if less than 65 bits of the codeword are revealed to the eavesdropper. At most one message bit may be leaked, if less than 95 bits are revealed and so on.

From the table, we notice that for larger \( r \) the generalized Hamming weight bounds are increasing by 1, when \( r \) is increased by 1. In other words, we can roughly expect one more message bit to be leaked for every additional revealed bit of the transmitted codeword, after a certain number of bits.

**IV. SIMULATION STUDIES**

Simulation results are provided in this section for some EG codes. Both type-I and type-II wiretap models have been simulated. For each set of revealed bit positions \( I \), the equivocation is defined to be \( \text{rank}(G_I) \), where \( G \) is the \( n-k \times n \) generator matrix of the code \( C \) used in the coset encoding scheme. Notice that equivocation is bounded between 0 and \( n-k \).

For type-I channels, the results are plotted as average equivocation versus the erasure probability of the wiretapper’s channel. For type-II channels, average equivocation is plotted against the number of erasures.

Though simulation results cannot be used as rigorous proofs of the security objective in wiretap channels, they are useful in validating and illustrating the theoretical inferences. Also, for very small blocklengths, when all possible combinations of erasure locations can be exhausted, numerical verification of theoretical bounds is possible.

Fig. 2 shows the equivocation graph, when the code \( C \) is fixed to be the (15,11) Hamming code. The generalized Hamming weights of the dual (15,11) Hamming code are known to be (8,12,14,15). As expected, the steps in the graph of Fig. 2 occur at these numbers. Since the blocklength is very small, we have exhausted all possible erasure locations. Therefore, this plot is an accurate verification of (6) for the (15,11) Hamming code.

Fig. 3 shows the equivocation graph for the (63,37) EG code in the wiretap-II model. We notice a gradual increase in equivocation for small erasure probabilities and a step behavior for larger erasure probabilities. Seen from the right, the first

Fig. 4 shows the equivocation graph for the (63,37) EG code in the wiretap-I model. We notice a gradual increase in equivocation for small erasure probabilities and a step behavior for larger erasure probabilities. Seen from the right, the first
down-step in equivocation occurs at an erasure probability of about 0.8.

In summary, the simulations for small blocklengths show that the performance of EG codes in wiretap channels of types I and II are as predicted by the Generalized Hamming weights. Since the LDPC threshold analysis will probably hold for very long blocklengths, simulations results for verifying the security thresholds are time-consuming and difficult to run.

V. CONCLUSION

EG-LDPC codes hold a lot of promise in providing information-theoretic security in the erasure wiretap channel model. The existence of a LDPC matrix for these codes enables analysis and use in the wiretap-I mode. The cyclic structure enables study of security properties in the wiretap-II model. In both models, encoding and decoding are of low complexity and involve only shift register operations.

As future work, the generalized Hamming weights of these codes can be bounded better. The analysis in wiretap-I channel model for finite blocklength will provide better insight into the working of these codes.

REFERENCES