Generating a Binary Symmetric Channel for Wiretap Codes

Willie K. Harrison*, Telmo Fernandes†, Marco A. C. Gomes†, and João P. Vilela‡
*Department of Electrical and Computer Engineering, Brigham Young University, Provo, UT, USA
†Instituto de Telecomunicações, Department of Electrical and Computer Engineering, University of Coimbra, Portugal
‡CISUC and Department of Informatics Engineering, University of Coimbra, Portugal
Emails: willie.harrison@byu.edu, telmofrnnds@gmail.com, marco@co.it.pt, jpvilela@dei.uc.pt

***Abstract*—In this paper, we fill a void between information theoretic security and practical coding over the Gaussian wiretap channel using a three-stage encoder/decoder technique. Security is measured using Kullback-Leibler divergence and resolvability techniques along with a limited number of practical assumptions regarding the eavesdropper's decoder. The results specify a general coding recipe for obtaining both secure and reliable communications over the Gaussian wiretap channel, and one specific set of concatenated codes is presented as a test case for the sake of providing simulation-based evaluation of security and reliability over the network. It is shown that there exists a threshold in signal-to-noise (SNR) ratio over a Gaussian channel, such that receivers experiencing SNR below the threshold have no practical hope of receiving information about the message when the three-stage coding technique is applied. Results further indicate that the two innermost encoding stages successfully approximate a binary symmetric channel, allowing the outermost encoding stage (e.g., a wiretap code) to focus solely on secrecy coding over this approximated channel.**

***Index Terms*—Physical-layer security, Gaussian wiretap channel, practical secrecy coding.**

EDICS: CIT-PHY, CIT-PHY-COD, CIT-INF-SECC.

I. Introduction

Physical-layer security has been advancing at a rapid pace of late. The origins of the field can be traced back to Shannon [1] and Wyner [2], while some of the more recent advances are highlighted in [3], [4]. In this paper, we focus on coding for secrecy, and highlight the need for a hybrid security standard for real networks and finite blocklength codes. Semantic secrecy and strong secrecy are now the metrics of choice among information theorists [5], [3], [6], while security gap and bit-error rate are supreme among practical researchers [7], [8]. Unfortunately, it can be shown that both of these methods fall short of measuring secrecy when finite blocklength codes are deployed in real networks.

This work was partially funded by the following entities and projects: the US National Science Foundation (Grant Award Number 1761280), the FLAD project INCISE (Interference and Coding for Secrecy), project SWING2 (PTDC/EEI-TEL/3684/2014), funded by Fundos Europeus Estruturais e de Investimento (FEEI) through Programa Operacional Competitividade e Internacionalização - COMPETE 2020 and by National Funds from FCT - Fundação para a Ciência e a Tecnologia, through projects POCI-01-0145-FEDER-016753 and UID/EEA/50008/2013.

No explicit wiretap codes are currently known that can achieve both reliability and information theoretic security (weak, strong, or semantic) over a Gaussian wiretap channel, despite the recent works on lattice coding, which show the existence of such codes [9], but do not show how to design them [10]. It is furthermore true, with the exception of very few works [11], [12], that the security analysis of wiretap codes tends to focus on the infinite blocklength regime. The current status of practical information theoretic security coding for the wiretap channel is that we yet lack knowledge of how to code at finite blocklengths over real-world channels, as almost all known code constructions achieve secrecy only over a discrete memoryless channel (DMC)-based wiretap model, often when the legitimate receiver's channel is assumed to be noiseless. For extensive summaries of the state-of-the-art in explicit code constructions for information theoretic security, the reader is directed to [3], [5], [13] and references therein.

Security gap results do offer explicit code designs at finite blocklength over Gaussian and fading channels [7], [8], and works that highlight these techniques tend to simulate their performance as well. However, the security gap relies on the average bit/block error rate to quantify security, which is far from the preferred strong and semantic secrecy measures used in information theoretic approaches. One of the main drawbacks of the security gap is that it does not measure correlation between the message and the received data, which could be exploited to attack the system. Information theoretic metrics measure correlation by definition. Security gap results are also only indicative of the average performance. For smaller blocklength codes, a full understanding of the distribution of error rates would be more valuable [14], [15]. For a summary of practical physical-layer security coding, we direct the reader to [4].

In this work, we present a general concatenated coding technique that can be used to achieve reliable and secure communications over the wiretap channel when both the main and eavesdropper's channels are additive white Gaussian noise (AWGN) channels. Our technique is a hybrid approach between the two existing competing strategies. The information theoretic approach to coding for secrecy is to employ a nested wiretap code structure, mapping each message to one of several codewords at random to increase the confusion of an eavesdropper. The security gap approach to coding for secrecy

is to achieve an extremely steep waterfall region in the bit-error rate (BER) curve of a code so that only a small advantage in channel quality is required over the eavesdropper to achieve both reliability and security in practice. Our approach uses ideas from both of these techniques. The process requires three layers of coding altogether; the inner-most layer is used to fine tune a reliability threshold for the intended receiver, while the second layer then propagates any remaining errors at the output of the inner-most decoder. The main novelty of our approach lies in using the two inner layers of coding over an AWGN channel to *generate* an effective noisy DMC for any receiver with signal-to-noise ratio (SNR) below the reliability threshold. The outer-most code is then a nested wiretap code (e.g., as in [2], [11], [16]), that has been shown to be able to achieve information theoretic security over many DMC models, as we provide with our two inner layers of coding.

Security guarantees are made using a combination of a limited number of reasonable decoder assumptions at the eavesdropper, and *resolvability* calculations with signals along the eavesdropper's receiver chain. To be clear, we adopt a similar approach as in [17], where a network is deemed secure if

$$\lim_{n\to\infty} \mathrm{dist}(p_{MA}; p_M p_A) = 0, \tag{1}$$

where, M is the message to be securely communicated, $A = Z^n$ is the eavesdropper's length-n observation directly at the output of the channel (see also Fig. 1), n is the blocklength of an encoding process at Alice, and $\mathrm{dist}(\cdot\,;\cdot)$ could be any meaningful distance measure between two distributions. When the distance between the distributions p_{MA} and $p_M p_A$ goes to zero, this indicates that the signals M and A are tending towards statistical independence, and thus the joint distribution and the product of the marginals become indistinguishable, or *unresolvable*. A small number of modifications to (1) are necessary to apply the technique in our case, and we will use similar measures when analyzing the effectiveness of a specific set of concatenated codes in *generating* a DMC for the eavesdropper.

The rest of the paper is organized as follows. Section II further motivates the approach being taken in this paper for achieving both reliable and secure communications over the Gaussian wiretap channel. The system model assumed by the paper and the specific security constraint imposed on the model are presented, and some additional related work is reviewed. In Section III, the novel three-stage coding technique is given generally, while in Section IV, we provide a specific set of codes that together form an implementation of the system. Theoretical and simulation-based analysis of the three-stage encoding/decoding technique is given in Section V, where it is shown that the inner two coding layers can be used to effectively generate a DMC for the eavesdropper, and a coset-based secrecy code at the outer coding layer provides practical secrecy. Results are given in Section VI, and the paper is concluded in Section VII.

II. Setup and Motivation

A. System Model

In this work, we consider a modern version of the Gaussian wiretap channel model originally presented in [18]. Notationally, capital letters are random variables, or random vectors where a superscript indicates the length of the vector. In Fig. 1, we see a network with three players; a source (Alice), a destination (Bob), and an eavesdropper (Eve). Alice encodes a message M with an underlying discrete alphabet $\mathcal{M} = \{1, 2, \ldots, |\mathcal{M}|\}$ into a length-n vector of symbols X^n, which is broadcast over two channels: a main channel, and an eavesdropper's channel. Both channels are AWGN channels, with additive noise sequences N_B^n for the main channel and N_E^n for the eavesdropper's channel. The two channels are assumed to be independent with noise variances σ_B^2 and σ_E^2, respectively. Then Bob receives

$$Y^n = X^n + N_B^n, \tag{2}$$

and Eve receives

$$Z^n = X^n + N_E^n. \tag{3}$$

Bob's (Eve's) decoder attempts to recover the message and has as its output $\hat{M}$ ($\tilde{M}$). The communications goal is to broadcast the data under two constraints:

1) $\Pr(M \neq \hat{M}) < \delta$, (reliability constraint) and
2) $\mathrm{dist}(p_{MA}; p_M p_A) < \epsilon$, (security constraint).

We assume that both δ and ϵ are very small positive real numbers, and for more meaningful operation of the system, we'd like them to be as small as possible. Notice that since we consider finite blocklength codes, we do not allow the security constraint to be evaluated in the limit as $n \to \infty$. The random variable A is a placeholder that can be filled in with any of a number of variables that represent signals along Eve's receiver chain. For this work, we consider the Kullback-Leibler (KL) divergence (or *relative entropy*) [19] as the distance metric of choice, i.e.,

$$\mathrm{dist}(p_{MA}; p_M p_A) = \mathbb{D}(p_{MA} || p_M p_A) \tag{4}$$

$$= \int_{a\in\mathcal{A}} \sum_{m\in\mathcal{M}} p_{MA} \log_2 \frac{p_{MA}}{p_M p_A}\, \mathrm{d}a \tag{5}$$

when A is a continuous random variable, and

$$\mathrm{dist}(p_{MA}; p_M p_A) = \mathbb{D}(p_{MA} || p_M p_A) \tag{6}$$

$$= \sum_{a\in\mathcal{A}} \sum_{m\in\mathcal{M}} p_{MA} \log_2 \frac{p_{MA}}{p_M p_A} \tag{7}$$

when A is a discrete random variable. In both cases [17],

$$\mathbb{D}(p_{MA} || p_M p_A) = \mathbb{I}(M; A) = \mathbb{H}(M) - \mathbb{H}(M|A), \tag{8}$$

where $\mathbb{I}(\cdot;\cdot)$ and $\mathbb{H}(\cdot)$ are the well-known mutual information and entropy functions, respectively [19].

Since the encoder is a three-stage process, A can be set to be any of the three decoder outputs with tradeoffs between the complexity of the calculation of (8) and the meaningfulness of security results for each choice. Depending on the decoding methods, some signals in the chain will have continuous distribution functions, while some will have

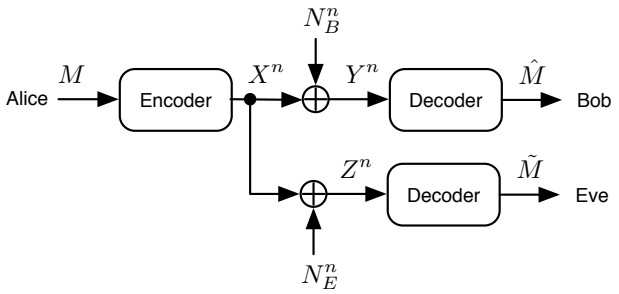

Fig. 1. The Gaussian wiretap channel model. Alice attempts to communicate reliably to Bob while keeping her messages secure against eavesdropping by Eve. Both channels of communication are Gaussian.

discrete distribution functions. The traditional approach is to let $A = Z^n$, as the data processing theorem [19] indicates that $\mathbb{I}(M; Z^n) \geq \mathbb{I}(M; A)$ for any signal A at the output of a decoding stage in the eavesdropper's receiver chain. The lefthand side of the inequality gives the strongest security guarantee, and some form of this expression is used in determining weak, strong, and semantic secrecy [5], [13]; however, the righthand side may give meaningful practical measures and be easier to simulate, particularly when A is a discrete random variable.

B. Additional Related Work

It has been noted by others that one of the main issues that continues to plague secrecy coding is perhaps our failure to adequately address security in the finite blocklength regime, particularly over practical channel models such as Gaussian and fading, where the friendly parties are not assumed to have noiseless channels ([4] Chapter 1, [5]).

There have been other attempts to overcome the shortcomings of coding over continuous channels, although not necessarily at finite blocklengths. In [5], [13], [20], the authors note that every binary-input communication channel can be modeled as a binary erasure channel (BEC) followed in series by an additional channel. Since we know how to code for secrecy over an eavesdropper's BEC, we can then simply ignore the additional channel, and code for the BEC. It is well known [3] that if a code achieves a level of equivocation $\mathbb{H}(M|Z^n)$ over a channel, then the equivocation can be no less over channels that are stochastically degraded[1] with respect to the original channel. However, this technique often delivers pessimistic results, indicating that an eavesdropper's SNR must be far below a level where signal detection could even be achieved to guarantee semantic (or strong) secrecy [5]. However, at these SNR levels, we have a hard time justifying the label of *eavesdropper* on Eve, since she would be hard pressed to even detect the legitimate communications in practice.

There also exists another approach to quantifying physical-layer security that involves bounding the equivocation as a function of blocklength using Fano's inequality [21], [22], but these bounds prove to be incredibly loose for small blocklength codes, and hence have limited utility in the finite (and short) blocklength regime.

[1]Channel one is *stochastically degraded* with respect to channel two, if there exists a channel three such that channel one's noise parameters are identical to the noise parameters of the combined channel made by serial concatenation of channel two with channel three [3].

III. Three-Stage Approach to Secrecy Coding

In this paper, we achieve both reliability and security over the Gaussian wiretap channel shown in Fig. 1 using a three-stage encoding process. The high-level procedure is outlined in Fig. 2, where we see a message M as the overall input to the system by Alice, and three encoder blocks at the transmitter. The outputs of the respective encoders are U^{n_u}, V^{n_v}, and X^n. If we assume the message is uniform over $\mathcal{M}$, then the rate of the Stage 1 encoder is

$$R_1 = \frac{\log_2 |\mathcal{M}|}{n_u}, \tag{9}$$

the rate of the Stage 2 encoder is

$$R_2 = \frac{n_u}{n_v}, \tag{10}$$

and the rate of the Stage 3 encoder is

$$R_3 = \frac{n_v}{n}. \tag{11}$$

Overall, if the alphabets of all three output signals are identical, then the three-stage encoder has a rate of

$$R = R_1 R_2 R_3 = \frac{\log_2 |\mathcal{M}|}{n}, \tag{12}$$

which is measured in bits per channel use, assuming one element of the vector X^n is transmitted with exactly one use of the Gaussian wiretap channel. Bob (Eve) receives Y^n (Z^n) as before, and the outputs of the decoder stages are labeled to correspond to their associating signals at the transmitter; that is, Bob's (Eve's) estimates of U^{n_u} and V^{n_v} are $\hat{U}^{n_u}$ and $\hat{V}^{n_v}$ ($\tilde{U}^{n_u}$ and $\tilde{V}^{n_v}$), respectively, and finally, $\hat{M}$ ($\tilde{M}$) is Bob's (Eve's) estimate of M. When writing the lower-case realizations of these random vectors, we omit the superscripts.

The types of codes used at the distinct stages of the encoding/decoding process and the goals of each stage are given as follows. Decoders are presented using Bob's variables, but we assume that Eve also has access to the decoders, and can use them in the same way that Bob does.

S1: The code employed at Stage 1 is a secrecy code (e.g., as in [5], [13]) capable of achieving information theoretic security over a DMC. The respective encoder and decoder functions for the Stage 1 code are defined generally as

$$u = \phi_1(m), \tag{13}$$
$$\hat{m} = \psi_1(\hat{u}). \tag{14}$$

Goal: ensure that $\mathrm{dist}(p_{M\tilde{M}}, p_M p_{\tilde{M}}) < \epsilon$ by wrapping the other two stages of coding in a randomized nested code structure for guaranteeing confusion at the eavesdropper.

S2: The code employed at Stage 2 is an error propagation code (e.g., a scrambler, interleaver, or some type of

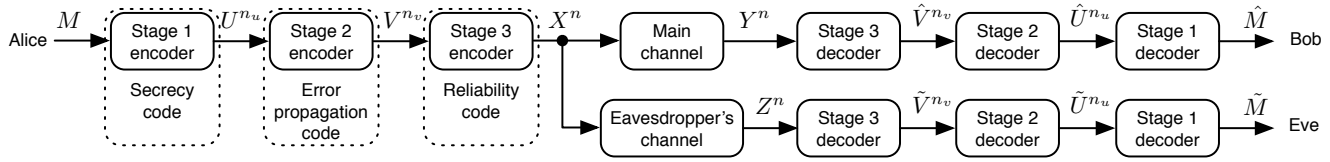

Fig. 2. Three-stage coding approach to secure communications. Stage 3 is fine-tuned to provide error-free communication to Bob, Stage 2 propagates any remaining errors throughout the output bits, Stage 1 wraps the entire procedure in a code known to achieve information theoretic security.

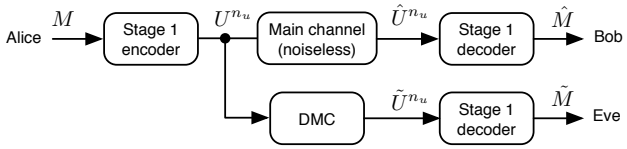

Fig. 3. Stages 2 and 3 and the Gaussian channel from Fig. 2 combine to generate a noiseless channel for Bob, and an effective DMC for Eve. For binary codes, Eve's generated channel is a BSC.

hashing function). The respective encoder and decoder functions for the Stage 2 code are

$$v = \phi_2(u), \tag{15}$$
$$\hat{u} = \psi_2(\hat{v}). \tag{16}$$

Goal: ensure that $\mathrm{dist}(p_{M\tilde{U}^{n_u}}, p_M p_{\tilde{U}^{n_u}}) < \epsilon$ by spreading any errors remaining from the Stage 3 decoder throughout the vector of symbols $\tilde{U}^{n_u}$ in Eve's receiver chain.

S3: The code employed at Stage 3 is an error correcting code. The respective encoder and decoder functions for the Stage 3 code are

$$x = \phi_3(v), \tag{17}$$
$$\hat{v} = \psi_3(y). \tag{18}$$

Goal: ensure $\Pr(V^{n_v} \neq \hat{V}^{n_v}) < \delta'$, for δ' sufficient to guarantee $\Pr(M \neq \hat{M}) < \delta$ through careful tuning of code parameters so as to avoid all errors being corrected in $\tilde{V}^{n_v}$.

These goals are chosen because the quantity $\mathrm{dist}(p_{MZ^n}; p_M p_{Z^n})$ is particularly difficult to calculate for sufficiently complicated encoding procedures, and the closed-form expression for p_{MZ^n} may be unknown. This same issue plagues all current works that propose explicit coding schemes over the Gaussian wiretap channel, which is one of the main reasons practical works focus on bit-error rate rather than information theoretic security. For us to maintain any useful definition of security, we therefore must assume that Eve employs the best known decoder at Stage 1, Stage 2, and Stage 3.

IV. Code Choices for an Implementation

In this section we define a set of coding algorithms that can be used in the three-stage encoder/decoder methodology of this paper. We note that these are not the only codes that can be employed within the three-stage scheme, but the simulation results of this paper will focus on only these codes. For this particular implementation, all codes are assumed to be binary, and therefore, signals and encoding operations are assumed to occur over the binary finite field $\mathbb{F}_2$ unless otherwise stated.

A. Coset Coding

The Stage 1 code can be taken to be a coset code as originally presented in [2], [23]. This technique was later used to provide the first explicit code construction capable of achieving weak secrecy over the binary erasure wiretap channel (BEWC) in [16], and the scheme has been used as the basis for additional coding constructions that achieve both strong and semantic secrecy over DMC wiretap models [5], [13]. In essence, each message $m \in \mathcal{M} = \{1, 2, \ldots, 2^k\}$ is assigned a unique coset of an $(n_u, n_u - k)$ binary linear block code. The 2^k cosets are labeled as $C_0, C_1, \ldots, C_{2^k-1}$, where C_0 is the linear block code whose $(n_u - k) \times n_u$ generator matrix is called G. Each coset is comprised of 2^{n_u-k} codewords [24], and a message m is encoded by choosing uniformly at random one of the codewords from C_m. To be specific,

$$\phi_1(m) = \begin{bmatrix} m & m' \end{bmatrix} \begin{bmatrix} G' \\ G \end{bmatrix} = u, \tag{19}$$

where G' is a $k \times n_u$ matrix whose rows are chosen so that $G^* = \begin{bmatrix} G' \\ G \end{bmatrix}$ has full rank in $\mathbb{F}_2$. Let H be the $k \times n_u$ parity check matrix of C_0. In [11] it was shown that the rows of G' can also be chosen so that the syndrome $s = uH^T$ of u for $\{u : \phi(m) = u\}$ is equal to m when m is written in binary form. This should not surprise us since syndrome decoders for error-correcting linear block codes make corrections as a function of the coset [24]. The variable m' is an $(n_u - k)$-length binary vector chosen uniformly at random from $\mathbb{F}_2^{n_u-k}$. The result of the encoding can be summarized as: m chooses the coset, and m' chooses the random representative of the coset. We will assume the decoder to be a simple syndrome calculation, and we choose the rows of G' to let $m = s$. Then the decoder function at Stage 1 is

$$\psi_1(\hat{u}) = s = \hat{u}H^T = \hat{m}. \tag{20}$$

Note that the secrecy code at Stage 1 has no error correction capability due to all vectors in $\mathbb{F}_2^{n_u}$ belonging to exactly one of the 2^k cosets. The variable U^{n_u} is a discrete random vector with elements in $\mathbb{F}_2^{n_u}$ for this choice of a Stage 1 code.

Note also that this code can be applied as presented here, with n_u as the blocklength of the secrecy code, or the blocklength of the secrecy code can be taken to be smaller,

and several secrecy codewords can be concatenated to form U^{n_u}. For any coset code, the encoder and decoder require at most additional computations on the order of n_u^3 from the matrix multiplications. The encoder of some coset codes can be made more efficient as is shown in [16] by adjusting the code matrices similarly as in [25].

B. Interleaved Coding for Secrecy with a Hidden Key

The Stage 2 and Stage 3 codes can be derived from the interleaved coding for secrecy with a hidden key (ICSHK) scheme that was originally presented in [15]. The technique is based on a combination of key-based interleaving and a powerful systematic punctured error-correcting code for key distribution and secure communications. The ICSHK scheme is fully described by Fig. 4. At the transmitter, a binary, length-n_k secret key K is generated uniformly at random, which is then mapped to a specific permutation Π_k. When the interleaver performs the specified permutation of the input symbols U^{n_u}, the output is a shuffled codeword U_i of length n_u. The secret key K and the interleaved codeword U_i are then concatenated and passed on to the Stage 3 encoder. The Stage 2 encoder can thus be written as

$$\phi_2(u) = \begin{bmatrix} k & \Pi_k(u) \end{bmatrix} = \begin{bmatrix} k & u_i \end{bmatrix} = v, \tag{21}$$

where k, u, u_i, and v are realizations of the random variables K, U^{n_u}, U_i, and V^{n_v}. From this, we can deduce that the length of the key is $n_k = (n_v - n_u)$.

The Stage 3 code is a powerful systematic binary (n, n_v) punctured block code $\mathcal{C}$. The systematic code is selected to code the input $V^{n_v} = \begin{bmatrix} K & U_i \end{bmatrix}$ resulting in $\begin{bmatrix} K & U_i & P_b \end{bmatrix}$, where P_b are the parity bits, and the input appears explicitly in the output, as is true for all systematic codes. Let the unpunctured rate of this code be denoted $\mathcal{R}$. Then it is not difficult to show that P_b is $\frac{n_v(1-\mathcal{R})}{\mathcal{R}}$ bits long. This codeword is then punctured by removing the secret key bits K prior to transmission, resulting in a *hidden key* that can be recovered as long as the Stage 3 decoder operates without error. Let the $n_v \times \frac{n_v}{\mathcal{R}}$ generator matrix of $\mathcal{C}$ be denoted $\mathcal{G}$, and the Stage 3 encoder can then be written as

$$\phi_3(v) = [v\mathcal{G}]_{(n_k+1:\frac{n_v}{\mathcal{R}})} = x, \tag{22}$$

where the subscript notation indicates that all but the first n_k bits are kept (not punctured) from the multiplication of v with $\mathcal{G}$. A new key is generated for each input word u, meaning n_k key bits of information are effectively embedded within the parity bits P_b by the encoding with code $\mathcal{C}$. The codeword that is finally transmitted over the Gaussian wiretap channel is then essentially

$$X^n = \begin{bmatrix} U_i & P_b \end{bmatrix}, \tag{23}$$

meaning that the code rate $\mathcal{R}$ of $\mathcal{C}$ should be chosen so that

$$\mathcal{R} = \frac{n_v}{n - n_u + n_v}, \tag{24}$$

which is determined by analyzing the lengths of the two vectors that make up X^n. Notice we say that X^n is *essentially* as shown in (23). This is because technically X^n is the modulated set of symbols that are transmitted over the channel. If we assume binary phase-shift keying (BPSK) modulation for this implementation, then the spirit of the encoder is preserved as 0 and 1 map to $+1$ and -1, respectively. For this paper, $\mathcal{C}$ is chosen to be a low-density parity-check (LDPC) code of appropriate blocklength and rate.

The decoder function chosen for Stage 3 is a soft-decision density evolution decoder, which is known to allow LDPC codes to approach the capacity of many communication channels [24], [26]. We assume that both Bob and Eve have access to this best-known decoder, and define

$$\psi_3(y) = \begin{bmatrix} \hat{k} & \hat{u}_i \end{bmatrix} = \hat{v} \tag{25}$$

to be the output of the density evolution decoder, returning a vector of soft decoded bits. Although the decoder returns soft information about both $\hat{K}$ and $\hat{U}_i$, there are no soft-decision algorithms for deinterleaving a vector. Thus, hard decisions must be made about $\hat{K}$, so as to apply the best estimate of the inverse interleaver, which achieves the proper shuffling only when $\hat{K} = K$. We thus define the Stage 2 decoder as

$$\psi_2(\hat{v}) = \Pi_{\hat{k}}^{-1}(\hat{u}_i) = \hat{u}. \tag{26}$$

At this point, we still have soft information about U^{n_u}, with no way to further exploit it since the Stage 1 code is a secrecy code with no correction capabilities. This scheme is efficient because it requires only an error-correcting code and an interleaver. The interleaver is essentially a standard permutation block (P-box) from cryptography, and can even be dealt with in hardware [27]. Several efficient error-control codes exist [24], [25], [26], and it is reasonable to restrict ourselves to these when designing the three-stage coding scheme.

V. System Analysis

A. Three-Stage Encoder

Analysis of the three-stage scheme with codes as defined in Section IV is not trivial. Consider the encoder, and let us identify joint distributions between M and signals along the transmitter chain of Fig. 2 with reference to Fig. 4 for the specific code types we have chosen to highlight. Recall that ϕ_i indicates the Stage i encoder, while ψ_i indicates the Stage i decoder. If we assume that M is uniform over $\mathcal{M} = \{1, 2, \ldots, 2^k\}$, then it is fairly straightforward to convince ourselves that the output of ϕ_1, U^{n_u}, is also uniform as

$$p_U(u) = \sum_{m \in \mathcal{M}} p_{U|M}(u|m) p_M(m) = \frac{1}{2^{n_u}}, \forall u \in \mathbb{F}_2^{n_u}, \tag{27}$$

where $p_{U|M}(u|m)$ is uniform over the coset corresponding to the message, and $p_M(m) = \frac{1}{2^k}$ for all $m \in \mathcal{M}$. The joint distribution between M and U^{n_u} is given by

$$p_{MU}(m, u) = \begin{cases} \frac{1}{2^{n_u}}, & \text{if } uH^T = m \\ 0, & \text{otherwise.} \end{cases} \tag{28}$$

The distribution on the output of ϕ_2, $V^{n_v} = \begin{bmatrix} K & U_i \end{bmatrix}$, is more difficult, as there is no guarantee that the key space is large

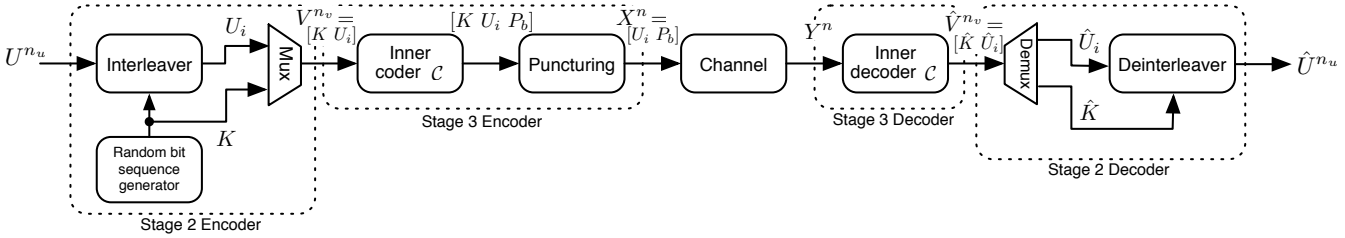

Fig. 4. High-level overview of the interleaved coding for secrecy with a hidden key (ICSHK) scheme using Bob's receiver variables. Eve's receiver is assumed to be identical, but with signals marked with a tilde symbol, rather than a hat symbol.

enough to allow for all possible permutations of U^{n_u}. The choice of the key size $n_k = (n_v - n_u)$ determines which is bigger, the total number of keys 2^{n_k}, or the total number of permutations over n_u bits (the factorial of n_u, $n_u!$). No matter the relationship, however, $p_V(v)$ is still uniform, because for every permutation of n_u bits, a uniform distribution in U simply maps to another uniform distribution over the same space in U_i. Since the keys are chosen uniformly at random, both portions of V^{n_v}, U_i and K, are distributed uniformly. The joint distribution

$$p_{MV}(m,v) = \sum_{u \in \mathbb{F}_2^{u_n}} p_{V|U}(v|u) p_{MU}(m,u), \tag{29}$$

requires us to know $p_{V|U}(v|u)$, which is found by sweeping the key space $\mathcal{K}$ to identify all legitimate (u,v) pairs. This cannot be uniform in general, but may approximate the uniform distribution if the number of keys is close to the number of possible permutations. For this, we need

$$n_k \approx \log_2(n_u!). \tag{30}$$

Since the right-hand side of (30) is only rarely an integer, and since n_k grows much faster than n_u in this relationship, this assignment is undesirable. In practice, n_k will likely be smaller than $\log_2(n_u!)$, and hence, the scheme chooses one of 2^{n_k} possible permutations at the interleaver, uniformly at random, and a subset of possible permutations are chosen with probability zero. Clearly the encoder mapping is getting more complicated, and our ability to express joint distributions in closed form is beginning to deteriorate.

To make matters more difficult, in practice one may select one or more of the codes so that the blocklengths at various stages do not really "fit" each other. For instance, the secrecy code could be chosen so that n_u is much smaller than it should be to match sizing constraints with the remaining encoders. When this is done, codewords are concatenated to make blocks of the proper size for further processing, resulting in encoder output distributions that are challenging to generalize.

Finally, we consider the distribution on X^n, and note that it must be uniform over the codewords since the mapping from V^{n_v} to X^n is deterministic, one-to-one, and onto (and hence, invertible). The joint distribution over M and X^n is

$$p_{MX}(m,x) = \sum_{u \in \mathbb{F}_2^{u_n}} \sum_{v \in \mathbb{F}_2^{u_v}} p_{X|V}(x|v) p_{V|U}(v|u) p_{MU}(m,u), \tag{31}$$

where

$$p_{X|V}(x|v) = \begin{cases} 1, & \text{if } \phi_3(v) = x, \\ 0, & \text{otherwise.} \end{cases} \tag{32}$$

In essence, the three-stage mapping from M to X^n for a fixed m produces a set of possible codewords x, each one mappable from m for at least one set of randomly generated (m', k). Although ϕ_1 and ϕ_2 choose m' and k, respectively, uniformly at random, and X^n is uniform itself, however, the joint distribution $p_{MX}(m,x)$ is not uniform and it preserves the one-to-many mapping as a function of the coset arrangement in ϕ_1, and the key space $\mathcal{K}$ in ϕ_2. Since both $p_M(m)$ and $p_X(x)$ are uniform, then the closer $p_{MX}(m,x)$ tends to uniform, the closer $\mathbb{D}(p_{MX}(m,x)||p_M(m)p_X(x)) = \mathbb{I}(M;X^n)$ will be to zero. The perfect secrecy condition [1] is achieved when $\mathbb{I}(M;X^n) = 0$, indicating that the degree to which $\mathbb{D}(p_{MX}(m,x)||p_M(m)p_X(x))$ approaches zero is also the degree to which the encoding scheme approaches perfect secrecy.

B. Three-Stage Decoder

The reliability constraint is achieved if the soft-decision density evolution decoder ψ_3, provides sufficiently error-free data for Bob. Since the next decoder in the receiver chain is an error propagator, error events at Bob and Eve will be costly. Of course, the true goal is to choose the Stage-3 code $\mathcal{C}$ to provide error-free data for Bob, but not for Eve. In practice, this means that Bob's signal-to-noise ratio should be tuned to the code, or vice versa, and that the data can be secured against Eve if and only if Bob's channel quality exceeds that of Eve's. In other words, the secrecy capacity must be greater than zero to achieve both constraints, as is true for all physical-layer security schemes.

The security analysis we wish to complete to verify our security constraint requires us to know joint distributions $p_{MA}(m,a)$ for A set to each of Z^n, $\tilde{V}^{n_v}$, $\tilde{U}^{n_u}$, and $\tilde{M}$. After adding Gaussian noise to the modulated version of X^n, we consider $p_{MZ}(m,z)$. While $p_{XZ}(x,z)$ is fairly straightforward to obtain [19], [28], mapping that relation back to a joint distribution on M and Z^n is more challenging. The varying design choices in the three-stage encoder make the security analysis at Eve's receiver fairly complex. At this point, we suggest another approach that combines theory and simulation to check the security constraint for practical eavesdropping scenarios.

In the next two respective subsections of the paper, we consider some adaptations on blocklength requirements to facilitate faster simulation analysis, and adopt a new line of inquiry that yields greater insights on the problem. This new direction of analysis addresses the possibility that was brought up earlier when comparing Figs. 2 and 3. Namely, we consider that the key-based interleaving and error control coding with puncturing, collectively known as the ICSHK, along with a Gaussian channel, may provide characteristics of a DMC for the eavesdropper and a noiseless channel for Bob, over which one can employ secrecy coding. This possibility was originally investigated in [15], but is taken to new levels here.

C. Small Blocklength at the Secrecy Code

In this section, we consider our three-stage encoder when n_u is small so as to allow us to perform simulation-based analysis of the technique over the Gaussian channel. We do not require n_v nor n to be small, so we will allow multiple codewords at U to be concatenated to form a larger word for processing at the Stage 2 and Stage 3 encoders. To accomplish this for fixed m, we simply repeat the message enough times to do the encoding.

One of the divergences that we wish to measure is $\mathbb{D}\left(p_{M\tilde{U}}(m,\tilde{u})||p_M(m)p_{\tilde{U}}(\tilde{u})\right)$, and the desired outcome is that the divergence is very small, indicating that the joint distribution between M and $\tilde{U}$ is "close" to the independent case. Note that

$$\mathbb{D}\left(p(m,\tilde{u})||p(m)p(\tilde{u})\right) = \sum_{m\in\mathcal{M}}\sum_{\tilde{u}\in\tilde{\mathcal{U}}} p(m,\tilde{u})\log_2\frac{p(\tilde{u}|m)}{p(\tilde{u})}, \tag{33}$$

and that if for every fixed message $m \in \mathcal{M}$, $p_{\tilde{U}|m}(\tilde{u}|m)$ is identical, then this distribution becomes $p_{\tilde{U}|M}(\tilde{u}|m)$ in general, and likewise becomes $p_{\tilde{U}}(\tilde{u})$, which implies that (33) is zero. Thus, it is sufficient to show that the distribution on $\tilde{U}$ does not change for any possible fixed m.

Let us explore this idea theoretically by considering the expected weight of U given $M = m$, some fixed message. The coset-based secrecy encoder randomly selects a message from the coset that corresponds to the syndrome equal to m. We will make use of the following fact to deduce approximate distributions on U_i, and U given any fixed m, and hence, also deduce distributions on $\tilde{U}_i$ and $\tilde{U}$.

Lemma 1. The sum of the weight of codewords in any coset of an $(n, n-k)$ binary linear block code is identical, and is equal to $n2^{n-k-1}$.

Proof. It is assumed that the generator matrix G for the $(n, n-k)$ binary linear block code does not have any all-zero columns, which is a fair assumption since the addition of an all-zero column artificially inflates the blocklength without adding any capability to the code. Consider the ith bit in all codewords. This bit is guaranteed to be zero for the all-zero codeword, and guaranteed to be one for at least one codeword c since the ith column in G is not all-zero. Consider any other codeword c' and observe the ith bit. Then recognize that the ith bit of the codeword $(c \oplus c')$ must have the opposite value (either zero or one) from the ith bit of c'. Since all codewords can be added to c to produce other codewords, a complete matching of all codewords can be made by considering pairs of codewords (c_1, c_2) such that $c_1 = c \oplus c_2$.

To extend this result to any coset, recognize that a coset of the linear block code is simply an offset that can be attained by adding any codeword from the coset to all codewords in the linear block code. Thus, either the ith bit is flipped for all codewords in the mapping from the linear code to the coset, or it is not. Either way, the ith bit is one in exactly half of the codewords in the coset, and zero in the other half. □

The distribution on the weight of U given m is, therefore, identical for all m, and when interleaving is considered, as long as the key is chosen uniformly, then the distribution on U_i is identical for all m, meaning $p_{U_i|m}(u_i|m)$, $p_{V|m}(v|m)$, $p_{X|m}(x|m)$, and $p_{A|m}(a|m)$ for A set to any signal in the receiver chain of Bob or Eve, are all equal for all possible $m \in \mathcal{M}$. Recognize that in general, this can only be an approximate result, however, because the key space does not allow generally for all interleaving mappings to be chosen uniformly at random.

In Eve's decoder, as long as the key is not recovered perfectly, then the wrong de-interleaving mapping will be used. In [15], it was shown that the distribution in the number of errors after deinterleaving in the ICSHK scheme is tightly concentrated around half of the bits when Eve's SNR is such that the probability of at least one residual error after the Stage 3 decoder is close to one, and these bit locations appear to be randomly distributed throughout the codeword. Thus, although we still have soft information at $\tilde{U}_i$, we must make hard decisions on the bits in $\tilde{K}$ to perform the deinterleaving, and unless the key bits are all correct, the probability that any bit in $\tilde{U}$ is correct is close to 0.5. The end result is that (33) can be expected to be very close to zero when the eavesdropper's estimate of the interleaving key $\tilde{K}$ is incorrect.

D. Generating a Discrete Memoryless Channel

Although we are not entirely in the dark regarding evaluation of $\mathbb{D}(p_{MA}(m,a)||p_M(m)p_A(a))$ for A set to the signals in Eve's receiver chain, we may ask if the process can be simplified by analyzing the implications of Figs. 2 and 3. In other words, could the ICSHK algorithm as shown in Fig. 4 really mimic the statistics of a DMC; in particular, a BSC? Towards answering that question, we extend our previous results in [15], [29], where it was noted that for a BSC:

1) bit flips across the channel should occur independently,
2) bit flips should occur with some fixed probability p, and
3) soft information should not be usable at the output of the channel.

The works of [15], [29] went to some lengths to show through simulation that these properties were met in practice by the ICSHK approach. Our analysis here goes further to satisfy the claim that the ICSHK scheme, properly deployed, generates an effective BSC for Eve, and an effective noiseless channel for Bob.

Let us consider the scheme in Fig. 4 within the channel model illustrated in Fig. 5. In Fig. 5, we show the ICSHK

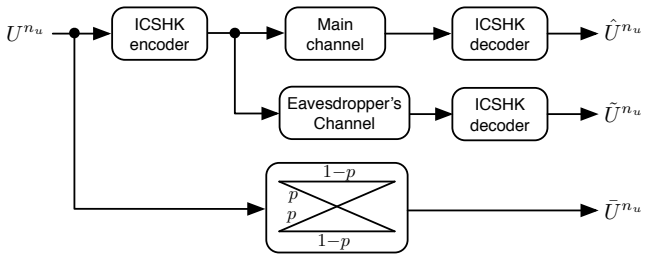

Fig. 5. A binary symmetric channel approximation to the eavesdropper's inner two stages of coding plus the channel is depicted here. We also present signal names to aid the reader in the technical analysis.

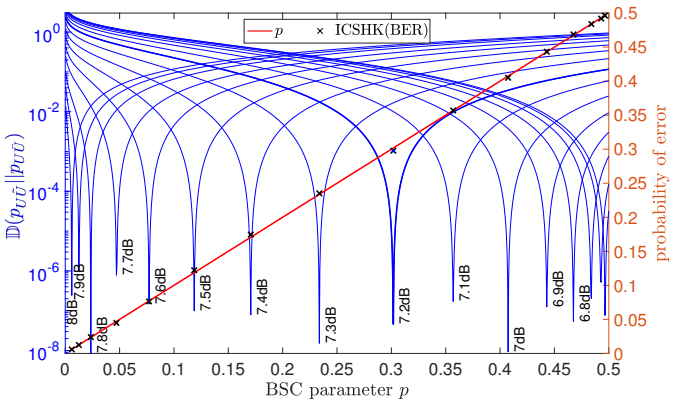

Fig. 6. Kullback-Leibler divergence $\mathbb{D}(p_{U\tilde{U}}||p_{U\bar{U}})$. Minima in curves show where the BSC model in Fig. 5 best approximates the ICSHK scheme operating over the Gaussian wiretap channel with BPSK modulation and $\sigma_E^2 = N_0/2$.

portion of the three-stage coding system along with the main and eavesdropper's channels. We also show an additional simple channel model, where the probability of a bit flip is given as p, and we label the output of the BSC(p) as $\bar{U}^{n_u}$. The comparison between the eavesdropper's receiver chain and the simple BSC model is done by calculating

$$\begin{aligned}\mathbb{D}(p_{U\tilde{U}}||p_{U\bar{U}}) &= \sum_u \sum_{u'} p_{U\tilde{U}}(u,u') \log_2 \frac{p_{U\tilde{U}}(u,u')}{p_{U\bar{U}}(u,u')} \\ &= \sum_u \sum_{u'} p_{U\tilde{U}}(u,u') \log_2 \frac{p_{\tilde{U}|u}(u'|u)}{p_{\bar{U}|u}(u'|u)}. \end{aligned} \quad (34)$$

Computing this quantity as both u and u' vary over all possible values of U^{n_u} is cumbersome, and results in unusable calculations without long simulations to accurately estimate probabilities near zero, as $\log_2 \frac{x}{0} = \infty$ for $x \neq 0$. To prevent this issue, we consider the expression as both u and u' simply vary over $\mathbb{F}_2$, which is justified by noting the seemingly independent nature of errors at the output of the ICSHK system [15], [29]. Since the distribution of the BSC input/output bits is theoretically known, its simulation is not necessary; but the distribution of input/output bits for the ICSHK case is simulated using a computer program.

In Fig. 6 we show calculations of (34) for 16 different cases of the ICSHK scheme. Gaussian noise in the eavesdropper's channel is assumed to have variance $\sigma_E^2 = N_0/2$. These cases are for values of 6.5 dB to 8.0 dB in the Gaussian channel's E_b/N_0 operating point in increments of 0.1 dB. In each case, we compute (34) between the target test case for the ICSHK scheme and all possible distributions based on the BSC(p) model, i.e., letting p range between zero and 0.5. We note that Fig. 6 shows a very close match in each case with a pronounced minimum value of the divergence, and as the SNR in the Gaussian channels goes down, the matching p value goes up as expected. These divergence calculations dip below 10^{-6} in the minimum case, indicating very close matches between the distributions.

Using a simpler technique, we find nearly identical results, as is also captured in Fig. 6. Suppose, we simply simulate the ICSHK scheme at a given E_b/N_0, and calculate the BER at that operating point. It turns out that this value matches the minimizing crossover probability p in the BSC almost exactly. These strong matches seem to verify our three-stage encoder/decoder system as an enabling approach for achieving practical security over the Gaussian wiretap model.

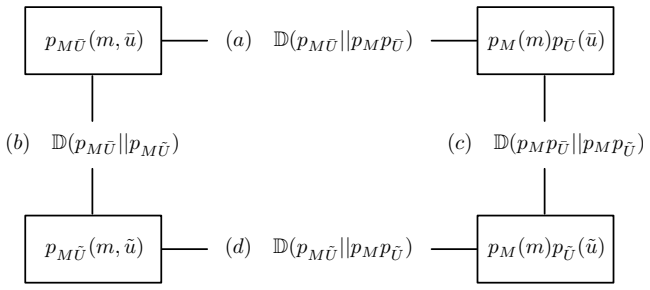

Fig. 7. Important distributions for security proofs and their KL divergences. We wish to make divergence (d) small to guarantee security for the three-stage coding scheme, as it is equal to the mutual information $\mathbb{I}(M;\tilde{U})$. Divergence (a) is equal to the mutual information $\mathbb{I}(M;\bar{U})$. Divergence (b) is simulated and understood to be small from Fig. 6. Divergence (c) can be shown to be zero. The implications of these divergences can then be levied on divergence (d).

E. Coset Coding as the Outer Code

The simulation results in Fig. 6 show very close matches between the joint distributions of $p_{U\tilde{U}}(u,\tilde{u})$ and $p_{U\bar{U}}(u,\bar{u})$ at certain operating points for the Gaussian and binary symmetric channels, respectively. We would like to convert this result to knowledge about $\mathbb{D}(p_{M\tilde{U}}||p_M p_{\tilde{U}})$. Recall, it is this quantity, and others like it, that allow us to address the security constraints of our system. Consider Fig. 7, where we note the following four divergence calculations:

(a) $\mathbb{D}(p_{M\bar{U}}||p_M p_{\bar{U}})$,
(b) $\mathbb{D}(p_{M\bar{U}}||p_{M\tilde{U}})$,
(c) $\mathbb{D}(p_M p_{\bar{U}}||p_M p_{\tilde{U}})$, and
(d) $\mathbb{D}(p_{M\tilde{U}}||p_M p_{\tilde{U}})$.

We know that there exist wiretap codes [5], [13] that can make divergence (a) go to zero as blocklength $n_u \to \infty$. This is the well-known strong secrecy constraint. For a fixed blocklength, we may write this as

$$\mathbb{D}(p_{M\bar{U}}||p_M p_{\bar{U}}) < \epsilon_a, \quad (35)$$

for some small $\epsilon_a > 0$. Divergence (b) can be bounded using our results in Fig. 6. Recall that these simulations imply that for a good match of channel parameters, then

$$\mathbb{D}(p_{U\tilde{U}}||p_{U\bar{U}}) < \epsilon_b, \tag{36}$$

for some small $\epsilon_b > 0$. From our simulations, ϵ_b appears to be no bigger than 10^{-6}, and this can likely be decreased with greater precision in the simulation.

Lemma 2. If $\mathbb{D}(p_{U\tilde{U}}||p_{U\bar{U}}) < \epsilon_b$, then $\mathbb{D}(p_{M\tilde{U}}||p_{M\bar{U}}) < \epsilon_b$, when codes are chosen for the three-stage coding scheme as in Section IV.

Proof. Note that

$$\mathbb{D}(p_{U\tilde{U}}||p_{U\bar{U}}) = \sum_u \sum_{u'} p_{U\tilde{U}}(u,u') \log_2 \frac{p_{\tilde{U}|u}(u'|u)}{p_{\bar{U}|u}(u'|u)} \tag{37}$$

$$= \sum_u \sum_{u'} \sum_m p_{U\tilde{U}M}(u,u',m) \log_2 \frac{p_{\tilde{U}|u,m}(u'|u,m)}{p_{\bar{U}|u,m}(u'|u,m)} \tag{38}$$

$$= \mathbb{D}(p_{MU\tilde{U}}||p_{MU\bar{U}}) \tag{39}$$

$$\geq \mathbb{D}(p_{M\tilde{U}}||p_{M\bar{U}}) \tag{40}$$

$$\implies \mathbb{D}(p_{M\tilde{U}}||p_{M\bar{U}}) < \epsilon_b, \tag{41}$$

which completes the constraint on divergence (b). Here (38) is due to the facts that $M \to U \to \tilde{U}$, and that $p_{MU\tilde{U}}(m,u,u') = 0$ when u is not in the mth coset (i.e., when $\psi_1(u) \neq m$). The expression in (40) is a direct application of the well-known equality [19]

$$\mathbb{D}(p(x,y)||q(x,y)) = \mathbb{D}(p(x)||q(x)) + \mathbb{D}(p(y|x)||q(y|x)) \tag{42}$$

and the fact that the divergence between any two distributions is nonnegative. □

Lemma 3. When a three stage coding scheme is deployed over an eavesdropper's Gaussian channel with code choices as in Section IV, and messages are assumed to be chosen uniformly at random, then $\mathbb{D}(p_M p_{\bar{U}}||p_M p_{\tilde{U}}) = 0$.

Proof. We first write

$$\mathbb{D}(p_M p_{\bar{U}}||p_M p_{\tilde{U}}) = \sum_m \sum_{u'} p_M(m) p_{\bar{U}}(u') \log_2 \frac{p_M(m) p_{\bar{U}}(u')}{p_M(m) p_{\tilde{U}}(u')}. \tag{43}$$

Marginal distributions on $\bar{U}$ and $\tilde{U}$ are both uniform over the same alphabet by symmetry of the problem, making this divergence (which is given as divergence (c)) exactly zero. □

At this point, we would like to declare victory; however, the KL divergence is not a true distance metric in the strictest sense of the word because it fails symmetry and the triangle inequality. Thus, it is not true that divergence (d) must be less than the sum of divergences (a), (b), and (c), and we can make no clear mathematical statement about divergence (d) as a function of ϵ_a and ϵ_b. However, the KL divergence is *like* a distance metric in that $\mathbb{D}(p(x)||q(x)) = 0$ iff $p(x) = q(x) \forall x$. Since divergences (a), (b), and (c) in Fig. 7 are all very close to zero, or exactly zero in one case, then we recognize that the distributions must be similar from step to step around the box of distributions in the figure. Although it is possible for divergence (d) to be greater than $\epsilon_a + \epsilon_b$, it is not possible for it to be much greater, as long as both ϵ_a and ϵ_b are very small.

Let us rewrite divergence (a) as

$$\mathbb{D}(p_{M\bar{U}}||p_M p_{\bar{U}}) = \sum_m \sum_{u'} p_{M\bar{U}}(m,u') \times \log_2 \left[\frac{p_{M\bar{U}}(m,u')}{p_M(m) p_{\bar{U}}(u')} \frac{p_{M\tilde{U}}(m,u')}{p_{M\tilde{U}}(m,u')} \right] \tag{44}$$

$$= \mathbb{D}(p_{M\bar{U}}||p_{M\tilde{U}}) + \sum_m \sum_{u'} p_{M\bar{U}}(m,u') \times \log_2 \frac{p_{M\tilde{U}}(m,u')}{p_M(m) p_{\tilde{U}}(u')}. \tag{45}$$

The entire expression of (45) is less than ϵ_a by (35), and the first term in (45) is less than ϵ_b by (36). The second term in (45) can be rewritten as

$$\sum_m \sum_{u'} \left(\frac{p_{M\bar{U}}(m,u')}{p_{M\tilde{U}}(m,u')} \right) p_{M\tilde{U}}(m,u') \log_2 \frac{p_{M\tilde{U}}(m,u')}{p_M(m) p_{\tilde{U}}(u')}. \tag{46}$$

Note that for the simple case where $p_{M\bar{U}}(m,u') = p_{M\tilde{U}}(m,u')$ for all m, u', then this term is exactly divergence (d), and it can then be bounded as a function of ϵ_a and ϵ_b. Thus, the smallness of divergence (d) is reliant on the match between the two joint distributions $p_{M\bar{U}}(m,u')$ and $p_{M\tilde{U}}(m,u')$. Measuring the divergence between the two is a reasonable approach to imply the closeness of the match.

This provides yet more evidence that $\mathbb{D}(p_{M\tilde{U}}||p_M p_{\tilde{U}})$ is small when wiretap codes are used that make $\mathbb{D}(p_{M\bar{U}}||p_M p_{\bar{U}})$ small. Note that for this to be a meaningful security metric, we must guarantee that the eavesdropper's mapping from Z^n to $\tilde{V}^{n_v}$ to $\tilde{U}^{n_u}$ includes best possible decoders for ψ_3 and ψ_2.

VI. Results

In this section, we choose specific codes for Stages 1, 2, and 3 in the three-stage coding scheme, and provide some simulation results. Let the linear block code at Stage 1 be the $(7,4)$ Hamming code so that $\mathcal{M} = \{0,1,\ldots,7\}$, messages are converted to their three-bit binary representations, and then encoded into $n_u = 7$ bit secrecy codewords. These length-7 codewords are buffered 174 at a time as input blocks for Stage 2 encoding. This gives an effective n_u value of 1218 bits, representing 522 message bits per block. The key size for these simulations was chosen to be 62 bits, yielding 2^{62} possible unique interleavers at Stage 2. Since finding the correct key is only part of what is required to correctly decode the message, this amounts to at least 62 bits of cryptographic strength when brute force on the key is tried [29], [30], [31]. The confusion on interleaved message bits provides even greater strength under such an attack scenario, and may prevent attackers from recognizing the correct key when it is tried. The shuffled bits of U_i and the key bits are then appended to make input blocks for the Stage 3 encoder. This yields $n_v = 1280$ bits, and requires the Stage 3 encoder to have dimension 1280. We choose a $(1536, 1280)$ irregular LDPC code as the Stage 3 encoder, and puncture the bits associated with the key in the systematic

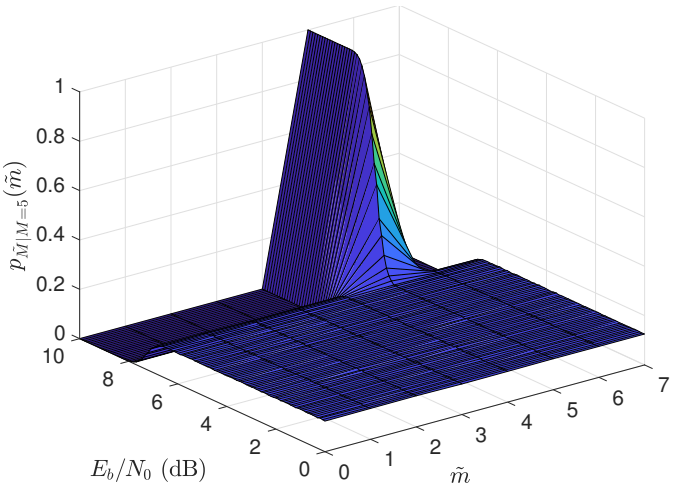

Fig. 8. The conditional distribution $p_{\tilde{M}|M=5}(\tilde{m})$ as a function of E_b/N_0 in a Gaussian channel when all three stages of encoding and decoding are employed. When all conditional distributions of this type are uniform, where the conditioning can be with respect to any valid message, then Eve's decoder has no information about the message. For this case, Eve's decoder outputs no information about M when $E_b/N_0 \leq 6.5$ dB. The system delivers the message reliably when $E_b/N_0 \geq 8$ dB.

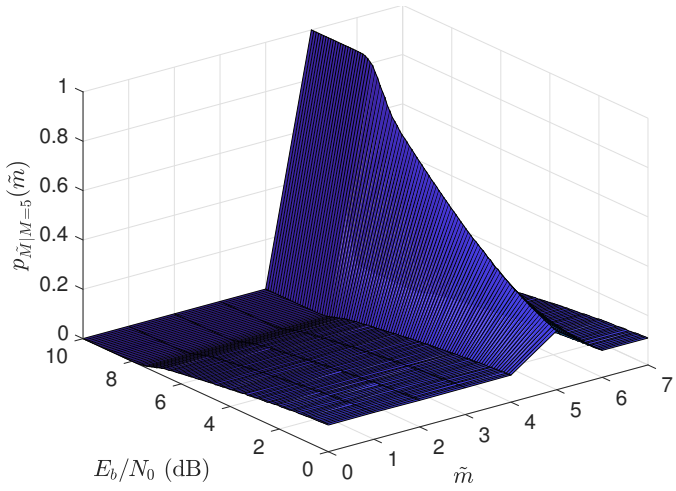

Fig. 9. The conditional distribution $p_{\tilde{M}|M=5}(\tilde{m})$ as a function of E_b/N_0 in a Gaussian channel when Stage 2 encoding and decoding is left out of the three-stage coding scheme. The information about the message is not well hidden, even when $E_b/N_0 = 0$ dB.

output prior to transmission. Thus, the total blocklength of the three-stage encoder is $n = 1474$, and the rate of the three-stage encoder is $R = 522/1474 \approx 0.3541$.

In Fig. 8, we show $p_{\tilde{M}|M=5}(\tilde{m})$ for the three-stage scheme over a range of E_b/N_0 values. Note from Fig. 6 that the ICSHK scheme generates BSC(p) for p ranging from 0 to 0.5 when E_b/N_0 ranges from 6.5 to 8 dB. It is, therefore, not surprising to see the change in the shape of $p_{\tilde{M}|M=5}(\tilde{m})$ in Fig. 8 occur over the same range of E_b/N_0. At E_b/N_0 below 6.5 dB, the distribution is uniform, while at E_b/N_0 above 8 dB, the true message is found with probability one. This same shape occurs when conditioning on any message $m \in \mathcal{M}$; i.e., at E_b/N_0 below 6.5 dB, Stages 2 and 3 generate an effective BSC(0.5) channel, which removes information about M, and at E_b/N_0 above 8 dB, Stages 2 and 3 generate an effectively noise-free channel. To see the role of the Stage 2 encoder and decoder more plainly in confusing the eavesdropper, we also display the same results when Stage 2 is left out in Fig. 9. Note here that we do not get the sharp transition in $p_{\tilde{M}|M=5}(\tilde{m})$ over a range of 1.5 dB in E_b/N_0, but rather information leaks to the eavesdropper about the message even at 0 dB. While an interleaver seems to work fine as the Stage 2 encoder, we suppose that some type of keyed hash function would provide stronger secrecy guarantees in practice. Note that as $p_{\tilde{M}|m}(\tilde{m})$ approaches uniform,

$$\mathbb{D}\left(p(m,\tilde{m})||p(m)p(\tilde{m})\right) \to 0, \tag{47}$$

indicating no information leakage at the output of ψ_3.

Finally, in Fig. 10, we simply plot the BER for five cases to gain additional intuition about desirable BER curve shapes for practical secrecy coding. Note that when all three stages of coding are employed, we observe an incredibly sharp waterfall region in the BER curve, as expected, and the sharpness of the waterfall lessens significantly when Stage 2 encoding and decoding is left out of the three-stage scheme. When a smaller LDPC code is used at Stage 3, the sharpness of the waterfall also decreases, although not as severely as when Stage 2 is removed. Both the larger and the smaller LDPC codes are chosen from the worldwide interoperability for microwave access (WiMAX) standard IEEE 802.16e [32]. The smaller code is a $(576, 480)$ irregular code, and we choose the key size to be 24 bits, so that the overall rate of the three-stage coding scheme is as close as possible to the case with the larger $(1536, 1280)$ code. The differences in the waterfall regions that we see from changing the size of the LDPC code at Stage 3 indicate that surface curves like unto Figs. 8 and 9 also exist for the smaller code case, but they are nearly identical in shape and form. The only difference is the slope of the transition from reliable to secure regions of operations in E_b/N_0. Finally, the uncoded BPSK curve is included as a reference.

VII. Conclusion

In conclusion, we propose a novel three-stage encoding scheme for physical-layer security over the Gaussian wiretap channel. While exact information theoretic claims of security are not given with this technique, practical measures of security are shown to be possible. We demonstrate how the KL divergence between several distributions changes as a function of E_b/N_0 over the Gaussian channel, and highlight the effectiveness of the ICSHK scheme at generating an effective discrete memoryless wiretap channel, such that operation at higher E_b/N_0 duplicates the performance of a noise-free channel and operation at slightly lower E_b/N_0 generates a BSC(p). Stage 1 secrecy coding (e.g., a wiretap code) can then deliver effectively secure and reliable communications.

References

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656–715, 1948.

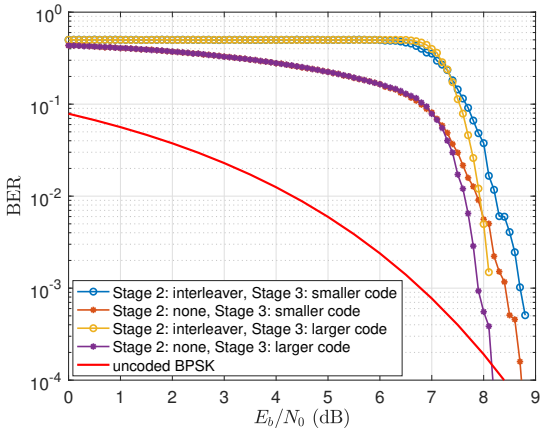

Fig. 10. BER results for the three-stage coding scheme with the $(7, 4)$ Hamming coset code at Stage 1, interleaving at Stage 2, and punctured LDPC coding at Stage 3. The smaller Stage 3 code under test is the LDPC WiMAX $(576, 480)$ code, which is paired with a key size of $n_k = 24$ bits at Stage 2. The larger Stage 3 code under test is the LDPC WiMAX $(1536, 1280)$ code, which is paired with a key size of $n_k = 62$ bits. The overall rate of the three-stage coding scheme is nearly identical for both cases, $R \approx 0.354$.

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

[3] M. Bloch and J. Barros, *Physical Layer Security : From Information Theory to Security Engineering*. Cambridge University Press, 2011.

[4] M. Baldi and S. Tomasin, Eds., *Physical and Data-Link Security Techniques for Future Communication Systems*, ser. Lecture Notes in Electrical Engineering. Switzerland: Springer International Publishing, 2016, vol. 358.

[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, October 2015.

[6] M. Bellare, S. Tessaro, and A. Vardy, "Semantic security for the wiretap channel," in *Advances in Cryptology CRYPTO 2012*, ser. Lecture Notes in Computer Science, R. Safavi-Naini and R. Canetti, Eds., vol. 7417. Springer Berlin Heidelberg, 2012, pp. 294–311, hard-copy.

[7] D. Klinc, J. Ha, S. W. McLaughlin, J. Barros, and B.-J. Kwak, "LDPC codes for the Gaussian wiretap channel," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 532–540, 2011.

[8] M. Baldi, M. Bianchi, and F. Chiaraluce, "Coding with scrambling, concatenation, and HARQ for the AWGN wire-tap channel: A security gap analysis," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 883–894, 2012.

[9] C. Ling, L. Luzzi, J. C. Belfiore, and D. Stehlé, "Semantically secure lattice codes for the Gaussian wiretap channel," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 6399–6416, Oct. 2014.

[10] F. Oggier, P. Solé, and J. C. Belfiore, "Lattice codes for the wiretap Gaussian channel: Construction and analysis," *IEEE Trans. Inf. Theory*, vol. 62, no. 10, pp. 5690–5708, Oct. 2016.

[11] J. Pfister, M. Gomes, J. P. Vilela, and W. K. Harrison, "Quantifying equivocation for finite blocklength wiretap codes," in *Proc. IEEE Int. Conf. Communications (ICC)*, Paris, France, June 2017, pp. 1–6.

[12] K. Zhang, M. Tomlinson, M. Z. Ahmed, M. Ambroze, and M. R. D. Rodrigues, "Best binary equivocation code construction for syndrome coding," *IET Communications*, vol. 8, no. 10, pp. 1696–1704, July 2014.

[13] 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, Sep. 2013.

[14] J. P. Vilela, M. Gomes, W. K. Harrison, D. Sarmento, and F. Dias, "Interleaved concatenated coding for secrecy in the finite blocklength regime," *IEEE Signal Processing Letters*, vol. 23, no. 3, pp. 356–360, Mar. 2016.

[15] D. Sarmento, J. Vilela, W. K. Harrison, and M. Gomes, "Interleaved coding for secrecy with a hidden key," in *2015 IEEE Globecom Workshops (GC Wkshps)*, Dec 2015, pp. 1–6.

[16] 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, August 2007.

[17] M. R. Bloch and J. N. Laneman, "Strong secrecy from channel resolvability," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8077–8098, Dec. 2013.

[18] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Transactions on Information Theory*, vol. 24, no. 4, pp. 451–456, July 1978.

[19] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ: John Wiley & Sons, Inc., 2006.

[20] R. Liu, Y. Liang, H. V. Poor, and P. Spasojevic, "Secure nested codes for type II wiretap channels," in *Proc. IEEE Information Theory Workshop (ITW)*, Lake Tahoe, CA, Sep. 2007, pp. 337–342.

[21] C. W. Wong, T. Wong, and J. Shea, "Secret-sharing LDPC codes for the BPSK-constrained Gaussian wiretap channel," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 551 –564, sept. 2011.

[22] M. Baldi, G. Ricciutelli, N. Maturo, and F. Chiaraluce, "Performance assessment and design of finite length LDPC codes for the Gaussian wiretap channel," in *Proc. IEEE Int. Conf. Communication Workshop (ICCW)*, June 2015, pp. 435–440.

[23] L. H. Ozarow and A. D. Wyner, "Wiretap channel II," *AT&T Bell Laboratories Technical Journal*, vol. 63, no. 10, pp. 2135–2157, December 1984.

[24] T. K. Moon, *Error Correction Coding : Mathematical Methods And Algorithms*. John Wiley & Sons, 2005.

[25] T. J. Richardson and R. L. Urbanke, "Efficient encoding of low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 638–656, Feb. 2001.

[26] T. Richardson and R. Urbanke, *Modern Coding Theory*. New York, NY: Cambridge University Press, 2008.

[27] C. Martins, T. Fernandes, M. Gomes, and J. Vilela, "Testbed implementation and evaluation of interleaved and scrambled coding for physical-layer security," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, June 2018, pp. 1–6.

[28] M. Rice, *Digital Communications: A Discrete-time Approach*. Pearson/Prentice Hall, 2009. [Online]. Available: https://books.google.com/books?id=EB3r7JtXlWwC

[29] W. K. Harrison, D. Sarmento, J. P. Vilela, and M. Gomes, "Analysis of short blocklength codes for secrecy," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 255, pp. 1–15, Oct. 2018. [Online]. Available: https://doi.org/10.1186/s13638-018-1276-1

[30] E. Barker, "Guideline for using cryptographic standards in the federal government: Cryptographic mechanisms," National Institute of Standards and Technology (NIST), Tech. Rep., Aug. 2016. [Online]. Available: http://dx.doi.org/10.6028/NIST.SP.800-175B

[31] N. D. Jorstad and L. T. S. Jr., "Cryptographic algorithm metrics," Institute for Defense Analyses (IDA): Science and Technology Division, Tech. Rep., Jan. 1997.

[32] *Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems - Amendment for Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands*, IEEE Std. 802.16e, 2005.