Partial Encryption after Encoding for Security and Reliability in Data Systems

Preprint · February 2022

CITATIONS
0

READS
57

4 authors, including:

Alejandro Cohen
Technion - Israel Institute of Technology
74 PUBLICATIONS **385** CITATIONS
SEE PROFILE

Rafael Gregorio Lucas D'Oliveira
Clemson University
60 PUBLICATIONS **3,370** CITATIONS
SEE PROFILE

Muriel Médard
Massachusetts Institute of Technology
879 PUBLICATIONS **32,981** CITATIONS
SEE PROFILE

Partial Encryption after Encoding for Security and Reliability in Data Systems

Alejandro Cohen*, Rafael G. L. D'Oliveira†, Ken R. Duffy‡, and Muriel Médard†
*Faculty of Electrical and Computer Engineering, Technion, Israel, Email: alecohen@technion.ac.il
†RLE, Massachusetts Institute of Technology, USA, Emails: {rafaeld, medard}@mit.edu
‡Hamilton Institute, Maynooth University, Ireland, Email: ken.duffy@mu.ie

***Abstract*—We consider the problem of secure and reliable communication over a noisy multipath network. Previous work considering a noiseless version of our problem proposed a hybrid universal network coding cryptosystem (HUNCC). By combining an information-theoretically secure encoder together with partial encryption, HUNCC is able to obtain security guarantees, even in the presence of an all-observing eavesdropper. In this paper, we propose a version of HUNCC for noisy channels (N-HUNCC). This modification requires four main novelties. First, we present a network coding construction which is jointly, individually secure and error-correcting. Second, we introduce a new security definition which is a computational analogue of individual security, which we call individual indistinguishability under chosen ciphertext attack (individual IND-CCA1), and show that N-HUNCC satisfies it. Third, we present a noise based decoder for N-HUNCC, which permits the decoding of the encoded-then-encrypted data. Finally, we discuss how to select parameters for N-HUNCC and its error-correcting capabilities.**

I. Introduction

We consider the problem of secure and reliable communication over a noisy multipath network. A transmitter, Alice, wishes to transmit confidential messages to a legitimate receiver, Bob, over multiple noisy communication links, in the presence of two types of possible eavesdropper's, Eve. Weak Eve can obtain noiseless information transmitted over a subset of the paths, while strong Eve can obtain information over all paths, as illustrated in Figure 1. While Bob's links are noisy, we assume Eve may obtain noiseless observation of Alice transmissions. Thus, in contrast to the techniques utilized in physical layer security [1], we do not rely on the noise in the network for any security purposes.

In [2], a noiseless version of our setting was studied. In that setting, the authors proposed a hybrid universal network coding cryptosystem (HUNCC) which combines information-theoretic security with a computationally secure cryptosystem (see [3], [4] for a comparison with other approaches). HUNCC works by first premixing the data using a particular type of secure network coding scheme [5] and then partially encrypting the mixed data before transmitting it across the (noiseless) untrusted multipath network. Thus, obtaining information-theoretic security against a weak eavesdropper which does not observe all communication links while still guaranteeing computational security against a strong eavesdropper which observes all communication.

In this paper, we introduce a variation of HUNCC for noisy channels (N-HUNCC), i.e., we modify HUNCC so that it can be applied to noisy communication links. This requires four main novelties. First, we present a secure network coding scheme, which extends [5] into a code with both security and error-correcting capabilities against weak Eve. We note that this construction is of independent interest to those working on network information-theoretic security. Second, against a strong Eve, we introduce a new stronger notion of security than the individual computational security proposed in [2], which we call individual indistinguishability under chosen ciphertext attack (Individual IND-CCA1). This stronger notion can also be readily applied to the settings in [2], [6]. Third, we provide a novel joint decryption-decoding scheme that combine error correction using an efficient Guessing Random Additive Noise Decoding (GRAND) [7], with decryption in an intermediate stage of GRAND decoding algorithm as illustrated in Figure 2. Finally, we discuss how to select parameters for N-HUNCC and its error-correcting capabilities.

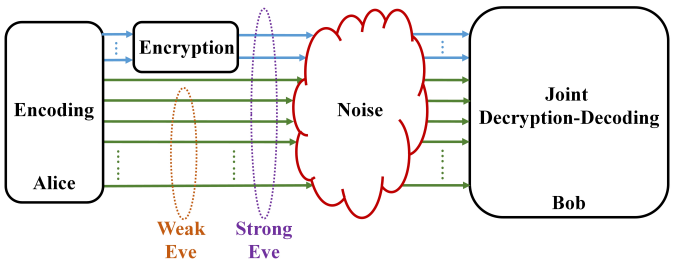

Fig. 1: Secure reliable communication over noisy multipath network with l paths, one source, Alice, one legitimate destination, Bob, and two types of possible eavesdropper's, Eve, weak and strong, which can obtain the noiseless information transmitted over $w < l$ or all the l paths, respectively.

The structure of this work is as follows. In Section II, we describe the system model. The security notations we use in this work are defined in Section III. The proposed secure and reliable N-HUNCC scheme with our main results are presented in Section IV. Section V describes the partial encryption scheme against strong Eve with the security analysis. The joint decryption encoded data is described in Section VI. In Appendix A, we present the construction of the individual secure scheme. Finally, we conclude this work in Section VII.

II. Setting

We consider a setting where a transmitter, Alice, wishes to transmit confidential messages $\mathbf{M} = [M_1, \ldots, M_{k_u}] \in \mathbb{F}_q^{k_u}$ to a legitimate receiver, Bob, over ℓ communication links, in the presence of an eavesdropper, Eve. While Bob's links are noisy, we assume Eve may obtain noiseless observation of

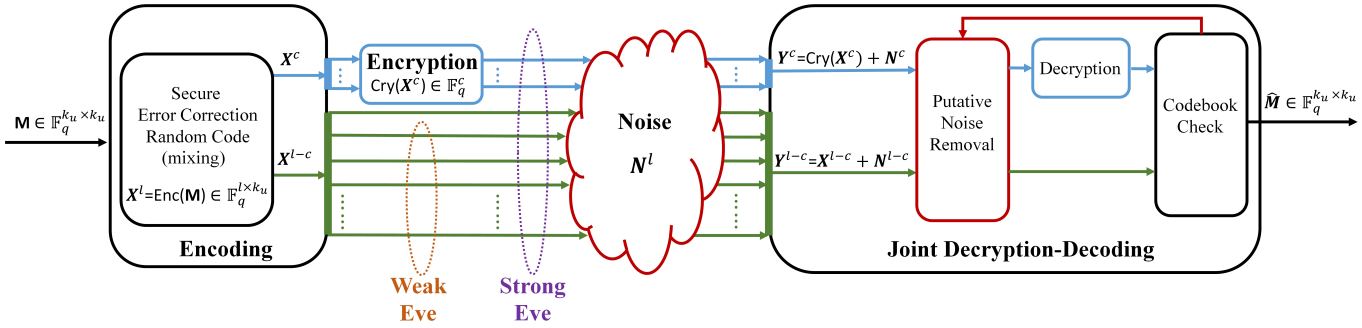

Fig. 2: Joint secure-reliable coding cryptosystem with partial encryption after encoding.

Alice transmissions. For the noisy channel at Bob, we consider the independent binary symmetric channel (BSC) with a bit flip probability of $p < \frac{1}{2}$.

Denote by $\mathbf{Y} = [Y_1, \ldots, Y_\ell]$ the vector of encoded messages (packets) that Alice transmits to Bob. We consider two types of Eve. A weak Eve, which only observes a subset, $w < \ell$, of the packets sent through the network, but which is computationally unbounded. And a strong Eve, which observes all such packets, but is computationally bounded. We denote these observations by $\mathbf{Z}_{E_w}$ and $\mathbf{Z}_{E_s}$, respectively.

III. Security Definitions

In this section, we define the security notations and guarantees we consider against both types of Eve, weak and strong.

A. Security Against a Weak Eve

Against a computationally unbounded weak Eve we use the notion of individual security.

Definition 1 (Individual Security)**.** Let $\mathbf{M} = [M_1, \ldots, M_{k_u}] \in \mathbb{F}_q^{k_u}$ be the confidential messages Alice wishes for Bob to receive, and $\mathbf{Y} = [Y_1, \ldots, Y_\ell]$ be the encoded messages. We say the encoding is (ℓ, w)-individually secure if for every $\omega \subset \{1, \ldots, \ell\}$ such that $|\omega| = w$, it holds that $H(M_i|\mathbf{Z}_\omega) = H(M_i)$, for every $i \in \{1, \ldots, \ell\}$, where $\mathbf{Z}_\omega = [Y_i]_{i \in \omega}$.

Thus, if Eve observes at most w packets, she is unable to learn anything (in a statistical sense) about any individual message $M_1, \ldots, M_{k_u} \in \mathbb{F}_q^{k_u}$. Since this notion is information-theoretic, it is independent of Eve's computational power. This notion was introduced in [8] to increase the efficiency of secure communication system in terms of data rates. Individual security has recently been considered for many applications due to the high efficiency it allows [9]–[15].

B. Security Against a Strong Eve

Against a strong Eve which observes all communications but is computationally bounded we use the notion of indistinguishability under chosen ciphertext attack (IND-CCA1). We start by giving the definition of a public-key cryptosystem.

Definition 2. A public-key cryptosystem consists of three algorithms:

- A key generation algorithm $\mathrm{Gen}(\kappa)$ which takes as input a security parameter κ and generates a public key p_k and a secret key s_k.
- An encryption algorithm $\mathrm{Enc}(m, p_k)$ which takes as input a message m belonging to some set of messages $\mathcal{M}$ and the public key p_k and then outputs a ciphertext c belonging to some set of ciphertexts $\mathcal{C}$.
- A polynomial time decryption algorithm $\mathrm{Dec}(c, s_k)$ which takes as input a ciphertext $c = \mathrm{Enc}(m, p_k)$ and the secret key s_k and outputs the original message m.

The encryption algorithm may be probabilistic and indeed must, in order to satisfy the following security constraint.

Definition 3 (IND-CCA1)**.** Indistinguishability under chosen ciphertext attack (IND-CCA1) is defined by the following game between an adversary and a challenger.

1) The challenger generates a key pair $\mathrm{Gen}(\kappa) = (p_k, s_k)$ for some security parameter κ and shares the public key p_k with the adversary.
2) The adversary may send a polynomial amount of ciphertexts to the challenger and receive back their decryptions. They may also perform a polynomial amount of operations.
3) The adversary chooses two challenge messages m_1^* and m_2^*, and sends them to the challenger.
4) The challenger chooses $i \in \{1, 2\}$ uniformly at random.
5) The challenger sends the challenge ciphertext $c^* = \mathrm{Enc}(m_i, p_k)$ to the adversary.
6) The adversary performs a polynomial amount of operations before outputting a guess for whether $b = 1$ or $b = 2$ and sends it to the challenger. If the adversary guesses correctly, they win.

The cryptosystem is indistinguishability under chosen ciphertext attack if any adversary has only a negligible advantage over a uniformly random guess of b, i.e. if they win the game with probability $\frac{1}{2} + \varepsilon(\kappa)$, where $\varepsilon(\kappa)$ is such that for every positive integer d, there exists an integer κ_d such that for all $\kappa > \kappa_d$, it holds that $\varepsilon(\kappa) < \frac{1}{\kappa^d}$. Such a function $\varepsilon(\kappa)$ is called a *negligible function*.

We now present a new notion of security which combines Definitions 1 and 3.

Definition 4 (Individual IND-CCA1)**.** Let the set of messages be $\mathcal{M} = \mathbb{F}_q^{k_u}$. Thus, each message $m = (m_1, \ldots, m_{k_u})$. We refer to each m_i as an individual message. Then, individual indistinguishability under chosen ciphertext attack (Individual IND-CCA1) is defined by the following game between an adversary and a challenger.

1) The challenger generates a key pair $\mathrm{Gen}(\kappa) = (p_k, s_k)$ for some security parameter κ and shares the public key p_k with the adversary.
2) The adversary may send a polynomial amount of ciphertexts to the challenger and receive back their decryptions. They may also perform a polynomial amount of operations.
3) The adversary chooses an index $j^* \in \{1, \ldots, k_u\}$ and two challenge individual messages $m_{j^*}^1$ and $m_{j^*}^2$, and sends them to the challenger.
4) The challenger chooses $i \in \{1, 2\}$ uniformly at random.
5) The challenger chooses $k_u - 1$ individual messages m_j, for $j \in \{1, \ldots, k_u\} - \{j^*\}$ uniformly at random and then constructs the message $m = (m_1, \ldots, m_{k_u})$, where $m_{j_*} = m_{j^*}^i$.
6) The challenger sends the challenge ciphertext $c^* = \mathrm{Enc}(m, p_k)$ to the adversary.
7) The adversary performs a polynomial amount of operations before outputting a guess for whether $b = 1$ or $b = 2$ and sends it to the challenger. If the adversary guesses correctly, they win.

The cryptosystem is individually indistinguishable under chosen ciphertext attack if any adversary has only a negligible advantage over a uniformly random guess of b, i.e. if they win with probability $\frac{1}{2} + \varepsilon(\kappa)$, where $\varepsilon(\kappa)$ is a negligible function.

Individual IND-CCA1 is thus a computational analogue of individual security. It guarantees that an adversary can only learn a negligible amount of information about any individual message. It is thus suited for the same settings where individual security can be applied, but where Eve might observe all communication, rendering individual security useless.

IV. Joint Secure-Reliable Coding Cryptosystem (Main Results)

In this section, we present our scheme N-HUNCC and our main results. In Figure 2, we illustrate how the scheme operates on a noisy multipath network with ℓ communication links. The noise we consider at Bob is, for each link, an independent BSC with a sum rate, for each channel transmission, of at most $\nu = \sum_{i=1}^{\ell} H(p)$. We note, however, that we allow Eve to obtain noiseless observations of the transmissions over the links, i.e., no noise in the channel is used for security purposes.

N-HUNCC follows the main ideas proposed in [2] in which one can encrypt only part of the data transmitted. However, unlike [2], here we assume that Bob may obtain noise observation of the data transmitted. We present four main novelties: 1) a new random code design with error correction, 2) that N-HUNCC is individually IND-CCA1 secure, 3) a novel joint decryption-decoding scheme, 4) a discussion on parameter selection and error-correcting capabilities. We detail those in Appendix A, and Sections V and VI, respectively. We note that the individual IND-CCA1 security proof can be readily applied to the settings in [2], [6]. We now provide a high level description of our proposed scheme.

We start by looking at the encoding process at Alice. The messages are encoded using an (ℓ, w)-individual secure random code as given in Appendix A. The number of individually secret messages against weak Eve is given by $k_s \leq \ell - \nu - w - 2k_u\varepsilon$. In the encoding process, using a random code, Alice encodes each of the i-th columns in the messages matrix $\mathbf{M}$ independently. Thus the encoder is given by

$$E : \mathbf{M}(i) \in \mathbb{F}_q^{k_u} \to \mathbf{X}(i) \in \mathbb{F}_q^{\ell},$$

which maps the i-th column $\mathbf{M}(i)$ in the massage matrix to the i-th column $\mathbf{X}(i)$ in the codeword matrix. Then, each row in the codeword matrix is transmitted in the ℓ independent links.

Using this secure coding scheme against a weak Eve, our first main result is the following achievability theorem.

Theorem 1. *Assume a noisy BSC multipath communication (ℓ, ν, w). N-HUNCC's encoder delivers, with high probability, $k_u \leq \ell - \nu - \varepsilon$ messages at the legitimate decoder, while keeping a weak eavesdropper which observes $w < \ell - \nu$ noiseless links ignorant with respect to any set of $k_s \leq k_u - w - 2k_u\varepsilon$ messages individually, such that $I(\boldsymbol{M}^{k_s}; \boldsymbol{Z}^w) \leq \varepsilon_\ell$, whenever $\nu \leq \sum_{i=1}^{\ell} H(p)$ and $k_u\varepsilon = o(k_u)$.*

The construction of the individual secure network code follows almost directly from [5, Section IV] considering carefully the increased size of the codewords required due to channel noise. Due to the space limitation the construction together with the proofs of reliability and individual secrecy are deferred to [16].

Now, in order to obtain security against a strong Eve, as opposed to the traditional approach where all communication links must be encrypted, we show that we only need to encrypt a portion $c = \ell - w$ of the communication links, where $\ell = k_u + \nu + w + 2k_u\varepsilon$. Without loss of generality, we let the links indexed by $1, \ldots, c$ to be the encrypted ones. The encryption by the cryptosystem at each i-th column of the c links is given by

$$\mathrm{Crypt}_1 : \mathbf{X}(i) \in \mathbb{F}_q^{\ell - w} \to \mathbf{Y}(i) \in \mathbb{F}_q^{\ell - w + r}, \tag{1}$$

which by adding an extra r symbols (depending on the encryption used) encrypts the i-th column $\mathbf{X}(i)$ in the codeword matrix into row $\mathbf{Y}(i)$ in the matrix obtained at Bob. The extra r bits at the outcome of the cryptosystem per column are concatenated to be transmitted over the c encrypted links. We note that, because we encrypt after encoding, the encryption can be performed at any stage in the system as long as it occurs before a strong Eve's observation. This is opposed to traditional schemes which require the encryption to essentially be done at Alice, or some equivalent of her, since all messages are needed for the error correction encoding which is usually applied after encryption.

Our next main result shows that N-HUNCC is secure against a computationally bounded strong eavesdropper.

Theorem 2. *In the setting of Theorem 1, let* Crypt_1 *be a IND-CCA1 secure cryptosystem used as described in* (1)*. Then, N-HUNCC is individually IND-CCA1 secure.*

The proof of the theorem is given in Section V.

At Bob, the joint decryption-decoding scheme is given by

$$D : [\mathbf{Y} \in \mathbb{F}_q^{c+r\times k_u}; \quad \mathbf{Y} \in \mathbb{F}_q^{\ell-c\times k_u}] \to \hat{\mathbf{M}} \in \mathbb{F}_q^{k_u\times k_u},$$

which maps the outcome noisy channel $\mathbf{Y}$ to $\hat{\mathbf{M}}$.

To decode the messages, Bob utilizes a modified version of the GRAND decoder [7]. Given the noisy channel outcome $\mathbf{Y}$, Bob orders the noise sequences from most likely to least likely. He then goes through the list, subtracting the noise from $\mathbf{Y}$ and then preforming the decryption on the first c rows. After this, he checks if all columns of the decrypted matrix are elements of the codebook. The first time this occurs Bob decodes that message. As shown in [7], this procedure is a Maximum Likelihood (ML) decoder. We discuss this decoder more in detail in Section VI.

We now discuss the parameter selection for N-HUNCC in order to be reliable on a BSC with error probability p. The main challenge is to deal with how the encryption affects the error correction capabilities of the code. If the output of the encryption is uniformly random, we have the following result.

Theorem 3. *In the setting of Theorem 2, suppose the output of* Crypt_1 *is uniformly distributed. Then, if* $\frac{k_u+r_0}{\ell+r}$ *is less than the capacity of the BSC channel, N-HUNCC can asymptotically transmit at arbitrarily low probability of errors at a rate of* $\frac{k_u}{\ell+r}$*. Here,* r_0 *denotes the amount of randomness in* Crypt_1*.*

The proof of the theorem is given in Section VI.

In practice it might be hard to enforce a uniform output on the encryption. However, there are cryptosystems with outputs which are computationally indistinguishable from a uniform distribution [17], [18]. Thus, we expect using such cryptosystems as an inner crypto function Crypt_1 should make N-HUNCC perform as described in Theorem 3. Indeed, most, if not all (to the best of our knowledge), practical implementations of random codes are actually pseudorandom.

We finish this section with the following remarks.

Remark 1. The security guarantee in Theorem 2 can be readily applied to noise-less HUNCC in order to obtain individual IND-CCA1 security for the settings in [2], [6], for example.

Remark 2. The arguments in Theorem 3 can be readily applied to scenarios in which one encodes data for reliability before using any encryption with an output which is indistinguishable from uniformly random. The Advanced Encryption Standard (AES), for example, consistently passes output uniformity tests [19], [20].

V. Partial Encryption Against a Strong Eve

In this section we provide the proposed partial encryption scheme on $\ell - w$ links using a cryptosystem after encoding as given in Appendix A. The encryption process and the security analysis against a strong Eve are detailed in Subsections V-A and V-B, respectively.

A. Encypting Encoded Data

The encoding construction in N-HUNCC, (see Appendix A) , starts with an encoding function $\mathrm{Enc} : \mathbb{F}_q^{k_u} \to \mathbb{F}_q^{\ell}$. To satisfy the Individual IND-CCA1 security as given in Definition 4, by only encrypting $\ell - w$ links, we consider an IND-CCA1 cryptosystem. Furthermore, for reliability, discussed in Section VI, we consider such cryptosystems with pseudorandom output, e.g., as given in [17].

In order to satisfy IND-CCA1 security, the cryptosystem adds randomness (possibly through some form of padding) of size r_0 and ultimately increases the output by $r \geq r_0$. This encryption is given by the deterministic injective function $\mathrm{Crypt}_1 : \mathbb{F}_q^{\ell+r_0-w} \to \mathbb{F}_q^{\ell-w+r}$, where $r_0 \leq r$ represents the randomized part of the cryptosystem algorithm and is therefore, not shared with Bob. We may also refer (by abuse of notation) to the cryptosystem as the probabilistic function $\mathrm{Crypt}_1 : \mathbb{F}_q^{\ell-w} \to \mathbb{F}_q^{\ell-w+r}$ where the randomness r_0 is implicit in the function.

N-HUNCC, then, consists in the cryptographic scheme

$$\mathrm{Crypt}_2 = (\mathrm{Crypt}_1 \circ \pi_1 \circ \mathrm{Enc}) \times (\pi_2 \circ \mathrm{Enc}) : \mathbb{F}_q^{k_u+r_0} \to \mathbb{F}_q^{\ell+r},$$

where $\pi_1 : \mathbb{F}_q^{\ell+r_0} \to \mathbb{F}_q^{\ell+r_0-w}$ is the projection of the first $\ell + r_0 - w$ entries and $\pi_2 : \mathbb{F}_q^{\ell} \to \mathbb{F}_q^{w}$ is the projection of the last w entries.[1] In other words, N-HUNCC consists in taking the message with k_u symbols, encoding it with the individual code described in [16, Appendix A] to obtain ℓ symbols. It then inputs the first $\ell - w$ of these symbols into Crypt_1 together with r_0 randomly chosen symbols needed for the cryptosystem, obtaining $\ell - w + r$ encrypted symbols. The output of Crypt_2 is then the concatenation of the $\ell - w + r$ encrypted symbols with the w unencrypted ones.

B. Security Against A Strong Eve (A Proof For Theorem 2)

Here we show that Crypt_2 and thus N-HUNCC is individually IND-CCA1 secure as given in Definition 4.

For each i-th column in the massage matrix $\mathbf{M}$, the encoding function of the individual secure random code is given by

$$\mathrm{Enc}(M_1(i), \ldots, M_{k_s}(i), M_{k_s+1}(i), \ldots, M_{k_u}(i)),$$

where the first k_s coordinates determine the bin $b(i)$ in the codebook and the last $k_u - k_s$ determine the position within the bin $e(i)$ the codeword is selected from, as described in [16, Appendix A]. We now play the security game described in Definition 4. We assume that the adversary chooses a $j \in [1, k_s]$ (the other case will follow analogously but for position instead of bin). The adversary then chooses two messages $M_j^1(i)$ and $M_j^2(i)$.

We show the stronger statement that even if we give the adversary the other entries $M_1(i), \ldots, M_{k_s}(i)$, he is not able to distinguish between the two bins

[1]The Cartesian product of two functions is $(f \times g)(x, y) = (f(x), g(x))$.

$b_1(i) = (M_1(i), \ldots, M_i^1(i), \ldots, M_{k_s}(i))$ and $b_2(i) = (M_1(i), \ldots, M_i^2(i), \ldots, M_{k_s}(i))$. The challenger still chooses $e(i) = (M_{k_s+1}(i), \ldots, M_{k_u}(i))$ uniformly at random and does not share it with the adversary. The challenger then chooses $j \in \{1, 2\}$ uniformly at random and sends back the cyphertext $c_j = \mathrm{Crypt}_2(b_j(i), e(i))$ to the adversary. Both crypto bins $b_1(i)$ and $b_2(i)$ have a total of $2^{w+k_u\varepsilon}$ codewords each of size $\ell+r$ as shown in [16], of which the first $\ell+r-w$ symbols are encrypted by Crypt_2 and the last w symbols are not. From knowing the last w symbols of c_j the adversary is able to reduce the number of the possible codewords in the bins to B_1 and B_2 respectively, with each cypher having probability $\frac{1}{B_1+B_2}$. As shown in [5, Section IV.B], the probability that the actual number of the possible codewords in the bins deviates from the average by more than ε' is bounded for sufficient large k_u, so that, for any bin $b_j, j \in \{1, \ldots, 2^{k_u-w}\}$, it holds that $(1-\varepsilon')2^{k_u\varepsilon} \leq B_j \leq (1+\varepsilon')2^{k_u\varepsilon}$ with high probability.

The last step now consists on the adversary trying to distinguish the messages via the information leaked by the encrypted part. The best possible case for the adversary (and worse for the challenger) is the case where bin 1 has as high probability as possible, which corresponds to $B_1 \geq B_2$ with as much difference as possible, and such that the distinguishability from the Crypt_1 is as high as possible. Thus, we consider the case where all the cyphertexts in bin 1 have probability $p_c + \varepsilon_c$ and the cyphertexts in bin 2 have probability $p_c - \varepsilon_c$, thus aligning the probabilities so that they concentrate in bin 1. Since $B_1(p_c + \varepsilon_c) + B_2(p_c - \varepsilon_c) = 1$, it follows that

$$p_c = \frac{1 - (B_1 - B_2)\varepsilon_c}{B_1 + B_2} \tag{2}$$

From the IND-CCA1 security of Crypt_1 it follows that for every $d \in \mathbb{N}$, there exists an k_d such that $k_u \geq k_d$ implies in

$$\frac{p_c + \varepsilon_c}{(p_c + \varepsilon_c) + (p_c - \varepsilon_c)} - \frac{1}{2} \leq \frac{1}{k^d}. \tag{3}$$

Hence, substituting (2) in (3) we have

$$\varepsilon_c \leq \frac{2}{k^d(B_1 + B_2) + 2(B_1 - B_2)} \tag{4}$$

Now we show that the difference between the probability of the correct bin being bin 1 and $\frac{1}{2}$ is negligible. Indeed,

$$\begin{aligned} \Pr[\text{bin 1}] - \frac{1}{2} &= B_1(p_c + \varepsilon_c) \\ &\leq \frac{k^d(B_1 - B_2) + 2(B_1 + B_2)}{2k^d(B_1 + B_2) + 4(B_1 - B_2)}. \end{aligned} \tag{5}$$

We select the bins with the highest deviation possible of codewords in the bins, as described above, to analyze the worst case. Thus, for $\varepsilon' = \frac{1}{k_u^t}$ and any $t \geq 2$, we have

$$\frac{B_1}{B_2} = \frac{1 + \varepsilon'}{1 - \varepsilon'} = \frac{k_u^t + 1}{k_u^t - 1}. \tag{6}$$

Hence, substituting (6) in (5) we obtain

$$\Pr[\text{bin 1}] - \frac{1}{2} \leq \frac{1}{2k_u^t + \frac{4}{k^d}} + \frac{1}{k^d + \frac{2}{k_u^t}},$$

which for every d' can be made smaller than $\frac{1}{k^{d'}}$ for large enough k_u by choosing an appropriate d and taking t to grow more than a constant, e.g. $t = \log(k_u)$.

VI. Decrypting Encoded Data (A Proof For Theorem 3)

In this section, we analyze the proposed joint decryption-decoding scheme as presented in Section IV using GRAND [7]. GRAND algorithms operate by sequentially inverting putative noise effects from the demodulated received sequence and querying if what remains is in the code-book [7], [21]–[23]. If those noise effects are queried in order from most-likely to least likely, the first instance where a code-book member is found is an ML decoding [7]. If the code is unstructured and stored in a dictionary, a code-book query corresponds to a tree-search with a complexity that is logarithmic in the code-length. If the code is linear in any finite field, code-book membership can be determined by a matrix multiplication and comparison. For encrypted, encoded data, only one extra step is required: the effect of each putative noise sequence is removed from the encrypted data, which is then decrypted and the resulting sequence tested for code-book membership.

Our encoding construction starts with a random code $\mathcal{C}_0 = \mathrm{Enc}(\mathbb{F}_q^{k_u}) \subseteq \mathbb{F}_q^{\ell}$ of size q^{k_u}. When Considering the randomness from Crypt_1 we obtain a code $\mathcal{C}_1 = \mathcal{C}_0 \times \mathbb{F}_q^r \subseteq \mathbb{F}_q^{\ell+r_0}$ of size $q^{k_u+r_0}$. That is, every original message $m \in \mathbb{F}^{k_u}$ corresponds to q^{r_0} possible codewords in $\mathcal{C}_1$. Finally, after applying Crypt_2 we obtain a code $\mathcal{C}_2 \subseteq \mathbb{F}_q^{\ell+r}$ with the same size $q^{k_u+r_0}$ of $\mathcal{C}_1$. The last w symbols of a codeword of $\mathcal{C}_2$ are uniformly distributed. The first $\ell-w$ symbols are the output of Crypt_1. If this output is uniform, then we have a random code $\mathcal{C}_2 \subseteq \mathbb{F}_q^{\ell+r}$ of size $q^{k_u+r_0}$ with the property that multiple (more precisely q^{r_0}) codewords decode to the same message. Let us suppose that we want the code to act as a regular code, i.e. treating each codeword of the q^{r_0} same-message codewords as if they corresponded to distinct messages. In this worse scenario, the code would asymptotically transmit at arbitrarily low probability of errors if [7], [24]

$$\frac{k_u + r_0}{\ell + r} < \text{Capacity of the BSC channel.} \tag{7}$$

However, when setting the parameters to satisfy (7), the effective rate of the code is given by $\frac{k_u}{\ell+r}$.

Moreover, we can compute an error exponent for our joint decryption-decoding scheme. Consider a communication on one of the coded-packets, $Y(i)$, expressed as a binary string $Y^b(i)$. It is transmitted over a BSC and impacted by a binary noise effect $N^b(i)$ resulting in a received signal $N^b(i) = Y^b(i) + N^b(i)$, where addition is $\mathbb{F}_2$ and $N^b(i)$ is a string of independent Bernoulli p random variables. Let $\hat{Y^b}(i)$ be the GRAND-decoding estimate of $Y^b(i)$. If a code-book of rate $R \leq \frac{k_u+r_0}{\ell+r}$ is selected uniformly, then the likelihood of an erroneous decoding decays exponentially in $n = \ell + r$ with Gallager's error exponent [7][Proposition 1]. That is,

$$\lim_{n\to\infty} \frac{1}{n} \log P\left(\hat{Y^b}(i) \neq Y^b(i)\right) = -\varepsilon(R, p),$$

where the exact form for $\varepsilon(R,p)$ can be identified as follows. Define the Rényi entropy of the BSC noise process with bit flip probability p and parameter $\alpha \in (0,1) \cup (1,\infty)$ to be

$$H_\alpha = \frac{1}{1-\alpha} \log \left(p^\alpha + (1-p)^\alpha \right),$$

with H_1 being the Shannon entropy and min entropy denoted $H_{\min} = -\log(\max(p, 1-p))$. Then defining

$$\Lambda^N(\alpha) = \begin{cases} \alpha H_{1/(1+\alpha)} & \text{for } \alpha \in (-1,\infty) \\ -H_{\min} & \text{for } \alpha \leq -1, \end{cases}$$

$I^N(x) = \sup_\alpha(\alpha x - \Lambda^N(\alpha))$ and x^* such that $d/dx I^N(x)|_{x=x^*} = 1$, then the error exponent can be identified as

$$\varepsilon(R,p) = \begin{cases} 1 - R - H_{1/2} \text{ if } R \in (0, 1-x^*) \\ I^N(1-R) \text{ if } R \in [1-x^*, 1-H(p)), \end{cases}$$

and the error exponent for the full system follows from an application of the principle of the largest term [25, Lemma 1.2.15].

Lemma 1. *The error exponent for the likelihood that one or more of the decodings is erroneous is given by*

$$\lim_{n\to\infty} \frac{1}{n} \log P \left(\bigcup_{i=1}^{\ell} \left\{ \hat{Y^b}(i) \neq Y^b(i) \right\} \right) = -\varepsilon(R,p).$$

VII. Conclusions

In this work, we suggest a noisy hybrid universal network coding cryptosystem that can be applied to noisy communications systems. The proposed cryptosystem is secure against a strong eavesdropper under a new security notion we introduce of Individual IND-CCA1. This notion of security can be readily applied to other HUNCC solutions offered in the literature with partial encryption [2], [6]. Finally, we present a joint decryption-decoding scheme that combines error correction using GRAND with decryption in an intermediate stage.

References

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

[2] A. Cohen, R. G. L. D'Oliveira, S. Salamatian, and M. Médard, "Network coding-based post-quantum cryptography," *IEEE Journal on Selected Areas in Information Theory*, 2021.

[3] A. Kumar, C. Ottaviani, S. S. Gill, and R. Buyya, "Securing the future internet of things with post-quantum cryptography," *Security and Privacy*, p. e200, 2021.

[4] S. S. Gill, A. Kumar, H. Singh, M. Singh, K. Kaur, M. Usman, and R. Buyya, "Quantum computing: A taxonomy, systematic review and future directions," *Software: Practice and Experience*, vol. 52, no. 1, pp. 66–114, 2022.

[5] A. Cohen, A. Cohen, M. Médard, and O. Gurewitz, "Secure multi-source multicast," *IEEE Transactions on Communications*, vol. 67, no. 1, pp. 708–723, 2018.

[6] R. G. L. D'Oliveira, A. Cohen, J. Robinson, T. Stahlbuhk, and M. Médard, "Post-quantum security for ultra-reliable low-latency heterogeneous networks," in *MILCOM 2021-2021 IEEE Military Communications Conference (MILCOM)*. IEEE, 2021, pp. 933–938.

[7] K. R. Duffy, J. Li, and M. Médard, "Capacity-achieving guessing random additive noise decoding," *IEEE Trans. Inf. Theory*, vol. 65, no. 7, pp. 4023–4040, 2019.

[8] D. Kobayashi, H. Yamamoto, and T. Ogawa, "Secure multiplex coding attaining channel capacity in wiretap channels," *IEEE transactions on information theory*, vol. 59, no. 12, pp. 8131–8143, 2013.

[9] K. Bhattad and K. R. Narayanan, "Weakly secure network coding," *NetCod, Apr*, vol. 104, 2005.

[10] D. Silva and F. R. Kschischang, "Universal weakly secure network coding," in *Networking and Information Theory, 2009. ITW 2009. IEEE Information Theory Workshop on*. IEEE, 2009, pp. 281–285.

[11] A. S. Mansour, R. F. Schaefer, and H. Boche, "Secrecy measures for broadcast channels with receiver side information: Joint vs individual," in *Information Theory Workshop (ITW), 2014 IEEE*. IEEE, 2014, pp. 426–430.

[12] ——, "The individual secrecy capacity of degraded multi-receiver wiretap broadcast channels," in *2015 IEEE International Conference on Communications (ICC)*. IEEE, 2015, pp. 4181–4186.

[13] ——, "On the individual secrecy capacity regions of the general, degraded and gaussian multi-receiver wiretap broadcast channel," *IEEE Transactions on Information and Security, 2016*, vol. 11, no. 9, pp. 2107–2122, 2016.

[14] Y. Chen, O. O. Koyluoglu, and A. Sezgin, "On the individual secrecy rate region for the broadcast channel with an external eavesdropper," in *2015 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2015, pp. 1347–1351.

[15] K. Jiang, T. Jing, F. Zhang, Y. Huo, and Z. Li, "Zf-sic based individual secrecy in simo multiple access wiretap channel," *IEEE Access*, vol. 5, pp. 7244–7253, 2017.

[16] A. Cohen, R. G. L. D'Oliveira, K. R. Duffy, and M. Médard, "Partial encryption after encoding for security and reliability in data systems," *arXiv preprint*, 2022.

[17] B. Möller, "A public-key encryption scheme with pseudo-random ciphertexts," in *European Symposium on Research in Computer Security*. Springer, 2004, pp. 335–351.

[18] D. J. Bernstein, "Introduction to post-quantum cryptography," in *Post-quantum cryptography*. Springer, 2009, pp. 1–14.

[19] P. Hellekalek and S. Wegenkittl, "Empirical evidence concerning AES," *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, vol. 13, no. 4, pp. 322–333, 2003.

[20] P. L'ecuyer and R. Simard, "TestU01: AC library for empirical testing of random number generators," *ACM Transactions on Mathematical Software (TOMS)*, vol. 33, no. 4, pp. 1–40, 2007.

[21] A. Solomon, K. R. Duffy, and M. Médard, "Soft maximum likelihood decoding using GRAND," in *IEEE ICC*, 2020.

[22] K. R. Duffy, "Ordered reliability bits guessing random additive noise decoding," in *IEEE ICASSP*, 2021, pp. 8268–8272.

[23] K. R. Duffy, M. Médard, and W. An, "Guessing random additive noise decoding with symbol reliability information (SRGRAND)," *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 3–18, 2022.

[24] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[25] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*. Springer, 2009.

Appendix A
Secure Random Codes Against a Weak Eve (A Proof For Theorem 1)

Here we provide the individual secure binary random code used at Alice to encode the message matrix $\mathbf{M}$. We propose a secure code for weak Eve which can observe the information of at most $w = \ell - \nu - k_s - 2k_u\varepsilon$ links. That is, we assume a degraded channel at weak Eve, with $p(y,z|x) = p(y|x)p(z|y)$. The individual security is obtained for $k_s \leq k_u - w - 2k_u\varepsilon$ messages form the k_u messages decoded at Bob correctly with high probability for $\nu \leq \sum_{i=1}^{\ell} H(p)$. We may now turn to the detailed construction and analysis.

a) Codebook Generation: Set $\Delta = 2^{w+k_u\varepsilon}$ and $k_u = \ell - \nu - \varepsilon$. Let $P(x) \sim Bernoulli(1/2)$. Using a distribution $P(X^\ell) = \prod_{j=1}^{\ell} P(x_j)$, for each possible column $M_1(i); \ldots; M_{k_s}(i)$ in the message matrix, that is, $2^{k_u-(w+k_u\varepsilon)}$

possibilities, generate Δ independent and identically distributed codewords $x^\ell(e)$, $1 \leq e \leq \Delta$. Thus, we have $2^{k_u-(w+k_u\varepsilon)}$ bins, each of size $2^{w+k_u\varepsilon}$. Note that the length of the columns in the bins is ℓ, thus the codebook matrix is increased by $\nu+\varepsilon$ compared to the message matrix $\mathbf{M}$.

b) Encoding: The encoder selects, for each column i of bits $M_1(i);\ldots;M_{k_s}(i)$, one codeword, $x^\ell(e(i))$, from the bin indexed by $M_1(i);\ldots;M_{k_s}(i)$, where $e(i) = M_{k_s+1}(i);\ldots;M_{k_u}(i)$. That is, $k_s = k_u - (w + k_u\varepsilon)$ bits of the column choose the bin, and the remaining $w-\varepsilon$ bits choose the codeword within the bin.

c) Reliability: Bob maps $\mathbf{Y}_s$ back to $\mathbf{M}_s$, as per column $1 \leq i \leq c$, the index of the bin in which the codeword $\mathbf{Y}_s$(i) resides, using for example GRAND decoder for noise ν [7], is $M_1(i);\ldots;M_{k_s}(i)$ and the index of the codeword location in that bin is $M_{k_s+1}(i);\ldots;M_{k_u}(i)$. The analysis on the probability of successfully decoding $\mathbf{M}(i)$ from $\mathbf{Y}(i)$ is a direct consequence using standard analysis of random coding [1, Section 3.4].

As for the information leakage at the weak eavesdropper, to show that k_s-individual security constraint is met, the proof is follows directly form [5, Section IV.B]. That is, for each column $i \in \{1,\ldots,k_u\}$, as long we choose $k_u\varepsilon$ to be an integer, $I(\mathbf{M}^{k_s}(i);\mathbf{Z}^w(i)) = O(k_u^{-t+1})$ for any $t \geq 2$.