CVQKD Reconciliation Based on Distributional Transform and Slepian-Wolf Coding

Rávilla R. S. Leite and Francisco M. de Assis

***Abstract*— In this article, we present an information reconciliation protocol designed for CVQKD based on the Distributional Transform for variable quantization and LDPC Slepian-Wolf coding. The decoding employed a modified Bit-Flipping algorithm to recover the error vector $\boldsymbol{E}$ between the correlated binary sequences of Alice and Bob. Numerical results were obtained by applying an LDPC code with a rate of $1/2$, in conjunction with Bit-Flipping and Sum-Product decoding. The results encourage the application of the technique using LDPC codes optimized for CVQKD applications.**

***Keywords*— CVQKD, LDPC Codes, Distributional Transform, Bit-Flipping.**

I. Introduction

Quantum Key Distribution (QKD), initially proposed by Bennett and Brassard [1], aims to enable the sharing of secret information between two distant and legitimate parties, Alice and Bob, by providing a key with unconditional security against an eavesdropper, Eve, relying on fundamental principles of quantum physics, especially the uncertainty principle and the no-cloning theorem [2], [3].

The QKD protocol can be implemented in two main ways: DVQKD (Discrete-Variable QKD) and CVQKD (Continuous-Variable QKD). In DVQKD, the key information is encoded through the phase or polarization of single photons [1], [4], [5] and single-photon detection is performed on the receiver side. In contrast, CVQKD encodes the key information in the quadratures of the electromagnetic field of non-orthogonal coherent states, and the signals measured by Bob are continuous [6], [7], [8], [9]. CVQKD offers better prospects for practical implementation since it requires standard optical communication technology instead of single-photon detection and provides higher secret key rates than DVQKD [8], [3], and can also utilize continuous (Gaussian) or discrete modulations.

Gaussian modulation protocols occur in four main stages [10], [8]: (1) state distribution and measurement; (2) parameter estimation; (3) information reconciliation; and (4) privacy amplification. This work focuses on the information reconciliation stage, where the quantization of the continuous variables remaining after discarding the values measured by Bob with the wrong bases and error correction occurs through an authenticated classical channel, assumed to be error-free. After quantization, each generated subchannel can be separately treated with multi-level coding (MLC), applying LDPC codes according to the channel capacity to perform error correction close to Shannon's capacity [11], [12], [3].

The final secret key rate (SKR) can be written as $\Delta I = \beta I_{AB} - \chi_{BE}$, where β is the reconciliation efficiency, which depends on the quantizantion and error correcting efficiency, I_{AB} is the the classical mutual information between Alice and Bob, and χ_{BE} is the Holevo information between Bob and Eve, i.e. the maximum amount of information about the raw key accessible by Eve [3]. Based on these concepts, it is possible to see how reconciliation efficiency affects the secret key rate at the end of the protocol.

This paper is structured as follows: Section II presents the fundamental concepts of the quantization technique used and some parameters obtained through simulations. Section III describes the application of Slepian-Wolf coding and Bit-Flipping-based decoding to the first two channels obtained from the quantization of the raw key's continuous variables. In Section IV, the results of Bit-Flipping and Belief-Propagation decodings for the sequences of the first quantization channel, encoded by the same LDPC matrix in a reverse reconciliation scheme, are compared. Finally, Section V presents the conclusions and future work.

II. Quantization of Variables by the Distributional Transform

After the quantum stage, Alice and Bob communicate through an authenticated classical channel, assumed to be error-free, to perform error correction and privacy amplification. The application of error-correcting codes only takes place after the quantization of the continuous values of the raw keys, which are those resulting after discarding the values that Bob measured with the wrong bases (sifting phase) from both sequences. Thus, the raw keys of Alice and Bob are two correlated Gaussian sequences, $\boldsymbol{A} = A_1, \cdots, A_n$ and $\boldsymbol{B} = B_1, \cdots, B_n$, respectively, with mutual information greater than 0, i. e., $I(A, B) > 0$ [13], [14]. For the continuous values of the raw key, N realizations of $A \sim \mathcal{N}(0, 1)$ were generated, corresponding to Alices's modulated states, as well as N realizations of $Z \sim \mathcal{N}(0, \sigma_Z^2)$, so that $B = A + Z$, since a quantum channel with additive Gaussian noise was assumed [15], [11].

Although the SEC (Slice Error Correction) [13], [12] or MD (Multidimensional) Reconciliation [16], [17], [15] are more commonly used to extract bit sequences from continuous valued data, in this work, the quantization of the raw key values was performed using the protocol proposed by Araújo, Assis and Albert in [18] based on an important *Lemma* from arithmetic coding [19]:

Rávilla R. S. Leite and Francisco M. de Assis, Department of Electrical Engineering, Federal University of Campina Grande (UFCG), Campina Grande-PB, Brazil. E-mails: ravilla.leite@ee.ufcg.edu.br, fmarcos@dee.ufcg.edu.br.
This work was partially financed by CNPq. (311680/2022-4 and 140827/2022-6)

Lemma 1: Let V be a random variable with a continuous distribution function F_V, define $U = F_V(V)$. So, U is uniformly distributed on $[0, 1]$.

This Lemma is known in Copula Theory as Distributional Transform and allows to transform a random variable by its cumulative distribution function, leading to a uniform distribution on the unit interval [11]. With this assumption and consideering that the bits in the binary expansion of a random variable with uniform distribution on $[0, 1]$ are independent and Bernoulli $(\frac{1}{2})$, Alice can calculate $X = \Phi(A) \sim$ unif$[0, 1]$, and similar for Bob, $Y = \Phi(B) \sim$ unif$[0, 1]$ to map the raw key elements on the unit interval, and after apply the binary expansion to the resulting values with ℓ bits of resolution. Dias and Assis called this technique Distributional Transform Expansion, DTE [11], and each resolution bit induces a BSC channel of the type $\mathcal{D}_j(X_1), ..., \mathcal{D}_j(X_n)$ and $\mathcal{D}_j(Y_1), ..., \mathcal{D}_j(Y_n)$, for $j = 1, \ldots, \ell$.

Let a be a realization of the Gaussian random variable $\boldsymbol{A}$ and $x = \Phi(a) \in [0, 1]$ the cumulative probability of a, a realization of $\boldsymbol{X}$, the first digit in the binary expansion $(\mathcal{D}_1(X_i))$ announces if $x < \frac{1}{2}$ or $x \geq \frac{1}{2}$, or in terms of the continuous values, if $a < 0$ or $a \geq 0$. This threshold can be observed in Figure 1 (in terms of the continuous value of a) and in Figure 2 (in terms of CDF) named as "Threshold 1". To get the value of the second bit $\mathcal{D}_2(X_i)$, the intervals referring to $\mathcal{D}_1(X_i)$ must be divided into two parts, generating four subintervals in the unit interval, as illustrated in Figures 1 and 2 as "Threshold 2" and "3". Looking at just the first half of the graph, we have that if $\Phi(A_i) \in [0, 1/4)$, $\mathcal{D}_2(X_i) = 0$ and if $\Phi(A_i) \in [1/4, 1/2)$, $\mathcal{D}_2(X_i) = 1$. Similarly for the second half of the unit interval.

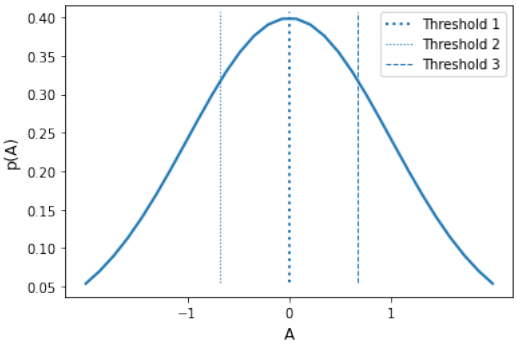

Fig. 1. Probability density function of the Normal Distribution. *Threshold 1* corresponds to $\Phi^{-1}(1/4)$ and *Threshold 2* to $\Phi^{-1}(3/4)$.

For the following channels, the intervals mentioned above must be subdivided by 2 for each bit of resolution you want to add, resulting in 2^ℓ subintervals. In this way, the decision regions become narrower and narrower, making the probability of inversion in the less significant channels greater. In [18] the probability of error for the first 6 quantization channels as a function of SNR is shown.

We quantized 1920 realizations of $\boldsymbol{X}$ and $\boldsymbol{Y}$, with 3-bit resolution, and calculated the capacities of the induced sub-channels, based on the probability of error between $\boldsymbol{X}$ and $\boldsymbol{Y}$, as shown in Table I. Note that induced sub-channels

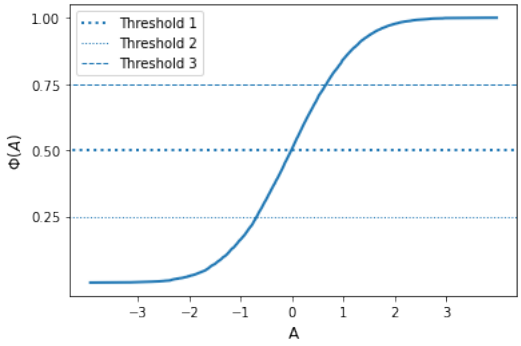

Fig. 2. Cumulative distribution function of the Normal Distribution. *Threshold 1* corresponds to $\Phi(a) = 1/4$ and *Threshold 2* to $\Phi(a) = 3/4$.

have different capacities, namely $C_j = 1 - \mathcal{H}(\alpha_j)$ where $\mathcal{H}(x) = -x\log(x) - (1-x)\log(1-x)$ stands for the binary entropy parameter $\alpha \in (0, 1)$. In encoding the sequences of each subchannel, a code with a rate compatible with this capacity, i.e., less than C_j, should be used. The block-length, in turn, must be compatible with the length of the $\boldsymbol{H}$ matrix used.

TABLE I
CAPACITY OF THE FIRST THREE SUB-CHANNELS OBTAINED WITH QUANTIZATION.

SNR (dB)	Channel Capacity (1)	Channel Capacity (2)	Channel Capacity (3)
0	0.1881	0.0216	0.0034
4	0.3212	0.0480	0.0074
8	0.4696	0.1250	0.0165
12	0.6033	0.2736	0.0492
16	0.7123	0.4422	0.1576
20	0.7975	0.5905	0.3316
24	0.8573	0.7040	0.4967
28	0.9019	0.7921	0.6349
32	0.9328	0.8544	0.7390
36	0.9543	0.8989	0.8169
40	0.9693	0.9315	0.8730

III. BIT-FLIPPING DECODING FOR THE FIRST TWO SUB-CHANNELS OF QUANTIZATION

Here, the first two sub-channels obtained after quantization were used in the simulations. In the first channel, we have $\boldsymbol{X} = (\mathcal{D}_1(X_1), ..., \mathcal{D}_1(X_n))$ and $\boldsymbol{Y} = (\mathcal{D}_1(Y_1), ..., \mathcal{D}_1(Y_n))$. In the simulation of the second channel, we have $\boldsymbol{X} = (\mathcal{D}_2(X_1), ..., \mathcal{D}_2(X_n))$ and $\boldsymbol{Y} = (\mathcal{D}_2(Y_1), ..., \mathcal{D}_2(Y_n))$. Figure 3 illustrates the average distance between $\boldsymbol{X}$ and $\boldsymbol{Y}$ for the two channels, after 1000 executions of the algorithm, as a function of SNR. Figure 4 shows the histograms of the number of occurrences of these distances for three SNR values belonging to the region where the discrepancy between the two channels is greatest, obtained after 1000 runs of the algorithm. The occurrences were grouped according to 20 distance intervals.

Once the capacity of each sub-channel has been calculated as a function of the desired SNR, it is possible to choose the LDPC matrix to be used in the information reconciliation. Considering a reverse reconciliation scheme, in which Alice needs to correct her sequence to match Bob's, simulations

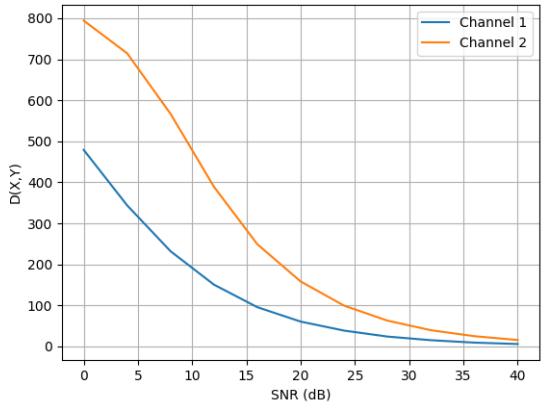

Fig. 3. Average Hamming distance between $\boldsymbol{X}$ and $\boldsymbol{Y}$ as a function of SNR, before decoding, for channels 1 and 2, after 1000 executions of the algorithm.

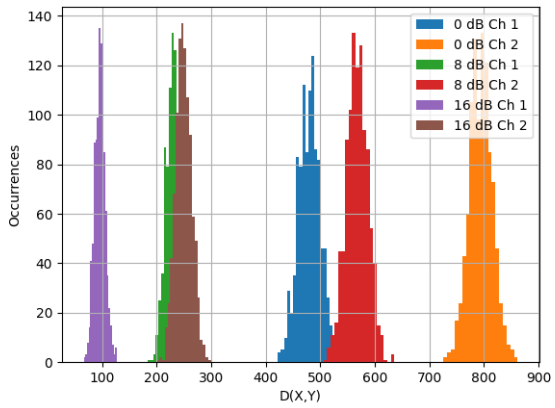

Fig. 4. Histograms with the Hamming distances between $\boldsymbol{X}$ and $\boldsymbol{Y}$ grouped into 20 intervals, based on SNR, for channels 1 and 2, after 1000 executions of the algorithm.

were performed with Slepian-Wolf coding [20], [2] and a modification of the Bit-Flipping algorithm, proposed in [21] (submitted for publication), for the binary sequences of the first two channels of interest.

In Slepian-Wolf coding [20], [22], Bob computes $S(\boldsymbol{Y}) = \boldsymbol{Y}\boldsymbol{H}^T$ and sends it to Alice through an authenticated classical channel, assumed to be error-free. It is important to note that in this type of analysis, thermal noise and other losses that may be imposed by the classical channel are irrelevant, and neither $\boldsymbol{X}$ nor $\boldsymbol{Y}$ are codewords. Alice can then reconstruct Bob's sequence from $\boldsymbol{X}$ and $S(\boldsymbol{Y})$. In [20] a modification of the sum-product algorithm, based on belief propagation, is applied. This modification includes the term $(1-2s_j)$, where s_j corresponds to the j-th component of Bob's syndrome, in the calculation of the likelihood ratios sent by the parity nodes to Alice's variable nodes. In this decoding process, the algorithm modifies Alice's sequence so that it matches Bob's. Decoding ends when the syndrome of the new sequence matches Bob's syndrome or a maximum number of iterations is reached.

In the modification of the Bit-Flipping algorithm presented in [21] (submitted for publication), the goal is to reconstruct the error vector between $\boldsymbol{X}$ and $\boldsymbol{Y}$ from the syndrome $S(\boldsymbol{E}) = S(\boldsymbol{X}) \oplus S(\boldsymbol{Y})$, and then, add it to Alice's sequence to match Bob's. The proposed modification to the algorithm is based on flipping the bits of the null sequence that are connected to the maximum number of errors in $S(\boldsymbol{E})$ (with value 1). The decoding stops when a sequence with a syndrome equal to $S(\boldsymbol{E})$ is obtained or a maximum number of iterations is reached.

The LDPC code used is described by its parity-check matrix $\boldsymbol{H}$, with dimensions 960×1920 and rate $r = 1/2$, of irregular type and quasi-cyclic, available in the G.hn standard [23]. The sequences $\boldsymbol{X}$ e $\boldsymbol{Y}$ had a length of $N = 1920$, compatible with the length of the parity-check matrix used. Figure 5 shows the success rate for the two channels. A decoding is considered successful when a vector, starting from the null vector, is obtained with a syndrome equal to $S(\boldsymbol{E})$. Since the same LDPC matrix was used to encode both channels, it can be seen in Table I that only between 8 and 12 dB does the channel capacity exceed the code rate $(1/2)$, so successful decodings begin to appear in this range. In the second channel, the capacity only exceeds the code rate between 16 and 20 dB, so no successful decodings occur before this range.

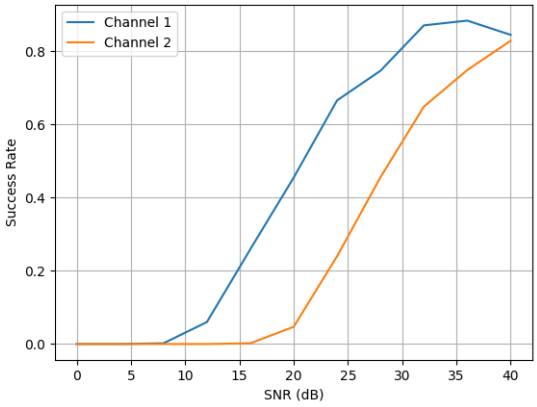

Fig. 5. Success rate after decoding as a function of SNR, for channels 1 and 2.

Figure 6 presents the average number of iterations for the algorithm to converge as a function of SNR, that is, for a successful decoding to occur. A gap between the curves is observed, due to the fact that in the second channel, successful decodings begin to occur only at higher SNRs, along with an increase followed by a subsequent decrease in the number of iterations for convergence. This may occur because, at lower SNRs, the algorithm finds an error vector far from the desired one, but with the same syndrome. At higher SNRs, around 40 dB, the average number of iterations is also lower because the difference between $\boldsymbol{X}$ and $\boldsymbol{Y}$ is small.

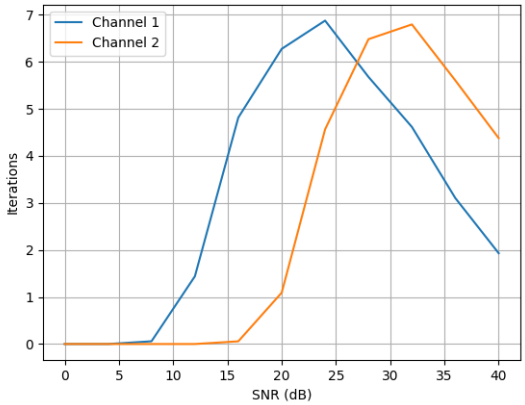

Fig. 6. Average number of necessary iterations to match the syndromes S($\boldsymbol{X}$) and S($\boldsymbol{Y}$), in channels 1 and 2.

IV. Comparison between Bit-Flipping and Sum-Product Decoding

Lastly, the simulations of channel 1 were replicated, following Slepian-Wolf coding, using the same LDPC matrix as before, now applying the Sum-Product algorithm with the modification presented in [20] for decoding. A total of 500 executions of the modified Bit-Flipping and Sum-Product algorithms were carried out for eight SNR values, starting from 12 dB, values for which the Channel Capacity exceeds the code rate. The average number of successful decodings for both algorithms (Figure 7), as well as the average number of iterations for convergence (Figure 8) and the distance between $\boldsymbol{X}$ and $\boldsymbol{Y}$ after decoding (Figure 9) were obtained.

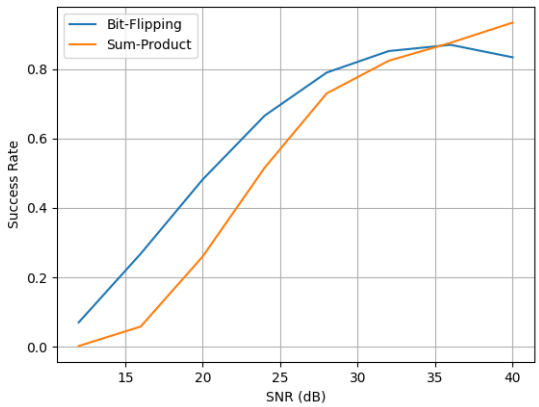

Fig. 7. Success rate after Bit-Flipping and Sum-Product decodings, as a function of SNR, for channel 1.

The success rate for the Bit-Flipping algorithm was higher up to 35 dB, although this algorithm requires a greater number of iterations for convergence. However, the computational cost of Bit-Flipping is much lower than that of the Sum-Product algorithm, so the number of iterations alone does not imply greater decoding speed. Thus, further studies into the application of specific LDPC matrices for CVQKD reconciliation according to the proposal presented in [21] become interesting, especially from the perspective of reducing the computational cost of decoding, in order to improve the reconciliation efficiency of the CVQKD protocol. However, a deeper investigation into the frame error rate (FER) between the blocks of Alice and Bob's sequences used in the reconciliation is also needed. The Sum-Product algorithm showed greater proximity between the sequence obtained after decoding and $\boldsymbol{Y}$.

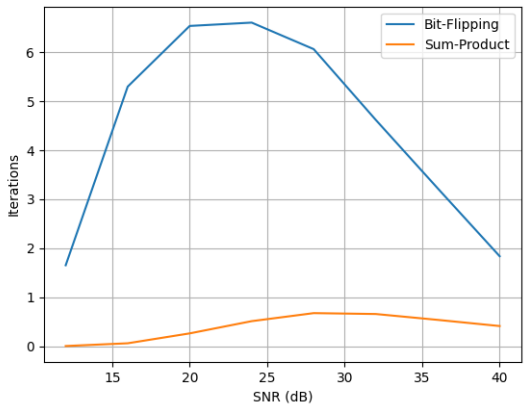

Fig. 8. Average number of necessary iterations to match the syndromes S($\boldsymbol{X}$) and S($\boldsymbol{Y}$), after Bit-Flipping and Sum-Product decodings, for channel 1.

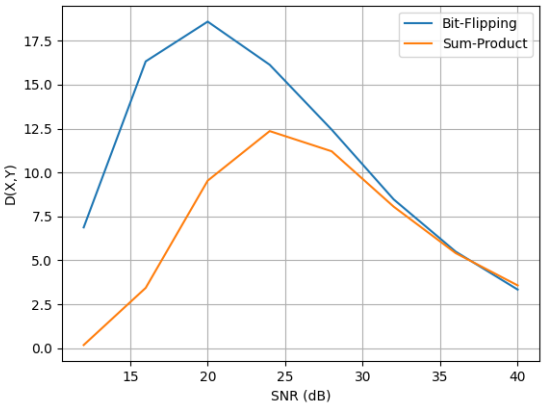

Fig. 9. Difference between the new sequence after Bit-Flipping and Sum-Product decodings and $\boldsymbol{Y}$.

V. Conclusions

In this article, the variable quantization technique proposed in [18], based on the Distributional Transform, was applied together with LDPC Slepian-Wolf coding [20], [22]. In decoding, a modification on the Bit-Flipping algorithm was used to reconstruct the error vector between Alice and Bob from the null vector and the difference between the syndromes $S(\boldsymbol{X})$ and $S(\boldsymbol{Y})$.

Simulations were conducted separately for the first two channels obtained after quantization, using an LDPC matrix

with a rate $r = 1/2$, and the numerical results were graphically presented. Additionally, simulations with the first channel were performed using two different decoding algorithms: the Modified Bit-Flipping and the Sum-Product as presented in [20], in order to evaluate some of the main performance differences between the two techniques, since the Bit-Flipping algorithm has a much lower computational cost than the other.

The Bit-Flipping algorithm applied to the proposed error vector reconstruction from the null vector showed a higher success rate than the Sum-Product algorithm in the 10 to 15 dB range, despite requiring a greater number of iterations for convergence. However, the post-decoding sequence obtained with the Sum-Product algorithm exhibited fewer errors relative to $\boldsymbol{Y}$. Nonetheless, the results encourage further research on the feasibility of applying Bit-Flipping-based reconciliation and error vector reconstruction to CVQKD protocols as a way to reduce the computational cost of decoding. To do this, it will be necessary to use LDPC matrices that are more appropriated for CVQKD protocols, i.e., with lower rates and larger dimensions.

ACKNOWLEDGEMENTS

The authors would like to thank CNPq and COPELE - UFCG for their financial support and IQuanta for the opportunity to carry out this research and for the availability of structure and material support.

REFERENCES

[1] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," *International Conference on Computers, Systems, and Signals Processing*, pp. 175–179, 1984.

[2] K. Kasai, R. Matsumoto, and K. Sakaniwa, "Information reconciliation for qkd with rate-compatible non-binary ldpc codes," in *2010 International Symposium On Information Theory Its Applications*, pp. 922–927, 2010.

[3] S. Bai, Zengliang; Yang and Y. Li, "High-efficiency reconciliation for continuous variable quantum key distribution," *Japanese Journal of Applied Physics*, Apr 2017.

[4] C. H. Bennet, "Quantum cryptography using any two nonorthogonal states," *Physical Review Letters*, vol. 68, pp. 3121–3124, may 1992.

[5] A. K. Ekert, "Quantum cryptography based on bell's theorem," *Physical Review Letters*, vol. 67, pp. 661–663, august 1991.

[6] F. Grosshans and P. Grangier, "Continuous variable quantum cryptography using coherent states," *Phys. Rev. Lett.*, vol. 88, p. 057902, Jan 2002.

[7] N. J. W. J. T.-B. R. Grosshans, Frédéric; Cerf and P. Grangier, "Virtual entanglement and reconciliation protocols for quantum cryptography with continuous variables," *Quantum Information and Computation*, vol. 0, no. 0, 2003.

[8] P. D. E. Ghorai, Shouvik; Grangier and A. Leverrier, "Asymptotic security of continuous-variable quantum key distribution with a discrete modulation," *Physical Review X*, vol. 9, june 2019.

[9] C. Weedbrook, A. M. Lance, W. P. Bowen, T. Symul, T. C. Ralph, and P. K. Lam, "Quantum cryptography without switching," *Phys. Rev. Lett.*, vol. 93, p. 170504, Oct 2004.

[10] E. Diamanti and A. Leverrier, "Distributing secret keys with quantum continuous variables: Principle, security and implementations," *Entropy*, august 2015.

[11] M. A. Dias and F. M. d. Assis, "Distributional transform based information reconciliation," april 2022.

[12] P. Jouguet, D. Elkouss, and S. Kunz-Jacques, "High-bit-rate continuous-variable quantum key distribution," *Phys. Rev. A*, vol. 90, p. 042329, Oct 2014.

[13] J. Assche, Gilles Van; Cardinal and N. J. Cerf, "Reconciliation of a quantum-distributed gaussian key," *IEEE Transactions on Information Theory*, vol. 50, Feb 2004.

[14] L. M. C. d. Araújo, "Novo método de quantização para protocolos de reconciliação de chaves secretas geradas quanticamente utilizando códigos ldpc no sentido slepian-wolf," 2017.

[15] C. Z. L. M. Milicevic, Mario; Feng and P. G. Gulak, "Key reconciliation with low-density parity-check codes for long-distance quantum cryptography," *ArXiv*, 2017.

[16] A. Leverrier, R. Alléaume, J. Boutros, G. Zémor, and P. Grangier, "Multidimensional reconciliation for a continuous-variable quantum key distribution," *Phys. Rev. A*, vol. 77, p. 042325, Apr 2008.

[17] Y. L. Z. W. X. Wang, Pu; Zhang and Y. Li, "Discrete-modulation continuous-variable quantum key distribution with a high key rate," *New Journal of Physics*, feb 2023.

[18] F. M. d. Araújo, Laryssa M. C. de; Assis and B. B. Albert, "Novo protocolo de reconciliação de chaves secretas geradas quanticamente utilizando códigos ldpc no sentido slepian-wolf," in *XXXVI Simpósio Brasileiro de Telecomunicações e Processamento de Sinais*, 2018.

[19] T. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006.

[20] Z. Liveris, Angelos D.; Xiong and C. N. Georghiades, "Compression of binary sources with side information at the decoder using ldpc codes," *IEEE Communications Letters*, vol. 6, oct 2002.

[21] R. R. S. Leite and F. M. d. Assis, "Cvqkd reconciliation with slepian-wolf ldpc coding and bit-flipping decoding," in *Submetido à publicação*.

[22] R. D. J. Limei, Guo; Qi and H. Duan, "Qkd iterative information reconciliation based on ldpc codes," *International Journal of Theoretical Physics*, 2020.

[23] *Unified high-speed wire-line based home networking transceivers - System architecture and physical layer specification*, jun 2010.