Huffman Distribution Matching for CV-QKD

Caroline da Silva Morais Alves[1], Gabrielly da Silva Roman[1], Micael Andrade Dias[2,1,*]

[1]*QuIIN – Quantum Industrial Innovation, Centro de Competência EMBRAPII CIMATEC em Tecnologias Quânticas, SENAI CIMATEC, Av. Orlando Gomes 1845, 41650-010, Salvador, BA, Brasil.*

[2]*Department of Electrical and Photonics Engineering, Technical University of Denmark, 2800 Lyngby, Denmark*
[*]*Corresponding author: mandi@dtu.dk*

Abstract: This paper examines the application of probabilistic shaping techniques, specifically Huffman Shaping (HS) and Geometric Huffman Coding (GHC), to optimize probability distributions in Continuous-Variable Quantum Key Distribution (CV-QKD) systems. The study focuses on the use of probabilistic shaping methods, in particular HS and GHC, to generate non-uniform discrete constellations. A comparative analysis is carried out between HS and GHC for generating dyadic probability distributions that approximate target probability mass functions (PMFs). The performance of the algorithms was evaluated using the Kullback–Leibler (KL) divergence and the variational distance. Numerical results show that GHC consistently achieves a lower KL divergence than HS for all M-PAM constellations and both target distributions, indicating a better approximation of the desired distribution. Although the variational distance shows a less pronounced difference, GHC still attains smaller distances. Furthermore, the paper establishes a crucial connection between classical and quantum distance measures, showing that for commuting quantum states (such as those prepared in this context), the quantum trace distance reduces to the classical trace distance (or half the variational distance). This validates the application of classical distribution matching algorithms in the analysis of quantum information problems.

Keywords: Quantum Key Distribution, Discrete Modulation, Probabilistic Shaping, Huffman Coding.
Abbreviations: CV-QKD, GHS, HS.

1. Introduction

A quantum key distribution (QKD) protocol is a cryptographic primitive that makes use of an insecure, noisy quantum channel and an authenticated, noiseless public channel to generate random secret cryptographic keys by the exchange of non-orthogonal quantum states through the quantum channel, and by applying classical processing using the classical channel. Typically, the legitimate parties are Alice and Bob, the transmitter and receiver, respectively, and the protocol is assumed to run in the presence of an eavesdropper, Eve [1].

QKD was initially implemented using discrete variables (DV), wherein information is encoded in the polarization states of single photons [2]. Subsequently, continuous-variable (CV) protocols were developed, encoding information in the quadratures of the electromagnetic field [3]. The primary advantage of CV-QKD protocols lies in their ease of integration with existing fiber optic network infrastructures, utilizing coherent detection (homodyne or heterodyne) with commercial components [4].

Although Gaussian modulated protocols are well developed, either in the theoretical security proofs [5] and in the implementations [6], they have a fundamental limitation due to the requirement of a theoretically continuous modulation, whereas analog-to-digital (ADCs) and digital-to-analog converters (DACs) have finite precision [7]. As an alternative, CV-QKD protocols can use discrete modulation schemes, leveraging techniques

well-known from classical communication. Discrete modulated CV-QKD protocols have been proposed with Phase Shift Keying modulation (PSK), Amplitude Phase Shift Keying (APSK) and Quadrature Amplitude Modulation (QAM) [8, 9, 10]. It is known, from classical information theory results, that such signaling schemes must approach a Gaussian distribution to optimize the system capacity [11].

To generate non-uniform discrete constellations, probabilistic shaping techniques can be employed to assign higher occurrence probabilities to lower-energy symbols, thereby approximating the constellation distribution to an ideal Gaussian. This process is referred to as distribution matching. Examples include Huffman Shaping (HS) [12] and the Geometric Huffman Coding (GHC) [13], both of which have been extensively studied in classical optical communications and are readily adaptable to the continuous-variable QKD context.

This work presents a comparative analysis of HS and GHC for the generation of dyadic probability distributions that accurately approximate target probability mass functions (PMFs) in the context of discrete constellations. The evaluation considers binomial and Gauss–Hermite target distributions, with performance assessed through the Kullback–Leibler divergence (KL) and the variational distance.

The rest of the paper is organized as follows. section 2 gives an overview of CV-QKD protocols with discrete modulation and the practical need for distribution matching. In subsubsection 3.1.1, the distribution matching techniques are detailed, and the numerical results are presented in section 4. The final considerations are given in section 6.

2. CV-QKD Protocols with Discrete Modulation

A CV-QKD protocol consists of four general steps: quantum communication (state preparation, transmission, and measurement), parameter estimation, information reconciliation, and privacy amplification. For a CV-QKD protocol using homodyne measurements, there will be an intermediate step before parameter estimation when Bob announces which quadrature was measured in each round. This stage is called sifting.

The discrete modulation takes place in the state preparation stage, when Alice prepares coherent states from the ensemble $\{|\alpha\rangle, p(\alpha)\}$ where $\alpha \in \mathbb{C}$ is the complex amplitude of the coherent state. For optimal performance, that is, to approximate the capacity of a Gaussian modulated protocol, the probability distribution $p(\alpha)$ must be designed to result in a mixed state $\rho = \sum p(\alpha)\,|\alpha\rangle\langle\alpha|$ close to a Gaussian state. The choice of $p(\alpha)$ is not arbitrary, and it must be optimized for each scenario [14]. Also, several probability distributions are known to converge to a Gaussian distribution and approximate the classical capacity of communication channels [11].

In this way, practical implementations of CV-QKD protocols with discrete modulation need to apply distribution-matching algorithms. Instead of a theoretical black-box that generates numbers drawn following the target distribution $p(\alpha)$, a probabilistic constellation shaping architecture applies a distribution matching algorithm that takes as input a sequence of random bits and outputs a sequence of symbol following an approximation of $p(\alpha)$.

3. Huffman Distribution Matching

For probabilistic constellation shaping, Huffman-based distribution matching maps symbols to codes that approximate a dyadic distribution, defined by $p_i = 2^{-l_i}$, where $l_i \in \mathbb{N}$ is the code length. The algorithm requires a non-uniform constellation near a Gaussian distribution [12], such as Gauss–Hermite or Random Walk (binomial) [1].

Table 1 shows the results of the symbol-to-codeword mapping process for a 6-PAM constellation based on a target Gauss-Hermite distribution.

Table 1: Example of applying Huffman coding for mapping uniform PAM constellations to non-uniform ones.

PAM-Symbols	Gauss-Hermite PMF	Dyadic PMF	Codewords
-3.3243	0.0026	0.125	$0,0,0$
-1.8892	0.0886	0.125	$0,0,1$
-0.6167	0.4088	0.25	$0,1$
0.6167	0.4088	0.25	$1,1$
1.8892	0.0886	0.125	$1,0,1$
3.3243	0.0026	0.125	$1,0,0$

3.1. The Huffman Coding

The Huffman algorithm constructs an optimal prefix code (no codeword is a prefix of another) for a given probability distribution. The goal is to assign a variable-length binary code to each input symbol so that no code is a prefix of another (ensuring unique and instantaneous decoding) and the average code length (in bits per symbol) is as small as possible. The central idea is to assign shorter codes to the most frequent symbols and longer codes to the less frequent ones [15].

The encoding process involves iteratively combining the two least probable symbols into a single "supersymbol", with its probability being the sum of the probabilities of the combined symbols $(x^{'} = x_m + x_{m-1})$. This process continues until only one symbol remains.

Since Huffman Coding (HC) is designed to find the code lengths l_i that minimize the average length, by mathematical equivalence, it also minimizes the KL divergence $D_{KL}(x\|p)$, where x is the target distribution and p is the dyadic probability distribution. However, for communication channels, we need to minimize $D_{KL}(p\|x)$. The reason lies in the direct relationship between KL divergence and information loss: when we deviate from the optimal distribution, we lose transmission capacity, and this loss is quantified by $D_{KL}(p\|x)$. This is the fundamental difference between HC and GHC.

3.1.1. The Huffman Shaping

HS reverses the conventional use of HC by constructing a code from the target probability dis-

tribution p of modulation symbols in an M-ary constellation, rather than encoding source symbols [12].

3.2. Geometric Huffman Coding

GHC addresses this channel-oriented criterion by producing a dyadic distribution optimally close to the target, rather than simply minimizing average code length. For this purpose, the GHC algorithm updates the probabilities based on the geometric mean during the construction of the Huffman tree. Its computational complexity remains $\mathcal{O}(m \log m)$, where m is the number of input symbols, matching that of standard Huffman coding [13].

Assuming a vector $x = (x_1, x_2, \ldots, x_m)$ with non-negative entries, ordered such that $(x_1 \geq x_2 \geq \cdots \geq x_m)$ [13]:

$$x' = \begin{cases} x_{m-1}, & \text{se } x_{m-1} \geq 4x_m \\ 2\sqrt{x_m x_{m-1}}, & \text{se } x_{m-1} < 4x_m. \end{cases} \quad (1)$$

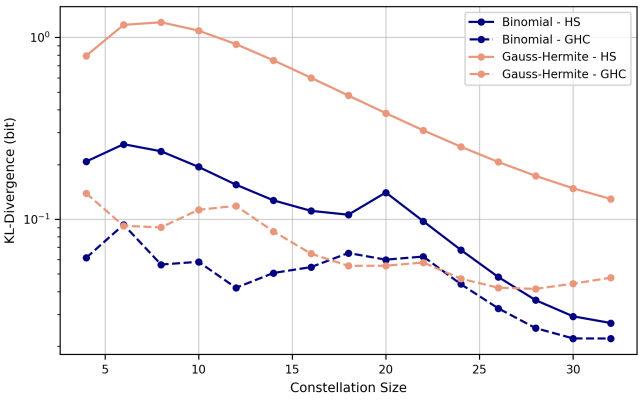
Böcherer and Mathar show that, for discrete memoryless channels, the search for good dyadic input PMFs is equivalent to minimizing $D_{KL}(p \parallel p^*)$ over all dyadic PMFs p, where p^* represents the PMF achieving channel capacity. Moreover, GHC asymptotically achieves channel capacity when the block length $k \to \infty$, as the normalized KL divergence $D_{KL}\left(p^{(k)} \parallel p^{*(k)}\right)/k$ converges to zero. [13]

4. Numerical Results

To evaluate the effectiveness of HS and GHC in constructing dyadic PMFs that approximate target distributions, binomial and Gauss–Hermite distributions were generated for M-PAM constellations, with M even and $M \in \{4, 6, 8, \ldots, 32\}$. Subsequently, the dyadic PMFs obtained from each method were compared to the target distributions using KL divergence and variational distance, enabling an assessment of each algorithm's efficiency across different PAM constellation scenarios.

In Figure 1, we plot the KL divergence between the target distributions and those obtained with HS and GHC. For all M-PAM constellations and both target distributions, it can be observed that GHC yields a systematically lower D_{KL} than HS, indicating a better approximation of the desired distribution.

Figure 1: KL divergence for M-PAM using Binomial and Gauss–Hermite with HS and GHC.



For the Binomial distribution, the divergences obtained by GHC have an average of ≈ 0.05 bit, with

a maximum observed divergence of ≈ 0.09 bit. In comparison, HS has an average value of ≈ 0.12 bit, and a maximum observed divergence of ≈ 0.26 bit.

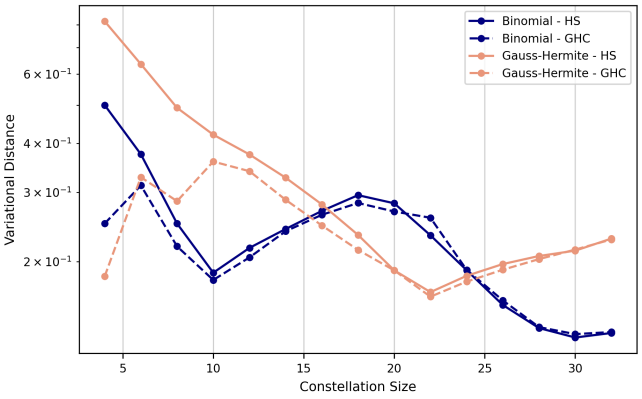
For the Gauss-Hermite distribution, the gap between the techniques is more pronounced. GHC achieves an average result of ≈ 0.07 bit, with a maximum observed divergence of ≈ 0.14 bit. In contrast, HS reaches a maximum D_{KL} of ≈ 1.2 bits for small constellations, gradually decreasing as the constellation size increases, with an average of ≈ 0.57 bit.

For the variational distance, the difference between the GHC and HS methods is less pronounced compared to the KL divergence, as shown in Figure 2. Although GHC achieves smaller variational distances, the curves converge more closely, especially for the Binomial distribution. For the Binomial distribution, the average distance for GHC is ≈ 0.22, with a maximum observed distance of ≈ 0.31. In contrast, HS has an average of ≈ 0.24 and a maximum observed distance of ≈ 0.5. For the Gauss-Hermite distribution, GHC achieves an average of ≈ 0.24, whereas HS has an average of ≈ 0.33.

5. Classical and Quantum Trace Distances

In the field of information theory, both classical and quantum, the ability to quantify the closeness between different information states is fundamental. In this context, the trace distance emerges as

Figure 2: Variational distance for M-PAM using Binomial and Gauss–Hermite with HS and GHC.



one of the most important and widely used measures.

In the context of classical information theory, the trace distance is a measure used to compare the closeness between two probability distributions, p_x and q_x, defined over the same set of indices x. This measure is defined by [16]:

$$D(p_x, q_x) = \frac{1}{2} \sum_x |p_x - q_x|. \tag{2}$$

This quantity provides an operational interpretation that quantifies the distinguishability between two probability distributions. An interesting fact about the trace distance is its similarity to the concept of variational distance. The variational distance between two probability distributions $t = (t_1, t_2, \ldots)$ and $\hat{t} = (\hat{t}_1, \hat{t}_2, \ldots)$ is given by [17]:

$$|\mathbf{t} - \hat{\mathbf{t}}|_1 = \sum_{\mathbf{i}} |\mathbf{t_i} - \hat{\mathbf{t}}_\mathbf{i}|. \tag{3}$$

The result of this distance measures how well a distribution $\hat{\mathbf{t}}$ approximates a target distribution $\mathbf{t}$.

There are similarities between the classical trace distance and the variational distance: both are based on the L_1 distance and quantify the "distance" between two classical probability distributions. However, there is a difference in how they are defined in the literature. The classical trace distance is defined by including a normalization factor of $1/2$, which makes its value range from 0 (identical distributions) to 1 (perfectly distinguishable distributions). The variational distance, on the other hand, does not include the normalization factor.

It is in this context that we introduce the concept of quantum trace distance. Mathematically, the quantum trace distance is the generalization of the classical trace distance to the domain of quantum states, represented by density operators ρ and σ, and is defined as [16]:

$$D(\rho,\sigma) \equiv \frac{1}{2}\mathrm{tr}|\rho-\sigma|. \tag{4}$$

The definition presented in (4) is a generalization of the definition for probability distributions in (2). The relevance of this generalization becomes particularly evident when the density operators ρ and σ commute [16]. The commutator between two operators A and B is defined as $[A,B] \equiv AB - BA$. We say that two operators commute if $AB - BA = 0$, which implies $AB = BA$ [16]. In quantum terms, this means that the order in which the operations are performed does not matter.

If ρ and σ commute, they can be diagonalized in the same orthonormal basis [16]. In this specific case, the quantum trace distance reduces to the classical trace distance between the eigenvalues of ρ and σ [16]. If

$$\rho = \sum_i r_i|i\rangle\langle i| \text{ e } \sigma = \sum_i s_i|i\rangle\langle i| \tag{5}$$

(r_i and s_i are the eigenvalues corresponding to the respective eigenstates $|i\rangle$), then [16]:

$$D(\rho,\sigma) = D(r_i,s_i). \tag{6}$$

Therefore, the commutation between quantum states ensures that the calculation of the quantum trace distance reduces to the classical trace distance between the probability distributions of the eigenvalues. Note that:

$$D(r_i,s_i) = \frac{1}{2}\sum_i |r_i - s_i| = \frac{1}{2}\|r-s\|_1. \tag{7}$$

The second equality follows directly from the definition of the variational distance (3). Therefore, for commuting quantum states, the quantum trace distance is exactly the variational distance of the eigenvalues corresponding to the respective eigenstates $|i\rangle$:

$$D(\rho,\sigma) = D(r_i,s_i) = \frac{1}{2}\|r-s\|_1. \tag{8}$$

This result is of utmost importance for understanding distribution matching algorithms. In such algorithms, a target distribution $p(x)$ is approxi-

mated by a distribution $q(x)$. The quality of this approximation is often measured by the variational distance $\|p-q\|_1$ [17], which is equivalent to the classical trace distance.

Consider the case in which we have a mixture of states ρ_x and a distribution matching process yields the distribution $q(x)$. We say that, at the transmitter output, we prepare $\rho_x^{'}$ instead of ρ_x, where $\rho_x = \sum_x p(x)|x\rangle\langle x|$ e $\rho_x^{'} = \sum_x q(x)|x\rangle\langle x|$. In these expressions, the states $|x\rangle$ are vectors of an orthonormal basis. This means that if we represent these operators as matrices in this basis $|x\rangle$, they will be diagonal matrices. The (i,i) element of the matrix $\rho_x\rho_x^{'}$ will be $p(i)\cdot q(i)$. And the (i,i) element of the matrix $\rho_x^{'}\rho_x$ will be $q(i)\cdot p(i)$. Since scalar multiplication is commutative ($p(i)\cdot q(i) = q(i)\cdot p(i)$), all the diagonal elements of the two products will be identical. All the off-diagonal elements of both products will be zero, since the original matrices are diagonal. Therefore, $\rho_x\rho_x^{'} = \rho_x^{'}\rho_x$. This directly implies that the commutator is zero: $[\rho_x, \rho_x^{'}] = 0$.

A direct consequence of ρ_x and $\rho_x^{'}$ commuting is given by the equalities in equation (8). Therefore, if ρ_x and $\rho_x^{'}$ commute, we have:

$$D(\rho_x, \rho_x^{'}) = D(p,q) = \frac{1}{2}\|p-q\|_1. \qquad (9)$$

Therefore, if we have a distribution algorithm that converges to the target distribution, the constellation or mixture of quantum states resulting from this approximated distribution will converge to the target mixture of quantum states. This result serves as a kind of guarantee that the classical distribution matching algorithms presented here make sense when we analyze the problem from the perspective of quantum information.

6. Conclusion

This study investigated the optimization of probability distributions in CV-QKD systems through probabilistic shaping techniques: HS and GHC. The motivation lies in the need to use discrete modulations in CV-QKD due to the finite precision of digital converters, aiming to approximate the constellation distribution to an ideal Gaussian.

Numerical results showed that GHC consistently outperformed HS in approximating the target distributions. GHC achieved a significantly lower KL divergence (e.g., an average of $\backsim$ 0.05 bits for binomial and $\backsim$ 0.07 bits for Gauss–Hermite, compared to $\backsim$ 0.12 bits and $\backsim$ 0.57 bits for HS, respectively), indicating better adherence to the desired distribution.

Additionally, we discussed the connection between classical and quantum distance measures. This connection validates the application of classical distribution matching algorithms in the analysis and optimization of quantum information problems. It ensures that if a classical algorithm converges to the target distribution, the resulting mixture of quantum states will also converge to the

target mixture of quantum states.

In summary, the work highlights GHC as a more effective technique for probabilistic shaping in CV-QKD and provides the theoretical foundation that justifies the use of classical information theory tools to optimize quantum communication systems.

Acknowlegement

This work was partially funded by the project *Analysis and Development of Distribution Matching Algorithms for CV-QKD*, supported by QuIIN – Quantum Industrial Innovation, the EMBRAPII CIMATEC Competence Center in Quantum Technologies, with financial resources from the PPI IoT/Industry 4.0 of the MCTI, grant number 053/2023, signed with EMBRAPII; and by EU HORIZON 2023 Marie Skłodowska-Curie Actions Postdoctoral Fellowships under project number 101153602 (COCoVaQ).

References

[1] Eneet Kaur, Saikat Guha, and Mark M. Wilde. Asymptotic security of discrete-modulation protocols for continuous-variable quantum key distribution. *Physical Review A*, 103(1), January 2021.

[2] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical Computer Science*, 560:7–11, 2014. Theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84.

[3] Frédéric Grosshans and Philippe Grangier. Continuous variable quantum cryptography using coherent states. *Physical Review Letters*, 88(5), January 2002.

[4] Eleni Diamanti and Anthony Leverrier. Distributing Secret Keys with Quantum Continuous Variables: Principle, Security and Implementations. *Entropy*, 17(12):6072–6092, August 2015.

[5] Anthony Leverrier. Composable security proof for continuous-variable quantum key distribution with coherent states. *Phys. Rev. Lett.*, 114(7), 2015.

[6] Adnan A. E. Hajomer, Ivan Derkach, Nitin Jain, Hou-Man Chin, Ulrik L. Andersen, and Tobias Gehring. Long-distance continuous-variable quantum key distribution over 100-km fiber with local local oscillator. *Science Advances*, 10(1):eadi9474, January 2024.

[7] Feihu Xu, Xiongfeng Ma, Qiang Zhang, Hoi-Kwong Lo, and Jian-Wei Pan. Secure quantum key distribution with realistic devices. *Reviews of Modern Physics*, 92(2), May 2020.

[8] Panagiotis Papanastasiou, Cosmo Lupo, Christian Weedbrook, and Stefano Pirandola. Quantum key distribution with phase-encoded coherent states: Asymptotic security analysis in thermal-loss channels. *Physical Review A*, 98(1):012340, July 2018.

[9] Ivan B. Djordjevic. Optimized-Eight-State CV-QKD Protocol Outperforming Gaussian Modulation Based Protocols. *IEEE Photonics Journal*, 11(4):1–10, August 2019.

[10] Margarida Almeida, Daniel Pereira, Margarida Facão, Armando N. Pinto, and Nuno A. Silva. Reconciliation Efficiency Impact on Discrete Modulated CV-QKD Systems Key Rates. *Journal of Lightwave Technology*, 41(19):6134–6141, October 2023.

[11] Yihong Wu and Sergio Verdú. The impact of constellation cardinality on gaussian channel capacity. *2010 48th Annual Allerton Conference on Communication, Control, and Computing, Allerton 2010*, pages 620–628, 2010.

[12] G. Ungerboeck. *Huffman Shaping*, pages 299–313. Springer US, Boston, MA, 2002.

[13] G. Böcherer and R. Mathar. Matching dyadic distributions to channels. In *2011 Data Compression Conference*, pages 23–32, 2011.

[14] Michele N. Notarnicola, Stefano Olivares, Enrico Forestieri, Emanuele Parente, Luca Potì, and Marco Secondini. Probabilistic amplitude shaping for continuous-variable quantum key distribution with discrete modulation over a wiretap channel. *IEEE Transactions on Communications*, 72(1):375–386, 2024.

[15] Thomas M Cover. *Elements of information theory*. John Wiley & Sons, 1999.

[16] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

[17] Georg Böcherer and Bernhard C Geiger. Optimal quantization for distribution synthesis. *IEEE Transactions on Information Theory*, 62(11):6162–6172, 2016.