On Thresholding Quantizer Design for Mutual Information Maximization: Optimal Structures and Algorithms

Thuan Nguyen
School of Electrical and
Computer Engineering
Oregon State University
Corvallis, OR, 97331
Email: nguyeth9@oregonstate.edu

Thinh Nguyen
School of Electrical and
Computer Engineering
Oregon State University
Corvallis, 97331
Email: thinhq@eecs.oregonstate.edu

***Abstract*—Consider a channel having the discrete input X that is corrupted by a continuous noise to produce the continuous-valued output U. A thresholding quantizer is then used to quantize the continuous-valued output U to the final discrete output V. One wants to design a thresholding quantizer that maximizes the mutual information between the input and the final quantized output $I(X;V)$. In this paper, the structure of optimal thresholding quantizer is established that finally results in two efficient algorithms having the time complexities $O(NM + K\log^2(NM))$ for finding the local optimal quantizer and $O(KM\log(NM))$ for finding the global optimal quantizer where N, M, K are the size of input X, received output U and quantized output V, respectively. Both theoretical and numerical results are provided to verify our contributions.**

Keyword: channel quantization, mutual information maximization, threshold, partition, optimization.

I. Introduction

Recently, designing a quantizer that maximizes the mutual information between the input and the quantized output is of great interest for many wireless applications. In particular, this type of quantizers is an important component in the design of low density parity check (LDPC) codes and polar codes [1], [2], [3]. Consequently, in recent years, there is a rich literature on finding such quantizers [4]–[13]. In general, the problem of maximizing the mutual information over all the possible quantizers is a hard problem. A naive exhaustive search results in the time complexity of $O(K^M)$ which can quickly become computationally intractable even for the modest values of M and K. To solve this problem, two well-known approaches were proposed: (1) Lloyd's algorithm for finding the local solutions and (2) the dynamic programming for finding a globally optimal solutions.

In the first approach, similar to the well-known Lloyd's algorithm that was proposed nearly fifty years ago [14], Zhang and Kurkoski proposed a k-means algorithm for finding a locally optimal quantizer with the complexity of $O(TNKM)$ where T is the number of iterations [4]. This approach can result in a locally optimal which can be far away from a globally optimal solution. Moreover, in some special cases [15], T is super polynomial which severely degrades the performance of k-means algorithm. On the other hand, under a certain condition i.e., the binary input $N = 2$, the optimal quantized outputs are the contiguous intervals [5], [16] [6], [11]. Thus, the well-known dynamic programming can be applied to find the global optimal solution in a polynomial time complexity $O(NKM^2)$. Based on the matrix searching, SMAWK algorithm [17] can further reduce the time complexity of dynamic programming to $O(NKM)$ under some specified conditions [11], [12], [18]. It is worth noting that the two above methods have a long history which were proposed for finding the optimal quantizer that minimizes the Euclidean distortion [14], [19], [20].

Although the dynamic programming and SMAWK algorithm can provide a polynomial time complexity, in [19], Wu wondered about other algorithms that can find the global solution even faster than $O(NKM)$. He also suggested that the hope about these algorithms cannot be placed on the bottom-up dynamic programming and SMAWK algorithm because the matrix searching style of SMAWK algorithm is already optimal. As the effort to answer the question of Wu, we propose a completely different approach based on the structure of optimal quantizer that results in two efficient algorithms having the time complexities of $O(NM + K\log^2(NM))$ for finding the local optimal quantizer and $O(KM\log(NM))$ for finding the global solution. Consequently, a locally optimal solution can be found in TK time faster than the method proposed by Zhang and Kurkoski [4] and the global solution is still faster than the traditional dynamic programming and other approaches in [5], [6]. Interestingly, the uniqueness of the locally optimal quantizer that minimizes the Euclidean distortion was discovered [21]. However, the condition for uniqueness of locally optimal quantizer that maximizes the mutual information is still an open problem. Thus, if one can find some certain conditions to guarantee that the locally optimal quantizer is unique i.e., any locally optimal solution is the globally optimal solution then our method can find the globally optimal quantizer in $O(NM + K\log^2(NM))$.

978-1-7281-5207-3/20/$31.00 ©2020 IEEE

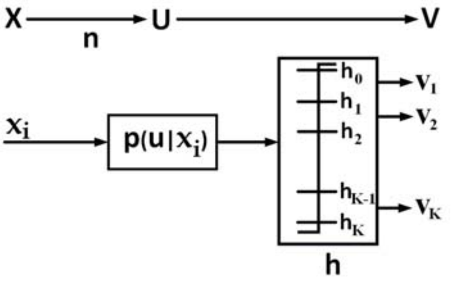

Figure 1: A DMTC having N inputs and K quantized outputs using a quantizer $Q \overset{\triangle}{=} h$ having $K+1$ thresholds.

II. Problem Formulation

Fig. 1 illustrates a thresholding quantizer in a discrete memoryless thresholding channel (DMTC). The input set is discrete consisting of N transmitted symbols $X = \{x_1, x_2, \dots, x_N\}$ having a given input p.m.f $p_X = \{p_{x_1}, p_{x_2} \dots, p_{x_N}\}$. Due to a continuous noise, the received signal $u \in U = \mathbb{R}$ is modeled via the conditional density $p_{U|X}(u|x_i)$. In practice, one can limit U to a finite range i.e., $U = [-A, A]$. We also note that $p_{U|X}(u|x_i)$ can have different statistics associated with each transmitted signal x_i. In the special case where $u_i = x_i + n_i$ with n_i's being i.i.d, then $p_{U|X}(u|x_i)$ is simply a shifted version of $p_{U|X}(u|x_j), \forall i, j$. The quantized outputs v is obtained by quantizing u into K discrete outputs $v_i \in V = \{v_1, v_2, \dots, v_K\}$ using a quantizer Q. Quantizer Q consists of $K+1$ thresholds $h = \{h_0 = -\infty \leq h_1 \leq h_2 \leq \dots \leq h_{K-1} \leq h_K = +\infty\}$ such that $Q(u) = v_i$, if $h_{i-1} \leq u < h_i$ or $v_i = [h_{i-1}, h_i)$. An optimal quantizer Q^* is one that maximizes the mutual information $I(X;V)$, i.e, Q^* is a solution to the following optimization problem:

$$C = \max_Q I(X;V), \tag{1}$$

An important structure of the thresholding quantizer $Q(u)$ is that $Q(u)$ maps all u in every disjoint interval (h_{i-1}, h_i) to a distinct v_i which is similar to the traditional quantization that minimizes Euclidean distortion [19], [20]. The advantage of thresholding quantizer is that it has a simple circuit implementation and is suitable for the decoder that uses Pulse Amplitude Modulation (PAM) where the output symbols are resolved based on the magnitude of received signals [22]. In this paper, we will focus on maximizing the mutual information $I(X;V)$ over all the possible thresholding quantizers, i.e., finding the optimal values of $h^* = \{h_0, h_1^*, \dots, h_{K-1}^*, h_K\}$.

III. Optimality conditions

In this section, we establish some mathematical preliminaries that will be used to develop the proposed algorithms.

A. Notations

Define $p_X = \{p_{x_1}, p_{x_2} \dots, p_{x_N}\} = \{p_1, p_2, \dots, p_N\}$ i.e., $p_j = p_{x_j}$ and define $\phi_i(u) = p_{U|X}(u|x_i)$ as the conditional noise density of u given the transmitted input x_i. Due to $Q(u) = v_i$, if $h_{i-1} \leq u < h_i$ or $v_i = [h_{i-1}, h_i)$, then

$$p(v_i|x_j) = \int_{u=h_{i-1}}^{h_i} \phi_j(u) du, \tag{2}$$

$$p(v_i) = \sum_{j=1}^{N} p_j p(v_i|x_j), \tag{3}$$

$$p(x_j|v_i) = \frac{p_j p(v_i|x_j)}{\sum_{q=1}^{N} p_q p(v_i|x_q)}, \tag{4}$$

$$p(x_j|u) = \frac{p_j \phi_j(u)}{\sum_{q=1}^{N} p_q \phi_q(u)}. \tag{5}$$

For convenient, we also define two probability vectors x_{v_i} and x_u by

$$x_{v_i} = [p(x_1|v_i), p(x_2|v_i), \dots, (x_N|v_i)], \tag{6}$$

$$x_u = [p(x_1|u), p(x_2|u), \dots, (x_N|u)]. \tag{7}$$

Note that from (2), (4) and (6), x_{v_i} is a function of h_{i-1} and h_i. Now, define $F_j(a) = \int_{-\infty}^{a} \phi_j(u) du$, then $p(v_i|x_j) = F_j(h_i) - F_j(h_{i-1})$. We first discretize U into M disjoint parts having the same width ϵ, then $U = \{u_1, u_2, \dots, u_M\}$, $M = \frac{|U|}{\epsilon}$. If one pre-computes and stores $F_j(u_t)$ for $\forall$ $j = 1, 2, \dots, N$ and $t = 1, 2, \dots, M$ in $O(NM)$, then $p(v_i|x_j)$, $p(v_i)$, $p(x_j|v_i)$ can be determined in $O(1)$.

B. Optimality condition

We first begin with some definitions.

Definition 1. Kullback-Leibler (KL) divergence of two probability vectors $a = (a_1, a_2, \dots, a_N)$ and $b = (b_1, b_2, \dots, b_N)$ of the same outcome set is defined as

$$D(a||b) = \sum_{i=1}^{N} a_i \log(\frac{a_i}{b_i}). \tag{8}$$

Definition 2. A channel is a dominated conditional distribution channel if all the distributions $\phi_i(u)$ satisfies:

$$\frac{\phi_i(u)}{\phi_j(u)} \geq \frac{\phi_i(u')}{\phi_j(u')}, \tag{9}$$

for $\forall$ $i \leq j$ and $u \leq u'$.

In practice, the inequality (9) is not too restricted. For example, in typical communication scenarios [16], [23] where the noise is additive, i.e., $u = x_i + n$, the inequality (9) holds for a variety of common noise distributions such as normal distribution, exponential distribution, gamma distribution, uniform distribution, and more generally, all log-concave distributions (Corollary 2 [16]). Now, we are ready to show the main results.

Theorem 1. For an optimal quantizer Q^* of DMTC having optimal threshold $h = \{h_0, h_1^*, \dots, h_{K-1}^*, h_K\}$, then

$$D(x_{u=h_i^*}||x_{v_i}) = D(x_{u=h_i^*}||x_{v_{i+1}}), \forall i, \tag{10}$$

where x_{v_i} and x_u are defined in (6) and (7), respectively.

Proof. Please see the Appendix A. □

Theorem 2. For a given thresholds h_{i-2} and h_{i-1} that generates $v_{i-1} = [h_{i-2}, h_{i-1})$, if the conditional density $\phi_j(u)$ satisfies (9), then existing a unique h_i that generates $v_i = [h_{i-1}, h_i)$ such that

$$D(x_{u=h_{i-1}}||x_{v_{i-1}}) = D(x_{u=h_{i-1}}||x_{v_i}). \qquad (11)$$

Proof. Please see our extension version. □

Since h_i is unique by Theorem 2, for a given h_{i-2} and h_{i-1}, h_i can be found in $O(\log(NM))$ using the bisection method where N is the size of vectors x_u, x_{v_i} and $M = \frac{|U|}{\epsilon}$ where ϵ is the solution resolution. For convenient, the detail of bisection method is provided at Sec. IV-C.

IV. Algorithms

Based on the structure of an optimal quantizer, we propose two efficient algorithms, one to determine a globally optimal solution with time complexity $O(KM\log(NM))$ and the other to find a locally optimal quantizer with time complexity of $O(NM + K\log^2(NM))$.

A. Algorithm 1

Algorithm 1 $O(KM\log(NM))$ time complexity algorithm finding the global optimal quantizer.

```
Input: $N$, $K$, $p_X$, $U$, $\phi_i(u)$, $\epsilon$, $I(X;V)_{opt} = 0$.
For $h_1^* \in U$.
        Finding $h_2^*, h_3^*, \ldots, h_{K-1}^*$ using bisection method.
        Computing $I(X;V)_{h_1^*}$.
            If $I(X;V)_{h_1^*} > I(X;V)_{opt}$:
                $I(X;V)_{opt} = I(X;V)_{h_1^*}$.
                $h^* = \{h_0, h_1^*, \ldots, h_{K-1}^*, h_K\}$.
            End If
End For
Output: $I(X;V)_{opt}$, $h^* = \{h_0, h_1^*, \ldots, h_{K-1}^*, h_K\}$.
```

Algorithm 1 finds the global optimal quantizer in $O(KM\log(NM))$ time complexity as follows. Consider an optimal quantizer Q^* (local or global) having the optimal thresholds $h^* = \{h_0, h_1^*, h_2^*, \ldots, h_{K-1}^*, h_K\}$. From Theorem 1, we have

$$\begin{aligned} D(x_{u=h_1^*}||x_{v_1}) &= D(x_{u=h_1^*}||x_{v_2}), \\ D(x_{u=h_2^*}||x_{v_2}) &= D(x_{u=h_2^*}||x_{v_3}), \\ \ldots &= \ldots \\ D(x_{u=h_{K-1}^*}||x_{v_{K-1}}) &= D(x_{u=h_{K-1}^*}||x_{v_K}). \end{aligned} \qquad (12)$$

Since h_0 is given, from Theorem 2, given h_1^*, h_2^* is unique. Similarly, given h_2^*, h_3^* is unique, and so on. Thus, each optimal threshold vector $h^* = \{h_0, h_1^*, \ldots, h_{K-1}^*, h_K\}$ is a function of a single optimal threshold h_1^*. From (6), x_{v_i} is a function of h_{i-1}^* and h_i^*, therefore x_{v_i} is a function of h_1^*. Now, by back substituting through all the equations in (12), at the last equation, h_1^* is a root of $\delta(h_1^*)$,

$$\delta(h_1^*) = D(x_{u=h_{K-1}^*}||x_{v_{K-1}}) - D(x_{u=h_{K-1}^*}||x_{v_K}) = 0. \qquad (13)$$

By performing an exhaustive search over all the possible values of $h_1^* \in U$, Algorithm 1 can find all of the locally optimal quantizers. From these local optimal quantizers, the global optimal quantizer can be determined.

Complexity: noting that for a given h_0 and h_1^*, h_2^* can be solved using the bisection method in $O(\log NM)$ time complexity. Next, h_3^* can be solved using bisection method for a given of h_1^* and h_2^* and so on. Thus, the total complexity of Algorithm 1 is $O(KM\log(NM))$.

B. Algorithm 2

From the analysis in Algorithm 1, finding all the locally optimal quantizer is equivalent to determining all the roots of (13) respect to variable h_1. Various fast root-finding algorithms can be deployed to find a root of an equation i.e., Newton's method, Secant method, and Durand-Kerner method [24]. It is not known whether there is an efficient algorithm that guarantees to find all the roots of an arbitrary equation. However, it can be verified that by setting $h_1^* = h_0$ will result in $\delta(h_1^*) < 0$ and setting $h_1^* = h_K$ will make $\delta(h_1^*) > 0$. Thus, we can use the bisection algorithm to find a single root h_1^* with two initial values h_0 and h_K. The Algorithm 2 below finds the local optimal quantizer in $O(NM + K\log^2(NM))$.

Algorithm 2 $O(NM + K\log^2(NM))$ time complexity algorithm finding the local optimal quantizer.

```
Input: $N$, $K$, $p_X$, $U$, $\phi_i(u)$, $\epsilon$, $\theta$ is a small number that
controls the precise of root-finding.
Initialization: $a = h_0$, $b = h_K$, $t = 0$, $M = \frac{|U|}{\epsilon}$.
While $t \le \log_2(M)$ or $|\delta(h_1^t)| > \theta$, then $h_1^t = \frac{a+b}{2}$.
    For a given $h_1^t$, finding other thresholds
    $\{h_2^t, \ldots, h_{K-1}^t\}$ and $x_{v_1}^t, x_{v_2}^t, \ldots, x_{v_K}^t$.
    Compute
        $\delta(h_1^t) = D(x_{u=h_{K-1}^t}||x_{v_{K-1}}^t) - D(x_{u=h_{K-1}^t}||x_{v_K}^t)$.
    If $\delta(h_1^t) > 0$ then $b = h_1^t$
    Else $a = h_1^t$
    $t = t + 1$
End while
Output: For a given $h^t = \{h_0, h_1^t, \ldots, h_{K-1}^t, h_K\}$, com-
pute $I^t(X;V)$.
```

Complexity: From Theorem 2, for a given h_1^t, other $h_2^t, \ldots, h_{K-1}^t$ can be found in $O((K-2)\log(NM))$ by solving (12) using the bisection method. Now, we have one more outer loop to find the optimal h_1^t in $O(\log(NM))$ that solves (13). To evaluate $\delta(h_1^t)$, we also need to evaluate x_{v_i} $\forall\ i$ that takes $O(NM)$ as discussed in Section III-A. Thus, the time complexity of Algorithm 2 to find a locally optimal solution is $O(NM + K\log^2(NM))$.

Running times of different algorithms (seconds)		
Algorithm	**Example 1**	**Example 2**
Algorithm 2	5.61	6.21
k-means	18.77	20.05
Algorithm 1	105.10	195.07
Dynamic programming based	209.67	317.80
Exhaustive search	2975.28	32177.20

Table I: Running times in seconds of different algorithms used in Example 1 and Example 2.

C. Bisection algorithm

If for some $x_1 < x_2$ and if $F(x_1) < 0$ and $F(x_2) > 0$, to solve the equation $F(x) = 0$ over the interval $[x_1, x_2]$, one can evaluate $F(\frac{x_1+x_2}{2})$ to determine whether it is larger or smaller than 0. If it is larger than 0, we repeat the process on the interval $[x_1, \frac{x_1+x_2}{2}]$. Otherwise, we repeat the process on the interval $[\frac{x_1+x_2}{2}, x_2]$. The process repeats until the solution is found, i.e., within some ϵ away from zero. The complexity of bisection algorithm is $O(\log D)$ where $D = \dfrac{x_2 - x_1}{\epsilon}$. We note that there exist multiple fast root-finding algorithms such as Newton's method, Secant method, and Durand-Kerner method [24]. These methods often require additional assumptions, e.g., smoothness, closed-form expressions for first and second derivatives to enable faster convergence. To that end, we use the bisection method in this paper for its simplicity.

V. Numerical Evaluations

In this section, we compare the performances in terms of run-time and accuracy of the proposed Algorithm 1 and 2 against those of exhaustive search, k-means algorithm [4], and the dynamic programming based algorithm [5], [12].

Example V.1. *We consider a communication system with an additive noise, i.e.,* $u = x_i + n_i$ *where* n_i *are i.i.d normal distribution* $N(0, 5)$ *and* $x_i \in \{-1, 1, 3, 5\}$, $p_i = 1/4$, $\forall$ *i. As a result,* $\phi_i(u) = N(\mu_i, \sigma_i)$ *where* $\mu_i = x_i$ *and* $\sigma_i = 5$ $\forall$ *i. The final output* $V = \{v_1, v_2, v_3, v_4\}$ *is obtained by quantizing* U *using a quantizer* Q *having 5 thresholds* $h = \{h_0, h_1, h_2, h_3, h_4\}$. *First, we re-normalize* $\phi_i(u)$ *to be nonzero in* $U = [-8, 12]$ *and discretize* U *to 200 bins having the same width* $\epsilon = 0.1$. *Therefore,* $h_0 = -8$, $h_4 = 12$, $M = 200$ *and* $N = K = 4$. *Next, we simultaneously run the proposed Algorithms 1 and 2, exhaustive search, k-means algorithm [4] and the dynamic programming based algorithm [5], [12].*

Example V.2. *Similar to the previous example, we consider a communication system with an additive i.i.d noise following a normal distribution* $N(0, 5)$. *However,* $x_i \in \{-1, 1, 3, 5, 7\}$ *and* $p_i = 1/5$, $\forall$ *i. The final output* $V = \{v_1, v_2, v_3, v_4, v_5\}$ *is quantized from* $U = [-8, 12]$ *width* $\epsilon = 0.1$ *using a quantizer* Q *having 6 thresholds* $h = \{h_0, h_1, h_2, h_3, h_4, h_5\}$. *Therefore,* $h_0 = -8$, $h_5 = 12$, $M = 200$ *and* $N = K = 5$. *Next, we simultaneously run the same five algorithms as in Example 1.*

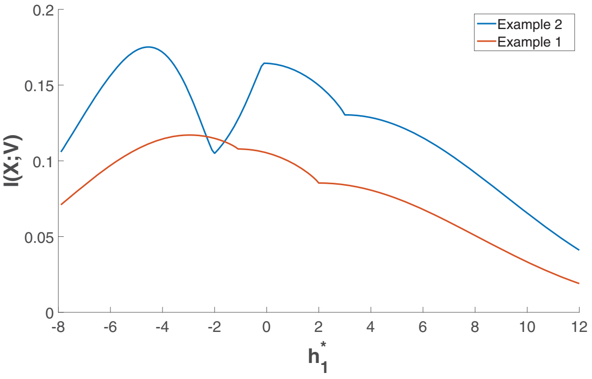

Figure 2: Mutual information as a function of h_1^* for Examples 1 and 2.

Table I shows the running times for five Algorithms on the two examples. For Example 1, the proposed Algorithm 2 is faster than the k-means algorithm while the proposed Algorithm 1 is faster than the dynamic programming based algorithm and exhaustive search as theoretically predicted. In this example, a locally optimal solution is exactly the same as the globally optimal solution since there is a unique maximum as seen in Fig. 2. Consequently, all five algorithms yield the same optimal mutual information $I^*(X; V) = 0.11702$.

As for Example 2, a locally optimal solution is different from the globally optimal solution as seen in Fig. 2. In particular, Algorithm 1 and the dynamic programming based algorithm provide the same globally optimal mutual information at $I^*(X; V) = 0.17514$ while Algorithm 2 and k-means algorithm only obtain the locally optimal solution at $I(X; V) = 0.16444$. As predicted, Algorithm 2 is fastest as shown in Table I, but it can only find a locally optimal solution.

We note that each optimal threshold vector $h^* = \{h_0, h_1^*, \dots, h_{K-1}^*, h_K\}$ is a function of a single optimal threshold h_1^*. Thus, Algorithm 1 can find all the optimal threshold vectors (therefore all the locally optimal quantizers) by searching exhaustively over all the possible values of $h_1 \in U$. Figure. 2 illustrates the mutual information $I(X; V)$ as a function of h_1 in Example 1 and Example 2. As seen, in Example 1, there is only one locally optimal quantizer, and therefore it is also a globally optimal quantizer. As such, all algorithms obtained the globally optimal solution. In a general scenario, Example 2 shows that there exist multiple locally optimal threshold vectors or multiple locally optimal quantizers. Consequently, only Algorithm 1, the dynamic programming based algorithm, and the exhaustive search can guarantee to find a globally optimal solution.

VI. Discussion and Open Problem

Algorithm 1 finds the global solution in $O(KM \log(NM))$, which is faster than the traditional dynamic programming [5] and other methods in [6]. However, it is still slower than the matrix searching SMAWK algorithm [11], [12]. On the other hand, Algorithm 2 is the fastest state of art for finding a locally optimal quantizer that is TK times faster than the traditional k-means algorithm [4].

Interestingly, the uniqueness of the locally optimal quantizer that minimizes certain distortion metric (e.g., convex) was discovered [21]. Thus, any locally optimal quantizer is the globally optimal quantizer in this scenario. However, the uniqueness of locally optimal quantizer that maximizes mutual information between input and quantized output is still an open problem. Thus, if one can find the conditions such that the locally optimal quantizer is *unique*, Algorithm 2 can determine the globally optimal quantizer in $O(NM + K\log^2(NM))$.

VII. Conclusion

In this paper, we proposed two efficient algorithms that can find the globally and locally optimal quantizers with the time complexities of $O(KM\log(NM))$ and $O(NM + K\log^2(NM))$ where N, M, K are the size of input X, received output U and quantized output V. Our techniques are based on the structure of optimal thresholding quantizer and fast root-finding bisection algorithms. Both theoretical and numerical results are presented to verify our contributions.

References

[1] Francisco Javier Cuadros Romero and Brian M Kurkoski. Decoding ldpc codes with mutual information-maximizing lookup tables. In *2015 IEEE International Symposium on Information Theory (ISIT)*, pages 426–430. IEEE, 2015.

[2] Brian M Kurkoski, Kazuhiko Yamaguchi, and Kingo Kobayashi. Noise thresholds for discrete ldpc decoding mappings. In *IEEE GLOBECOM 2008-2008 IEEE Global Telecommunications Conference*, pages 1–5. IEEE, 2008.

[3] Ido Tal and Alexander Vardy. How to construct polar codes. *arXiv preprint arXiv:1105.6164*, 2011.

[4] Jiuyang Alan Zhang and Brian M Kurkoski. Low-complexity quantization of discrete memoryless channels. In *2016 International Symposium on Information Theory and Its Applications (ISITA)*, pages 448–452. IEEE, 2016.

[5] Brian M Kurkoski and Hideki Yagi. Quantization of binary-input discrete memoryless channels. *IEEE Transactions on Information Theory*, 60(8):4544–4552, 2014.

[6] Harish Vangala, Emanuele Viterbo, and Yi Hong. Quantization of binary input dmc at optimal mutual information using constrained shortest path problem. In *2015 22nd International Conference on Telecommunications (ICT)*, pages 151–155. IEEE, 2015.

[7] Thuan Nguyen and Thinh Nguyen. Communication-channel optimized partition. *arXiv preprint arXiv:2001.01708*, 2020.

[8] Thuan Nguyen and Thinh Nguyen. Minimizing impurity partition under constraints. *arXiv preprint arXiv:1912.13141*, 2019.

[9] Thuan Nguyen and Thinh Nguyen. Optimal quantizer structure for binary discrete input continuous output channels under an arbitrary quantized-output constraint. *arXiv preprint arXiv:2001.02999*, 2020.

[10] Thuan Nguyen and Thinh Nguyen. Single-bit quantization capacity of binary-input continuous-output channels. *arXiv preprint arXiv:2001.01842*, 2020.

[11] Ken-ichi Iwata and Shin-ya Ozawa. Quantizer design for outputs of binary-input discrete memoryless channels using smawk algorithm. In *2014 IEEE International Symposium on Information Theory*, pages 191–195. IEEE, 2014.

[12] Xuan He, Kui Cai, Wentu Song, and Zhen Mei. Dynamic programming for discrete memoryless channel quantization. *arXiv preprint arXiv:1901.01659*, 2019.

[13] Rudolf Mathar and Meik Dörpinghaus. Threshold optimization for capacity-achieving discrete input one-bit output quantization. In *2013 IEEE International Symposium on Information Theory*, pages 1999–2003. IEEE, 2013.

[14] Stuart Lloyd. Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2):129–137, 1982.

[15] David Arthur and Sergei Vassilvitskii. How slow is the k-means method? In *Symposium on computational geometry*, volume 6, pages 1–10, 2006.

[16] Thuan Nguyen and Thinh Nguyen. On the uniqueness of binary quantizers for maximizing mutual information. *arXiv preprint arXiv:2001.01836*, 2020.

[17] Alok Aggarwal, Maria M Klawe, Shlomo Moran, Peter Shor, and Robert Wilber. Geometric applications of a matrix-searching algorithm. *Algorithmica*, 2(1-4):195–208, 1987.

[18] Thuan Nguyen and Thinh Nguyen. Entropy-constrained maximizing mutual information quantization. *arXiv preprint arXiv:2001.01830*, 2020.

[19] Xiaolin Wu. Optimal quantization by matrix searching. *Journal of algorithms*, 12(4):663–673, 1991.

[20] D Sharma. Design of absolutely optimal quantizers for a wide class of distortion measures. *IEEE Transactions on Information Theory*, 24(6):693–702, 1978.

[21] John Kieffer. Uniqueness of locally optimal quantizer for log-concave density and convex error weighting function. *IEEE Transactions on Information Theory*, 29(1):42–47, 1983.

[22] Thuan Nguyen, Yu-Jung Chu, and Thinh Nguyen. On the capacities of discrete memoryless thresholding channels. In *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, pages 1–5. IEEE, 2018.

[23] Brian M Kurkoski and Hideki Yagi. Single-bit quantization of binary-input, continuous-output channels. In *2017 IEEE International Symposium on Information Theory (ISIT)*, pages 2088–2092. IEEE, 2017.

[24] Kaj Madsen. A root-finding algorithm based on newton's method. *BIT Numerical Mathematics*, 13(1):71–75, 1973.

Appendix

A. Proof of Theorem 1

Due to limited space, we sketch the proof as follows. Noting that $I(X;V) = H(X) - H(X|V)$ and $H(X)$ is given, maximizing $I(X;V)$ is equivalent to minimizing $H(X|V)$ [4].

$$\begin{aligned} H(X|V) &= \sum_{i=1}^{K} p(v_i) H(X|v_i) \\ &= \sum_{i=1}^{K} p(v_i) \sum_{j=1}^{N} [-p(x_j|v_i)\log p(x_j|v_i)]. \end{aligned}$$

Using (2), (3) and (4) and taking $\dfrac{dH(X|V)}{dh_i}$ respect to variable h_i and set it to zero, we have

$$\sum_{j=1}^{N} \frac{p_(v_i|x_j)}{dh_i} \log p(x_j|v_i) + \sum_{j=1}^{N} \frac{p(v_{i+1}|x_j)}{dh_i} \log p(x_j|v_{i+1}) = 0. \tag{14}$$

However, from (2)

$$\frac{p(v_i|x_j)}{dh_i} = u_j(h_i), \tag{15}$$

$$\frac{p(v_{i+1}|x_j)}{dh_i} = -u_j(h_i). \tag{16}$$

Now, by substituting (15) and (16) into (14) and dividing both size to $\sum_{j=1}^{N} u_j(h_i)$, we have

$$\sum_{j=1}^{N} p(x_j|h_i) \log \frac{p(x_j|h_i)}{p(x_j|v_i)} = \sum_{j=1}^{N} p(x_j|h_i) \log \frac{p(x_j|h_i)}{p(x_j|v_{i+1})},$$

which is equivalent to (10).