Channel Quantizers that Maximize Random Coding Exponents for Binary-Input Memoryless Channels

Hideki Yagi
Center for Frontier Science and Engineering
The University of Electro-Communications
Chofu-shi, Tokyo 182-8585, Japan
Email: yagi@ice.uec.ac.jp

Brian M. Kurkoski
Dept. of Information and Communication Engineering
The University of Electro-Communications
Chofu-shi, Tokyo 182-8585, Japan
Email: kurkoski@ice.uec.ac.jp

***Abstract*—The problem of finding the optimum output quantizer for a given discrete memoryless channel is investigated, where the quantizer output has fewer values than the channel output. While mutual information has received attention as an objective function for optimization, the focus of this paper is use of the random coding exponent, which was originally derived by Gallager, as criteria. Two problems are addressed, where one problem is a partial problem of the other. The main result is a quantizer design algorithm, and a proof that it finds the optimum quantizer in the partial problem. The quantizer design algorithm is based on a dynamic programming approach, and is an extension of a mutual-information maximization method. For the binary-input case, it is shown that the optimum quantizer can be found with complexity that is polynomial in the number of channel outputs.**

I. Introduction

A fundamental problem for designers of digital circuits that implement communication algorithms is how to quantize numerical values with as few bits as possible. There is inevitably a performance-complexity tradeoff: decreasing the number of bits used to represent these values will decrease complexity, at the expense of performance. In practice, many circuit designers use fairly ad hoc approaches, for example, using a uniform quantizer and optimizing the step size.

Recently, however, there have been efforts to give an information theoretic foundation to the problem of channel quantization. The most fundamental question is how to optimally quantize the output of a channel, and recent work deals with optimality in the sense of maximizing *mutual information*. For the binary-input AWGN channel, it is straightforward to optimally quantize to three discrete output levels by maximizing over a single parameter [1]. For more than three outputs, non-uniform quantization provides higher mutual information than uniform quantization [2]. For a fixed quantizer, upper bounds on the capacity exist, and "locally optimal" quantization methods appear effective [3]. But by changing the focus from continuous-output channels to discrete output-channels, it becomes possible to find the globally optimal quantizer, using a dynamic programming approach [4]. Thus, mutual information is a well-studied metric for optimizing channel quantizers.

This research was supported in part by JST's Special Coordination Funds for Promoting Science and Technology, MEXT under Grant-in-Aid for Young Scientists (B) No. 22760270 and Grant-in-Aid for Scientific Research (C) No. 21560388.

The focus of this paper is an information theoretic metric not previously considered for channel quantization: the *random coding exponent*. There exists at least one code of block length n and rate R for which the probability of decoding error, P_{e}, can be upper bounded by:

$$P_{\mathrm{e}} \quad \leq \quad 2^{-nE_r(R)}, \tag{1}$$

where $E_r(R)$ is the random coding exponent [5]. It is well-known that the random coding exponent is a tight lower bound on the true error exponent for a range of rates near the channel capacity. The maximization of the random coding exponent may be interpreted as the minimization of the probability of decoding error for a given discrete memoryless channel (DMC), which may be a finely quantized version of a continuous-output channel, and a given R.

Whereas mutual information is a metric concerned with asymptotically long block lengths, the random coding error exponent allows us to form an upper bound on the error probability of maximum likelihood decoding for fixed block length codes, e.g., an ensemble of random codes [5]. When the code length is sufficiently large (but can be fixed), then the bound (1) also applies to an ensemble of non-binary low-density parity-check (LDPC) codes [6]. Since channel quantization algorithms can be applied to implementation of LDPC decoders [7], the random coding exponent may ultimately be a more suitable metric for finite-length codes.

The main result of this paper is to show how to find the quantizer for a DMC which maximizes metrics based on the random coding exponent. The quantizer reduces the number of discrete outputs to a smaller number of discrete quantizer outputs, for a DMC, an input distribution, and a rate R. Following previous work [4], a dynamic programming approach is used. However, since the metric is the random coding exponent and not mutual information, the core technical contribution of this paper is new necessary conditions on optimality. Whereas this paper is restricted to DMCs, a continuous output channel can be approximated with arbitrarily small discrepancy by a finely quantized channel.

It should be noted that in the 1960s and 1970s the *channel cutoff rate* was considered as an information theoretic criterion. In addition, mutual information can be expressed in terms of the random coding exponent, as will be discussed in the following section. From these observations, the problem considered in this paper can be viewed as a unified approach for the design of good quantizers. This is a step in the direction of a framework which can maximize the random

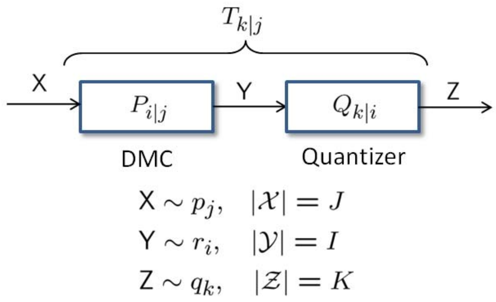

Fig. 1: A discrete memoryless channel followed by a quantizer. Given p_j, $P_{i|j}$, $0 \le \rho \le 1$, and K, find $Q_{k|i}$ which maximizes $\tilde{E}_0(\rho; Q)$.

coding exponent, the cut-off rate, and mutual information for any binary-input DMC.

The outline of this paper is as follows. In Sect. II background and the main result are given; the main result is a Theorem which states that the optimal quantizer can be found with polynomial complexity. Also, the relationship among the random coding exponent, the cut-off rate and mutual information is given. Then, Sect. III gives necessary conditions for optimality, which are needed in proving the Theorem. Sect. IV gives the quantizer design algorithm itself.

II. Main Results

A. Background

Let us consider a DMC followed by an output quantizer as shown in Fig. 1. Let X, Y and Z be random variables corresponding to the input, the output, and the quantized symbol, respectively. The alphabets of these random variables are denoted by $\mathcal{X}, \mathcal{Y}$, and $\mathcal{Z}$, and the alphabet sizes are J, I and K, respectively. Define probability distributions as follows:

$$\begin{aligned} p_j &= \Pr(\mathsf{X} = j), \quad j = 1, \ldots, J, \\ r_i &= \Pr(\mathsf{Y} = i), \quad i = 1, \ldots, I, \\ q_k &= \Pr(\mathsf{Z} = k), \quad k = 1, \ldots, K, \\ P_{i|j} &= \Pr(\mathsf{Y} = i | \mathsf{X} = j), \\ Q_{k|i} &= \Pr(\mathsf{Z} = k | \mathsf{Y} = i), \text{ and} \\ T_{k|j} &= \Pr(\mathsf{Z} = k | \mathsf{X} = j) = \sum_i Q_{k|i} P_{i|j}. \end{aligned}$$

We denote probability distributions $\{p_j | j = 1, \ldots, J\}$ and $\{Q_{k|i} | k = 1, \ldots, K, i = 1, \ldots, I\}$ by p and Q, respectively. The sum $\sum_i$, etc. means the sum over the whole alphabet $\sum_{i=1}^{I}$, etc. Except for Lemma 2 (given in the next section), this paper makes the restriction that $J = 2$.

B. Random Coding Exponent and Main Result

For a given DMC, an input distribution p_j, and a rate R, define

$$E_0(\rho, T) = -\log \sum_k \Big[\sum_j p_j T_{k|j}^{\frac{1}{1+\rho}} \Big]^{1+\rho} \tag{2}$$

for $0 \le \rho \le 1$. This function was introduced by Gallager [5], and is sometimes called the *Gallager function*. It has been shown that $E_0(\rho, T)$ is a minus logarithm of a function which is convex (lower convex) in p_j and concave (upper convex) in $T_{k|j}$. Since the probability distribution of the quantized channel T is a function of the quantizer Q, the Gallager function can also be expressed as

$$\tilde{E}_0(\rho, Q) = -\log \sum_k \Big[\sum_j p_j \Big(\sum_i P_{i|j} Q_{k|i} \Big)^{\frac{1}{1+\rho}} \Big]^{1+\rho} \tag{3}$$

for $0 \le \rho \le 1$. Since any affine transform of an argument does not affect the convexity (concavity) of the function, this function is also also a minus logarithm of a concave function in $Q_{k|i}$ for fixed $P_{i|j}$. The random coding exponent is defined as

$$E_r(R) = \max_p \max_{0 \le \rho \le 1} \big\{ \tilde{E}_0(\rho, Q) - \rho R \big\}, \tag{4}$$

where the first maximization is taken over all input probability distributions. For fixed p, ρ and R, the random coding exponent for the original DMC is $E_0(\rho, P) - \rho R$, whereas that for the quantized channel is $\tilde{E}_0(\rho, Q) - \rho R$. For simplicity, we also call $\tilde{E}_0(\rho, Q) - \rho R$ the random coding exponent for fixed p, ρ, and R. For any $0 \le \rho \le 1$, there is a relation between these exponents,

$$E_0(\rho, P) - \rho R \ge \tilde{E}_0(\rho, Q) - \rho R. \tag{5}$$

This is confirmed as follows: first, we define

$$F_0(\rho, P) = \sum_i \Big[\sum_j p_j P_{i|j}^{\frac{1}{1+\rho}} \Big]^{1+\rho}. \tag{6}$$

By a variant of Minkowski's inequality [5, p. 524], we can easily show

$$F_0(\rho, P) \quad \le \quad \sum_k \Big[\sum_j p_j \Big(\sum_i Q_{k|i} P_{i|j} \Big)^{\frac{1}{1+\rho}} \Big]^{1+\rho}. \tag{7}$$

Taking the minus logarithm on both sides, we obtain (5).

Our ultimate goal is to find the quantizer Q which maximizes $E_r(R)$ for a given R. The key lies in the derivation of a solution in a simplified problem; for given P, p, R, and K, find the quantizer Q which maximizes the random coding exponent $\tilde{E}_0(\rho, Q) - \rho R$ for any given ρ. This problem is equivalent to maximizing the Gallager function $\tilde{E}_0(\rho, Q)$ itself when R and ρ are fixed.

The following theorem is the main result of this paper.

Theorem. For any binary-input DMC, there exists an algorithm with complexity at most I^3 that finds the quantizer Q^* satisfying

$$Q^* = \arg\max_{Q \in \mathcal{Q}} \tilde{E}_0(\rho, Q), \tag{8}$$

for any given input probability distribution p and $0 \le \rho \le 1$.

Remark 1. In order to prove the Theorem, we will show Lemma 3 in Sect. III-B. Using Lemma 3, it can also be shown that the number of candidate solutions for the quantizer that maximizes $E_r(R)$ for a given R is less than I^{K-1}, so we can find such a quantizer with polynomial-order complexity of I. From this perspective, the simplified problem for maximizing the Gallager function is fundamental, and so will be considered in detail. ■

C. Connection with Other Criteria

While this paper concentrates on the random coding exponent, in this subsection the close relationship among the random coding exponent, the cut-off rate and mutual information is illustrated.

It has been shown in [5] that

$$I(\mathsf{X};\mathsf{Z}) = \lim_{\rho \to +0} \frac{\tilde{E}_0(\rho, Q)}{\rho}. \tag{9}$$

The exponent $\tilde{E}_0(\rho, Q) - \rho R$ for a fixed ρ is decreasing in R, and as R increases, the optimum ρ^* that maximizes the random coding exponent (4) monotonously decreases. When R reaches the mutual information $I(\mathsf{X};\mathsf{Z})$, then the optimum ρ^* is zero. When we let ρ be sufficiently small, our problem corresponds to maximizing mutual information.

The channel cutoff rate was suggested as a criterion for designing quantizers by Wozencraft and Kennedy [8] in the 1960s. Massey [9] went on to emphasize the importance of the cutoff rate as a receiver design criterion, and gave an algorithm to find a channel quantizer which maximizes the cutoff rate for the binary-input AWGN channel. Lee [10] extended these results to channels with non-binary inputs. For a given DMC, the cut-off rate for an input distribution p_j is defined as

$$R_0(Q) = -\log \sum_k \Big[\sum_j p_j \Big(\sum_i P_{i|j} Q_{k|i} \Big)^{\frac{1}{2}} \Big]^2. \tag{10}$$

It is clear that restricting our problem to the case $\rho = 1$ reduces the problem to maximizing the cut-off rate.

III. Concave Minimization and a Necessary Condition for Optimality

A. Concave Minimization

Since only Q can be altered for maximization of $\tilde{E}_0(\rho, Q)$ in this problem, it suffices to find a quantizer Q which minimizes

$$\tilde{F}_0(\rho, Q) = \sum_k \Big[\sum_j p_j \Big(\sum_i Q_{k|i} P_{i|j} \Big)^{\frac{1}{1+\rho}} \Big]^{1+\rho}. \tag{11}$$

The function $\tilde{F}_0(\rho, Q)$ is regarded as a new objective function to be *minimized*. Thus, we want to find

$$Q^* = \arg \min_{Q \in \mathcal{Q}} \tilde{F}_0(\rho, Q), \tag{12}$$

where the left hand side is equal to that of (8). The new objective function $\tilde{F}_0(\rho, Q)$ is concave in $Q_{k|i}$. The problem considered here is *concave minimization* [11].

Lemma 1. [11, Theorem 1.19] A concave (convex) function $f : S \to \mathbb{R}$ attains its global minimum (maximum) over a convex set S at an extreme point of S.

Thus, in the context of concave minimization, the optimal solution is one of extreme points of the (convex) feasible region. Since the feasible region of our problem is the set of all probability distributions $\mathcal{Q}$, its extreme points correspond to $I \times K$ matrices each of whose rows has only one element 1. We shall refer to these matrices as *deterministic matrices*. We have the following lemma.

Lemma 2. For any DMC and any K, the optimal quantizer Q^* is deterministic. That is, $Q^*_{k|i} \in \{0, 1\}$, for all i and k.

Proof: The lemma can be proved by considering the fact that all the extreme points of the feasible region correspond to deterministic matrices. ■

Thus, it suffices to find an optimum quantizer in the class of deterministic quantizers, which has also been shown in [12] for the AWGN channel. To illustrate the implication of Lemma 2, we show the following example.

Example 1. Consider the binary-input errors and erasure channel, with the transition matrix:

$$P = \begin{bmatrix} 1-a-b & a & b \\ b & a & 1-a-b \end{bmatrix}, \tag{13}$$

for $a, b \geq 0$ and $a + b \leq 1$. This channel is *symmetric* in the sense of [5], so the uniform input distribution gives the largest random coding exponent and mutual information for the original DMC. Suppose the three outputs, called 0, erasure and 1, are to be quantized to two levels ($K = 2$). One might expect that symmetry should be maintained by mapping the erasure symbol to the two output symbols with probability 0.5 each. However, as Lemma 2 shows, this probabilistic assignment has larger objective function (i.e., smaller random coding exponent) than mapping the erasure symbol to either 0 or 1 with probably one. Thus, there are two optimum quantizers; one quantizer divides the output alphabet as $\{1, 2\}$ and $\{3\}$ and the other quantizer divides the output alphabet as $\{1\}$ and $\{2, 3\}$. These optimal quantizers lack symmetry between the channel input and quantizer output. If one wants to maintain the symmetry, one suggestion is to *time-share* these optimum quantizers with the same fraction of time. It is easy to see that probability distribution of the quantizer outputs is uniform in this case. ■

The number of deterministic quantizers (deterministic matrices) is K^I. A naive optimization approach would be to search over all K^I candidate solutions, which is searching over all deterministic quantizers. This has complexity which is exponential in I. However, a more efficient algorithm exists for maximizing the mutual information for the uniform input distribution as shown in [13], and it will be shown that this also applies to the problem considered in this paper; maximizing the Gallager function with non-uniform input distributions. The key observations are (i) the same necessary condition for optimum quantizers under the mutual information criteria is valid, and (ii) the efficient algorithm developed in [13], can also be used to minimize the Gallager function.

B. Necessary Condition for Optimality

Without any loss of generality for binary-input channels, it is assumed that the outputs of the DMC are indexed in such a way that the following is satisfied:

$$\frac{P_{1|1}}{P_{1|2}} < \frac{P_{2|1}}{P_{2|2}} < \cdots < \frac{P_{I-1|1}}{P_{I-1|2}} < \frac{P_{I|1}}{P_{I|2}}. \tag{14}$$

Here, strict inequalities are used because if $\frac{P_{i_0|1}}{P_{i_0|2}} = \frac{P_{i_0+1|1}}{P_{i_0+1|2}}$ for some $i_0 \in \mathcal{Y}$, then outputs i_0 and $i_0 + 1$ can be combined to a single output with $\frac{P_{i_0|1}+P_{i_0+1|1}}{P_{i_0|2}+P_{i_0+1|2}}$, to form a new channel with $I-1$ outputs. The original channel and the new channel have the same value of the objective function.

Lemma 3 stated below gives a necessary condition for the quantizer to be optimal. It is key for proving that the quantizer design algorithm proposed in [13], produces the optimal quantizer with the new objective function $\tilde{F}(\rho, Q)$, as well. Let $\mathcal{A}_k, k = 1, \ldots, K$, be the set of channel output symbols quantized to k.

Lemma 3. For the quantizer which minimizes the objective function $\tilde{F}(\rho, Q)$, $\mathcal{A}_k, k = 1, \ldots, K$, consists of consecutive channel outputs, when the channel outputs are sorted according to (14).

Proof: The quantization problem considered here is an example of *impurity-minimization partitions* from machine learning [15]. Defining the backward channel $\tilde{P}_{j|i} = \Pr(\mathsf{X} = j|\mathsf{Y} = i)$ for the given DMC $P_{i|j}$ and a fixed input distribution p_j, the sorting condition (14) can also be expressed as

$$\tilde{P}_{1|1} < \tilde{P}_{1|2} < \cdots < \tilde{P}_{1|I}. \tag{15}$$

Theorem 1 in [15] states that, if the objective function is concave in $\tilde{T}_{j|k} = \Pr(\mathsf{X} = j|\mathsf{Z} = k)$, then the optimal K partitions consist of K convex subsets of $\{\tilde{P}_{1|i} \mid i = 1, \ldots, I\}$. Here, a partition corresponds to the set of channel outputs that map to a single quantizer output. It is easily shown that the objective function $\tilde{F}(\rho, Q)$ is concave in $\tilde{T}_{j|k}$ using a variant of Minkowski's inequality [5, p. 524]. Restricting to binary-input DMCs, it is clear that the lemma holds. ■

Lemma 3 implies that, if (14) holds, then each quantizer output of the optimum quantizer consists of consecutive channel outputs. We call these quantizers *consecutive quantizers*. It is readily seen that the number of consecutive quantizers is $\binom{I-1}{K-1}$, and a brute-force search requires $O(I^{K-1})$ complexity. Since the number of deterministic quantizers is $O(K^I)$, this reduces the candidate solutions significantly (recall that we consider the situation in which I is large while K is relatively small). When K is large, such a brute-force search becomes infeasible. In the next section, we show that an algorithm for the finding optimum quantizer with complexity at most $O(I^3)$.

Remark 2. We use the fact that the channel quantization problem can be viewed as an instance of impurity-minimization partitions from machine learning. While restricted to binary-input DMCs, one can extend the claim of Lemma 3 to memoryless channels with any finite input alphabet and (possibly continuous) output alphabet using [15, Theorem1]. ■

Remark 3. Given the parameters p^* and ρ^* that gives $E_r(R)$ for a fixed R, the addressed problem is equivalent to finding the quantizer that maximizes $E_r(R)$, and hence, the optimum quantizer should also be a consecutive quantizer from Lemma 3. The number of candidate solutions is also $\binom{I-1}{K-1} \leq I^{K-1}$. For each candidate Q, we can perform the maximization in ρ by taking the partial derivative [5, Sect. 5.6] and in p by the Arimoto-Blahut algorithm [16], [17]. The complexity for this maximization does not depend on I, and we can find the quantizer that maximizes $E_r(R)$ in $O(I^{K-1})$ time. ■

C. Examples

We show some examples of optimum quantizers that maximize the objective function and mutual information [13] for a given DMC.

Consider a class of binary-input DMCs with $I = 5$, and the output symbols are quantized into $K = 3$ quantizer outputs. According to Lemma 3, the set of candidates for the optimum quantizer is $\{Q_1, ..., Q_6\}$ shown in Table I. Here, we have $\binom{I-1}{K-1} = 6$ quantizers satisfying the necessary condition of Lemma 3.

TABLE I: The set of candidate quantizers with $I = 5$ and $K = 3$ satisfying the necessary condition given by Lemma 3.

Quantizer	$\mathcal{A}_1$	$\mathcal{A}_2$	$\mathcal{A}_3$
Q_1	{1,2,3}	{4}	{5}
Q_2	{1,2}	{3,4}	{5}
Q_3	{1,2}	{3}	{4,5}
Q_4	{1}	{2,3,4}	{5}
Q_5	{1}	{2,3}	{4,5}
Q_6	{1}	{2}	{3,4,5}

TABLE II: The optimum quantizers under the two criteria with the objective value $\tilde{E}_0(\rho, Q) - \rho R$ (denoted by "RCE val.") with $\rho = 1/2$. "MI criterion" stands for the mutual information criterion, while "RCE criterion" means the random coding exponent criterion. The rightmost column shows the difference of $\tilde{E}_0(\rho, Q) - \rho R$.

DMC #	MI criterion		RCE criterion		difference
	quantizer	RCE val.	quantizer	RCE val.	
1	Q_3	0.09272	Q_2	0.10804	0.01532
2	Q_2	0.07598	Q_4	0.09276	0.01678
3	Q_6	0.09857	Q_5	0.11338	0.01481
4	Q_2	0.05333	Q_6	0.06418	0.01084

We randomly generate DMCs satisfying the sorting condition (14), and we show examples in which the optimum quantizer that maximizes the objective function $\tilde{E}_0(\rho, Q) - \rho R$ is different from the optimum one that maximizes mutual information $I(\mathsf{X}; \mathsf{Z})$. In order to see the effect of two different criteria, we choose $R = I(\mathsf{X}; \mathsf{Z})/2$ and $\rho = 1/2$ for convenience, where $I(\mathsf{X}; \mathsf{Z})$ is attained by the optimum quantizer under the mutual information criterion. In Table II, the optimum quantizers under the two criteria are shown with their objective value $\tilde{E}_0(\rho, Q) - \rho R$. The rightmost column shows the difference of $\tilde{E}_0(\rho, Q) - \rho R$ for respective optimum quantizers, and these values show that the error probability of ML decoding for the optimum quantizer under the random coding exponent criterion decays much faster than that for the optimum quantizer under the mutual information criterion, as the block length n increases. To show the difference between the optimum quantizers under the two criteria, the curves for $\tilde{E}_0(\rho^*, Q^*) - \rho^* R$ are depicted in Fig. 2. Here, the maximization in ρ is performed for the quantizers shown in Table II at each rate.

IV. Efficient Quantizer Design Algorithm

An effective algorithm for finding the optimum quantizer requiring time and space complexity at most $O(I^3)$ is presented. Before presenting the algorithm, we give some preliminaries.

A. Partial Objective Function

For a deterministic quantizer, which is of interest due to Lemma 2, the objective function can be expressed as

$$\tilde{F}_0(\rho, Q) = \sum_k \left[\sum_j p_j \Big(\sum_{i \in \mathcal{A}_k} P_{i|j} \Big)^{\frac{1}{1+\rho}} \right]^{1+\rho}, \tag{16}$$

since $Q_{k|i} = 1$ iff $i \in \mathcal{A}_k$.

Under the quantization mapping from channel outputs to quantizer outputs, the preimage of quantizer output m is $\mathcal{A}_m$. The partial objective function ι_m for this output is:

$$\iota_m = \left[\sum_j p_j \Big(\sum_{i \in \mathcal{A}_m} P_{i|j} \Big)^{\frac{1}{1+\rho}} \right]^{1+\rho}, \tag{17}$$

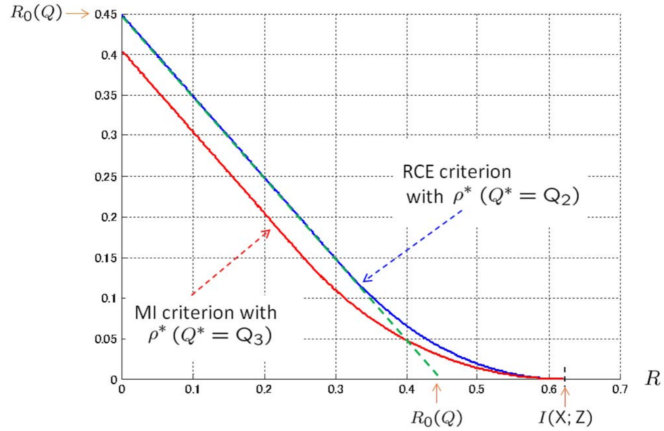

Fig. 2: An example of exponents $\tilde{E}_0(\rho^*, Q^*) - \rho^* R$ for optimum quantizers over DMC 1 under the two criteria. The maximization in ρ is performed for two quantizers at each rate.

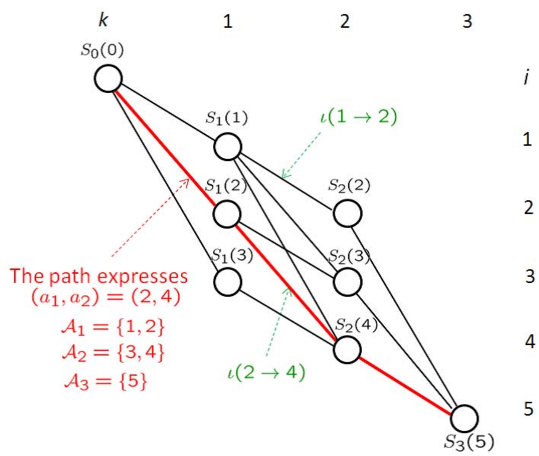

Fig. 3: Trellis-type illustration showing the relationship between state metrics for $I = 5$ and $K = 3$.

so the total objective function is

$$\tilde{F}_0(\rho, Q) = \sum_k \iota_k. \tag{18}$$

Further, let consecutive channel outputs $a' + 1$ to a, with $a' < a \leq I$, be assigned to a single quantizer output k. That is, $\mathcal{A}_k = \{a_{k-1} + 1, \ldots, a_k\}$. For these sets of consecutive outputs, we denote by $\iota(a_{k-1} \rightarrow a_k)$ the partial objective unction ι_k, i.e.,

$$\iota(a_{k-1} \rightarrow a_k) = \iota_k \tag{19}$$

for this assignment. The values of $\iota(a \rightarrow a')$ for all a, a' such that $1 \leq a < a' \leq I$ will be used as metrics of a quantizer design algorithm.

B. Quantizer Design Algorithm

The Quantization Algorithm, proposed in [13], is a quantizer design algorithm and is the realization of a dynamic program. Although the metrics here are different from those in [13], we do not need to change the algorithm itself. Here, we briefly review the algorithm.

The algorithm has a state value $S_k(i)$, which is the minimum value of the partial objective function when channel outputs 1 to i are quantized to quantizer outputs 1 to k. This can be computed recursively by conditioning on the state value at time index $k - 1$:

$$S_k(a) = \min_{a'} \Big(S_{k-1}(a') + \iota(a' \rightarrow a) \Big), \tag{20}$$

where the minimization is taken over $a' \in \{k-1, \ldots, a-1\}$. Clearly, $S_K(I)$ is the minimum value of the total objective function. The path $S_0(0), S_1(a_1), \ldots, S_K(a_K)$ which gives the minimum of the total objective function corresponds to the optimum quantizer whose boundaries are $\{a_1, a_2, \ldots, a_K\}$. The relationship between the states metrics are illustrated in a trellis-type diagram in Fig. 3, for $I = 5$ and $K = 3$. In the trellis-type diagram, a metric value $\iota(a' \rightarrow a)$ calculated by (17) and (19) is assigned to the edge from $S_k(a')$ to $S_{k+1}(a)$ for all pair (a', a) such that $a' \in \{0, 1, \ldots, I-1\}$ and $a \in \{a'+1, \ldots, t\}$ (where $t = \min\{a'+1+I-K, I\}$) at each $k = 1, \ldots, I-1$. The flow of the algorithm is the same as in [13], and hence is omitted.

The complexity of this algorithm is not greater than I^3 [4], and this complexity reduces that required for a brute-force approach, which is $O(I^{K-1})$ from Lemma 3. This complexity result, along with the proof of optimality, proves the Theorem.

References

[1] X. Ma, X. Zhang, H. Yu, and A. Kavcic, "Optimal quantization for soft-decision decoding revisited," in *International Symposium on Information Theory and its Applications*, (Xian, China), Oct. 2002.

[2] A. D. Liveris and C. N. Georghiades, "On quantization of low-density parity-check coded channel measurements," in *Proceedings IEEE Global Telecommunications Conference*, (San Francisco, CA, USA), pp. 1649–1653, Dec. 2003.

[3] J. Singh, O. Dabeer, and U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Trans. on Commun.*, vol. 57, pp. 3629–3639, Dec. 2009.

[4] B. M. Kurkoski and H. Yagi, "Quantization of binary-input discrete memoryless channels, with applications to LDPC decoding," available http://arxiv.org/abs/1107.5637.

[5] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: Wiley, 1968.

[6] A. Bennatan and D. Burshtein, "On the application of LDPC codes to arbitrary discrete-memoryless channels," *IEEE Trans. on Inform. Theory*, vol. 50, pp. 417 – 438, Mar. 2004.

[7] B. Kurkoski, K. Yamaguchi, and K. Kobayashi, "Noise thresholds for discrete LDPC decoding mappings," in *Proceedings IEEE Global Telecommunications Conference*, (New Oreleans, USA), pp. 1–5, Nov.–Dec. 2008.

[8] J. M. Wozencraft and R. S. Kennedy, "Modulation and demodulation for probabilistic coding," *IEEE Trans. on Inform. Theory*, vol. 12, pp. 291–297, July 1966.

[9] J. L. Massey, "Coding and modulation in digital communications," in *Proceedings of International Zurich Seminar on Digital Communications*, (Zurich, Switzerland), pp. E2(1)–E2(4), 1974.

[10] L. Lee, "On optimal soft-decision demodulation," *IEEE Trans. on Inform. Theory*, vol. 22, pp. 437 – 444, July 1976.

[11] R. Horst, P. M. Pardalos, and N. V. Thoai, *Introduction to Global Optimization*. Kluwer Academic Publishers, 1995.

[12] J. Singh, O. Dabeer, and U. Madhow, "Communication limits with low precision analog-to-digital conversion at the receiver," in *Proceedings IEEE International Conference on Communications*, (Glasgow, Scotland), pp. 6269–6274, Jun. 2007.

[13] B. Kurkoski and H. Yagi, "Concatenation of a discrete memoryless channel and a quantizer," in *Proceedings of the IEEE Information Theory Workshop*, (Cairo, Egypt), pp. 160–164, Jan. 2010.

[14] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[15] D. Burshtein, V. D. Pietra, D. Kanevsky, and A. Nadas, "Minimum impurity partitions," *The Annals of Statistics*, vol. 20, no. 3, pp. 1637–1646, 1992.

[16] S. Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Trans. on Inform. Theory*, vol. 18, pp. 14–20, Jan. 1972.

[17] R. E. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Trans. on Inform. Theory*, vol. 18, pp. 460–473, July 1972.