Rate-Constrained Shaping Codes for Structured Sources

Yi Liu, *Student Member, IEEE,* Pengfei Huang, *Member, IEEE,*
Alexander W. Bergman, and Paul H. Siegel, *Life Fellow, IEEE*

***Abstract*—Shaping codes are used to encode information for use on channels with cost constraints. Applications include data transmission with a power constraint and, more recently, data storage on flash memories with a constraint on memory cell wear. In the latter application, system requirements often impose a rate constraint. In this paper, we study rate-constrained fixed-to-variable length shaping codes for noiseless, memoryless costly channels and general i.i.d. sources. The analysis relies on the theory of word-valued sources. We establish a relationship between the code expansion factor – the ratio of the expected codeword length to the length of the input source word – and the minimum average symbol cost. We then determine the expansion factor that minimizes the average cost per source symbol (total cost), corresponding to a conventional optimal source code with cost. An equivalence is established between codes minimizing average symbol cost and codes minimizing total cost, and a separation theorem is proved, showing that optimal shaping can be achieved by a concatenation of optimal compression and optimal shaping for a uniform i.i.d. source. Shaping codes often incorporate, either explicitly or implicitly, some form of non-equiprobable signaling. We use our results to further explore the connections between shaping codes and codes that map a sequence of i.i.d. source symbols into an output sequence of symbols that are approximately independent and distributed according to a specified target distribution, such as distribution matching (DM) codes. Optimal DM codes are characterized in terms of a new performance measure - generalized expansion factor (GEF) - motivated by the costly channel perspective. The GEF is used to study DM codes that minimize informational divergence and normalized informational divergence.**

***Index Terms*—Source coding, flash memory, data compression, costly channel, shaping codes, distribution matching.**

Portions of this paper were presented at the 8th Annual Non-Volatile Memories Workshop, La Jolla, CA, March 12–14, 2017, the IEEE International Symposium on Information Theory, Aachen, Germany, June 25-30, 2017, and the 9th Annual Non-Volatile Memories Workshop, La Jolla, CA, March 11–13, 2018.

Y. Liu and P. H. Siegel are with the Center for Memory and Recording Research, Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA 92093, USA (e-mail: yil333@ucsd.edu; psiegel@ucsd.edu).

P. Huang was with the Center for Memory and Recording Research, Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA 92093, USA. He is now with Western Digital Corporation, Milpitas, CA 95035, USA (e-mail: pehuangucsd@gmail.com).

A. W. Bergman was with the Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA 92093, USA. He is now with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA (e-mail: awb@stanford.edu).

I. Introduction

Shaping codes are used to encode information for use on channels with a cost constraint. A prominent application is in data transmission with a power constraint, where constellation shaping is achieved by addressing into a suitably designed multidimensional constellation or, equivalently, by incorporating, either explicitly or implicitly, some form of non-equiprobable signaling. An excellent reference on this topic is Fischer [14].

More recently, shaping codes have been proposed for use in data storage on flash memories subject to a constraint on memory cell wear. In that application, storage system requirements often impose a rate constraint, and the data source may be structured, rather than unconstrained. Motivated by this scenario, this paper investigates information-theoretic properties and design of rate-constrained fixed-to-variable length shaping codes for noiseless, memoryless costly channels and general i.i.d. sources. The analysis relies on the theory of word-valued sources developed in Nishiara and Morita [45]. Our primary interest is in the design of codes that minimize the average cost per code symbol for a given expansion factor – the ratio of the expected codeword length to the length of the input source word – which we refer to as the *type-I shaping problem*. We also consider the well-studied problem of designing codes that minimize average cost per code symbol, or *total cost*, which we refer to as the *type-II shaping problem*.

The word-valued source analysis provides a natural link between shaping codes and codes that efficiently map a sequence of i.i.d. source symbols into an output sequence of symbols that are approximately independent and distributed according to a specified target distribution. Such codes have been studied in the context of random number generating source codes by Han and Uchida [22] and as distribution matching (DM) codes by Böcherer and Mathar [8], Böcherer [5], Amjad and Böcherer [3], Böcherer and Amjad [6], Schulte and Böcherer [51] and Schulte and Steiner [52]. Our shaping code analysis suggests a new performance measure - generalized expansion factor (GEF) - for fixed-to-variable length DM codes which we use to study codes that minimize informational divergence and normalized informational divergence from a shaping code perspective.

There is a substantial literature on shaping codes and, more recently, a body of work relating to DM codes. Therefore, before summarizing our results in more detail, we provide a brief review of relevant work in both of these areas as a

framework for our contributions.

A. Shaping Codes

1) Codes minimizing total cost: The problem of coding for noiseless costly channels, or source coding with unequal symbol costs, traces its conceptual origins to Shannon's 1948 paper that launched the study of information theory [53]. In that paper, Shannon considered the problem of transmitting information over a telegraph channel. The channel symbols – dots and dashes – have different time durations, which can be interpreted as transmission costs. Shannon determined the symbol probabilities that maximize the data transmission rate with integer symbol costs. This result was then generalized to arbitrary positive symbol costs by Krause [31] and Csiszár [12].

Several researchers have considered the problem of designing codes for costly channels with an i.i.d. source. Most of this work has emphasized construction of codes that minimize average cost per source symbol, which we refer to as *total cost*, without an explicit rate constraint. In Karp [27], costly channel coding was studied from an algebraic perspective, and the problem of designing a shaping code to minimize total cost was recast as an integer programming problem. However, this code design approach is not computationally practical, and the algorithm proposed to reduce the complexity will result in sub-optimal results. In Golin et al. [18], a dynamic programming solution for this integer programming problem was proposed, providing a polynomial time bound on complexity. Other approaches using tree-based constructions were proposed in [31], Melhorn [43], and Csiszár and Körner [13]. They all constructed asymptotically optimal prefix-free variable-length shaping codes. A universal coding scheme based on types was also introduced in [13].

A special case, corresponding to a uniform i.i.d. source in which all codewords are equally likely to occur, was studied by Varn [58], who proposed a variable-length code construction that minimizes the average codeword cost for a fixed codebook size. This coding technique was then incorporated into a universal coding scheme in Iwata [24], which combines LZ78 compression with Varn coding. Later in this paper, we generalize the Iwata scheme, which can be viewed as an embodiment of a separation theorem proved in Section III, and further explore properties and applications of Varn codes.

A generalization of Huffman coding for unequal symbol costs was proposed in Gilbert [17]. In Guazzo [19], practical arithmetic coding was introduced. This coding technique was then generalized by Savari and Gallager [50] and its properties, such as optimality and coding delay, were analyzed. However, the analysis is based on infinite precision arithmetic coding, which cannot be realized in practice.

In Böcherer and Mathar [8] and Böcherer [5], a variable-to-fixed length code construction called geometric Huffman coding was used to design codes for an i.i.d. uniform source that asymptotically minimize the total cost of a noiseless channel with unequal symbol durations. (This construction matches codeword probabilities to dyadic symbol distributions that optimally approximate the optimal symbol distribution.)

We emphasize that all of the codes mentioned above considered the problem of minimizing cost per source symbol, i.e., total cost, with no explicit consideration of rate. The dependence of total cost on code rate was not thoroughly investigated.

2) Rate-constrained codes minimizing average cost: The problem of designing rate-constrained codes for costly channels has received less attention. For a finite-state channel with associated symbol/transition costs and an average cost constraint, the maximum entropy of a stationary Markov chain, along with the entropy-maximizing symbol/transition probabilities, can be found in McEliece and Rodemich [41], Justesen and Høholdt [26], and Khandekar, McEliece, and Rodemich [28]. In McEliece [40], the special case corresponding to a memoryless channel is addressed. Later in Böcherer [5], both memoryless channels and generalizations to channels with memory were addressed.

Böcherer and Mathar [8] and Böcherer [5] apply the geometric Huffman coding approach to design variable-to-fixed length codes that match codeword probabilities to dyadic symbol distributions that approximate the entropy-maximizing probability mass function for memoryless costly channels subject to an average cost constraint, thereby asymptotically achieving the maximum rate.

The state-splitting algorithm [2], which was developed to construct finite-state codes for constrained channels, has been extended for application to construction of codes for costly channels. Heegard, Marcus, and Siegel [23] studied a class of channels with average runlength constraints, which represent a special case of noiseless channels with a cost constraint. They constructed variable-to-variable length synchronous codes using state-splitting techniques adapted for channels with variable-length symbols. Khayrallah and Neuhoff [29] and McLaughlin and Khayrallah [42] construct fixed-to-fixed length and variable-to-fixed length codes based on state-splitting methods for magnetic recording and constellation shaping applications. Krachkovsky et al. [30] determine a costly channel model matched to a Markov source and construct corresponding codes using enumerative techniques for application to transmission over an intersymbol-interference channel. All of these works strive to construct codes that come close to the capacity-cost functions originally presented in [40], [41], and [26].

Other recent work relating to this problem has been motivated by non-volatile memory applications, so we briefly describe the corresponding costly channel model. NAND flash memory uses floating-gate transistors, commonly referred to as *cells*, to store information in the form of different cell voltage levels. The flash memory cells gradually wear out with repeated writing and erasing, referred to as program/erase cycling, and the damage caused by the cycling is dependent on the programmed voltage levels [34], [35]. The costly channel model associates to each cell voltage level a wear cost reflecting

the extent of the damage induced by writing that level.

Recently, in [25], Jagmohan et al. proposed *endurance coding*, intended for shaping of programmed data for flash memories. For a given cost model and a specified target code rate, the optimal distribution of cell levels that minimizes the average cost was determined analytically, reproducing the results in the references cited above. For single bit per cell (SLC) flash memory, with associated level costs of 0 and 1, greedy enumerative codes that minimize the number of cells with cost 1 were designed and evaluated in terms of the rate-cost trade-off. However, endurance coding is intended for uniform i.i.d. source data. For structured source data, which would include a general i.i.d. source, the idea of combining source compression with endurance coding was proposed, but the relationship between the code performance and the code rate for arbitrary sources was not thoroughly studied.

In Sharon et al. [54], low-complexity, rate-1, fixed-length *direct shaping codes* for structured data were proposed for use on SLC flash memory. The code construction used a greedy approach based upon an adaptively-built encoding dictionary that does not require knowledge of the source statistics. This construction was extended to a direct shaping code compatible with two-bit per cell (MLC) flash memory operation by Liu et al. in [34], [35]. However, it was proved in Liu and Siegel [37] that direct shaping codes are in general suboptimal. (Our experimenal results in Section VII contain a comparison of a shaping scheme motivated by our analysis to a direct shaping code on MLC flash memory.)

3) Summary of contributions on shaping codes: In this paper, our goal is to systematically study the fundamental performance limits of fixed-to-variable length shaping codes from a rate and distribution perspective. We first use known properties of word-valued sources to determine the symbol occurrence probability of shaping code output sequences (Lemma 4). We then derive an upper bound on the code sequence entropy rate (Lemma 5). Using these results, we are able to reduce the problem of minimizing average code symbol cost subject to a constraint on the code rate to an optimization problem for an i.i.d. process. This problem can be viewed as the dual problem to the entropy-maximization problem considered in the prior literature. We refer to this minimization problem as the *type-I shaping problem*, and we call shaping codes that achieve the minimum average cost for a given rate *optimal type-I shaping codes.* We develop a theoretical bound on the trade-off between the rate – or more precisely, the corresponding *expansion factor* – and the average cost of a type-I shaping code (Theorem 6). We then study shaping codes that minimize total cost (minimum average cost per source symbol). We refer to the problem of minimizing the total cost as the *type-II shaping problem* and shaping codes that achieve the minimum total cost are referred to as *optimal type-II shaping codes*. We derive the relationship between the code expansion factor and the total cost and determine the optimal expansion factor (Theorem 7). We then prove an equivalence theorem showing that an optimal type-I shaping code can be realized using an optimal type-II shaping code for another suitably chosen costly channel model (Theorem 8). We can therefore solve the type-I shaping problem using known coding techniques such as generalized Shannon-Fano codes [13] . A consequence of the analysis is a separation theorem for type-II shaping codes, which states that optimal shaping can be achieved by a concatenation of lossless compression and optimal shaping for a uniform i.i.d. source. This provides an alternative architecture for implementing asymptotically optimal shaping codes using, for example, Varn codes. Finally, we prove a separation theorem for type-I shaping codes with given expansion factor, using a careful analysis of the behavior of the minimum average cost as a function of the expansion factor.

B. Distribution Matching (DM) Codes

1) Applications of DM codes to shaping: The application of non-equiprobable signaling in the context of coding with a cost constraint reflects the interesting interplay between shaping codes and DM-type codes (in the broad sense of codes that map an i.i.d. sequence of source symbols to an output sequence of symbols that are approximately independent and distributed according to $\{P_i\}$). Beginning with the work on constellation shaping, there have been a number of applications of DM-type codes to coding for a costly channel.

In [15], signal constellations with non-uniform symbol probabilities were used for efficient modulation on band-limited channels. Noting the conceptually dual nature of non-equiprobable signaling and source coding, as articulated in [14, Chapter 4], several authors proposed the use of "reverse" source codes derived from, for example, Huffman codes, Tunstall codes, and arithmetic codes as approximate DM codes for shaping applications. See, for example, works by Kschischang and Pasupathy [32], Ungerboeck [57], and Abrahams [1]. Limitations in trying to establish a precise duality between source codes and DM codes are discussed in [8], [3], Baur and Böcherer [4], Lempel et al. [33], and [5, Section 3].

Gallager [16, p. 208] proposed a method of generating symbols with a biased distribution to be combined with linear coding as an approach to achieving capacity of an asymmetric channel. This idea was incorporated into a general scheme that can use capacity-achieving codes for symmetric channels, such as polar codes, to achieve the capacity of arbitrary discrete memoryless asymmetric channels in Mondelli et al. [44].

In [10], Böcherer et al. propose a scheme that combines DM codes (such as constant composition codes) with systematic error correction codes. This scheme can be regarded as a simplification of the bootstrap scheme in Böcherer and Mathar [7], which concatenates the check bits generated by the systematic ECC encoder with the following information bits and applies a DM encoder to them. In [44], the authors also proved that the bootstrap scheme, which they refer to as a chaining construction, can be used to achieve the capacity of any discrete memoryless asymmetric channel.

2) DM codes with optimality measures: In Han [21] and Visweswariah et al. [59], it was shown that an optimal variable-length source code can be regarded as an optimal variable-length DM code for a uniform distribution. The criterion for optimality was the vanishing of a form of normalized conditional Kullback-Leibler (KL) divergence between a subset of codewords of fixed length and words generated i.i.d. with the target distribution, asymptotically in the block length. This result was further developed in Han and Uchida [22], where an optimal variable-length source code with cost, meaning a code that minimizes total cost, was shown to be an optimal DM code. The maximum achievable rate of non-prefix-free DM codes was discussed in Uchida [56].

In [8], dyadic probability mass functions with some optimality properties were used to match the capacity-achieving probability distribution of a discrete memoryless channel, and variable-to-fixed length geometric Huffman codes, mentioned earlier, were used as DM codes. Normalized informational divergence – defined as the KL-divergence between a codeword probability distribution and the distribution of the codewords when generated i.i.d. by the target distribution, normalized by the codeword length – was introduced as the DM code optimality measure. It was then proved that geometric Huffman coding is asymptotically optimal, in the sense that the normalized informational divergence converges to zero as the codeword length increases. Other fixed-length DM codes with vanishing normalized informational divergence were presented in Ramabadran [46] and [51].

Constellation shaping techniques have also been adapted for use in DM coding. For example, a divergence-optimal DM code based on shell mapping was presented in [52] and a DM-like code using trellis-based enumerative amplitude shaping was presented by Gültekin et al. [20].

In [3], the notions of informational divergence and normalized information divergence were extended to measure the performance of fixed-to-variable length codes. Optimality of complete Tunstall code trees with respect to minimizing informational divergence was proved, a result we extend in Section VI. An efficient algorithm for finding binary DM codes that minimize the normalized informational divergence, based on an iterative adaptation of binary Tunstall coding, was presented, and asymptotic optimality with increasing block length was established. In [6] the relationship between normalized information divergence of a DM code and its rate was studied, a topic that we further address in Section VI.

3) Summary of contributions on DM codes: In this paper, we systematically study the problem of designing optimal fixed-to-variable length, prefix-free DM codes from the perspective of word-valued sources and shaping codes. The degree of distribution matching is measured by the KL-divergence between the distribution on word-valued source output sequences and the distribution on those sequences generated i.i.d. according to the target distribution. Vanishing asymptotic normalized KL-divergence at the sequence level, suggested by the approach in [45] and also studied by Soriaga [55], is used as the criterion for optimality. We first characterize the expansion factor of an optimal DM code for a general i.i.d. source (Theorem 12). We then show that an optimal type-II shaping code for a cost model determined by the negative logarithm of a target distribution is an optimal DM code for that distribution (Theorem 13). (This "self-information" cost model was also used in [52] to design information divergence optimal fixed-to-fixed DM codes using shell mapping.)

The connection between shaping codes and DM codes suggests another measure for evaluating DM code performance, which we refer to as *generalized expansion factor* (GEF). We establish a lower bound on the generalized expansion factor, and show that a code that achieves the lower bound is an optimal DM code (Theorem 15). This implies that Varn codes are asymptotically optimal DM codes for a uniform i.i.d. source. Using the GEF, we also extend the separation theorem of shaping codes to DM codes (Theorem 16).

Finally, we discuss relationships between different DM code performance measures. We show that for a DM code with fixed codebook size, minimizing the GEF is equivalent to minimizing the informational divergence introduced in [3], leading to the conclusion that Varn codes designed for the appropriate cost model minimize informational divergence (Theorem 17), generalizing a result for binary Tunstall codes in [3]. We also give an explicit description of the relationship between the normalized informational divergence of a DM code and its expansion factor (Theorem 18 and Remark 18), refining a bound in [6].

C. Organization of the Paper

The remainder of the paper is organized as follows. In Section II, we use known properties of word-valued sources to determine the symbol occurrence probability of shaping code output sequences and the lower bound on the symbol distribution entropy. In Section III, we analyze the distribution, cost, and rate properties of fixed-to-variable length shaping codes. The analysis is then used to prove the equivalence theorem and separation theorem. In Section IV, we establish the equivalence between optimal distribution matching codes and optimal shaping codes. Section V introduces the generalized expansion factor and proves the separation theorem for DM codes. Section VI compares different DM code performance measures. In Section VII, we apply a shaping scheme motivated by our theoretical results to a multilevel flash memory and we show simulation results illustrating the application of Varn codes to DM coding. Section VIII concludes the paper.

II. Information-theoretic Preliminaries

A. Basic Model

First, we fix some notation. Let $\mathbf{X} = X_1 X_2 \ldots$, where $X_i \sim X$ for all i, be an i.i.d. source with alphabet $\mathcal{X} = \{\alpha_1, \ldots, \alpha_u\}$. We use $|\mathcal{X}|$ to denote the size of the alphabet and use $P(x^*)$ to denote the probability of any finite sequence x^*. Let $\mathcal{Y} = \{\beta_1, \ldots, \beta_v\}$ be an alphabet and $\mathcal{Y}^*$ be the set of all finite

sequences over $\mathcal{Y}$, including the null string λ of length 0. Each β_i is associated with a cost C_i. Without loss of generality, we assume that $0 \leqslant C_1 \leqslant C_2 \leqslant \ldots C_v$, and we also assume that there exists at least one pair of costs, C_i and C_j, such that $C_i \neq C_j$. We use a cost vector $\mathcal{C} = [C_1, C_2, \ldots, C_v]$ to represent the cost associated with alphabet $\mathcal{Y}$.

A general shaping code is defined as a prefix-free variable-length mapping $\phi : \mathcal{X}^q \to \mathcal{Y}^*$ which maps a length-q data word x_1^q to a variable-length codeword y^*. We use $\mathbf{Y}$ to denote the process $\phi(\mathbf{X}^q)$, where $\mathbf{X}^q$ is the vector process $X_1^q, X_{q+1}^{2q}, \ldots$. The entropy rate of the process $\mathbf{Y}$ is

$$H(\mathbf{Y}) = \lim_{n\to\infty} \frac{1}{n} H(Y_1 Y_2 \ldots Y_n). \tag{1}$$

We denote the length of a codeword $\phi(x_1^q)$ by $L(\phi(x_1^q))$ and the expected length of codewords generated by a sequence of length-q source words is given by

$$E(L) = \sum_{x_1^q \in \mathcal{X}^q} P(x_1^q) L(\phi(x_1^q)). \tag{2}$$

The *expansion factor* is defined as the ratio of the expected codeword length to the length of the input source word, namely

$$f = E[L]/q. \tag{3}$$

Remark 1. Endurance codes and direct shaping codes can be treated as special cases of this general class of shaping codes. Endurance codes are used when the source has a uniform i.i.d. distribution, with entropy rate $H(\mathbf{X}) = \log_2 |\mathcal{X}|$. A length-$m$ direct shaping code is a shaping code with $q = 1$, $f = 1$, where both $\mathbf{X}$ and $\mathbf{Y}$ have alphabet size 2^m. □

The pair $\mathbf{X}$ and ϕ form a word valued source, as defined in [45]. The following theorem, proved in [45], gives the entropy rate of the shaping code $\phi(\mathbf{X}^q)$.

Theorem 1. *For a prefix-free variable-length code $\mathbf{Y} = \phi(\mathbf{X}^q)$ such that $H(\mathbf{X}^q) < \infty$ and $E(L) < \infty$, the entropy rate of the encoder output satisfies*

$$H(\mathbf{Y}) = \frac{H(\mathbf{X}^q)}{E(L)} = \frac{qH(\mathbf{X})}{E(L)}. \tag{4}$$

□

B. Asymptotic Symbol Occurrence Probability

For simplicity and without loss of generality, we assume $q = 1$. The mapping is $\phi : \mathcal{X} \to \mathcal{Y}^*$. Let y_1^l denote the first l symbols of $\phi(\mathbf{X})$. We assume the cost is independent and additive, so the cost of sequence y_1^l can be expressed as

$$W(y_i^l) = \sum_{i=1}^{v} N_i(y_1^l) C_i \tag{5}$$

where $N_i(y_1^l)$ stands for the number of occurrences of β_i in sequence y_1^l. The cost per code symbol is therefore $\sum_i N_i(y_1^l) C_i / l$. Let

$$Q(y_1^l) = Pr\{Y_1^l = y_1^l\} \tag{6}$$

denote the probability distribution of Y_1^l. The expected cost per symbol of a length-l shaping code sequence is

$$\begin{aligned} W_l &= \sum_{y_1^l \in \mathcal{Y}^l} Q(y_1^l) W(y_i^l)/l \\ &= \sum_{i=1}^{v} \sum_{y_1^l \in \mathcal{Y}^l} Q(y_1^l) N_i(y_1^l) C_i / l. \end{aligned} \tag{7}$$

The asymptotic expected cost per symbol, or *average cost* of a shaping code is

$$A(\phi(\mathbf{X})) = \lim_{l\to\infty} W_l. \tag{8}$$

Let

$$\hat{P}_i = \lim_{l\to\infty} \sum_{y_1^l} Q(y_1^l) N_i(y_1^l)/l = \lim_{l\to\infty} \frac{E(N_i(Y_1^l))}{l} \tag{9}$$

be the asymptotic probability of occurrence of β_i. Then the average cost of a shaping code can be expressed as

$$A(\phi(\mathbf{X})) = \sum_i \hat{P}_i C_i. \tag{10}$$

In the rest of this subsection, we will show how to calculate $\hat{P}_i$. Define the prefix operator π as $y_1^n \pi^i = y_1^{n-i}$ for $0 \leqslant i < n$ and $y_1^n \pi^i = \lambda$ for $i \geqslant n$. Let $\pi\{y^*\}$ denote the set of all the prefixes of a sequence y^*. We denote by $\mathcal{G}_\phi(y_1^l)$ the set of all sequences $x^* \in \mathcal{X}^*$ such that y_1^l is a prefix of $\phi(x^*)$ but not of $\phi(x^*\pi)$. That is,

$$\mathcal{G}_\phi(y_1^l) = \{x^* \in \mathcal{X}^* | y_1^l \in \pi\{\phi(x^*)\} \wedge |\phi(x^*\pi)| < l\} \tag{11}$$

and the distribution of y_1^l can be expressed as

$$Q(y_1^l) = \sum_{x^* \in \mathcal{G}_\phi(y_1^l)} P(x^*). \tag{12}$$

We define by M_l the minimum length of a sequence x_1^n such that $|\phi(x_1^n)| \geqslant l$ and let S_{M_l} be the length of $\phi(x_1^{M_l})$. Note that

$$S_{M_l - 1} < l \leqslant S_{M_l}. \tag{13}$$

According to [45], the random variable M_l satisfies the property of being a *stopping rule* for the sequence of i.i.d. random variables $\{\phi(X^\infty)\}$. Wald's equality [60] then implies that

$$E(N_i(\phi(X_1^{M_l}))) = E(N_i(\phi(X)))E(M_l). \tag{14}$$

The following two lemmas were proved in [45].

Lemma 2. *Given a nonnegative-valued function f, let $F_i = f(X_i)$. If $E(F) < \infty$, then*

$$\lim_{l\to\infty} \frac{E(F_{M_l})}{l} = 0. \tag{15}$$

□

Remark 2. The previous lemma is not obvious, because even when $E(F) < \infty$, $E(F_{M_l})$ is not necessarily equal to $E(F)$. □

Lemma 3. *If $E[L] < \infty$, then*

$$\lim_{l\to\infty} \frac{E[M_l]}{l} = \frac{1}{E(L)}. \tag{16}$$

□

Using these results, we derive a lemma which tells us how to calculate the asymptotic occurrence probability of the encoder output process $\mathbf{Y}$.

Lemma 4. *For a prefix-free variable-length code $\phi : \mathcal{X}^q \to \mathcal{Y}^*$ such that $E(N_i(\phi(X^q))) < \infty$ for all symbols β_i and $E(L) < \infty$, the asymptotic probability of occurrence $\hat{P}_i$ of β_i is given by*

$$\hat{P}_i = E(N_i(\phi(X^q)))\frac{1}{E(L)}. \tag{17}$$

Proof: See Appendix A. ■

It is easy to check that $\sum_i \hat{P}_i = 1$, so this distribution is well defined.

C. Lower Bound on Symbol Distribution Entropy

Consider a prefix-free variable-length code ϕ as in Lemma 4. Let $\hat{Y}_1^l$ denote an i.i.d. sequence of length l and with distribution $\{\hat{P}_i\}$. The probability of a length-l sequence y_1^l with respect to this distribution is $\hat{P}(y_1^l) = \prod_i \hat{P}_i^{N_i(y_1^l)}$. The Kullback-Leibler (KL) divergence (also known as the KL-distance or relative entropy) [11] is a measure of the inefficiency caused by an approximation. The KL-divergence between Y_1^l and $\hat{Y}_1^l$ is

$$D(Y_1^l||\hat{Y}_1^l) = \sum_{y_1^l\in\mathcal{Y}^l} Q(y_1^l)\log_2\frac{Q(y_1^l)}{\hat{P}(y_1^l)}. \tag{18}$$

The following lemma provides a lower bound on the symbol distribution entropy.

Lemma 5. *The asymptotic normalized KL-divergence satisfies*

$$\lim_{l\to\infty}\frac{1}{l}D(Y_1^l||\hat{Y}_1^l) = H(\hat{Y}) - H(\mathbf{Y}), \tag{19}$$

where $H(\hat{Y}) = -\sum_i \hat{P}_i \log_2 \hat{P}_i$. Specifically,

$$H(\hat{Y}) \geqslant H(\mathbf{Y}). \tag{20}$$

Proof: We rewrite the $D(Y_1^l||\hat{Y}_1^l)$ as

$$\begin{aligned} D(Y_1^l||\hat{Y}_1^l) &= \sum_{y_1^l\in\mathcal{Y}^l} Q(y_1^l)\log_2\frac{Q(y_1^l)}{\hat{P}(y_1^l)} \\ &= \sum_{y_1^l\in\mathcal{Y}^l} Q(y_1^l)\log_2 Q(y_1^l) - \sum_{y_1^l\in\mathcal{Y}^l} Q(y_1^l)\log_2\hat{P}(y_1^l) \\ &= -H(Y_1^l) - \sum_{y_1^l\in\mathcal{Y}^l} Q(y_1^l)\log_2\hat{P}(y_1^l). \end{aligned} \tag{21}$$

The second term of the right-hand side of this equation is

$$\begin{aligned} \sum_{y_1^l\in\mathcal{Y}^l} Q(y_1^l)\log_2\hat{P}(y_1^l) &= \sum_{y_1^l\in\mathcal{Y}^l} Q(y_1^l)\log_2\prod_i \hat{P}_i^{N_i(y_1^l)} \\ &= \sum_{y_1^l\in\mathcal{Y}^l} Q(y_1^l)\sum_i N_i(y_1^l)\log_2\hat{P}_i \\ &= \sum_i \log_2\hat{P}_i \sum_{y_1^l\in\mathcal{Y}^l} Q(y_1^l)N_i(y_1^l). \end{aligned} \tag{22}$$

Combining equations (21) and (22), we have

$$\begin{aligned} &\lim_{l\to\infty}\frac{1}{l}D(Y_1^l||\hat{Y}_1^l) \\ &= -\lim_{l\to\infty}\frac{1}{l}H(Y_1^l) - \sum_i \log_2\hat{P}_i \lim_{l\to\infty}\sum_{y_1^l\in\mathcal{Y}^l}\frac{Q(y_1^l)N_i(y_1^l)}{l} \\ &= -H(\mathbf{Y}) - \sum_i \hat{P}_i\log_2\hat{P}_i = H(\hat{Y}) - H(\mathbf{Y}). \end{aligned} \tag{23}$$

Using the fact that $D(Y_1^l||\hat{Y}_1^l) \geqslant 0$, we have

$$H(\hat{Y}) \geqslant H(\mathbf{Y}). \tag{24}$$

This completes the proof. ■

Remark 3. We note that the entropy rate bound (24) also follows from applying the independence bound for joint entropy of a random vector, invoking concavity of the entropy function, and taking the limit in the length of the vector. □

Remark 4. From the proof, we see that $H(\hat{Y}) = H(\mathbf{Y})$ implies $\lim_{l\to\infty}\frac{1}{l}D(Y^l||\hat{Y}^l) = 0$. Therefore, the codeword sequence $Y_1Y_2\cdots$ approximates an i.i.d. sequence generated by $\hat{Y}$. □

Example 1. Consider a uniform i.i.d. binary source $\mathbf{X}$ and a prefix-free variable-length code defined by the mapping $\{00\to 000, 01\to 001, 10\to 01, 11\to 1\}$. The occurrence probabilities of symbols 0 and 1 are $2/3$ and $1/3$, respectively. The symbol distribution entropy is

$$H(\hat{Y}) = -\frac{1}{3}\log_2\frac{1}{3} - \frac{2}{3}\log_2\frac{2}{3} \simeq 0.9183. \tag{25}$$

The entropy rate of the shaping code sequence is

$$H(\mathbf{Y}) = \frac{H(\mathbf{X}^2)}{E(L)} = \frac{2}{2.25} = 0.8889. \tag{26}$$

We see that, in this case, $H(\mathbf{Y}) < H(\hat{Y})$. □

III. Optimal Shaping Codes

A. Cost Minimizing Probability Distribution

In this subsection, we discuss the properties of optimal shaping codes. We consider two scenarios. First, we analyze shaping codes that minimize the average cost with a given expansion factor. We then analyze shaping codes that minimize the expected cost per source symbol, or total cost.

We refer to the first minimization problem as the *type-I shaping problem*, and we call shaping codes that achieve the

minimum average cost for a given expansion factor *optimal type-I shaping codes*. The following theorem gives a lower bound on the average cost and the corresponding asymptotic symbol occurrence probabilities.

Theorem 6. *Given the source* $\mathbf{X}$ *and cost vector* $\mathcal{C}$*, the average cost of a type-I shaping code* $\phi : \mathcal{X}^q \to \mathcal{Y}^*$ *with expansion factor* f *is lower bounded by* $\sum_i \hat{P}_i C_i$*, with*

$$\hat{P}_i = \frac{1}{N} 2^{-\mu C_i}, \tag{27}$$

where N *is a normalization constant such that* $\sum_i \hat{P}_i = 1$ *and* μ *is a non-negative constant such that* $\sum_i -\hat{P}_i \log \hat{P}_i = H(\mathbf{X})/f$*.*

Proof: From Theorem 1 and Lemma 5, we see that, for a shaping code ϕ with expansion factor f, the following inequality holds:

$$H(\hat{Y}) \geqslant H(\mathbf{Y}) = \frac{qH(\mathbf{X})}{E(L)} = \frac{H(\mathbf{X})}{f}. \tag{28}$$

To calculate the minimum possible average cost, we must solve the optimization problem:

$$\begin{aligned} \underset{\hat{P}_i}{\text{minimize}} \quad & \sum_i \hat{P}_i C_i \\ \text{subject to} \quad & H(\hat{Y}) \geqslant \frac{H(\mathbf{X})}{f} \\ & \sum_i \hat{P}_i = 1. \end{aligned} \tag{29}$$

We divide this optimization problem into two parts. First, we fix $H(\hat{Y})$ and find the optimal symbol occurrence probabilities. Then we find the optimal $H(\hat{Y})$ to minimize the average cost. The optimization problem then becomes

$$\begin{aligned} \underset{H(\hat{Y})}{\text{minimize}} \quad \underset{\hat{P}_i}{\text{minimize}} \quad & \sum_i \hat{P}_i C_i \\ \text{subject to} \quad & H(\hat{Y}) \geqslant \frac{H(\mathbf{X})}{f} \\ & \sum_i \hat{P}_i = 1. \end{aligned} \tag{30}$$

If we fix $H(\hat{Y})$, we can solve the optimization problem by using the method of Lagrange multipliers. The solution is

$$\hat{P}_i = \frac{1}{N} 2^{-\mu C_i} \tag{31}$$

where $N = \sum_i 2^{-\mu C_i}$ is a normalization constant and μ is a non-negative constant such that

$$H(\hat{Y}) = \sum_i -\hat{P}_i \log_2 \hat{P}_i. \tag{32}$$

Note that $\mu = 0$ if and only if $H(\hat{Y}) = \log_2 |\mathcal{Y}|$. For simplicity, let h denote $H(\hat{Y})$. Then μ and N are functions of h, which we denote by $N \stackrel{\text{def}}{=} N(h)$ and $\mu \stackrel{\text{def}}{=} \mu(h)$, respectively. Let $C(h) = \sum_i \frac{C_i}{N(h)} 2^{-\mu(h) C_i}$ be the minimum cost, given that $h \geqslant \frac{H(\mathbf{X})}{f}$. From (32), we see that

$$C(h) = \frac{h - \log_2 N}{\mu} \quad \text{when } \mu > 0. \tag{33}$$

The optimization problem we have reduced to here, minimizing the average cost of a probability mass function subject to a lower bound on entropy, is dual to the problem considered in prior work such as [40, Problem 1.8] and [5, Sec. 5.2], which is a special case of results in [41], [26], and [28]. The relationship between entropy rate and average cost discussed in these papers has the same functional form as (33). We can apply the analysis in [5, Sec. 5.2], to conclude that

$$\frac{\mathrm{d}h}{\mathrm{d}C} = \mu \Rightarrow \frac{\mathrm{d}C}{\mathrm{d}h} > 0 \quad \text{when } \mu > 0. \tag{34}$$

Therefore, the minimum cost for a shaping code with expansion factor f is achieved when $h = H(\hat{Y}) = \frac{H(X)}{f}$.

Note that we have minimized average cost by optimizing the asymptotic symbol occurrence probability $\hat{P}_i$ of a prefix-free variable-length mapping whose output entropy rate is fixed, without consideration of whether the output sequence coincides with an i.i.d. sequence. ■

Remark 5. If the source $\mathbf{X}$ has a uniform distribution, then μ satisfies $-f \sum_i \hat{P}_i \log \hat{P}_i = \log_2 |\mathcal{X}|$. Thus, we recover the result in [25] characterizing endurance codes with minimum average cost. □

Remark 6. When the minimum average cost is achieved, we have $H(\hat{Y}) = H(\mathbf{Y})$. Thus, the codeword sequence approximates an i.i.d. sequence generated by distribution $\{\hat{P}_i\}$ (see Remark 4). □

Given a prefix-free variable-length shaping code $\phi : \mathcal{X}^q \to \mathcal{Y}^*$, assume that after nq source symbols are encoded, the codeword sequence is $\phi(x_1^{nq})$. As in equations (7), (8), (9), we formally define the expected cost per source symbol, or *total cost* of a shaping code as

$$\begin{aligned} T(\phi(\mathbf{X}^q)) &= \frac{\sum_i E(N_i(\phi(X^{nq})))C_i}{nq} = \frac{\sum_i E(N_i(\phi(X^q)))C_i}{q} \\ &= \frac{E(L)}{q} \frac{\sum_i E(N_i(\phi(X^q)))C_i}{E(L)} = f \sum_{i=1}^{v} \hat{P}_i C_i. \end{aligned} \tag{35}$$

We refer to the problem of minimizing the total cost as the *type-II shaping problem*. Shaping codes that achieve the minimum total cost are referred to as *optimal type-II shaping codes*. The corresponding optimization problem is as follows:

$$\begin{aligned} \underset{\hat{P}_i, f}{\text{minimize}} \quad & f \sum_{i=1}^{v} \hat{P}_i C_i \\ \text{subject to} \quad & H(\hat{Y}) \geqslant H(\mathbf{Y}) = \frac{H(\mathbf{X})}{f} \\ & \sum_i \hat{P}_i = 1. \end{aligned} \tag{36}$$

Using Theorem 6, we can calculate the total cost as a function of the expansion factor f. Fig. 1 shows the total cost curve for a quaternary source and code alphabet, a uniformly distributed source $\mathbf{X}$, and cost vector $\mathcal{C} = [1, 2, 3, 4]$. There is an optimal value of f and corresponding minimum total cost.

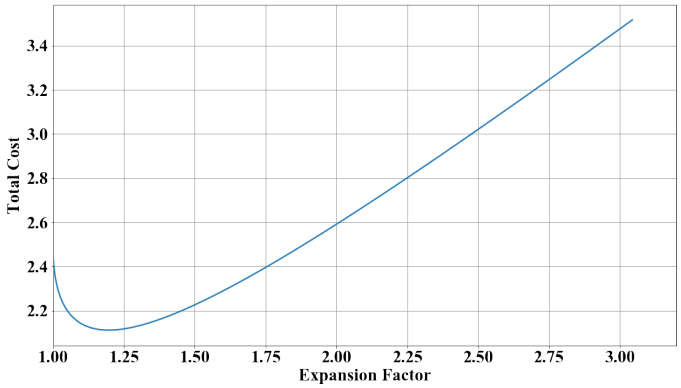


Fig. 1: Total cost versus f for random source with $\mathcal{C} = [1, 2, 3, 4]$.

We now determine the minimum achievable total cost of a shaping code.

Theorem 7. *Given the source* $\mathbf{X}$ *and cost vector* $\mathcal{C}$*, if* $C_1 \neq 0$*, then the minimum total cost of a shaping code* $\phi : \mathcal{X}^q \to \mathcal{Y}^*$ *is given by* $f \sum_i \hat{P}_i C_i$*, where* $\hat{P}_i = 2^{-\mu C_i}$*,* μ *is a positive constant selected such that* $\sum_i 2^{-\mu C_i} = 1$*. The corresponding expansion factor* f *is*

$$f = \frac{H(\mathbf{X})}{-\sum_i \hat{P}_i \log_2 \hat{P}_i}. \tag{37}$$

If $C_1 = 0$*, then the total cost is a decreasing function of* f*.*

Proof: See Appendix B. ∎

Remark 7. For a positive cost vector $\mathcal{C}$, the minimum achievable total cost is

$$\begin{aligned} T(\phi(\mathbf{X}^q)) &= f \sum_i \hat{P}_i C_i = \frac{H(\mathbf{X})}{-\sum_i \hat{P}_i \log_2 \hat{P}_i} \sum_i \hat{P}_i C_i \\ &= \frac{H(\mathbf{X})}{-\sum_i \hat{P}_i \log_2 2^{-\mu C_i}} \sum_i \hat{P}_i C_i = \frac{H(\mathbf{X})}{\mu \sum_i \hat{P}_i C_i} \sum_i \hat{P}_i C_i \\ &= \frac{H(\mathbf{X})}{\mu}. \end{aligned} \tag{38}$$

The capacity of a noiseless finite-state costly channel, which is essentially the inverse of the minimum total cost, was considered in [41], [26], [28], [9], and [5, Sec 6] from combinatorial and probabilistic perspectives. Equivalences between the combinatorial and probabilistic definitions of capacity were established, extending the original results of Shannon.

In [13, Theorem 4.4] and [22, Theorem 1], the minimum total cost of a prefix-free variable-length code was determined. In [5, Section 4], an asymptotically optimal variable-to-fixed length code for binary uniform i.i.d. source that minimizes total cost was proposed. However, these works did not address the code expansion factor and asymptotic symbol occurrence probability corresponding to the minimum total cost.

In [40, Problems 1.8], [39] and [5, Sec. 5.2], the relationship between the maximum entropy of a probability mass function on an alphabet with cost subject to an average cost constraint was discussed. However, these works did not explore the functional relationship between the total cost and the expansion factor of a code. Here, using the word-valued source perspective, we establish the relationship between the total cost of a rate-constrained prefix-free code and its expansion factor. This relationship plays an important role in the proof of the separation theorem (Theorem 10) in Appendix D. We also address the special case of zero lowest cost, i.e., $C_1 = 0$, in which no global minimum can be reached. □

Remark 8. If we only apply optimal lossless compression to the source $\mathbf{X}$, the code sequence has a uniform distribution. Therefore, we have $\mu = 0$ and $N = |\mathcal{Y}| > 1$. This implies that simply applying compression to the source data is not the best way to reduce the total cost. □

B. Optimal Data Shaping Code Design

Many previous works investigated type-II shaping code design. For example, see [27], [58], and [13] . In this subsection, we consider the problem of designing an optimal type-I shaping code by transforming this problem into a type-II shaping problem. Combining Theorems 6 and 7, we can prove the following equivalence theorem.

Theorem 8. *A code that achieves the minimum total cost for cost vector* $\mathcal{C}'$ *also achieves minimum average cost for cost vector* $\mathcal{C}$ *and expansion factor* f *if*

$$C'_i = -\log_2 \hat{P}_i, \tag{39}$$

where $\{\hat{P}_i\}$ *are the probabilities minimizing average cost for the cost vector* $\mathcal{C}$ *and expansion factor* f*.*

Proof: First we consider the optimal type-II shaping code $\phi : \mathcal{X}^q \to \mathcal{Y}^*$ with cost vector $\mathcal{C}'$. By Theorem 7, this code generates codeword sequence with probability of occurrence $P'_i = 2^{-\mu C'_i}$, where μ satisfies the equation

$$\sum_i 2^{-\mu C'_i} = 1. \tag{40}$$

Since $C'_i = -\log_2 \hat{P}_i$, it is easy to check that the solution of equation (40) is $\mu = 1$. This means when the minimum total cost is achieved, the probability of occurrence of codeword sequence is $P'_i = 2^{-C'_i} = \hat{P}_i$ and the expansion factor of this code is

$$f' = \frac{H(X)}{-\sum_i P'_i \log_2 P'_i} = f \tag{41}$$

Referring to Theorem 6, we see that ϕ is also optimal with respect to minimizing average cost with cost vector $\mathcal{C}$ and expansion factor f. ∎

When designing a type-I shaping code with expansion factor f and cost vector $\mathcal{C}$, we can first calculate the desired distribution $\{\hat{P}_i\}$, then transform this problem into a type-II shaping code problem for the channel with symbol cost $\{C'_i = -\log_2 \hat{P}_i\}$. Thus we can apply known type-II shaping code algorithms to solve this problem.

Remark 9. For an arbitrary i.i.d. source and a positive cost vector $\mathcal{C}$, generalized Shannon-Fano codes [13, Theorem 4.4] are tree-based variable-length codes, $\phi : \mathcal{X}^q \to \mathcal{Y}^*$, whose total cost is upper bounded by

$$\begin{aligned} T(\phi) &< \frac{H(\mathbf{X})}{\mu} + \frac{\max_i\{C_i\}}{q} \\ &\to \frac{H(\mathbf{X})}{\mu} \quad \text{as } q \to \infty. \end{aligned} \tag{42}$$

This coding scheme includes dividing the $[0,1]$ interval based on $P(x_1^q)$ and calculating $2^{-\mu W}$, where W is the cost of a codeword. This construction may become impractical when q is large. □

Remark 10. For a uniform i.i.d. source and a positive cost vector $\mathcal{C}$, Varn codes [58] are tree-based, variable-length codes $\phi_K : \mathcal{X}^{\log_{|\mathcal{X}|} K} \to \mathcal{Y}^*$ that minimize total cost for a specified codebook size K. In [49], bounds were established on the average *codeword* cost for the Varn code with codebook size K, denoted $C(\phi_K)$. Specifically,

$$\frac{\log_2 K}{\mu} \leqslant C(\phi_K) \leqslant \frac{\log_2 K}{\mu} + \max_i\{C_i\}. \tag{43}$$

Dividing by $\log_{|\mathcal{X}|} K$, we see that the total cost of the Varn code with codebook size K is bounded by

$$\frac{\log_2 |\mathcal{X}|}{\mu} \leqslant T(\phi_K) \leqslant \frac{\log_2 |\mathcal{X}|}{\mu} + \frac{\max_i\{C_i\}}{\log_{|\mathcal{X}|} K}. \tag{44}$$

Therefore

$$\lim_{K\to\infty} T(\phi_K) = \frac{\log_2 |\mathcal{X}|}{\mu} \tag{45}$$

which implies that Varn codes are asymptotically optimal type-II shaping codes (see Remark 7). □

We now present a separation theorem for type-II shaping codes. It states that minimum total cost can be achieved by a concatenation of optimal lossless compression with an optimal type-II shaping code for a uniform i.i.d. source. The proof uses a construction based on typical sequences.

Theorem 9. *Given the source* $\mathbf{X}$ *and cost vector* $\mathcal{C}$*, the minimum total cost can be achieved by a concatenation of an optimal lossless compression code with an optimal type-II shaping code for a uniform i.i.d. source.*

Proof: See Appendix C. ■

An example of an optimal type-II shaping scheme that illustrates Theorem 9 was described by Iwata in [24]. It uses a concatenation of an LZ78 code and a Varn code as outer and inner codes, respectively.

There is also a separation theorem for type-I shaping codes, stating that minimum average cost for a given expansion factor can be achieved by a concatenation of optimal lossless compression with an optimal type-I shaping code for a uniform i.i.d. source and suitable expansion factor. The proof relies on the type-II separation theorem and the equivalence between type-II and type-I shaping codes established in Theorem 8. It requires an analysis of the behavior of the total cost function in the vicinity of the expansion factor that minimizes total cost.

Theorem 10. *Given the source* $\mathbf{X}$*, cost vector* $\mathcal{C}$ *and expansion factor* f*, the minimum average cost can be achieved by a concatenation of an optimal lossless compression code with a binary optimal type-I shaping code for uniform i.i.d. source and expansion factor*

$$f' = \frac{f}{H(\mathbf{X})}. \tag{46}$$

Proof: See Appendix D. ■

IV. Distribution Matching Code Design

Given a target distribution $\{P_i\}$, distribution matching (DM) considers the problem of mapping an i.i.d. sequence of source symbols to an output sequence of symbols that are approximately independent and distributed according to $\{P_i\}$. An optimal DM code must satisfy two conditions: the codeword sequence has symbol occurrence probabilities $\hat{P}_i = P_i$, and the output sequence looks like an i.i.d. sequence. We measure the latter property using the asymptotic normalized KL-divergence defined in Lemma 5.

It has been shown in Theorems 6 and 7 that an optimal shaping code will generate an output sequence such that $\lim_{l\to\infty} \frac{1}{l} D(Y_1^l || \hat{Y}_1^l) = 0$. Thus the output sequence approximates an i.i.d. sequence with symbol occurrence probability distribution $\{\hat{P}_i\}$. This implies that we can solve the distribution matching problem by designing a corresponding shaping code. In this section, we consider the problem of designing optimal DM codes. We first formulate the problem of *generating an i.i.d. sequence* and then show the connection between DM codes and shaping codes. We then propose a *generalized expansion factor* to measure the performance of a DM code. A comparison of DM code performance measures is also presented.

A. Problem Formulation

We use the asymptotic normalized Kullback-Leibler divergence [55] to formally define an optimal DM code ϕ for distribution $\{P_i\}$.

Definition 11. *A variable-length mapping* $\phi : \mathcal{X}^q \to \mathcal{Y}^*$ *is an optimal DM code for distribution* $\{P_i\}$ *if*

$$\lim_{l\to\infty} \frac{1}{l} D(Y_1^l || \tilde{Y}_1^l) = 0, \tag{47}$$

where $\tilde{\mathbf{Y}}$ *is an i.i.d. process with distribution* $\{P_i\}$. □

By combining Theorem 1 and Lemma 5, we can prove the following theorem.

Theorem 12. *The expansion factor of a mapping satisfies the lower bound*

$$f = \frac{H(\mathbf{X})}{H(\mathbf{Y})} \geqslant \frac{H(\mathbf{X})}{H(\hat{Y})} \tag{48}$$

with equality if and only if $\lim_{l\to\infty}\frac{1}{l}D(Y_1^l||\hat{Y}_1^l) = H(\hat{Y}) - H(\mathbf{Y}) = 0$. *When* $f = \frac{H(\mathbf{X})}{H(\hat{Y})}$, *this code is an optimal DM code for distribution* $\{P_i\}$. □

Remark 11. Assuming this mapping is an optimal compression, the compression ratio g is

$$g = \frac{H(\mathbf{X})}{\log_2|\mathcal{Y}|}. \tag{49}$$

By Theorem 1 and Lemma 5, we have

$$H(\hat{Y}) \geqslant H(\mathbf{Y}) = \frac{H(\mathbf{X})}{g} = \log_2|\mathcal{Y}|. \tag{50}$$

Since $H(\hat{Y}) \leqslant \log_2|\mathcal{Y}|$, we know that $H(\hat{Y}) = H(\mathbf{Y}) = \log_2|\mathcal{Y}|$ and

$$\lim_{l\to\infty}\frac{1}{l}D(Y_1^l||\hat{Y}_1^l) = 0. \tag{51}$$

This implies the codeword sequence looks i.i.d. and has probability of occurrence

$$\hat{P}_i = \frac{1}{|\mathcal{Y}|} \quad \text{for all } i. \tag{52}$$

This proves the well-known fact that the output of an optimal compression approximates a uniform i.i.d. sequence [59], [21], [22]. □

Let $\tilde{\mathbf{Y}}$ be the i.i.d. process with distribution $\{P_i\}$. As in the derivation of (23) in Lemma 5, we find

$$\begin{aligned}
&\lim_{l\to\infty}\frac{1}{l}D(Y_1^l||\tilde{Y}_1^l)\\
&= -\lim_{l\to\infty}\frac{1}{l}H(Y_1^l) - \sum_i \log_2 P_i \lim_{l\to\infty}\sum_{y_1^l\in\mathcal{Y}^l}\frac{Q(y_1^l)N_i(y_1^l)}{l}\\
&= -H(\mathbf{Y}) - \sum_i \hat{P}_i\log_2 P_i = -\sum_i \hat{P}_i\log_2 P_i - \frac{H(\mathbf{X})}{f}\\
&= \frac{-f\sum_i \hat{P}_i\log_2 P_i - H(\mathbf{X})}{f}.
\end{aligned} \tag{53}$$

From Theorem 7, we know that for a channel with cost $\{C_i = -\log_2 P_i\}$, the total cost $-f\sum_i \hat{P}_i\log_2 P_i$ is lower bounded by $H(\mathbf{X})$. The shaping code that achieves this lower bound has the following two properties:

- The probability of occurrence of symbol β_i satisfies $\hat{P}_i = P_i$ for all β_i,
- The asymptotic normalized KL-divergence between $\mathbf{Y}$ and $\hat{Y}$ satisfies

$$\lim_{l\to\infty}\frac{1}{l}D(Y_1^l||\hat{Y}_1^l) = 0. \tag{54}$$

This implies that this code generates a sequence that approximates an i.i.d. sequence with distribution $\{P_i\}$. This analysis also implies that the expansion factor of an optimal DM code is

$$f_{\text{opt}} = \frac{H(\mathbf{X})}{-\sum_i \hat{P}_i\log_2\hat{P}_i} = \frac{H(\mathbf{X})}{-\sum_i P_i\log_2 P_i}. \tag{55}$$

We summarize in the following theorem the relationship between optimal shaping codes and optimal DM codes, extending the result in [22] by explicitly showing the optimal expansion factor.

Theorem 13. *The optimal type-II shaping code with cost vector* $\mathcal{C}$, *or the equivalent type-I shaping code from Theorem* 8, *is an optimal DM code for distribution* $\{P_i\}$ *if*

$$C_i = -\log_2 P_i \tag{56}$$

for every symbol β_i, *in the sense that*

$$\lim_{l\to\infty}\frac{1}{l}D(Y_1^l||\tilde{Y}_1^l) = 0. \tag{57}$$

The expansion factor of this optimal DM code is

$$f_{opt} = \frac{H(\mathbf{X})}{-\sum_i P_i\log_2 P_i}. \tag{58}$$

□

Remark 12. Shell mapping was used in [52] to design fixed-length DM codes with uniformly distributed input bits. The shell mapper that minimizes informational divergence (introduced later in Section VI-A) uses the "self-information" weight function $C_i = -\log_2 P_i$ and the optimal expansion factor is determined by a search. Theorem 13 considers a more general variable-length DM code with arbitrary i.i.d. source and characterizes the optimal expansion factor. Codes minimizing informational divergence are discussed further in Section VI-A. □

V. Generalized Expansion Factor

The relationship between optimal shaping codes and optimal DM codes was established above. The total cost of the shaping code suggests an alternative performance measure for a DM code which will be useful when analyzing the optimality of a shaping-based DM code construction and in proving a separation theorem for DM codes. Specifically, we define the *generalized expansion factor* (GEF) of a prefix-free variable-length code as follows.

Definition 14. *Given a prefix-free variable-length code* $\phi : \mathcal{X}^q \to \mathcal{Y}^*$ *and a set of positive real numbers* $\{P_1, P_2, \ldots, P_v\}$ *such that* $\sum P_i = 1$, *the* **generalized expansion factor** *of this code is defined as*

$$F(\phi, P_1, \ldots, P_v) = -f\frac{\sum_i \hat{P}_i\log_2 P_i}{\log_2|\mathcal{Y}|} \tag{59}$$

where f *is the code expansion factor and* $\{\hat{P}_i\}$ *is the asymptotic symbol occurrence probability distribution.* □

For simplicity, we sometimes use F to represent $F(\phi, P_1, \ldots, P_v)$. The following theorem shows that F can be used to evaluate an optimal DM code.

Theorem 15. *Given a prefix-free variable-length code* $\phi : \mathcal{X}^q \to \mathcal{Y}^*$ *and a set of positive real numbers* $\{P_1, P_2, \ldots, P_v\}$

such that $\sum P_i = 1$, *the generalized expansion factor of this mapping is lower bounded by*

$$F(\phi, P_1, \ldots, P_v) \geqslant \frac{H(\mathbf{X})}{\log_2 |\mathcal{Y}|}. \tag{60}$$

If $F = \frac{H(\mathbf{X})}{\log_2 |\mathcal{Y}|}$, *this mapping is an optimal DM code for the target distribution* $\{P_i\}$, *in the sense that*

$$\lim_{l\to\infty} \frac{1}{l} D(Y_1^l || \tilde{Y}_1^l) = 0. \tag{61}$$

□

Proof: Assume symbol β_i in the codeword sequence has cost $C_i = -\log_2 P_i$. The total cost of this mapping is

$$T(\phi) = f \sum_i \hat{P}_i C_i = -f \sum_i \hat{P}_i \log_2 P_i. \tag{62}$$

Comparing equations (59) and (62) we have

$$F(\phi, P_1, \ldots, P_v) = \frac{T(\phi)}{\log_2 |\mathcal{Y}|}. \tag{63}$$

This indicates that the GEF of a DM code is equivalent to its total cost when applying it to a costly channel with cost $C_i = -\log_2 P_i$. From Theorem 7, we know the total cost of a prefix-free mapping satisfies the lower bound

$$T(\phi) \geqslant \frac{H(\mathbf{X})}{\mu} \tag{64}$$

where μ is a constant such that $\sum 2^{-\mu C_i} = 1$. Since $C_i = -\log_2 P_i$, it is easy to check that $\mu = 1$ and

$$F(\phi, P_1, \ldots, P_v) = \frac{T(\phi)}{\log_2 |\mathcal{Y}|} \geqslant \frac{H(\mathbf{X})}{\log_2 |\mathcal{Y}|}. \tag{65}$$

When the minimum GEF is achieved, this code is also an optimal type-II shaping code with $C_i = -\log_2 P_i$. Theorem 13 then implies that this code is an optimal DM code for the target distribution $\{P_i\}$, in the sense that

$$\lim_{l\to\infty} \frac{1}{l} D(Y_1^l || \tilde{Y}_1^l) = 0. \tag{66}$$

■

Remark 13. As shown in Remark 10, for a uniform i.i.d. source and a cost vector $\mathcal{C}$, a Varn code $\phi_K : \mathcal{X}^{\log_{|\mathcal{X}|} K} \to \mathcal{Y}^*$ is an asymptotically optimal type-II shaping code. If the costs are given by $C_i = -\log_2 P_i$, where $\sum_i P_i = 1$, the total cost is bounded by

$$\log_2 |\mathcal{X}| \leqslant T(\phi_K) \leqslant \log_2 |\mathcal{X}| + \frac{\max_i\{C_i\}}{\log_{|\mathcal{X}|} K}. \tag{67}$$

Equation (63) implies that for the target distribution $\{P_i\}$, Varn codes minimize GEF for a specified codebook size K. Thus, a Varn code can be regarded as a DM code with generalized expansion factor bounded by

$$\frac{\log_2 |\mathcal{X}|}{\log_2 |\mathcal{Y}|} \leqslant F \leqslant \frac{\log_2 |\mathcal{X}|}{\log_2 |\mathcal{Y}|}\left(1 + \frac{\max_i\{-\log_2 P_i\}}{\log_2 K}\right). \tag{68}$$

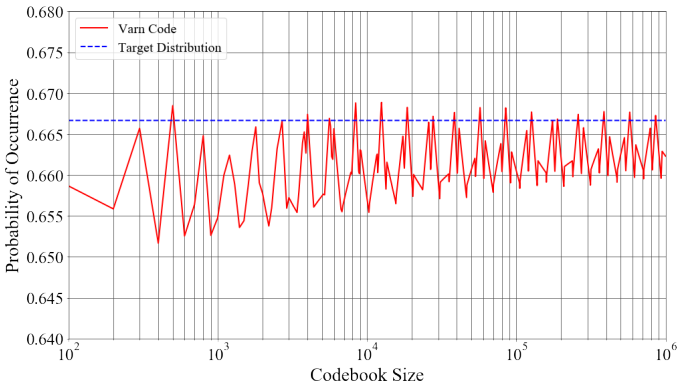


Fig. 2: Probability of occurrence $\hat{P}_0$ of a Varn code for the target distribution $\{2/3, 1/3\}$.

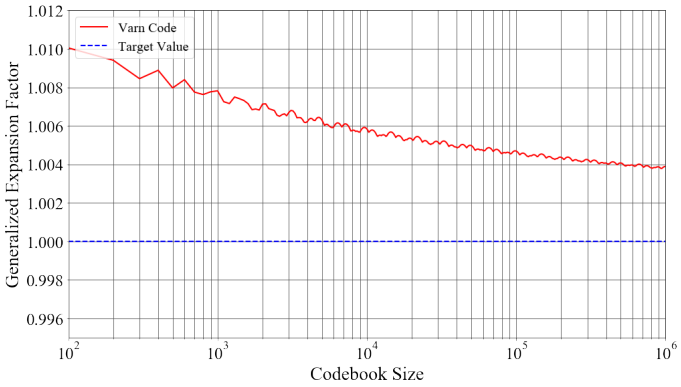


Fig. 3: Generalized expansion factor of a Varn code for the target distribution $\{2/3, 1/3\}$.

Therefore, we have

$$\lim_{K\to\infty} F(\phi_K, P_1, \ldots, P_v) = \frac{\log_2 |\mathcal{X}|}{\log_2 |\mathcal{Y}|} \tag{69}$$

which implies that Varn codes are asymptotically optimal DM codes. Fig. 2 and Fig. 3 show the probability of occurrence and generalized expansion factor of binary Varn codes (i.e., with $\mathcal{X} = \mathcal{Y} = \{0,1\}$) for a target distribution $P_0 = 2/3$, $P_1 = 1/3$. As the codebook size K increases, we see that the probability of occurrence $\hat{P}_0$ approaches the target distribution value $P_0 = 2/3$ and the generalized expansion factor approaches the theoretical lower bound 1. □

The separation theorem for shaping codes in Theorem 9 now extends naturally to DM codes.

Theorem 16. *An optimal DM code can be constructed by a concatenation of optimal lossless compression with an optimal DM code for a uniform i.i.d. source, in the sense that the minimum generalized expansion factor can be achieved by such a concatenation.* □

Remark 14. When $P_1 = P_2 = \cdots = P_v = \frac{1}{|\mathcal{Y}|}$, the generalized

expansion factor reduces to

$$F = f\frac{\sum_i \hat{P}_i \log_2 |\mathcal{Y}|}{\log_2 |\mathcal{Y}|} = f. \tag{70}$$

This provides the motivation for designating F by this name.□

Remark 15. We use an example to illustrate the difference between the generalized expansion factor and the normalized conditional divergence introduced in [22] and [21] when the encoder has finite length. Given a ternary source with alphabet $\mathcal{X} = \{\alpha_1, \alpha_2, \alpha_3\}$ and probability distribution $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\}$, consider two codes defined by the mappings

$$\begin{aligned} &\Phi_1 : \{\alpha_1 \to 0, \alpha_2 \to 10, \alpha_3 \to 11\}, \\ &\Phi_2 : \{\alpha_1 \to 00, \alpha_2 \to 10, \alpha_3 \to 11\}. \end{aligned} \tag{71}$$

Their generalized expansion factors for target distribution $\{1/2, 1/2\}$ are

$$F_1 = \frac{3}{2} < F_2 = 2. \tag{72}$$

This suggests that Φ_1 is a better approximation of an optimal DM code for target distribution $\{\frac{1}{2}, \frac{1}{2}\}$ (in fact, Φ_1 is an optimal DM code). The normalized conditional divergences are

$$\begin{aligned} &D(\Phi_1(X)||V|I) = \\ &\frac{1}{2}(\log_2 \frac{1}{1/2}) + \frac{1}{2}(\frac{1}{2}\log_2 \frac{1/2}{1/4} + \frac{1}{2}\log_2 \frac{1/2}{1/4}) = 1 \end{aligned} \tag{73}$$

$$\begin{aligned} &D(\Phi_2(X)||V|I) \\ &= \frac{1}{2}\log_2 \frac{1/2}{1/4} + \frac{1}{4}\log_2 \frac{1/4}{1/4} + \frac{1}{4}\log_2 \frac{1/4}{1/4} = \frac{1}{2}. \end{aligned} \tag{74}$$

We find that $D(\Phi_1(X)||V|I) > D(\Phi_2(X)||V|I)$, which suggests the opposite conclusion that Φ_2 would be a better approximation of the optimal DM code. □

VI. Comparison of DM Performance Measures

In this section, we use a shaping code perspective to study DM codes whose performance is measured using informational divergence and normalized informational divergence.

A. Generalized Expansion Factor and Informational Divergence

In this subsection we study the relationship between the generalized expansion factor and the informational divergence introduced in [3], which is also used as a performance measure for DM codes.

Consider a variable-length code $\phi : \mathcal{X}^{\log_{|\mathcal{X}|} K} \to \mathcal{Y}^*$ with codebook size K. We use $\mathcal{L}$ to denote the set of all codewords generated by this mapping. The leaf probability, or the probability of a codeword y_1^l, is defined as

$$P^{\mathcal{L}}(y_1^l) = P(y_1)P(y_2)\dots P(y_l) = \prod P_i^{N_i(y_1^l)}. \tag{75}$$

This is also the probability of sequence y_1^l generated by an i.i.d. source with distribution $\{P_i\}$. The true probability of codeword y_1^l is the probability of the corresponding source sequence $\phi^{-1}(y_1^l)$. The *informational divergence* (I-divergence) between these two distributions is defined as

$$I = \sum_{y_1^l \in \mathcal{L}} P(\phi^{-1}(y_1^l)) \log_2 \frac{P(\phi^{-1}(y_1^l))}{P^{\mathcal{L}}(y_1^l)}. \tag{76}$$

Now we use the same code for type-II shaping. We set the cost of each symbol to be $C_i = -\log_2 P_i$. The cost of codeword y_1^l is

$$\begin{aligned} W(y_1^l) &= \sum C_i N_i(y_1^l) = -\sum \log_2 P_i^{N_i(y_1^l)} \\ &= -\log_2 \prod P_i^{N_i(y_1^l)} = -\log_2 P^{\mathcal{L}}(y_1^l) \end{aligned} \tag{77}$$

and the total cost of this shaping code, or equivalently the GEF, is

$$\begin{aligned} F(\phi, P_1, \dots, P_v) &= \frac{T(\phi)}{\log_2 |\mathcal{Y}|} \\ &= \frac{1}{\log_{|\mathcal{X}|} K \log_2 |\mathcal{Y}|} \sum_{y_1^l \in \mathcal{L}} P(\phi^{-1}(y_1^l)) W(y_1^l) \\ &= -\frac{\log_2 |\mathcal{X}|}{\log_2 K \log_2 |\mathcal{Y}|} \sum_{y_1^l \in \mathcal{L}} P(\phi^{-1}(y_1^l)) \log_2 P^{\mathcal{L}}(y_1^l). \end{aligned} \tag{78}$$

The I-divergence of this code can then be expressed in terms of its GEF, namely

$$\begin{aligned} I &= \sum_{y_1^l \in \mathcal{L}} P(\phi^{-1}(y_1^l)) \log_2 \frac{P(\phi^{-1}(y_1^l))}{P^{\mathcal{L}}(y_1^l)} \\ &= \sum_{y_1^l \in \mathcal{L}} P(\phi^{-1}(y_1^l)) \log_2 P(\phi^{-1}(y_1^l)) \\ &\quad - \sum_{y_1^l \in \mathcal{L}} P(\phi^{-1}(y_1^l)) \log_2 P^{\mathcal{L}}(y_1^l) \\ &= F\frac{\log_2 K \log_2 |\mathcal{Y}|}{\log_2 |\mathcal{X}|} - H(\mathbf{X}^{\log_{|\mathcal{X}|} K}) \\ &= (F - \frac{H(\mathbf{X})}{\log_2 |\mathcal{Y}|})\frac{\log_2 K \log_2 |\mathcal{Y}|}{\log_2 |\mathcal{X}|}. \end{aligned} \tag{79}$$

Since $\log_2 K$ is a constant, minimizing I is equivalent to minimizing F. This equation shows the relationship between I-divergence and GEF, and also highlights the duality between costly channel coding and DM coding. As shown in Remark 13, Varn codes minimize GEF for a uniform i.i.d. source. Therefore we can conclude the following optimality theorem for Varn codes.

Theorem 17. *Let $\{P_i\}$ be a target distribution. A code $\phi : \mathcal{X}^{\log_{|\mathcal{X}|} K} \to \mathcal{Y}^*$ for a uniform i.i.d. source that minimizes I-divergence is given by a Varn code designed for costs $C_i = -\log_2 P_i$.* □

Remark 16. In [49], Savari showed that Varn codes and reverse Tunstall codes are identical when finding exhaustive prefix-free codes (i.e., when $(K-1)/(|\mathcal{Y}|-1)$ is an integer). Specifically, a Tunstall code designed to compress distribution $\{P_i\}$ and

a Varn code designed for costly channel $\{C_i = -\log_2 P_i\}$ generate identical code trees. Therefore a reverse Tunstall code minimizes the I-divergence when the target distribution is binary (i.e., when $|\mathcal{Y}| = 2$). This was also proved in [3, Proposition 1] using a different method.

However, for non-exhaustive codes this equivalence does not exist, and it remains unknown whether a reverse Tunstall code minimizes I-divergence when the target distribution is non-binary (i.e., when $|\mathcal{Y}| > 2$). Therefore Theorem 17 can be viewed as a generalization of [3, Proposition 1]. □

B. Type-I Shaping Problem and Normalized I-Divergence

Another measure for DM codes used in [3] is normalized I-divergence. In this subsection, we study its properties using the perspective of the type-I shaping problem. Normalized I-divergence is defined as

$$\mathscr{I} = \frac{I}{E(L)}. \tag{80}$$

Using (79), we rewrite this as

$$\begin{aligned}\mathscr{I} &= (F - \frac{H(\mathbf{X})}{\log_2|\mathcal{Y}|})\frac{\log_2 K \log_2|\mathcal{Y}|}{E(L)\log_2|\mathcal{X}|} \\ &= (-f\frac{\sum_i \hat{P}_i \log_2 P_i}{\log_2|\mathcal{Y}|} - \frac{H(\mathbf{X})}{\log_2|\mathcal{Y}|})\frac{\log_2|\mathcal{Y}|}{f} \\ &= \sum_i \hat{P}_i C_i - \frac{H(\mathbf{X})}{f}\end{aligned} \tag{81}$$

where $C_i = -\log_2 P_i$. From equations (53) and (81) we see that asymptotic normalized KL-divergence and normalized I-divergence are identical for i.i.d. distribution matching.

We divide the problem of finding the minimum $\mathscr{I}$ into two parts. First we fix the expansion factor f and find the minimum achievable $\mathscr{I}$, denoted by $\mathscr{I}_{\min}(f)$. Then we find the optimal f to minimize $\mathscr{I}_{\min}(f)$. The result is found by noting the similarity to the type-I shaping problem and invoking Theorem 6.

Theorem 18. *Let ϕ be a prefix-free variable-length mapping with expansion factor f. Let $\{P_i\}$ be the target distribution and set $C_i = -\log_2 P_i$. The minimum normalized I-divergence $\mathscr{I}_{min}(f)$ with fixed f is*

$$\mathscr{I}_{min}(f) = \sum_i \hat{P}_i C_i - \frac{H(\mathbf{X})}{f}, \tag{82}$$

where $\hat{P}_i = \frac{2^{-\mu C_i}}{\sum_i 2^{-\mu C_i}}$ and $H(\hat{Y}) = -\sum_i \hat{P}_i \log_2 \hat{P}_i = H(\mathbf{X})/f$.

Proof: We must solve the following optimization problem, which is closely related to the type-I shaping problem.

$$\begin{aligned}\underset{\hat{P}_i}{\text{minimize}} \quad & \sum_i \hat{P}_i C_i - \frac{H(\mathbf{X})}{f} \\ \text{subject to} \quad & H(\hat{Y}) \geqslant \frac{H(\mathbf{X})}{f} \\ & \sum_i \hat{P}_i = 1.\end{aligned} \tag{83}$$

From Theorem 6, we immediately have

$$\mathscr{I}_{\min}(f) = \sum_i \hat{P}_i C_i - \frac{H(\mathbf{X})}{f} = \sum_i \hat{P}_i C_i - H(\hat{Y}), \tag{84}$$

where $\hat{P}_i = \frac{2^{-\mu C_i}}{\sum_j 2^{-\mu C_j}}$ and $H(\hat{Y}) = -\sum_i \hat{P}_i \log_2 \hat{P}_i = H(\mathbf{X})/f$. ■

The next proposition determines the derivative of $\mathscr{I}_{\min}(f)$ and finds the optimal expansion factor, f_{opt}, that minimizes $\mathscr{I}_{\min}(f)$.

Proposition 19 *The first derivative of $\mathscr{I}_{min}(f)$ is*

$$\frac{\mathrm{d}\mathscr{I}_{min}}{\mathrm{d}f} = \frac{H(\mathbf{X})}{f^2}\frac{\mu-1}{\mu} \quad \mu > 0. \tag{85}$$

Let $f_{opt} = -H(\mathbf{X})/\sum_i P_i \log_2 P_i$. Then $\mathscr{I}_{min}(f)$ is continuous, strictly monotone decreasing on $[\frac{H(\mathbf{X})}{\log_2|\mathcal{Y}|}, f_{opt})$ (or, for $\mu \in [0,1)$) and continuous, strictly monotone increasing on $(f_{opt}, +\infty)$ (or, for $\mu \in (1,\infty)$). When $f = f_{opt}$, $\mathscr{I}_{min}(f_{opt}) = 0$.

Proof: We have studied the behavior of minimum total cost with fixed f in Appendices B and D. Here we use the same technique to study $\mathscr{I}_{\min}(f)$. Note that $\mathscr{I}_{\min}(f)$ is a function of μ. The derivative $\mathrm{d}f/\mathrm{d}\mu$ is already given in equation (124), and it is easy to check that

$$\frac{\mathrm{d}\mathscr{I}_{\min}}{\mathrm{d}\mu} = \frac{(\mu-1)\ln 2 \sum_{i<j} 2^{-\mu(C_i+C_j)}(C_i - C_j)^2}{N^2}. \tag{86}$$

Applying the chain rule along with (86) and (124), we have

$$\frac{\mathrm{d}\mathscr{I}_{\min}}{\mathrm{d}f} = \frac{H(\mathbf{X})}{f^2}\frac{\mu-1}{\mu}. \tag{87}$$

Let $f_{\text{opt}} = \frac{H(\mathbf{X})}{-\sum_i P_i \log_2 P_i}$ (or, equivalently, let $\mu = 1$). Equations (86) and (87) imply that $\mathscr{I}_{\min}(f)$ is continuous, strictly monotone decreasing on $[\frac{H(\mathbf{X})}{\log_2|\mathcal{Y}|}, f_{\text{opt}})$ (or, for $\mu \in [0,1)$) and strictly monotone increasing on $(f_{\text{opt}}, +\infty)$ (or, for $\mu \in (1,\infty)$). The minimum of $\mathscr{I}_{\min}(f)$ is achieved when $f = f_{\text{opt}}$. We have

$$\begin{aligned}\mathscr{I}_{\min}(f_{\text{opt}}) &= \sum_i \hat{P}_i C_i - \frac{H(\mathbf{X})}{f_{\text{opt}}} \\ &= \sum_i \frac{2^{-C_i}}{\sum_j 2^{-C_j}} C_i + \sum_i P_i \log_2 P_i \\ &= -\sum_i \frac{P_i}{\sum_j P_j}\log_2 P_i + \sum_i P_i \log_2 P_i = 0.\end{aligned} \tag{88}$$

This completes the proof. ■

Remark 17. In [5, Sec 5.1], the author studied the minimum KL-divergence between a pmf $\{\hat{P}_i\}$ and the target distribution $\{P_i\}$, where each $\hat{P}_i$ is associated with a cost w_i and the average cost of the pmf is upper bounded by C. The analysis is similar

to the analysis in Proposition 19 if we specialize to the case where $w_i = -\log_2 P_i = C_i$. The KL-divergence is

$$\begin{aligned} D = D(\hat{P}_i||P_i) &= \sum_i \hat{P}_i \log_2 \frac{\hat{P}_i}{P_i} \\ &= -\sum_i \hat{P}_i \log_2 P_i + \sum_i \hat{P}_i \log_2 \hat{P}_i \\ &= \sum_i \hat{P}_i C_i - H(\hat{P}). \end{aligned} \tag{89}$$

By combining equation (81) with Lemma 5, we have

$$\begin{aligned} \mathscr{I} &= \sum_i \hat{P}_i C_i - \frac{H(\mathbf{X})}{f} = \sum_i \hat{P}_i C_i - H(\mathbf{Y}) \\ &\geqslant \sum_i \hat{P}_i C_i - H(\hat{P}) = D, \end{aligned} \tag{90}$$

with equality if and only if the output process generated by ϕ approximates an i.i.d. process (Remark 4).

The minimum KL-divergence with average cost upper bounded by a specified average cost C, denoted by $D(C)$, was also studied in [5, Sec 5.1]. The pmf that achieves $D(C)$ is

$$\hat{P}_i = \frac{2^{-\mu C_i}}{\sum_j 2^{-\mu C_j}}, \quad \sum_i \hat{P}_i C_i = C, \tag{91}$$

when $C \leqslant \sum P_i C_i$, or $\mu \geqslant 1$. When $C > \sum P_i C_i$, by setting $\mu = 1$, we have $\hat{P}_i = P_i$, $\sum_i \hat{P}_i C_i < C$, and

$$D(C) = \sum_i \hat{P}_i C_i - H(\hat{P}) = 0. \tag{92}$$

By comparing equations (82) and (91), we conclude that

$$\mathscr{I}_{\min}(f) = D(C), \tag{93}$$

where

$$C = \sum_i \hat{P}_i C_i \quad f = \frac{H(\mathbf{X})}{-\sum_i \hat{P}_i \log_2 \hat{P}_i} \quad \hat{P}_i = \frac{2^{-\mu C_i}}{\sum_j 2^{-\mu C_j}}, \tag{94}$$

when $\mu \geqslant 1$, or equivalently when $C \leqslant \sum P_i C_i$ and $f \geqslant f_{\text{opt}}$. □

The derivative $\mathrm{d}D(C)/\mathrm{d}C$ was found in [5, Sec. 5.1]. Using the chain rule, we know that

$$\frac{\mathrm{d}D(C)}{\mathrm{d}C} = \begin{cases} \frac{\mathrm{d}\mathscr{I}_{\min}}{\mathrm{d}f} \frac{\mathrm{d}f}{\mathrm{d}C} & \text{when } C \leqslant \sum P_i C_i \ (\mu \geqslant 1), \\ 0 & \text{when } C > \sum P_i C_i. \end{cases} \tag{95}$$

From (34) we have

$$\frac{\mathrm{d}h}{\mathrm{d}C} = \frac{\mathrm{d}\frac{H(\mathbf{X})}{f}}{\mathrm{d}C} = \mu \quad \Rightarrow \quad \frac{\mathrm{d}f}{\mathrm{d}C} = -\frac{\mu f^2}{H(\mathbf{X})}. \tag{96}$$

By combining this with (85), we have

$$\frac{\mathrm{d}D(C)}{\mathrm{d}C} = \begin{cases} 1-\mu & \text{when } C \leqslant \sum P_i C_i \ (\mu \geqslant 1), \\ 0 & \text{when } C > \sum P_i C_i. \end{cases} \tag{97}$$

Therefore Proposition 19 allows us to recover the derivative of $D(C)$.

Remark 18. In [6, Sec. V], a bound on the rate of a prefix-free variable-length DM code in the vicinity of $\mathscr{I} = 0$ was given. Proposition 19 gives an explicit relationship between $\mathscr{I}$ and the code rate over a wider range of rates. □

Remark 19. In [55, Section 6.2.2], Soriaga considered the case where system requirements dictate that the expansion factor of the DM code encoder can not exceed f_0, where $f_0 < f_{opt}$. In such a case, the code cannot be an optimal DM code for target distribution $\{P_i\}$. We may try to approximate an optimal DM code by designing a code with $f \leqslant f_0$ that minimizes the asymptotic normalized KL-divergence, $\lim_{l\to\infty} \frac{1}{l} D(Y_1^l || \tilde{Y}_1^l)$. We denote by $\mathscr{D}(f_0)$ the minimum possible value of this divergence. The relationship between $\mathscr{D}(f_0)$ and f_0 for a finite-order Markov target distribution was given in [55] and the result is also applicable to the i.i.d. case considered here. Since the asymptotic normalized KL-divergence for a fixed f is lower bounded by $\mathscr{I}_{\min}(f)$ (Theorem 18) and $\mathscr{I}_{\min}(f)$ is strictly monotone decreasing when $f \leqslant f_0 < f_{opt}$ (Proposition 19), we have

$$\begin{aligned} \mathscr{D}(f_0) = \mathscr{I}_{\min}(f_0) &= \sum_i \hat{P}_i C_i - \frac{H(\mathbf{X})}{f_0} \\ &= \sum_i \hat{P}_i \frac{(-\log_2 \hat{P}_i - \log_2 N)}{\mu} - \frac{H(\mathbf{X})}{f_0} \\ &= \frac{-\sum_i \hat{P}_i \log_2 \hat{P}_i}{\mu} - \sum_i P_i \frac{\log_2 N}{\mu} - \frac{H(\mathbf{X})}{f_0} \\ &= \frac{H(\mathbf{X})}{\mu f_0} - \frac{\log_2 N}{\mu} - \frac{H(\mathbf{X})}{f_0}, \end{aligned} \tag{98}$$

where μ and N are constants such that $N = \sum_i 2^{-\mu C_i}$ and $\sum_i -\hat{P}_i \log \hat{P}_i = H(\mathbf{X})/f_0$, with $\hat{P}_i = \frac{1}{N} 2^{-\mu C_i}$. Because the code that achieves this lower bound with expansion factor f_0 is an optimal type-I shaping code, based on the equivalence theorem we can extend the result in [55, Section 6.2.2] by concluding that this code is an optimal DM code for target distribution $\{\hat{P}_i\}$. □

Remark 20. In Theorem 18, we have shown that when $\mathscr{I} \to 0$, then $f \to f_{\text{opt}}$. This implies that the GEF of the code satisfies

$$F = \frac{f\mathscr{I} + H(\mathbf{X})}{\log_2 |\mathcal{Y}|} \to \frac{H(\mathbf{X})}{\log_2 |\mathcal{Y}|}. \tag{99}$$

Similarly, as shown in Appendix D, when $F \to H(\mathbf{X})/\log_2 |\mathcal{Y}|$ (or, equivalently, when total cost $T \to H(\mathbf{X})$), then $f \to f_{\text{opt}}$ and

$$\mathscr{I} = \frac{F - \frac{H(\mathbf{X})}{\log_2 |\mathcal{Y}|}}{f} \to 0. \tag{100}$$

In view of the equivalence between asymptotic normalized KL-divergence and $\mathscr{I}$, these observations extend Theorem 15 by providing bounds on asymptotic normalized KL-divergence in the vicinity of $F = H(\mathbf{X})/\log_2 |\mathcal{Y}|$. □

VII. Experimental Results

A. Optimal Data Shaping Code for MLC Flash Memory

We evaluated the performance of shaping codes on a multilevel-cell (MLC) NAND flash memory. In MLC flash, the cells are arranged in a rectangular array (also called a *block*) and each row of cells is called a *wordline*. The cells can be programmed to four different voltage levels, denoted $\{0, 1, 2, 3\}$, so each cell can store two bits of information. It was shown in [35], [34] that MLC flash memory can be modeled as a costly channel with alphabet $\{0, 1, 2, 3\}$, where the cost of the erase level 0 can be taken to be $C_0 = 0$. Using the methodology described in [35], the cost vector for the memory was found empirically to be

$$\mathcal{C} = [0, 0.58, 0.87, 1.29]. \tag{101}$$

From Theorem 7, we know that the total cost is a decreasing function of the expansion factor. To assess the performance of optimal shaping, and to permit a comparison to the direct-shaping code in [35], [34], we applied a rate-1, type-I shaping code to the ASCII representation of the English-language text of *The Count of Monte Cristo*. The "optimal" shaping scheme was designed according to the principles suggested by the equivalence theorem and separation theorem. We first compressed the file using the LZ77 algorithm. The observed compression rate was $g = 1/2.740$. We then used Theorem 6 to compute the target symbol occurrence probabilities of a shaping code that minimizes average cost for a uniform i.i.d. source, a cost vector $\mathcal{C}$, and the expansion factor $f' = f/g = 2.740$. The resulting symbol occurrence probability distribution was given by

$$\hat{P} = [0.8606, 0.0989, 0.0335, 0.0070]. \tag{102}$$

Using Theorem 8, we computed the costs for the equivalent code that minimizes total cost, yielding the cost vector

$$\mathcal{C}' = [0.2167, 3.3378, 4.8983, 7.1585]. \tag{103}$$

We constructed a Varn code with codebook size $K = 256$ based on the cost vector $\mathcal{C}'$. This code is a length-8, type-II shaping code and the concatenation of the compression and the Varn code is a rate-1, type-I shaping code. The expansion factor of the Varn code is 2.768, which is close to the expansion factor of the optimal type-II shaping code for cost vector $\mathcal{C}'$, where $f_{opt} = 2.737$. Its codeword length distribution is shown in Fig. 4.

To characterize the performance of the designed shaping code, we performed a program/erase (P/E) cycling experiment on the MLC flash memory by repeating the following steps, which collectively represent one P/E cycle. The experiment was conducted with the uncoded source data, and then with the output data from the shaping code.

- Erase the MLC flash memory block.
- Program the MLC flash memory.
- For each successive programming cycle, "rotate" the data, so the data that was written on the ith wordline is written on the $(i+1)$st wordline, wrapping around the last wordline to the first wordline.
- After every 100 P/E cycles, erase the block and program pseudo-random data. Then perform a read operation, record bit errors, and calculate the bit error rate.

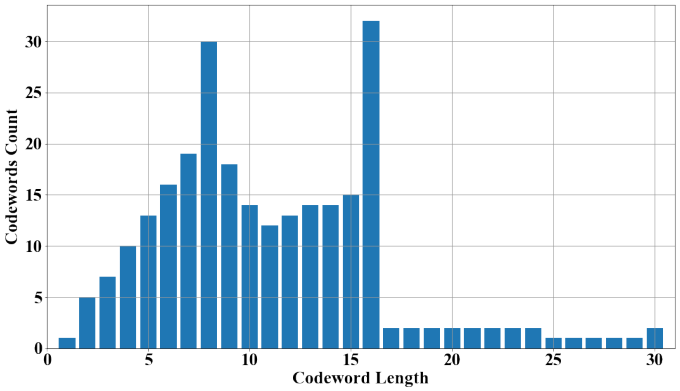
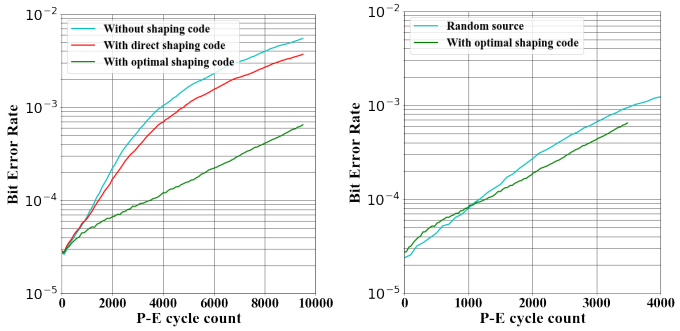


Fig. 4: Codeword length distribution of Varn code with the codebook size $K = 256$ for English-language text.

Fig. 5(a) shows the average bit error rates (BERs) for the uncoded source data, the direct shaping code [35], and the optimal shaping code. The results indicate that the optimal shaping code provides a significant increase in the memory lifetime compared to no shaping and direct shaping.



(a) (b)

Fig. 5: BER performance for English-language text.

As a way of comparing the performance of optimal shaping to that of data compression alone, we rescaled the P/E cycle count of the shaping code by the compression ratio 2.740 and compared the result to P/E cycling of pseudo-random data. This corresponds to a BER comparison based upon the total amount of source data stored in the memory. The results, shown in Fig. 5(b), indicate that the performance of optimal shaping is superior to data compression alone as a function of total source data written.

A similar experiment was conducted for a Chinese-language text, *Collected Works of Lu Xun, Volumes 1–4*, represented using UTF-16LE encoding. We constructed a Varn code with codebook size $K = 256$ based on the cost vector

$$\mathcal{C}' = [0.4222, 2.6647, 3.7860, 5.4099]. \tag{104}$$

The expansion factor of the the Varn code was 1.751, which is close to the expansion factor of the optimal type-II shaping code, $f_{opt} = 1.759$. Its codeword length distribution is shown in Fig. 6. The BER results are shown in Fig. 7(a) and Fig. 7(b).

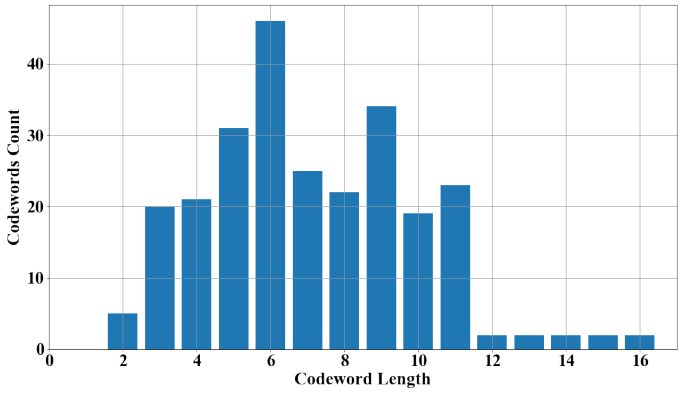
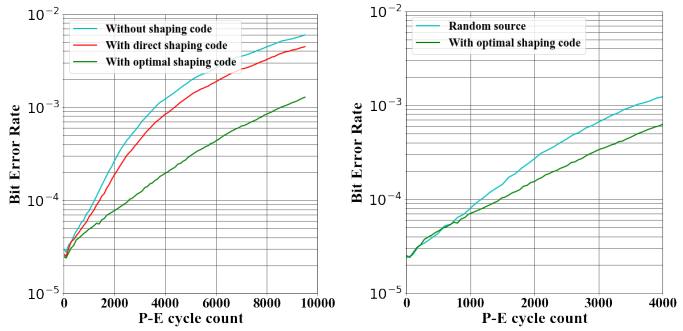


Fig. 6: Codeword length distribution of Varn code with codebook size $K = 256$ for Chinese-language text.



(a) (b)

Fig. 7: BER performance for Chinese-language text.

B. Varn Codes for Distribution Matching

Remark 13 shows that the upper bound on the generalized expansion factor of Varn codes decreases as the codebook size increases. This suggests that as the codebook size of a Varn code increases, the approximation to an optimal DM code should improve. In this subsection, we empirically tested this premise by constructing Varn codes with codebook size $K = 100$, 1000 and 10000, respectively, for a target distribution $\{P_0, P_1\} = \{\frac{2}{3}, \frac{1}{3}\}$. The measure of goodness we used here was similar to the serial test in [48, Section 2.11], namely KL-divergences for patterns of increasing length. Codeword sequences with 10000 codewords were generated using the random number sequence collected from [47]. The first 71514 bits in codeword sequences were used for comparison (71514 was the length of the codeword sequence generated by the Varn code with codebook size $K = 100$). The probability of occurrence of length 1, 2 and 3 patterns was calculated. For example, we define the probability of occurrence of '10' (P'_{10}) and '101' (P'_{101}) in codeword sequence y_1^l as

$$P'_{10} = \frac{\{\text{the number of subsequences } y_i^{i+1} = \text{'10'}\}}{l-1}, \quad (105)$$

$$P'_{101} = \frac{\{\text{the number of subsequences } y_i^{i+2} = \text{'101'}\}}{l-2}. \quad (106)$$

The first-, second-, and third-order KL-divergences between P' and distribution $\{P_0, P_1\} = \{\frac{2}{3}, \frac{1}{3}\}$ were calculated, using the following definitions:

$$I_1 = \sum_{i\in\{0,1\}} P'_i \log_2 \frac{P'_i}{P_i} \quad (107)$$

$$I_2 = \sum_{i\in\{0,1\}} \sum_{j\in\{0,1\}} P'_{ij} \log_2 \frac{P'_{ij}}{P_i P_j} \quad (108)$$

$$I_3 = \sum_{i\in\{0,1\}} \sum_{j\in\{0,1\}} \sum_{k\in\{0,1\}} P'_{ijk} \log_2 \frac{P'_{ijk}}{P_i P_j P_k}. \quad (109)$$

The results are shown in Table I. The divergences decrease as K increases, indicating that the approximation to an i.i.d. sequence with target distribution $\{P_0, P_1\} = \{\frac{2}{3}, \frac{1}{3}\}$ is improving.

	P'_0	I_1	I_2	I_3
$K = 100$	0.6447	0.0015	0.0032	0.0055
$K = 1000$	0.6498	0.00091	0.0018	0.0027
$K = 10000$	0.6602	0.00014	0.00027	0.00028

TABLE I: First- , second-, and third-order KL-divergence

VIII. Conclusion

In this paper, we studied information-theoretic properties and performance limits of a general class of shaping codes. We determined the asymptotic symbol occurrence probability distribution, and used it to determine the minimum achievable average cost for a type-I shaping code. Using these results, we determined the minimum total cost and optimal expansion factor for a type-II shaping code. A consequence of this analysis is an equivalence theorem, stating that a type-I shaping code with a given expansion factor and a cost vector can be realized by a type-II shaping code. We then proved a separation theorem stating that optimal shaping can be achieved by a concatenation of optimal lossless compression and optimal shaping for a uniform i.i.d. source. Experimental results showed that optimal shaping can provide a significant increase in flash memory lifetime when applied to English-language and Chinese-language texts, providing total data capacity greater than that achieved by data compression alone.

We also studied properties of prefix-free variable-length distribution matching (DM) codes from the perspective of shaping. We characterized optimal DM codes in terms of the asymptotic normalized divergence and showed that when the divergence equals zero, a DM code encoder generates a codeword sequence that looks i.i.d., with symbol occurrence probability equal to the target distribution. We showed that optimal type-II shaping codes can be used to construct optimal DM codes. This

suggested the definition of the generalized expansion factor as a performance measure for DM codes and implied a separation theorem for DM codes. We also established the relationship between the generalized expansion factor and the informational divergence of a DM code. The relationship between the type-I shaping problem and the minimization of normalized informational divergence was also studied. Simulation results showed an increase in distribution matching performance of Varn codes designed for a Bernoulli distribution as the codebook size increases.

ACKNOWLEDGMENT

This work was supported in part by National Science Foundation (NSF) Grant CCF-1619053. The authors acknowledge the helpful comments of the anonymous reviewers and the associate editor.

REFERENCES

[1] J. Abrahams, "Variable-length unequal cost parsing and coding for shaping," *IEEE Trans. Inf. Theory*, vol. 44, no. 4, pp. 1648–1649, Jul. 1998.

[2] R. Adler, D. Coppersmith, and M. Hassner, "Algorithms for sliding block codes," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 1, pp. 5–22, Jan. 1983.

[3] R. A. Amjad and G. Böcherer, "Fixed-to-variable length distribution matching," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Istanbul, Turkey, Jul. 2013, pp. 1511–1515.

[4] S. Baur and G. Bcherer, "Arithmetic distribution matching," in *Proc. Int. ITG Conf. Source Channel Coding (SCC)*, Hamburg, Germany, Feb. 2-5, 2015, pp. 1–6.

[5] G. Böcherer, "Capacity-achieving probabilistic shaping for noisy and noiseless channels," PhD thesis, RWTH Aachen University, 2012. [Online]. Available: http://www.georg-boecherer.de/capacityAchievingShaping.pdf.

[6] G. Böcherer and R. A. Amjad, "Informational divergence and entropy rate on rooted trees with probabilities," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2014, pp. 176–180.

[7] G. Böcherer and R. Mathar, "LDPC codes with zero shaping gap," in *Proc. IEEE Inf. Theory Workshop (ITW)*, 2011.

[8] G. Böcherer and R. Mathar, "Matching dyadic distributions to channels," in *Proc. Data Compression Conf. (DCC)*, 2011, pp. 23–32.

[9] G. Böcherer, V. C. da Rocha Jr., C. Pimentel, and R. Mathar, "On the capacity of constrained systems," in *Proc. Int. ITG Conf. Source Channel Coding (SCC)*, 2010.

[10] G. Böcherer, F. Steiner, and P. Schulte, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4651–4665, Dec. 2015.

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.

[12] I. Csiszár, "Simple proofs of some theorems on noiseless channels," *Inf. Contr.*, vol. 14, pp. 285–298, 1969.

[13] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge, U.K.: Cambridge University Press, 2011.

[14] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission*. Hoboken, NJ, USA: Wiley, 2002.

[15] G. Forney, R. Gallager, G. Lang, F. Longstaff and S. Qureshi, "Efficient modulation for band-limited channels," *IEEE J. Select. Areas Commun.*, vol. 2, no. 5, pp. 632–647, Sep. 1984.

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

[17] E. N. Gilbert, "Coding with digits of unequal cost," *IEEE Trans. Inf. Theory*, vol. 41, no. 2, pp. 596–600, Mar. 1995.

[18] M. J. Golin and G. Rote, "A dynamic programming algorithm for constructing optimal prefix-free codes with unequal letter costs," *IEEE Trans. Inf. Theory*, vol. 44, no. 5, pp. 1770–1781, Sep. 1998.

[19] M. Guazzo, "A general minimum-redundancy source-coding algorithm," *IEEE Trans. Inf. Theory*, vol. 26, no. 1, pp. 15–25, Jan. 1980.

[20] Y. C. Gültekin, W. J. van Houtum, S. Şerbetli, and F. M. Willems, "Constellation shaping for IEEE 802.11," in *Proc. IEEE Int. Symp. Personal, Indoor, Mobile Radio Commun. (PIMRC)*, 2017.

[21] T. Han, *Information-Spectrum Methods in Information Theory*. Berlin, Germany: Springer-Verlag, 2003, (originally published by Baifukan 1998 in Japanese).

[22] T. Han and O. Uchida, "Source code with cost as a nonuniform random number generator," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 712–717, Mar. 2000.

[23] C. D. Heegard, B. H. Marcus, and P. H. Siegel,"Variable-length state splitting with applications to average runlength-constrained (ARC) codes," *IEEE Trans. Inf. Theory*, vol. 37, no. 3, pp. 759–777, May 1991.

[24] K. Iwata, M. Morii, and T. Uyematsu, "An efficient universal coding algorithm for noiseless channel with symbols of unequal cost," *IEICE Trans. Fundamentals*, vol. E80-A, no. 11, pp. 2232–2237, Nov. 1997.

[25] A. Jagmohan, M. Franceschini, L. A. Lastras-Montaño and J. Karidis, "Adaptive endurance coding for NAND Flash," in *Proc. IEEE GLOBECOM Workshops*, Dec. 2010, pp. 1841–1845.

[26] J. Justesen and T. Høholdt, "Maxentropic Markov chains," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 4, pp. 665–667, Jul. 1984.

[27] R. Karp, "Minimum-redundancy coding for the discrete noiseless channel," *IRE IEEE Trans. Inf. Theory*, vol. 7, no. 1, pp. 27–38, Jan. 1961.

[28] A. Khandekar, R. J. McEliece, and E. Rodemich, "The discrete noiseless channel revisited," in *Proc. 1999 Int. Symp. Communication Theory and Applications*, pp. 115–137, 1999.

[29] A. S. Khayrallah and D. L. Neuhoff, "Coding for channels with cost constraints," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 854–867, May 1996.

[30] V. Y. Krachkovsky, R. Karabed, S. Yang, and B. A. Wilson, "On modulation coding for channels with cost constraints," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Honolulu, Hawaii, Jun. 29-Jul. 4, 2014, pp. 421–425.

[31] R. M. Krause, "Channels which transmit letters of unequal duration," *Inf. Contr.*, vol. 55, pp. 13–24, 1962.

[32] F. R. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 913–929, May 1993. May 1993.

[33] A. Lempel, S. Even, and M. Cohn, "An algorithm for optimal prefix parsing of a noiseless and memoryless channel," *IEEE Trans. Inf. Theory*, vol. 19, no. 2, pp. 208–214, Mar. 1973.

[34] Y. Liu and P. H. Siegel, "Shaping codes for structured data," *7th Annual Non-Volatile Memories Workshop (NVMW)*, La Jolla, CA, Mar. 6–8, 2016.

[35] Y. Liu and P. H. Siegel, "Shaping codes for structured data," in *Proc. IEEE Globecom*, Washington, D.C., Dec. 4-8, 2016, pp. 1–5.

[36] Y. Liu, P. Huang and P. H. Siegel, "Performance of shaping codes for flash memory," *8th Annual Non-Volatile Memories Workshop (NVMW)*, La Jolla, CA, Mar. 12–14, 2017.

[37] Y. Liu, P. Huang and P. H. Siegel, "Performance of optimal data shaping codes," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 25-30, 2017, pp. 1003–1007.

[38] Y. Liu, P. Huang, A. W. Bergman and P. H. Siegel, "Optimal data shaping code design," *9th Annual Non-Volatile Memories Workshop (NVMW)*, La Jolla, CA, Mar. 11-13, 2018.

[39] R. S. Marcus, "Discrete noiseless coding," Master's thesis, Massachusetts Institute of Technology, 1957.

[40] R. J. McEliece, *The Theory of Information and Coding*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2001.

[41] R. J. McEliece and E. R. Rodemich, "A maximum entropy Markov chain," in *Proc. 17th Conf. Inf. Sciences and Systems*, Johns Hopkins University, Mar. 1983, pp. 245–248.

[42] S. W. McLaughlin and A. S. Khayrallah, "Shaping codes constructed from cost-constrained graphs," *IEEE Trans. Inf. Theory*, vol. 43, no. 2, pp. 692–699, Mar. 1997.

[43] K. Mehlhorn, "An efficient algorithm for constructing nearly optimal prefix codes," *IEEE Trans. Inf. Theory*, vol. 26, no. 5, pp. 513–517, Sep. 1980.

[44] M. Mondelli, S. H. Hassani, and R. L. Urbanke, "How to achieve the capacity of asymmetric channels," *IEEE Trans. Inf. Theory*, vol. 64, no. 5, pp. 3371–3393, 2018.

[45] M. Nishiara and H. Morita, "On the AEP of word-valued sources," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 1116–1120, May 2000.

[46] T. V. Ramabadran, "A coding scheme for m-out-of-n codes," *IEEE Trans. Commun.*, vol. 38, no. 8, pp. 1156–1163, Aug. 1990.

[47] RANDOM.ORG, 2018. [online]. Available: https://www.random.org.
[48] A. Rukhin et al., "A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications," NIST Special Publication 800-22 Revision 1a, Apr. 2010.
[49] S. A. Savari, "Some notes on Varn coding," *IEEE Trans. Inf. Theory*, vol. 40, no. 1, pp. 181–186, Jan. 1994.
[50] S. A. Savari and R. G. Gallager, "Arithmetic coding for finite-state noiseless channels," *IEEE Trans. Inf. Theory*, vol. 40, no. 1, pp. 100–107, Jan. 1994.
[51] P. Schulte and G. Böcherer, "Constant composition distribution matching," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 430–434, Jan. 2016.
[52] P. Schulte and F. Steiner, "Divergence-optimal fixed-to-fixed length distribution matching with shell mapping," *IEEE Wireless Commun. Letters,* vol.8, no.2, pp. 620-623, Apr. 2019.
[53] C. E. Shannon, "A mathematical theory of communication, Part I, Part II," *Bell Syst. Tech. J,* vol. 27, pp. 379–423, 1948.
[54] E. Sharon, et al., "Data Shaping for Improving Endurance and Reliability in Sub-20nm NAND," presented at *Flash Memory Summit*, Santa Clara, CA, Aug. 4–7, 2014.
[55] J. Soriaga, *On Near-Capacity Code Design for Partial-Response Channels.* Ph.D. thesis, University of California, San Diego, La Jolla, CA, USA, Mar. 2005.
[56] O. Uchida, "Maximum generating rate of the variable-length nonuniform random number," in *Proc. IEEE Inf. Theory Workshop*, Cairns, Australia, Sep. 2-7, 2001, pp. 141–143.
[57] G. Ungerboeck, "Huffman shaping," in *Codes, Graphs, and Systems*, R. Blahut and R. Koetter, Eds. Springer, 2002, ch. 17, pp. 299–313.
[58] B. Varn, "Optimal variable length codes (arbitrary symbol cost and equal code word probability)," *Inform. Contr.*, vol. 19, pp. 289–301, 1971.
[59] K. Visweswariah, S. R. Kulkarni and S. Verdu, "Source codes as random number generators," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 462–471, Mar. 1998.
[60] A. Wald, *Sequential Analysis*. Courier Corporation, 1973.

APPENDIX A
PROOF OF LEMMA 4

Proof: In this proof, without loss of generality, we will assume $q = 1$. First, we evaluate the expectation of the sequence of random variable $\{N_i(\phi(X_1^{M_l}))\}_{l=1}^{\infty}$. Combining Lemma 3 with equation (14), we have

$$\begin{aligned}\lim_{l\to\infty}\frac{1}{l}E(N_i(\phi(X_1^{M_l}))) &= \lim_{l\to\infty}\frac{1}{l}E(N_i(\phi(X)))E(M_l)\\ &= E(N_i(\phi(X)))\lim_{l\to\infty}\frac{1}{l}E(M_l)\\ &= E(N_i(\phi(X)))\frac{1}{E(L)}.\end{aligned} \tag{110}$$

Similarly, we have

$$\begin{aligned}&\lim_{l\to\infty}\frac{1}{l}E(N_i(\phi(X_1^{M_l-1})))\\ &\quad= \lim_{l\to\infty}\frac{1}{l}E(N_i(\phi(X_1^{M_l})) - N_i(\phi(X_{M_l})))\\ &\quad= E(N_i(\phi(X)))\frac{1}{E(L)} - \lim_{l\to\infty}\frac{1}{l}E(N_i(\phi(X_{M_l})))\\ &\quad= E(N_i(\phi(X)))\frac{1}{E(L)},\end{aligned} \tag{111}$$

where $\lim_{l\to\infty}\frac{1}{l}E(N_i(\phi(X_{M_l}))) = 0$ follows from Lemma 2.

By definition,

$$E(N_i(\phi(X_1^{M_l}))) = \sum_{y_1^l}\sum_{x_1^{M_l}\in\mathcal{G}_\phi(y_1^l)} P(x_1^{M_l})N_i(\phi(x_1^{M_l})). \tag{112}$$

Since $S_{M_l} > l$, we have $N_i(\phi(x_1^{M_l})) \geqslant N_i(y_1^l)$ and $E(N_i(y_1^l))$ can be bounded as follows

$$\begin{aligned}E(N_i(Y_1^l)) &= \sum_{y_1^l} N_i(y_1^l)Q(y_1^l)\\ &= \sum_{y_1^l} N_i(y_1^l)\sum_{x_1^{M_l}\in\mathcal{G}_\phi(y_1^l)} P(x_1^{M_l})\\ &\leqslant \sum_{y_1^l} N_i(\phi(x_1^{M_l}))\sum_{x_1^{M_l}\in\mathcal{G}_\phi(y_1^l)} P(x_1^{M_l})\\ &= \sum_{y_1^l}\sum_{x_1^{M_l}\in\mathcal{G}_\phi(y_1^l)} P(x_1^{M_l})N_i(\phi(x_1^{M_l}))\\ &= E(N_i(\phi(X_1^{M_l}))).\end{aligned} \tag{113}$$

Similarly, $N_i(\phi(x_1^{M_l-1})) \leqslant N_i(\phi(y_1^l))$ and $E(N_i(Y_1^l))$ is lower bounded by

$$\begin{aligned}E(N_i(Y_1^l)) &= \sum_{y_1^l} N_i(y_1^l)\sum_{x_1^{M_l}\in\mathcal{G}_\phi(y_1^l)} P(x_1^{M_l})\\ &\geqslant \sum_{y_1^l}\sum_{x_1^{M_l}\in\mathcal{G}_\phi(y_1^l)} P(x_1^{M_l})N_i(\phi(x_1^{M_l-1}))\\ &= E(N_i(\phi(X_1^{M_l-1}))).\end{aligned} \tag{114}$$

Equations (113) and (114) imply that

$$\begin{aligned}\limsup_{l\to\infty}\frac{1}{l}E(N_i(Y_1^l)) &\leqslant \liminf_{l\to\infty}\frac{1}{l}E(N_i(\phi(X_1^{M_l})))\\ &= E(N_i(\phi(X)))\frac{1}{E(L)}\end{aligned} \tag{115}$$

and

$$\begin{aligned}\liminf_{l\to\infty}\frac{1}{l}E(N_i(Y_1^l)) &\geqslant \limsup_{l\to\infty}\frac{1}{l}E(N_i(\phi(X_1^{M_l-1})))\\ &= E(N_i(\phi(X)))\frac{1}{E(L)}.\end{aligned} \tag{116}$$

Thus we conclude that

$$\hat{P}_i = \lim_{l\to\infty}\frac{1}{l}E(N_i(Y_1^l)) = E(N_i(\phi(X)))\frac{1}{E(L)}. \tag{117}$$

■

APPENDIX B
PROOF OF THEOREM 7

Proof: We solve the optimization problem:

$$\begin{aligned}\underset{\hat{P}_i,f}{\text{minimize}}\quad & f\sum_i \hat{P}_i C_i\\ \text{subject to}\quad & H(\hat{Y}) \geqslant \frac{H(\mathbf{X})}{f}\\ & \sum_i \hat{P}_i = 1.\end{aligned} \tag{118}$$

We divide this optimization problem into two parts. First we fix expansion factor f and find the minimum achievable average cost using Theorem 6. Then we find the optimal f to minimize the total cost. The optimization problem then becomes:

$$\begin{aligned} \underset{f}{\text{minimize}} \quad & \underset{\hat{P}_i}{\text{minimize}} \quad \sum_i f\hat{P}_i C_i \\ \text{subject to} \quad & H(\hat{Y}) \geqslant \frac{H(\mathbf{X})}{f} \\ & \sum_i \hat{P}_i = 1. \end{aligned} \tag{119}$$

For fixed f, the symbol occurrence probability corresponding to the minimum average cost is

$$P_i' = \frac{1}{N} 2^{-\mu C_i}, \tag{120}$$

where μ is a constant such that

$$\frac{1}{f} = \frac{-\sum_i P_i' \log_2 P_i'}{\log_2 |\mathcal{X}|} \tag{121}$$

and N is the normalization factor

$$N = \sum_i 2^{-\mu C_i}. \tag{122}$$

Treating f as a function of μ, and calculating $\mathrm{d}f/\mathrm{d}\mu$, we see that $f(\mu)$ is a monotone increasing function of μ, with $f(0) = \frac{H(\mathbf{X})}{\log_2 |\mathcal{Y}|}$. Specifically, replacing the P_i' in (121) by (120) gives

$$f = \frac{NH(\mathbf{X})}{(\sum_i \mu C_i 2^{-\mu C_i}) + N \log_2 N}. \tag{123}$$

The derivative of f with respect to μ is

$$\frac{\mathrm{d}f}{\mathrm{d}\mu} = \frac{\mu \ln 2 H(\mathbf{X}) \sum_{i<j} 2^{-\mu(C_i+C_j)} (C_i - C_j)^2}{[(\sum_i \mu C_i 2^{-\mu C_i}) + N \log_2 N]^2}, \tag{124}$$

which is easily seen to be positive for $\mu > 0$.

The optimization problem is then equivalent to:

$$\begin{aligned} \text{minimize} \quad & T(\mu) = H(\mathbf{X}) \frac{\sum_i C_i 2^{-\mu C_i}}{\sum_i \mu C_i 2^{-\mu C_i} + N \log_2 N} \\ \text{subject to} \quad & \mu \geqslant 0. \end{aligned} \tag{125}$$

Calculating $\mathrm{d}T/\mathrm{d}\mu$, we see that its sign is the negative of the sign of $\log_2 N$. Specifically, we find that

$$\frac{\mathrm{d}T}{\mathrm{d}\mu} = -\ln 2 H(\mathbf{X}) \log_2 N \frac{N \sum_i C_i^2 2^{-\mu C_i} - (\sum_i C_i 2^{-\mu C_i})^2}{(\sum_i \mu C_i 2^{-\mu C_i} + N \log_2 N)^2}. \tag{126}$$

Observe that

$$\begin{aligned} & N \sum_i C_i^2 2^{-\mu C_i} - (\sum_i C_i 2^{-\mu C_i})^2 \\ &= \sum_{i,j} C_i^2 2^{-\mu(C_i+C_j)} - \sum_i C_i^2 2^{-2C_i} - \sum_{i<j} 2C_iC_j 2^{-\mu(C_i+C_j)} \\ &= \sum_{i<j} (C_i^2 + C_j^2 - 2C_iC_j) 2^{-\mu(C_i+C_j)} \\ &= \sum_{i<j} (C_i - C_j)^2 2^{-\mu(C_i+C_j)} > 0. \end{aligned} \tag{127}$$

Therefore,

$$\frac{\mathrm{d}T}{\mathrm{d}\mu} = -\ln 2 H(\mathbf{X}) \log_2 N \frac{\sum_{i<j} (C_i - C_j)^2 2^{-\mu(C_i+C_j)}}{(\sum_i \mu C_i 2^{-\mu C_i} + N \log_2 N)^2}. \tag{128}$$

The claimed relationship between the sign of $\mathrm{d}T/\mathrm{d}\mu$ and the sign of $\log_2 N$ is then evident.

It follows that if $C_1 = 0$, then $N = \sum_i 2^{-\mu C_i} > 1$, implying that $T(\mu)$ is a monotone decreasing function on $[0, \infty)$. On the other hand, if $C_1 > 0$, then when $\mu = 0$, we have $N = |\mathcal{Y}|$, so $\log_2 N > 0$. Since $\mathrm{d}N/\mathrm{d}\mu < 0$ for all μ, we conclude that T will decrease as μ increases, reaching a minimum at $N = 1$. Beyond that point, $\mathrm{d}N/\mathrm{d}\mu > 0$.

Thus, the corresponding expansion factor that achieves the minimum total cost is

$$f = \frac{H(\mathbf{X})}{-\sum_i \hat{P}_i \log_2 \hat{P}_i} \tag{129}$$

where $\hat{P}_i = 2^{-\mu C_i}$, and μ is a positive constant satisfying $\sum_i 2^{-\mu C_i} = 1$. ■

APPENDIX C
PROOF OF THEOREM 9

Before proving the separation theorem, we first design a type-II shaping code for the uniform source. The *average codeword cost* $C(K)$ for a Varn code with codebook size K is bounded by

$$\log_2 K/\mu \leqslant C(K) \leqslant \log_2 K/\mu + \max_i C_i \tag{130}$$

where μ is a constant such that

$$\sum_i 2^{-\mu C_i} = 1. \tag{131}$$

Even though the average codeword cost is bounded, it is not guaranteed that the cost of every codeword satisfies the same bound. In order to ensure this, we instead use a modified design which we refer to as a modified Varn code. Given the binary alphabets $\{0,1\}$ and $\mathcal{Y}$, a tree-based variable-length modified Varn code $\phi : \mathcal{X}^K \to \mathcal{Y}^*$ is designed as follows:

- Let $\nu = [(2^K - 1) \mod (|\mathcal{Y}| - 1)]$.
- If $\nu > 0$, let $\delta = |\mathcal{Y}| - 1 - \nu$. Else if $\nu = 0$, let $\delta = 0$. Set $M = 2^K + \delta$.
- Design an exhaustive Varn code with codebook size M.
- Trim down the tree by getting rid of the δ branches with largest cost.

We note that $\delta \leqslant |\mathcal{Y}| - 2$.

The following lemma gives an upper bound on the codeword cost in a modified Varn code.

Lemma 20. *Every codeword of the modified Varn code has cost upper bounded by*

$$W_i \leqslant \log_2 M/\mu + \max_i C_i. \tag{132}$$

□

Proof: Consider the internal node that was the last one to be expanded and suppose it has cost W_0. The cost of any leaf node is larger than W_0, so

$$W_i \geqslant W_0. \tag{133}$$

Since this internal node is the last one to be expanded, its cost is larger than that of any other internal node, implying

$$W_i \leqslant W_0 + \max C_i. \tag{134}$$

Since the tree is full, it is easy to check that

$$\sum_i 2^{-\mu W_i} = 1. \tag{135}$$

Combining equations (133) and (135), we have

$$1 = \sum_i 2^{-\mu W_i} \leqslant M 2^{-\mu W_0}. \tag{136}$$

This implies

$$W_0 \leqslant \log_2 M/\mu \tag{137}$$

and

$$W_i \leqslant W_0 + \max C_i \leqslant \log_2 M/\mu + \max C_i. \tag{138}$$

■

Now we are ready to prove the separation theorem.

Proof: Let's first define a constant

$$D = \left\lceil \frac{\log_2 |\mathcal{X}| + 2}{H(\mathbf{X})} \right\rceil. \tag{139}$$

For any given $\gamma > 0$, define

$$\epsilon = \min\{\frac{\mu\gamma}{2H(\mathbf{X})D}, \frac{1}{D}, \frac{\mu\gamma}{16}\} \tag{140}$$

and $\epsilon' = \frac{\mu\gamma}{16}$. Consider the typical set $A_\epsilon^{(q)}$ with respect to an i.i.d. source with entropy $H(\mathbf{X})$. There exists positive integer Q_1 such that when $q > Q_1$, $\Pr\{A_\epsilon^{(q)}\} > 1-\epsilon$. There are $\leqslant 2^{q(H(\mathbf{X})+\epsilon)}$ length-q sequences x^q in $A_\epsilon^{(q)}$, so we can use no more than $\lceil q(H+\epsilon)+1 \rceil$ bits to index them. We prefix all these sequences by a 0, giving a total length of $\lceil q(H(\mathbf{X})+\epsilon)+2 \rceil$ to represent each sequence in $A_\epsilon^{(q)}$. Similarly, we can index each sequence not in $A_\epsilon^{(q)}$ by using $\lceil q \log |\mathcal{X}| \rceil$ bits. Prefixing these indices by 1, we have a prefix-free code ψ_1 for all sequences in $\mathcal{X}^q$.

Now we construct a length-$\lceil q(H(\mathbf{X})+\epsilon)+2 \rceil$ modified Varn code $\psi_2 : \mathcal{X}^{\lceil q(H(\mathbf{X})+\epsilon)+2 \rceil} \to \mathcal{Y}^*$. We use this code to encode the codeword sequence generated by ψ_1. For every codeword in $A_\epsilon^{(q)}$, the cost is upper bounded by $\log_2 M/\mu + \max_i C_i$. For every codeword in the complement of $A_\epsilon^{(q)}$, which we denote by $B_\epsilon^{(q)}$, $\left\lceil \frac{\lceil q \log |\mathcal{X}|+1 \rceil}{\lceil q(H(\mathbf{X})+\epsilon)+2 \rceil} \right\rceil$ codewords in ψ_2 are needed. The total cost is upper bounded by $(\left\lceil \frac{\lceil q \log |\mathcal{X}|+1 \rceil}{\lceil q(H(\mathbf{X})+\epsilon)+2 \rceil} \right\rceil)(\log_2 M/\mu + \max_i C_i)$, where

$$\left\lceil \frac{\lceil q \log |\mathcal{X}| + 1 \rceil}{\lceil q(H(\mathbf{X})+\epsilon)+2 \rceil} \right\rceil \leqslant \left\lceil \frac{q \log |\mathcal{X}| + 2q}{qH(\mathbf{X})} \right\rceil = D. \tag{141}$$

Consider the concatenation of ψ_1 and ψ_2, denoted $\Psi = \Psi_2 \circ \Psi_1$, with $\Psi(\mathcal{X}^q) = \psi_2(\psi_1(\mathcal{X}^q))$. The total cost is

$$\begin{aligned}
T(\Psi) &= \frac{1}{q}(\sum_{x^q} P(x^q) W(x^q)) \\
&= \frac{1}{q}(\sum_{x^q \in A_\epsilon^{(q)}} P(x^q) W(x^q) + \sum_{x^q \in A_\epsilon^{(q)C}} P(x^q) W(x^q)) \\
&\leqslant \frac{1}{q}[\sum_{x^q \in A_\epsilon^{(q)}} P(x^q)(\log_2 M/\mu + \max_i C_i) \\
&+ \sum_{x^q \in A_\epsilon^{(q)C}} P(x^q) D(\log_2 M/\mu + \max_i C_i)] \\
&= \frac{1}{q}[\Pr\{A_\epsilon^{(q)}\}(\log_2 M/\mu + \max_i C_i) \\
&+ \Pr\{A_\epsilon^{(q)C}\} D(\log_2 M/\mu + \max_i C_i)] \\
&< \frac{1}{q}[(\log_2 M/\mu + \max_i C_i) \\
&+ \epsilon D(\log_2 M/\mu + \max_i C_i)] \\
&= (\frac{\log_2 M}{q\mu} + \frac{\max_i C_i}{q})(1+\epsilon D)
\end{aligned} \tag{142}$$

where

$$M = 2^{\lceil q(H(\mathbf{X})+\epsilon)+2 \rceil} + \delta. \tag{143}$$

Since $\delta \leqslant |\mathcal{Y}| - 2$, we can bound M by

$$M \leqslant 2^{q(H(\mathbf{X})+\epsilon)+3} + |\mathcal{Y}| \tag{144}$$

and by L'Höpital's rule, we have

$$\lim_{q\to\infty} \frac{\log_2(2^{q(H(\mathbf{X})+\epsilon)+3} + |\mathcal{Y}|)}{q} = H(\mathbf{X}) + \epsilon. \tag{145}$$

Therefore, there exists positive integer Q_2 such that, when $q > Q_2$,

$$\frac{\log_2 M}{q} < H(\mathbf{X}) + \epsilon + \epsilon'. \tag{146}$$

Thus, when $q > \max\{Q_1, Q_2\}$, the total cost of Ψ is upper bounded by

$$T(\Psi) < (\frac{H(X)+\epsilon+\epsilon'}{\mu} + \frac{\max_i C_i}{q})(1+\epsilon D). \tag{147}$$

Choose $Q_3 = \left\lceil \frac{8 \max_i C_i}{\gamma} \right\rceil$. When $q > \max\{Q_1, Q_2, Q_3\}$, we have

$$\begin{aligned}
T(\Psi) &< (\frac{H(\mathbf{X})+\epsilon+\epsilon'}{\mu} + \frac{\max_i C_i}{q})(1+\epsilon D) \\
&= \frac{H(\mathbf{X})}{\mu} + (\frac{\epsilon+\epsilon'}{\mu} + \frac{\max_i C_i}{q})(1+\epsilon D) + \epsilon \frac{H(X)}{\mu} D \\
&\leqslant \frac{H(\mathbf{X})}{\mu} + (\frac{\gamma}{8} + \frac{\gamma}{8})(1+1) + \frac{\gamma}{2} = \frac{H(\mathbf{X})}{\mu} + \gamma.
\end{aligned} \tag{148}$$

As shown in [11, Theorem 3.2.1], for any $\gamma' > 0$, there exists a $Q_4 > 0$ such that when $q > Q_4$, the average codeword length per input symbol of Ψ_1 satisfies

$$\frac{1}{q}E[L(\Psi_1)] \leqslant H(\mathbf{X}) + \gamma'. \tag{149}$$

The total cost of Ψ_2 for binary uniform i.i.d. source is upper bounded by

$$T'(\Psi_2) \leqslant \frac{\log_2 M}{\mu\lceil q(H(\mathbf{X})+\epsilon)+2\rceil} + \frac{\max_i C_i}{\mu\lceil q(H(\mathbf{X})+\epsilon)+2\rceil}. \tag{150}$$

By an argument similar to that used to derive the upper bound on $T(\Psi)$ in (148), we conclude that for any $\gamma'' > 0$, there exists a $Q_5 > 0$ such that when $q > Q_5$,

$$T'(\Psi_2) \leqslant \frac{1}{\mu} + \gamma''. \tag{151}$$

To summarize, given any $\gamma, \gamma', \gamma'' > 0$, for sufficiently large q, namely $q > \max\{Q_1, Q_2, Q_3, Q_4, Q_5\}$, we can find a data compression encoder Ψ_1 such that

$$\frac{1}{q}E[L(\Psi_1)] \leqslant H(\mathbf{X}) + \gamma' \tag{152}$$

and a code Ψ_2 for binary uniform i.i.d. source such that

$$T'(\Psi_2) \leqslant \frac{1}{\mu} + \gamma''. \tag{153}$$

The concatenation of Ψ_1 and Ψ_2 will generate a code $\Psi = \Psi_2 \circ \Psi_1$ that has total cost upper bounded by

$$T(\Psi) \leqslant \frac{H(\mathbf{X})}{\mu} + \gamma. \tag{154}$$

This finishes the proof of the separation theorem. ■

Remark 21. Besides modified Varn coding, any fixed-to-variable (or fixed-to-fixed) coding scheme for a uniform source, such as the constant composition distribution matching codes introduced in [51], can be used to prove the separation theorem, as long as the cost of the codeword sequence is bounded from above and the ratio of the upper bound to input sequence length is asymptotically optimal. □

Appendix D
Proof of Theorem 10

We know from Theorem 8 that the problem of designing an optimal type-I shaping code for channel with cost $\mathcal{C}$ and expansion f_0 is related to designing an optimal type-II shaping code for channel with cost $\mathcal{C}'$. The minimum total cost for channel with cost $\mathcal{C}'$ is $H(\mathbf{X})$ and the expansion factor for an optimal type-II shaping code for channel with cost $\mathcal{C}'$ is $f_0' = f_0$.

We first fix some notation. Given a code ϕ, denote by $T_{\mathcal{C}'}(\phi)$ the total cost of this code for channel with cost $\mathcal{C}'$. Denote by $f(\phi)$ its expansion factor and denote by $A_{\mathcal{C}}(\phi)$ and $A_{\mathcal{C}'}(\phi)$ the average cost of this code for channel with cost $\mathcal{C}$ and channel with cost $\mathcal{C}'$, respectively. The asymptotic

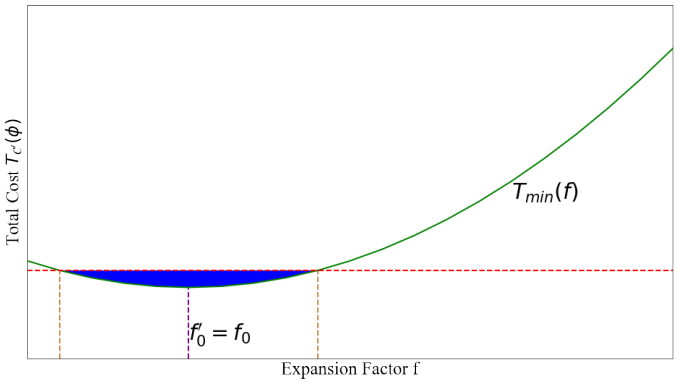


Fig. 8: Achievable total cost vs. expansion factor

symbol occurrence probability of ϕ is $\{P_i\}$ and $\hat{P}_i = \frac{2^{-\mu C_i}}{N}$ is the symbol occurrence probability of the optimal type-I shaping code.

Before we prove the separation theorem for type-I shaping code, we analyze the behavior of $f(\phi)$, $A_{\mathcal{C}}(\phi)$ and $A_{\mathcal{C}'}(\phi)$ when $T_{\mathcal{C}'}(\phi)$ approaches $H(X)$. The proof of Theorem 7 determines the achievable total cost on channel with cost $\mathcal{C}'$. By combining equations (124) and (128), we have

$$\frac{\mathrm{d}T_{\min}}{\mathrm{d}f} = -\frac{1}{\mu}\log_2 N. \tag{155}$$

This implies that function $T_{\min}(f)$ is continuous, strictly monotone decreasing on $[\frac{H(\mathbf{X})}{\log_2|\mathcal{Y}|}, f_0)$ and strictly monotone increasing on (f_0, ∞). This is shown schematically by the green curve $T_{\min}(f)$ in Fig. 8. Thus for any $\zeta > 0$, there exists a γ such that if there exists a code such that $T_{\mathcal{C}'}(\phi) < H(\mathbf{X}) + \gamma$, as indicated by the blue area in Fig. 8, then $|f(\phi) - f_0| < \zeta$. Such a code can always be found for q sufficiently large, for example, by using the generalized Shannon-Fano construction (see [13]).

Now for any $\eta > 0$, $\zeta > 0$, first choose $\zeta' = \min\{\frac{f_0}{2}, \zeta, \frac{\eta\mu f_0^2}{4H(X)}\}$. Choose γ' such that a type-II shaping code with total cost upper bounded by $H(\mathbf{X}) + \gamma'$ has expansion factor $|f - f_0| < \zeta'$. Choose $\gamma = \min\{\gamma', \frac{\mu\eta f_0}{4}\}$, then $A_{\mathcal{C}'}(\phi)$ is bounded by

$$\begin{aligned} A_{\mathcal{C}'}(\phi) &= \sum_i P_i C_i' = \frac{T_{\mathcal{C}'}(\phi)}{f} < \frac{H(\mathbf{X})+\gamma}{f_0-\zeta'} \\ &= \frac{H(\mathbf{X})}{f_0} + \frac{H(\mathbf{X})}{f_0}\frac{\zeta'}{f_0-\zeta'} + \frac{\gamma}{f_0-\zeta'} \\ &\leqslant \frac{H(\mathbf{X})}{f_0} + \frac{H(\mathbf{X})}{f_0}\frac{2\zeta'}{f_0} + \frac{2\gamma}{f_0} \\ &\leqslant \frac{H(\mathbf{X})}{f_0} + \frac{\mu\eta}{2} + \frac{\mu\eta}{2} = \frac{H(\mathbf{X})}{f_0} + \mu\eta. \end{aligned} \tag{156}$$

Since the channel cost C_i' is calculated by

$$C_i' = -\log_2 \hat{P}_i = -\log_2 \frac{2^{-\mu C_i}}{N} = \mu C_i + \log_2 N. \tag{157}$$

$A_{\mathcal{C}}(\phi)$ is upper bounded by

$$\begin{aligned} A_{\mathcal{C}}(\phi) &= \sum_i P_i C_i = (\sum_i P_i C_i' - \log_2 N)/\mu \\ &< (\frac{H(\mathbf{X})}{f_0 \mu} - \frac{\log_2 N}{\mu}) + \eta. \end{aligned} \tag{158}$$

To summarize, for any $\eta > 0$, $\zeta > 0$, there exists a $\gamma > 0$ such that if there exists a code ϕ such that $T_{\mathcal{C}'}(\phi) < H(\mathbf{X}) + \gamma$, this code has expansion factor

$$|f(\phi) - f_0| < \zeta \tag{159}$$

and the average cost $A_{\mathcal{C}}(\phi)$ is upper bounded by

$$A_{\mathcal{C}}(\phi) < (\frac{H(\mathbf{X})}{f_0 \mu} - \frac{\log_2 N}{\mu}) + \eta. \tag{160}$$

Now we present the proof of Theorem 10.

Proof: We consider the channel with cost $\mathcal{C}$ and equivalent channel with cost $\mathcal{C}'$. Given $f, \eta, \zeta, \eta', \zeta'$, and $\gamma' > 0$, for q sufficiently large, there exists a data compression encoder Ψ_1 such that the average codeword length

$$\frac{1}{q} E[L(\Psi_1)] \leqslant H(\mathbf{X}) + \gamma' \tag{161}$$

and a code Ψ_2 for a binary uniform i.i.d. source and costly channel with cost $\mathcal{C}'$ such that

$$T_{\mathcal{C}'}(\Psi_2) \leqslant 1 + \gamma''. \tag{162}$$

The expansion factor of an optimal type-II shaping code for binary uniform i.i.d. source and costly channel $\mathcal{C}'$ is

$$f' = \frac{f}{H(\mathbf{X})}. \tag{163}$$

So when γ'' is small enough, Ψ_2 has expansion factor

$$|f(\Psi_2) - \frac{f}{H(\mathbf{X})}| < \zeta' \tag{164}$$

and $A_{\mathcal{C}}(\Psi_2)$ is upper bounded by

$$\begin{aligned} A_{\mathcal{C}}(\Psi_2) &< (\frac{1}{\frac{f}{H(\mathbf{X})}\mu} - \frac{\log_2 N}{\mu}) + \eta' \\ &= (\frac{H(\mathbf{X})}{f\mu} - \frac{\log_2 N}{\mu}) + \eta'. \end{aligned} \tag{165}$$

The concatenation of Ψ_1 and Ψ_2 will generate a code $\Psi = \Psi_2 \circ \Psi_1$ that has total cost upper bounded by

$$T_{\mathcal{C}'}(\Psi) \leqslant H(\mathbf{X}) + \gamma. \tag{166}$$

When γ is small enough, Ψ has expansion factor

$$|f(\Psi) - f| < \zeta \tag{167}$$

and $A_{\mathcal{C}}(\Psi)$ is upper bounded by

$$A_{\mathcal{C}}(\Psi) < (\frac{H(\mathbf{X})}{f\mu} - \frac{\log_2 N}{\mu}) + \eta. \tag{168}$$

This completes the proof. ■

Yi Liu (S'16) received the B.S. degree in physics from Peking University, Beijing, China, in 2014. He is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, University of California, San Diego, where he is associated with the Center for Memory and Recording Research. His current research interests are coding for costly channel and non-volatile memories.

Pengfei Huang (S'13–M'20) received the B.E. degree in electrical engineering from Zhejiang University, Hangzhou, China, in 2010, the M.S. degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2013, and the Ph.D. degree in electrical engineering from the University of California San Diego, CA, USA, in 2018. From 2014 to 2018, he was associated with the Center for Memory and Recording Research. Since 2018, he has been with Western Digital Corporation, where he is engaged in research and development on enterprise solid-state drives. His current research interests are coding for distributed storage systems and non-volatile memories.

Alexander W. Bergman received the B.S. degree in Electrical Engineering from the University of California, San Diego in 2018. He is currently pursuing the Ph.D. degree in Electrical Engineering at Stanford University. His current research interests include 3D imaging and representation, computer vision, and machine learning.

Paul H. Siegel (M'82–SM'90–F'97–LF'19) received the S.B. and Ph.D. degrees in mathematics from Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 1975 and 1979, respectively. He held a Chaim Weizmann Postdoctoral Fellowship with the Courant Institute, New York University, New York, NY, USA. He was with the IBM Research Division, San Jose, CA, USA, from 1980 to 1995. He joined the faculty of the University of California, San Diego, CA, USA, in July 1995, where he is currently a Distinguished Professor of Electrical and Computer Engineering in the Jacobs School of Engineering. He is affiliated with the Center for Memory and Recording Research where he holds an Endowed Chair and served as Director from 2000 to 2011. His research interests include information theory and communications, particularly coding and modulation techniques, with applications to digital data storage and transmission. He was a Member of the Board of Governors of the IEEE Information Theory Society from 1991 to 1996 and again from 2009 to 2014. He served as Co-Guest Editor of the May 1991 Special Issue on "Coding for Storage Devices" of the IEEE Transactions on Information Theory. He served the same Transactions as Associate Editor for Coding Techniques from 1992 to 1995, and as Editor-in-Chief from July 2001 to July 2004. He was also Co-Guest Editor of the May/September 2001 two-part issue on "The Turbo Principle: From Theory to Practice" and the February 2016 issue on "Recent Advances in Capacity Approaching Codes" of the IEEE Journal on Selected Areas in Communications. He is a member of the National Academy of Engineering. He was the 2015 Padovani Lecturer of the IEEE Information Theory Society. He was the corecipient of the

2007 Best Paper Award in Signal Processing and Coding for Data Storage from the Data Storage Technical Committee of the IEEE Communications Society. He was the corecipient of the 1992 IEEE Information Theory Society Paper Award and the 1993 IEEE Communications Society Leonard G. Abraham Prize Paper Award.