

# Explicit Construction of Optimal Constant-Weight Codes for Identification via Channels

Sergio Verdú, *Fellow, IEEE*, and Victor K. Wei, *Member, IEEE*

**Abstract**—The identification coding theorems of Ahlswede and Dueck have shown that for any nonzero probabilities of missed and false identification, it is possible to transmit  $\exp(\exp(nR))$  messages with  $n$  uses of a noisy channel, where  $R$  is as close as desired to the Shannon capacity of the channel. That capability is achieved by the identification codes explicitly constructed with a three-layer concatenated constant-weight code used in conjunction with a channel transmission code of rate  $R$ .

**Index Terms**—Identification via channels, concatenated codes, Reed-Solomon codes, constant-weight codes.

## I. INTRODUCTION

IDENTIFICATION VIA CHANNELS is a new formulation of the problem of reliable information transmission through noisy channels introduced by Ahlswede and Dueck [1].

In Shannon's formulation, the decoder selects one message out of  $M$  possible ones upon reception of a noisy version of the transmitted  $n$ -symbol codeword. A rate  $R$  is achievable if no matter how small the probability of choosing an incorrect message, codes with rate  $(\log M)/n$  approaching  $R$  exist for all sufficiently large  $n$ . The maximum achievable rate is called the channel capacity  $C$ , and channel coding theorems show how to express the capacity of a channel in terms of its probabilistic description.

Unlike Shannon's formulation, in identification via channels the decoder selects a list of messages and the decoding reliability is measured in terms of two kinds of error probabilities: the probability that the transmitted message is not in the list (*missed identification*), and the probability that any message which is not the transmitted one is included in the list (*false identification*). As in the Shannon formulation, large codes with small error probabilities are sought. Note that if instead of constraining the false identification probability we were to constrain the list size, then the model would become the list decoding problem introduced by Elias [5], [6]. In list decoding, if the list size is allowed to grow as  $\exp(nE)$ , then rates as high as  $C + E$  are achievable but not higher [4, p. 196].

The fundamental limits of identification via channels are quite different. For any channel (not necessarily discrete or memoryless) Ahlswede and Dueck [1] showed that identifica-

tion (ID) codes can sustain (asymptotically) the transmission of  $\exp(\exp(nC))$  messages with  $n$  channel symbols while keeping the probabilities of missed and false identification arbitrarily small. They also showed a "soft converse" result stating that it is not possible to improve that performance for discrete memoryless channels when the error probabilities are constrained to vanish exponentially fast. The strong converse (i.e., for fixed error probabilities) was proved by Han and Verdú [9] for discrete memoryless channels, and in [8] for any finite-input channel that satisfies the strong converse to the channel coding theorem. The doubly exponential growth in the code size is made possible by the lack of a direct constraint on the list size, which in fact also turns out to grow doubly exponentially with the blocklength.

A practical application of identification via channels is the problem of *identification plus transmission* [9], where a central station wishes to transmit through a noisy channel a  $B$ -bit message to one of  $N$  terminals. Upon receipt of the noisy version of the transmitted codeword, every terminal decides whether it is the intended recipient of the message and if so it decodes the message. Identification plus transmission (IT) codes allow message transmission at channel capacity as long as  $\log \log N$  grows no faster than  $B$ , which represents a noticeable improvement over the performance limits achievable by encoding the address and the message separately, with an identification and a transmission code respectively.

The achievability proofs of the coding theorems in [1], [9] show existence of ID and IT codes whose sizes grow doubly exponentially with the blocklength and rate (second-order exponent) equal to the channel capacity. As in Shannon coding theorems, those proofs are existential and do not reveal how to construct codes that achieve those limits. In fact, no code has been reported so far to achieve double exponential growth at any positive rate, let alone at rates close to capacity. In this respect, identification via channels has mimicked the Shannon theory, where the coding theorems were discovered before nontrivial error correcting codes had been invented.<sup>1</sup>

In this paper, we answer the call of [1, p. 27] for explicit constructions of ID codes with positive (second-order) rates. We show that ID and IT codes can be constructed by concatenating transmission codes with binary constant-weight codes. The rate of the resulting identification code is upper bounded by the rate of transmission code. This bound holds with equality if the binary constant-weight code is *optimal for identification*. The main result of this paper is an explicit construction of binary constant-weight codes that are optimal for identification. Thus, not only positive

<sup>1</sup> With the possible exception of the Hamming code.

Manuscript received July 25, 1991; revised April 17, 1992. S. Verdú was supported in part by the Office of Naval Research under Grant N00014-90-J-1734, by the National Science Foundation under PYI Grant ECSE-8857689 and a Grant from Bellcore. This work was presented in part at the IEEE International Symposium on Information Theory, Budapest, Hungary, June 24-28, 1991.

S. Verdú is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544.

V. K. Wei is with Bellcore, 445 South Street, Morristown, NJ 07960.

IEEE Log Number 9203436.

identification rates are easily achieved, but we show that achieving rates near the identification capacity can be accomplished by constructing transmission codes with rates near the Shannon capacity—the central problem in coding theory. In particular, optimal constant-weight binary codes achieve the identification capacity of the noiseless channel. Another scheme to achieve the identification capacity of the noiseless channel (based on two successive prime-number encryptions) has been proposed in [2]. The blocklength- $n$  identification codes obtained under such a scheme require knowledge of the first  $2^n$  prime numbers. Therefore, unlike the construction proposed in this paper, the practicality of the scheme of [2] depends on the computational complexity of state-of-the-art prime-number generation.

Background material and definitions on identification via channels is contained in Section II along with the construction of ID and IT codes from binary constant-weight codes and arbitrary transmission codes. This construction works for any channel, not necessarily discrete or memoryless. It essentially decouples the explicit construction of identification codes for arbitrary channels into the construction of identification codes for the noiseless channel and the construction of transmission codes for the noisy channel. The conditions for the optimality of binary constant-weight codes (which achieve the identification capacity of the noiseless channel) are given in Section III. By focusing attention on binary constant-weight codes obtained from the concatenation of nonbinary codes and PPM codes, we give two explicit constructions of optimal constant-weight codes, one based on algebraic geometry codes, and the other one using a three-layer concatenated code, where the inner code is a PPM code and the outer codes are Reed-Solomon codes. The error exponents of those explicitly constructed codes are also investigated in Section III, and their performance is compared to the bounds in [1]. Finally, we show how the Gilbert–Varshamov bound can be used to give an existential proof of the direct identification coding theorem and of the error-exponent coding theorem for the noiseless channel.

## II. IDENTIFICATION AND BINARY CONSTANT-WEIGHT CODES

We consider arbitrary channels (not necessarily discrete or memoryless) with input alphabet  $A$  and output alphabet  $B$ . A channel is described by the sequence of conditional distributions  $\{W^n : A^n \rightarrow B^n\}_{n=1}^\infty$ .

**Definition 1 [1]:** An  $(n, N, \lambda_1, \lambda_2)$  ID code is a collection  $\{(Q_a, D_a), a = 1, \dots, N\}$  where  $Q_a$  is a probability distribution on  $A^n$  and  $D_a \subset B^n$ , such that

- 1)  $\int W^n(D_a|x)dQ_a(x) \geq 1 - \lambda_1, \quad a = 1, \dots, N$
- 2)  $\int W^n(D_b|x)dQ_a(x) \leq \lambda_2, \quad \text{for all } a \neq b$

The rate of an  $(n, N, \lambda_1, \lambda_2)$  ID code is defined as  $1/n \log \log N$ .

**Definition 2:**  $R$  is an achievable ID rate if for every  $\gamma > 0, 0 < \lambda_1 < 1, 0 < \lambda_2 < 1$  and for all sufficiently large  $n$ , there exist  $(n, N, \lambda_1, \lambda_2)$  ID codes with rate

$$\frac{1}{n} \log \log N > R - \gamma.$$

The ID capacity of the channel is the maximum achievable ID rate. The direct identification coding theorem of [1] is now stated.

**Theorem 1 [1]:** The ID capacity of any channel is greater than or equal its (Shannon) capacity.

As we discussed in the introduction, increasingly more general versions of the converse to Theorem 1 were proved in [1], [9], and [8], respectively.

A related class of codes arises in *identification plus transmission*.

**Definition 3 [9]:** An  $(n, N, M, \lambda_1, \lambda_2)$  IT code is a mapping

$$f : \{1, \dots, N\} \times \{1, \dots, M\} \rightarrow A^n$$

and a collection of subsets

$$\{D_{a,m} \subset B^n, a \in \{1, \dots, N\}, m \in \{1, \dots, M\}\}$$

such that for all  $a = 1, \dots, N$ ,

- 1)  $D_{a,m} \cap D_{a,l} = \emptyset$  if  $l \neq m$ ,
- 2)  $\frac{1}{M} \sum_{m=1}^M W^n(D_{a,m}|f(a, m)) \geq 1 - \lambda_1$ ,
- 3)  $\frac{1}{M} \sum_{m=1}^M W^n(D_b|f(a, m)) \leq \lambda_2$ , if  $b \neq a$ , where  $D_a \triangleq \bigcup_{m=1}^M D_{a,m}$ .

The rate-pair of an  $(n, N, \lambda_1, \lambda_2)$  IT code is  $(1/n \log M, 1/n \log \log N)$ .

**Definition 4 [9]:**  $(R, R')$  is an achievable IT rate-pair if for every  $\gamma > 0, 0 < \lambda_1 < 1, 0 < \lambda_2 < 1$ , and for all sufficiently large  $n$ , there exist  $(n, N, M, \lambda_1, \lambda_2)$  IT codes such that

$$\begin{aligned} \frac{1}{n} \log M &> R - \gamma, \\ \frac{1}{n} \log \log N &\geq R' - \gamma. \end{aligned}$$

The achievability part of the identification plus transmission coding theorem can be shown similarly to Theorem 1.

**Theorem 2 [9]:** For an arbitrary channel  $(C, C)$  is an achievable IT rate-pair.

The converse of this result [9] states that every achievable IT rate-pair  $(R, R')$  must satisfy  $\max\{R, R'\} \leq C$ . (Thus the set of achievable IT rate-pairs is the square  $(0, C)^2$ . This converse is known to hold for any finite-input channel and its proof is elementary, in contrast to that of the converse identification theorem.

It follows from Definitions 1 and 3 that, given any  $(n, N, M, \lambda_1, \lambda_2)$  IT code, we can immediately construct an  $(n, N, \lambda_1, \lambda_2)$  ID code by letting each distribution  $Q_a$  on  $A^n$  assign mass  $1/M$  to  $f(a, m)$ ,  $m = 1, \dots, M$ . Consequently, we henceforth focus attention on constructing optimal IT codes (whose rate-pairs approach the optimum point  $(C, C)$ ). Automatically, this will give optimal ID codes.

The class of IT codes we focus attention on in this paper is the result of linking a transmission (error-correcting) code with a binary constant-weight code. To describe those codes we use the following standard notation.

**Definition 5:** An  $(S, N, M, K)$  binary constant-weight code is a set of  $N$  binary  $S$ -tuples of Hamming weight  $M$  such that

the pairwise overlap (maximum number of coincident 1's for any pair of codewords) does not exceed  $K$ .

Any  $(S, N, M, K)$  binary constant-weight code can be described by an  $N \times M$  incidence matrix on  $\{1, \dots, S\}$  such that for every  $a \in \{1, \dots, N\}$ , the row  $(s(a, 1), \dots, s(a, M))$  gives the locations of the  $M$  1's in the  $a$ th codeword.

Of particular interest in the sequel will be three measures associated with binary constant-weight codes: the *weight factor*

$$\beta \triangleq \frac{\log M}{\log S}, \quad (1)$$

the *second-order rate* (as opposed to the (first-order) rate  $1/S \log N$ )

$$\rho \triangleq \frac{\log \log N}{\log S}, \quad (2)$$

and the *overlap fraction*

$$\mu \triangleq \frac{K}{M}.$$

Note that any  $(S, N, M, K)$  binary constant-weight code satisfies

$$\beta \leq 1 \quad (3)$$

and

$$\rho \leq 1, \quad (4)$$

where (4) follows from  $N \leq 2^S$  if we assume that the log basis is not smaller than 2.

**Definition 6:** An  $(n, S, \lambda)$  transmission code for channel  $\{W^n : A^n \rightarrow B^n\}$  is a collection  $\{(\phi(s), E_s) \in A^n \times 2^{B^n}, s = 1, \dots, S\}$  such that the subsets  $E_i$  are nonoverlapping and

$$W^n(E_i | \phi(i)) \geq 1 - \lambda.$$

The rate of an  $(n, S, \lambda)$  transmission code is  $1/n \log S$ . Nesting any  $(S, N, M, K)$  binary constant-weight code with any  $(n, S, \lambda)$  transmission code we construct an IT code as the next result shows.

**Proposition 1:** Given an  $(S, n, M, \mu M)$  binary constant-weight code  $\{s(a, m), a = 1, \dots, N; m = 1, \dots, M\}$  ( $0 \leq \mu \leq 1$ ) and an  $(n, S, \lambda)$  transmission code  $\{(\phi(s), E_s), s = 1, \dots, S\}$ , then the following is an  $(n, N, M, \lambda, \lambda + \mu)$  IT code:

$$\begin{aligned} f(a, m) &= \phi(s(a, m)) \quad a = 1, \dots, N, \quad m = 1, \dots, M, \\ D_{a,m} &= E_{s(a,m)}, \quad a = 1, \dots, N, \quad m = 1, \dots, M. \end{aligned}$$

*Proof:* Since  $\{E_s\}$  are disjoint, Condition 1) in Definition 3 is satisfied. Moreover, the error probabilities are easy to bound (cf. [1], [9]):

$$\begin{aligned} & \frac{1}{M} \sum_{m=1}^M W^n(D_{a,m} | f(a, m)) \\ &= \frac{1}{M} \sum_{m=1}^M W^n(E_{s(a,m)} | \phi(s(a, m))) \\ &\geq 1 - \lambda, \end{aligned}$$

and for all  $a \neq b$ , let  $A_a = \cup_{m=1}^M \{s(a, m)\}$  and consider

$$\begin{aligned} & \frac{1}{M} \sum_{m=1}^M W^n \left( \bigcup_{j=1}^M D_{b,j} | f(a, m) \right) \\ &= \frac{1}{M} \sum_{m: s(a,m) \in A_a \cap A_b} W^n \left( \bigcup_{j=1}^M D_{b,j} | f(a, m) \right) \\ & \quad + \frac{1}{M} \sum_{m: s(a,m) \in A_a - A_b} W^n \left( \bigcup_{j=1}^M D_{b,j} | f(a, m) \right) \\ &\leq \frac{1}{M} |A_a \cap A_b| + \frac{1}{M} \\ & \quad \cdot m : \sum_{s(a,m) \in A_a - A_b} W^n(E_{s(a,m)}^c | f(a, m)) \\ &\leq \mu + \lambda. \quad \square \end{aligned}$$

The rate-pair (cf. Definition 3) of the IT code constructed in Proposition 1 from a transmission code and a constant-weight code is equal to  $(\beta R, \rho R)$  where  $R$  is the rate of the transmission code and  $\beta$  and  $\rho$  are the weight factor and second-order rate of the constant-weight code. The converse to the Shannon coding theorem and (3) and (4) dictate that, in order to approach the performance promised by Theorem 2 with this class of IT codes, we need both transmission codes whose rates approach the channel capacity and binary constant-weight codes whose weight factor and second-order rate approach unity. This motivates the following definition.

### III. OPTIMAL BINARY CONSTANT-WEIGHT CODES

**Definition 7:** Consider a sequence  $\{C_i\}$ , where  $C_i$  is a  $(S_i, N_i, M_i, \mu_i M_i)$  binary constant-weight code with weight factor  $\beta_i$ , second-order rate  $\rho_i$  and pairwise overlap fraction  $\mu_i$ . We say that the sequence of codes  $\{C_i\}$  is *optimal for identification* if

$$\beta_i \rightarrow 1, \quad (5)$$

$$\rho_i \rightarrow 1, \quad (6)$$

$$\mu_i \rightarrow 0. \quad (7)$$

There is an important distinction between the problem of constructing codes that are optimal for identification and the code construction problem usually considered in the literature. In the former we aim to achieve high second-order rate while in the latter we aim to achieve high first-order rate. As we shall see later, our most explicit scheme for constructing optimal (for identification) codes achieves the maximum second-order rate (i.e., unity) while its asymptotic first-order rate is zero. Henceforth, *optimal* codes are optimal codes for identification, unless otherwise stated.

A nonconstructive proof of the existence of optimal (for identification) binary constant-weight codes can be obtained

using the Gilbert–Varshamov bound (see Appendix). However, the main result of this paper is an explicit construction of optimal binary constant-weight codes. To describe such a construction we need a few standard concepts in coding theory.

*Definition 8:* Let  $C_1$  and  $C_2$  be codes with blocklength  $n_i$ , size  $N_i$  and alphabet  $A_i$ ,  $i = 1, 2$ . If  $|A_2| = N_1$ , then the *concatenated* (or *nested*) code  $C = C_1 \circ C_2$  with blocklength  $n_1 n_2$ , size  $N_2$  and alphabet  $A_1$  is constructed using any one-to-one function  $h: A_2 \rightarrow C_1$ :

$$C = \{(h(y_1), \dots, h(y_{n_2})) : (y_1, \dots, y_{n_2}) \in C_2\}.$$

The code  $C_1$  is called the inner code and  $C_2$  is called the outer code. Naturally, concatenation can be nested, for example

$$C = (C_1 \circ C_2) \circ C_3 = C_1 \circ (C_2 \circ C_3) = C_1 \circ C_2 \circ C_3$$

An easy way to construct binary constant-weight codes is to concatenate a code whose alphabet is not binary with a *pulse position modulation* (PPM) code.

*Definition 9:* A  $(q, q, 1, 0)$  binary constant-weight code (which consists of all binary  $q$ -vectors of unit weight) is called a  $[q]$  PPM code.

Note that if we concatenate a  $[q]$  PPM code with an outer code with alphabet size  $q$ , blocklength  $n'$ , size  $N$  and minimum distance  $d$ , we obtain an  $(n'q, N, n', n' - d)$  binary constant-weight code. Conditions (5)–(7) will be satisfied, if and only if the outer code satisfies

$$\frac{\log q}{\log n'} \rightarrow 0, \quad (8)$$

$$\frac{\log \log N}{\log n'} \rightarrow 1, \quad (9)$$

and

$$\frac{d}{n'} \rightarrow 1. \quad (10)$$

Note that it is sufficient but not necessary for (9) to hold that the rate of the outer code be nonzero. Condition (10) can be viewed as the asymptotic orthogonality of the codebook.

Constant-weight codes that approach the Johnson upper bound have been exhibited in [13] by concatenating a PPM code with a Reed–Solomon code.

*Definition 10:*<sup>2</sup> Let  $q$  be a prime power and denote the elements of  $\text{GF}(q)$  by  $\{a_1, \dots, a_q\}$ . A  $[q, k]$  Reed–Solomon code ( $k < q$ ) is the set of  $q$ -vectors over  $\text{GF}(q)$ :

$$\{(p(a_1), \dots, p(a_q)) : p(x) \text{ is a polynomial of degree } < k \text{ with coefficients from } \text{GF}(q)\}.$$

Thus, the blocklength is  $q$ , the size is  $q^k$ , and the minimum distance is  $q - k + 1$  because if two polynomials of degree  $< k$  coincide at  $k$  or more places, then they are identical.

<sup>2</sup>Other equivalent definitions of Reed–Solomon codes (sometimes with minor variations on blocklength) can be found in standard coding theory textbooks. Besides error control this technique of redundancy coding via polynomial interpolation has found applications in several other areas. For example, in secrecy sharing [11], in transmission and storage in data networks [12], and in analog diversity coding [3].

We see immediately that condition (8) cannot be satisfied by a Reed–Solomon code. Therefore, no sequence of binary constant-weight codes each of which is constructed by concatenating a PPM code and a Reed–Solomon code can be optimal for identification. This is a result of the relatively short blocklength of Reed–Solomon codes compared to other block codes with the same alphabet, such as two concatenated Reed–Solomon codes, which, as we will see turn out to lead to optimal constant-weight binary codes. Before giving such an explicit construction with Reed–Solomon codes, let us see how it is possible to construct optimal codes for identification directly from a relatively new class of algebraic geometry codes.

Algebraic geometry is a branch of mathematics which provides powerful tools for solving various problems. Recently, asymptotically good error-correcting codes have been constructed from algebraic geometric methods which can even excel over the Gilbert–Varshamov bound. Constant-weight codes obtained by concatenating algebraic geometry codes and PPM codes have been studied in [7]. In particular, the literature contains a class of algebraic geometry codes that satisfies (8)–(10). Thus, they can be used to construct optimal codes for identification.

The algebraic geometry codes arising from modular curves in [10] have the parameters in the following result which is rephrased for our purposes.

*Theorem 3 [10]:* Let  $q = p^{2m}$  where  $p$  is a prime,  $R = 1 - (q^{1/2} - 1)^{-1} - \delta > 0$ . Then, for any  $\varepsilon > 0$ , there exists a sequence of  $q$ -ary codes of increasing lengths whose asymptotic rate is greater than or equal to  $R - \varepsilon$  and whose asymptotic ratio of minimum (Hamming) distance to length is greater than or equal to  $\delta - \varepsilon$ .

Therefore, given fixed  $q = p^{2m}$  and fixed  $R > 0$ , there exists a sequence of  $q$ -ary codes of blocklength  $n'$ , size  $N$ , and minimum distance  $d$  satisfying

$$\frac{\log q}{\log n'} \rightarrow 0, \\ \frac{\log \log N}{\log n'} = \frac{\log n' + \log R + \log \log q}{\log n'} \rightarrow 1,$$

and

$$\frac{d}{n'} \rightarrow 1 - \frac{1}{q^{1/2} - 1}.$$

Next, construct a sequence  $(q_i, R_i)$  with  $q_i \rightarrow \infty$  and  $R_i \rightarrow 0$ . Then, choosing a sufficiently long code for each  $(q_i, R_i)$ , we obtain a sequence of codes satisfying (8)–(10). For other aspects related to the capability of this scheme see [7].

The sequence of codes guaranteed in Theorem 3 can be explicitly constructed. In fact, the computational complexity of this construction is polynomial in terms of the blocklength. Therefore, the derived construction of optimal codes for identification can also be viewed as explicit, answering the call by Ahlswede and Dueck [1]. However, this result requires a nontrivial background on algebraic geometry and may not be directly accessible to most readers.

In the remainder of the paper, we present a conceptually simpler, more practically attractive and even more explicit construction of optimal codes for identification. This construction is a three-layer concatenation of a layer of PPM code and two layers of Reed–Solomon codes. It can be understood without much background.

**Proposition 2:** The  $[q, k, t]$  three-layer concatenated code  $C_1 \circ C_2 \circ C_3$ , with  $C_1 = [q]$  PPM,  $C_2 = [q, k]$  Reed–Solomon and  $C_3 = [q^k, q^t]$  Reed–Solomon, with  $t < k < q = \text{prime power}$  is a  $(q^{k+2}, q^{kq^t}, q^{k+1}, kq^k + q^{1+t})$  binary constant-weight code.

*Proof:* It follows directly from the foregoing discussion that  $C_1 \circ C_2$  is a  $(q^2, q^k, q, k-1)$  binary constant-weight code, and  $C_1 \circ C_2 \circ C_3$  is a  $(S, N, M, K)$  binary constant-weight code with  $S = q^{k+2}$ ,  $N = (q^k)^{q^t}$  and  $M = q^{k+1}$ . To upper bound  $K$ , fix any pair of codewords in the outer code. No more than  $(q^t - 1)$  positions may be identical. A position in which both codewords coincide contributes an overlap of  $q$  components in  $C_1 \circ C_2 \circ C_3$ , whereas any positions in which both codewords differ contributes an overlap of at most  $k - 1$  components. Therefore,  $K \leq (q^t - 1)q + q^k(k - 1)$ , and the result follows by dropping the negative terms in this expression.  $\square$

It now means to choose  $t, k$ , and  $q$  appropriately so that (5)–(7) are satisfied.

**Proposition 3:** Let  $C_i$  be a  $[q_i, k_i, t_i]$  three-layer concatenated code as in Proposition 2. The sequence of codes  $\{C_i\}$  is optimal for identification if

- a)  $t_i \rightarrow \infty$
- b)  $t_i/k_i \rightarrow 1$ ,
- c)  $k_i/q_i \rightarrow 0$ ,
- d)  $q_i^{t_i - k_i} \rightarrow 0$ .

*Proof:* It follows from (1), (2), and Proposition 2 that

$$\begin{aligned} \beta_i &= \frac{k_i + 1}{k_i + 2}, \\ \rho_i &= \frac{t_i}{k_i + 2} + \frac{\log k_i + \log \log q_i}{(k_i + 2) \log q_i}, \\ \mu_i &\leq \frac{k_i}{q_i} + q_i^{t_i - k_i}, \end{aligned} \quad (14)$$

whereupon the sufficiency of a)–d) for (5)–(7) is clear.

A simple sequence of parameters that satisfies the conditions in Proposition 3 is  $t_i = i$ ,  $k_i = i + 1$ , and  $q_i$  any increasing sequence of prime powers.

Note that the (first-order) rate of a  $[q, k, t]$  three-layer concatenated code is equal to  $q^{t-k}(k/q) (\log q)/q$ , and therefore it goes to zero for any sequence satisfying condition d).

In addition to the study of ID capacity, Ahlswede and Dueck studied the error exponents achievable for identification via channels. A triple  $(R, E_1, E_2)$  is called achievable if for all  $\delta > 0$  and all sufficiently large  $n$  there exist  $(n, N, \lambda_1, \lambda_2)$ –ID codes whose rate exceeds  $R - \delta$  and whose error exponents  $1/n \log(1/\lambda_i)$  exceed  $E_i - \delta$ . Although the region of achievable error exponents is not yet known, Ahlswede and Dueck gave an inner bound and an outer bound to that region.

([1], Theorem 2). The outer bound and some conditions under which it is known to be tight are given by the following theorem.

**Theorem 4:** If  $(R, E_1, E_2)$  is achievable and  $E_1 > 0$ , then

$$R + 2E_2 \leq C.$$

This bound is tight if either a) the channel is noiseless, or b)  $E_1 \rightarrow 0$ .

We will now compare the bound in Theorem 4 with the error exponents achievable with the codes constructed in this paper. In order to carry out this comparison, we will focus exclusive attention on  $N$  and will not be concerned with the growth of  $M$  as we did when studying IT codes.

**Proposition 4:** Let  $E(R)$  be the reliability function of the channel. The following error-exponent region can be achieved with three-layer concatenated codes:

$$\left\{ \left( \rho R, E(R), \min \left\{ E(R), \frac{R}{3}(1 - \rho) \right\}; 0 \leq R \leq C \right. \right. \\ \left. \left. ; 0 \leq \rho \leq 1 \right) \right\}.$$

*Proof:* Fix  $0 \leq \rho \leq 1$  and  $0 \leq R \leq C$ . The  $[q, k, k-1]$  three-layer concatenated code with

$$k = \frac{1 + 2\rho}{1 - \rho}$$

exhibits a second-order rate equal to

$$\rho + O\left(\frac{\log \log q}{\log q}\right),$$

whereas its overlap fraction and blocklength satisfy

$$\begin{aligned} \frac{\log 1/\mu}{\log S} &= \frac{1}{k+2} - O\left(\frac{1}{\log q}\right) \\ &= \frac{1 - \rho}{3} - O\left(\frac{1}{\log q}\right). \end{aligned}$$

Finally, letting  $q$  be arbitrarily large and coupling the three-layer concatenated code with a  $(n, \exp(nR), \lambda)$  transmission code, the error exponents of the resulting ID code are (Proposition 1)  $1/n \log(1/\lambda)$  and  $1/n \log(1/(\lambda + \mu))$ .  $\square$

In the special case of the noiseless channel, Proposition 4 states that the three-layer concatenated code achieves any triple  $(R, E_1, E_2)$  with  $R + 3E_2 = C$ , which falls short of the optimum error-exponent performance predicted by Theorem 4. In fact, even for noiseless channels, the search for ID codes which not only achieve ID capacity but the best possible error exponents remains open. Regarding the explicit construction with the algebraic geometry code of Theorem 3, it can be checked that it cannot achieve any triples  $(R, E_1, E_2)$  with  $E_2 > 0$ . However, it is possible to improve that result in Proposition 4 in the region where  $\rho < 1/4$  using the inner two layers of the concatenated code only.

**Proposition 5:** The following error exponent region is achievable

$$\left\{ \left( \rho R, E(R), \min \left\{ E(R), R \left( \frac{1}{2} - \rho \right) \right\}; 0 \leq R \leq C \right. \right. \\ \left. \left. ; 0 \leq \rho \leq 1/2 \right) \right\}. \quad (20)$$

*Proof:* The inner two layers of our concatenation scheme constitute a  $(q^2, q^k, q, k-1)$  binary constant-weight code. Fix  $0 < \rho < 1/2$  and select  $k$  and  $q$  such that

$$\rho = \frac{\log k}{2 \log q}.$$

Then, the second-order rate of the code and its overlap fraction satisfy, respectively,

$$\frac{\log \log N}{\log S} = \rho + O\left(\frac{\log \log q}{\log q}\right), \\ \frac{\log 1/\mu}{\log S} = \frac{1}{2} - \rho \frac{\log(k-1)}{\log k}.$$

Finally, letting  $q$  (and thus  $k$ ) go to infinity and using a transmission code with rate  $R$  the result follows.  $\square$

#### APPENDIX

The Gilbert–Varshamov bound and algebraic manipulations are used in this appendix to prove a result from which Theorem 1 and the tightness of the bound in Theorem 4 for noiseless channels follow as corollaries.

*Proposition A:* Let  $S$  be a positive integer,  $T$  and  $F$  be positive real numbers satisfying  $T+2F < 1$ . Then there exists an  $(S, N, S^{T+F}, S^T)$  binary constant-weight code with

$$\frac{\log \log N}{\log S} = T + O\left(\frac{\log \log S}{\log S}\right).$$

*Remark:* The code in Proposition A can be used to construct directly an ID code with rate  $T$  and false-identification error exponent  $F$  for the noiseless binary channel for any  $T$  and  $F$  such that  $T+2F < 1$ , which attains the outer bound of Theorem 4. For an arbitrary channel with reliability function  $E(R)$ , the code here can be used in combination with an optimal transmission code with rate  $R$  to construct an ID code achieving the triple  $(TR, E(R), \min\{E(R); FR\})$  for that channel. Thus, this appendix provides an alternative proof to Theorem 1 and to the inner bound in [1, Theorem 2] for noiseless channels. The purpose of our effort here is to shed some light on the intuition behind the problem by focusing on a special case and by deriving a simple proof for it. We do not prove the full strength of the direct part of [1, Theorem 2], which we do not believe can be proved by using the Gilbert–Vasharmov bound alone, or using constant-weight codes for that matter.

*Proof:* The Gilbert–Varshamov bound implies the existence of  $(S, N, M, K)$  binary constant-weight codes satisfying

$$N \geq \frac{\binom{S}{M}}{\sum_{i>K} \binom{M}{i} \binom{S-M}{M-i}}. \quad (A.1)$$

If  $K/M > M/S$ , the denominator can be further bounded in terms of the largest term in the summation as follows

$$\sum_{i>K} \binom{M}{i} \binom{S-M}{M-i} \leq (M-K) \binom{M}{K+1} \binom{S-M}{M-K-1}. \quad (A.2)$$

Henceforth, we will let  $M = S^{T+F}$  and  $K = S^T$ , which together with the assumption  $T+2F < 1$  implies  $K/M > M/S$ . Expanding the binomial coefficients in (A.1) and (A.2) we obtain

$$N \geq \frac{1}{M} \prod_{i=1}^{M-K-1} \frac{(S-K-i)(M-K-i)}{(M-K-i)(S-M+1-i)} \\ \cdot \prod_{i=0}^K \frac{(S-i)(K+1-i)}{(M-i)(M-i)} \\ \geq \frac{1}{M} \prod_{i=0}^K \frac{(S-i)(K+1-i)}{(M-i)(M-i)}. \quad (A.3)$$

Most, but not all, terms in the product are larger than one. We wish to show that the product is exponential in  $S^T$ . To that end, we partition the terms in the product into  $K/k$  groups by assembling every  $(K/k)$ th term together. Let  $\varepsilon' = 1 - T - 2F > 0$  and let  $\varepsilon = (k-1)\varepsilon' + (T+F) > 0$ , where  $k$  is an integer greater than 1 such that the partial product corresponding to the  $i$ th group,  $0 \leq i' < K/k$ , is

$$\prod_{j=1}^k \frac{(S-jK/k+i')(K+1-jK/k+i')}{(M-jK/k+i')(M-jK/k+i')} \\ \geq \frac{S-K}{(M-K)^2} \left[ \frac{(S-K+K/k)(1+K/k)}{(M-K+K/k)(M-K+K/k)} \right]^{k-1} \\ \geq \frac{S}{M^2} \left[ \frac{(S-K+K/k)(1+K/k)}{(M-K+K/k)(M-K+K/k)} \right]^{k-1} \\ \geq \frac{1}{M} \left( \frac{SKk}{kM^2} \right)^{k-1} \\ \geq k^{-(k-1)} S^{(k-1)\varepsilon' - (T+F)} \\ = k^{-(k-1)} S^\varepsilon, \quad (A.4)$$

where the foregoing inequalities follow from  $S \geq M$  and the fact that  $(S-x)(K+1-x)/(M-x)^2$  and  $(S-x)/(M-x)^2$  are nonincreasing and nondecreasing, respectively, in  $x$ ,  $0 \leq x \leq K$ . Then, (A.3) and (A.4) yield

$$N \geq \frac{1}{M} \prod_{i=0}^K \frac{(S-i)(K+1-i)}{(M-i)(M-i)} \geq \frac{1}{M} \left[ k^{-(k-1)} S^\varepsilon \right]^{K/k} \\ = \frac{1}{M} \left[ k^{-(k-1)} S^\varepsilon \right]^{S^T/k},$$

which implies

$$\frac{\log \log N}{\log S} = T + O\left(\frac{\log \log S}{\log S}\right). \quad \square$$

#### ACKNOWLEDGMENT

References [2], [7], and [13] were brought to the authors' attention by anonymous referees.

## REFERENCES

- [1] R. Ahlswede and G. Dueck, "Identification via channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 15–29, Jan. 1989.
- [2] R. Ahlswede and B. Verboven, "On identification via multiway channels with feedback," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1519–1526, Nov. 1991.
- [3] F. Ayanoglu, C. L. I., R. D. Gitlin, and J. E. Mazo, "Diversity coding: Using error control for self-healing communication networks," *Proc. 1990 INFOCOM*, vol. 1, pp. 95–104, San Francisco, CA, June 1990.
- [4] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic 1981.
- [5] P. Elias, "List decoding for noisy channels," *IRE WESCON Conv Record*, vol. 2, pp. 94–104, 1957.
- [6] ———, "Error-correcting codes for list decoding," *IEEE Trans. Inform. Theory*, vol. 37, pp. 5–12, Jan. 1991.
- [7] T. Ericson and V. A. Zinoviev, "An improvement of the Gilbert bound for constant weight codes," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 721–723, Sept. 1987.
- [8] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inform. Theory*, accepted for publication, May 1993.
- [9] ———, "New results in the theory and applications of identification via channels," *IEEE Trans. Inform. Theory*, vol. 38, pp. 14–25, Jan. 1992.
- [10] G. L. Katsman, M. A. Tsfasman, and S. G. Vladut, "Modular curves and codes with a polynomial construction," *IEEE Trans. Inform. Theory*, vol. IT-30, pt. II, pp. 353–355, Mar. 1984.
- [11] D. E. Knuth and A. C. Yao, "The complexity of random number generation," *Proceedings of Symposium New Directions and Recent Results in Algorithms and Complexity*. New York: Academic Press, 1976.
- [12] M. O. Rabin, "Efficient dispersal of information for security, load balancing and fault tolerance," *J. ACM*, vol. 36, pp. 336–348, Apr. 1989.
- [13] V. A. Zinoviev, "Cascade equal-weight codes and maximal packings," *Prob. Cont. Inform. Theory*, vol. 12, pp. 3–10, 1983.