

The Rate-Immediacy Barrier in Explicit Tree Code Constructions

Gil Cohen^{*} Leonard J. Schulman[†] Piyush Srivastava[‡]

Abstract

Since the introduction of tree codes by Schulman (STOC 1993), explicit construction of asymptotically good tree codes has remained a notorious challenge. A work by Cohen, Haeupler and Schulman (STOC 2018), as well as the state-of-the-art construction by Ben Yaacov, Cohen, and Yankovitz (STOC 2022) have achieved codes with rate $\Omega(1/\log \log n)$, exponentially improving upon the original rate $\Omega(1/\log n)$ construction of Evans, Klugerman and Schulman from 1994. All of these constructions rely, at least in part, on increasingly sophisticated methods of combining (block) error-correcting codes.

In this work, we identify a fundamental barrier to constructing tree codes using known techniques. We introduce a key property which we call *immediacy*, that, while not required by the original definition of tree codes, is shared by all known constructions and inherently arises in recursive combinations of error-correcting codes. Our main technical contribution is the proof of a *rate-immediacy trade-off*, which, in particular, implies that any tree code with constant distance and non-trivial immediacy must necessarily have vanishing rate. By applying our rate-immediacy trade-off to existing constructions, we establish that their known rate analyses are essentially optimal given their actual error-correction properties. More broadly, our work highlights the need for fundamentally new ideas—beyond the recursive use of error-correcting codes—to achieve substantial progress in explicitly constructing asymptotically good tree codes.

^{*}Department of Computer Science, Tel Aviv University, Tel Aviv, Israel. Email: gil@tauex.tau.ac.il. Supported by ERC starting grant 949499.

[†]E&AS Division, California Institute of Technology, Pasadena, USA. Email: schulman@caltech.edu. Supported in part by NSF Award 2321079.

[‡]Tata Institute of Fundamental Research, Mumbai, India. Email: piyush.srivastava@tifr.res.in. Supported in part by the Department of Atomic Energy, Government of India, under project nos. RTI4001 and RTI4014, by SERB MATRICS grant number MTR/2023/001547, and by the Infosys-Chandrasekharan virtual center for Random Geometry.

1 Introduction

Coding theory addresses the problem of communication over an imperfect channel. In the classic setting [Sha48, Ham50], Alice aims to communicate a message to Bob via a channel that may introduce errors. The central question is: how should Alice encode her message so that Bob can accurately recover it, provided that the number of errors is limited? This scenario motivates the notion of an error-correcting code. Formally, a function $C: \Sigma^k \rightarrow \Sigma^n$ is a *block error-correcting code* with distance δ and block length k if, for every pair of distinct strings $x, y \in \Sigma^k$, their respective encodings $C(x)$ and $C(y)$ differ in at least a δ fraction of positions, with respect to the Hamming distance. The *rate* of information transmission $\rho = \frac{k}{n}$ and the fraction of errors corrected, $\frac{\delta}{2}$, are competing parameters. A fundamental problem in coding theory is to construct explicit *asymptotically good codes*, i.e., codes that maintain constant distance $\delta > 0$ and constant rate $\rho > 0$. Here, by “explicit” we mean that the encoding function C can be computed in polynomial time. Justesen [Jus72] was the first to provide such an explicit construction. Since then, many explicit constructions have been developed (see, e.g., [TVZ82, SS96]).

While error-correcting codes solve the problem of sending a single message from Alice to Bob, there are scenarios involving dynamic interaction, where messages exchanged depend on previously communicated information. Interactive coding addresses this more intricate problem of enabling reliable interactive communication over imperfect channels. Standard error-correcting codes are insufficient for this task. Readers interested in this rapidly growing research field are encouraged to consult the comprehensive survey by Gelles [Gel17].

Tree codes are crucial combinatorial structures introduced in [Sch93, Sch96] to facilitate deterministic interactive coding. Analogous to error-correcting codes in single-message scenarios, tree codes are trees equipped with a specific distance property. To define this formally, we first introduce some notation. Let T be an infinite complete rooted binary tree, with edges labeled from an alphabet Σ . For two vertices u, v of equal depth, let their least common ancestor be denoted w . Let ℓ be the number of edges on the path from u (or v) to w . Define $p_u, p_v \in \Sigma^\ell$ as the sequences of symbols labeling the edges on the paths from w to u and from w to v , respectively. The quantity $h(u, v)$ denotes the relative Hamming distance between p_u and p_v . Intuitively, $h(u, v)$ measures how distinct the sequences labeling the paths to u and v are, disregarding their common prefix. A tree code enforces a minimum bound on this quantity:

Definition 1 (Tree codes [Sch93]). Let $\delta \in [0, 1]$ and let T be an infinite complete rooted binary tree. A labeling of the edges of T from an alphabet Σ is called a *tree code with distance δ* if, for every pair of vertices u, v at the same depth, it holds that $h(u, v) \geq \delta$. The *rate* of the tree code, denoted by ρ , is defined as $\frac{1}{\log_2 |\Sigma|}$.

An alternative definition found in the literature describes a tree code as a family $(T_n)_{n \in \mathbb{N}}$, where

each T_n is a rooted binary tree of *finite* depth n . Such a family is called a tree code with distance δ if each T_n has a distance of at least δ , as previously defined. Clearly, an infinite tree code naturally induces such a finite family when truncated at any given depth. Conversely, it has been shown that the opposite direction also holds [BYCY22], with only a constant degradation in parameters. Hence, we use these two definitions interchangeably throughout this informal discussion.

Initially, it was not clear whether an asymptotically good tree code—one with both positive rate and positive distance—existed. Schulman provided three distinct proofs showing that, for any constant $\delta < 1$, a tree code exists with alphabet size $|\Sigma| = O_\delta(1)$ that achieves distance δ . More recently, by adapting one of these approaches, it was proved that there is a tree code with just $|\Sigma| = 4$ symbols, namely a rate- $\frac{1}{2}$ tree code, and positive distance (specifically, $\delta > 0.136$) [CS20]. It was also observed in [CS20] that 3 symbols are insufficient for guaranteeing positive distance. However, all known existence proofs of asymptotically good tree codes are non-explicit, relying on probabilistic methods in non-trivial ways. Despite significant interest, the explicit construction of asymptotically good tree codes remains a notorious open problem.

Given these difficulties, researchers have naturally considered constructing tree codes allowing for vanishing rate, where the objective is nonetheless to minimize the rate deterioration as a function of the depth n .¹ A trivial construction which encodes the entire path from the root on each edge, achieves $\delta = 1$ but rate $\frac{1}{n}$. In an unpublished manuscript, Evans, Klugerman, and Schulman [EKS94] provided a construction with rate $\Omega(1/\log n)$. An improvement was made only fairly recently by Cohen, Haeupler and Schulman [CHS18], who constructed tree codes with rate $\Omega(1/\log \log n)$. A decoding algorithm to the latter tree code construction was devised by Narayanan and Weidner [NW20], who also suggested alternative constructions. Connections between [CHS18] and the work of Pudlák [Pud16] (see also [Yan25] for a related construction) were explored by Bhandari and Harsha [BH20]. Additionally, two distinct explicit constructions with constant rate were proposed by Moore and Schulman [MS14] and by Ben Yaacov, Cohen, and Narayanan [BYCN21], but their correctness hinges upon plausible but unproven conjectures.

The state-of-the-art construction by Ben Yaacov, Cohen, and Yankovitz [BYCY22] also achieves a rate of $\Omega(1/\log \log n)$ for constant distance δ , but it improves upon the dependence on the distance parameter. Specifically, while the construction of [CHS18] has a rate that is upper bounded by $O(1/\log \log n)$ regardless of the value of δ , the construction of [BYCY22] achieves rate $\Omega(1/(\sqrt{\delta} \cdot \log \log n))$. In particular, it can attain *constant* rate by compromising on the distance, achieving a distance of $\delta = \Omega(1/(\log \log n)^2)$. This related, “dual” problem of constructing a constant-rate tree code with slowly deteriorating distance was first studied by Gelles, Haeupler, Kol, Ron-Zewi, and Wigderson [GHKRZW16], who showed how to achieve distance $\Omega(1/\log n)$.

¹We sometimes also refer to the depth n of a tree code as its *transmission length*.

1.1 Immediacy

Given that we still have no explicit construction of asymptotically good tree codes—despite significant efforts for over three decades—it is important to better understand why current techniques fall short. Identifying a barrier may help steer research away from potential dead ends. More specifically, since tree codes induce standard block error-correcting codes, while the latter are heavily relied upon in all known constructions (except for the two conjecture-based approaches [MS14, BYCN21]), it is natural to ask: to what extent might constant-rate tree codes be constructed from block codes? Exploring this question may shed light on the fundamental nature of tree codes and clarify how “close” they truly are to their well-understood counterparts, block error-correcting codes.

The main contribution of this paper is identifying a fundamental barrier in the construction of tree codes. We observe that each of the existing explicit constructions of tree codes (except those relying on unproven conjectures) satisfies a stronger guarantee than strictly required for a tree code; we refer to this guarantee as *immediacy*. This stronger property arose inherently due to the recursive nature of existing constructions [EKS94, CHS18, NW20, BYCY22]. We prove that the immediacy property incurs a cost in rate. Specifically, we establish a quantitative *rate-immediacy trade-off*, precisely capturing the achievable rate given a code’s immediacy. Our results show that existing tree code constructions essentially achieve an optimal rate-immediacy balance, thus clarifying why their rate inevitably vanishes. More broadly, our work indicates that the typically employed framework, that of tiling standard error-correcting block codes in progressively sophisticated ways, cannot yield asymptotically good tree codes. Such strategies inherently satisfy the immediacy property, constraining their potential rate.

We formally define immediacy in detail later, as its precise definition is somewhat technical, involving a certain set system (see the definition of *immediacy codes* in Section 3). However, to capture the essence, we first provide an informal, simplified definition in which immediacy is represented by a monotone-increasing function $\text{Imm} : \mathbb{N} \rightarrow \mathbb{N}$. A tree code C is said to have immediacy Imm if, for every pair of messages $x, y \in \Sigma^n$ (where n is the transmission length), every index s where $x_s \neq y_s$, and each integer $k \geq 1$, the relative Hamming distance between $C(x)$ and $C(y)$, restricted to the interval $[s, s + \text{Imm}(k))$ (provided that $s + \text{Imm}(k)$ is no larger than the depth of the tree, if finite), is bounded below by δ .

The definition of a standard tree code imposes no non-trivial immediacy constraint. Disagreements between corresponding codewords are required to accumulate only starting from the *first* index s at which the two messages x and y differ. In contrast, a code admitting an immediacy function requires that disagreements accumulate starting from *every* index s for which $x_s \neq y_s$, at interval lengths determined by the immediacy function Imm . Put differently, unlike a standard tree code, a tree code with immediacy function Imm supports *cold-start decoding*. That is, if during

any interval $[s, s + \text{Imm}(k))$ the noise rate is below $\delta/2$, then the receiver can decode the s -th bit—even if the noise prior to time s was high and bits prior to s cannot yet be reconstructed. While this property may be advantageous in certain applications (and is typical of weakly-encoded communications such as human speech or packetized transmissions), our results show that it is an “expensive” feature that cannot be afforded when aiming to construct asymptotically good tree codes.

To take an example, it can be shown by inspection that the first tree code construction [EKS94], which has a rate of $\Omega(1/\log n)$ (where n is the depth of the tree), possesses a non-trivial immediacy property that corresponds roughly to $\text{Imm}(k) = 2^{\Theta(k)}$, which we denote throughout this informal presentation as $\exp(k)$. At this informal stage, it is convenient to regard Imm as a real-valued function—which in this case is $\text{Imm}(x) = \exp(x)$ —thus extending the original discrete definition. With this, our main result can be informally stated as follows:

Theorem 2 (Main result; informal). *Let T be a depth- n tree code with constant distance and immediacy function Imm . Then, subject to some technical conditions² on Imm , that are satisfied for the cases $\text{Imm}(k) = \exp(k)$ and $\text{Imm}(k) = \exp(\exp(k))$, the rate of T satisfies*

$$\rho = O\left(\frac{1}{\text{Imm}^{-1}(n)}\right).$$

Recall that for the construction given in [EKS94], the immediacy corresponds roughly to $\text{Imm}(x) = \exp(x)$ and so $\text{Imm}^{-1}(x) = \Theta(\log x)$. Thus, Theorem 2 indicates that the rate of such a construction cannot exceed $O(1/\log n)$ —which matches its known rate, up to the constant. Theorem 2 also explains the rate achieved by the state-of-the-art tree code constructions [CHS18, NW20, BYCY22]. Upon inspection, these constructions—similarly to [EKS94]—also rely on composition of error-correcting codes and possess a non-trivial, though exponentially weaker, immediacy that corresponds roughly to $\text{Imm}(x) = \exp(\exp(x))$. Thus, Theorem 2 indicates that the rate of these constructions cannot exceed $O(1/\log \log n)$, matching the proven rate of these codes.

Weaker immediacy properties We emphasize here, however, that the precise immediacy properties enjoyed by the [EKS94] and [CHS18] constructions are slightly weaker than that implied by an immediacy function, and so Theorem 2 does not *formally* apply to them. Roughly speaking, these constructions only admit cold-start decoding after a “lag”. We show in Section 5 that our general framework of *immediacy codes* and our general rate upper bound results in Theorems 8 and 9 discussed in Section 3.1 are powerful enough to extend to constructions with

²For instance, it turns out that exponential immediacy (i.e. $\text{Imm}(k) = \exp(k)$ as described above) implies *linear* immediacy. Consequently, Theorem 2 should be interpreted as applying only to exponential (or faster-growing) immediacy.

such weaker forms of immediacy. Thus, as shown in Section 5, the rate upper bounds indicated by Theorem 2 indeed apply to the above constructions. (Further, as we show in Section 3.2, constructions satisfying the simplified immediacy condition described here are also immediacy codes with appropriate parameters and our general rate upper bounds proved in Theorems 8 and 9 therefore also apply to them.)

More generally, immediacy should be kept in mind when constructing tree codes: immediacy properties of a proposed construction can be analyzed in order to obtain an upper bound on the rate it can achieve. (Further, the above discussion also shows that a rough analysis along the lines of Theorem 2 may serve as a useful guide before a formal analysis of the exact immediacy properties of a proposed construction.) It is worth noting that the two constructions whose analyses rely on unproven conjectures [MS14, BYCN21] do not appear to exhibit non-trivial immediacy. Therefore, they are not excluded from being asymptotically good codes by this work.

Rate–distance–immediacy trade-off We point out that our rate upper bounds apply also to codes with vanishing distance. In fact, for a general relative distance $\delta > 0$, the upper bound on the rate ρ stated in Theorem 2 exhibits, roughly, an inverse-linear dependence on δ . The precise details are more nuanced, and we direct the reader to the formal statement of our main result in Theorem 8, and to the computations leading to eqs. (10) and (11) (that show that fast growing immediacy functions imply an appropriate immediacy codes condition). The key takeaway, however, is that Theorem 2, in its full generality, yields a *rate–distance–immediacy trade-off*. For ease of presentation, we choose to focus on the trade-off between rate and immediacy in this introductory section (see Section 1.3 for further discussion).

1.2 Related work

Randomized vs. deterministic encoding schemes Randomized hashing schemes were employed for online and interactive communication [Sch92] before deterministic schemes were made possible by tree codes. Randomized schemes do not suffer a trade-off between rate and immediacy. Computationally efficient encoding and decoding was shown for such schemes in [BK12, BN13, GH14]; in a considerable accomplishment Brakerski, Kalai, and Saxena [BKS20] later showed that, with access only to explicit (i.e., polynomial-time encodable) tree codes, an interactive communication can be performed at the rate of those tree codes, while also performing all necessary decoding in polynomial time.

Tree code variants The original motivation for tree codes came from the question of developing analogues of Shannon’s channel coding theorem (for one-way communication) for situations involving interaction. Using the existence of tree codes, such an interactive coding theorem

was proved by Schulman [Sch96, Theorem 2]. Later work by Gelles, Moitra, and Sahai [GMS11] showed that such an interactive coding theorem can also be established using a weaker variant of tree codes that they called *potent tree codes*. On the other hand, stronger variants of tree codes featuring local testability were recently introduced by Moud, Rosen, and Rothblum [MRR25]. It is interesting to compare this local testability guarantee with the immediacy property of tree codes, as immediacy also provides a form of local guarantee, albeit of a seemingly different nature. Lastly, we mention that a signal-processing analogue of tree codes was introduced by Schulman and Srivastava [SS19], and a variant of tree codes, dubbed *palette-alternating tree codes* which allows one to bypass the $\frac{1}{2}$ -rate barrier was introduced by Cohen and Samocha [CS20].

1.3 Technical overview and organization of the paper

We follow the tradition of using information theoretic tools for bounding capabilities of codes. Information theory had its genesis in understanding fundamental limitations on source and channel coding in the stochastic error model. However, it has also been useful in understanding limitations on codes in the “Hamming model” of a bounded number of worst case errors, including in the setting of more modern coding-theoretic notions such as the notion of local decoding analyzed by Katz and Trevisan [KT00, Section 3.1], as well as its more delicate relaxed variant analyzed by Gur and Lachish [GL21] (see also [DGL23, Gol25]). In our application to understanding immediacy properties of tree codes, the crucial ingredient turns out to be a careful accounting of the “common information” between different parts of the codeword, performed simultaneously at different length scales.

The main conceptual contribution of this paper, introduced in Section 3, is the notion of *immediacy codes*. The notion of immediacy codes provides a more robust formulation of the informal notion of immediacy functions discussed in the introduction (the close connection between the two notions is further explored in Section 3.2), and also allows us to smoothly carry out the accounting of common information alluded to above. The main information theoretic tool we need for this accounting is a form of the data processing inequality (Lemma 10)³ to quantify “common information” between different parts of a code word. Our main technical result is that *immediacy codes cannot have high rates* (Theorem 8): this is proved in Section 4. The principle underlying the definition of immediacy codes and the proof of Theorem 8 is a careful accounting of the amount of “information” that needs to be duplicated in order to provide the immediacy guarantees that the code provides. We emphasize that once we have set up the definition of immediacy codes and the requisite form of the data processing inequality, the proof of Theorem 8 is technically quite simple.

³This corollary of the (proof of the) data processing inequality also underlies the formulation of the *Gács-Körner common information* (see, e.g., [KA10, Section III]), but we do not need this connection for this paper.

Finally, in Section 5, we apply our framework to study the immediacy properties of the tree code constructions proposed in [EKS94], [CHS18] and [GHRZW16]. As discussed earlier, the immediacy properties of these constructions differ slightly from the informal notion introduced above. For the first two constructions, we show that the rates achieved by these codes are optimal—up to constant factors—with respect to the immediacy guarantees they provide. In Section 5.1.1, we consider a method due to Braverman [Bra12] for constructing depth- n good tree codes in time sub-exponential in n , and show that the rate-runtime trade-off obtained by the best currently known version of this construction is essentially the best possible given the immediacy guarantees achieved by the construction.

The construction of [GHRZW16] attempts to tackle a different trade-off: it achieves an $\Omega(1/\log n)$ distance with constant rate. We show that for the immediacy guarantee that this construction provides, no constant rate tree code can improve the distance to $\omega\left(\frac{\log \log n}{\log n}\right)$. This establishes that the rate/distance trade-off achieved by this construction is close to the best possible given the immediacy guarantees it achieves.

We do not carry out a full analysis of the immediacy properties of the constructions in [NW20, BYCY22]. The construction of [NW20] employs the same “code tiling” structure as [CHS18]; the main difference lies in the “outer” large-alphabet tree code used. Since immediacy is determined primarily by the code tiling aspect, the analysis of [NW20] proceeds similarly to our analysis of [CHS18]. The construction of [BYCY22], in contrast, uses a somewhat different code-tiling strategy. Nevertheless, a variant of our analysis for [CHS18] can still be adapted to study its immediacy properties. Lastly, we note that the construction of [GHRZW16] is similar to that of [EKS94] with respect to the code-tiling component. As such, our result—when viewed not as a rate-immediacy trade-off but rather as a distance-immediacy trade-off—can be used to establish an upper bound on the distance achievable by the construction.

We begin in the next section with some technical preliminaries.

2 Preliminaries

Notation We will view strings s of length ℓ as indexed by integers in $\{1, 2, \dots, \ell\} =: [\ell]$. Given such a string, and a subset $A \subseteq [\ell]$ we denote by s_A the sub-sequence of s obtained by taking the letters at the indices in A . We will sometimes overload the usual notation for real intervals (e.g. $[a, b]$, $(a, b]$ etc.) to refer only to the *integers* in those intervals: this overloading should always be clear from the context. We denote $\log_2 x$ by $\lg x$ for any positive real x .

2.1 Information theory

We use standard information theoretic notation for entropy, conditional entropy and mutual information (see, e.g., [CT06]), which we now proceed to review. Given a random variable X taking values in some finite set Ω , its *entropy* $H(X)$ is defined as $-\sum_{\omega \in \Omega} \Pr[X = \omega] \cdot \lg \Pr[X = \omega]$ (so we measure entropy in bits), with the convention $0 \cdot \lg 0 = 0$. Given two random variables X and Y , both taking values in (possibly different) finite sets, the *conditional entropy* $H(X|Y)$ is defined as

$$H(X|Y) := - \sum_{x,y} \Pr[X = x, Y = y] \cdot \lg \Pr[X = x|Y = y]. \quad (1)$$

The *mutual information* $I(X : Y)$ is then defined as

$$I(X : Y) = H(X) - H(X|Y). \quad (2)$$

If Z is another random variable taking values in a finite set, the *conditional mutual information* $I(X : Y|Z)$ is defined as

$$I(X : Y|Z) = H(X|Z) - H(X|Y, Z). \quad (3)$$

We collect here some of the standard properties of these quantities, which we shall subsequently use without comment. Proofs of these properties can be found, e.g., in [CT06].

Proposition 3 (Properties of mutual information and entropy). *Let X, Y and Z be random variables taking values in finite sets $\Omega_1, \Omega_2, \Omega_3$ respectively. Then*

- (Non-negativity of the entropy) $0 \leq H(X) \leq \lg |\Omega_1|$.
- (Non-negativity of the mutual information) $I(X : Y) \geq 0$ and $I(X : Y|Z) \geq 0$.
- (Conditioning reduces entropy) $H(X|Y, Z) \leq H(X|Y) \leq H(X)$.
- (Chain rule for entropy) $H(X, Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$.
- (Sub-additivity of entropy) $H(X, Y) \leq H(X) + H(Y)$.
- (Chain rule for mutual information) $I(X, Y : Z) = I(X : Z) + I(Y : Z|X) = I(Y : Z) + I(X : Z|Y)$.
- (Conditional entropy of deterministic functions is zero) *If there is a function $f : \Omega_2 \rightarrow \Omega_1$ such that $f(Y) = X$, then $H(X|Y) = 0$.*

3 Immediacy codes

In this section, we first introduce the notion of *immediacy codes* (Definition 6), a formal abstraction of the informal notion of immediacy discussed in the introduction. To connect the two notions, we show in Section 3.2 how codes with the two kinds of immediacy functions discussed in the introduction satisfy the immediacy code property. In Section 4, we prove our main technical result (Theorem 8): a general rate upper bound for immediacy codes.

The motivation behind our notion of immediacy codes is to capture the *cold-start decoding* property implied by immediacy: if the noise in a given interval is small, then a code with immediacy properties should be able to decode at least part of the interval. For greater applicability of our framework, we would not like to impose such a condition on *all* intervals, and towards this goal we begin with a few combinatorial definitions. By a *partition* of a set U , we mean an ordered tuple of pairwise disjoint non-empty subsets of U whose union is U .

Definition 4 (Tagged partitions). A *tagged partition* S of $[n]$ is a partition (S_1, S_2, \dots, S_k) of $[n]$, along with a partition $(\text{left}(S_i), \text{right}(S_i))$ of each S_i , $1 \leq i \leq k$, into two non-empty sets.

Definition 5 (Laminar partitions). An (α, ℓ) -*laminar partition* of n is a tuple $P = (P_0, P_1, P_2, \dots, P_\ell)$ of partitions of $[n]$ into disjoint non-empty subsets, in which P_1, \dots, P_ℓ are tagged partitions, and which satisfies the following properties:

1. **(Size property)** For each $1 \leq i \leq \ell$, and for each $B \in P_i$, $|\text{left}(B)| \geq \alpha |B|$. (Here, $\text{left}(\cdot)$ is as in the definition of a tagged partition).
2. **(The laminar property)** For each partition P_i with $i \geq 1$, and any subset $B \in P_i$, there exist sub-collections of P_{i-1} that are partitions of $\text{left}(B)$ and $\text{right}(B)$, respectively, into pairwise disjoint non-empty sets (see fig. 1 for an illustration).

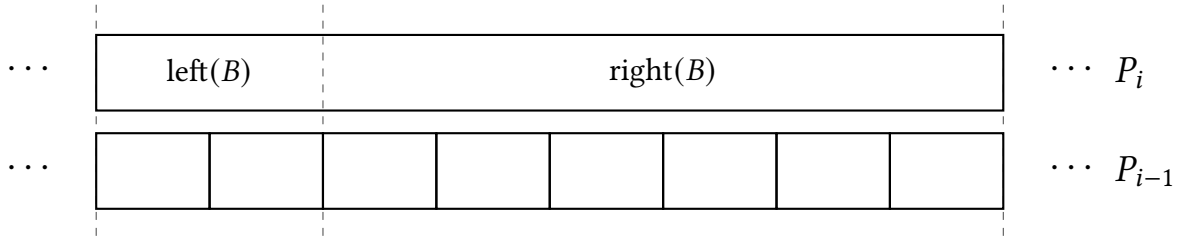


Figure 1: An illustration of laminar partitions

Definition 6 (Immediacy codes). An (α, ℓ) -*immediacy code* with code alphabet Σ and transmission length n is a function $c : \Sigma_{\text{in}}^n \rightarrow \Sigma^n$ along with an (α, ℓ) -laminar partition $P = (P_0, P_1, P_2, \dots, P_\ell)$ of $[n]$ such that

- **(The code decodes a large neighborhood)** If B is a subset in a partition $P_i \in P$ for $i \geq 1$, then for all $x, y \in \Sigma_{\text{in}}^n$, $x_{\text{left}(B)} \neq y_{\text{left}(B)}$ implies that $c(x)_{\text{right}(B)} \neq c(y)_{\text{right}(B)}$. In other words, there is a function $\phi_B : \Sigma^{|\text{right}(B)|} \rightarrow \Sigma_{\text{in}}^{|\text{left}(B)|}$ such that $\phi_B(c(x)_{\text{right}(B)}) = x_{\text{left}(B)}$ for all $x \in \Sigma_{\text{in}}^n$.

For some applications, the neighborhood decoding condition above may not hold for all subsets in all partitions. For handling these cases, we will also need the following weakening of the above definition.

Definition 7 (Deficient Immediacy code). A D -deficient (α, ℓ) -immediacy code with code alphabet Σ and transmission length n is a function $c : \Sigma_{\text{in}}^n \rightarrow \Sigma^n$ along with an (α, ℓ) -laminar partition $P = (P_0, P_1, P_2, \dots, P_\ell)$ of $[n]$ such that

1. There exist *deficiency subsets* S_i of P_i , for $1 \leq i \leq \ell$, such that the total size of all the subsets contained in the S_i is at most D . In symbols,

$$\sum_{i=1}^{\ell} \sum_{B \in S_i} |B| \leq D. \quad (4)$$

2. **(The code often decodes a large neighborhood)** If B is a subset in a partition $P_i \in P$ for $i \geq 1$ such that B is not an element of the corresponding deficiency set S_i , then for all $x, y \in \Sigma_{\text{in}}^n$, $x_{\text{left}(B)} \neq y_{\text{left}(B)}$ implies that $c(x)_{\text{right}(B)} \neq c(y)_{\text{right}(B)}$. In other words, there is a function $\phi_B : \Sigma^{|\text{right}(B)|} \rightarrow \Sigma_{\text{in}}^{|\text{left}(B)|}$ such that $\phi_B(c(x)_{\text{right}(B)}) = x_{\text{left}(B)}$ for all $x \in \Sigma_{\text{in}}^n$.

3.1 Rate upper bounds for immediacy codes

We can now state our main technical results, which give rate upper bounds for (deficient) immediacy codes.

Theorem 8 (Rate upper bound for immediacy codes). Let n be a positive integer, and suppose that $c : \Sigma_{\text{in}}^n \rightarrow \Sigma^n$ is an $(\alpha(n), \ell(n))$ -immediacy code. Then,

$$\lg |\Sigma| \geq \alpha(n)\ell(n) \lg |\Sigma_{\text{in}}|. \quad (5)$$

Theorem 9 (Rate upper bound for deficient immediacy codes). Let n be a positive integer, and suppose that $c : \Sigma_{\text{in}}^n \rightarrow \Sigma^n$ is a D -deficient $(\alpha(n), \ell(n))$ -immediacy code. Then,

$$\lg |\Sigma| \geq \alpha(n) \cdot \left(\ell(n) - \frac{D}{n} \right) \cdot \lg |\Sigma_{\text{in}}|. \quad (6)$$

Remark In Section 3.2, we describe how the notion of immediacy handled in the above theorems relates to the informal discussion in the Introduction in terms of immediacy functions. We then show how the informal statement of Theorem 2 can be formalized using the above theorems. Note, however, that the more abstract nature of the above theorems allows us to handle the subtler immediacy properties of the known tree code constructions we consider in Section 5.

3.2 Immediacy functions and immediacy codes

We now show that tree codes satisfying the immediacy conditions described in the introduction also satisfy the immediacy code condition, with parameters depending upon the distance and the immediacy function. In order to avoid technical issues such as divisibility, we make here some simplifying assumptions on the form of the immediacy function. We note that these assumptions apply to the examples of immediacy functions discussed in the introduction, and also that the framework of immediacy codes is flexible enough that one can work through a similar route even when they do not hold.

Let $T : \{0, 1\}^n \rightarrow \Sigma^n$ be a tree code with immediacy function Imm , distance parameter $\delta \in (0, 1)$, and suppose that its depth $n = 2 \cdot \text{Imm}(\ell t)$, where ℓ and t are positive integers, with $t = t(\delta)$ possibly depending upon the distance δ . Let $\kappa := \lfloor \lg(2/\delta) \rfloor$ be a positive integer so that $2^{-\kappa} < \delta \leq 2^{-\kappa+1}$. We assume that t can be chosen so that for each $1 \leq j \leq \ell$,

$$\begin{aligned} 2^{-\kappa} \cdot \text{Imm}(jt) \text{ is an integer, and} \\ 2 \cdot \text{Imm}((j-1)t) \text{ divides } 2^{-\kappa} \cdot \text{Imm}(jt) \end{aligned} \tag{7}$$

(We note that this is the main technical assumption that would need to be modified when working with immediacy functions other than those discussed in the introduction. For the two immediacy functions discussed in the introduction, we show how to choose such a t towards the end of this section.)

Given these preliminaries, we now show that T is an (α, ℓ) -immediacy code with $\alpha := 2^{-(\kappa+1)}$ and ℓ as above. We first show that with the divisibility condition in eq. (7), we naturally obtain an (α, ℓ) laminar partition $(P_0, P_1, P_2, \dots, P_\ell)$. For each $0 \leq j \leq \ell$, partition $[n] = [2\text{Imm}(\ell t)]$ into consecutive blocks of length $2\text{Imm}(jt)$ each, and let these blocks constitute the tagged partition P_j . Given a block $B \in P_j$ for $j \geq 1$, we define $\text{left}(B)$ to be the set of the leftmost $2^{-\kappa} \cdot \text{Imm}(jt)$ integers in B and $\text{right}(B)$ to be the rightmost $2\text{Imm}(jt) - |\text{left}(B)|$ integers in B . Given the divisibility conditions in eq. (7), we can then verify by a direct computation that both the laminar and size properties in the definition of an (α, ℓ) -laminar partition are satisfied by this construction. In particular, $\text{left}(B)$ and $\text{right}(B)$ are disjoint unions of blocks in P_{j-1} .

We now show that T is also an immediacy code with respect to the tagged partition $(P_0, P_1, \dots, P_\ell)$.

Fix any $1 \leq j \leq \ell$ and any block B in the tagged partition P_j . We only need to show that if $x, y \in \{0, 1\}^n$ differ on some index in $\text{left}(B)$, then $T(x)$ and $T(y)$ differ on some index in $\text{right}(B)$. For any such x, y , let i be the smallest index in B on which x and y differ, and let S be the interval $[i, i + \text{Imm}(jt))$ of length $\text{Imm}(jt)$. By the choice of x and y , we have $i \in \text{left}(B)$. Since $|\text{left}(B)| = \text{Imm}(jt) \cdot 2^{-\kappa}$, $|B| = 2\text{Imm}(jt)$, and $\kappa \geq 1$, it then follows that S is contained in B , and intersects $\text{right}(B)$. Further, by the definition of the immediacy function, the Hamming distance between $T(x)_S$ and $T(y)_S$ must be at least

$$\delta \cdot \text{Imm}(jt) > 2^{-\kappa} \text{Imm}(jt) = |\text{left}(B)|.$$

Thus, since this Hamming distance is greater than the length of $\text{left}(B)$, it must be the case that $T(x)_S$ and $T(y)_S$ differ also on some index in $\text{right}(B)$, as we wanted to prove. This proves that T is an (α, ℓ) -immediacy code, with $\alpha = 2^{-(\kappa+1)}$, as chosen above.

Rate upper bounds with immediacy functions Using the rate upper bound for (α, ℓ) -immediacy codes given by our main result (Theorem 8), we thus get that if Σ is the output alphabet of T , it must be the case that (recall that $n = 2 \cdot \text{Imm}(\ell t(\delta))$)

$$\lg |\Sigma| \geq \frac{\ell}{2^{1+\kappa}} = \frac{\text{Imm}^{-1}(n/2)}{t(\delta) \cdot 2^{1+\kappa}} \geq \frac{\delta}{4t(\delta)} \text{Imm}^{-1}(n/2). \quad (8)$$

Put differently, the rate ρ of such a tree code satisfies

$$\rho \leq \frac{4t(\delta)}{\delta \cdot \text{Imm}^{-1}(n/2)}. \quad (9)$$

(We remark here that since we work in the setting of constant distance δ , we have not attempted to optimize the dependence on δ in the above computation.)

Choosing $t(\delta)$. We now show how to choose $t(\delta)$ so as to satisfy the divisibility conditions of eq. (7), for the two immediacy functions considered in the introduction.

Case 1: $\text{Imm}(k) = 2^k$. Put $t(\delta) = 1 + \kappa = \lfloor \lg(4/\delta) \rfloor$. A direct computation shows that the condition in eq. (7) holds with this choice. The rate upper bound in eq. (9) becomes

$$\rho \leq \frac{4 \lg(4/\delta)}{\delta \cdot \lg(n/2)}. \quad (10)$$

Case 2: $\text{Imm}(k) = 2^{2^k}$. Put $t(\delta) := \lceil \lg(\kappa + 2) \rceil = \lceil \lg \lfloor \lg(8/\delta) \rfloor \rceil$. Again, a direct computation shows that the condition in eq. (7) holds with this choice. The rate upper bound in eq. (9)

becomes

$$\rho \leq \frac{4 \cdot \lceil \lg \lg(8/\delta) \rceil}{\delta \cdot \lg \lg(n/2)}. \quad (11)$$

4 Rate upper-bound for immediacy codes

In this section, we prove our main result: a rate upper bound for immediacy codes (Theorem 8). The main ingredient is a simple information theoretic tool, that we describe next. As stated in the introduction, given this tool and the definition of immediacy codes, the proof of the main result becomes quite simple.

4.1 Data processing inequality

We will need the following simple consequence of (the usual proof of) the data processing inequality. Similar inequalities arise also in studies of the Gács-Körner common information (see, e.g., [KA10]).

Lemma 10 (A consequence of the data processing inequality). *Let A, B, C be discrete random variables such that $H(A|B) = H(A|C) = 0$. Then $I(B : C) \geq H(A)$. In particular,*

$$H(B) + H(C) \geq H(B, C) + H(A).$$

Proof. Since A, B, C are discrete, the associated entropies and conditional entropies are non-negative. Combining this with the non-negativity of mutual information and conditional mutual information, we thus have

$$0 \leq I(A : C|B) \leq H(A|B) = 0. \quad (12)$$

We also have $I(A : C) = H(A) - H(A|C) = H(A)$. The chain rule for mutual information along with the non-negativity of the conditional mutual information then gives

$$\begin{aligned} H(A) = I(A : C) &\leq I(A : C) + I(B : C|A) \\ &= I(A, B : C) = I(B : C) + I(A : C|B) \stackrel{\text{eq. (12)}}{=} I(B : C), \end{aligned}$$

which proves that $I(B : C) \geq H(A)$. The final claim follows since $H(B, C) = H(B) + H(C) - I(B : C)$. \square

4.2 Proof of the main result

We now prove our main result (Theorem 8). We restate the result here for convenience of reference: please also see the remark at the end of Section 3.1 for further discussion.

Theorem 8 (Rate upper bound for immediacy codes). *Let n be a positive integer, and suppose that $c : \Sigma_{\text{in}}^n \rightarrow \Sigma^n$ is an $(\alpha(n), \ell(n))$ -immediacy code. Then,*

$$\lg |\Sigma| \geq \alpha(n)\ell(n) \lg |\Sigma_{\text{in}}|. \quad (5)$$

Proof. To simplify notation, we shorten $\ell(n)$ to ℓ in the following. Let $P = (P_0, P_1, \dots, P_\ell)$ be the laminar partition associated with the immediacy code c . Modify the code c so that it is systematic, in the sense that $c(x)_k$ determines x_k for each $1 \leq k \leq n$: this can be done by changing the alphabet to $\Sigma \times \Sigma_{\text{in}}$, and simply concatenating x_k to each $c(x)_k$. Note that even after the modification, the code c continues to be an immediacy code, with the same associated laminar partition P . However, the output alphabet of the code now is Σ' where $\Sigma' = \Sigma \times \Sigma_{\text{in}}$.

Let X be uniformly distributed over Σ_{in}^n , and define $Y = c(X)$. We then have $H(X) = H(Y) = n \cdot \lg |\Sigma_{\text{in}}|$, where we measure entropy in bits. Now, consider any subset B that is an element of one of the partitions P_i , for some $i \geq 1$. Since c is systematic, we then have $H(X_{\text{left}(B)} | Y_{\text{left}(B)}) = H(X_{\text{left}(B)} | c(X)_{\text{left}(B)}) = 0$. Further, by the neighborhood decoding property of the immediacy code c , we also have

$$H(X_{\text{left}(B)} | Y_{\text{right}(B)}) = H(\phi_B(c(X)_{\text{right}(B)}) | c(X)_{\text{right}(B)}) = 0. \quad (13)$$

(Here, ϕ_B is as in the definition of the neighborhood decoding property of an immediacy code.) Applying Lemma 10, we then have

$$H(Y_B) = H(Y_{\text{left}(B)}, Y_{\text{right}(B)}) \leq H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)}) - H(X_{\text{left}(B)}) \quad (14)$$

$$= H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)}) - |\text{left}(B)| \cdot \lg |\Sigma_{\text{in}}| \quad (15)$$

$$\leq H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)}) - \alpha(n)|B| \cdot \lg |\Sigma_{\text{in}}|, \quad (16)$$

where the last inequality follows since $|\text{left}(B)| \geq \alpha(n)|B|$, by the size property of $(\alpha(n), \ell(n))$ -laminar partitions. Note that eq. (16) holds for every subset B that is part of some partition P_i , $i \geq 1$, of P . Define, for $0 \leq i \leq \ell$

$$T_i := \sum_{B \in P_i} H(Y_B). \quad (17)$$

We then have, for $1 \leq i \leq \ell$,

$$T_i = \sum_{B \in P_i} H(Y_B) \stackrel{\text{eq. (16)}}{\leq} -\alpha(n)n \cdot \lg |\Sigma_{\text{in}}| + \sum_{B \in P_i} (H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)})) \quad (18)$$

$$\leq -\alpha(n)n \cdot \lg |\Sigma_{\text{in}}| + \sum_{B' \in P_{i-1}} H(Y_{B'}) = T_{i-1} - \alpha(n)n \cdot \lg |\Sigma_{\text{in}}|. \quad (19)$$

where the second inequality follows from the laminar property of the laminar partition P (which ensures that each of $\text{left}(B)$ and $\text{right}(B)$ are disjoint unions of sets in the partition P_{i-1} , for every set $B \in P_i$), along with the sub-additivity of entropy: for any two random variables Z_1 and Z_2 , $H(Z_1, Z_2) \leq H(Z_1) + H(Z_2)$. By induction, eq. (19) thus gives

$$T_\ell \leq T_0 - \alpha(n)n\ell \cdot \lg |\Sigma_{\text{in}}|. \quad (20)$$

The subadditivity of entropy also gives $T_\ell \geq H(Y) = n \cdot \lg |\Sigma_{\text{in}}|$, and

$$T_0 \leq \sum_{i=1}^n H(Y_i) \leq n \cdot (\lg |\Sigma_{\text{in}}| + \lg |\Sigma|), \text{ since each } Y_i \text{ has support } \Sigma' = \Sigma \times \Sigma_{\text{in}}. \quad (21)$$

Substituting these in eq. (20) gives $\lg |\Sigma| \geq \alpha(n)\ell \cdot \lg |\Sigma_{\text{in}}|$. \square

The proof above is robust to certain small perturbations to the definition of an immediacy code. In particular, it can be easily adapted to establish rate upper bounds for D -deficient immediacy codes as well, as we show below. We restate the corresponding result (Theorem 9) below for ease of reference.

Theorem 9 (Rate upper bound for deficient immediacy codes). *Let n be a positive integer, and suppose that $c : \Sigma_{\text{in}}^n \rightarrow \Sigma^n$ is a D -deficient $(\alpha(n), \ell(n))$ -immediacy code. Then,*

$$\lg |\Sigma| \geq \alpha(n) \cdot \left(\ell(n) - \frac{D}{n} \right) \cdot \lg |\Sigma_{\text{in}}|. \quad (6)$$

Proof. The proof is a minor modification of the proof of Theorem 8, but we include all the details for completeness. Let $P = (P_0, P_1, \dots, P_\ell)$ be the laminar partition associated with the D -deficient immediacy code c , and let the deficiency subsets $S_i \subseteq P_i$ be as in the definition (Definition 7). Modify the code c so that it is systematic, in the sense that $c(x)_k$ determines x_k for each $1 \leq k \leq n$: this can be done by changing the alphabet to $\Sigma \times \Sigma_{\text{in}}$, and simply concatenating x_k to each $c(x)_k$. Note that even after the modification, the code c continues to be a D -deficient immediacy code, with the same associated laminar partition P and the same associated deficiency sets. However, the output alphabet of the code now is Σ' where $\Sigma' = \Sigma \times \Sigma_{\text{in}}$.

Let X be uniformly distributed over Σ_{in}^n , and define $Y = c(X)$, so that $H(X) = H(Y) = n \cdot \lg |\Sigma_{\text{in}}|$. Now, consider any subset B that is an element of one of the partitions P_i , for some $i \geq 1$. Since c is systematic, we then have $H(X_{\text{left}(B)} | Y_{\text{left}(B)}) = H(X_{\text{left}(B)} | c(X)_{\text{left}(B)}) = 0$. Further, in case $B \notin S_i$, then by the neighborhood decoding property of the deficient immediacy code c , we also have

$$H(X_{\text{left}(B)} | Y_{\text{right}(B)}) = H(\phi_B(c(X)_{\text{right}(B)}) | c(X)_{\text{right}(B)}) = 0. \quad (22)$$

(Here, ϕ_B is as in the definition of the neighborhood property of a deficient immediacy code.) Applying Lemma 10, we then have

$$H(Y_B) = H(Y_{\text{left}(B)}, Y_{\text{right}(B)}) \leq H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)}) - H(X_{\text{left}(B)}) \quad (23)$$

$$= H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)}) - |\text{left}(B)| \cdot \lg |\Sigma_{\text{in}}| \quad (24)$$

$$\leq H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)}) - \alpha(n)|B| \cdot \lg |\Sigma_{\text{in}}|, \quad (25)$$

where the last inequality follows since $\text{left}(B) \geq \alpha(n)|B|$, by the size property of $(\alpha(n), \ell(n))$ -laminar partitions. Note that when $i \geq 1$, eq. (25) holds for every subset $B \in P_i \setminus S_i$. The only difference with the proof of Theorem 8 is that when $B \in P_i$ is an element of the deficiency set S_i , we however only have (by the sub-additivity of the entropy):

$$H(Y_B) = H(Y_{\text{left}(B)}, Y_{\text{right}(B)}) \leq H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)}). \quad (26)$$

Define, for $0 \leq i \leq \ell$

$$T_i := \sum_{B \in P_i} H(Y_B). \quad (27)$$

We then have, for $1 \leq i \leq \ell$,

$$\begin{aligned} T_i &= \sum_{B \in P_i} H(Y_B) \stackrel{\text{eqs. (25) and (26)}}{\leq} -\alpha(n)n \cdot \lg |\Sigma_{\text{in}}| + \alpha(n) \cdot \lg |\Sigma_{\text{in}}| \cdot \sum_{B \in S_i} |B| \\ &\quad + \sum_{B \in P_i} (H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)})) \end{aligned} \quad (28)$$

$$\leq -\alpha(n)n \cdot \lg |\Sigma_{\text{in}}| + \alpha(n) \cdot \lg |\Sigma_{\text{in}}| \cdot \sum_{B \in S_i} |B| + \sum_{B' \in P_{i-1}} H(Y_{B'}) \quad (29)$$

$$= T_{i-1} - \alpha(n)n \cdot \lg |\Sigma_{\text{in}}| + \alpha(n) \cdot \lg |\Sigma_{\text{in}}| \sum_{B \in S_i} |B|. \quad (30)$$

where the second inequality follows from the laminar property of the laminar partition P (which ensures that each of $\text{left}(B)$ and $\text{right}(B)$ are disjoint unions of sets in the partition P_{i-1} , for every

set $B \in P_i$), along with the sub-additivity of entropy. By induction, eq. (30) thus gives

$$T_\ell \leq T_0 - \alpha(n)n\ell \cdot \lg |\Sigma_{\text{in}}| + \alpha(n) \cdot \lg |\Sigma_{\text{in}}| \cdot \sum_{i=1}^{\ell} \sum_{B \in S_i} |B| \leq T_0 - \alpha(n) \cdot \lg |\Sigma_{\text{in}}| \cdot (n\ell - D), \quad (31)$$

where the last inequality uses the fact that the code is only D -deficient (eq. (4) of Definition 7). The subadditivity of entropy yields $T_\ell \geq H(Y) = n \cdot \lg |\Sigma_{\text{in}}|$, and also that

$$T_0 \leq \sum_{i=1}^n H(Y_i) \leq n \cdot (\lg |\Sigma_{\text{in}}| + \lg |\Sigma|), \text{ since each } Y_i \text{ has support } \Sigma' = \Sigma \times \Sigma_{\text{in}}. \quad (32)$$

Substituting these in eq. (31) gives $\lg |\Sigma| \geq \alpha(n) \cdot \left(\ell - \frac{D}{n}\right) \cdot \lg |\Sigma_{\text{in}}|$. \square

5 Rate upper-bounds for immediacy properties of known constructions

In this section, we describe the immediacy properties enjoyed by the tree code constructions of [CHS18], [EKS94] and [GHKRZW16]. The first two are similar to—but not the same as—the informal notion of immediacy discussed in the introduction. However, as we show below, our general framework of immediacy codes developed in Section 3 still applies, and allows us to conclude that the $\Omega(1/\log \log n)$ and $\Omega(1/\log n)$ rates achieved by these constructions are the best possible (up to constant factors) given the immediacy properties they achieve. The third construction, due to [GHKRZW16], tackles a different trade-off: what is the best (even if vanishing, as n increases) distance one can achieve for a tree code if one imposes the condition that the rate has to be constant. For this construction we show that the distance $\Omega(1/\log n)$ distance it achieves with a constant rate is tight up to a $\log \log n$ factor for the immediacy guarantee that this construction provides.

5.1 The CHS construction

We start with the construction of [CHS18] (which for brevity we shall call the CHS construction). To describe the immediacy properties of this construction, we first import some of the notation set up in [CHS18]. In the following, references to theorems, pages etc. in [CHS18] refer to the ECC version of [CHS18]. Define a sequence of length scales (following the proof of Theorem 1.1 in [CHS18, p. 16]) given by $\ell_1 := 2^{20}$ and $\ell_{i+1} := \ell_i^2/2^{10}$ for $i \geq 1$; thus $\ell_i = 2^{10} \cdot 32^{2^i}$ for every positive integer i . We consider a transmission length n of the form ℓ_{m+1} for some positive integer m . Thus, n is divisible by each ℓ_i , $1 \leq i \leq m$, and no divisibility issues arise in the following

description. We also note for future use that for $i \geq 1$

$$16\sqrt{\ell_i} \leq \frac{3\ell_i}{8}. \quad (33)$$

For each length scale ℓ_i , $1 \leq i \leq m+1$ as above, we consider the partition P_{i-1} of $[n]$ into $2n/\ell_i$ consecutive disjoint intervals of length $\ell_i/2$ each. In agreement with the terminology in [CHS18], we refer to the elements of the P_i as *blocks*. For a block $B \in P_i$ for $i \geq 1$, we denote by $\text{left}(B)$ the set consisting of the first $|B|/4$ positions in B , and by $\text{right}(B)$ the set consisting of the last $3|B|/4$ positions in B . With these definitions, $P := (P_0, P_1, P_2, \dots, P_m)$ is a $(1/4, m)$ -laminar partition of $[n]$. In particular, for any $B \in P_i$, $i \geq 1$, $\text{left}(B)$ and $\text{right}(B)$ are disjoint unions of blocks in P_{i-1} , because $\ell_i/2$ divides $\ell_{i+1}/8$.

The CHS immediacy condition We can now describe the immediacy-like property satisfied by the CHS tree code construction. Denote their code by T . Let x and y be two strings in $\{0, 1\}^n$ which differ at a position $s' \in [n]$. Then, for any i such that $2 \leq i \leq m+1$ such that s' is *not* in the rightmost block in P_{i-1} (i.e., such that $s' \leq n - \ell_i/2$), the following is true. Let $s(i)$ denote the *leftmost* position in the unique block $B \in P_{i-1}$ containing s' such that $x_{s(i)} \neq y_{s(i)}$. Then, for any d satisfying

$$\ell_{i-1}/2 = 16\sqrt{\ell_i} \leq d \leq \ell_i/2, \quad (34)$$

it holds that

$$\Delta_{\text{Hamming}}(T(x)_{I(d)}, T(y)_{I(d)}) \geq d/3, \quad (35)$$

where $I(d)$ denotes the interval

$$I(d) := [s(i), s(i) + d]. \quad (36)$$

The above condition is implicit in the proof given in [CHS18]: in particular, it follows by substituting $s(i)$ above in the role played by the “split” s in the proof of Claim 6.4 of [CHS18].

We then have the following consequence of the above immediacy condition.

Lemma 11. *Let m be a positive integer. Let the sequence ℓ_1, ℓ_2, \dots , the positive integer $n = \ell_{m+1}$, and the $(1/4, m)$ -laminar partition $P = (P_0, P_1, P_2, \dots, P_m)$ of n be as defined above. Let $T : \{0, 1\}^n \rightarrow \Sigma^n$ be a code satisfying the CHS immediacy condition described above. Consider any block B in a partition $P_{i-1} \in P$, $2 \leq i \leq m+1$, so that B is not the rightmost block in P_{i-1} . Then, there exists a function ϕ_B such that for any $x \in \{0, 1\}^n$, $\phi_B(c(x)_{\text{right}(B)}) = x_{\text{left}(B)}$. In other words, the prefix $x_{\text{left}(B)}$ of x_B is uniquely determined given the suffix $c(x)_{\text{right}(B)}$ of $c(x)_B$.*

Proof. It is sufficient to show that if $x, y \in \{0, 1\}^n$ are such that $x_{\text{left}(B)} \neq y_{\text{left}(B)}$ then it must be the case that $c(x)_{\text{right}(B)} \neq c(y)_{\text{right}(B)}$. Consider therefore $x, y \in \{0, 1\}^n$ such that $x_{\text{left}(B)} \neq y_{\text{left}(B)}$.

Thus, there must be a position s' in $\text{left}(B)$ such that $x_{s'} \neq y_{s'}$. Let $s(i)$ be the leftmost such position in $\text{left}(B)$. Choose d so that the interval

$$I(d) = [s(i), s(i) + d] = B \cap [s, n],$$

i.e., so that $I(d)$ is the suffix of B starting at $s(i)$. Since $|B| = \ell_i/2$, this d is of the form $\ell_i/2 - k(i)$, where $k(i) \geq 1$ is the relative index, counting starting with one from the left, of $s(i)$ within B . Since $s(i) \in \text{left}(B)$ and $\text{left}(B)$ consists of the leftmost $|B|/4 = \ell_i/8$ positions in B , it is also the case that $k(i) \leq \ell_i/8$. Given eq. (33), d therefore satisfies the lower bound required in eq. (34) since $k(i) \leq \ell_i/8$ and $i \geq 2$. It also satisfies the upper bound required in eq. (34) by construction. Since B is, by hypothesis, not the rightmost block in P_{i-1} , eq. (35) then yields that

$$\Delta_{\text{Hamming}}(c(x)_{I(d)}, c(y)_{I(d)}) \geq d/3 = \frac{\ell_i}{8} + \frac{\ell_i/8 - k(i)}{3}. \quad (37)$$

Suppose, if possible, that $c(x)_{\text{right}(B)} = c(y)_{\text{right}(B)}$, so that $\Delta_{\text{Hamming}}(c(x)_{\text{right}(B)}, c(y)_{\text{right}(B)}) = 0$. Let S denote the suffix of $\text{left}(B)$ starting from the position $s(i)$, so that $I(d) = S \sqcup \text{right}(B)$. We then have

$$\Delta_{\text{Hamming}}(c(x)_{I(d)}, c(y)_{I(d)}) = \Delta_{\text{Hamming}}(c(x)_S, c(y)_S) + \Delta_{\text{Hamming}}(c(x)_{\text{right}(B)}, c(y)_{\text{right}(B)}) \quad (38)$$

$$= \Delta_{\text{Hamming}}(c(x)_S, c(y)_S) \leq |S| = \frac{\ell_i}{8} - k(i) + 1. \quad (39)$$

However, this contradicts eq. (37) since $\ell_i \geq 32$ and $k(i) \geq 1$. It must therefore be the case that $c(x)_{\text{right}(B)} \neq c(y)_{\text{right}(B)}$. \square

It is an easy consequence of the above lemma that any binary tree code satisfying the CHS immediacy condition is a deficient immediacy code, with a small deficiency parameter.

Proposition 12. *Let m be a positive integer. Let the sequence ℓ_1, ℓ_2, \dots , the positive integer $n = \ell_{m+1}$, and the $(1/4, m)$ -laminar partition $P = (P_0, P_1, P_2, \dots, P_m)$ of n be as defined above. Let $T : \{0, 1\}^n \rightarrow \Sigma^n$ be a code satisfying the CHS immediacy condition. Then T is an n -deficient $(1/4, m)$ -immediacy code.*

Proof. For each $1 \leq i \leq m$ define S_i to consist only of the rightmost block in P_i . Then, it follows directly from Lemma 11 that T is a D -deficient $(1/4, m)$ -immediacy code with deficiency sets S_1, S_2, \dots, S_m , provided D satisfies

$$D \geq \sum_{i=1}^m \sum_{B \in S_i} |B|. \quad (40)$$

Now, the latter quantity can be bounded from above as

$$\frac{n}{2} \sum_{i=1}^m \frac{\ell_{i+1}}{\ell_{m+1}} \leq n, \quad (41)$$

using the growth condition on the ℓ_i . Thus the choice $D = n$ works. \square

Applying Theorem 9, we thus get that the output alphabet size of any binary tree code satisfying the CHS immediacy condition with transmission length n as above satisfies

$$\lg |\Sigma| \geq \frac{m-1}{4} = \Omega(\lg \lg n), \quad (42)$$

since $n = 2^{10} \cdot 32^{2^{m+1}}$. Thus, the rate ρ of such a code must satisfy $\rho = O(1/\lg \lg n)$.

5.1.1 Subexponential time constructions of good tree codes

Braverman [Bra12] presented a generic procedure to explicitly construct good tree codes in sub-exponential time. More formally, given $\epsilon \in (0, 1)$, Braverman's procedure runs in time $\exp(O(n^\epsilon))$ and outputs a depth n tree code that has constant rate (depending upon ϵ) and fixed constant distance. At a high level, Braverman's construction works by starting with a brute force $\exp(O(n^\epsilon))$ -time construction of a tree code of depth n^ϵ , and then using a sequence of (efficiently implementable) tilings of this intermediate tree code, augmented with error-correcting codes, in order to handle error-propagation at different length scales. However, as described in [Bra12], this construction achieves a final alphabet size that is at least exponential in $1/\epsilon$, so that the rate achieved is no better than $O(1/\epsilon)$.

We first review here a refinement of the above construction, mentioned in passing in [CHS18, remarks following Theorem 1.1], that constructs a constant distance tree code with rate $\Theta(1/\log(1/\epsilon))$ and depth n in time $\exp(O(n^\epsilon))$. We then describe the immediacy properties of this construction, and show that no code with these immediacy properties can have rate better than $\Omega(1/\log(1/\epsilon))$. The arguments are very similar to those carried out above for the immediacy properties of the CHS construction.

Review of a subexponential-time constructible good tree code We use the notation from [CHS18] imported above. Let k be a fixed positive integer parameter (corresponding roughly to $\log(1/\epsilon)$). Consider an n of the form $\ell_{m+1} = 2^{10} \cdot 32^{2^{m+1}}$ as above. The polynomial time (in n) construction of a tree code of depth n given in the proof of Theorem 1.1 of [CHS18] consists principally of a concatenation of constant rate intermediate codes C_i , for each $1 \leq i \leq m+1$, where the i th code corresponds to length scale $\ell_i = 2^{10} \cdot 32^{2^i}$. In consequence, the resulting final

code has rate $\Theta(1/\lg \lg n)$.

To get a better rate at the cost of an increase in the runtime of the construction, start by replacing the above construction by a concatenation of only the codes $C_{m-k+1}, C_{m-k+2}, \dots, C_{m+1}$. Following the proof of Theorem 1.1 of [CHS18], one then gets that this resulting intermediate code is, in the terminology of [CHS18], a *lagged tree code* (i.e., where error propagation from an error starts after a certain lag) with lag $16\sqrt{\ell_{m-k+1}} = \ell_{m-k}/2$. Applying a variant of the idea from [Bra12], one can then remove the lag as follows. First construct, in time $\exp(O(\ell_{m-k}))$, a constant rate and constant distance tree code of depth ℓ_{m-k+1} . Then, concatenate to the above intermediate construction two tilings of this good tree code staggered by a shift of $\ell_{m-k+1}/2$ (following, e.g., Lemma 5 of [Bra12]). The resulting construction is a tree code of constant distance and rate $\Theta(1)/k$, obtained in time $\exp(O(\ell_{m-k+1})) = \exp(O(n^{1/2^k}))$. Reparameterizing with $\epsilon = 1/2^k$, we see that the rate is $\Theta(1)/\lg(1/\epsilon)$ and the runtime of the construction is $\exp(O(n^\epsilon))$.

Immediacy properties We now consider the immediacy properties of the the above construction. The arguments before Lemma 11, when applied to above variant of the CHS construction, show that the immediacy conditions described in eqs. (34) to (36) hold for this version for all i satisfying $m - k + 2 \leq i \leq m + 1$. (For the original CHS construction, these conditions held for all i such that $2 \leq i \leq m + 1$.)

Rate upper bound We now show that any tree code satisfying these immediacy properties cannot have rate better than $O(1/k)$. To do this, we repeat the proof of Lemma 11 and Proposition 12 with the above modified version of the CHS immediacy condition to obtain that any such code must be an n -deficient $(1/4, k)$ -immediacy code with respect to the $(1/4, k)$ -laminar partition $(P_{m-k}, P_{m-k+1}, \dots, P_m)$ (where the P_i are as defined in the above description of the original CHS construction, following eq. (33)). Applying Theorem 9, we therefore get that the output alphabet size of any depth n binary tree code satisfying the above variant of the CHS immediacy condition satisfies

$$\lg |\Sigma| \geq \frac{k-1}{4}. \quad (43)$$

Thus, the rate ρ of such a code must satisfy $\rho = O(1/k)$.

5.2 The EKS construction

We now turn to the construction of [EKS94] (which for brevity we shall call the EKS construction.) A version of this construction can be described as follows. Let $\text{ECC}_\ell : \{0, 1\}^\ell \rightarrow \left\{ \{0, 1\}^b \right\}^\ell$ denote an error-correcting block code with distance $\delta \in (0, 1)$. (It is known that there is a choice of $b = b(\delta)$, independent of the block-length ℓ , so that such a code exists for each positive integer

ℓ , and is such that ECC_ℓ can be computed in time polynomial in ℓ : see, e.g., [CHS18, Lemma 3.2 in the ECCC version].) Note that ECC_1 can be chosen to be the repetition code of length b . For simplicity, we describe the construction of the EKS tree code T when the depth $n = 2^k$ is a power of 2. For $0 \leq i \leq k$, let P_i be the partition of $[n]$ consisting of $n/2^i$ consecutive disjoint blocks of length 2^i each, and we write $P_{i,j} = \{(j-1)2^i + 1, \dots, j \cdot 2^i\}$. For a message $x \in \{0, 1\}^n$ we let $x_{P_{i,j}} \in \{0, 1\}^{2^i}$ denote the restriction of x to the interval $P_{i,j}$. To describe the encoding $T(x)$ of a message $x \in \{0, 1\}^n$, we form the following $(k+1) \times n$ table with entries in $\{0, 1\}^b$:

1. The first row consists of the concatenation of $\text{ECC}_1(x_j)$ for $j = 1, \dots, n$.
2. The i th row, for $2 \leq i \leq k+1$, consists of $(0^b)^{2^{i-2}}$ followed by the concatenation of $\text{ECC}_{2^{i-2}}(x_{P_{i-2,j}})$ for $j = 1, \dots, n \cdot 2^{2-i} - 1$. Informally, the code blocks in the i th row are “right-shifted” by 2^{i-2} .

The encoding $T(x)$ is defined by letting its j th character, for $1 \leq j \leq n$, be the j th column (from the left) of this table. Note that because of the “right shift” operation in item 2, this construction satisfies the “online” requirement for tree codes. The output alphabet of T is $\{0, 1\}^{b(k+1)}$.

The EKS immediacy condition From the construction, it is easy to see that the EKS construction satisfies the following version of immediacy. Denote the code by T , and let $x, y \in \{0, 1\}^n$ be two strings that differ at a position $s' \in [n]$. Let $0 \leq \ell < k$ be any integer such that $s' \leq n - 2^\ell = 2^k - 2^\ell$. Define $s = s(\ell) := \lceil s'/2^\ell \rceil \cdot 2^\ell$ to be the smallest multiple of 2^ℓ that is at least as large as s' . Then

$$\Delta_{\text{Hamming}}(T(x)_{(s, s+2^\ell)}, T(y)_{(s, s+2^\ell)}) \geq \delta \cdot 2^\ell > 0. \quad (44)$$

We now show that any code satisfying the immediacy condition described by eq. (44) must be an $(\Omega(1), \Omega(\log n))$ -immediacy code. To see this, we consider the partitions P_i , $0 \leq i \leq k$, of $[n] = [2^k]$ considered above, and convert each of them into a tagged partition by defining, for each $B \in P_i$ with $i \geq 1$, $\text{left}(B)$ to be the leftmost 2^{i-1} entries in B and $\text{right}(B)$ to be the rightmost 2^{i-1} entries in B . With this construction, it is immediate that $(P_0, P_1, P_2, \dots, P_k)$ is a $(1/2, k)$ -laminar partition. Since the code T satisfies the immediacy condition in eq. (44), it is also immediate that it satisfies the neighborhood decoding property for each $B \in P_i$ with $i \geq 1$. We thus see that any tree code T with depth $n = 2^k$ and satisfying the immediacy condition given by eq. (44) must be a $(1/2, k)$ -immediacy code. Applying Theorem 8, we thus see that the output alphabet Σ of such a code must satisfy

$$|\Sigma| \geq \frac{k}{2} = \Omega(\lg n), \quad (45)$$

since $n = 2^k$. Thus the rate ρ of such a code must be $O(1/\lg n)$.

5.3 The GHKRZW construction

We now turn to a tree code construction by Gelles, Haeupler, Kol, Ron-Zewi and Wigderson [GHKRZW16] (which we shall refer to as the GHKRZW construction). While related to the EKS construction discussed in Section 5.2, this construction is qualitatively different from the previous constructions considered in this paper in that it insists upon a *constant*, non-vanishing rate even if at the cost of a vanishing *distance* parameter. We show here that for the immediacy guarantee the GHKRZW construction achieves, and given its requirement of a constant rate, the $\Omega(1/\log n)$ vanishing distance it achieves is tight up to a $\log \log n$ factor.

The GHKRZW construction (described in [GHKRZW16, Section 5.1]) demonstrates the following: there exists a positive integer k_0 , such that for every $\epsilon \in (0, 1)$ there is a tree code T , with transmission length n , input alphabet $\Sigma_{\text{in}} := \{0, 1\}^{(\lg n)/\epsilon}$, output alphabet $\Sigma := \{0, 1\}^{(1+1/\epsilon) \cdot \lg n}$ (so that its rate is $1/(1 + \epsilon)$) and distance parameter δ is $\frac{\epsilon}{32k_0 \cdot \lg n}$. Here, it is convenient to assume (as we do now) that $\lg n$, k_0 and $1/\epsilon$ are all powers of two, and that $k_0 \geq 16$. (The notation and conventions here are essentially exactly those adopted in [GHKRZW16, Section 5.1]).

The GHKRZW immediacy condition The distance proof for the GHKRZW construction, given in Section 5.2 of [GHKRZW16], in fact establishes the following immediacy property. Let k_0, ϵ be parameters not growing with the transmission length n (as above) and let $m := (k_0/\epsilon) \cdot \lg n$ be as defined in [GHKRZW16]; note that m is also a power of two by the assumptions adopted above. Let t be any integer satisfying $\lg m \leq t \leq \lg n - 1$. Suppose that $x, y \in \Sigma_{\text{in}}$ differ at a position $i \leq n - 2^t$, and let $i_0 = i_0(t) := \lfloor (i - 1)2^{-t} \rfloor \cdot 2^t$ be the largest multiple of 2^t that is smaller than i , as defined in [GHKRZW16, Section 5.2]. Then,

$$\Delta_{\text{Hamming}}(T(x)_{(i_0, i_0+2^{t+1}]}, T(y)_{(i_0, i_0+2^{t+1}]}) \geq \delta \cdot 2^{t+1}. \quad (46)$$

To prove this, one repeats the proof in Section 5.2 of [GHKRZW16], ignoring the parameter j in that proof, and also ignoring the condition, stipulated in the proof, that i is the *first* position on which x and y differ. The last paragraph on p. 1930 in [GHKRZW16], combined with the first display on p. 1931 (neither of which use the latter condition that i is the first position on which x and y differ), then yields (a slightly stronger version of) the above condition, after we note that δ' in that proof is at least 8δ , where δ , as above, is the distance parameter achieved by the GHKRZW construction.⁴

We now show that any constant rate tree code satisfying the immediacy condition specified by eq. (46) cannot have distance much better than the GHKRZW construction. The proof is similar to

⁴For the convenience of the reader, we note here that there is a minor typographical error on p. 1930 of [GHKRZW16], where the “max” in the first display in Section 5.2 should actually be “min”.

the case of an exponential immediacy function analyzed in Section 3.2.

As in that proof, our goal is to show that such a code is an immediacy code with appropriate parameters. We begin by defining the corresponding laminar partition. Let $\delta \in (0, 1)$ be the distance parameter (possibly dependent upon n) of such a code. Let $\kappa := \lfloor \lg(2/\delta) \rfloor$ be a positive integer so that $2^{-\kappa} < \delta \leq 2^{-\kappa+1}$. Set

$$\ell := 1 + \lfloor (\lg(n/(2m)))/\kappa \rfloor. \quad (47)$$

Let P_0 be the partition of $[n]$ into n singletons, and for $1 \leq i \leq \ell$, define P_i to be the partition of $[n]$ into blocks of consecutive integers, each of length $n/2^{\kappa \cdot (\ell-i)}$. Thus, for $1 \leq i \leq \ell$, any set B in P_i is an interval of the form $(b \cdot 2^{t+1}, (b+1) \cdot 2^{t+1}]$, where b is an integer and $t = \lg n - \kappa \cdot (\ell - i) - 1$. For $1 \leq i \leq \ell$, convert P_i into a tagged partition by defining, for each $B \in P_i$, $\text{left}(B)$ to be the set of the smallest $2^{-\kappa} \cdot |B|$ elements in B (so that $\text{right}(B) = B \setminus \text{left}(B)$). It is easy to verify that the above construction gives a $(2^{-\kappa}, \ell)$ -laminar partition of $[n]$.

We now show that any code satisfying the immediacy condition of eq. (46) must be a $(2^{-\kappa}, \ell)$ -immediacy code. We use the laminar partition defined above. Fix $1 \leq j \leq \ell$ and $B \in P_j$. Let $t = \lg n - \kappa \cdot (\ell - j) - 1$ be as above (note that t satisfies $\lg m \leq t \leq \lg n - 1$), so that $B = (b \cdot 2^{t+1}, (b+1) \cdot 2^{t+1}]$ for some non-negative integer b . Consider any $x, y \in \Sigma_{\text{in}}^n$ that differ on $\text{left}(B) = (b \cdot 2^{t+1}, (b+2^{-\kappa}) \cdot 2^{t+1}]$. Thus, there exists $i \in \text{left}(B) \subseteq (b \cdot 2^{t+1}, (b+1/2) \cdot 2^{t+1}]$ such that $x_i \neq y_i$. The immediacy condition of eq. (46) then gives that

$$\Delta_{\text{Hamming}}(T(x)_B, T(y)_B) \geq \delta \cdot 2^{t+1} > 2^{-\kappa} \cdot 2^{t+1} = |\text{left}(B)|. \quad (48)$$

Here the second inequality uses $\delta > 2^{-\kappa}$. Thus, we see that the Hamming distance between $T(x)_B$ and $T(y)_B$ is strictly more than the size of $\text{left}(B)$ whenever $x_{\text{left}(B)}$ and $y_{\text{left}(B)}$ differ. This implies that $T(x)_{\text{right}(B)}$ and $T(y)_{\text{right}(B)}$ must differ whenever $x_{\text{left}(B)}$ and $y_{\text{left}(B)}$ differ. This establishes that T is a $(2^{-\kappa}, \ell)$ -immediacy code.

We now apply Theorem 8 to see that a tree code $T : \Sigma_{\text{in}}^n \rightarrow \Sigma^n$ satisfying the immediacy condition of eq. (46) must therefore satisfy

$$\frac{\lg |\Sigma|}{\lg |\Sigma_{\text{in}}|} \geq 2^{-\kappa} \cdot \ell \stackrel{\text{eq. (47)}}{\geq} \frac{\delta \cdot \lg \frac{n}{2m}}{2 \cdot \lg(2/\delta)}. \quad (49)$$

Recall that $m = (k_0/\epsilon) \cdot \lg n$, where k_0, ϵ do not grow with n . Thus, eq. (49) shows that any constant rate tree code satisfying the immediacy condition eq. (46) of the GHKRZW construction must have a distance parameter δ satisfying

$$\frac{\delta}{\log(1/\delta)} = O(1/\log n). \quad (50)$$

In particular, a code with the immediacy guarantee that the GHKRZW construction achieves cannot achieve a distance parameter δ that is $\omega\left(\frac{\log \log n}{\log n}\right)$.

References

- [BH20] S. Bhandari and P. Harsha. A note on the explicit constructions of tree codes over polylogarithmic-sized alphabet, February 2020. [arXiv:2002.08231](#).
- [BK12] Z. Brakerski and Y. T. Kalai. Efficient interactive coding against adversarial noise. In *Proc. IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 160–166, 2012. DOI: [10.1109/FOCS.2012.22](#).
- [BKS20] Z. Brakerski, Y. T. Kalai, and R. R. Saxena. Deterministic and efficient interactive coding from hard-to-decode tree codes. In *Proc. IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 446–457, 2020. DOI: [10.1109/FOCS46700.2020.00049](#).
- [BN13] Z. Brakerski and M. Naor. Fast algorithms for interactive coding. In *Proc. ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 443–456, 2013. DOI: [10.1137/1.9781611973105.32](#).
- [Bra12] M. Braverman. Towards deterministic tree code constructions. In *Proc. Innovations in Theoretical Computer Science Conference (ITCS)*, pages 161–167, 2012. DOI: [10.1145/2090236.2090250](#).
- [BYCN21] I. Ben Yaacov, G. Cohen, and A. K. Narayanan. Candidate tree codes via Pascal determinant cubes. In *Proc. Approximation, Randomization, and Combinatorial Optimization, APPROX/RANDOM*, volume 207 of *LIPIcs*, pages 54:1–54:22, 2021. DOI: [10.4230/LIPICS.APPROX/RANDOM.2021.54](#).
- [BYCY22] I. Ben Yaacov, G. Cohen, and T. Yankovitz. Explicit binary tree codes with sub-logarithmic size alphabet. In *Proc. ACM Symposium on Theory of Computing (STOC)*, STOC, page 595–608, 2022. DOI: [10.1145/3519935.3520033](#).
- [CHS18] G. Cohen, B. Haeupler, and L. J. Schulman. Explicit binary tree codes with polylogarithmic size alphabet. In *Proc. ACM Symposium on Theory of Computing (STOC)*, pages 535–544, 2018. DOI: [10.1145/3188745.3188928](#). Full version available on ECCO at <https://eccc.weizmann.ac.il/report/2018/032/>.

- [CS20] G. Cohen and S. Samocha. Palette-alternating tree codes. In *Proc. Computational Complexity Conference (CCC)*, volume 169 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 11:1–11:29, 2020. DOI: [10.4230/LIPIcs.CCC.2020.11](https://doi.org/10.4230/LIPIcs.CCC.2020.11).
- [CT06] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, second edition, 2006.
- [DGL23] M. Dall’Agnol, T. Gur, and O. Lachish. A structural theorem for local algorithms with applications to coding, testing, and verification. *SIAM J. Comput.*, 52(6):1413–1463, 2023. DOI: [10.1137/21M1422781](https://doi.org/10.1137/21M1422781).
- [EKS94] W. Evans, M. Klugerman, and L. J. Schulman. Postscript to “Coding for interactive communication”. URL: <http://users.cms.caltech.edu/~schulman/Papers/intercodingpostscript.txt>. Unpublished work, 1994.
- [Gel17] R. Gelles. Coding for interactive communication: A survey. *Foundations and Trends in Theoretical Computer Science*, 13(1-2):1–157, 2017. DOI: [10.1561/0400000079](https://doi.org/10.1561/0400000079).
- [GH14] M. Ghaffari and B. Haeupler. Optimal Error Rates for Interactive Coding II: Efficiency and List Decoding. In *Proc. IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 394–403, 2014. DOI: [10.1109/FOCS.2014.49](https://doi.org/10.1109/FOCS.2014.49).
- [GHKRZW16] R. Gelles, B. Haeupler, G. Kol, N. Ron-Zewi, and A. Wigderson. Towards optimal deterministic coding for interactive communication. In *Proc. Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1922–1936, 2016. DOI: [10.1137/1.9781611974331.ch135](https://doi.org/10.1137/1.9781611974331.ch135).
- [GL21] T. Gur and O. Lachish. On the power of relaxed local decoding algorithms. *SIAM Journal on Computing*, 50(2):788–813, 2021. DOI: [10.1137/19M1307834](https://doi.org/10.1137/19M1307834).
- [GMS11] R. Gelles, A. Moitra, and A. Sahai. Efficient and explicit coding for interactive communication. In *Proc. IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 768–777, 2011. DOI: [10.1109/FOCS.2011.51](https://doi.org/10.1109/FOCS.2011.51).
- [Gol25] O. Goldreich. On the lower bound on the length of relaxed locally decodable codes. In *Computational Complexity and Local Algorithms : On the Interplay Between Randomness and Computation*, pages 151–192. Springer Nature Switzerland, 2025. DOI: [10.1007/978-3-031-88946-2_11](https://doi.org/10.1007/978-3-031-88946-2_11). Preliminary version: Electronic Colloquium on Computational Complexity (ECCC): [TR23-064](https://eccc.hawaii.edu/2023/TR23-064).
- [Ham50] R. W. Hamming. Error detecting and error correcting codes. *The Bell System Technical Journal*, 29(2):147–160, 1950. DOI: [10.1002/j.1538-7305.1950.tb00463.x](https://doi.org/10.1002/j.1538-7305.1950.tb00463.x).

- [Jus72] J. Justesen. Class of constructive asymptotically good algebraic codes. *IEEE Transactions on Information Theory*, 18(5):652–656, 1972. DOI: [10.1109/tit.1972.1054893](https://doi.org/10.1109/tit.1972.1054893).
- [KA10] S. Kamath and V. Anantharam. A new dual to the Gács-Körner common information defined via the Gray-Wyner system. In *Proc. Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1340–1346, September 2010. DOI: [10.1109/ALLERTON.2010.5707069](https://doi.org/10.1109/ALLERTON.2010.5707069).
- [KT00] J. Katz and L. Trevisan. On the efficiency of local decoding procedures for error-correcting codes. In *Proc. ACM Symposium on Theory of Computing (STOC)*, pages 80–86, May 2000. DOI: [10.1145/335305.335315](https://doi.org/10.1145/335305.335315).
- [MRR25] T. Moud, A. Rosen, and R. Rothblum. Locally testable tree codes. In *Proc. ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 5523–5559, 2025. DOI: [10.1137/1.9781611978322.189](https://doi.org/10.1137/1.9781611978322.189).
- [MS14] C. Moore and L. J. Schulman. Tree codes and a conjecture on exponential sums. In *Proc. Innovations in Theoretical Computer Science Conference (ITCS)*, pages 145–154, 2014. DOI: [10.1145/2554797.2554813](https://doi.org/10.1145/2554797.2554813).
- [NW20] A. K. Narayanan and M. Weidner. On decoding Cohen-Haeupler-Schulman tree codes. In S. Chawla, editor, *Proc. ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1337–1356, 2020. DOI: [10.1137/1.9781611975994.81](https://doi.org/10.1137/1.9781611975994.81).
- [Pud16] P. Pudlák. Linear tree codes and the problem of explicit constructions. *Linear Algebra and its Applications*, 490:124–144, 2016. DOI: [10.1016/j.laa.2015.10.030](https://doi.org/10.1016/j.laa.2015.10.030).
- [Sch92] L. J. Schulman. Communication on noisy channels: a coding theorem for computation. *Proc. IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 724–733, 1992. DOI: [10.1109/SFCS.1992.267778](https://doi.org/10.1109/SFCS.1992.267778).
- [Sch93] L. J. Schulman. Deterministic coding for interactive communication. In *Proc. ACM Symposium on Theory of Computing (STOC)*, pages 747–756, 1993. DOI: [10.1145/167088.167279](https://doi.org/10.1145/167088.167279).
- [Sch96] L. J. Schulman. Coding for interactive communication. *IEEE Transactions on Information Theory*, 42(6):1745–1756, 1996. DOI: [10.1109/18.556671](https://doi.org/10.1109/18.556671).
- [Sha48] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, July 1948. DOI: [10.1002/j.1538-7305.1948.tb01338.x](https://doi.org/10.1002/j.1538-7305.1948.tb01338.x).

- [SS96] M. Sipser and D. A. Spielman. Expander codes. *IEEE Transactions on Information Theory*, 42(6):1710–1722, 1996. DOI: [10.1109/18.556667](https://doi.org/10.1109/18.556667).
- [SS19] L. J. Schulman and P. Srivastava. Online codes for analog signals. *IEEE Transactions on Information Theory*, 65(10):6633–6649, 2019. DOI: [10.1109/TIT.2019.2919632](https://doi.org/10.1109/TIT.2019.2919632).
- [TVZ82] M. A. Tsfasman, S. G. Vladut, and T. Zink. Modular curves, Shimura curves, and Goppa codes, better than Varshamov-Gilbert bound. *Mathematische Nachrichten*, 109(1):21–28, 1982. DOI: [10.1002/mana.19821090103](https://doi.org/10.1002/mana.19821090103).
- [Yan25] T. Yankovitz. CHS-alike $1/O(\log \log n)$ -rate tree codes from elementary binary shifts. Electronic Colloquium on Computational Complexity (ECCC): TR25-167, November 2025. URL: <https://eccc.weizmann.ac.il/report/2025/167/>.