Electronic Colloquium on Computational Complexity, Report No. 46 (2025)

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 such codes has remained a notorious challenge. While the construction of asymptotically-good explicit tree codes continues to be elusive, 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 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 current 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 from recursive combinations of error-correcting codes. Our main technical contribution is the proof of a *rate–immediacy tradeoff*, 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 tradeoff to existing constructions, we establish that their known rate analyses are essentially optimal. 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 no. RTI4001, by SERB MATRICS grant number MTR/2023/001547, and by the Infosys-Chandrasekharan virtual center for Random Geometry.

ISSN 1433-8092

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 $C\colon \Sigma^k \to \Sigma^n$ is an *error-correcting code* with distance δ 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] 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—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 $\Omega(1/\log n)$.

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

1.1 Immediacy

Given the lack of progress in constructing tree codes over the past few decades—despite significant efforts—it is important to better understand why current techniques fall short. Identifying a barrier could help steer research away from potential dead ends. More specifically, since tree codes induce standard block error-correcting codes, and these codes 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 can 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, which we refer to as *immediacy*. This stronger property arises inherently due to the recursive nature of existing constructions [EKS94, CHS18, NW20, BYCY22]. While immediacy might be desirable in certain contexts, we prove that it incurs a significant cost in rate. Specifically, we establish a quantitative *rate-immediacy tradeoff*, 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 framework typically employed—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 Section 3). However, to capture the essence, we first provide an informal, simplified definition in which immediacy is represented by a monotone-increasing function $\mathrm{Imm} : \mathbb{N} \to \mathbb{N}$. A tree code C is said to have immediacy Imm if, for every pair of messages x, y, 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 $[s, s + \mathrm{Imm}(k))$ (provided that $s + \mathrm{Imm}(k)$ is no larger than the depth of T), is bounded below by δ.

Note that 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, immediacy 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, during any interval indicated by the immediacy function, provided the adversary does not excessively interfere with transmissions in this interval, the receiver can decode the s-th bit, even if earlier parts of the message cannot yet be reconstructed. While this property may be advantageous in certain applications, our results show that it is an "overly expensive" feature that cannot be afforded when aiming to construct asymptotically good tree codes.

By inspection, it can be shown 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 $\mathrm{Imm}(k) = 2^{\Theta(k)}$, which we denote throughout this informal presentation as $\exp(k)$.[2] At this informal stage, it is convenient to regard Imm as a real-valued function—which in this case is $\mathrm{Imm}(x) = \exp(x)$—thus extending the original discrete definition. With this, our main result can be informally stated as follows (see also the remarks below for some subtleties that are omitted here for simplicity):

Theorem 1 (Main result; informal)**.** *Let T be a depth-n tree code with constant distance and immediacy function* Imm*. Then, the rate of T satisfies*

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

Recall that for the construction given in [EKS94], we have $\mathrm{Imm}(x) = \exp(x)$ and so $\mathrm{Imm}^{-1}(x) = \Theta(\log x)$. Thus, Theorem 1 implies that the rate of this construction cannot exceed $O(1/\log n)$, thereby explaining its known rate. Theorem 1 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 $\mathrm{Imm}(x) = \exp(\exp(x))$.[3] Thus, Theorem 1 implies that the rate of these constructions cannot exceed $O(1/\log\log n)$, matching the proven rate of these codes. More generally, immediacy should be kept in mind when constructing tree codes, and Theorem 1 can be applied to any proposed construction to obtain a quantitative upper bound on the achievable rate. 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.

Subtleties omitted from the informal discussion. The precise formulation of our main result, presented in Theorem 4, builds on a more intricate notion that we refer to as *immediacy codes* (see Definition 4). The connection between this formal definition and the simplified notion of immediacy functions introduced earlier in this section is clarified in Section 3.1. While the

[2] The precise immediacy property enjoyed by the [EKS94] construction is slightly different, and starts only after a "lag": see Section 5. However, it is helpful to temporarily ignore this technical issue in order to understand the main ideas.

[3] Again, the precise immediacy property enjoyed by the [CHS18] construction is slightly different, and starts only after a "lag": see Section 5. Again, it is helpful to temporarily ignore this technical issue.

core idea of our main result is captured by Theorem 1, there are important subtleties to note. For instance, it turns out that exponential immediacy (i.e. $\mathrm{Imm}(k) = \exp(k)$ as described above) implies *linear* immediacy. Consequently, Theorem 1 should be interpreted as applying only to exponential (or faster-growing) immediacy.

Rate–distance–immediacy tradeoff. We point out that Theorem 1 also applies to codes with vanishing distance. In fact, for a general relative distance $\delta > 0$, the upper bound on the rate ρ stated in Theorem 1 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 4, and to the computations leading to eqs. (8) and (9). The key takeaway, however, is that Theorem 1, in its full generality, yields a *rate–distance–immediacy tradeoff*. For ease of presentation, we choose to focus on the tradeoff between rate and immediacy in this introductory section (see Section 1.2 for further discussion).

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 tradeoff between rate and immediacy, and offer computationally efficient encoding and decoding. It was therefore a very considerable accomplishment of Brakerski, Kalai, and Saxena [BKS20] to show that, using only constant-rate explicit tree codes, all encoding and decoding tasks of the parties can be performed in polynomial time. The method uses a rotating cast of tree codes in order to simulate some of the properties which one could get very simply, if one were using a code with an immediacy property. Our result thus “in hindsight” explains why the authors had to resort to a complex encoding protocol, rather than simply running their desired communication pattern over an underlying, tree-code-equipped channel.

Tree code variants. We conclude this section by noting that the interactive coding theorem can also be established using a weaker variant of tree codes known as *potent tree codes* which was introduced by Gelles, Moitra, and Sahai [GMS11]. 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].

1.2 Technical overview and organization of the paper

We follow the time-honoured 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 limitation 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, Gol23]). 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 3.1), 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 3)[4] to quantify "common information" between different parts of a code word. Our main technical result is that *immediacy codes cannot have high rates* (Theorem 4): this is proved in Section 4. 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 this result is technically quite simple.

Finally, in Section 5, we apply our framework to study the immediacy properties of the tree code constructions proposed in [EKS94], [CHS18] and [GHKRZW16]. 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. The construction of [GHKRZW16] attempts to tackle a different tradeoff: 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 $\omega\left(\frac{\log\log n}{\log n}\right)$.

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

[4] This 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.

strategy. Nevertheless, a variant of our analysis for [CHS18] can still be readily adapted to study its immediacy properties. Lastly, we note that the construction of [GHKRZW16] is similar to that of [EKS94] with respect to the code-tiling component. As such, our result—when viewed not as a rate–immediacy tradeoff but rather as a distance–immediacy tradeoff—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 $\{1, 2, \ldots, \ell\} =: [\ell]$. 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). Given a tuple $(A_1, A_2, \ldots, A_k)$ of random variables, $H((A_1, A_2, \ldots, A_k))$ is often written as $H(A_1, A_2, \ldots, A_k)$ (and even just as $H(A_1A_2 \ldots A_k)$ when there is no risk of confusion with multiplication). Given two random variables X and Y, both taking values in (possibly different) finite sets, the *conditional entropy* $H(X|Y)$ is defined as

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

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

$$I(X : Y) = H(X) - H(X|Y). \tag{2}$$

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

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

We collect below some of the standard properties of these quantities. The proofs of these properties are easy, and can be found, e.g., in [CT06].

Proposition 2 (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)* $0 \le H(X) \le \lg |\Omega_1|$.
- *(Non-negativity of the mutual information)* $I(X : Y) \ge 0$ *and* $I(X : Y|Z) \ge 0$.
- *(Conditioning reduces entropy)* $H(X|Y, Z) \le H(X|Y) \le H(X)$.
- *(Chain rule for entropy)* $H(X, Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$.
- *(Sub-additivity of entropy)* $H(X, Y) \le H(X) + H(Y)$.
- *(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 $f : \Omega_2 \to \Omega_1$ such that $f(Y) = X$, then $H(X|Y) = 0$.*

In this paper, we will often use these properties without much comment.

3 Immediacy codes

In this section, we first introduce the notion of *immediacy codes* (Definition 4), a formal abstraction of the informal notion of immediacy discussed in the introduction. To connect the two notions, we show in Section 3.1 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 4): a general rate upper bound for immediacy codes.

We begin with a few combinatorial definitions that go into the definition of immediacy codes. 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 2 (Tagged partitions). A *tagged partition* S of $[n]$ is a partition $(S_1, S_2, \ldots, S_k)$ of $[n]$, along with an ordered partition $(\mathrm{left}(S_i), \mathrm{right}(S_i))$ of each S_i, $1 \le i \le k$, into two non-empty sets.

Definition 3 (Laminar partitions). An (α, ℓ)*-laminar partition* of n is a tuple $P = (P_0, P_1, P_2, \ldots, P_\ell)$ of partitions of $[n]$ into disjoint non-nonempty subsets, in which $P_1, \ldots, P_\ell$ are tagged partitions, and which satisfies the following properties:

1. **(Size property)** For each $1 \le i \le \ell$, and for each $B \in P_i$, $|\mathrm{left}(B)| \ge \alpha\, |B|$. (Here, $\mathrm{left}(\cdot)$ 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 $\mathrm{left}(B)$ and $\mathrm{right}(B)$, respectively, into pairwise disjoint non-empty sets (see fig. 1 for an illustration).

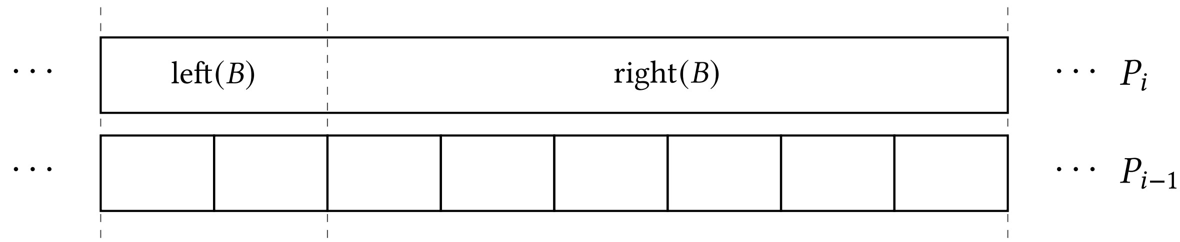

Figure 1: An illustration of laminar partitions

Definition 4 (Immediacy codes). An (α, ℓ)-*immediacy code* with code alphabet Σ and transmission length n is a function $c : \Sigma_{\mathrm{in}}^n \to \Sigma^n$ along with an (α, ℓ)-laminar partition $P = (P_0, P_1, P_2, \ldots, P_\ell)$ 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 $x, y \in \Sigma_{\mathrm{in}}^n$, $x_{\mathrm{left}(B)} \neq y_{\mathrm{left}(B)}$ implies that $c(x)_{\mathrm{right}(B)} \neq c(y)_{\mathrm{right}(B)}$. In other words, there is a function $\phi_B : \Sigma^{|\mathrm{right}(B)|} \to \Sigma_{\mathrm{in}}^{|\mathrm{left}(B)|}$ such that $\phi_B(c(x)_{\mathrm{right}(B)}) = x_{\mathrm{left}(B)}$ for all $x \in \Sigma_{\mathrm{in}}^n$. (Here, $\mathrm{left}(B)$ and $\mathrm{right}(B)$ are as in the definition of a tagged partition.)

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 5 (Deficient Immediacy code). A D-deficient (α, ℓ)-immediacy code with code alphabet Σ and transmission length n is a function $c : \Sigma_{\mathrm{in}}^n \to \Sigma^n$ along with an (α, ℓ)-laminar partition $P = (P_0, P_1, P_2, \ldots, P_\ell)$ 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,

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

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 $x, y \in \Sigma_{\mathrm{in}}^n$, $x_{\mathrm{left}(B)} \neq y_{\mathrm{left}(B)}$ implies that $c(x)_{\mathrm{right}(B)} \neq c(y)_{\mathrm{right}(B)}$. In other words, there is a function $\phi_B : \Sigma^{|\mathrm{right}(B)|} \to \Sigma_{\mathrm{in}}^{|\mathrm{left}(B)|}$ such that $\phi_B(c(x)_{\mathrm{right}(B)}) = x_{\mathrm{left}(B)}$ for all $x \in \Sigma_{\mathrm{in}}^n$. (Here, $\mathrm{left}(B)$ and $\mathrm{right}(B)$ are as in the definition of a tagged partition.)

3.1 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 $T : \{0,1\}^n \to \Sigma^n$ be a tree code with immediacy function Imm, distance parameter $\delta \in (0,1)$, and suppose that its depth $n = 2 \cdot \mathrm{Imm}(\ell t)$, 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 $2^{-\kappa} < \delta \le 2^{-\kappa+1}$. We assume that t can be chosen so that for each $1 \le j \le \ell$,

$$\begin{gathered} 2^{-\kappa} \cdot \mathrm{Imm}(jt) \text{ is an integer, and} \\ 2 \cdot \mathrm{Imm}((j-1)t) \text{ divides } 2^{-\kappa} \cdot \mathrm{Imm}(jt) \end{gathered} \tag{5}$$

(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. (5), we naturally obtain an (α, ℓ) laminar partition $(P_0, P_1, P_2, \ldots, P_\ell)$. For each $0 \le j \le \ell$, partition $[n] = [2\mathrm{Imm}(\ell t)]$ into consecutive blocks of length $2\mathrm{Imm}(jt)$ each, and let these blocks constitute the tagged partition P_j. Given a block $B \in P_j$ for $j \ge 1$, we define $\mathrm{left}(B)$ to be the set of the leftmost $2^{-\kappa} \cdot \mathrm{Imm}(jt)$ integers in B and $\mathrm{right}(B)$ to be the rightmost $2\mathrm{Imm}(jt) - |\mathrm{left}(B)|$ integers in B. Given the divisibility conditions in eq. (5), 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, $\mathrm{left}(B)$ and $\mathrm{right}(B)$ 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 $(P_0, P_1, \ldots, P_\ell)$. Fix any $1 \le j \le \ell$ 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 $\mathrm{left}(B)$, then $T(x)$ and $T(y)$ differ on some index in $\mathrm{right}(B)$. 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 $[i, i + \mathrm{Imm}(jt))$ of length $\mathrm{Imm}(jt)$. By the choice of x and y, we have $i \in \mathrm{left}(B)$. Since $|\mathrm{left}(B)| = \mathrm{Imm}(jt) \cdot 2^{-\kappa}$, $|B| = 2\mathrm{Imm}(jt)$, and $\kappa \ge 1$, it then follows that S is contained in B, and intersects $\mathrm{right}(B)$. Further, by the definition of the immediacy function, the Hamming distance between $T(x)_S$ and $T(y)_S$ must be at least

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

Thus, since this Hamming distance is greater than the length of left(B), it must be the case that $T(x)_S$ and $T(y)_S$ differ also on some index in 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.

Anticipating the rate upper bound for (α, ℓ)-immediacy codes proved below in our main result (Theorem 4), we thus get that if Σ is the output alphabet of T, it must be the case that (recall that $n = 2 \cdot \mathrm{Imm}(\ell t(\delta))$)

$$\lg |\Sigma| \geq \frac{\ell}{2^{1+\kappa}} = \frac{\mathrm{Imm}^{-1}(n/2)}{t(\delta) \cdot 2^{1+\kappa}} \geq \frac{\delta}{4t(\delta)}\mathrm{Imm}^{-1}(n/2). \tag{6}$$

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

$$\rho \leq \frac{4t(\delta)}{\delta \cdot \mathrm{Imm}^{-1}(n/2)}. \tag{7}$$

(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. (5), for the two immediacy functions considered in the introduction.

Case 1: $\mathrm{Imm}(k) = 2^k$**.** Put $t(\delta) = 1+\kappa = \lfloor \lg(4/\delta) \rfloor$. A direct computation shows that the condition in eq. (5) holds with this choice. The rate upper bound in eq. (7) becomes

$$\rho \leq \frac{4 \lg(4/\delta)}{\delta \cdot \lg(n/2)}. \tag{8}$$

Case 2: $\mathrm{Imm}(k) = 2^{2^k}$**.** 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. (5) holds with this choice. The rate upper bound in eq. (7) becomes

$$\rho \leq \frac{4 \cdot \lceil \lg \lg(8/\delta) \rceil}{\delta \cdot \lg \lg(n/2)}. \tag{9}$$

4 Rate upper-bound for immediacy codes

In this section, we prove our main result: a rate upper bound for immediacy codes (Theorem 4). 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 3 (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

$$0 \leq I(A : C|B) \leq H(A|B) = 0. \tag{10}$$

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

$$\begin{aligned} H(A) &= I(A : C) \leq I(A : C) + I(B : C|A) \\ &= I(A, B : C) \overset{\text{eq. (10)}}{=} I(B : C) + I(A : C|B) = I(B : C), \end{aligned}$$

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

4.2 Statement and proof of the main result

The following is our main result, formalizing Theorem 1.

Theorem 4 (Rate upper bound for immediacy codes). *Let n be a positive integer, and suppose that $c : \Sigma_{\text{in}}^n \to \Sigma^n$ is an $(\alpha(n), \ell(n))$-immediacy code . Then,*

$$\lg |\Sigma| \geq \alpha(n)\ell(n) \lg |\Sigma_{\text{in}}| . \tag{11}$$

Proof. To simplify notation, we shorten $\ell(n)$ to ℓ in the following. Let $P = (P_0, P_1, \ldots, P_\ell)$ 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 $1 \le k \le n$: 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 $i \ge 1$. Since c is systematic, we than 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

$$H(X_{\text{left}(B)} | Y_{\text{right}(B)}) = H(\phi_B(c(X)_{\text{right}(B)}) | c(X)_{\text{right}(B)}) = 0. \tag{12}$$

(Here, ϕ_B is as in the definition of the neighborhood decoding property of an immediacy code.) Applying Lemma 3, we then have

$$\begin{aligned} H(Y_B) = H(Y_{\text{left}(B)}, Y_{\text{right}(B)}) &\le H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)}) - H(X_{\text{left}(B)}) && (13)\\ &= H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)}) - |\text{left}(B)| \cdot \lg |\Sigma_{\text{in}}| && (14)\\ &\le H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)}) - \alpha(n)|B| \cdot \lg |\Sigma_{\text{in}}|, && (15) \end{aligned}$$

where the last inequality follows since $\text{left}(B) \ge \alpha(n)|B|$, by the size property of $(\alpha(n), \ell(n))$-laminar partitions. Note that eq. (15) holds for every subset B that is part of some partition P_i, $i \ge 1$, of P. Define, for $0 \le i \le \ell$

$$T_i := \sum_{B \in P_i} H(Y_B). \tag{16}$$

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

$$\begin{aligned} T_i = \sum_{B \in P_i} H(Y_B) &\overset{\text{eq. (15)}}{\le} -\alpha(n)n \cdot \lg |\Sigma_{\text{in}}| + \sum_{B \in P_i} \left(H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)}) \right) && (17)\\ &\le -\alpha(n)n \cdot \lg |\Sigma_{\text{in}}| + \sum_{B' \in P_{i-1}} H(Y_{B'}) = T_{i-1} - \alpha(n)n \cdot \lg |\Sigma_{\text{in}}|. && (18) \end{aligned}$$

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. (18) thus gives

$$T_\ell \leq T_0 - \alpha(n) n\ell \cdot \lg |\Sigma_{\text{in}}| \,. \tag{19}$$

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

$$T_0 \leq \sum_{i=1}^{n} H(Y_i) \leq n \cdot (\lg |\Sigma_{\text{in}}| + \lg |\Sigma|), \text{ since each } Y_i \text{ has support } \Sigma' = \Sigma \times \Sigma_{\text{in}}. \tag{20}$$

Substituting these in eq. (19) gives $\lg |\Sigma| \geq \alpha(n)\ell \cdot \lg |\Sigma_{\text{in}}|$. □

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.

Theorem 5 (Rate upper bound for deficient immediacy codes). *Let n be a positive integer, and suppose that $c : \Sigma_{\text{in}}^n \to \Sigma^n$ is a D-deficient $(\alpha(n), \ell(n))$-immediacy code . Then,*

$$\lg |\Sigma| \geq \alpha(n) \cdot \left(\ell(n) - \frac{D}{n}\right) \cdot \lg |\Sigma_{\text{in}}| \,. \tag{21}$$

Proof. The proof is a minor modification of the proof of Theorem 4, but we include all the details for completeness. Let $P = (P_0, P_1, \ldots, P_\ell)$ 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 5). 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 than 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

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

(Here, ϕ_B is as in the definition of the neighborhood property of a deficient immediacy code.)

Applying Lemma 3, we then have

$$
\begin{aligned}
H(Y_B) = H(Y_{\text{left}(B)}, Y_{\text{right}(B)}) &\le H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)}) - H(X_{\text{left}(B)}) && (23)\\
&= H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)}) - |\text{left}(B)| \cdot \lg|\Sigma_{\text{in}}| && (24)\\
&\le H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)}) - \alpha(n)|B| \cdot \lg|\Sigma_{\text{in}}|\,, && (25)
\end{aligned}
$$

where the last inequality follows since $\text{left}(B) \ge \alpha(n)|B|$, by the size property of $(\alpha(n), \ell(n))$-laminar partitions. Note that when $i \ge 1$, eq. (25) holds for every subset $B \in P_i \setminus S_i$. The only difference with the proof of Theorem 4 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):

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

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

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

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

$$
\begin{aligned}
T_i = \sum_{B\in P_i} H(Y_B) &\overset{\text{eqs. (25) and (26)}}{\le} -\alpha(n)n \cdot \lg|\Sigma_{\text{in}}| + \alpha(n)\cdot \lg|\Sigma_{\text{in}}| \cdot \sum_{B\in S_i}|B| \\
&\quad + \sum_{B\in P_i}\big(H(Y_{\text{left}(B)}) + H(Y_{\text{right}(B)})\big) && (28)\\
&\le -\alpha(n)n\cdot \lg|\Sigma_{\text{in}}| + \alpha(n)\cdot\lg|\Sigma_{\text{in}}| \cdot \sum_{B\in S_i}|B| + \sum_{B'\in P_{i-1}} H(Y_{B'}) && (29)\\
&= T_{i-1} - \alpha(n)n\cdot\lg|\Sigma_{\text{in}}| + \alpha(n)\cdot\lg|\Sigma_{\text{in}}| \sum_{B\in S_i}|B|\,. && (30)
\end{aligned}
$$

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

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

where the last inequality uses the fact that the code is only D-deficient (eq. (4) of Definition 5). The subadditivity of entropy yields $T_\ell \ge H(Y) = n \cdot \lg|\Sigma_{\text{in}}|$, and also that

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

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

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 ECCC 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 $i \geq 1$

$$16\sqrt{\ell_i} \leq \frac{3\ell_i}{8}. \tag{33}$$

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, $P := (P_0, P_1, P_2, \ldots, P_m)$ 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 $2 \le i \le m+1$ such that s' is *not* in the rightmost block in P_{i-1} (i.e., such that $s' \le n - \ell_i/2$), the following is true. Let $s(i)$ denote the *leftmost* position in the unique block $B \in P_{i-1}$ containing s' such that $x_{s(i)} \ne y_{s(i)}$. Then, for any d satisfying

$$\ell_{i-1}/2 = 16\sqrt{\ell_i} \le d \le \ell_i/2, \tag{34}$$

it holds that

$$\Delta_{\text{Hamming}}\left(T(x)_{I(d)}, T(y)_{I(d)}\right) \ge d/3, \tag{35}$$

where $I(d)$ denotes the interval

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

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 6. *Let m be a positive integer. Let the sequence $\ell_1, \ell_2, \ldots$, the positive integer $n = \ell_{m+1}$, and the $(1/4, m)$-laminar partition $P = (P_0, P_1, P_2, \ldots, P_m)$ of n be as defined above. Let $T : \{0,1\}^n \to \Sigma^n$ be a code satisfying the CHS immediacy condition described above. Consider any block B in a partition $P_{i-1} \in P$, $2 \le i \le m+1$, 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 $x_{\text{left}(B)} \ne y_{\text{left}(B)}$ then it must be the case that $c(x)_{\text{right}(B)} \ne c(y)_{\text{right}(B)}$. Consider therefore $x, y \in \{0,1\}^n$ such that $x_{\text{left}(B)} \ne y_{\text{left}(B)}$. Thus, there must be a position s' in $\text{left}(B)$ such that $x_{s'} \ne y_{s'}$. 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 $k(i) \ge 1$ 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 $k(i) \le \ell_i/8$. Given eq. (33), d therefore satisfies the lower bound required in eq. (34) since $k(i) \le \ell_i/8$ and $i \ge 2$. 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

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

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

$$\begin{aligned}\Delta_{\text{Hamming}}\left(c(x)_{I(d)}, c(y)_{I(d)}\right) &= \Delta_{\text{Hamming}}\left(c(x)_S, c(y)_S\right) + \Delta_{\text{Hamming}}\left(c(x)_{\text{right}(B)}, c(y)_{\text{right}(B)}\right) &(38)\\ &= \Delta_{\text{Hamming}}\left(c(x)_S, c(y)_S\right) \leq |S| = \frac{\ell_i}{8} - k(i) + 1. &(39)\end{aligned}$$

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

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 7. *Let m be a positive integer. Let the sequence $\ell_1, \ell_2, \ldots$, the positive integer $n = \ell_{m+1}$, and the $(1/4, m)$-laminar partition $P = (P_0, P_1, P_2, \ldots, P_m)$ of n be as defined above. Let $T : \{0,1\}^n \to \Sigma^n$ 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 6 that T is a D-deficient $(1/4, m)$-immediacy code with deficiency sets $S_1, S_2, \ldots, S_m$, provided D satisfies

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

Now, the latter quantity can be bounded from above as

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

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

Applying Theorem 5, 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

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

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

5.2 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 $\mathrm{ECC}_\ell : \{0,1\}^\ell \to \left\{\{0,1\}^b\right\}^\ell$ 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 $0 \le i \le k$, let P_i be the partition of $[n]$ consisting of $n/2^i$ consecutive disjoint blocks of length 2^i each, and we write $P_{i,j} = \{(j-1)2^i + 1, \ldots, j \cdot 2^i\}$. 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 $\mathrm{ECC}_1(x_j)$ for $j = 1, \ldots, n$.

2. The ith row, for $2 \le i \le k+1$, consists of $(0^b)^{2^{i-2}}$ followed by the concatenation of $\mathrm{ECC}_{2^{i-2}}(x_{P_{i-2,j}})$ for $j = 1, \ldots, n \cdot 2^{2-i} - 1$. Informally, the code blocks in the ith row are "right-shifted" by 2^{i-2}.

The encoding $T(x)$ is defined by letting its jth character, for $1 \le j \le n$, be the jth 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 $0 \le \ell < k$ be any integer such that $s' \le n - 2^\ell = 2^k - 2^\ell$. 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

$$\Delta_{\mathrm{Hamming}}(T(x)_{(s,s+2^\ell]}, T(y)_{(s,s+2^\ell]}) \ge \delta \cdot 2^\ell > 0. \tag{43}$$

We now show that any code satisfying the immediacy condition described by eq. (43) must be an $(\Omega(1), \Omega(\log n))$-immediacy code. To see this, we consider the partitions P_i, $0 \le i \le k$, of $[n] = [2^k]$ considered above, and convert each of them into a tagged partition by defining, for each $B \in P_i$ with $i \ge 1$, $\mathrm{left}(B)$ to be the leftmost 2^{i-1} entries in B and $\mathrm{right}(B)$ to be the rightmost 2^{i-1} entries in B. With this construction, it is immediate that $(P_0, P_1, P_2, \ldots, P_k)$ is a $(1/2, k)$-laminar partition. Since the code T satisfies the immediacy condition in eq. (43), it is also immediate that it satisfies the neighborhood decoding property for each $B \in P_i$ with $i \ge 1$. We thus see that any tree code T with depth $n = 2^k$ and satisfying the immediacy condition given by eq. (43) must be a $(1/2, k)$-immediacy code. Applying Theorem 4, we thus see that the output alphabet Σ of such a code must satisfy

$$|\Sigma| \geq \frac{k}{2} = \Omega(\lg n), \tag{44}$$

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, for brevity, 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 $\Sigma := \{0, 1\}^{1+(\lg n)/\epsilon}$ (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,

$$\Delta_{\text{Hamming}} \left(T(x)_{(i_0, i_0+2^{t+1}]}, T(y)_{(i_0, i_0+2^{t+1}]} \right) \geq \delta \cdot 2^{t+1}. \tag{45}$$

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.[5]

We now show that any constant rate tree code satisfying the immediacy condition specified by eq. (45) cannot have distance much better than the GHKRZW construction. The proof is similar to the case of an exponential immediacy function analyzed in Section 3.1.

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

$$\ell := 1 + \lfloor (\lg (n/(2m)))/\kappa \rfloor . \tag{46}$$

Let P_0 be the partition of $[n]$ into n singletons, and for $1 \leq i \leq \ell$, define P_i to the partition of $[n]$ into blocks of consecutive integers, each of length $n/2^{\kappa \cdot (\ell - i)}$. Thus, for $1 \leq i \leq i+1$, 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 $B \in P_i$, left(B) to be the set of the smallest $2^{-\kappa} \cdot |B|$ elements in B (so that right$(B) = B \setminus$ 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. (45) 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 left$(B) = (b \cdot 2^{t+1}, (b + 2^{-\kappa}) \cdot 2^{t+1}]$. Thus, there exists $i \in$ 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. (45) then gives that

$$\Delta_{\text{Hamming}} (T(x)_B, T(y)_B) \geq \delta \cdot 2^{t+1} > 2^{-\kappa} \cdot 2^{t+1} = |\text{left}(B)| . \tag{47}$$

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 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 4 to see that a tree code $T : \Sigma_{\text{in}}^n \to \Sigma^n$ satisfying the immediacy

[5]For 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".

condition of eq. (45) must therefore satisfy

$$\frac{\lg|\Sigma|}{\lg|\Sigma_{\text{in}}|} \geq 2^{-\kappa}\cdot\ell \overset{\text{eq. (46)}}{\geq} \frac{\delta\cdot\lg\frac{n}{2m}}{2\cdot\lg(2/\delta)}. \tag{48}$$

Recall that $m = (k_0/\epsilon)\cdot\lg n$, where k_0, ϵ do not grow with n. Thus, eq. (48) shows that any constant rate tree code satisfying the immediacy condition eq. (45) of the GHKRZW construction must have a distance parameter δ satisfying

$$\frac{\delta}{\log(1/\delta)} = O(1/\log n). \tag{49}$$

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.

[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.

[BYCN21] I. Ben Yaacov, G. Cohen, and A. K. Narayanan. Candidate tree codes via Pascal determinant cubes. In M. Wootters and L. Sanità, editors, *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 sublogarithmic 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 ECCC 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.

[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.

[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.

[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.

[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.

[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.

[Gol23] O. Goldreich. On the lower bound on the length of relaxed locally decodable codes. In *Electron. Colloquium Comput. Complex., TR23-064*, volume 9, pages 10–20, 2023.

[Ham50] R. W. Hamming. Error detecting and error correcting codes. *Bell System Technical Journal*, 29(2):147–160, 1950.

[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.

[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.

[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.

[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.

[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.

[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.

[Pud16] P. Pudlák. Linear tree codes and the problem of explicit constructions. *Linear Algebra and its Applications*, 490:124–144, 2016.

[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.

[Sch93] L. J. Schulman. Deterministic coding for interactive communication. In *Proc. ACM Symposium on Theory of Computing (STOC)*, pages 747–756, 1993.

[Sch96] L. J. Schulman. Coding for interactive communication. *IEEE Transactions on Information Theory*, 42(6):1745–1756, 1996.

[Sha48] C. E. Shannon. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55, 2001. Originally appeared in *Bell System Tech. J.* 27:379–423, 623–656, 1948.

[SS96] M. Sipser and D. A. Spielman. Expander codes. *IEEE Transactions on Information Theory*, 42(6):1710–1722, 1996. DOI: 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.

[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.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il