Electronic Colloquium on Computational Complexity, Report No. 151 (2025)

Constant Rate Codes for Adaptive Broadcasts Do Not Exist

Klim Efremenko*
Ben-Gurion University

Gillat Kol†
Princeton University

Dmitry Paramonov‡
Princeton University

Raghuvansh R. Saxena§
Tata Institute of Fundamental Research

Abstract

Can the *n-party broadcast channel*, where any symbol sent by one party is received by all, be made resilient to noise with low overhead? Namely, is it possible to construct *interactive error-correcting codes* that convert any protocol designed for the noiseless broadcast channel into one that works over the noisy broadcast channel and is not much longer than the original protocol?

[EKS18, STOC 2018] showed that such interactive codes with constant multiplicative overhead are possible under the assumption that the noiseless protocol being simulated is *non-adaptive*, meaning that it is restricted to have a pre-determined order of turns. Their noise resilient simulating protocols, however, require adaptivity, where each party can decide whether or not to broadcast given all the information available to them, including their input and received transcript. The question of whether such a simulation is possible for general, potentially adaptive, noiseless protocols was left open.

We resolve this question negatively, proving that any interactive code that converts adaptive noiseless broadcast protocols into adaptive broadcast protocols resilient to stochastic errors must incur a multiplicative overhead of $\Omega(\log n / \log\log n)$, which is nearly tight.

*klimefrem@gmail.com
†gillat.kol@gmail.com
‡dp20@cs.princeton.edu
§raghuvansh.saxena@gmail.com. Supported by the Department of Atomic Energy, Government of India, under project no. RTI4001.

ISSN 1433-8092

Contents

1 Introduction

A set of n parties, each holding a private input, wish to communicate over a noisy binary *broadcast* channel that flips each bit with probability $\epsilon > 0$, independently. Is it possible to design error-correcting codes with constant rate for this setting? In this context, "error-correcting codes" means a scheme that converts a protocol intended for a noiseless broadcast channel into one that ensures the same output with high probability, even in the presence of noise. By "constant rate", we mean that the length of the noise-resilient simulating protocol is linear in the length of the noiseless protocol.

Broadcast channels are an abstraction of highly distributed *wireless* systems and the feasibility of high-rate codes for such channels has been explored since the 1980s [Gam87, Gal88], with both positive and negative results. Recently, it was shown that constant-rate codes are impossible in the *non-adaptive* broadcast setting [EKPS21b]. Non-adaptive (*a.k.a*, *oblivious* or *static*) protocols are a restricted class of protocols where the order of turns is fixed in advance and does not depend on the parties' inputs or their received transcript.

While non-adaptive protocols are useful, they do not fully leverage the capabilities of many practical wireless broadcast channels, and in fact, communication-efficient protocols for certain central problems are adaptive (*e.g.*, the well-known Decay protocol for computing network size [BGI92]). In adaptive protocols, parties can decide whether to broadcast based on all available information, including their input and the received transcripts. These protocols are known to be much more powerful than non-adaptive protocols in the context of error correcting codes [Hae14, GHS14, GH14, AGS16, EKS18, EKS20a, EKS21]. For example, [EKS18] were able to circumvent the [EKPS21b] lower bound and design constant-rate codes for settings where the noiseless protocol is non-adaptive, assuming that the simulation protocol is allowed to be adaptive. The general case, where both the noiseless and simulation protocols can be adaptive, was left open.

A partial answer to the question of adaptive-to-adaptive simulation was provided by [EKPS23], which showed that under the stochastic *message drops* (*a.k.a*, *erasures*) model[1] proposed by [CHHHZ19], constant-rate codes exist. [EKPS23] also note that most interactive coding lower bounds for multi-party protocols under stochastic corruptions (bit-flips) extend to the erasure model (*e.g.*, [BEGH16, EKS19, EKPS21b]), suggesting that proving a lower bound for corruptions may require new techniques.

1.1 Our Result

We resolve the general case of adaptive-to-adaptive simulation over the binary broadcast channel with corruptions, proving that constant-rate codes do not exist for this channel.

Theorem 1.1 (Informal, see Theorem 3.2)**.** *Let $\epsilon \in (0, 1/2]$, $n \geq 1$, and $T(n) = n^{\Theta(1)}$. There exists a deterministic adaptive protocol Π with $T(n)$ rounds over the n-party (noiseless)*

[1] In this model, each party receives the broadcast bit with probability $1 - \epsilon$, independently, and receives '$\perp$' to indicate an erasure with probability ϵ.

broadcast channel, such that any randomized adaptive protocol that simulates Π *over the* n*-party* ϵ*-noisy broadcast channel with constant error probability has* $\Omega(T(n) \cdot \frac{\log n}{\log \log n})$ *rounds.*

The overhead in Theorem 1.1 is optimal up to $\mathcal{O}(\log \log n)$ factors, as polynomial-length protocols can be simulated with $\Omega(\log n)$ overhead by repeating each round $\Theta(\log n)$ times and having each party take the majority of the symbols it receives. This process effectively reduces the error rate to $\frac{1}{\text{poly}(n)}$.

Ruling out adaptive simulations. Theorem 1.1 proves a lower bound against adaptive simulation protocols. Such lower bounds are tricky, as when a message is corrupted, it may not only cause the parties to alter the content of their future messages, but it can also lead them to change the set of rounds in which they choose to broadcast. For instance, they can decide to dynamically allocate more rounds to the parties that were corrupted the most (see, *e.g.*, [GHS14]).

A key technical challenge in proving lower bounds against adaptive simulations is that most techniques for proving communication lower bounds rely on the assumption that, at any point during the execution of the protocol, conditioned on the execution so far, the set of possible inputs forms a *combinatorial rectangle* (in other words, a 'product set'). This holds true for multi-party non-adaptive protocols, as exactly one party broadcasts in each round and the identity of this party is known. However, in adaptive protocols, new dependencies between the inputs can arise, implying that they are no longer a combinatorial rectangle.

The hard-to-simulate adaptive protocol. Lower bounds on the rate of interactive codes are often proved for the *pointer-chasing* problem (see Section 2.1.2). While pointer-chasing is "complete" for non-adaptive protocols, in the sense that every non-adaptive protocol can be viewed as a pointer chasing protocol, the pointer-chasing problem admits a non-adaptive protocol and therefore is subject to the [EKS18] scheme. Thus, it is not a good candidate for a hard-to-simulate protocol in our case.

Instead, the noiseless protocol used by the proof of Theorem 1.1 is a protocol for the *pointer-chasing with unique secrets* communication problem. In this problem, the previously communicated edges in the underlying pointer-chasing tree do not determine the next party to broadcast an edge, but merely select a large *group* of parties that contains this next party. In the noiseless protocol, selecting the next party out of the group is done via very short adaptive sub-protocol that identifies the group member that has the "*unique secret*" (see Section 2), but requires more communication in the noisy setting.

The adaptive broadcast model. We now describe the model assumed by Theorem 1.1. An adaptive protocol over the n-party (noiseless) broadcast channel is a communication protocol where n parties communicate in synchronous rounds. In each round, each party can choose to either *broadcast* a bit or remain *silent*. If exactly one party broadcasts a bit, each party receives the broadcast bit. However, if more than one party broadcasts in a given

round (a *collision*), or if no party broadcasts (a *silent round*), all parties receive the '$\perp$' symbol. This model is known as the *(single-hop) collision-as-silence radio networks* model, as the same '$\perp$' symbol is received in both collision and silent rounds, and it is the most commonly used collision-handling model in the literature[2].

When running an adaptive protocol over an n-party ϵ-noisy broadcast channel, in each round, each party receives the symbol from $\{0, 1, \perp\}$ that it would have received if this round had been run over the n-party (noiseless) broadcast channel with probability $1 - \epsilon$, and a random symbol from $\{0, 1, \perp\}$ with probability ϵ, independently for each round and party.

1.2 Related Work

Interactive coding. *Interactive error-correcting codes* encode interactive communication protocols designed for noiseless channels into protocols that can also work over noisy channels. The study of interactive codes began with a landmark paper by Schulman [Sch92], which focused on two-party protocols and sparked numerous follow-up works. Over the last decade, interactive codes for multi-party distributed channels have garnered significant attention. This includes codes for *peer-to-peer* channels [RS94, JKL15, HS16, ABE+16, BEGH16, GK19, GKR19] and codes for various *broadcast* channels [Gal88, Yao97, KM05, FK00, New04, GKS08, CHHHZ17, EKS18, CHHHZ19, EKS19, EKS20b, AGL20, EKPS21a, MG21, EKPS21b, EKPS23, EKP+24]. Our work contributes to the literature on the latter.

In the context of coding for broadcast channels, prior work has explored both adaptive and non-adaptive protocols. We next survey the most relevant results on simulating non-adaptive/adaptive protocols by non-adaptive/adaptive noise resilient protocols over the broadcast channel. Naturally, if the noiseless protocols being considered are adaptive, the simulation should also be adaptive. However, simulations of non-adaptive noiseless protocols by adaptive noise-resilient protocols have also been considered.

Non-adaptive to non-adaptive simulation. The study of the noise resilience of broadcast channels was initiated by El Gamal [Gam87], who introduced the *noisy broadcast* model. This is a noisy version of the *non-adaptive broadcast* model (*a.k.a.*, the *shared blackboard* model), where a set of n parties, each holding a private input, communicate in synchronized rounds. In each round, a pre-specified party broadcasts a bit to all the other parties. However, the bit received by each party is randomly flipped with some fixed constant probability $\epsilon > 0$, independently for each party and round. This model was later popularized by [Yao97] as a simple abstraction for understanding the impact of noise on highly distributed wireless systems.

El Gamal posed the following challenge: How many rounds are needed to solve the *bit-exchange* problem, where each party has a bit input and needs to learn the inputs of

[2]Another widely used model is the *collision detection* model, where collisions and silence are received as different symbols. We define our channel as collision-as-silence, but Theorem 1.1 also applies to the collision detection model and other related models.

all other parties? Gallager [Gal88] gave an elegant $\mathcal{O}(n \log \log n)$-round protocol for this problem, which was later proved to be optimal by [GKS08]. The bit-exchange problem under erasure noise was considered by [GHM18], who gave an $\mathcal{O}(n \log^* n)$-round protocol. The bit-exchange problem is equivalent to computing the identity function, and the round complexity of other specific n-bit functions, such as OR, MAJORITY, and PARITY, has been studied under related noise models [Yao97, KM05, FK00, New04, GKS08].

The general case of simulating *any* non-adaptive protocol with a noise-resilient non-adaptive protocol was recently studied by [EKPS21b]. Their main result shows that for protocols of length polynomial in n, such a simulation requires $\tilde{\Theta}(\sqrt{\log n})$ multiplicative overhead in the number of round. We note that the question of non-adaptive to non-adaptive simulation with low overhead and the question of adaptive-to-adaptive simulation with low overhead are incomparable. While the simulation protocols in the latter case are more powerful, so are the noiseless protocols they attempt to simulate.

Non-adaptive to adaptive simulation. [EKS18] gave a scheme for converting any non-adaptive noiseless protocol into an adaptive, noise-resilient one with only a constant multiplicative overhead. Theorem 1.1 shows that their scheme cannot be extended to simulate adaptive protocols as well.

Adaptive to adaptive simulation with erasures. [EKPS23] designed a constant-rate scheme for converting any adaptive noiseless protocol into an adaptive, noise-resilient one with only a constant multiplicative overhead under *erasures*. Theorem 1.1 shows that their result cannot be extended to handle corruptions.

2 Proof Sketch

In this section, we give a detailed outline of the proof of Theorem 1.1. In Section 2.1, we motivate the design of our hard-to-simulate problem, *pointer-chasing with unique secrets*, by discussing the limitations of the non-adaptive to adaptive simulation of [EKS18] (this subsection can be skipped as the rest of the section is self-contained). We then outline our lower bound proof for this problem.

2.1 Designing the Hard-to-Simulate Protocol

2.1.1 The [EKS18] Scheme

The scheme in [EKS18] employs the *rewind-if-error* framework, which was originally developed for the two-party interactive coding setting [Sch92]. Rewind-if-error coding schemes involve multiple iterations, with each iteration consisting of two phases: a *simulation phase*, in which a small number of rounds of the noiseless protocol are executed, and a *consistency check phase*, where the parties check whether they have the same received transcript or if an

error occurred (*e.g.*, by comparing hashes of their received transcripts). If the check phase succeeds, the parties proceed with the simulation; otherwise, they *rewind* and re-simulate the last few rounds.

One key reason the [EKS18] scheme fails when applied to adaptive protocols is due to *repeated rewinds*: With a noise rate of ϵ, we can expect approximately ϵn parties to receive an incorrect bit in each round of the simulation phase. Since ϵ is constant, $\epsilon n \gg 1$. This means that the consistency check phase will almost always fail and trigger a rewind, preventing any progress from being made.

While the total number of parties, n, is large, the [EKS18] simulation successfully bypassed the repeated rewinds problem. This is achieved by observing that the non-adaptivity of Π could be used to identify a small subset S of parties that *critically* need to know the simulated transcript. These parties are those that will broadcast in the rounds immediately following the current one. The other parties broadcast later and thus have more time to decode the bit broadcast in the current round. [EKS18] showed that it suffices to ensure that the parties in S are not affected by noise. However, in the adaptive case, where *any* of the n parties may broadcast next, this approach is not feasible.

2.1.2 Failed Attempt: Weakly Adaptive Pointer-Chasing

Recall that the *pointer-chasing* communication problem involves an underlying (possibly non-binary) tree. Each layer of this tree is "owned" by one of the communicating parties, and each party gets a single edge going out of each of the vertices in the layers they own. The parties' objective is to find a root-to-leaf path that only uses edges in the union of their input edge sets. As discussed in Section 1.1, while lower bounds against non-adaptive simulations are typically shown for the pointer-chasing protocol[3], this protocol is non-adaptive and therefore can be simulated with low overhead using the [EKS18] scheme.

Building on the discussion above regarding the shortcomings of the [EKS18] scheme, our hard-to-simulate protocol should ensure that the identity of the next party to send an edge is unknown. To achieve this, we can modify the standard pointer-chasing problem so that different parties own different vertices at the same level. Specifically, assume that each party owns one vertex in the underlying pointer-chasing tree and has an outgoing edge from this vertex. In this setup, the identity of the second party to send an edge is only determined after the party owning the root node sends the first edge.

While this version of the pointer-chasing protocol is adaptive, as the order of turns is not pre-determined, it is only "*weakly adaptive*" in the sense that exactly one party broadcasts in each round and the order of turns only depends on the transcript (and not on the private inputs of the parties, unless those inputs have been communicated). A more careful analysis of the [EKS18] scheme shows that it can be expected to work for such weakly adaptive noiseless protocols. The reason is that, although the identity of the party that speaks in

[3]The owner of the root v_0 broadcasts their edge (v_0, v_1), then the owner of v_1 broadcasts their edge (v_1, v_2), *etc.*

round $t+1$ is unknown ahead of time, after round t, the identity of this party is known to all. Thus, if the simulation until round t was (roughly) correct, round $t+1$ effectively becomes (close to) non-adaptive, and the idea of [EKS18] still applies.

2.1.3 Hard-to-Simulate: Pointer-Chasing with Unique Secrets

The previous attempt at changing the pointing-chasing problem failed as the identity of the next speaker is determined by the transcript so far. This will no longer be the case if the identity of the next speaker also depends on their private input. Motivated by this, in our pointer-chasing with unique secrets problem, the previous edges communicated only determine a large "*group*" (set of size $n^{\Theta(1)}$) of parties that contains the party who has the next edge, and act as a *disguise* to the identity of this party.

Unique secrets. We next explain how the party that broadcast the next edge is selected out of the group, while making sure that its identity is not implied by the communication. An easy way to do so is to select a random party $i \neq 1$ in the group and a random bit b, and give parties i and party 1 (who we refer to as the "*leader*" of the group) the bit b, while all the other parties get $\bar{b}$. We refer to those additional input bits as "*secrets*" (recall each party also has an edge as input). Observe that there is exactly one party whose secret is unique among the non-leader parties. We call this party the "*unique party*". Also observe that the leader knows the unique secret, but not the identity of the unique party.

In the noiseless protocol, the leader broadcasts their secret, which immediately allows the party with the matching secret to know that they are the unique party. This party then broadcasts their edge to determine the next group. Note that the identity of the unique party cannot be deduced from the communication and requires knowing the private secret of that party. Therefore, this protocol is not weakly adaptive.

Longer secrets. Finding the unique party can also be done by a different short adaptive sub-protocol that does not rely on the leader: In the first round, all non-leader parties with secret 0 broadcast 0, and in round 2, all non-leader parties with secret 1 broadcast 1. Notice that if, for instance, $b = 1$ is the unique secret, the unique party will be the only one broadcasting in round 2, while all other non-leader parties will broadcast in round 1, leading to a collision. It is straightforward to deduce the unique secret from the protocol's transcript—only one of the rounds will have a non-'$\perp$' value, and this value is the unique secret.

To simplify the lower bound proof, we aim to eliminate this second sub-protocol for identifying the unique party. To achieve this, we assign each party a secret that is a bit string of length $\Theta(\log \log n)$, rather than a single bit, while still ensuring there is one unique party whose secret matches the leader's. Each of the other possible secrets is given to many other parties (there are $n^{\Theta(1)}$ parties in a group and only $\operatorname{poly} \log n$ possible secrets). Note that the second protocol can be adapted to work with any set of secrets $\{s_1, \ldots, s_k\}$, by

having parties with secret s_ℓ broadcast in round ℓ. However, the length of this protocol increases with k, the size of the set. By choosing secrets of length $\Theta(\log \log n)$, we make the length of the protocol $k = \text{poly} \log n$, which is too large.

Why is it hard to simulate? Intuitively, the difficulty in solving the pointer-chasing with unique secrets problem over a noisy channel arises from the fact that when the leader of the current group broadcasts their secret, each bit of the broadcast is likely to be received incorrectly by a constant fraction of the parties. Given that the secret is relatively short and there are many parties in the group, it is expected that several parties j will receive a noisy version of the leader's secret that coincidentally matches their own secret. Each of these parties j may mistakenly believe they are the unique party and proceed to broadcast their own edge, leading to collisions and preventing meaningful information transmission[4].

The remainder of this section gives a detailed outline of our lower bound proof.

2.2 Lower Bound

As discussed above, for our lower bound, we consider an instance of pointer-chasing with unique secrets. Here is a more formal description of this problem: The n parties are divided into $B = n^{\Theta(1)}$ *blocks* and each block is divided into $G = (\log n)^{\Theta(1)}$ *groups*. One of the players in every group is the *leader* of the group (for concreteness, the first party in the group), and each party has as input a "*secret*" s and a "*key*" $k \in [G]$. It is promised that in every group, there is exactly one other party with the same secret as the leader. We call this party the *unique party*. The key of the unique party is called the *groupkey* and it identifies a group in the next block. The goal of the parties is to start from the first group in the first block and compute its groupkey, then compute the groupkey of the group it identifies in the second block, and so on for all the B blocks.

Our lower bound shows that a protocol with $o(B \cdot \log n)$ rounds over the noisy channel cannot compute all the B groupkeys. Very roughly, we show this by proving that in any chunk of $T_{\mathsf{Ch}} = \Theta(\log n)$ rounds, the protocol can only identify at most a constant number of groupkeys in expectation. As the total number of chunks is $o(B)$, so is the total number of groupkeys identified by the protocol, and the lower bound follows. To see why any chunk can only identify a constant number of groupkeys, consider first what happens in the first chunk of the protocol. We will later generalize this to all the other chunks.

Analyzing the first chunk. We show that after the first chunk, which is the first T_{Ch} rounds of the protocol, the parties do not know the groupkey of the first group. Assuming this for now, note that this means that the parties do not know which group is the "correct" group in the future blocks, which implies that even if they identify groupkeys of some of the groups in the future blocks, they are unlikely to be relevant to the output of the protocol.

[4]We point out that, as suggested by [EKPS23], under erasures, a simulation protocol with constant overhead is possible, as such parties j will no longer be misled into thinking they are the unique party.

A bit more formally, if the groupkey of the first group is completely random after the first chunk, each of the groups in the future blocks is likely to be correct with probability $\frac{1}{G}$, and so is any groupkey computed for these groups. As the number of groupkeys computed cannot exceed the length of the chunk, we get that the total number of relevant groupkeys computed for the future blocks is at most $\frac{T_{\mathsf{Ch}}}{G} = o(1)$. Combined with the groupkey of the first block, we get that this chunk only revealed $1 + o(1)$ many relevant groupkeys, as desired.

We now argue why the parties do not know the groupkey of the first group in the first T_{Ch} many rounds of the protocol. For this, consider each round $z \in [T_{\mathsf{Ch}}]$ and consider for all parties j in the first group, what are the possible inputs and transcripts for which party j will speak in round z. We say that a secret s is *energizing* (for round z) if for polynomially many parties j, there exists an input of the form $(s, \cdot)$ and a received transcript that will make party j broadcast[5] in round z. We consider two cases:

1. There is at most one value of s that is energizing, call it s^*. In this case, for all other values of s, very few parties with secret s can potentially speak in round z. Thus, if the leader's secret is anything other than s^*, it is unlikely that the unique party (who is a uniformly random party among all non-leader parties in the group) will speak in round z. This implies that unless the leader's secret is s^*, the groupkey is not revealed in round z, with high probability.

2. There are at least two different values of s that are energizing. In this case, fix any secret s' for the leader and consider the energizing value of s that is not s'. As our groups are large, there are polynomially many parties whose input is the input $(s, \cdot)$ that can make them broadcast, and as the chunks are small, we also have that a polynomially large number of these parties will get the transcript that makes them broadcast on this input, with high probability. Thus, regardless of the leader's secret, this round z will have at least two parties broadcasting, resulting in a collision and the groupkey will not be revealed.

Combining these two cases, we get that the only way the groupkey can possibly be reveal in this chunk is if the leader's secret is s^* (as defined in Item 1) for some round z in this chunk. As the number T_{Ch} of rounds is small and the leader's secret is uniformly random, this is unlikely, and we are done.

Analyzing the other chunks. The first chunk in the protocol was easy to analyze as the distribution of the parties' inputs was known before the chunk. Thus, it was known that the leader's secret is uniformly random, the unique player is uniformly random independently of the leader's secret, *etc.* This may not be true in the later chunks, as the transcript so far may have distorted the distribution of the parties' inputs before the future chunk. Roughly

[5] This is done by going over all possible inputs and received transcripts for party j, and therefore overcounts the number of parties that may broadcast in round z (as their actual input may not be the one that will make them broadcast). Our analysis works with this overcount as well.

speaking, the way we analyze these chunks is to show that our proof for the first chunk is "robust", in the sense that it goes through even if the distributions are slightly distorted.

To formalize this, for every group in every block, we keep track of how much its distribution differs from the initial distribution, measured in terms of KL divergence (relative entropy). For the groups whose distributions are close, *i.e.*, whose relative entropy is upper bounded by $\frac{1}{\text{poly}\log n}$, we show that the analysis above still goes through. On the other hand, for the groups whose distributions have changed a lot, we argue that these groups must be few in number. This is because whenever the relative entropy is large, it means the protocol has revealed a lot of "information" about this group. However, the total amount of information revealed cannot exceed the length of the protocol, and thus the number of groups for which it is large must be small.

We handle these groups separately using the fact that these groups are independent of the groupkey the parties are trying to compute in this chunk. Thus, these small number of revealed groups are unlikely to be correct, and contribute little to the computation the parties are trying to perform.

Revealing blocks. To finish this sketch, we note that the actual distribution we consider differs slightly from the distribution above. Specifically, in our proof we also require that the groupkeys for all the groups in the same block are distinct, ensuring that it possible for every group in the next block to be the correct one. This is because if not, then it could be possible for the parties to find out the groups that cannot possibly be correct in the future blocks, and focus only on the other groups. This could make the task much easier, and render the lower bound impossible.

However, ensuring that every group in any block has a distinct groupkey leads to another difficulty. Namely, if one knows the groupkey of any group in a block, one also knows that the groupkeys of the other groups are different from this one. Similarly, if one knows the groupkey of a lot of the groups in a block, then one has a lot of information about the groupkeys of the remaining blocks, making them easier for the protocol to solve. In our analysis, we show that we can afford to ignore all the blocks for which we know a lot of groupkeys. In other words, whenever it is the case that we know groupkeys for many groups in any block, we will reveal the entire block. This would leave us with blocks for which we have a small number of groupkeys, meaning the remaining groupkey are still very random, and the previous argument would go through.

3 Model and Preliminaries

Let ε denote the empty string or the empty tuple. For integers a, b, we let $[a, b]$ denote the set of all integers i satisfying $a \leq i \leq b$. We change the square brackets to parenthesis when the inequality is strict and define the notations (a, b), $[a, b)$, and $(a, b]$. As usual we abbreviate $(0, n] = \{1, 2, \ldots, n\}$ to $[n]$.

3.1 Concentration Inequalities

Lemma 3.1 (Multiplicative Chernoff bound)**.** *Suppose $X_1, \cdots, X_n$ are independent random variables taking values in $[0,1]$. Let X denote their sum and let $\mu = \mathbb{E}[X]$ denote the sum's expected value. Then,*

$$\Pr(X \geq \mu \cdot (1+\delta)) \leq \mathrm{e}^{-\frac{\delta^2 \mu}{2+\delta}}, \qquad \forall 0 \leq \delta,$$
$$\Pr(X \leq \mu \cdot (1-\delta)) \leq \mathrm{e}^{-\frac{\delta^2 \mu}{2}}, \qquad \forall 0 \leq \delta \leq 1.$$

3.2 The Noisy Broadcast Model

The noisy broadcast model is defined by a number $n > 0$ of parties and a noise parameter $\epsilon \in [0,1]$. When $\epsilon = 0$, we say that the broadcast model is *noiseless* and may drop ϵ from the notation. A (deterministic) protocol over the (n,ϵ)-noisy broadcast model is defined by a tuple:

$$\Pi = \Big(T, (\mathcal{X}_i)_{i\in[n]}, \mathcal{Y}, (\mathsf{msg}_i)_{i\in[n]}, \mathsf{out}\Big), \tag{1}$$

where: (1) $T = \|\Pi\| > 0$ is a parameter denoting the length of the protocol. (2) For all $i \in [n]$, $\mathcal{X}_i$ is the set of inputs of party i. (3) $\mathcal{Y}$ is the set of possible outputs for the protocol. (4) For all $i \in [n]$, $\mathsf{msg}_i : \mathcal{X}_i \times (\{0,1,\perp\})^* \to \{0,1,\perp\}$ is a function that computes the message sent by party i based on its input and the received transcript so far. (5) $\mathsf{out} : (\{0,1,\perp\})^T \to \mathcal{Y}$ is a function that computes the output of the protocol from its transcript. We suppress items on the right hand side of Equation (1) when they are clear from context. We define a randomized protocol to be a distribution over deterministic protocols.

Execution of a protocol. We now define the execution of a protocol in the presence of random noise. Fix a protocol Π as above and let $N = (N_t^i)_{i\in[n],t\in[T]}$ be a noise vector such that for all $i \in [n]$ and $t \in [T]$, the value $N_t^i = \star$ with probability $1-\epsilon$ and with probability ϵ, we have that N_t^i is a uniformly random symbol from the set $\{0,1,\perp\}$. Furthermore, N_t^i is sampled independently for all $i \in [n]$ and $t \in [T]$.

Given such a Π and N, the protocol Π starts with all parties having an inputs $x_i \in \mathcal{X}_i$ and proceed in T rounds, maintaining the invariant that before round $t \in [T]$, all parties $i \in [n]$ have received a transcript $\Pi_{<t}^i \in \{0,1,\perp\}^{t-1}$. In round $t \in [T]$ all parties $i \in [n]$ compute $\mathsf{msg}_i(x_i, \Pi_{<t}^i) \in \{0,1,\perp\}$. We say party i speaks (or broadcasts) in round t if $\mathsf{msg}_i(x_i, \Pi_{<t}^i) \neq \perp$. If the number of parties speaking in round t is different from 1, define $\Pi_t^{\mathsf{true}} = \perp$. Otherwise, there is exactly one party $i^* \in [n]$ speaking in round t and we define $\Pi_t^{\mathsf{true}} = \mathsf{msg}_{i^*}\big(x_{i^*}, \Pi_{<t}^{i^*}\big)$. For all parties $i \in [n]$, the symbol Π_t^i received by party i in round t equals Π_t^{true} if $N_t^i = \star$ and equals N_t^i otherwise. All parties $i \in [n]$ append Π_t^i to $\Pi_{<t}^i$ and continue executing the protocol. After T rounds are over, party 1 outputs $\mathsf{out}\big(\Pi_{\leq T}^1\big)$.

Observe that the execution of the protocol is determined by the inputs $X = (x_i)_{i\in[n]}$ and the noise N. In particular, the output of party 1 is determined by X and N. Due to this,

we sometimes write the output of the protocol as $\mathsf{out}_\Pi(X, N)$.

3.3 The Hard Instance

We divide the n parties into $B = n^{0.1}$ blocks of $\frac{n}{B}$ parties each. Each block of parties is further subdivided into $G = (\log n)^{10}$ groups of $\frac{n}{BG}$ parties each. For convenience, we define $m = \frac{n}{BG} - 1 \geq n^{0.89}$ so that there are $m + 1$ parties in any group. Thus, each group can be given a unique index $(b, g) \in [B] \times [G]$ and each party can be given a unique index $i = (b, g, j) \in [B] \times [G] \times [0, m]$. we interpret i to be a number in $[n]$ in the natural way. We omit writing $[B] \times [G]$ and $[B] \times [G] \times [0, m]$ when it is clear from context. For all groups (b, g), we define the party $(b, g, 0)$ to be the "leader" of the group (b, g). Let $S = G$ be a parameter. For all parties $i \in [n]$, the input of party i is a pair $x_i = (s_i, k_i) \in [S] \times [G]$ consisting of a "secret" and a "groupkey". We will only consider inputs that satisfy the following two promises:

1. For all groups (b, g), there is exactly one party in the group (b, g) whose secret matches the leader. Formally, we have:

$$\forall (b, g) : \qquad |\{j \in [m] \mid s_{b,g,j} = s_{b,g,0}\}| = 1. \tag{2}$$

   We will use $\mathsf{uq}_{b,g} \in [m]$ to denote the unique value in the set above and call player $\left(b, g, \mathsf{uq}_{b,g}\right)$ the unique player in the group (b, g). Also, we define $\mathsf{gkey}_{b,g} = k_{b,g,\mathsf{uq}_{b,g}}$ to be the groupkey of the unique player in the group (b, g), and also call it the groupkey of the group (b, g).

2. We also require that $\mathsf{gkey}_{b,g}$ is different for all the groups in the same block. This is equivalent to saying that the values $\mathsf{gkey}_{b,g}$ for the groups in any block form a permutation over $[G]$. Formally,

$$\forall (b, g) \neq (b, g') : \qquad \mathsf{gkey}_{b,g} \neq \mathsf{gkey}_{b,g'}. \tag{3}$$

We define the distribution $\mathcal{D}$ to be the uniform distribution on all inputs $X \in ([S] \times [G])^n$ that satisfy Equations (2) and (3). We write X_b when we restrict attention to inputs in a given block $b \in [B]$ and $X_{b,g}$ when we restrict attention to inputs in the group (b, g). Observe that the random variables $(X_b)_{b \in [B]}$ are mutually independent.

We now define the function gkeys the parties want to compute. Roughly speaking, this is just the sequence of groupkeys of all groups starting from the first group in the first block. Formally[6], for any inputs $X \in ([S] \times [G])^n$ for the parties, any group (b, g) and any

[6]While the parties need to compute the groupkeys starting from the first group in the first block, our formal definition is more general and assumes they are starting from an arbitrary group (b, g). This means that the "base case" is $b' = b - 1$.

coordinate $b' \in [b-1, B]$, we recursively define:

$$\mathsf{gkeys}_{b'}(X, b, g) = \begin{cases} g, & \text{if } b' = b-1 \\ \mathsf{gkey}_{b', \mathsf{gkeys}_{b'-1}(X,b,g)}, & \text{otherwise} \end{cases}. \tag{4}$$

We omit writing b, g when $b = g = 1$ and adopt the convention that $\mathsf{gkeys}_B(X, B+1, g) = g$ for all $g \in [G]$ and all inputs X.

3.4 Main Result

Let $n > 0$. Observe that when party i has input x_i as defined in Section 3.3, there exists an $\mathcal{O}(B \cdot \log\log n)$ round protocol that computes[7] $\mathsf{gkeys}_{[B]}(X)$ when the noise parameter $\epsilon = 0$. The protocol divides the rounds into B blocks of $\mathcal{O}(\log\log n)$ rounds each and satisfies the property that at the end of block b, for all $0 \le b \le B$, party 1 knows $\mathsf{gkeys}_{[b]}(X)$.

This property is trivially satisfied for $b = 0$. Suppose that $b > 0$ and the property is satisfied for $b-1$. As the protocol is noiseless, all the parties have the same transcript and they all know $\mathsf{gkeys}_{[b-1]}(X)$ before block b. Let $g = \mathsf{gkeys}_{b-1}(X)$. In block b, the leader $(b, g, 0)$ of the group (b, g) takes $\mathcal{O}(\log\log n)$ rounds to broadcast $s_{b,g,0}$. Then, all the other parties, *i.e.*, party (b, g, j) for $j \in [m]$, check if $s_{b,g,j} = s_{b,g,0}$ which they just received. If this check passes (which only happens for party $\mathsf{uq}_{b,g}$ by Equation (2)), party j takes $\mathcal{O}(\log\log n)$ rounds to broadcast $k_{b,g,j} = \mathsf{gkey}_{b,g}$. By Equation (4), this equals $\mathsf{gkeys}_b(X)$, as desired.

With this protocol, in order to show Theorem 1.1, it suffices to show that any noisy protocol computing $\mathsf{gkeys}_{[B]}(X)$ requires $\Omega(B \cdot \log n)$ rounds. This is captured in the theorem below, which implies Theorem 1.1. We state the theorem below with $\epsilon = \frac{3}{10}$ but this choice only affects the constants in the theorem statement.

Theorem 3.2. *Fix $n > 0$ large enough and $\epsilon = \frac{3}{10}$. For any (possibly randomized) protocol Π in the (n, ϵ)-noisy broadcast channel with $\|\Pi\| \le 10^{-6} \cdot B \cdot \log n$, it holds that:*

$$\Pr_{X \sim \mathcal{D}, N \sim \mathcal{N}} \left(\mathsf{out}_\Pi(X, N) = \mathsf{gkeys}_{[B]}(X) \right) \le 0.1.$$

4 The Lower Bound: Proof of Theorem 3.2

We devote this section to proving Theorem 3.2. Let $n > 0$ be large enough and Π be a protocol satisfying $T = \|\Pi\| \le 10^{-6} \cdot B \cdot \log n$ as in the theorem statement. As we work with a specific distribution $\mathcal{D}$, we can assume without loss of generality that Π is deterministic. Recall that $\mathsf{N} \sim \mathcal{N}$ is the random variable for the noise in the channel and is independent of $\mathcal{D}$. All our probabilities and events would be defined in terms of the distribution $\mathcal{D} \times \mathcal{N}$. We omit writing the distributions when it is clear from context.

[7] Throughout this work, for a tuple $x = (x_1, \ldots, x_n)$ and a set $S \subseteq [n]$, we use $x_S = (x_i)_{i \in S}$ to denote the coordinates for the tuple that are in S.

We will use sans-serif letters $\mathsf{X}, \mathsf{s}, \mathsf{k}$, *etc.* to denote random variables and the corresponding letters to denote the realizations. For example, the random variable corresponding to the input of a party $i \in [n]$ is denoted using $\mathsf{x}_i = (\mathsf{s}_i, \mathsf{k}_i)$. Observe from our definitions that the random variables X_b are mutually independent. As all the randomness in our setting comes from the distributions $\mathcal{D}$ and $\mathcal{N}$, once we fix an input $X \sim \mathcal{D}$ and a noise vector $N \sim \mathcal{N}$, we also fix the entire execution of the protocol Π (as it is determined by X and N). Additionally, for all $t \in [T]$, fixing $X \sim \mathcal{D}$ and a noise vector $N_{\leq t} \sim \mathcal{N}_{\leq t}$ fixes the execution of Π in the first t rounds and also fixes what every party is broadcasting in round $t+1$. Thus, for all $t \in [T]$, we can define:

$$\mathsf{Spk}_t(X, N_{<t}) = \left\{ i \mid \mathsf{msg}_i\left(x_i, \Pi^i_{<t}\right) \neq \perp \right\}, \tag{5}$$

to be the set of parties speaking in round t. For our analysis, we divide the protocol Π into chunks of $T_{\mathsf{Ch}} = \frac{1}{500} \cdot \log n$ rounds each and let $C = T/T_{\mathsf{Ch}} \leq \frac{B}{2000}$ denote the total number of chunks. This means that any round $t \in [T]$ can be equivalently written as a pair $(c, z) \in [C] \times [T_{\mathsf{Ch}}]$, where c is the current chunk and z in the index of t in the current chunk. We use these two interchangeably.

Let $c \in [C]$ be a chunk and define $\widehat{T_{\mathsf{Ch}}} = 2 \cdot T_{\mathsf{Ch}}$ and $L = \widehat{T_{\mathsf{Ch}}} + BG$. We now define several functions of the inputs X for the parties and the noise $N_{\leq c \cdot T_{\mathsf{Ch}}}$ in the first c chunks. For notational convenience, we keep the dependence on X and $N_{\leq c \cdot T_{\mathsf{Ch}}}$ implicit. We will define five functions, namely, $\left(\Phi_{(c,l)}, \Phi^{\mathsf{sm}}_{(c,l)}, \Phi^{\mathsf{gk}}_{(c,l)}, \Gamma_{(c,l)}, \Upsilon_{(c,l)}\right)_{l \in [L]}$. Intuitively, these functions capture the information we "reveal" or "condition on" in our proof to ensure independence between certain random variables. For each round $z \in [T_{\mathsf{Ch}}]$, we reveal up to two parties that are speaking in that round along with their inputs. This information is enough to determine if the round is a silent round or a collision round and if it is neither, the bit sent in that round. Because of this factor of two, we have that $\widehat{T_{\mathsf{Ch}}}$ is twice T_{Ch}. After the chunk is over, we also reveal all the groups that are not sufficiently random. These can be up to BG in number so L is that much larger than $\widehat{T_{\mathsf{Ch}}}$. The exact information we need to reveal is different for different parts of the proof and is captured in the functions that are formally defined inductively as follows. For any $l \in [L]$, suppose that the functions have been defined for all $l' < l$.

1. For all $l \in \left[\widehat{T_{\mathsf{Ch}}}\right]$, if l is odd, let $i = (b, g, j)$ be the smallest element in $\mathsf{Spk}_{(c, \lceil l/2 \rceil)}\left(X, N_{<(c, \lceil l/2 \rceil)}\right)$ (if it exists, or equivalently, if the set is non-empty). Set:

$$\Phi_{(c,l)} = ((b, g), X_{b,g}), \qquad \Phi^{\mathsf{sm}}_{(c,l)} = (i, x_i), \qquad \Phi^{\mathsf{gk}}_{(c,l)} = \left(i, x_i, \mathsf{gkey}_{b,g}\right). \tag{6}$$

   If the set is empty, then we set each element in the tuples above to $\perp$. The definition for even l is analogous except that we use the second smallest element instead of the smallest.

2. For all $l \in \left(\widehat{T_{\mathsf{Ch}}}, L\right]$, if there exists a group $(b, g) \notin \Gamma_{<(c,l)}$ that satisfies one of the

following conditions (we use $\mathcal{D}_{b,g}$ to denote the marginal distribution of $\mathcal{D}$ corresponding to the group (b,g)):

$$\begin{aligned} \mathbb{D}\Big(\big(\mathcal{D} \mid \Upsilon_{<(c,l)}\big)_{b,g} \;||\; \mathcal{D}_{b,g}\Big) &\geq \frac{1}{(\log n)^4}, \\ \big|\Gamma_{<(c,l)} \cap (\{b\} \times [G])\big| &\geq \frac{G}{4}, \end{aligned} \tag{7}$$

then, we let (b,g) denote the smallest such group and define

$$\Phi_{(c,l)} = ((b,g), X_{b,g}), \qquad \Phi^{\mathsf{sm}}_{(c,l)} = (\perp,\perp), \qquad \Phi^{\mathsf{gk}}_{(c,l)} = (\perp,\perp,\perp). \tag{8}$$

If no such group exists, we set each element in the tuples above to $\perp$.

In both cases, we define:

$$\begin{aligned} \Gamma_{(c,l)} &= \big(\Gamma_{<(c,l)} \cup \{\Phi_{(c,l),1}\}\big) \setminus \{\perp\}, \\ \Upsilon_{(c,l)} &= \big(\Phi_{(c,l)}, N_{\leq(c,\min(\lceil l/2 \rceil, T_{\mathsf{Ch}}))}\big). \end{aligned} \tag{9}$$

We will view these functions as random variables determined by the inputs X and $\mathsf{N}_{\leq c \cdot T_{\mathsf{Ch}}}$. To emphasize this, we will use sans-serif letters $\Phi_{(c,l)}$, *etc.* when we look at them as random variables and the corresponding normal letters $\Phi_{(c,l)}$ for their realizations. Because of our definitions, we have:

Lemma 4.1. *Let $c \in [C]$ be a chunk and $\Upsilon_{\leq(c,L)}$ be arbitrary*[8]*. For all groups $(b,g) \notin \Gamma_{\leq(c,L)}$, we have that:*

$$\mathbb{D}\Big(\big(\mathcal{D} \mid \Upsilon_{\leq(c,L)}\big)_{b,g} \;||\; \mathcal{D}_{b,g}\Big) < \frac{1}{(\log n)^4} \qquad \textit{and} \qquad \big|\Gamma_{\leq(c,L)} \cap (\{b\} \times [G])\big| < \frac{G}{4}.$$

Proof. Observe from Equation (7) that if $\Phi_{(c,l)} = (\perp,\perp)$ for some $l \in \left(\widehat{T_{\mathsf{Ch}}}, L\right]$, then $\Phi_{(c,l')} = (\perp,\perp)$ for all $l' \in [l,L]$. On the other hand, if $\Phi_{(c,l)} \neq (\perp,\perp)$, then a new element is added to $\Gamma_{\leq(c,l)}$. As $L = \widehat{T_{\mathsf{Ch}}} + BG$, we have that $(b,g) \notin \Gamma_{\leq(c,L)}$ implies that $\Phi_{(c,L)} = (\perp,\perp)$. It follows that both conditions in Equation (7) are not satisfied for the group (b,g) and we have the lemma. □

A potential function. We now define a potential function and state our main lemma. For all blocks $b \in [B]$, inputs X, and subsets $\mathscr{G} \subseteq [B] \times [G]$ of groups, we define:

$$\Psi_b(X,\mathscr{G}) = \begin{cases} 1, & \text{if } (b, \mathsf{gkeys}_{b-1}(X)) \in \mathscr{G} \\ \min\Big(1, \frac{1}{(\log n)^{10}} \cdot |\mathscr{G} \cap (\{b\} \times [G])|\Big), & \text{otherwise} \end{cases}. \tag{10}$$

[8] Recall that fixing $\Upsilon_{\leq(c,L)}$ also fixes $\Phi_{\leq(c,l)}$ and $\Gamma_{\leq(c,L)}$.

Our main lemmas are:

Lemma 4.2. *For all $0 \le c \le C$, we have:*

$$\sum_{b=1}^{B} \mathbb{E}\big[\Psi_b\big(\mathsf{X}, \Gamma_{\le(c,L)}\big)\big] \le 2c + \frac{1}{(\log n)^8} \cdot \mathbb{E}\big[\big|\Gamma_{\le(c,L)}\big|\big].$$

Lemma 4.3. *For all $0 \le c \le C$, we have:*

$$\mathbb{E}\big[\big|\Gamma_{\le(c,L)}\big|\big] \le 20c \cdot (\log n)^7.$$

Before showing Lemmas 4.2 and 4.3, we first show why they imply Theorem 3.2. The rest of this section is dedicated to proving Lemma 4.2.

Proof of Theorem 3.2 assuming Lemmas 4.2 and 4.3. From Lemmas 4.2 and 4.3, we have that $\sum_{b=1}^{B} \mathbb{E}\big[\Psi_b\big(\mathsf{X}, \Gamma_{\le(C,L)}\big)\big] \le 3C \le \frac{B}{500}$. It follows that there exists $b \in [B]$ such that $\Pr\big(\Psi_b\big(\mathsf{X}, \Gamma_{\le(C,L)}\big) = 1\big) \le \frac{1}{500}$. From Equation (10), we have that $\Pr\big((b, \mathsf{gkeys}_{b-1}(\mathsf{X})) \in \Gamma_{\le(C,L)}\big) \le \frac{1}{500}$. For the rest of this proof, we fix this b and define the event $\mathcal{E}$ to be the event that $(b, \mathsf{gkeys}_{b-1}(\mathsf{X})) \in \Gamma_{\le(C,L)}$. Note that whether or not $\mathcal{E}$ occurs is determined by the pair $\big(\mathsf{X}_{<b}, \Upsilon_{\le(C,L)}\big)$ allowing us to view $\mathcal{E}$ as the set of pairs for which it occurs. By a union bound, we have:

$$\begin{aligned}
&\Pr\big(\mathsf{out}_\Pi(\mathsf{X}, \mathsf{N}) = \mathsf{gkeys}_{[B]}(\mathsf{X})\big) \\
&\le \frac{1}{500} + \sum_{\big(X_{<b}, \Upsilon_{\le(C,L)}\big) \notin \mathcal{E}} \Pr\big(X_{<b}, \Upsilon_{\le(C,L)}\big) \cdot \Pr\big(\mathsf{out}_\Pi(\mathsf{X}, \mathsf{N}) = \mathsf{gkeys}_{[B]}(\mathsf{X}) \mid X_{<b}, \Upsilon_{\le(C,L)}\big).
\end{aligned}$$

It suffices to bound each conditional probability term above by 0.05. Fix an arbitrary $\big(X_{<b}, \Upsilon_{\le(C,L)}\big) \notin \mathcal{E}$ and consider the corresponding term above. Note that fixing $X_{<b}$ fixes the value of $\mathsf{gkeys}_{b-1}(\mathsf{X})$ and let g denote the fixed value. Moreover, observe that fixing $\Upsilon_{\le(C,L)}$ fixes the transcript received by all the parties during the execution of the protocol (and thus the output) and that, conditioned on $\Upsilon_{\le(C,L)}$, the random variable $\mathsf{gkey}_{b,g}$ is independent of $\mathsf{X}_{<b}$. (This is formally shown in Lemma 4.16 below.) Letting k be the b-th coordinate of this fixed output, we have by Equation (4) that:

$$\begin{aligned}
\Pr\big(\mathsf{out}_\Pi(\mathsf{X}, \mathsf{N}) = \mathsf{gkeys}_{[B]}(\mathsf{X}) \mid X_{<b}, \Upsilon_{\le(C,L)}\big) &\le \Pr\big(\mathsf{gkey}_{b,g} = k \mid X_{<b}, \Upsilon_{\le(C,L)}\big) \\
&= \Pr\big(\mathsf{gkey}_{b,g} = k \mid \Upsilon_{\le(C,L)}\big).
\end{aligned}$$

As $\mathcal{E}$ does not occur, we have $(b, g) \notin \Gamma_{\le(C,L)}$ and by Lemma 4.1, it follows that we have $\mathbb{D}\Big(\big(\mathcal{D} \mid \Upsilon_{\le(C,L)}\big)_{b,g} \;||\; \mathcal{D}_{b,g}\Big) < \frac{1}{(\log n)^4}$. From Fact A.15, this gives that for all $k \in [G]$, we have:

$$\Pr\big(\mathsf{out}_\Pi(\mathsf{X}, \mathsf{N}) = \mathsf{gkeys}_{[B]}(\mathsf{X}) \mid X_{<b}, \Upsilon_{\le(C,L)}\big) \le \Pr\big(\mathsf{gkey}_{b,g} = k \mid X_{<b}, \Upsilon_{\le(C,L)}\big)$$

$$\leq \Pr\big(\mathsf{gkey}_{b,g} = k\big) + \frac{1}{(\log n)^2}$$
$$\leq 0.05.$$

□

4.1 Technical Lemmas

Lemma 4.4. *Let $0 \leq c \leq C$ and $\Upsilon_{\leq(c,L)}$ be an arbitrary realization[9] of $\Upsilon_{\leq(c,L)}$. It holds that:*

$$\frac{1}{5} \cdot \left|\Gamma_{\leq(c,L)}\right| - c \cdot \widehat{T_{\mathsf{Ch}}} \leq (\log n)^4 \cdot \sum_{c'=1}^{c} \sum_{l=\widehat{T_{\mathsf{Ch}}}+1}^{L} \mathbb{D}\Big(\big(\mathcal{D} \mid \Upsilon_{<(c',l)}\big)_{\Phi_{(c',l),1}} \;||\; \mathcal{D}_{\Phi_{(c',l),1}}\Big).$$

Proof. We analyze how the set $\Gamma_{\leq(c',l)}$ changes. Define the set:

$$\mathscr{G} = \big\{(c',l) \in [c] \times [L] \mid \Phi_{(c',l),1} \notin \Gamma_{<(c',l)} \cup \{\perp\}\big\},$$

and note from Equation (9) that $|\mathscr{G}| = \left|\Gamma_{\leq(c,L)}\right|$. For every $(c',l) \in \mathscr{G}$, if $l \notin \left[\widehat{T_{\mathsf{Ch}}}\right]$, we have that one of the two conditions in Equation (7) must hold. We define:

$$\mathscr{G}' = \left\{(c',l) \in \mathscr{G} \mid l \notin \left[\widehat{T_{\mathsf{Ch}}}\right] \wedge \left|\Gamma_{<(c',l)} \cap \big(\{\Phi_{(c',l),1,1}\} \times [G]\big)\right| \geq \frac{G}{4}\right\},$$

to be the subset of $\mathscr{G}$, where the second condition holds. Now, observe that, for any $(c',l) \in \mathscr{G}'$, there exists $\frac{G}{4}$ values in $\mathscr{G} \setminus \mathscr{G}'$ that are both determined by $\Phi_{(c',l),1,1}$ and different for different values of $\Phi_{(c',l),1,1}$. This means that $\frac{1}{4} \cdot |\mathscr{G}'| \leq |\mathscr{G} \setminus \mathscr{G}'|$. It follows that:

$$\frac{1}{5} \cdot \left|\Gamma_{\leq(c,L)}\right| = \frac{1}{5} \cdot |\mathscr{G}| \leq |\mathscr{G} \setminus \mathscr{G}'|.$$

Finally, note that for all $(c',l) \in \mathscr{G} \setminus \mathscr{G}'$ for which $l \notin \left[\widehat{T_{\mathsf{Ch}}}\right]$, the first condition in Equation (7) must hold. This means that for all these (c',l), we have:

$$\mathbb{D}\Big(\big(\mathcal{D} \mid \Upsilon_{<(c',l)}\big)_{\Phi_{(c',l),1}} \;||\; \mathcal{D}_{\Phi_{(c',l),1}}\Big) \geq \frac{1}{(\log n)^4}.$$

As the number of such values is at least $|\mathscr{G} \setminus \mathscr{G}'| - c \cdot \widehat{T_{\mathsf{Ch}}}$ and the KL divergence is non-negative (Lemma A.11), the lemma follows. □

Lemma 4.5. *For all blocks $b \in [B]$, inputs X, and subsets $\mathscr{G}, \mathscr{G}' \subseteq [B] \times [G]$ of groups, we have:*

$$\Psi_b(X, \mathscr{G} \cup \mathscr{G}') \leq \Psi_b(X, \mathscr{G}) + \Psi_b(X, \mathscr{G}').$$

[9] Observe that fixing any such realization fixes the value of $\Gamma_{\leq(c,L)}$ and $\Phi_{\leq(c,L)}$.

Proof. This is straightforward if $(b, \mathsf{gkeys}_{b-1}(X)) \in \mathscr{G} \cup \mathscr{G}'$ so we assume otherwise. We get:

$$\begin{aligned}
\Psi_b(X, \mathscr{G} \cup \mathscr{G}') &= \min\left(1, \frac{1}{(\log n)^{10}} \cdot |(\mathscr{G} \cup \mathscr{G}') \cap (\{b\} \times [G])|\right) \\
&\leq \min\left(1, \frac{1}{(\log n)^{10}} \cdot |\mathscr{G} \cap (\{b\} \times [G])| + \frac{1}{(\log n)^{10}} \cdot |\mathscr{G}' \cap (\{b\} \times [G])|\right) \\
&\leq \min\left(1, \frac{1}{(\log n)^{10}} \cdot |\mathscr{G} \cap (\{b\} \times [G])|\right) + \min\left(1, \frac{1}{(\log n)^{10}} \cdot |\mathscr{G}' \cap (\{b\} \times [G])|\right) \\
&\leq \Psi_b(X, \mathscr{G}) + \Psi_b(X, \mathscr{G}').
\end{aligned}$$

□

Lemma 4.6. *For all blocks $b \in [B]$, inputs X, and subsets $\mathscr{G} \subseteq [B] \times [G]$ of groups, we have:*

$$\Psi_b(X, \mathscr{G}) \leq \mathbb{1}((b, \mathsf{gkeys}_{b-1}(X)) \in \mathscr{G}) + \frac{1}{(\log n)^{10}} \cdot |\mathscr{G} \cap (\{b\} \times [G])|.$$

Proof. Direct from Equation (10). □

Corollary 4.7 (Corollary of Lemmas 4.5 and 4.6)**.** *For all blocks $b \in [B]$, inputs X, and subsets $\mathscr{G}, \mathscr{G}' \subseteq [B] \times [G]$ of groups, we have:*

$$\Psi_b(X, \mathscr{G} \cup \mathscr{G}') - \Psi_b(X, \mathscr{G}) \leq \mathbb{1}((b, \mathsf{gkeys}_{b-1}(X)) \in \mathscr{G}' \setminus \mathscr{G}) + \frac{1}{(\log n)^{10}} \cdot |(\mathscr{G}' \setminus \mathscr{G}) \cap (\{b\} \times [G])|.$$

4.1.1 The Marginal Distribution for One Group

Throughout this section, for all $s \in [S]$, we define the set $\mathscr{X}_{|s} = \{s\} \times [G]$. Thus, we have $\overline{\mathscr{X}_{|s}} = ([S] \setminus \{s\}) \times [G]$. Also, for all groups (b, g) and all $s' \in [S]$, and $j' \in [m]$, we define the event $\mathcal{E}^{b,g}_{s',j'}$ to be the event that $\mathsf{s}_{b,g,0} = s' \wedge \mathsf{uq}_{b,g} = j'$.

Lemma 4.8. *Let (b, g) be a group, $s' \in [S]$, and $j' \in [m]$. We have $\Pr\left(\mathcal{E}^{b,g}_{s',j'}\right) = \frac{1}{mS}$. Further, if for all $j \in [0, m]$, we have a set of inputs $\mathscr{X}_j \subseteq [S] \times [G]$, then:*

$$\Pr\left(\forall j \in [0, m] : \mathsf{x}_{b,g,j} \in \mathscr{X}_j \mid \mathcal{E}^{b,g}_{s',j'}\right) = \frac{\left|\mathscr{X}_0 \cap \mathscr{X}_{|s'}\right|}{G} \cdot \frac{\left|\mathscr{X}_{j'} \cap \mathscr{X}_{|s'}\right|}{G} \cdot \prod_{j \neq j' \in [m]} \frac{\left|\mathscr{X}_j \cap \overline{\mathscr{X}_{|s'}}\right|}{G \cdot (S-1)}.$$

Moreover, conditioned on $\mathcal{E}^{b,g}_{s',j'}$, the random variables $(\mathsf{X}_{b,g,j})_{j \neq j' \in [m]}$ are mutually independent and are independent of all the other inputs.

Proof. The first part is by symmetry. For the second part, note that Items 1 and 2 of our promise on the distribution $\mathcal{D}$ implies that sampling from $\mathcal{D}$ conditioned on $\mathcal{E}^{b,g}_{s',j'}$ is the same as sampling from the uniform distribution conditioned on $\mathsf{s}_{b,g,0} = \mathsf{s}_{b,g,j'} = s'$ and $\mathsf{s}_{b,g,j} \neq s'$ for all $j \neq j' \in [m]$. For the "moreover" part, note that Item 2 of our promise only depends on the input of player (b, g, j') in group (b, g). □

Lemma 4.9. *Let (b, g) be a group and $j \in [0, m]$. Let $\mathscr{X}_j \subseteq [S] \times [G]$ be a set of inputs. We have:*

$$\Pr(\mathsf{x}_{b,g,j} \in \mathscr{X}_j) \leq \frac{1}{GS} \cdot |\mathscr{X}_j|.$$

Proof. If $j = 0$, we have from Lemma 4.8 that:

$$\Pr(\mathsf{x}_{b,g,0} \in \mathscr{X}_0) \leq \frac{1}{S} \cdot \sum_{s'=1}^{S} \frac{\left|\mathscr{X}_0 \cap \mathscr{X}_{|s'}\right|}{G} = \frac{|\mathscr{X}_0|}{GS}.$$

Otherwise, we have from Lemma 4.8 that:

$$\begin{aligned}
\Pr(\mathsf{x}_{b,g,j} \in \mathscr{X}_j) &\leq \frac{1}{mS} \cdot \sum_{s'=1}^{S} \left(\frac{\left|\mathscr{X}_j \cap \mathscr{X}_{|s'}\right|}{G} + (m-1) \cdot \frac{\left|\mathscr{X}_j \cap \overline{\mathscr{X}_{|s'}}\right|}{G \cdot (S-1)} \right) \\
&\leq \frac{1}{mGS} \cdot |\mathscr{X}_j| + \frac{m-1}{mGS} \cdot |\mathscr{X}_j| \\
&\leq \frac{1}{GS} \cdot |\mathscr{X}_j|.
\end{aligned}$$

□

Lemma 4.10. *Let (b, g) be a group and for all $j \in [m]$, let $\mathscr{X}_j \subseteq [S] \times [G]$ be a set of inputs. We have:*

$$2 \cdot \ln\left(\frac{S}{S-1} \cdot \Pr(\forall j \in [m] : \mathsf{x}_{b,g,j} \in \mathscr{X}_j) - \frac{1}{S-1} \right) \leq 2 - \frac{1}{GS} \cdot \sum_{j=1}^{m} |\overline{\mathscr{X}_j}|.$$

Proof. We have:

$$\begin{aligned}
\Pr(\forall j \in [m] : \mathsf{x}_{b,g,j} \in \mathscr{X}_j) &= \sum_{s'=1}^{S} \sum_{j'=1}^{m} \Pr\left(\mathcal{E}^{b,g}_{s',j'}\right) \cdot \Pr\left(\forall j \in [m] : \mathsf{x}_{b,g,j} \in \mathscr{X}_j \mid \mathcal{E}^{b,g}_{s',j'}\right) \\
&\leq \frac{1}{mS} \cdot \sum_{s'=1}^{S} \sum_{j'=1}^{m} \prod_{j \neq j' \in [m]} \frac{\left|\mathscr{X}_j \cap \overline{\mathscr{X}_{|s'}}\right|}{G \cdot (S-1)} && \text{(Lemma 4.8)} \\
&\leq \frac{1}{mS} \cdot \sum_{s'=1}^{S} \sum_{j'=1}^{m} \mathrm{e}^{-\sum_{j \neq j' \in [m]} \frac{\left|\overline{\mathscr{X}_{|s'}} \setminus \mathscr{X}_j\right|}{G \cdot (S-1)}} && \text{(As } x \leq \mathrm{e}^{x-1} \text{ for all } x\text{)} \\
&\leq \frac{1}{S} \cdot \sum_{s'=1}^{S} \max_{j' \in [m]} \mathrm{e}^{-\sum_{j \neq j' \in [m]} \frac{\left|\overline{\mathscr{X}_{|s'}} \setminus \mathscr{X}_j\right|}{G \cdot (S-1)}}.
\end{aligned}$$

To continue, for all s', let $\tau(s')$ be the term corresponding to s' above and note that $\tau(s') \leq 1$ for all $s' \in [S]$. Order all the s' in decreasing order of $\tau(s')$ (breaking ties arbitrarily). For

both s' in the first two positions in this order we have:

$$\begin{aligned}\Pr(\forall j \in [m] : \mathsf{x}_{b,g,j} \in \mathscr{X}_j) &\leq \frac{1}{S} + \frac{S-1}{S} \cdot \sum_{s'=1}^{S} \max_{j' \in [m]} \mathrm{e}^{-\sum_{j \neq j' \in [m]} \frac{\left|\overline{\mathscr{X}_{|s'}} \setminus \mathscr{X}_j\right|}{G \cdot (S-1)}} \\ &\leq \frac{1}{S} + \frac{S-1}{S} \cdot \mathrm{e}^{1-\sum_{j=1}^{m} \frac{\left|\overline{\mathscr{X}_{|s'}} \setminus \mathscr{X}_j\right|}{G \cdot (S-1)}} \\ &\leq \frac{1}{S} + \frac{S-1}{S} \cdot \mathrm{e}^{1-\sum_{j=1}^{m} \frac{\left|\overline{\mathscr{X}_{|s'}} \setminus \mathscr{X}_j\right|}{GS}}.\end{aligned}$$

Rearranging and adding for both these s', we get:

$$2 \cdot \ln\left(\frac{S}{S-1} \cdot \Pr(\forall j \in [m] : \mathsf{x}_{b,g,j} \in \mathscr{X}_j) - \frac{1}{S-1}\right) \leq 2 - \frac{1}{GS} \cdot \sum_{j=1}^{m} \left|\overline{\mathscr{X}_j}\right|.$$

□

Lemma 4.11. *Let (b, g) be a group and for all $j \in [0, m]$, let $\mathscr{X}_j \subseteq [S] \times [G]$ be a set of inputs. Suppose that $\sum_{j=1}^{m} \left|\overline{\mathscr{X}_j}\right| \leq 5 \cdot GS$ and that $|\mathscr{X}_j| \geq \frac{GS}{2}$ for all $j \in [0, m]$. Then, for all $g' \in [G]$ and all sets $\mathscr{S} \subseteq [S]$ with $|\mathscr{S}| \geq \frac{3S}{4}$, we have:*

$$\mathbb{H}_\infty\left(\mathsf{uq}_{b,g} \mid \mathsf{s}_{b,g,0} \in \mathscr{S} \wedge \mathsf{gkey}_{b,g} = g' \wedge \forall j \in [0, m] : \mathsf{x}_{b,g,j} \in \mathscr{X}_j\right) \geq 0.9 \cdot \log m.$$

Proof. Note that, for all $j' \in [m]$, we have:

$$\begin{aligned}&\Pr\left(\mathsf{uq}_{b,g} = j' \wedge \mathsf{s}_{b,g,0} \in \mathscr{S} \wedge \mathsf{gkey}_{b,g} = g' \wedge \forall j \in [0, m] : \mathsf{x}_{b,g,j} \in \mathscr{X}_j\right) \\ &\quad = \sum_{s' \in \mathscr{S}} \Pr\left(\mathcal{E}^{b,g}_{s',j'} \wedge \mathsf{gkey}_{b,g} = g' \wedge \forall j \in [0, m] : \mathsf{x}_{b,g,j} \in \mathscr{X}_j\right) \\ &\quad = \frac{1}{mS} \cdot \sum_{s' \in \mathscr{S}} \frac{\left|\mathscr{X}_0 \cap \mathscr{X}_{|s'}\right|}{G} \cdot \frac{\mathbb{1}((s', g') \in \mathscr{X}_{j'})}{G} \cdot \prod_{j \neq j' \in [m]} \frac{\left|\mathscr{X}_j \cap \overline{\mathscr{X}_{|s'}}\right|}{G \cdot (S-1)}. \qquad \text{(Lemma 4.8)}\end{aligned}$$

Now, as $|\mathscr{X}_j| \geq \frac{GS}{2}$ for all $j \in [0, m]$ and $S = (\log n)^{10}$, we have that

$$\left|\mathscr{X}_j \cap \overline{\mathscr{X}_{|s'}}\right| \geq \left|\overline{\mathscr{X}_{|s'}}\right| \cdot \left(1 - \frac{\left|\overline{\mathscr{X}_j}\right|}{\left|\overline{\mathscr{X}_{|s'}}\right|}\right) \geq \left|\overline{\mathscr{X}_{|s'}}\right| \cdot \mathrm{e}^{-2 \cdot \frac{\left|\overline{\mathscr{X}_j}\right|}{\left|\overline{\mathscr{X}_{|s'}}\right|}} \geq \left|\overline{\mathscr{X}_{|s'}}\right| \cdot \mathrm{e}^{-\frac{4}{GS} \cdot \left|\overline{\mathscr{X}_j}\right|}.$$

This is because $1 - x \geq \mathrm{e}^{-2x}$ for all $0 \leq x \leq 0.75$. From this, we have:

$$\begin{aligned}&\Pr\left(\mathsf{uq}_{b,g} = j' \wedge \mathsf{s}_{b,g,0} \in \mathscr{S} \wedge \mathsf{gkey}_{b,g} = g' \wedge \forall j \in [0, m] : \mathsf{x}_{b,g,j} \in \mathscr{X}_j\right) \\ &\qquad \geq \frac{1}{mSG^2} \cdot \mathrm{e}^{-\frac{4}{GS} \cdot \sum_{j=1}^{m} \left|\overline{\mathscr{X}_j}\right|} \cdot \sum_{s' \in \mathscr{S}} \mathbb{1}((s', g') \in \mathscr{X}_{j'}) \cdot \left|\mathscr{X}_0 \cap \mathscr{X}_{|s'}\right|.\end{aligned}$$

Using this lower bound, we get for all $\tilde{m} \in [m]$ that:

$$\begin{aligned}
&\Pr\big(\mathsf{uq}_{b,g} = \tilde{m} \mid \mathsf{s}_{b,g,0} \in \mathscr{S} \wedge \mathsf{gkey}_{b,g} = g' \wedge \forall j \in [0,m] : \mathsf{x}_{b,g,j} \in \mathscr{X}_j\big) \\
&\quad \le \frac{\Pr\big(\mathsf{uq}_{b,g} = \tilde{m}\big)}{\sum_{j'=1}^{m} \Pr\big(\mathsf{uq}_{b,g} = j' \wedge \mathsf{s}_{b,g,0} \in \mathscr{S} \wedge \mathsf{gkey}_{b,g} = g' \wedge \forall j \in [0,m] : \mathsf{x}_{b,g,j} \in \mathscr{X}_j\big)} \\
&\quad \le \frac{SG^2}{\sum_{j'=1}^{m} \sum_{s' \in \mathscr{S}} \mathbb{1}((s',g') \in \mathscr{X}_{j'}) \cdot \left|\mathscr{X}_0 \cap \mathscr{X}_{|s'}\right|} \cdot \mathrm{e}^{\frac{4}{GS} \cdot \sum_{j=1}^{m} \left|\overline{\mathscr{X}_j}\right|} \\
&\quad \le \frac{2 \cdot SG^2}{m \cdot \sum_{s' \in \mathscr{S}} \left|\mathscr{X}_0 \cap \mathscr{X}_{|s'}\right|} \cdot \mathrm{e}^{20} && (\text{As } \textstyle\sum_{j=1}^{m} \left|\overline{\mathscr{X}_j}\right| \le 5 \cdot GS) \\
&\quad \le \frac{G}{m} \cdot 2^{50}. && (\text{As } |\mathscr{S}| \ge \tfrac{3S}{4} \text{ and } |\mathscr{X}_0| \ge \tfrac{GS}{2})
\end{aligned}$$

The lemma follows. □

Lemma 4.12. *For all $i \in [n]$, let $\mathscr{X}_i \subseteq [S] \times [G]$ be a set of inputs and (b,g) be a group. Suppose that $\overline{\mathscr{X}_{b,g,j}} = \emptyset$ for at least $m - m^{2/3}$ many $j \in [0,m]$. Then, for all sets $J \subseteq [m]$ with $|J| \ge 10 \cdot m^{2/3}$ and all $s \ne s' \in [S]$, $k \in [G]$, and $j' \in [m]$, we have:*

$$\Pr\left(\sum_{j \in J} \mathbb{1}(\mathsf{x}_{b,g,j} = (s,k)) < m^{1/3} \mid \mathcal{E}^{b,g}_{s',j'} \wedge \forall i \in [n] : \mathsf{x}_i \in \mathscr{X}_i\right) \le \mathrm{e}^{-\frac{1}{8} \cdot m^{1/3}}.$$

Proof. By our assumptions, there exists a subset $J' \subseteq J$ with $|J'| \ge 8 \cdot m^{2/3}$ such that $j' \notin J'$ and $\overline{\mathscr{X}_{b,g,j}} = \emptyset$ for all $j \in J'$. We have:

$$\begin{aligned}
&\Pr\left(\sum_{j \in J} \mathbb{1}(\mathsf{x}_{b,g,j} = (s,k)) < m^{1/3} \mid \mathcal{E}^{b,g}_{s',j'} \wedge \forall i \in [n] : \mathsf{x}_i \in \mathscr{X}_i\right) \\
&\quad \le \Pr\left(\sum_{j \in J'} \mathbb{1}(\mathsf{x}_{b,g,j} = (s,k)) < m^{1/3} \mid \mathcal{E}^{b,g}_{s',j'} \wedge \forall i \in [n] : \mathsf{x}_i \in \mathscr{X}_i\right) \\
&\quad \le \Pr\left(\sum_{j \in J'} \mathbb{1}(\mathsf{x}_{b,g,j} = (s,k)) < m^{1/3} \mid \mathcal{E}^{b,g}_{s',j'}\right) && (\text{Lemma 4.8}) \\
&\quad \le \mathrm{e}^{-\frac{1}{8} \cdot m^{1/3}} && (\text{As } S = (\log n)^{10} \text{ and Lemmas 3.1 and 4.8})
\end{aligned}$$

□

4.1.2 Information Theory Lemmas

Lemma 4.13. *Let $n > 0$ and $\mathsf{X} = (\mathsf{X}_1, \ldots, \mathsf{X}_n)$ be identically distributed (possibly correlated) random variables and $\mathscr{H} = \log(|\mathsf{supp}(\mathsf{X}_1)|)$. Let Y be another random variable such that for all $y \in \mathsf{supp}(\mathsf{Y})$, there exists a set $S_y \subseteq [n]$ such that setting $\mathsf{Y} = y$ fixes the value of X_i for*

all $i \in S_y$. We have:

$$\mathbb{H}(\mathsf{X} \mid \mathsf{Y}) \leq \mathscr{H} \cdot (n - \mathbb{E}[|S_y|]).$$

Proof. We have:

$$\begin{aligned}
\mathbb{H}(\mathsf{X} \mid \mathsf{Y}) &= \sum_{y \in \mathsf{supp}(\mathsf{Y})} \Pr(y) \cdot \mathbb{H}(\mathsf{X} \mid y) && \text{(Definition A.2)} \\
&= \sum_{y \in \mathsf{supp}(\mathsf{Y})} \Pr(y) \cdot \mathbb{H}\Big(\mathsf{X}_{\overline{S_y}} \mid y\Big) \\
&\leq \sum_{y \in \mathsf{supp}(\mathsf{Y})} \Pr(y) \cdot \sum_{i \notin S_y} \mathbb{H}(\mathsf{X}_i \mid y) && \text{(Corollary A.5)} \\
&\leq \sum_{y \in \mathsf{supp}(\mathsf{Y})} \Pr(y) \cdot \mathscr{H} \cdot (n - |S_y|) \\
& && \text{(Lemma A.6 and that } \mathsf{X}_i \text{ are identically distributed)} \\
&= \mathscr{H} \cdot (n - \mathbb{E}[|S_y|]).
\end{aligned}$$

□

Lemma 4.14. *Let $n > 0$ and $\mathsf{X} = (\mathsf{X}_1, \ldots, \mathsf{X}_n)$ be identically distributed (possibly correlated) random variables and $\mathscr{H} = \log(|\mathsf{supp}(\mathsf{X}_1)|)$. Let $\mathsf{S}, \mathsf{T} \subseteq [n]$ be (set-valued) random variables such that $\mathsf{S} \subseteq \mathsf{T}$ almost surely. Let Y, Z be random variables such that Z determines T and for all $y \in \mathsf{supp}(\mathsf{Y})$, there exists a set $S_y \subseteq [n]$ such that setting $\mathsf{Y} = y$ implies $\mathsf{S} = S_y$ and fixes the value of X_i for all $i \in S_y$. We have:*

$$\mathbb{H}(\mathsf{X} \mid \mathsf{Y}, \mathsf{Z}) \leq \mathscr{H} \cdot \mathbb{E}[|\mathsf{T}| - |\mathsf{S}|] + \sum_{y,z} \Pr(y, z) \cdot \mathbb{H}(\mathsf{X}_{\overline{T}} \mid y, z).$$

Proof. We have:

$$\begin{aligned}
\mathbb{H}(\mathsf{X} \mid \mathsf{Y}, \mathsf{Z}) &= \sum_{y,z} \Pr(y, z) \cdot \mathbb{H}(\mathsf{X} \mid y, z) && \text{(Definition A.2)} \\
&= \sum_{y,z} \Pr(y, z) \cdot \mathbb{H}\Big(\mathsf{X}_{\overline{S_y}} \mid y, z\Big) \\
&\leq \sum_{y,z} \Pr(y, z) \cdot \mathbb{H}\big(\mathsf{X}_{T \setminus S_y} \mid y, z\big) + \sum_{y,z} \Pr(y, z) \cdot \mathbb{H}(\mathsf{X}_{\overline{T}} \mid y, z) && \text{(Corollary A.5)} \\
&\leq \mathscr{H} \cdot \mathbb{E}[|\mathsf{T}| - |\mathsf{S}|] + \sum_{y,z} \Pr(y, z) \cdot \mathbb{H}(\mathsf{X}_{\overline{T}} \mid y, z). && \text{(Lemma A.6)}
\end{aligned}$$

□

4.1.3 Lemmas about Our Random Variables

Lemma 4.15. *Let $\mathscr{G} \subseteq [B] \times [G]$ be a set of groups and $\left(\mathsf{gk}_{b,g}\right)_{(b,g)\in\mathscr{G}}$ be elements of $[G]$. Define the event $\mathcal{E}$ to be the event that for all $(b,g) \in \mathscr{G}$, we have $\mathsf{gkey}_{b,g} = \mathsf{gk}_{b,g}$. The $(|\mathscr{G}|+1)$ random variables $\left(\left(\mathsf{X}_{b,g}\right)_{(b,g)\in\mathscr{G}}, \mathsf{X}_{\overline{\mathscr{G}}}\right)$ are mutually independent conditioned on $\mathcal{E}$.*

Proof. Proof by induction on $|\mathscr{G}|$. The base case $|\mathscr{G}| = 0$ is trivial. For the inductive case, consider an arbitrary $\mathscr{G}$ with $|\mathscr{G}| > 0$. Let (b^*, g^*) be the smallest element of $\mathscr{G}$ and define $\mathscr{G}' = \mathscr{G} \setminus \{(b^*, g^*)\}$. Define the event $\mathcal{E}'$ to be the event that for all $(b,g) \in \mathscr{G}'$, we have $\mathsf{gkey}_{b,g} = \mathsf{gk}_{b,g}$ and the event $\mathcal{E}^*$ to be the event that $\mathsf{gkey}_{b^*,g^*} = \mathsf{gk}_{b^*,g^*}$ so that we have $\mathcal{E} = \mathcal{E}' \wedge \mathcal{E}^*$. First, note that for all $(b,g) \in \mathscr{G}'$, we have:

$$\Pr(X_{b,g} \mid \mathcal{E}) = \Pr(X_{b,g} \mid \mathcal{E}', \mathcal{E}^*) = \Pr(X_{b,g} \mid \mathcal{E}'),$$

as whether or not $\mathcal{E}^*$ occurs is determined by X_{b^*,g^*} which is independent of $\mathsf{X}_{b,g}$ conditioned on $\mathcal{E}'$ (by our induction hypothesis). Next, note that:

$$\begin{aligned}
\Pr(X \mid \mathcal{E}) &= \Pr(X \mid \mathcal{E}', \mathcal{E}^*) \\
&= \frac{\Pr(X, \mathcal{E}^* \mid \mathcal{E}')}{\Pr(\mathcal{E}^* \mid \mathcal{E}')} \\
&= \frac{\Pr(X_{b^*,g^*}, \mathcal{E}^* \mid \mathcal{E}') \cdot \Pr(X_{\overline{\mathscr{G}}} \mid X_{b^*,g^*}, \mathcal{E}^*, \mathcal{E}')}{\Pr(\mathcal{E}^* \mid \mathcal{E}')} \cdot \prod_{(b,g)\in\mathscr{G}'} \Pr(X_{b,g} \mid \mathcal{E}') \\
&\quad \text{(Induction hypothesis and whether } \mathcal{E}^* \text{ occurs is determined by } \mathsf{X}_{b^*,g^*}) \\
&= \frac{\Pr(X_{b^*,g^*}, \mathcal{E}^* \mid \mathcal{E}') \cdot \Pr(X_{\overline{\mathscr{G}}} \mid X_{b^*,g^*}, \mathcal{E}^*, \mathcal{E}')}{\Pr(\mathcal{E}^* \mid \mathcal{E}')} \cdot \prod_{(b,g)\in\mathscr{G}'} \Pr(X_{b,g} \mid \mathcal{E}).
\end{aligned}$$

Now, note that our distribution $\mathcal{D}$ has the property that, conditioned on $X_{b^*,g^*}, \mathcal{E}^*, \mathcal{E}'$, the random variable $\mathsf{X}_{\overline{\mathscr{G}}}$ is just a uniformly random element of set that is independent of X_{b^*,g^*}. We get:

$$\begin{aligned}
\Pr(X \mid \mathcal{E}) &= \frac{\Pr(X_{b^*,g^*}, \mathcal{E}^* \mid \mathcal{E}') \cdot \Pr(X_{\overline{\mathscr{G}}} \mid \mathcal{E}^*, \mathcal{E}')}{\Pr(\mathcal{E}^* \mid \mathcal{E}')} \cdot \prod_{(b,g)\in\mathscr{G}'} \Pr(X_{b,g} \mid \mathcal{E}) \\
&= \Pr(X_{b^*,g^*} \mid \mathcal{E}) \cdot \Pr(X_{\overline{\mathscr{G}}} \mid \mathcal{E}) \cdot \prod_{(b,g)\in\mathscr{G}'} \Pr(X_{b,g} \mid \mathcal{E}) \\
&= \Pr(X_{\overline{\mathscr{G}}} \mid \mathcal{E}) \cdot \prod_{(b,g)\in\mathscr{G}} \Pr(X_{b,g} \mid \mathcal{E}).
\end{aligned}$$

□

Lemma 4.16. *Let $c \in [C]$ and $l \in [L]$. For any realization $\Upsilon_{\leq(c,l)}$ of $\Upsilon_{\leq(c,l)}$ there exists a noise vector $N_{\leq(c,\min(\lceil l/2\rceil, T_{\mathsf{Ch}}))}$ and sets $(\mathscr{X}_i)_{i\in[n]}$ such that the following equivalence between*

events holds:

$$\Upsilon_{\leq(c,l)} = \Upsilon_{\leq(c,l)} \equiv \mathsf{N}_{\leq(c,\min(\lceil l/2\rceil,T_{\mathsf{Ch}}))} = N_{\leq(c,\min(\lceil l/2\rceil,T_{\mathsf{Ch}}))} \wedge \forall i \in [n] : \mathsf{x}_i \in \mathscr{X}_i.$$

Moreover, fixing any realization $\Upsilon_{\leq(c,l)}$ also fixes the transcript received by all the parties in the first $(c, \min(\lfloor l/2\rfloor, T_{\mathsf{Ch}}))$ rounds.

Proof. Proof by induction on (c, l). The base case is straightforward. We show the result for (c, l) assuming it holds for smaller values. First, consider the case that $l \in \left(\widehat{T_{\mathsf{Ch}}}, L\right]$. In this case, note from Equation (8) that, conditioned on $\Upsilon_{<(c,l)}$, the first coordinate of $\Phi_{(c,l)}$ is already fixed. If this fixed value is $\perp$, the second coordinate is also fixed to $\perp$ and we are done by the induction hypothesis. Otherwise, letting this fixed value be (b, g), we get that for all $\Upsilon_{\leq(c,l)}$, there exists a value $X_{b,g}$ such that conditioned on $\Upsilon_{<(c,l)}$, the event $\Upsilon_{(c,l)} = \Upsilon_{(c,l)}$ is the same as $\mathsf{X}_{b,g} = X_{b,g}$. We get:

$$\begin{aligned}\Upsilon_{\leq(c,l)} = \Upsilon_{\leq(c,l)} &\equiv \Upsilon_{<(c,l)} = \Upsilon_{<(c,l)} \wedge \Upsilon_{(c,l)} = \Upsilon_{(c,l)} \\ &\equiv \Upsilon_{<(c,l)} = \Upsilon_{<(c,l)} \wedge \mathsf{X}_{b,g} = X_{b,g}.\end{aligned}$$

We are done by the induction hypothesis which also implies the "moreover" part. Now, consider the case $l \in \left[\widehat{T_{\mathsf{Ch}}}\right]$. In this case, we assume that l is even as the proof when l is odd is analogous. Let (b, g) be the last group in $\Upsilon_{\leq(c,l)}$. That is, we have:

$$\Upsilon_{\leq(c,l)} = \left(\Upsilon_{\leq(c,l-1)}, \left(((b, g), X_{b,g}), N_{\leq(c,\min(\lceil l/2\rceil,T_{\mathsf{Ch}}))}\right)\right).$$

The group (b, g) may be $\perp$, in which case we assume it is the pair $(B + 1, G + 1)$. Observe from Equation (6) that the event $\Upsilon_{(c,l),1} = (b, g)$ is the same as saying that none of the parties in the groups between $\Phi_{(c,l-1),1}$ and (b, g) spoke in round $(c, l/2)$ and some party in group (b, g) did speak. As the transcript received by all that parties in $(c, l/2 - 1)$ rounds is fixed by the induction hypothesis, we can define a set $\mathscr{X}_i'$ for all parties i in groups between $\Phi_{(c,l-1),1}$ and (b, g), such that party i does not speak if and only if $\mathsf{x}_i \in \mathscr{X}_i'$. This takes care of the former condition, while the latter is subsumed by the event $\mathsf{X}_{b,g} = X_{b,g}$. Defining $\mathscr{X}_i'$ to be the set of all inputs for all other i, we have:

$$\begin{aligned}\Upsilon_{\leq(c,l)} = \Upsilon_{\leq(c,l)} &\equiv \Upsilon_{<(c,l)} = \Upsilon_{<(c,l)} \wedge \Upsilon_{(c,l)} = \Upsilon_{(c,l)} \\ &\equiv \Upsilon_{<(c,l)} = \Upsilon_{<(c,l)} \wedge \mathsf{X}_{b,g} = X_{b,g} \wedge \forall i \in [n] : \mathsf{x}_i \in \mathscr{X}_i'.\end{aligned}$$

We are done by the induction hypothesis. For the "moreover" part, observe that fixing $\Upsilon_{\leq(c,l)}$, fixes the symbols sent by the parties in groups upto (b, g) in the round $(c, l/2)$. As $\Upsilon_{\leq(c,l)}$ also fixes the noise in this round, the moreover part follows. □

Lemma 4.17. *For all $c \in [C]$ and $l \in [L]$, the random variable $\Upsilon_{\leq(c,l)}$ determines the random variables $\Phi^{\mathsf{sm}}_{\leq(c,l)}$ and $\Phi^{\mathsf{gk}}_{\leq(c,l)}$.*

Proof. Note by Lemma 4.16 that $\Upsilon_{\leq(c,l)}$ fixes the transcript received by all the parties in the first $(c, \min(\lfloor l/2 \rfloor, T_{\mathsf{Ch}}))$ rounds and from Equation (9) that $\Upsilon_{\leq(c,l)}$ fixes $\Phi_{\leq(c,l)}$. Combining, we get the result. □

Lemma 4.18. *Let $c \in [C]$ and $l \in \left[\widehat{T_{\mathsf{Ch}}}\right]$. For any $\Delta^{\mathsf{sm}}_{(c,l)} = \left(\Upsilon_{\leq(c-1,L)}, \Phi^{\mathsf{sm}}_{\leq(c,l)}, N_{\leq(c,\lceil l/2 \rceil)}\right)$, there exist sets $(\mathscr{X}_i)_{i\in[n]}$ such that, letting $\mathbf{\Delta}_{(c,l)}$ be the random variable associated with $\Delta^{\mathsf{sm}}_{(c,l)}$, the following equivalence between events holds:*

$$\mathbf{\Delta}_{(c,l)} = \Delta^{\mathsf{sm}}_{(c,l)} \equiv \mathsf{N}_{\leq(c,\lceil l/2 \rceil)} = N_{\leq(c,\lceil l/2 \rceil)} \wedge \forall i \in [n] : \mathsf{x}_i \in \mathscr{X}_i.$$

Moreover, fixing any $\Delta^{\mathsf{sm}}_{(c,l)}$ also fixes the transcript $\Pi^i_{\leq(c,\lfloor l/2 \rfloor)}$ received by all the parties $i \in [n]$ in the first $(c, \lfloor l/2 \rfloor)$ rounds. Finally, for all $i \in [n]$, the set $\mathscr{X}_i$ equals the set promised by Lemma 4.16 for $\Upsilon_{\leq(c-1,L)}$ unless there exists $l' \in [l]$ and inputs (s,k) such that $\mathsf{msg}_i\left((s,k), \Pi^i_{<(c,\lceil l'/2 \rceil)}\right) \neq \perp$ and either $\Phi^{\mathsf{sm}}_{(c,l'),1} = \perp$ or $i \leq \Phi^{\mathsf{sm}}_{(c,l'),1}$

Proof. Proof by induction on l. For the base case $l = 0$, note by Lemma 4.17 that $\Upsilon_{\leq(c-1,L)}$ determines $\Delta^{\mathsf{sm}}_{(c,l)}$ and we are done by Lemma 4.16. For the inductive step, we proceed similarly to the proof of Lemma 4.16. Assume that l is even as the proof when l is odd is analogous. Let $i = (b, g, j)$ be the last player in $\Phi^{\mathsf{sm}}_{\leq(c,l)}$. That is, we have:

$$\Phi^{\mathsf{sm}}_{\leq(c,l)} = \left(\Phi^{\mathsf{sm}}_{\leq(c,l-1)}, (i, x_i)\right).$$

The player i may be $\perp$, in which case we assume it is the player $n+1$. Observe from Equation (6) that the event $\Phi^{\mathsf{sm}}_{(c,l),1} = i$ is the same as saying that none of the parties between $\Phi^{\mathsf{sm}}_{(c,l-1),1}$ and i spoke in round $(c, l/2)$ and party i did speak. As the transcript received by all the parties in the first $(c, l/2-1)$ rounds is fixed by the induction hypothesis, we can define a set $\mathscr{X}'_{i'}$ for all parties i' between $\Phi^{\mathsf{sm}}_{(c,l-1),1}$ and i such that party i' does not speak if and only if $\mathsf{x}_{i'} \in \mathscr{X}'_{i'}$. Defining, $\mathscr{X}'_{i'}$ to be the set of all inputs for all other i', we have that:

$$\begin{aligned}\mathbf{\Delta}_{(c,l)} = \Delta^{\mathsf{sm}}_{(c,l)} &\equiv \mathbf{\Delta}_{(c,l-1)} = \Delta^{\mathsf{sm}}_{(c,l-1)} \wedge \mathbf{\Phi}^{\mathsf{sm}}_{(c,l)} = (i, x_i) \\ &\equiv \mathbf{\Delta}_{(c,l-1)} = \Delta^{\mathsf{sm}}_{(c,l-1)} \wedge \mathsf{x}_i = x_i \wedge \forall i' \in [n] : \mathsf{x}_{i'} \in \mathscr{X}'_{i'}.\end{aligned}$$

We are done by the induction hypothesis. For the "moreover" part, observe that fixing $\Delta^{\mathsf{sm}}_{(c,l)}$ fixes the symbols sent by the parties upto i in the round $(c, l/2)$. As $\Delta^{\mathsf{sm}}_{(c,l)}$ also fixes the noise in this round, we are done. To finish, note that the "finally" part is because of the definition of $\mathscr{X}'_{i'}$. □

Lemma 4.19. *Let $c \in [C]$ and $l \in \left[\widehat{T_{\mathsf{Ch}}}\right]$. For any $\Delta^{\mathsf{gk}}_{(c,l)} = \left(\Upsilon_{\leq(c-1,L)}, \Phi^{\mathsf{gk}}_{\leq(c,l)}, N_{\leq(c,\lceil l/2 \rceil)}\right)$ (fixing which fixes the value of $\Gamma_{\leq(c,l)}$), we have that, conditioned on $\Delta^{\mathsf{gk}}_{(c,l)}$, the random variables $\mathsf{X}_{\overline{\Gamma_{\leq(c,l)}}}$ and $\Upsilon_{\leq(c,l)}$ are independent.*

Proof. Observe that conditioned on $\Delta^{\mathsf{gk}}_{(c,l)}$ also fixes $\Gamma_{\leq(c-1,L)}$ and let $\Gamma_{\leq(c-1,L)}$ be the fixed value. Let $\Gamma^*_{(c,l)} = \Gamma_{\leq(c,l)} \setminus \Gamma_{\leq(c-1,L)}$ for convenience and let $\Delta^{\mathsf{sm}}_{(c,l)}$ be as in Lemma 4.18.

Next, note from Equation (6) that $\Delta^{\mathsf{gk}}_{(c,l)}$ is just $\Delta^{\mathsf{sm}}_{(c,l)}$ along with some information about the groupkeys for some groups. Any such group lies in $\Gamma_{\leq(c,l)}$ by definition and if it also lies in $\Gamma_{\leq(c-1,L)}$, the information about the groupkey is implied by $\Upsilon_{\leq(c-1,L)}$. We get that there are values $\left(\mathsf{gk}_{b,g}\right)_{(b,g)\in\Gamma^*_{(c,l)}}$ such that, letting $\mathbf{\Delta}^{\mathsf{gk}}_{(c,l)}$ be the random variable associated with $\Delta^{\mathsf{gk}}_{(c,l)}$, we have:

$$\Delta^{\mathsf{gk}}_{(c,l)} = \Delta^{\mathsf{gk}}_{(c,l)} \equiv \Delta^{\mathsf{sm}}_{(c,l)} = \Delta^{\mathsf{sm}}_{(c,l)} \wedge \forall (b,g) \in \Gamma^*_{(c,l)} : \mathsf{gkey}_{b,g} = \mathsf{gk}_{b,g} \tag{11}$$

By a similar reasoning, we have that, conditioned on $\Delta^{\mathsf{gk}}_{(c,l)}$, the random variable $\Upsilon_{\leq(c,l)}$ is equivalent to the random variable $\mathsf{X}_{\Gamma^*_{(c,l)}}$. The lemma now follows as we have for $X_{\overline{\Gamma_{\leq(c,l)}}}$ and $X_{\Gamma^*_{(c,l)}}$ that:

$$\begin{aligned}
\Pr\left(X_{\overline{\Gamma_{\leq(c,l)}}} \mid \Delta^{\mathsf{gk}}_{(c,l)}\right) &= \Pr\left(X_{\overline{\Gamma_{\leq(c,l)}}} \mid \Delta^{\mathsf{sm}}_{(c,l)}, \bigwedge_{(b,g)\in\Gamma^*_{(c,l)}} \mathsf{gkey}_{b,g} = \mathsf{gk}_{b,g}\right) && \text{(Equation (11))} \\
&= \Pr\left(X_{\overline{\Gamma_{\leq(c,l)}}} \mid N_{\leq(c,\lceil l/2\rceil)}, \bigwedge_{i\in[n]} \mathsf{x}_i \in \mathscr{X}_i, \bigwedge_{(b,g)\in\Gamma^*_{(c,l)}} \mathsf{gkey}_{b,g} = \mathsf{gk}_{b,g}\right) \\
&\qquad \text{(Letting } \mathscr{X}_i \text{ be as promised by Lemma 4.18)} \\
&= \Pr\left(X_{\overline{\Gamma_{\leq(c,l)}}} \mid \bigwedge_{i\in[n]} \mathsf{x}_i \in \mathscr{X}_i, \bigwedge_{(b,g)\in\Gamma^*_{(c,l)}} \mathsf{gkey}_{b,g} = \mathsf{gk}_{b,g}\right) \\
&\qquad \text{(As } \mathsf{X} \text{ is independent of } \mathsf{N}) \\
&= \Pr\left(X_{\overline{\Gamma_{\leq(c,l)}}} \mid X_{\Gamma^*_{(c,l)}}, \bigwedge_{i\in[n]} \mathsf{x}_i \in \mathscr{X}_i, \bigwedge_{(b,g)\in\Gamma^*_{(c,l)}} \mathsf{gkey}_{b,g} = \mathsf{gk}_{b,g}\right) \\
&\qquad \text{(Lemma 4.15)} \\
&= \Pr\left(X_{\overline{\Gamma_{\leq(c,l)}}} \mid X_{\Gamma^*_{(c,l)}}, N_{\leq(c,\lceil l/2\rceil)}, \bigwedge_{i\in[n]} \mathsf{x}_i \in \mathscr{X}_i, \bigwedge_{(b,g)\in\Gamma^*_{(c,l)}} \mathsf{gkey}_{b,g} = \mathsf{gk}_{b,g}\right) \\
&\qquad \text{(As } \mathsf{X} \text{ is independent of } \mathsf{N}) \\
&= \Pr\left(X_{\overline{\Gamma_{\leq(c,l)}}} \mid X_{\Gamma^*_{(c,l)}}, \Delta^{\mathsf{sm}}_{(c,l)}, \bigwedge_{(b,g)\in\Gamma^*_{(c,l)}} \mathsf{gkey}_{b,g} = \mathsf{gk}_{b,g}\right) && \text{(Lemma 4.18)} \\
&= \Pr\left(X_{\overline{\Gamma_{\leq(c,l)}}} \mid X_{\Gamma^*_{(c,l)}}, \Delta^{\mathsf{gk}}_{(c,l)}\right). && \text{(Equation (11))}
\end{aligned}$$

□

4.2 Proof of Lemma 4.2

We prove Lemma 4.2 by induction on c. The base case $c = 0$ can be easily verified. We prove the lemma for $c > 0$ assuming it holds for $c - 1$. By the induction hypothesis, it is enough to show that:

$$\sum_{b=1}^{B} \mathbb{E}\left[\Psi_b\left(\mathsf{X}, \Gamma_{\leq(c,L)}\right) - \Psi_b\left(\mathsf{X}, \Gamma_{\leq(c-1,L)}\right)\right] \leq 2 + \frac{1}{(\log n)^8} \cdot \mathbb{E}\left[\left|\Gamma_{\leq(c,L)}\right| - \left|\Gamma_{\leq(c-1,L)}\right|\right].$$

We will show this holds even when conditioned on any realization $\Upsilon_{\leq(c-1,L)}$ of $\Upsilon_{\leq(c-1,L)}$. In fact, we will show by (backwards) induction that, for all $0 \leq \tilde{b} \leq B$, and any inputs $X_{\leq\tilde{b}}$ for the players in the first $\tilde{b}$ blocks, we have that:

$$\begin{aligned}\sum_{b=\tilde{b}+1}^{B} &\mathbb{E}\left[\Psi_b\left(\mathsf{X}, \Gamma_{\leq(c,L)}\right) - \Psi_b\left(\mathsf{X}, \Gamma_{\leq(c-1,L)}\right) \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}\right] \\ &\leq 2 - 2^{1+2\cdot\left(\tilde{b}-B\right)} + \frac{1}{(\log n)^8} \cdot \mathbb{E}\left[\left|\Gamma_{\leq(c,L)}\right| - \left|\Gamma_{\leq(c-1,L)}\right| \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}\right].\end{aligned} \tag{12}$$

The base case $\tilde{b} = B$ is straightforward. We show it for $0 \leq \tilde{b} < B$ assuming it holds for $\tilde{b} + 1$. Fix an arbitrary $\Upsilon_{\leq(c-1,L)}$ and $X_{\leq\tilde{b}}$ and observe that this also fixes $\Gamma_{\leq(c-1,L)}$ and $\mathsf{gkeys}_{\tilde{b}}(\mathsf{X})$. Let $\Gamma_{\leq(c-1,L)}$ and $\tilde{g}$ be the values these are fixed to. First, consider the case when $\left(\tilde{b}+1, \tilde{g}\right) \in \Gamma_{\leq(c-1,L)}$. In this case, it follows from Equation (10) that the term corresponding to $b = \tilde{b} + 1$ vanishes and we are done by the induction hypothesis. Henceforth, we assume that $\left(\tilde{b}+1, \tilde{g}\right) \notin \Gamma_{\leq(c-1,L)}$.

Defining the event $\mathcal{E}_1$. We now define an event $\mathcal{E}_1$ that (under our conditioning) is determined by $\mathsf{X}_{\tilde{b}+1}$. For this, recall from Lemma 4.16 that fixing $\Upsilon_{\leq(c-1,L)}$ fixes the transcript received by all the parties $i \in [n]$ in the first $(c-1)$ chunks and denote this transcript by $\Pi^i_{\leq(c-1)\cdot T_{\mathsf{Ch}}}$. For all $z \in [T_{\mathsf{Ch}}]$, $g \in [G]$ satisfying $\left(\tilde{b}+1, g\right) \notin \Gamma_{\leq(c-1,L)}$, and inputs (s,k) for one party, define the set:

$$\mathsf{hits}_{z,g}(s,k) = \left\{ j \in [m] \mid \exists \Pi' \in \{0,1,\perp\}^{z-1} : \mathsf{msg}_{\tilde{b}+1,g,j}\left((s,k), \Pi^{\tilde{b}+1,g,j}_{\leq(c-1)\cdot T_{\mathsf{Ch}}} \circ \Pi'\right) \neq \perp \right\}. \tag{13}$$

For all $z \in [T_{\mathsf{Ch}}]$, we define a tuple $(g,(s,k))$ to be z-heavy if $|\mathsf{hits}_{z,g}(s,k)| > 10 \cdot m^{2/3}$ and z-light otherwise. Also, for all $j \in \mathsf{hits}_{z,g}(s,k)$, we let $\Pi'_{z,g,(s,k)}(j)$ to be an arbitrary Π' satisfying the condition in Equation (13). For our proof, we will look at the first $10 \cdot m^{2/3}$ elements of $\mathsf{hits}_{z,g}(s,k)$. Define:

$$\mathsf{hits}^{\mathsf{few}}_{z,g}(s,k) = \begin{cases} \mathsf{hits}_{z,g}(s,k), & \text{if } (g,(s,k)) \text{ is } z\text{-light} \\ \min\left(\left\{J \subseteq \mathsf{hits}_{z,g}(s,k) \mid |J| = 10 \cdot m^{2/3}\right\}\right), & \text{if } (g,(s,k)) \text{ is } z\text{-heavy} \end{cases}. \tag{14}$$

Next, for all $z \in [T_{\mathsf{Ch}}]$, if there exists a tuple $(g,(s,k))$ that is z-heavy, define $\left(g^{(z)},\left(s^{(z)},k^{(z)}\right)\right)$ to be the z-heavy tuple $(g,(s,k))$ minimizing $\left(g,\max\left(\mathsf{hits}^{\mathsf{few}}_{z,g}(s,k)\right)\right)$ and let $j^{(z)}$ be the second coordinate of the minimum value (the first coordinate must be $g^{(z)}$) and $i^{(z)} = \left(\tilde{b}+1,g^{(z)},j^{(z)}\right)$. Otherwise, define these values to be $\perp$. Define the event $\mathcal{E}_1$ as:

Definition 4.20. *Define $\mathcal{E}_1$ to be the event that*

$$\exists z \in [T_{\mathsf{Ch}}] : g^{(z)} \neq \perp \wedge \mathsf{s}_{\tilde{b}+1,g^{(z)},0} = s^{(z)}.$$

For all $g \in [G]$ such that $\left(\tilde{b}+1,g\right) \notin \Gamma_{\leq(c-1,L)}$, define the set $\mathscr{S}_g = \left\{s^{(z)} \mid z \in [T_{\mathsf{Ch}}], g^{(z)} = g\right\}$. With this definition observe that the event $\mathcal{E}_1$ is equivalent to the event that there exists such a g with $\mathsf{s}_{\tilde{b}+1,g,0} \in \mathscr{S}_g$.

Lemma 4.21. *We have:*

$$\Pr\left(\mathcal{E}_1 \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}\right) \leq \frac{1}{4}.$$

Proof. By a union bound, it suffices to fix an arbitrary $z \in [T_{\mathsf{Ch}}]$ such that $g^{(z)} \neq \perp$ and show that:

$$\Pr\left(\mathsf{s}_{\tilde{b}+1,g^{(z)},0} = s^{(z)} \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}\right) \leq \frac{1}{4 \cdot T_{\mathsf{Ch}}}.$$

For this, note that due to Lemma 4.16 and the fact that $(\mathsf{X}_b)_{b\in[B]}$ and N are mutually independent, we have can remove $X_{\leq\tilde{b}}$ from the conditioning above. Also, note from $g^{(z)} \neq \perp$ that $\left(\tilde{b}+1,g^{(z)}\right) \notin \Gamma_{\leq(c-1,L)}$. This gives:

$$\begin{aligned}
\Pr\left(\mathsf{s}_{\tilde{b}+1,g^{(z)},0} = s^{(z)} \mid \Upsilon_{\leq(c-1,L)}\right) &\leq \Pr\left(\mathsf{s}_{\tilde{b}+1,g^{(z)},0} = s^{(z)}\right) + \left\|\left(\mathcal{D} \mid \Upsilon_{\leq(c-1,L)}\right)_{\tilde{b}+1,g^{(z)}} - \mathcal{D}_{\tilde{b}+1,g^{(z)}}\right\|_{\mathrm{TV}} \\
&\leq \Pr\left(\mathsf{s}_{\tilde{b}+1,g^{(z)},0} = s^{(z)}\right) + \frac{1}{(\log n)^2} \\
&\qquad\qquad \text{(Lemma 4.1 and Fact A.15)} \\
&\leq \frac{1}{4 \cdot T_{\mathsf{Ch}}}. \qquad \text{(As } T_{\mathsf{Ch}} = \tfrac{1}{500} \cdot \log n \text{ and } S = (10 \cdot \log n)^{10}\text{)}
\end{aligned}$$

□

Upper bounding the left side of Equation (12) when $\mathcal{E}_1$ occurs. In order to prove Equation (12), we will use the law of total expectation and upper bound the expectation on the left side conditioned on $\mathcal{E}_1$ and also upper bound it conditioned on $\overline{\mathcal{E}_1}$. Recall that whether or not $\mathcal{E}_1$ occurs is determined by $\mathsf{X}_{\tilde{b}+1}$ allowing us to view $\mathcal{E}_1$ as just a set of values of $\mathsf{X}_{\tilde{b}+1}$ for which it occurs. We have:

$$\sum_{b=\tilde{b}+1}^{B} \mathbb{E}\left[\Psi_b\left(\mathsf{X},\Gamma_{\leq(c,L)}\right) - \Psi_b\left(\mathsf{X},\Gamma_{\leq(c-1,L)}\right) \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \mathcal{E}_1\right]$$

$$
\begin{aligned}
&= \sum_{X_{\tilde{b}+1} \in \mathcal{E}_1} \Pr\left(X_{\tilde{b}+1} \mid \Upsilon_{\leq (c-1,L)}, X_{\leq \tilde{b}}, \mathcal{E}_1\right) \\
&\qquad \times \sum_{b=\tilde{b}+1}^{B} \mathbb{E}\left[\Psi_b\left(\mathsf{X}, \Gamma_{\leq (c,L)}\right) - \Psi_b\left(\mathsf{X}, \Gamma_{\leq (c-1,L)}\right) \mid \Upsilon_{\leq (c-1,L)}, X_{\leq \tilde{b}+1}\right] \\
&\leq \sum_{X_{\tilde{b}+1} \in \mathcal{E}_1} \Pr\left(X_{\tilde{b}+1} \mid \Upsilon_{\leq (c-1,L)}, X_{\leq \tilde{b}}, \mathcal{E}_1\right) \\
&\qquad \times \left(3 - 2^{1+2\cdot\left(\tilde{b}+1-B\right)} + \frac{1}{(\log n)^8} \cdot \mathbb{E}\left[\left|\Gamma_{\leq (c,L)}\right| - \left|\Gamma_{\leq (c-1,L)}\right| \mid \Upsilon_{\leq (c-1,L)}, X_{\leq \tilde{b}+1}\right]\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(Induction hypothesis)} \\
&\leq 3 - 2^{1+2\cdot\left(\tilde{b}+1-B\right)} + \frac{1}{(\log n)^8} \cdot \mathbb{E}\left[\left|\Gamma_{\leq (c,L)}\right| - \left|\Gamma_{\leq (c-1,L)}\right| \mid \Upsilon_{\leq (c-1,L)}, X_{\leq \tilde{b}}, \mathcal{E}_1\right].
\end{aligned}
$$

Upper bounding the left side of Equation (12) when $\mathcal{E}_1$ does not occur. Recall the definitions in Equations (13) and (14) and Definition 4.20 and let $\Gamma^* = \left(\left\{\tilde{b}+1\right\} \times [G]\right) \setminus \Gamma_{\leq (c-1,L)}$ for convenience. As $\left(\tilde{b}+1, \tilde{g}\right) \in \Gamma^*$, we have that Γ^* is non-empty. From Lemma 4.1, it follows that $|\Gamma^*| > \frac{3G}{4}$ and for all $\left(\tilde{b}+1, g\right) \in \Gamma^*$, we have:

$$
\mathbb{D}\left(\left(\mathcal{D} \mid \Upsilon_{\leq (c-1,L)}\right)_{\tilde{b}+1,g} \,||\, \mathcal{D}_{\tilde{b}+1,g}\right) < \frac{1}{(\log n)^4}.
$$

Let $N_{\leq (c-1)\cdot T_{\mathsf{Ch}}}$ and sets $(\mathscr{X}_i)_{i\in[n]}$ be as promised by Lemma 4.16 for $\Upsilon_{\leq (c-1,L)}$. From Lemma A.12, this implies that $\Pr\left(\forall j \in [0,m] : \mathsf{x}_{\tilde{b}+1,g,j} \in \mathscr{X}_{\tilde{b}+1,g,j}\right) > \frac{3}{4}$ for all $\left(\tilde{b}+1, g\right) \in \Gamma^*$. Plugging this into Lemmas 4.9 and 4.10, we get that for all $\left(\tilde{b}+1, g\right) \in \Gamma^*$, it holds that:

$$
\sum_{j=1}^{m} \left|\overline{\mathscr{X}_{\tilde{b}+1,g,j}}\right| \leq 4 \cdot GS \qquad \text{and} \qquad \forall j \in [0,m] : \quad \left|\overline{\mathscr{X}_{\tilde{b}+1,g,j}}\right| \leq \frac{GS}{2}.
$$

From the former, we conclude that for all $\left(\tilde{b}+1, g\right) \in \Gamma^*$, it holds that $\overline{\mathscr{X}_{\tilde{b}+1,g,j}} = \emptyset$ at least $m - 4 \cdot GS$ many values of $j \in [m]$. Now, define:

$$
\mathscr{I} = \left\{i \in [n] \mid \overline{\mathscr{X}_i} \neq \emptyset\right\} \cup \left\{\left(\tilde{b}+1, g, j\right) \mid \left(\tilde{b}+1, g\right) \in \Gamma^*, j \in \bigcup_{z\in[T_{\mathsf{Ch}}]} \bigcup_{(s,k)} \mathsf{hits}^{\mathsf{few}}_{z,g}(s,k)\right\}. \tag{15}
$$

Let $\mathsf{U}_{\tilde{b}+1} = \left(\mathsf{uq}_{\tilde{b}+1,g}\right)_{g\in[G]}$ be the random variable determining the unique players in block $\tilde{b}+1$. Now, define the event $\mathcal{E}_2$ that occurs if and only if we have a realization Λ of $\Lambda = \left(\mathsf{U}_{\tilde{b}+1}, \mathsf{X}_{\mathscr{I}}, \mathsf{X}_{>\tilde{b}+1}, \mathsf{N}_{\leq c\cdot T_{\mathsf{Ch}}}, \Phi^{\mathsf{sm}}_{\leq \left(c, \widehat{T_{\mathsf{Ch}}}\right)}\right)$ for which at least one of the following two conditions hold:

1. There exists $g \in [G]$ such that $\left(\tilde{b}+1, g\right) \in \Gamma^*$ and $\left(\tilde{b}+1, g, \mathsf{uq}_{\tilde{b}+1,g}\right) \in \mathscr{I}$.

2. There exists $z \in [T_{\mathsf{Ch}}]$ such that $g^{(z)} \neq \perp$ and we have:

$$\sum_{j \in \mathsf{hits}^{\mathsf{few}}_{z,g^{(z)}}\left(s^{(z)},k^{(z)}\right)} \mathbb{1}\left(\mathsf{x}_{\tilde{b}+1,g^{(z)},j} = \left(s^{(z)},k^{(z)}\right) \wedge \mathsf{N}^{\tilde{b}+1,g^{(z)},j}_{((c-1,T_{\mathsf{Ch}}),(c,z))} = \Pi'_{z,g^{(z)},\left(s^{(z)},k^{(z)}\right)}(j)\right) < 2.$$

Observe that whether or not $\mathcal{E}_2$ occurs is determined by Λ allowing us to view $\mathcal{E}_2$ as the set of all Λ for which it occurs. We claim the following bound on the probability of $\mathcal{E}_2$.

Lemma 4.22. *We have:*

$$\Pr\left(\mathcal{E}_2 \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}\right) \leq \frac{1}{4B}.$$

We defer the proof of Lemma 4.22 to Section 4.2.1. Assuming it for now, we claim that:

Lemma 4.23. *For all $\Lambda \notin \mathcal{E}_2$, we have that:*

$$\mathbb{H}_\infty\left(\mathsf{gkey}_{\tilde{b}+1,\tilde{g}} \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}, \Lambda\right) \geq \log G - 1.$$

Proof. Fix an arbitrary $\Lambda \notin \mathcal{E}_2$ and let $\Lambda = \left(U_{\tilde{b}+1}, X_{\mathscr{I}}, X_{>\tilde{b}+1}, N_{\leq c\cdot T_{\mathsf{Ch}}}, \Phi^{\mathsf{sm}}_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\right)$. Recall the notation $\Delta^{\mathsf{sm}}_{\left(c,\widehat{T_{\mathsf{Ch}}}\right)}$ from Lemma 4.18 and let $(\mathscr{Y}_i)_{i\in[n]}$ be the sets promised by Lemma 4.18 for $\Delta^{\mathsf{sm}}_{\left(c,\widehat{T_{\mathsf{Ch}}}\right)}$. We get:

$$\begin{aligned}
&\mathbb{H}_\infty\left(\mathsf{gkey}_{\tilde{b}+1,\tilde{g}} \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}, \Lambda\right) \\
&\qquad = \mathbb{H}_\infty\left(\mathsf{gkey}_{\tilde{b}+1,\tilde{g}} \mid X_{\leq\tilde{b}}, X_{>\tilde{b}+1}, X_{\mathscr{I}}, \overline{\mathcal{E}_1}, U_{\tilde{b}+1}, N_{\leq c\cdot T_{\mathsf{Ch}}} \wedge \forall i \in [n] : \mathsf{x}_i \in \mathscr{Y}_i\right),
\end{aligned}$$

where, for all $i \in [n]$, we have $\mathscr{Y}_i = \mathscr{X}_i$ unless there exists $l' \in \left[\widehat{T_{\mathsf{Ch}}}\right]$ and inputs (s,k) such that $\mathsf{msg}_i\left((s,k), \Pi^i_{<(c,\lceil l'/2\rceil)}\right) \neq \perp$ and either $\Phi^{\mathsf{sm}}_{(c,l'),1} = \perp$ or $i \leq \Phi^{\mathsf{sm}}_{(c,l'),1}$. In particular, we claim that we have $\mathscr{Y}_i = \mathscr{X}_i$ for all $i \in \mathscr{I}'$ where we define the set $\mathscr{I}' = \left\{\left(\tilde{b}+1, g, \mathsf{uq}_{\tilde{b}+1,g}\right) \mid \left(\tilde{b}+1, g\right) \in \Gamma^*\right\}$.

Indeed, fix an arbitrary $g \in [G]$ such that $\left(\tilde{b}+1, g\right) \in \Gamma^*$ and let $i = \left(\tilde{b}+1, g, \mathsf{uq}_{\tilde{b}+1,g}\right)$ for convenience. Also fix an arbitrary $l' \in \left[\widehat{T_{\mathsf{Ch}}}\right]$ and inputs (s,k) and let $z = \lceil l'/2\rceil$. If $\mathsf{msg}_i\left((s,k), \Pi^i_{<(c,\lceil l'/2\rceil)}\right) = \perp$, then there is nothing to show. Otherwise, we have from Equation (13) that $\mathsf{uq}_{\tilde{b}+1,g} \in \mathsf{hits}_{z,g}(s,k)$. Now, note that $\mathsf{uq}_{\tilde{b}+1,g} \notin \mathsf{hits}^{\mathsf{few}}_{z,g}(s,k)$ as otherwise, we have $i \in \mathscr{I}$ contradicting Item 1 in the definition of $\mathcal{E}_2$. This means that $\mathsf{hits}_{z,g}(s,k) \neq \mathsf{hits}^{\mathsf{few}}_{z,g}(s,k)$ which implies that $(g,(s,k))$ is z-heavy. It follows that we have $\left(g^{(z)}, \max\left(\mathsf{hits}^{\mathsf{few}}_{z,g^{(z)}}\left(s^{(z)},k^{(z)}\right)\right)\right) < \left(g, \mathsf{uq}_{\tilde{b}+1,g}\right)$. Combining this with the fact that Item 2 in

the definition of $\mathcal{E}_2$ is false, we have that there are at least two parties smaller than i that are speaking in round (c, z). This means that $\Phi^{\mathsf{sm}}_{(c,l'),1} \neq \perp$ and $\Phi^{\mathsf{sm}}_{(c,l'),1} < i$, and we are done.

Having proved the claim, we again use the fact that Item 1 in the definition of $\mathcal{E}_2$ to get that for all $i \in \mathscr{I}'$, we have $\overline{\mathscr{Y}_i} = \emptyset$. We will show the min-entropy bound hold even under a stronger conditioning where we condition on the inputs of all players not in $\mathscr{I}'$, the value $(s_i)_{i \in \mathscr{I}'}$ and the noise $N_{\leq c \cdot T_{\mathsf{Ch}}}$. In other words, the only randomness remaining in the inputs is the randomness $(\mathsf{k}_i)_{i \in \mathscr{I}'}$. Observe that this conditioning is indeed stronger. Under this stronger conditioning, we use the fact that $\overline{\mathscr{Y}_i} = \emptyset$ for all $i \in \mathscr{I}'$ to get that the distribution of $\mathsf{gkey}_{\tilde{b}+1,\tilde{g}}$ is uniform over a set of size at least $\frac{3G}{4}$. The lemma follows. □

To upper bounding the left side of Equation (12), note that we have from the law of total expectation that:

$$\begin{aligned}
&\sum_{b=\tilde{b}+1}^{B} \mathbb{E}\big[\Psi_b\big(\mathsf{X}, \Gamma_{\leq(c,L)}\big) - \Psi_b\big(\mathsf{X}, \Gamma_{\leq(c-1,L)}\big) \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}\big] \\
&\qquad \leq B \cdot \Pr\big(\mathcal{E}_2 \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}\big) + \sum_{\Lambda \notin \mathcal{E}_2} \Pr\big(\Lambda \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}\big) \\
&\qquad\qquad \times \sum_{b=\tilde{b}+1}^{B} \mathbb{E}\big[\Psi_b\big(\mathsf{X}, \Gamma_{\leq(c,L)}\big) - \Psi_b\big(\mathsf{X}, \Gamma_{\leq(c-1,L)}\big) \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}, \Lambda\big].
\end{aligned} \tag{16}$$

We focus on the last expectation above for an arbitrary $\Lambda \notin \mathcal{E}_2$. We have from Corollary 4.7 that:

$$\begin{aligned}
&\sum_{b=\tilde{b}+1}^{B} \mathbb{E}\big[\Psi_b\big(\mathsf{X}, \Gamma_{\leq(c,L)}\big) - \Psi_b\big(\mathsf{X}, \Gamma_{\leq(c-1,L)}\big) \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}, \Lambda\big] \\
&\qquad \leq 1 + \sum_{b=\tilde{b}+2}^{B} \mathbb{E}\big[\mathbb{1}\big((b, \mathsf{gkeys}_{b-1}(\mathsf{X})) \in \Gamma_{\leq(c,L)} \setminus \Gamma_{\leq(c-1,L)}\big) \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}, \Lambda\big] \\
&\qquad\qquad + \frac{1}{(\log n)^{10}} \cdot \sum_{b=\tilde{b}+2}^{B} \mathbb{E}\big[\big|\big(\Gamma_{\leq(c,L)} \setminus \Gamma_{\leq(c-1,L)}\big) \cap (\{b\} \times [G])\big| \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}, \Lambda\big].
\end{aligned}$$

Now, we claim that for all $b \in \left[\tilde{b}+2, B\right]$, our conditioning fixes the value of $\big(\Gamma_{\leq(c,L)} \setminus \Gamma_{\leq(c-1,L)}\big) \cap (\{b\} \times [G])$. That is, it fixes the groups in block b that are added to $\Gamma_{\leq(c,L)}$. In fact, we will show that not only does it fix the groups that are added, it also fixes the order in which these groups are added. Indeed, for any prefix of this order, Equation (7) says the next element to be added is determined by X and $\left(\Upsilon_{\leq(c-1,L)}, \Phi^{\mathsf{sm}}_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}, N_{\leq c \cdot T_{\mathsf{Ch}}}\right)$. As this tuple forms a "rectangle" (see Lemma 4.18), the next element is determined by X_b, which is fixed by our

conditioning and we are done. Let this fixed value be $\mathscr{G}_b$. We get:

$$\sum_{b=\tilde{b}+1}^{B} \mathbb{E}\big[\Psi_b\big(\mathsf{X}, \Gamma_{\le(c,L)}\big) - \Psi_b\big(\mathsf{X}, \Gamma_{\le(c-1,L)}\big) \mid \Upsilon_{\le(c-1,L)}, X_{\le\tilde{b}}, \overline{\mathcal{E}_1}, \Lambda\big]$$
$$\le 1 + \frac{1}{(\log n)^{10}} \cdot \sum_{b=\tilde{b}+2}^{B} |\mathscr{G}_b| + \sum_{b=\tilde{b}+2}^{B} \sum_{(b,g')\in\mathscr{G}_b} \Pr\big(\mathsf{gkeys}_{b-1}(\mathsf{X}) = g' \mid \Upsilon_{\le(c-1,L)}, X_{\le\tilde{b}}, \overline{\mathcal{E}_1}, \Lambda\big).$$

Now, use Equation (4) and Item 2 of our promise on the distribution $\mathcal{D}$ and our conditioning to get that for all $(b, g') \in \mathscr{G}_b$ there exists a unique $g^* \in [G]$ such that $\mathsf{gkeys}_{b-1}(\mathsf{X}) = g'$ only if $\mathsf{gkeys}_{\tilde{b}+1}(\mathsf{X}) = \mathsf{gkey}_{\tilde{b}+1,\tilde{g}} = g^*$. Using this, we get that:

$$\sum_{b=\tilde{b}+1}^{B} \mathbb{E}\big[\Psi_b\big(\mathsf{X}, \Gamma_{\le(c,L)}\big) - \Psi_b\big(\mathsf{X}, \Gamma_{\le(c-1,L)}\big) \mid \Upsilon_{\le(c-1,L)}, X_{\le\tilde{b}}, \overline{\mathcal{E}_1}, \Lambda\big]$$
$$\le 1 + \frac{1}{(\log n)^{10}} \cdot \sum_{b=\tilde{b}+2}^{B} |\mathscr{G}_b| + \frac{2}{G} \cdot \sum_{b=\tilde{b}+2}^{B} |\mathscr{G}_b| \qquad \text{(Lemma 4.23)}$$
$$\le 1 + \frac{1}{(\log n)^{8}} \cdot \sum_{b=\tilde{b}+2}^{B} |\mathscr{G}_b|. \qquad \text{(As } G = (10 \cdot \log n)^{10}\text{)}$$

Plugging this bound into Equation (16), we get:

$$\sum_{b=\tilde{b}+1}^{B} \mathbb{E}\big[\Psi_b\big(\mathsf{X}, \Gamma_{\le(c,L)}\big) - \Psi_b\big(\mathsf{X}, \Gamma_{\le(c-1,L)}\big) \mid \Upsilon_{\le(c-1,L)}, X_{\le\tilde{b}}, \overline{\mathcal{E}_1}\big]$$
$$\le B \cdot \Pr\big(\mathcal{E}_2 \mid \Upsilon_{\le(c-1,L)}, X_{\le\tilde{b}}, \overline{\mathcal{E}_1}\big)$$
$$+ \sum_{\Lambda\notin\mathcal{E}_2} \Pr\big(\Lambda \mid \Upsilon_{\le(c-1,L)}, X_{\le\tilde{b}}, \overline{\mathcal{E}_1}\big) \cdot \left(1 + \frac{1}{(\log n)^{8}} \cdot \sum_{b=\tilde{b}+2}^{B} |\mathscr{G}_b|\right)$$
$$\le 1 + B \cdot \Pr\big(\mathcal{E}_2 \mid \Upsilon_{\le(c-1,L)}, X_{\le\tilde{b}}, \overline{\mathcal{E}_1}\big) + \frac{1}{(\log n)^{8}} \cdot \Pr\big(\overline{\mathcal{E}_2} \mid \Upsilon_{\le(c-1,L)}, X_{\le\tilde{b}}, \overline{\mathcal{E}_1}\big)$$
$$\times \mathbb{E}\big[\big|\Gamma_{\le(c,L)}\big| - \big|\Gamma_{\le(c-1,L)}\big| \mid \Upsilon_{\le(c-1,L)}, X_{\le\tilde{b}}, \overline{\mathcal{E}_1}, \overline{\mathcal{E}_2}\big]$$
$$\le 1 + B \cdot \Pr\big(\mathcal{E}_2 \mid \Upsilon_{\le(c-1,L)}, X_{\le\tilde{b}}, \overline{\mathcal{E}_1}\big)$$
$$+ \frac{1}{(\log n)^{8}} \cdot \mathbb{E}\big[\big|\Gamma_{\le(c,L)}\big| - \big|\Gamma_{\le(c-1,L)}\big| \mid \Upsilon_{\le(c-1,L)}, X_{\le\tilde{b}}, \overline{\mathcal{E}_1}\big]$$
$$\le 1 + \frac{1}{4} + \frac{1}{(\log n)^{8}} \cdot \mathbb{E}\big[\big|\Gamma_{\le(c,L)}\big| - \big|\Gamma_{\le(c-1,L)}\big| \mid \Upsilon_{\le(c-1,L)}, X_{\le\tilde{b}}, \overline{\mathcal{E}_1}\big]. \qquad \text{(Lemma 4.22)}$$

Using the bounds above. We now use the two bounds for the left side of Equation (12) proved above to get:

$$\begin{aligned}
&\sum_{b=\tilde{b}+1}^{B} \mathbb{E}\big[\Psi_b\big(\mathsf{X}, \Gamma_{\leq(c,L)}\big) - \Psi_b\big(\mathsf{X}, \Gamma_{\leq(c-1,L)}\big) \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}\big] \\
&\quad \leq \Pr\big(\mathcal{E}_1 \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}\big) \cdot \sum_{b=\tilde{b}+1}^{B} \mathbb{E}\big[\Psi_b\big(\mathsf{X}, \Gamma_{\leq(c,L)}\big) - \Psi_b\big(\mathsf{X}, \Gamma_{\leq(c-1,L)}\big) \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \mathcal{E}_1\big] \\
&\qquad + \Pr\big(\overline{\mathcal{E}_1} \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}\big) \cdot \sum_{b=\tilde{b}+1}^{B} \mathbb{E}\big[\Psi_b\big(\mathsf{X}, \Gamma_{\leq(c,L)}\big) - \Psi_b\big(\mathsf{X}, \Gamma_{\leq(c-1,L)}\big) \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}\big] \\
&\quad \leq 1 + \frac{1}{4} + \left(3 - 2^{1+2\cdot(\tilde{b}+1-B)}\right) \cdot \frac{1}{4} + \frac{1}{(\log n)^8} \cdot \mathbb{E}\big[\big|\Gamma_{\leq(c,L)}\big| - \big|\Gamma_{\leq(c-1,L)}\big| \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}\big] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(Lemma 4.21)} \\
&\quad \leq 2 - 2^{1+2\cdot(\tilde{b}-B)} + \frac{1}{(\log n)^8} \cdot \mathbb{E}\big[\big|\Gamma_{\leq(c,L)}\big| - \big|\Gamma_{\leq(c-1,L)}\big| \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}\big].
\end{aligned}$$

This proves Equation (12) and thus, also proves Lemma 4.2.

4.2.1 Proof of Lemma 4.22

We now show Lemma 4.22. By a union bound, it suffices to upper bound the probability of Items 1 and 2 in the definition of $\mathcal{E}_2$ separately.

Upper bounding the probability of Item 1. By another union bound, it suffices to consider a fixed $g \in [G]$ such that $\left(\tilde{b}+1, g\right) \in \Gamma^*$. Let $\mathscr{S}_g$ be as in Definition 4.20 and note from $S = (\log n)^{10}$ that $\left|\overline{\mathscr{S}_g}\right| \geq \frac{3S}{4}$. By Lemma 4.11, we have that for all $g' \in [G]$, it holds that:

$$\mathbb{H}_\infty\big(\mathsf{uq}_{\tilde{b}+1,g} \mid \mathsf{s}_{\tilde{b}+1,g,0} \in \overline{\mathscr{S}_g} \wedge \mathsf{gkey}_{\tilde{b}+1,g} = g' \wedge \forall j \in [0,m] : \mathsf{x}_{\tilde{b}+1,g,j} \in \mathscr{X}_{\tilde{b}+1,g,j}\big) \geq 0.9 \cdot \log m.$$

Using Lemma 4.15 and the equivalent definition of $\mathcal{E}_1$ in Definition 4.20, we get that for all $g' \in [G]$:

$$\mathbb{H}_\infty\big(\mathsf{uq}_{\tilde{b}+1,g} \mid \overline{\mathcal{E}_1} \wedge \mathsf{gkey}_{\tilde{b}+1,g} = g' \wedge \forall j \in [0,m] : \mathsf{x}_i \in \mathscr{X}_i\big) \geq 0.9 \cdot \log m.$$

As this holds for all $g' \in [G]$ and as X is independent of N, we have:

$$\mathbb{H}_\infty\big(\mathsf{uq}_{\tilde{b}+1,g} \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}\big) \geq 0.9 \cdot \log m. \tag{17}$$

Item 1 now follows as we get:

$$\begin{aligned}
&\Pr\Big(\Big(\tilde{b}+1, g, \mathsf{uq}_{\tilde{b}+1,g}\Big) \in \mathscr{I} \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}\Big) \\
&\quad\leq \sum_{i\in\mathscr{I}} \Pr\Big(\Big(\tilde{b}+1, g, \mathsf{uq}_{\tilde{b}+1,g}\Big) = i \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}\Big) \\
&\quad\leq \Big|\mathscr{I} \cap \Big(\Big\{\Big(\tilde{b}+1, g\Big)\Big\} \times [m]\Big)\Big| \cdot m^{-0.9} && \text{(Equation (17))} \\
&\quad\leq \big(5GS + 10 \cdot T_{\mathsf{Ch}} \cdot GS \cdot m^{2/3}\big) \cdot m^{-0.9} && \text{(Equation (15))} \\
&\quad\leq m^{-0.2}.
\end{aligned}$$

Upper bounding the probability of Item 2. By another union bound, it suffices to consider a fixed $z \in [T_{\mathsf{Ch}}]$ such that $g^{(z)} \neq \perp$. Let $\mathcal{E}^{(z)}$ be the event in Item 2 for this z. Also, define that event $\mathcal{E}'^{(z)}$ as:

$$\sum_{j\in\mathsf{hits}^{\mathsf{few}}_{z,g^{(z)}}\left(s^{(z)},k^{(z)}\right)} \mathbb{1}\big(\mathsf{x}_{\tilde{b}+1,g^{(z)},j} = \big(s^{(z)}, k^{(z)}\big)\big) < m^{1/3}.$$

To upper bound the probability of $\mathcal{E}^{(z)}$ occuring, we first upper bound the probability of $\mathcal{E}'^{(z)}$ occuring. For this, for all $i \in [n]$, define the set $\mathscr{X}_i^{(z)}$ as follows: If i is in the first $\tilde{b}$ blocks, then define $\mathscr{X}_i^{(z)} = \{x_i\}$. Otherwise, if there exists $g' \neq g^{(z)}$ satisfying $\Big(\tilde{b}+1, g'\Big) \in \Gamma^*$, define $\mathscr{X}_i^{(z)} = \mathscr{X}_i \cap \big(\overline{\mathscr{S}_{g'}} \times [G]\big)$. If both these conditions fail, set $\mathscr{X}_i^{(z)} = \mathscr{X}_i$. We apply Lemma 4.12 on the sets $\mathscr{X}_i^{(z)}$ to get that for all $s' \in \overline{\mathscr{S}_{g^{(z)}}}$ and all $j' \in [m]$, we have:

$$\Pr\Big(\mathcal{E}'^{(z)} \mid \mathsf{s}_{\tilde{b}+1,g^{(z)},0} = s' \wedge \mathsf{uq}_{\tilde{b}+1,g^{(z)}} = j' \wedge \forall i \in [n] : \mathsf{x}_i \in \mathscr{X}_i^{(z)}\Big) \leq \mathrm{e}^{-\frac{1}{8}\cdot m^{1/3}}.$$

As $s' \in \overline{\mathscr{S}_{g^{(z)}}}$ and $j' \in [m]$ were arbitrary, we get:

$$\Pr\Big(\mathcal{E}'^{(z)} \mid \mathsf{s}_{\tilde{b}+1,g^{(z)},0} \in \overline{\mathscr{S}_{g^{(z)}}} \wedge \forall i \in [n] : \mathsf{x}_i \in \mathscr{X}_i^{(z)}\Big) \leq \mathrm{e}^{-\frac{1}{8}\cdot m^{1/3}}.$$

From the definition of $\mathscr{X}_i^{(z)}$ and using Definition 4.20, we get:

$$\Pr\big(\mathcal{E}'^{(z)} \mid X_{\leq\tilde{b}} \wedge \overline{\mathcal{E}_1} \wedge \forall i \in [n] : \mathsf{x}_i \in \mathscr{X}_i\big) \leq \mathrm{e}^{-\frac{1}{8}\cdot m^{1/3}}.$$

As X is independent of N, we have $\Pr\big(\mathcal{E}'^{(z)} \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}\big) \leq \mathrm{e}^{-\frac{1}{8}\cdot m^{1/3}}$. Using a union bound, this gives:

$$\Pr\big(\mathcal{E}^{(z)} \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}\big) \leq \mathrm{e}^{-\frac{1}{8}\cdot m^{1/3}} + \Pr\Big(\mathcal{E}^{(z)} \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}, \overline{\mathcal{E}'^{(z)}}\Big).$$

We now analyze the second term above. For this, note that the event we condition on is determined by X and the noise $\mathsf{N}_{\leq(c-1)\cdot T_{\mathsf{Ch}}}$ in the first $(c-1)$ chunks. Thus, in order to

upper bound this probability it suffices to fix an arbitrary X and $N_{\leq(c-1)\cdot T_{\mathsf{Ch}}}$ such that our conditioning occurs and upper bound $\Pr\left(\mathcal{E}^{(z)} \mid X, N_{\leq(c-1)\cdot T_{\mathsf{Ch}}}\right)$. For this, note that as $\mathcal{E}'^{(z)}$ does not occur, we have a set $J \subseteq \mathsf{hits}^{\mathsf{few}}_{z,g^{(z)}}\left(s^{(z)}, k^{(z)}\right)$ with $|J| \geq m^{1/3}$ such that $x_{\tilde{b}+1,g^{(z)},j} = \left(s^{(z)}, k^{(z)}\right)$ for all $j \in J$. We get:

$$\begin{aligned}
&\Pr\left(\mathcal{E}^{(z)} \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}, \overline{\mathcal{E}'^{(z)}}\right) \\
&\quad\leq \Pr\left(\sum_{j\in J} \mathbb{1}\left(\mathsf{N}^{\tilde{b}+1,g^{(z)},j}_{((c-1,T_{\mathsf{Ch}}),(c,z))} = \Pi'_{z,g^{(z)},\left(s^{(z)},k^{(z)}\right)}(j)\right) < 2 \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}, \overline{\mathcal{E}'^{(z)}}\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (\text{Item 2}) \\
&\quad\leq 2^{-n^{0.2}}. \qquad\qquad\qquad\qquad\qquad\qquad (\text{Lemma 3.1 and as } T_{\mathsf{Ch}} = \tfrac{1}{500}\cdot\log n)
\end{aligned}$$

Plugging in, we get:

$$\Pr\left(\mathcal{E}^{(z)} \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}\right) \leq \mathrm{e}^{-\frac{1}{8}\cdot m^{1/3}} + 2^{-n^{0.2}} \leq 2^{-n^{0.15}}.$$

Finishing the proof. To finish, we combine both the parts above and get $\Pr\left(\mathcal{E}_2 \mid \Upsilon_{\leq(c-1,L)}, X_{\leq\tilde{b}}, \overline{\mathcal{E}_1}\right) \leq m^{-0.15} \leq \frac{1}{4B}$, as desired.

4.3 Proof of Lemma 4.3

This section is devoted to the proof of Lemma 4.3. For this, define $\mathscr{H} = \mathbb{H}(X_{1,1})$ to be the entropy of the inputs of the first group of players. By symmetry, this is also the entropy of any group of players. We will prove that for all $0 \leq c \leq C$, we have the following entropy upper bound:

$$\mathscr{H}\cdot BG - \mathbb{E}\left[\left|\Gamma_{\leq(c,L)}\right|\right]\cdot\left(\mathscr{H} - \frac{1}{5\cdot(\log n)^4}\right) - 9c\cdot T_{\mathsf{Ch}}\cdot\log n \leq \mathbb{H}\left(\mathsf{X} \mid \Upsilon_{\leq(c,L)}\right). \tag{18}$$

This is enough as we have from Lemma 4.13 that $\mathbb{H}\left(\mathsf{X} \mid \Upsilon_{\leq(c,L)}\right) \leq \mathscr{H}\cdot\left(BG - \mathbb{E}\left[\left|\Gamma_{\leq(c,L)}\right|\right]\right)$. Lemma 4.3 now follows by a simple rearrangement. To prove Equation (18), note from Lemma 4.4 that it suffices to show that, for all $0 \leq c \leq C$, we have:

$$\begin{aligned}
&\sum_{\Upsilon_{\leq(c,L)}} \Pr\left(\Upsilon_{\leq(c,L)}\right)\cdot\sum_{c'=1}^{c}\sum_{l=\widehat{T_{\mathsf{Ch}}}+1}^{L} \mathbb{D}\left(\left(\mathcal{D} \mid \Upsilon_{<(c',l)}\right)_{\Phi_{(c',l),1}} \,||\, \mathcal{D}_{\Phi_{(c',l),1}}\right) \\
&\qquad\qquad\qquad\qquad \leq 8c\cdot T_{\mathsf{Ch}}\cdot\log n - \mathscr{H}\cdot\left(BG - \mathbb{E}\left[\left|\Gamma_{\leq(c,L)}\right|\right]\right) + \mathbb{H}\left(\mathsf{X} \mid \Upsilon_{\leq(c,L)}\right).
\end{aligned}$$

We prove this by induction on c. The base case $c = 0$ is straightforward. We show the statement for $c > 0$ assuming that it holds for $c - 1$. By the induction hypothesis, it suffices

to show that:

$$\sum_{\Upsilon_{\le(c,L)}} \Pr\big(\Upsilon_{\le(c,L)}\big) \cdot \sum_{l=\widehat{T_{\mathsf{Ch}}}+1}^{L} \mathbb{D}\Big(\big(\mathcal{D} \mid \Upsilon_{<(c,l)}\big)_{\Phi_{(c,l),1}} \,||\, \mathcal{D}_{\Phi_{(c,l),1}}\Big) \\ \le 8 \cdot T_{\mathsf{Ch}} \cdot \log n + \mathscr{H} \cdot \mathbb{E}\big[\big|\Gamma_{\le(c,L)}\big| - \big|\Gamma_{\le(c-1,L)}\big|\big] - \mathbb{I}\big(\mathsf{X} : \Upsilon_{\le(c,L)} \mid \Upsilon_{\le(c-1,L)}\big). \tag{19}$$

Analyzing the range $l \in \left(\widehat{T_{\mathsf{Ch}}}, L\right]$. We divide the proof of Equation (19) into two parts and first consider the values $l \in \left(\widehat{T_{\mathsf{Ch}}}, L\right]$. For any such l, note that fixing any $\Upsilon_{\le(c-1,L)}$ also fixes the value of $\Gamma_{<(c,L)}$ and $\Gamma_{\le(c,L)}$. Let $\Gamma^*_{(c,l)} = \Gamma_{\le(c,L)} \setminus \Gamma_{<(c,L)}$ for convenience and observe that conditioned on $\Upsilon_{\le(c-1,L)}$, the random variable $\Upsilon_{\le(c,l)}$ determines and is determined by $\mathsf{X}_{\Gamma^*_{(c,l)}}$. This means that:

$$\begin{aligned}
\mathbb{I}\big(\mathsf{X} : \Upsilon_{\le(c,l)} \mid \Upsilon_{<(c,l)}\big) &= \mathbb{H}\big(\Upsilon_{\le(c,l)} \mid \Upsilon_{<(c,l)}\big) - \mathbb{H}\big(\Upsilon_{\le(c,l)} \mid \mathsf{X}, \Upsilon_{<(c,l)}\big) && \text{(Definition A.7)} \\
&= \mathbb{H}\big(\Upsilon_{\le(c,l)} \mid \Upsilon_{<(c,l)}\big) \\
&= \sum_{\Upsilon_{<(c,l)}} \Pr\big(\Upsilon_{<(c,l)}\big) \cdot \sum_{X_{\Gamma^*_{(c,l)}}} \Pr\Big(X_{\Gamma^*_{(c,l)}} \mid \Upsilon_{<(c,l)}\Big) \cdot \log \frac{1}{\Pr\Big(X_{\Gamma^*_{(c,l)}} \mid \Upsilon_{<(c,l)}\Big)} \\
& && \text{(Definition A.2)} \\
&= \sum_{\Upsilon_{<(c,l)}} \Pr\big(\Upsilon_{<(c,l)}\big) \cdot \mathscr{H} \cdot \big(\big|\Gamma_{\le(c,l)}\big| - \big|\Gamma_{<(c,l)}\big|\big) \\
&\quad - \sum_{\Upsilon_{<(c,l)}} \Pr\big(\Upsilon_{<(c,l)}\big) \cdot \mathbb{D}\Big(\big(\mathcal{D} \mid \Upsilon_{<(c,l)}\big)_{\Phi_{(c,l),1}} \,||\, \mathcal{D}_{\Phi_{(c,l),1}}\Big) \\
& && \text{(Definition A.10)} \\
&= \mathscr{H} \cdot \mathbb{E}\big[\big|\Gamma_{\le(c,l)}\big| - \big|\Gamma_{<(c,l)}\big|\big] \\
&\quad - \sum_{\Upsilon_{<(c,l)}} \Pr\big(\Upsilon_{<(c,l)}\big) \cdot \mathbb{D}\Big(\big(\mathcal{D} \mid \Upsilon_{<(c,l)}\big)_{\Phi_{(c,l),1}} \,||\, \mathcal{D}_{\Phi_{(c,l),1}}\Big).
\end{aligned}$$

Adding this for all $l \in \left(\widehat{T_{\mathsf{Ch}}}, L\right]$ gives:

$$\mathbb{I}\Big(\mathsf{X} : \Upsilon_{\le(c,L)} \mid \Upsilon_{\le\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\Big) = \mathscr{H} \cdot \mathbb{E}\Big[\big|\Gamma_{\le(c,L)}\big| - \Big|\Gamma_{\le\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\Big|\Big] \\ - \sum_{\Upsilon_{\le(c,L)}} \Pr\big(\Upsilon_{\le(c,L)}\big) \cdot \sum_{l=\widehat{T_{\mathsf{Ch}}}+1}^{L} \mathbb{D}\Big(\big(\mathcal{D} \mid \Upsilon_{<(c,l)}\big)_{\Phi_{(c,l),1}} \,||\, \mathcal{D}_{\Phi_{(c,l),1}}\Big).$$

Plugging this into Equation (19), we have that it suffices to show:

$$\mathbb{I}\Big(\mathsf{X} : \Upsilon_{\le\left(c,\widehat{T_{\mathsf{Ch}}}\right)} \mid \Upsilon_{\le(c-1,L)}\Big) \le 8 \cdot T_{\mathsf{Ch}} \cdot \log n + \mathscr{H} \cdot \mathbb{E}\Big[\Big|\Gamma_{\le\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\Big| - \big|\Gamma_{\le(c-1,L)}\big|\Big]. \tag{20}$$

Proving Equation (20). We now focus on showing Equation (20). We have:

$$
\begin{aligned}
&\mathbb{I}\Big(\mathsf{X} : \Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)} \mid \Upsilon_{\leq(c-1,L)}\Big) \\
&\quad = \mathbb{H}\big(\mathsf{X} \mid \Upsilon_{\leq(c-1,L)}\big) - \mathbb{H}\Big(\mathsf{X} \mid \Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\Big) && \text{(Definition A.7)} \\
&\quad = \mathbb{H}\big(\mathsf{X} \mid \Upsilon_{\leq(c-1,L)}, \mathsf{N}_{\leq c\cdot T_{\mathsf{Ch}}}\big) - \mathbb{H}\Big(\mathsf{X} \mid \Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\Big) \\
&\qquad\qquad \text{(Lemmas 4.16 and A.4 and as noise is independent across rounds)} \\
&\quad = \mathbb{I}\left(\mathsf{X} : \Phi^{\mathsf{gk}}_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)} \mid \Upsilon_{\leq(c-1,L)}, \mathsf{N}_{\leq c\cdot T_{\mathsf{Ch}}}\right) + \mathbb{H}\left(\mathsf{X} \mid \Upsilon_{\leq(c-1,L)}, \Phi^{\mathsf{gk}}_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}, \mathsf{N}_{\leq c\cdot T_{\mathsf{Ch}}}\right) - \mathbb{H}\Big(\mathsf{X} \mid \Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\Big) \\
&\qquad\qquad \text{(Definition A.7)} \\
&\quad \leq \mathbb{H}\left(\Phi^{\mathsf{gk}}_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)} \mid \Upsilon_{\leq(c-1,L)}\right) + \mathbb{H}\left(\mathsf{X} \mid \Upsilon_{\leq(c-1,L)}, \Phi^{\mathsf{gk}}_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}, \mathsf{N}_{\leq c\cdot T_{\mathsf{Ch}}}\right) - \mathbb{H}\Big(\mathsf{X} \mid \Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\Big) \\
&\qquad\qquad \text{(Lemma A.4 and Fact A.8)} \\
&\quad \leq 8 \cdot T_{\mathsf{Ch}} \cdot \log n + \mathbb{H}\left(\mathsf{X} \mid \Upsilon_{\leq(c-1,L)}, \Phi^{\mathsf{gk}}_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}, \mathsf{N}_{\leq c\cdot T_{\mathsf{Ch}}}\right) - \mathbb{H}\Big(\mathsf{X} \mid \Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\Big). \\
&\qquad\qquad \text{(Lemmas 4.17 and A.6)}
\end{aligned}
$$

Analyzing the second term above, note that:

$$
\begin{aligned}
&\mathbb{H}\left(\mathsf{X} \mid \Upsilon_{\leq(c-1,L)}, \Phi^{\mathsf{gk}}_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}, \mathsf{N}_{\leq c\cdot T_{\mathsf{Ch}}}\right) \\
&\quad \leq \mathscr{H} \cdot \mathbb{E}\left[\left|\Gamma_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\right| - \left|\Gamma_{\leq(c-1,L)}\right|\right] \\
&\qquad + \sum_{\substack{\Upsilon_{\leq(c-1,L)} \\ \Phi^{\mathsf{gk}}_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)} \\ N_{\leq c\cdot T_{\mathsf{Ch}}}}} \Pr\left(\Upsilon_{\leq(c-1,L)}, \Phi^{\mathsf{gk}}_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}, N_{\leq c\cdot T_{\mathsf{Ch}}}\right) \cdot \mathbb{H}\left(\mathsf{X}_{\overline{\Gamma_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}}} \mid \Upsilon_{\leq(c-1,L)}, \Phi^{\mathsf{gk}}_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}, N_{\leq c\cdot T_{\mathsf{Ch}}}\right) \\
&\qquad\qquad \text{(Lemma 4.14)} \\
&\quad \leq \mathscr{H} \cdot \mathbb{E}\left[\left|\Gamma_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\right| - \left|\Gamma_{\leq(c-1,L)}\right|\right] \\
&\qquad + \sum_{\Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}} \Pr\left(\Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\right) \cdot \mathbb{H}\left(\mathsf{X}_{\overline{\Gamma_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}}} \mid \Upsilon_{\leq(c-1,L)}, \Phi^{\mathsf{gk}}_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}, N_{\leq c\cdot T_{\mathsf{Ch}}}\right) \\
&\qquad\qquad \text{(Lemma 4.17)} \\
&\quad \leq \mathscr{H} \cdot \mathbb{E}\left[\left|\Gamma_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\right| - \left|\Gamma_{\leq(c-1,L)}\right|\right] + \sum_{\Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}} \Pr\left(\Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\right) \cdot \mathbb{H}\left(\mathsf{X}_{\overline{\Gamma_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}}} \mid \Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\right) \\
&\qquad\qquad \text{(Lemma 4.19)} \\
&\quad \leq \mathscr{H} \cdot \mathbb{E}\left[\left|\Gamma_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\right| - \left|\Gamma_{\leq(c-1,L)}\right|\right] + \sum_{\Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}} \Pr\left(\Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\right) \cdot \mathbb{H}\left(\mathsf{X} \mid \Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\right)
\end{aligned}
$$

$$\leq \mathscr{H} \cdot \mathbb{E}\Big[\Big|\Gamma_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\Big| - \big|\Gamma_{\leq(c-1,L)}\big|\Big] + \mathbb{H}\Big(\mathsf{X} \mid \Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\Big). \qquad \text{(Definition A.2)}$$

Plugging in, we get

$$\mathbb{I}\Big(\mathsf{X} : \Upsilon_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)} \mid \Upsilon_{\leq(c-1,L)}\Big) \leq 8 \cdot T_{\mathsf{Ch}} \cdot \log n + \mathscr{H} \cdot \mathbb{E}\Big[\Big|\Gamma_{\leq\left(c,\widehat{T_{\mathsf{Ch}}}\right)}\Big| - \big|\Gamma_{\leq(c-1,L)}\big|\Big],$$

which is just Equation (20), and the proof is complete.

References

[ABE+16] Noga Alon, Mark Braverman, Klim Efremenko, Ran Gelles, and Bernhard Haeupler. Reliable communication over highly connected noisy networks. In *Symposium on Principles of Distributed Computing (DISC)*, pages 165–173. ACM, 2016. 3

[AGL20] Yagel Ashkenazi, Ran Gelles, and Amir Leshem. Brief announcement: Noisy beeping networks. In *Symposium on Principles of Distributed Computing (PODC)*, pages 458–460, 2020. 3

[AGS16] Shweta Agrawal, Ran Gelles, and Amit Sahai. Adaptive protocols for interactive communication. In *Information Theory (ISIT)*, pages 595–599. IEEE, 2016. 1

[BEGH16] Mark Braverman, Klim Efremenko, Ran Gelles, and Bernhard Haeupler. Constant-rate coding for multiparty interactive communication is impossible. In *Symposium on Theory of Computing (STOC)*, pages 999–1010. ACM, 2016. 1, 3

[BGI92] Reuven Bar-Yehuda, Oded Goldreich, and Alon Itai. On the time-complexity of broadcast in multi-hop radio networks: An exponential gap between determinism and randomization. *Journal of Computer and System Sciences*, 45(1):104–126, 1992. 1

[CHHHZ17] Keren Censor-Hillel, Bernhard Haeupler, D Ellis Hershkowitz, and Goran Zuzic. Broadcasting in noisy radio networks. In *Symposium on Principles of Distributed Computing (PODC)*, pages 33–42, 2017. 3

[CHHHZ19] Keren Censor-Hillel, Bernhard Haeupler, D Ellis Hershkowitz, and Goran Zuzic. Erasure correction for noisy radio networks. In *International Symposium on Distributed Computing (DISC)*, 2019. 1, 3

[EKP+24] Klim Efremenko, Gillat Kol, Dmitry Paramonov, Ran Raz, and Raghuvansh R. Saxena. Information dissemination via broadcasts in the presence of adversarial noise. In *Computational Complexity Conference (CCC)*, volume 300, pages 19:1–19:33, 2024. 3

[EKPS21a] Klim Efremenko, Gillat Kol, Dmitry Paramonov, and Raghuvansh R. Saxena. Computation over the noisy broadcast channel with malicious parties. In *Innovations in Theoretical Computer Science Conference, (ITCS)*, volume 185, pages 82:1–82:19, 2021. 3

[EKPS21b] Klim Efremenko, Gillat Kol, Dmitry Paramonov, and Raghuvansh R. Saxena. Tight bounds for general computation in noisy broadcast networks. In *Symposium on Foundations of Computer Science (FOCS)*, pages 634–645, 2021. 1, 3, 4

[EKPS23] Klim Efremenko, Gillat Kol, Dmitry Paramonov, and Raghuvansh R. Saxena. Protecting single-hop radio networks from message drops. In *International Colloquium on Automata, Languages, and Programming (ICALP)*, volume 261, pages 53:1–53:20, 2023. 1, 3, 4, 7

[EKS18] Klim Efremenko, Gillat Kol, and Raghuvansh Saxena. Interactive coding over the noisy broadcast channel. In *Symposium on Theory of Computing (STOC)*, pages 507–520. ACM, 2018. 0, i, 1, 2, 3, 4, 5, 6

[EKS19] Klim Efremenko, Gillat Kol, and Raghuvansh Saxena. Radio network coding requires logarithmic overhead. In *Foundations of Computer Science (FOCS)*, pages 348–369, 2019. 1, 3

[EKS20a] Klim Efremenko, Gillat Kol, and Raghuvansh R. Saxena. Interactive error resilience beyond 2/7. In *Symposium on Theory of Computing (STOC)*, pages 565–578, 2020. 1

[EKS20b] Klim Efremenko, Gillat Kol, and Raghuvansh R. Saxena. Noisy beeps. In Yuval Emek and Christian Cachin, editors, *Symposium on Principles of Distributed Computing (PODC)*, pages 418–427, 2020. 3

[EKS21] Klim Efremenko, Gillat Kol, and Raghuvansh R. Saxena. Optimal error resilience of adaptive message exchange. In *Symposium on Theory of Computing (STOC)*, pages 1235–1247, 2021. 1

[FK00] Uriel Feige and Joe Kilian. Finding OR in a noisy broadcast network. *Information Processing Letters*, 73(1-2):69–75, 2000. 3, 4

[Gal88] Robert G. Gallager. Finding parity in a simple broadcast network. *IEEE Transactions on Information Theory*, 34(2):176–180, 1988. 1, 3, 4

[Gam87] Abbas El Gamal. Open problems presented at the 1984 workshop on specific problems in communication and computation sponsored by bell communication research. *"Open Problems in Communication and Computation", by Thomas M. Cover and B. Gopinath (editors). Springer-Verlag*, 1987. 1, 3

[GH14] Mohsen Ghaffari and Bernhard Haeupler. Optimal Error Rates for Interactive Coding II: Efficiency and List Decoding. In *Foundations of Computer Science (FOCS)*, FOCS, pages 394–403, 2014. 1

[GHM18] Ofer Grossman, Bernhard Haeupler, and Sidhanth Mohanty. Algorithms for noisy broadcast with erasures. In *Colloquium on Automata, Languages, and Programming (ICALP)*, volume 107 of *LIPIcs*, pages 153:1–153:12, 2018. 4

[GHS14] Mohsen Ghaffari, Bernhard Haeupler, and Madhu Sudan. Optimal error rates for interactive coding i: Adaptivity and other settings. In *Symposium on Theory of computing (STOC)*, pages 794–803, 2014. 1, 2

[GK19] Ran Gelles and Yael T Kalai. Constant-rate interactive coding is impossible, even in constant-degree networks. *IEEE Transactions on Information Theory*, 65(6):3812–3829, 2019. 3

[GKR19] Ran Gelles, Yael Tauman Kalai, and Govind Ramnarayan. Efficient multiparty interactive coding for insertions, deletions, and substitutions. In *Symposium on Principles of Distributed Computing (PODC)*, pages 137–146, 2019. 3

[GKS08] Navin Goyal, Guy Kindler, and Michael Saks. Lower bounds for the noisy broadcast problem. *SIAM Journal on Computing*, 37(6):1806–1841, 2008. 3, 4

[Hae14] Bernhard Haeupler. Interactive channel capacity revisited. In *Foundations of Computer Science (FOCS)*, pages 226–235. IEEE, 2014. 1

[HS16] William M. Hoza and Leonard J. Schulman. The adversarial noise threshold for distributed protocols. In *Symposium on Discrete Algorithms (SODA)*, pages 240–258, 2016. 3

[JKL15] Abhishek Jain, Yael Tauman Kalai, and Allison Bishop Lewko. Interactive coding for multiparty protocols. In *Symposium on Theory of computing (STOC)*, pages 1–10, 2015. 3

[KM05] Eyal Kushilevitz and Yishay Mansour. Computation in noisy radio networks. *SIAM Journal on Discrete Mathematics (SIDMA)*, 19(1):96–108, 2005. 3, 4

[MG21] Manuj Mukherjee and Ran Gelles. Multiparty interactive coding over networks of intersecting broadcast links. *IEEE Journal on Selected Areas in Information Theory*, 2(4):1078–1092, 2021. 3

[New04] Ilan Newman. Computing in fault tolerance broadcast networks. In *Computational Complexity Conference (CCC)*, pages 113–122, 2004. 3, 4

[RS94] Sridhar Rajagopalan and Leonard J. Schulman. A coding theorem for distributed computation. In *Symposium on the Theory of Computing (STOC)*, pages 790–799, 1994. 3

[Sch92] Leonard J Schulman. Communication on noisy channels: A coding theorem for computation. In *Foundations of Computer Science (FOCS)*, pages 724–733. IEEE, 1992. 3, 4

[Yao97] Andrew Chi-Chih Yao. On the complexity of communication under noise. *invited talk in the 5th ISTCS Conference*, 1997. 3, 4

A Information Theory Preliminaries

Recall that we use sans-serif letters to denote random variables. We reserve E to denote an arbitrary event. All random variables will be assumed to be discrete and we shall adopt the convention $0 \log \frac{1}{0} = 0$. When it is clear from context, we may abbreviate the event $\mathsf{X} = x$ as just x. All logarithms are taken with base 2.

A.1 Entropy

Definition A.1 (Entropy)**.** *The (binary) entropy of* X *is defined as:*

$$\mathbb{H}(\mathsf{X}) = \sum_{x \in \mathsf{supp}(\mathsf{X})} \Pr(x) \cdot \log \frac{1}{\Pr(x)}.$$

The entropy of X *conditioned on* E *is defined as:*

$$\mathbb{H}(\mathsf{X} \mid E) = \sum_{x \in \mathsf{supp}(\mathsf{X})} \Pr(x \mid E) \cdot \log \frac{1}{\Pr(x \mid E)}.$$

Definition A.2 (Conditional Entropy)**.** *We define the conditional entropy of* X *given* Y *and* E *as:*

$$\mathbb{H}(\mathsf{X} \mid \mathsf{Y}, E) = \sum_{y \in \mathsf{supp}(\mathsf{Y})} \Pr(y \mid E) \cdot \mathbb{H}(\mathsf{X} \mid y, E).$$

Henceforth, we shall omit writing the $\mathsf{supp}(\cdot)$ when it is clear from context.

Lemma A.3 (Chain Rule for Entropy)**.** *It holds for all* X*,* Y*,* Z *and* E *that:*

$$\mathbb{H}(\mathsf{XY} \mid \mathsf{Z}, E) = \mathbb{H}(\mathsf{X} \mid \mathsf{Z}, E) + \mathbb{H}(\mathsf{Y} \mid \mathsf{X}, \mathsf{Z}, E).$$

Lemma A.4 (Conditioning reduces Entropy)**.** *It holds for all* X*,* Y*,* Z *and* E *that:*

$$\mathbb{H}(\mathsf{X} \mid \mathsf{Y}, \mathsf{Z}, E) \leq \mathbb{H}(\mathsf{X} \mid \mathsf{Z}, E).$$

Equality holds if and only if X *and* Y *are independent conditioned on* Z, E.

Corollary A.5 (Corollary of Lemmas A.3 and A.4)**.** *It holds for all* X, Y, Z *and* E *that:*

$$\mathbb{H}(\mathsf{XY} \mid \mathsf{Z}, E) = \mathbb{H}(\mathsf{X} \mid \mathsf{Z}, E) + \mathbb{H}(\mathsf{Y} \mid \mathsf{Z}, E).$$

Lemma A.6. *It holds for all* X *and* E *that:*

$$0 \leq \mathbb{H}(\mathsf{X} \mid E) \leq \log(|\mathsf{supp}(\mathsf{X})|).$$

The second inequality is tight if and only if X *conditioned on* E *is the uniform distribution over* $\mathsf{supp}(\mathsf{X})$.

A.2 Mutual Information

Definition A.7 (Mutual Information)**.** *The mutual information between* X *and* Y *is defined as:*

$$\mathbb{I}(\mathsf{X} : \mathsf{Y}) = \mathbb{H}(\mathsf{X}) - \mathbb{H}(\mathsf{X} \mid \mathsf{Y}) = \mathbb{H}(\mathsf{Y}) - \mathbb{H}(\mathsf{Y} \mid \mathsf{X}).$$

The mutual information between X *and* Y *conditioned on* Z *is defined as:*

$$\mathbb{I}(\mathsf{X} : \mathsf{Y} \mid \mathsf{Z}) = \mathbb{H}(\mathsf{X} \mid \mathsf{Z}) - \mathbb{H}(\mathsf{X} \mid \mathsf{YZ}) = \mathbb{H}(\mathsf{Y} \mid \mathsf{Z}) - \mathbb{H}(\mathsf{Y} \mid \mathsf{XZ}).$$

Fact A.8. *We have* $0 \leq \mathbb{I}(\mathsf{X} : \mathsf{Y} \mid \mathsf{Z}) \leq \mathbb{H}(\mathsf{X} \mid \mathsf{Z}) \leq \mathbb{H}(\mathsf{X})$.

Fact A.9 (Chain Rule for Mutual Information)**.** *If* W, X, Y, Z *are random variables, then*

$$\mathbb{I}(\mathsf{WX} : \mathsf{Y} \mid \mathsf{Z}) = \mathbb{I}(\mathsf{W} : \mathsf{Y} \mid \mathsf{Z}) + \mathbb{I}(\mathsf{X} : \mathsf{Y} \mid \mathsf{WZ}).$$

A.3 KL Divergence

Definition A.10 (KL Divergence)**.** *If* μ, ν *are two distributions over the same (finite) set* Ω*, the Kullback-Leibler (KL) Divergence between* μ *and* ν *is defined as:*

$$\mathbb{D}(\mu \mid\mid \nu) = \sum_{\omega \in \Omega} \mu(\omega) \cdot \log \frac{\mu(\omega)}{\nu(\omega)}.$$

For a finite non-empty set S, we shall use $\mathsf{unif}(S)$ to denote the uniform distribution over S. We omit S from the notation when it is clear from the context. We use $\mathsf{dist}(\mathsf{X} \mid E)$ to denote the distribution of the random variable X conditioned on the event E.

Lemma A.11. *Let* μ, ν *be two distributions with the same support* Ω*. We have* $\mathbb{D}(\mu \mid\mid \nu) \geq 0$.

Lemma A.12. *Let* X *be a random variable uniformly distributed over a set* Ω *and* $S_1 \subseteq S_2 \subseteq \Omega$ *be given. Let* E *be an event such that* X *conditioned on* E *is supported on* S_1*, we*

have:

$$\mathbb{D}(\mathsf{dist}(\mathsf{X} \mid E) \mid\mid \mathsf{dist}(\mathsf{X} \mid \mathsf{X} \in S_2)) \geq \log \frac{|S_2|}{|S_1|}.$$

Lemma A.13. *Let* X *be a random variable uniformly distributed over a set* Ω *and* $S_1 \subseteq S_2 \subseteq \Omega$ *be given:*

$$\mathbb{D}(\mathsf{dist}(\mathsf{X} \mid \mathsf{X} \in S_1) \mid\mid \mathsf{dist}(\mathsf{X} \mid \mathsf{X} \in S_2)) = \log \frac{|S_2|}{|S_1|}.$$

A.4 Total Variation Distance

Definition A.14 (Total variation distance)**.** *Let* μ, ν *be two distributions over the same (finite) set* Ω*. The total variation distance between* μ *and* ν *is defined as:*

$$\|\mu - \nu\|_{\mathrm{TV}} = \max_{\Omega' \subseteq \Omega} \sum_{\omega \in \Omega'} \mu(\omega) - \nu(\omega).$$

Fact A.15 (Pinsker's inequality)**.** *Let* μ, ν *be two distributions over the same set* Ω*. It holds that:*

$$\|\mu - \nu\|_{\mathrm{TV}} \leq \sqrt{\frac{1}{2} \cdot \mathbb{D}(\mu \mid\mid \nu)}.$$

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il