Channel Capacity for Adversaries with Computationally Bounded Observations

Eric Ruzomberka, Chih-Chun Wang and David J. Love
School of Electrical and Computer Engineering, Purdue University, West Lafayette, USA
email: {eruzombe,chihw,djlove}@purdue.edu

Abstract—We study reliable communication over point-to-point adversarial channels in which the adversary can observe the transmitted codeword via some function that takes the \( n \)-bit codeword as input and computes an \( r n \)-bit output for some given \( r \in [0, 1] \). We consider the scenario where the \( r n \)-bit observation is computationally bounded – the adversary is free to choose an arbitrary observation function as long as the function can be computed using a polynomial amount of computational resources. This observation-based restriction differs from conventional channel-based computational limitations, where in the later case, the resource limitation applies to the computation of the (adversarial) channel error. For all \( r \in [0, 1 - H(p)] \) where \( H(\cdot) \) is the binary entropy function and \( p \) is the adversary’s error budget, we characterize the capacity of the above channel. For this range of \( r \), we find that the capacity is identical to the obviously complete setting \( (r = 0) \). This result can be viewed as a generalization of known results on myopic adversaries and channels with active eavesdroppers for which the observation process depends on a fixed distribution and fixed-linear structure, respectively, that cannot be chosen arbitrarily by the adversary.

I. INTRODUCTION

Beginning with Shannon’s seminal paper [1], early channel coding research observed that fundamental coding limits are highly sensitive to channel modeling assumptions. This sensitivity is demonstrated by a gap in capacity between the two classical models: the Shannon model in which channel errors follow a known random distribution and the Hamming model in which error patterns are worst-case for some fixed number of bit errors. In the design of robust codes, the more conservative Hamming model is particularly attractive as it makes no assumptions about the channel distribution and thus any resulting conclusion is robust against a wide variety of channel imperfections. The downside of the Hamming model, however, is that it admits a smaller capacity than the Shannon model. In many cases, the gap in capacity is large [2].

A. Closing the gap

Recent research efforts have made progress in closing this gap by considering settings in between the two classical models. Ideally, the following two properties hold for a good channel model:

**Property 1:** The channel is mild in the sense that its capacity coincides with the Shannon model capacity.

**Property 2:** The channel inherits conservative aspects of the Hamming model. In particular, the channel may vary in an arbitrary manner unknown to the communicating parties.

In the following Section [2–8] we focus on two different approaches which have had success towards producing good channel models. These approaches have been to 1) bound the channel’s computing power (i.e., computational complexity) [3], [4] and 2) bound the information known to the channel about the communication scheme [5–13].

B. Complexity Bounded Channels and Oblivious Channels

Consider a transmitter Alice who wishes to communicate a message \( m \) from a set of \( M \) possible messages over a noisy channel to a receiver Bob. To protect the message from noise corruption, Alice encodes \( m \) into an \( n \) bit codeword \( x \) of rate \( R = (1/n) \log M \) and transmits \( x \) over the channel. The channel adds an \( n \)-bit error vector \( e \) to \( x \), and Bob receives the binary channel output \( y = x \oplus e \). The channel is controlled by an adversary who chooses \( e \) to prevent reliable (unique) decoding by Bob. For an error budget \( p \in (0, 1/2) \), the adversary can only induce \( pn \) bit flips, i.e., the Hamming weight of \( e \) must be bounded above by \( pn \).

In the computationally bounded model (first proposed by Lipton [3]), the adversary computes \( e \) using limited computational resources, e.g., via an algorithm that takes a finite number of computational steps. This model has the appeal of sufficiently describing practical channels, including channels with memory and channels governed by natural, but unknown processes. However, the computationally bounded model can be severe – an impossibility result of Guruswami and Smith [4] is that the model’s capacity can be less than the Shannon capacity, and can even be 0 when the latter is positive. Thus, Property 1 does not hold for the computationally bounded model.

Another existing approach is the partially oblivious model, where the adversary chooses \( e \) using incomplete side-information about the transmitted codeword \( x \). This model includes myopic channels, e.g., [5–7], causal channels, e.g., [3], [9]–[10], channels with active eavesdroppers, e.g., [13], and some arbitrarily varying channels (AVCs), e.g., [11], [12]. Although the model can vary between works, the setting usually has the following general structure: for \( r \in [0, 1] \) and some (deterministic) observation function \( f_n : \{0, 1\}^n \rightarrow \{0, 1\}^{rn} \), the adversary makes an \( r n \)-bit observation \( f_n(x) \) of codeword
\( x \) prior to choosing \( e \). The special cases \( r = 0 \) and \( r = 1 \) correspond to no information (i.e., completely oblivious) and perfect information (i.e., omniscient), respectively.

Property 1 can hold for the oblivious model when \( r \) is sufficiently small. However, Property 2 does not hold for many oblivious channels in the literature. For example, in the myopic channel model, the adversary randomly draws \( f_n \) from a known distribution. For Property 2 to hold, however, we must allow \( f_n \) to be arbitrarily chosen and require Alice and Bob to devise their communication scheme without knowledge of \( f_n \). This is equivalent to the adversary choosing a worst-case \( f_n \) for a fixed \( r \) — a model studied by Langberg [13] under the name of the \((1-r)\)-oblivious channel. The capacity of the \((1-r)\)-oblivious channel remains an open problem, where the best known lower bound is given by [13] and will be summarized in Section 1.

C. This Work

In this paper, we consider a channel model that has qualities of both the computationally bounded model and the partially oblivious model. We do so by requiring the adversary to observe \( x \) via an \( rn \)-bit observation function \( f_n \) that is computationally bounded.

Specifically, for fixed positive integers \( c \) and \( s \), the adversary chooses a sequence of observation functions \( f_n(x), \forall n \geq 1 \) that belongs to \( CPX(r,cn^s) \) — the set of observation functions with \( n \) input bits and \( rn \) output bits that can be computed by a Boolean circuit with at most \( cn^s \) gates. We allow the choice of \( f_n \) to be unknown to Alice or Bob. On the other hand, the \( f_n \) chosen by the adversary can depend on the codebook of Alice but cannot depend on the actual message being sent.

Using the observation function \( f_n \) of its choice, the adversary observes \( f_n(x) \) and chooses \( e \) with no computational bound. We refer to the above adversary as a \( CPX(r,cn^s) \)-oblivious adversary. By construction, Property 2 holds for a channel controlled by a \( CPX(r,cn^s) \)-oblivious adversary due to \( f_n \) being unknown to Alice or Bob.

Our imposed computational restriction is practical and sufficiently models realistic adversarial channels. A channel controlled by a \( CPX(r,cn^s) \)-oblivious adversary closely approximates a \((1-r)\)-oblivious channel (i.e., a channel controlled by a \( CPX(r,\infty) \)-oblivious adversary) without weakening the power of the adversary too much. Indeed, the adversary is quite strong. To illustrate its strength, if for a sequence of functions \( \{f_n\}_{n=1}^{\infty} \) and for \( c, s \geq 1 \) there exists a finite \( n_0 \) such that for all \( n \geq n_0 \), \( f_n \notin CPX(r,cn^s) \), then the sequence is widely regarded to be an infeasible computation [15]. The technical value of the computational constraint is to bound the number of observation functions that the adversary can choose from.

D. Results

We assume that Alice uses deterministic encoding and we consider capacity under the diminishing average error probability criterion in which the probability of decoding error is averaged over the message set. Under the above model, the Shannon capacity is \( C_{Shannon}(p) = 1 - H(p) \) where \( H(p) = -p \log p - (1-p) \log(1-p) \) is the binary entropy function [12], [13]. We remark that \( C_{Shannon}(p) \) is achievable in our model if the \( CPX(r,cn^s) \)-oblivious adversary ignores its observation \( f_n(x) \) and naively chooses \( e \) randomly from the set all possible error vectors with Hamming weight \( pn \).

For \( p \in (0,1/2), r \in [0,1] \) and positive integers \( c, s \), let \( C(p,r,cn^s) \) denote the capacity of a channel controlled by a \( CPX(r,cn^s) \)-oblivious adversary with error budget \( p \). Similarly, let \( C(p,r,\infty) \) denote the capacity of \((1-r)\)-oblivious channel. The following result shows that Property 1 holds for our model for a wide range of \( r \).

Theorem 1. For \( p \in (0,1/2), r \in [0, C_{Shannon}(p)) \), and positive integers \( c \) and \( s \), \( C(p,r,cn^s) = C(p,0,cn^s) = C(p,0,\infty) = C_{Shannon}(p) \).

We share a few remarks on the above theorem. When \( r < C_{Shannon}(p) = 1 - H(p) \), Theorem 1 implies that the adversary can do no better than to ignore its side-information \( f_n(x) \) and choose \( e \) randomly from the set of all \( n \)-bit vectors with Hamming weight \( pn \). Additionally, we note that the largest known lower bound on \( C(p,r,\infty) \) is \( 1 - r - H(p) \) for \( r \in \left[ 0, \frac{1 - H(p)}{2} \right] \) [13]. Since \( C(p,r,\infty) \) is a lower bound to \( C(p,r,cn^s) \), Theorem 1 significantly sharpens the best known lower bound of \( C(p,r,cn^s) \) to an exactly tight characterization. For \( r > C_{Shannon}(p) \), an immediate lower bound of \( C(p,r,cn^s) \) is given by the Gilbert-Varshamov (GV) bound.
messages whose corresponding codewords are contained in the
Bob performs list decoding by creating a list
$L \subseteq \mathbb{B}^n$ of size bounded above by
the result is generalized by Theorem 1 after observing that a
CPX scheme but not exact knowledge of the actual message $m_0$.
In particular, the adversary knows Alice’s codebook $C_n$ and
is partially oblivious to the transmitted codeword $x$. By
partially oblivious, we mean that for $r \in [0,1)$ and some
function $f_n : \{0,1\}^n \rightarrow \{0,1\}^{rn}$, the adversary observes
a realization $\psi$ of the random variable $\Psi = \Psi(m_0) = f_n(C_n(m_0)) = f_n(x)$. Due to the adversary’s computational
bound, for positive integers $c,s$, the adversary chooses $f_n$ from the set $\text{CPX}(r, cn^s)$ (we provide a rigorous definition
of CPX($r, cn^s$) in Section II-C) using its knowledge of $C_n$
but not the realization of $m_0$. The chosen function $f_n$ is not
revealed to Alice or Bob. Finally, the adversary chooses $e \in \mathbb{B}_pn(0)$ based on the knowledge of the codebook $C_n$ and
the observation $\Psi(m_0)$. We refer to the above adversary as
the CPX($r, cn^s$)-oblivious adversary with error budget $p$.

C. Adversary’s Complexity Constraint

For $r \in [0,1]$ and positive integers $c,s$, we precisely define the set $\text{CPX}(r, cn^s)$. Let $\mathcal{F}_{n,r}$ denote the set of all
Boolean functions of the form $f_n : \{0,1\}^n \rightarrow \{0,1\}^{rn}$. To
define CPX($r, cn^s$), we first define the circuit complexity of
a function $f_n \in \mathcal{F}_{n,r}$.

A Boolean circuit $B_n$ is an acyclic directed graph where
each node is either an input node (with in-degree 0) or a
logic gate (with in-degree 2). All nodes in $B_n$ have out-degree
1 with unbounded fan-out and each logic gate computes an
arbitrary Boolean function from $\{0,1\}^2$ to $\{0,1\}$. The size
of $B_n$ is the total number of nodes in $B_n$ (input nodes and
logic gates). Note that an observation function $f_n \in \mathcal{F}_{n,r}$
can be computed by some Boolean circuit that takes $n$ bits as
input and produces $rn$ bits as output. The circuit (size)
complexity of an observation function $f_n \in \mathcal{F}_{n,r}$ is the size of
the smallest size Boolean circuit $B_n$ that can compute $f_n$. We
define CPX($r, cn^s$) to be the set of all functions $f_n \in \mathcal{F}_{n,r}$
with a circuit complexity of at most $cn^s$. In modern complexity
theory, the study of circuit complexity is a common approach
to proving lower bounds on the complexity of certain problems
[15].

D. Capacity

For a fixed $[n,M]$ codebook $C_n$, the (average) probability
of decoding error $P_e(C_n)$ is defined as the maximum over all
$f_n \in \text{CPX}(r, cn^s)$ of the quantity
$$\mathbb{E}_\Psi \left[ \max_{e \in \mathbb{B}_pn(0)} \mathbb{P}(\hat{m}(e,m_0) \neq m_0 | \Psi(m_0) = \psi) \right]$$

The fact that $\Psi$ is a random variable follows from its dependency on
the random variable $m_0$.

However, the adversary knows that $m_0$ is drawn uniformly from $[M]$.
where the probability measure $P_{m_0} (\cdot)$ is w.r.t. the distribution of $m_0$, and the expectation $E_{\Psi} \left[ \cdot \right] = \sum_{\psi \in \{0,1\}^{n_1}} P(\Psi(m_0) = \psi)$. Given the above channel model, we can define achievable rate in the usual way.

**Definition 1** (Achievable Rate). For $p \in (0, 1/2)$, $r \in [0, 1]$, and positive integers $c, s$, a rate $R \in [0, 1]$ is said to be $(c, s)$-achievable if for any $\epsilon_c > 0$, there exists an $n_0$ such that for all $n \geq n_0$, there exists an $[n, M]$ codebook $C_n$ such that $P_e(C_n) \leq \epsilon_c$.

For $p \in (0, 1/2)$, $r \in [0, 1]$, and positive integers $c, s$, we define the capacity $C(p, r, cn^s)$ as the supremum of $(c, s)$-achievable rates.

### III. PROOF OUTLINE, OVERVIEW OF PROOF TECHNIQUE

In this section, we outline the proof of Theorem 1 and discuss an overview of our proof technique. A detailed proof of Theorem 1 can be found in Section IV.

#### A. Achievability Scheme

For our proof of Theorem 1, we construct a specific $C_n$.

**Encoder Construction:** Alice’s $[n, M]$ codebook $C_n$ is constructed as follows. Let $p \in (R, C_{Shannon}(p))$. Codebook $C_n$ is a concatenation of two codebooks: an outer $[pn, M]$ codebook $C_{out} : [M] \to \{0,1\}^{pn}$ and for $N = 2^{m}$, an inner $[n, N]$ codebook $C_{in} : [0,1]^m \to [0,1]^n$. Encoding proceeds as follows. First, Alice encodes $m_0$ with $C_{out}$ where we denote the resulting codeword as $C_{out}(m_0)$. Subsequently, Alice encodes $C_{out}(m_0)$ with $C_{in}$ where we denote the resulting codeword as $C_n = C_n(C_{out}(m_0))$. After encoding, Alice transmits the codeword $x = C_n(m_0)$ over the channel.

**Decoder Construction:** Bob’s list decoder is constructed as follows. Given the channel output $y$, Bob first performs list decoding by forming a list $L_{in}(m_0, e, C_{in})$ of all words $w \in \{0,1\}^{pn}$ such that $C_{in}(w)$ is contained in the ball $B_{pn}(y)$. After list decoding, Bob refines the list (i.e., Bob performs disambiguation) by removing all words $w \in L_{in}$ that are inconsistent with $C_{out}$; we say that a word $w$ is inconsistent with $C_{out}$ if $w \neq C_{out}(m)$ for any $m \in [M].$

Denote the refined list as $L_{out}$. After $L_{out}$ is refined, decoding decision is made according to the following rules. If $|L_{out}| = 1$, then we have exactly one $m \in [M]$ s.t. $C_{out}(m) \in L_{out}$, and the decoder outputs $\hat{m} = m$. If $L_{out}$ is empty or $|L_{out}| > 1$, then the decoder declares an error by setting $\hat{m}$ to an error symbol. We say that a decoding error occurs if $\hat{m} \neq m_0$. However, by the list decoding logic, $C_{out}(m_0)$ is guaranteed to be in $L_{out}$, and so the only non-trivial decoding error event occurs when $|L_{out}| > 1$.

**Probability of Error:** Given the above construction, the probability of decoding error $P_e(C_{out}, C_{in})$ can be written as

$$\max_{f_n \in \text{CPX}(r, cn^s)} \mathbb{E}_{\Psi} \left[ \max_{e \in B_{pn}(0)} \mathbb{P}_{m_0} (|L_{out}| > 1 | \Psi(m_0) = \psi) \right]$$

which in turn is equal to the maximum over all $f_n \in \text{CPX}(r, cn^s)$ of the quantity

$$\mathbb{E}_{\Psi} \left[ \max_{e \in B_{pn}(0)} \mathbb{P}_{m_0} \left( \bigcup_{i=1}^{\rho} \{ \Psi(m_i) = \psi \} \right) \right]$$

where $\psi = \max_{i=1}^{\rho} \mathbb{P}_{m_0} (|L_{out}| > 1 | \Psi(m_i) = \psi)$ for each $i = 1, \ldots, \rho$.

In this section, we outline the proof of Theorem 1 and discuss an overview of our proof technique. A detailed proof of Theorem 1 can be found in Section IV.

**Definition 1** (Achievable Rate). For $p \in (0, 1/2)$, $r \in [0, 1]$, and positive integers $c, s$, a rate $R \in [0, 1]$ is said to be $(c, s)$-achievable if for any $\epsilon_c > 0$, there exists an $n_0$ such that for all $n \geq n_0$, there exists an $[n, M]$ codebook $C_n$ such that $P_e(C_n) \leq \epsilon_c$.

For $p \in (0, 1/2)$, $r \in [0, 1]$, and positive integers $c, s$, we define the capacity $C(p, r, cn^s)$ as the supremum of $(c, s)$-achievable rates.

**B. Outline of the proof of Theorem 1**

**Random Coding Argument:** In the sequel, for $p \in (0, 1/2)$, $r \in [0, 1 - H(p)]$, positive integers $c, s$, and for any $\epsilon_c, \epsilon_r > 0$, we set the outer-code rate $\rho = C_{Shannon}(p) - \epsilon_r$ and the inner-out combined code rate $R = \rho - \epsilon_r$. Thus, rate $R$ can be arbitrarily close to $C_{Shannon}$. To prove Theorem 1, we show that the rate $R$ is $(c, s)$-achievable. We show this by using a random-coding argument in conjunction with the code construction presented in Section III-A.

The argument states that for any fixed $[pn, M]$ outer code $C_{out}$ that is a 1:1 function and for any $\epsilon_c > 0$, if there exists some $n_0$ such that for all $n \geq n_0$ there exists some non-empty set $G$ of $[n, M]$ codebooks such that for any $C_{in} \in G$ we have $P_e(C_{in}, C_{out}) \leq \epsilon_c$, then Theorem 1 holds.

**Setup:** In the sequel, we drop the dependency on $C_{out}$ from all notation due to the outer codebook being fixed. To apply the random-coding argument, we first apply a simple union bound to $P_e(C_{in})$ and bound it above by $P_{\text{ub}}(C_{in})$ defined as

$$\max_{f_n \in \text{CPX}(r, cn^s)} \sum_{i=1}^{N} \mathbb{E}_{\Psi} \left[ \max_{e \in B_{pn}(0)} q_i(f_n, \psi, e, C_{in}) \right]$$

where for $f_n \in \text{CPX}(r, cn^s)$, $\psi \in \{0,1\}^{pn}$, $e \in B_{pn}(0)$ and $i \in [N]$, we define

$$q_i(f_n, \psi, e, C_{in}) = \mathbb{P}(w_i \in L_{in} \cup \Psi(m_i) = \psi)$$

to be the probability that $w_i$ results in a decoding error.

To bound $P_{\text{ub}}$, we leverage the list decodable properties of the inner codebook.

**Definition 2.** For $\ell > 0$, an $[n, N]$ codebook $C_{in}$ is said to be $[l, p]$ list decodable if $|L_{in} \cap B_{pn}(y)| \leq \ell$ for every $y \in \{0,1\}^{pn}$.

For each $n$ large enough, to show that there exists an $[n, N]$ codebook $C_{in}$ such that $P_{\text{ub}}(C_{in}) \leq \epsilon_c$, it is sufficient to show that for some number $L = O(1/\epsilon_r)$ (independent of $n$), the probability (over the choice of inner codebook) that $C_{in}$ is not $[L, p]$ list decodable or $q_i(f_n, \psi, e, C_{in}) > \epsilon_c/L$ for some
We note the probability that $C_{in}$ is not $[L,p]$ list decodable is vanishing as $n \to \infty$ following known results on the list decodability of random codes (e.g., [8 Claim A.15]). Hence, following a union bound, we only need to show that with probability bounded away from 1, we have $q_{i}(f_{n},\psi,e,C_{in}) > \epsilon_{n}/L$ for some parameters as described above. When confusion can be avoided, we drop the notated dependencies and subscripts of $q_{i}(f_{n},\psi,e,C_{in})$ and simply write $q(C_{in})$ to emphasize the dependency on $C_{in}$.

**Analysis of $q(C_{in})$:** We fix $i \in [L]$, $f_{n} \in \text{CPX}(r,cn^{s})$, $\psi \in \{0,1\}^{r}$ and $e \in P_{pn}(0)$ and for $n = 1,2,\ldots$ we study the concentration of measure of $q(C_{in})$ around its expectation $E_{C_{in}}[q]$ (here, the expectation is w.r.t. the distribution of $C_{in}$). We do so by deriving concentration inequalities via the logarithmic Sobolev inequalities, e.g., [18]. This derivation is also known as the entropy method.

An example of a common inequality derived via the entropy method is as follows. Define the variation of $q(C_{in})$ as $V(C_{in}) = \sum_{j=1}^{N} E_{C_{in}}[q(C_{in} - q(C_{in}(j,z)))]^{2}$ where codebook $C_{in}(j,z)$ is equal to $C_{in}$ with the $j$th codeword replaced with the codeword $z$ uniformly distributed in $\{0,1\}^{r}$. The quantity $V(C_{in})$ captures how smoothly $q(C_{in})$ varies for incremental changes to $C_{in}$. For $a > 0$, we say that $q$ is $a$-smooth if for all $[n,N]$ codebooks $C_{in}$ we have $V(C_{in}) \leq a$.

**Proposition 1 ( [15 Corollary 3]).** Suppose there exists an $a = a(n) > 0$ such that $q$ is $a$-smooth. For $\lambda > 0$, we have

$$
\mathbb{P}_{C_{in}}(q - E_{C_{in}}[q] > \lambda) \leq \exp\left\{ \frac{\lambda^{2}}{4a} \right\}
$$

where the probability $\mathbb{P}_{C_{in}}$ is w.r.t. the distribution of $C_{in}$.

To apply Proposition 1, one can first set $\lambda = \epsilon_{n}/L - E_{C_{in}}[q]$, and then find a value of $a$ small enough such that $q$ is $a$-smooth and the R.H.S. of the inequality approaches 0 in the limit $n \to \infty$. In practice, this value of $a$ is difficult to find. The difficulty arises from the fact that for a small subset of $[n,N]$ codebooks $C_{in}$, the variation $V(C_{in})$ is large. Thus, Proposition 1 cannot be used directly.

**Handling large variation:** To address this issue, we take the following bootstrapping approach. We first approximate $q(C_{in})$ with a function $q'(C_{in})$ that is equal to $q(C_{in})$ with high probability over the choice of $C_{in}$ and has a small variation $V'(C_{in})$ for all $[n,N]$ codebooks $C_{in}$. We then imply the concentration of $q$ by showing that the approximation $q'$ is concentrated via an entropy-method-type concentration inequality that resembles Proposition 1.

To approximate $q(C_{in})$, we first define a set $T$ of typical $[n,N]$ codebooks such that $\mathbb{P}_{C_{in}}(C_{in} \notin T)$ is small. If $C_{in} \in T$, we define $q'(C_{in}) = q(C_{in})$. If otherwise $C_{in} \notin T$, we define $q'(C_{in})$ such that for sufficiently small $a > 0$, $q'$ is $a$-smooth.

Lastly, we briefly discuss the reasoning behind our choice of code construction. We remark that our construction and the resulting definition of $q$ help us to show that $V'(C_{in})$ is sufficiently small when $C_{in} \notin T$. Concatenated coding and list decoding/refinement allow us to isolate for $i = 1,\ldots,N$ the effect of the $i$th codeword of $C_{in}$ on $P_{i}(C_{in})$ and show $V'(C_{in}) = O(i) \forall C_{in} \notin T$. Without the construction, for non-typical codebooks, the function under analysis may have a variation equal to $O(2^{Rn})$ which is too large to apply our concentration inequalities.

**Prior work:** Our above approach is inspired by Langberg’s framework [13] to study concentration of measure when the function under analysis is non-smooth. The main technical contribution of [13] is to carefully define the typical set $T$ based on the codebooks’ list decodable properties in a way where one can then apply Vu’s martingale-type concentration inequalities for non-smooth functions [19]. We follow Langberg’s framework by also defining typicality in terms of list decodability. However, we use entropy-method-type concentration inequalities.

The major technical difference between our work and [13] lies at the definition of smoothness. For $a > 0$, reference [13] defines smoothness in terms of $a$-Lipschitz: for the quantity $W(C_{in}) = N \max_{j \in [N],z \in \{0,1\}^{r}} |q(C_{in}) - q(C_{in}(j,z))|^{2}$, $q$ is $a$-Lipschitz if for all $[n,N]$ codebooks $C_{in}$ we have $W(C_{in}) \leq a^{4}$.

We remark that for $a > 0$, $q$ is $a$-smooth if $q$ is $a$-Lipschitz, and thus $a$-Lipschitz is a stronger notion of smoothness than the notion used in our work.

The advantage to caracterizing smoothness using $a$-smooth as opposed to $a$-Lipschitz is apparent in the following observation. Suppose that for $a_{L} > 0$, $q$ is $a_{L}$-Lipschitz. Then one can usually find a smaller value $a_{S} \in (0,a_{L})$ such that $q$ is $a_{S}$-smooth, and in turn, leverage the $a_{S}$-smooth criterion to apply tighter concentration inequalities than those reliant on the $a_{L}$-Lipschitz criterion. Indeed, we take this approach to show that the probability that $q$ is not concentrated is at most $2^{-2^{\Omega(n)}}$, whereas bounds on the order of $2^{-\Omega(n)}$ are obtainable using $a_{L}$-Lipschitz together with the framework of [13]. The cost of this approach is that finding a smaller value $a_{S}$ can require significant effort compared to finding $a_{L}$. Indeed, our proof of Theorem 1 invests significant effort into finding $a_{S}$.

**Computational Bound:** Lastly, we briefly discuss the role that the adversary’s computational bound plays in our random-coding argument. Recall that to show a rate $R$ is $(c,s)$-achievable, it is sufficient to show that there exists one $[n,N]$ codebook $C_{in}$ such that quantity 1 is small for all $f_{n} \in \text{CPX}(r,cn^{s})$. Showing the existence of this codebook becomes difficult if the set $\text{CPX}(r,cn^{s})$ contains many functions. We simplify our search for this codebook by bounding the number of functions in the set $\text{CPX}(r,cn^{s})$. We remark that the set $\text{CPX}(r,cn^{s})$ can be shown to have $2^{\Omega(n)}$
functions, and therefore, $\text{CPX}(r, cn^s)$ is much smaller than the set of functions $\text{CPX}(r, \infty) = \mathcal{F}_{n,r}$ with unbounded circuit complexity which has $2^m n^{2s}$ functions.

IV. PROOF OF THEOREM 1

The following setup will be used throughout the proof of Theorem 1:

1) Fix $p \in (0, 1/2)$, $r \in (0, C_{\text{Shannon}}(p))$ and $c, s$ to be positive integers.
2) Fix parameters $\delta, \Delta \delta, \varepsilon_p, \varepsilon_R > 0$ such that the following 2 Conditions hold:
   - Condition 1. $r < -H(p) - \delta - \Delta \delta - \varepsilon_p - \varepsilon_R$
   - Condition 2. $\varepsilon_R \in (0, (\frac{\delta}{2s - \frac{1}{30}}))$

Set the outer-code rate $\rho = C_{\text{Shannon}}(p) - \varepsilon_p$ and the inner-outer combined code rate $R = \rho - \varepsilon_R$.

3) For blocklength $n = 1, 2, \ldots$, let the codebook $C_n$ be the code construction described in Section III-A.

4) Fix the $[m, M]$ outer codebook $C_{\text{out}}$ to be any 1:1 function from $\{0, 1\}^m$ to $\{0, 1\}^r$. Let the $[n, N]$ inner codebook $C_n$ drawn from distribution $\Omega[n, N]$. Note that $C_n = C_{\text{in}} \circ C_{\text{out}}$ is Alice’s $[n, M]$ codebook.

A. Preliminaries

The following preliminary results characterize the list-decodability properties of a randomly picked $C_n$.

**Lemma 1.** Let $\ell = \ell(n) > 0$ be $\omega(n)$ (i.e., $\lim_{n \to \infty} \ell(n)/n = \infty$). For large enough $n$, a codebook $C_n$ drawn from distribution $\Omega[n, N]$ is $[\ell, p]$ list decodeable w.p. greater than $1 - 2^{-\ell(n)/4}$. Proof is in Appendix B.

**Lemma 2 (Claim A.15).** For $L > 1/\varepsilon_p$ and for large enough $n$ (depending only on $\varepsilon_p$), a codebook $C_n$ drawn from distribution $\Omega[n, N]$ is $[L, p]$ list decodeable w.p. greater than $1 - \frac{1}{n}$.

**Lemma 3.** Let $A \subseteq [0, 1]^n$ and let $\mu = E_{C_{\text{in}}} [A \cap C_{\text{in}}] = 2^{-(1-\alpha)n} |A|$. For $t_L < \mu$ and $t_U > \mu$, and for a codebook $C_{\text{in}}$ drawn from distribution $\Omega[n, N]$ (and resulting $[n, M]$ codebook $C_n = C_{\text{in}} \circ C_{\text{out}}$), the set $A \cap C_{\text{in}}$ has bounded size w.h.p such that $|A \cap C_{\text{in}}| > t_L$ and $|A \cap C_{\text{in}}| > t_U$ w.p. at most $2 \exp\left(-\frac{(u-t_L)^2}{4u}ight)$ and $2 \exp\left(-\frac{(u-t_U)^2}{4u}ight)$, respectively. Proof is in Appendix C.

Lastly, we introduce some notation for describing the adversary’s observation. For the adversary’s observation function $f_n \in \mathcal{F}_{n,r}$ and realization $\psi = f_n(C_n(m_0))$, we define the observation set $O_{\psi} = \{ z \in \{0, 1\}^n : f_n(z) = \psi \}$. We note that observing $\Psi(m_0) = \psi$ is equivalent to knowing that $C_n(m_0) \in O_{\psi}$. Hence, the following two perspectives are equivalent: the adversary chooses $f_n \in \mathcal{F}_{n,r}$ and the adversary chooses a partition $\mathcal{O} = (O_1, \ldots, O_{2^r})$ of the space $\{0, 1\}^n$ consisting of $2^rn$ non-empty observation sets. With an abuse of notation, for $f_n \in \text{CPX}(r, cn^s)$, we write $\mathcal{O} \in \text{CPX}(r, cn^s)$ to denote the observation sets corresponding to $f_n$.

For $\mathcal{O} \in \text{CPX}(r, cn^s)$, $\psi \in \{0, 1\}^r$, $e \in B_{pn}(0)$, and $i \in [N]$, note that definition (4) of $q_i$ can be written as the sum over $m \in [M]$ of

$$\Pr_{m_0}(\mathcal{O}_i(\psi, e, C_n)), \psi \in \{0, 1\}^r, e \in B_{pn}(0),$$

Defining $\Phi_{i, m}(\mathcal{O}, \psi, e, C_n) = 1(\psi(m, e, C_n) \in \mathcal{I}_m \cap \mathcal{C}(m))$, and for large $n$, $\psi \in \{0, 1\}^r$, $e \in B_{pn}(0)$, the quantity $E_{C_{\text{in}}} [q_i(\mathcal{O}, \psi, e, C_{\text{in}})]$ is going to 0.

**Proof.** Let $\emptyset$ denote the empty set and let $i \in [N], \mathcal{O} \in \text{CPX}(r, cn^s), \psi \in \{0, 1\}^r, e \in B_{pn}(0)$. Observe that for any $[n, N]$ codebook $C_n$, $q_i(\mathcal{O}, \psi, e, C_{\text{in}})$ is bounded above by $\Pr_{m_0}(\mathcal{I}_m \cap \mathcal{C}(m_0, e, C_{\text{in}}) \neq \emptyset)$. Hence, we have

$$E_{C_{\text{in}}} [q_i(\mathcal{O}, \psi, e, C_{\text{in}})] \leq E_{C_{\text{in}}} [\Pr_{m_0}(\mathcal{I}_m \cap \mathcal{C}(m_0, e, C_{\text{in}}) \neq \emptyset)]$$

To prove Lemma 4, it is sufficient to show that for all $m \in [M]$ and all $i \in [N], \mathcal{O} \in \text{CPX}(r, cn^s), \psi \in \{0, 1\}^r, e \in B_{pn}(0)$, the quantity $E_{C_{\text{in}}} [1(\mathcal{I}_m \cap \mathcal{C}(m, e, C_{\text{in}}) \neq \emptyset)]$ is going to zero uniformly.

Let $m \in [M]$ and let $i \in [N], \mathcal{O} \in \text{CPX}(r, cn^s), \psi \in \{0, 1\}^r, e \in B_{pn}(0)$. By the definition of the set $\mathcal{I}_m$ and a simple union bound, the quantity $E_{C_{\text{in}}} [1(\mathcal{I}_m \cap \mathcal{C}(m, e, C_{\text{in}}) \neq \emptyset)]$ is bounded above by

$$\sum_{m' \in [M] \setminus \{m\}} \Pr_{C_{\text{in}}} (\mathcal{C}(m') \subseteq \mathcal{C}(m, e, C_{\text{in}}))$$

which in turn, by letting $\mathcal{E}$ be the event that $C_{\text{in}} = \{n^2 + 1, p\}$ list decodable and by the law of total probability, is bounded above by

$$\sum_{m' \in [M] \setminus \{m\}} [\Pr_{C_{\text{in}}} (\mathcal{C}(m') \subseteq \mathcal{C}(m, e, C_{\text{in}})) + \Pr_{C_{\text{in}}} (\mathcal{E}^c)]$$

Note that for $m' \in [M] \setminus \{m\}$, $\Pr_{C_{\text{in}}} (\mathcal{C}(m') \subseteq \mathcal{C}(m, e, C_{\text{in}}))$ is bounded above by $\frac{1}{M-1}$ following that the codeword $C_{\text{in}} \cup \mathcal{C}(m_0)$ can be one of at most $n^2$ codewords of $C_{\text{in}}$ randomly chosen from $\mathcal{N} - 1$ codewords (nb. we can exclude codeword $C_{\text{in}} \cup \mathcal{C}(m_0)$ contained in $\mathcal{L}(m, e, C_{\text{in}})$ by the list decodability properties of $C_{\text{in}}$. Also,
by Lemma [1] for large enough $n$, $P_{C_m}(S^n)$ is bounded above by $2^{-(n^2+1)/4}$. It follows that quantity (7) is bounded above by $\frac{\delta^2}{N-1}n^2 + (M-1)2^{-(n^2+1)/4}$ which in turn is going to zero as $n \to \infty$ independent of $m, i, \bar{O}, \psi, e$.

C. Approximation of $q_i$

To approximate $q_i(\bar{O}, \psi, e, C_m)$, we begin by defining a typical set of $[n,N]$ codebooks for which $q_i(\bar{O}, \psi, e, C_m)$ has a small variation for all typical $C_m$. We then define an approximation function $q'_i(\bar{O}, \psi, e, C_m)$ that is equal to $q_i(\bar{O}, \psi, e, C_m)$ if $C_m$ is typical and $q'_i(\bar{O}, \psi, e, C_m)$ has a well behaved variation if $C_m$ is not typical.

We first define typical codebooks. For the parameter $\delta > 0$, we define typical codebooks for all $\bar{O} \in \mathbb{CPX}(r, cn^n)$ and $\psi \in \{0,1\}^n$ such that the observation set $O_\psi$ is larger than $2^{(1-R)n}n^{2n}$. Hence, there is no need to define typical codebooks for these observation sets.

**Definition 3 (Typical Codebooks).** Suppose that $\bar{O} \in \mathbb{CPX}(r, cn^n)$ and $\psi \in \{0,1\}^n$ such that $|O_\psi| \geq 2^{(1-R)n}n^{2n}$. Set $\delta' \geq \delta$ to be the unique number such that $|O_\psi| = 2^{(1-R)n}n^{2n}$. Define

$$\ell(\bar{O}, \psi, \epsilon_R, \epsilon_p) = 2^{\frac{\delta'}{2}}n$$

$$t_L(\bar{O}, \psi, \epsilon_R, \epsilon_p) = 2^{-(1-R)n}|O_\psi| - 2^{\frac{\delta'}{2}n} + 2^{\frac{\delta'}{2}n}$$

and

$$t_U(\bar{O}, \psi, \epsilon_R, \epsilon_p) = 2^{-(1-R)n}|O_\psi| + 2^{\frac{\delta'}{2}n} - 2^{\frac{\delta'}{2}n}$$

An $[n,N]$ codebook $C_m$ is said to be typical w.r.t. the parameters $\bar{O}, \psi, \epsilon_R, \epsilon_p$ if $C_m$ is $[r,p]$ list decodable and $t_L \leq |O_\psi \cap C_m| \leq t_U$ where $C_m = C_m \circ C_{out}$. The typical set $T_{O_\psi}$ is the set of all $[n,N]$ codebooks that are typical w.r.t. $\bar{O}, \psi, \epsilon_R, \epsilon_p$.

**Definition 4 (Approximation function).** Suppose that $\bar{O} \in \mathbb{CPX}(r, cn^n)$ and $\psi \in \{0,1\}^n$ such that $|O_\psi| \geq 2^{(1-R)n}n^{2n}$. For $i \in [N]$, $e \in B_{pn}(0)$ and $[n,N]$ codebook $C_m$, define the approximation

$$q'_i(\bar{O}, \psi, e, C_m) = \frac{\Phi(\bar{O}, \psi, e, C_m)}{t(\bar{O}, \psi, C_m)}$$

(8)

where $t(\bar{O}, \psi, C_m) = \max\{|O_\psi \cap C_m|, t_L\}$. Notice that $q'_i(\bar{O}, \psi, e, C_m) \leq q_i(\bar{O}, \psi, e, C_m)$ with equality if $C_m \in T_{O_\psi}$. Furthermore, define the variation of $q'_i(\bar{O}, \psi, e, C_m)$ as

$$V'_i(\bar{O}, \psi, e, C_m) = \sum_{j=1}^{N} E_z \left[ \Delta'(j, z, C_m)^2 \right]$$

(9)

where

$$\Delta'(j, z, C_m) = |q'_i(\bar{O}, \psi, e, C_m) - q'_i(\bar{O}, \psi, e, C_m(j,z))|$$

and the expectation is taken over the random variable $z$ uniformly distributed in $\{0,1\}^n$.

The above definition of $q'_i$ is carefully set such that $V'_i(\bar{O}, \psi, e, C_m)$ is well behaved. This behavior is established in Section IV-E.

D. Combinatorial Results

In the sequel, unless otherwise stated, we fix integer $L > 1/\epsilon_p$, fix $(i, \bar{O}, \psi, e) \in \mathcal{P}(L)$, and allow only the $[n,N]$ codebook $C_m$ to vary. We drop the fixed variables from our notation. We write $q(C_m)$ to denote $q_i(\bar{O}, \psi, e, C_m)$. Similarly, we write $T$ to denote $T_{O_\psi}(\Phi(\bar{O}, \psi, e, C_m))$ to denote $\Phi(\bar{O}, \psi, e, C_m)$, $\phi_m(C_m)$ to denote $\phi_m(\bar{O}, \psi, e, C_m)$, $t(C_m)$ to denote $t(\bar{O}, \psi, C_m)$, $V(C_m)$ to denote $V_i(\bar{O}, \psi, e, C_m)$ and $V'(C_m)$ to denote $V'_i(\bar{O}, \psi, e, C_m)$.

For $k = 1, \ldots, N$, let the notation $j_k$ denote the the index

$$\text{int}(w_k(m, e, C_m)) \in [N]$$

For $m \in [M]$, let $y_m = C_m + e$.

**Claim 1.** Let $C_m$ be an $[n,N]$ codebook, $m \in [M]$, $z \notin B_{pn}(y_m)$ and $k \in [N]$ such that $w_k(m, e, C_m) \neq C_{out}(m)$. Let $C'_m$ denote the codebook $C_m(j_k, z)$. If $C_m(j_k) \notin B_{pn}(y_m)$ and

$$d(y_m, C_m(w_k(m, e, C_m))) > d(y_m, z),$$

then

$$d(y_m, C_m(w_k(m, e, C'_m))) > d(y_m, z).$$

(10)

**Proof.** Observe that $|C_m \cap B_{pn}(y_m)| \leq i-1$. Together with the fact $z \notin B_{pn}(y_m)$ and $C_m(j_k) \notin B_{pn}(y_m)$, it follows that $|C'_m \cap B_{pn}(y_m)| = |C_m \cap B_{pn}(y_m)| \leq i-1$. This implies equation (10).

**Claim 2.** Let $C_m$ be an $[n,N]$ codebook, $m \in [M]$, $z \in \{0,1\}^n$ and $k \in \{i+1, \ldots, N\}$ such that $w_k(m, e, C_m) \neq C_{out}(m)$. If

$$d(y_m, C_m(w_k(m, e, C_m))) < d(y_m, z),$$

then $w_k(m, e, C_m) = w_k(m, e, C_m(j_k, z))$.

**Proof.** Observe that by replacing the codeword $C_m(j_k)$ in the codebook $C_m$ with the word $z$, we do not change the position of the $i$ closest codewords in $C_m$ to word $y_m$. Hence, $w_i(m, e, C_m) = w_i(m, e, C_m(j_k, z))$.

**Claim 3.** Let $C_m$ be an $[n,N]$ codebook, $m \in [M]$, $z \notin B_{pn}(y_m)$ and $k \in [N]$ such that $w_k(m, e, C_m) \neq C_{out}(m)$. If either $k > i$ or $C_m(j_k) \notin B_{pn}(y_m)$, then

$$1(w_k(m, e, C_m) \in \mathcal{I}_m \cap \mathcal{L}_m(m, e, C_m))$$

$$= 1(w_k(m, e, C_m(j_k, z)) \in \mathcal{I}_m \cap \mathcal{L}_m(m, e, C_m(j_k, z))).$$

(11)

**Proof.** Consider 2 Cases. (Case 1): Suppose that $C_m(w_k(m, e, C_m)) \notin B_{pn}(y_m)$ (i.e., $w_k(m, e, C_m) \notin \mathcal{L}_m(m, e, C_m)$). From the condition that $k \leq i$ (or by assumption), it follows that $C_m(j_k) \notin B_{pn}(y_m)$, and in turn by Claim 1 we have for $C'_m = C_m(j_k, z)$ that $C'_m(w_k(m, e, C'_m)) \notin B_{pn}(y_m)$, and in turn, $w_i(m, e, C_m(j_k, z)) \notin \mathcal{L}_m(m, e, C_m(j_k, z))$. It follows that both sides of equation (11) are 0, and thus, Claim 3 holds in this Case. (Case 2): Suppose that $C_m(w_k(m, e, C_m)) \in B_{pn}(y_m)$ (i.e., $w_k(m, e, C_m) \in \mathcal{L}_m(m, e, C_m)$). Then $k > i$ and $d(y_m, C_m(w_k(m, e, C_m))) < d(y_m, z)$, and in turn, following Claim 2 we have that $w_i(m, e, C_m) = w_i(m, e, C_m(j_k, z))$.}
Furthermore, since the word $C_{in}(j_k, z) \circ C_{out}(m) + e$ is equal to $y_m$, word $w_k(m, e, C_{in}(j_k, z))$ is in $L_{in}(m, e, C_{in}(j_k, z))$. Thus, equation (11) holds, and in turn, Claim 3 holds in this Case.

Claim 4. Let $C_{in}$ be an $[n, N]$ codebook, $m \in [M]$, $z \notin B_{pn}(y_m)$ and $k \in [N]$ such that $w_k(m, e, C_{in}) \neq C_{out}(m)$. If either $k > i$ or $C_{in}(j_k) \notin B_{pn}(y_m)$, then $\phi_m(C_{in}) = \phi_m(C_{in}(j_k, z))$.

Proof. Claim 4 follows from Claim 3 and the observation that since $k \in [N]$ such that $w_k(m, e, C_{in}) \neq C_{out}(m)$, we have that $1(C_{in} \circ C_{out}(m) \in O_\psi) = 1(C_{in}(j_k, z) \circ C_{out}(m) \in O_\psi)$.

E. Smoothness of $q_i^c$

The goal of this subsection is to establish two bounds on $V'$. We say that a number $a_G > 0$ is a typical variation coefficient of $q'$ if $V'(C_{in}) \leq a_G$ for all $C_{in} \in T$. We say that a number $a_{\ell} > 0$ is a global variation coefficient of $q'$ if $V'(C_{in}) \leq a_{\ell}$ for all $[n, N]$ codebooks $C_{in}$. This subsection characterizes the smoothness of $q'$ by finding small typical and global variation coefficients that will later prove useful in establishing the concentration of $q'$.

The following Lemma shows that $q_i^c(C_{in})$ is smooth for all $C_{in} \in T$ in a Lipschitz sense. This fact will help us find a small typical variation coefficient of $q'$.

**Lemma 5.** Define

$$K = K_i(\bar{\psi}, \psi, e) = \frac{2\ell + 3}{t_L - 1}$$

If $C_{in} \in T$, then $q_i^c(C_{in})$ is $K$-Lipschitz, i.e., $\Delta'(j, z, C_{in}) \leq K$ for all $j \in [N]$, $z \in \{0, 1\}^n$.

Proof of Lemma 5. Let $C_{in} \in T$, $j \in [N]$ and $z \in \{0, 1\}^n$. We first count the number of messages $m \in [M]$ such that $\phi_m(C_{in}) \neq \phi_m(C_{in}(j, z))$. Since $C_{in}$ is $[\ell, p]$ list decodable, the $[n, N]$ codebook $C_{in}'$ resulting from a translation of $C_{in}$ by the vector $e$ (i.e., $C_{in}' = \{C_{in}(1 + e, C_{in}(2 + e, \ldots, C_{in}(N) + e)\}$ is also $[\ell, p]$ list decodable. Hence, there exists at most $2\ell$ messages $m_1, \ldots, m_{2\ell}$ such that $d(C_{in}(C_{out}(m_k)) + e, C_{in}(j)) \leq pn$ or $d(C_{in}(C_{out}(m_k)) + e, z) \leq pn$ for $k = 1, \ldots, 2\ell$. With the above observation, we can show the following claim.

Claim 5. For any message $m \in [M]$ that is not in the set $\{m_1, \ldots, m_{2\ell}\}$,

$$1\{w_k(m, e, C_{in}) \in L_{in}(m, e, C_{in})\} + 1\{w_k(m, e, C_{in}(j, z)) \in L_{in}(m, e, C_{in}(j, z))\}$$

We observe that Claim 5 is a special case of Claim 3. To see this, let $m \in [M] \backslash \{m_1, \ldots, m_{2\ell}\}$ and $y_m = C_{in} \circ C_{out}(m) + e$, and first observe that the ball $B_{pn}(y_m)$ contains neither $C_{in}(j)$ nor $z$. Let $k \in [N]$ such that $j = \text{int}(w_k(m, e, C_{in}))$. Since $C_{in}(j) \notin B_{pn}(y_m)$, it follows that $w_k(m, e, C_{in}) \neq C_{out}(m)$. These conditions are sufficient to satisfy the hypothesis of Claim 3.

Following Claim 5, the number of messages $m \in [M]$ such that $\phi_m(C_{in}) \neq \phi_m(C_{in}(j, z))$ is bounded above by $2\ell + 2$ (where the 2 is added to account for the 2 possible messages $m_1', m_2'$ such that for $k = 1, 2$, $I(C_{in} \circ C_{out}(m_i') \in O_\psi) \neq I(C_{in}(j, z) \circ C_{out}(m_i') \in O_\psi)$ which in turn may result in $\phi_m(C_{in}) \neq \phi_m(C_{in}(j, z))$). From the triangle inequality, it follows that $|\Phi(C_{in}) - \Phi(C_{in}(j, z))| \leq 2\ell + 2$.

We are now ready to prove Lemma 5 by upper bounding $\Phi(C_{in}) - \Phi(C_{in}(j, z))$. To do this, we walk through the upper bound of the quantity $\Phi(C_{in}) - \Phi(C_{in}(j, z))$. The upper bound of the negative of the above quantity follows the same approach. We have that

$$\Phi(C_{in}) - \Phi(C_{in}(j, z))$$

where inequality (a) follows from $t(C_{in}(j, z)) \leq t(C_{in}) + 1$, inequality (b) follows from $|\Phi(C_{in}) - \Phi(C_{in}(j, z))| \leq 2\ell + 2$ and inequality (c) follows from $|\Phi(C_{in}) - |O_\psi \cap C_{in}|| \leq t(C_{in})$. To complete the proof, we apply the bound $|O_\psi \cap C_{in}| \geq t_L$ to inequality (c) which follows from the fact that $C_{in} \in T$.

**Lemma 6.** (Global variation Coefficient). Suppose that $O_\psi$ is bounded in size such that for some $\delta' \geq \delta$ we have $|O_\psi| = 2(1 - \delta)n^22^n$. For any $[n, N]$ codebook $C_{in}$ and for large enough $n$ (that depends only on $\delta$ and $\epsilon_\psi$), $V'(C_{in}) \leq a_G = 5\ell + 14$.

Before proving Lemma 6, we prove the following useful inequality.

**Lemma 7.** For an $[n, N]$ codebook $C_{in}$, for $m \in [M]$ and for $z \notin B_{pn}(C_{in}(m) + e)$,

$$\sum_{j=1}^{N} |\phi_m(C_{in}) - \phi_m(C_{in}(j, z))| \leq i + 1.$$

Proof of Lemma 7. For $k = 1, \ldots, L$, let the notation $j_k$ denote the the index $\text{int}(w_k(m, e, C_{in})) \in [N]$. Following Claim 4, the quantity $\sum_{j=1}^{N} |\phi_m(C_{in}) - \phi_m(C_{in}(j, z))|$ is equal to

$$\sum_{k=1, \ldots, \ell} |\phi_m(C_{in}) - \phi_m(C_{in}(j_k, z))|,$$

which in turn is bounded above by $i + 1$.

We are now ready to prove Lemma 6.

**Proof of Lemma 6.** Recall that $V'(C_{in})$ is equal to

$$\sum_{j=1}^{N} \sum_{z \in \{0, 1\}^n} \left| \frac{\Phi(C_{in}) - \Phi(C_{in}(j, z))}{t(C_{in})} \right| 2^{-n}. \quad (14)$$

In expression (14), we can cut the two summations by partitioning the set \(\{ j \in [N] \}\) into \(\{ j \in [N] : C_{in}(j) \in O_{\psi} \}\) and \(\{ j \in [N] : C_{in}(j) \notin O_{\psi} \}\), and by partitioning the set \(\{ z \in \{0,1\}^n \}\) into \(\{ z \in \{0,1\}^n : z \in O_{\psi} \}\) and \(\{ z \in \{0,1\}^n : z \notin O_{\psi} \}\). Hence, we write \(V'(C_{in})\) as the sum of 4 terms:

\[
\begin{align*}
\sum_{j:C_{in}(j) \in O_{\psi}, z \in O_{\psi}} \frac{|\Phi(C_{in}) - \Phi(C_{in}(j,z))|}{t(C_{in})} 2^{-n} \\
+ \sum_{j:C_{in}(j) \in O_{\psi}, z \notin O_{\psi}} \frac{|\Phi(C_{in}) - \Phi(C_{in}(j,z))|}{t(C_{in})} 2^{-n} \\
+ \sum_{j:C_{in}(j) \notin O_{\psi}, z \in O_{\psi}} \frac{|\Phi(C_{in}) - \Phi(C_{in}(j,z))|}{t(C_{in})} 2^{-n} \\
+ \sum_{j:C_{in}(j) \notin O_{\psi}, z \notin O_{\psi}} \frac{|\Phi(C_{in}) - \Phi(C_{in}(j,z))|}{t(C_{in})} 2^{-n}.
\end{align*}
\]  

(15)

We separately bound term (15) through term (18).

**First Term:** We first bound term (15). Writing \(\Phi\) as a sum of \(\phi_m\) terms, term (15) is bounded above by

\[
\sum_{j:C_{in}(j) \in O_{\psi}, z \in O_{\psi}} \sum_{m \in \Omega_{\psi} \cap \psi} \frac{|\phi_m(C_{in}) - \phi_m(C_{in}(j,z))|}{t(C_{in})} 2^{-n}.
\]

which in turn can be bounded above by partitioning the set \(\{ z \in O_{\psi} \}\) into \(\{ z \in O_{\psi} \cap \Phi_{pn}(y) \}\) and \(\{ z \in O_{\psi} \cap \Phi_{pn}(y) \}\), and applying the following inequalities: \(|\Phi_{pn}(y)| \leq 2t_{(p,n)}\) and \(|O_{\psi} \cap C_n| \leq t(C_{in})\); the bound is as follows:

\[
\sum_{m:C_{in}(m) \in O_{\psi}, z \in O_{\psi} \cap \Phi_{pn}(y)} \sum_{j:C_{in}(j) \in O_{\psi}} \frac{|\phi_m(C_{in}) - \phi_m(C_{in}(j,z))|}{t(C_{in})} 2^{-m}
\]

(16)

which in turn is bounded above by \(i + 1 + 2^{-c_o n}\) following Lemma 7.

**Second term:** Next, we bound term (16). Let the notation \(j \in C_{out}\) denote an index \(j \in [N]\) that belongs to the set \(\{\text{int}(C_{out}(1)), \text{int}(C_{out}(2)), \ldots, \text{int}(C_{out}(M))\}\). Similarly, let the notation \(j \notin C_{out}\) denote the index \(j \in [N]\) that does not belong to the set \(\{\text{int}(C_{out}(1)), \text{int}(C_{out}(2)), \ldots, \text{int}(C_{out}(M))\}\). In term (16), we cut the summation over \(j\) by partitioning the set \(\{ C_{in}(j) \in O_{\psi} \}\) into the sets \(\{ C_{in}(j) \in O_{\psi}, j \notin C_{out} \}\) and \(\{ C_{in}(j) \in O_{\psi}, j \notin C_{out} \}\). Since \(t(C_{in}(j,z))\) is equal to \(t(C_{in})\) when \(j \notin C_{out}\), term (16) is equal to

\[
\sum_{j: C_{in}(j) \in O_{\psi}} \sum_{z \notin O_{\psi}} \frac{|\Phi(C_{in}) - \Phi(C_{in}(j,z))|}{t(C_{in})} 2^{-n} \\
+ \sum_{j: C_{in}(j) \in O_{\psi}} \sum_{z \notin O_{\psi}} \frac{|\Phi(C_{in}) - \Phi(C_{in}(j,z))|}{t(C_{in})} 2^{-n}.
\]

(17)

To bound quantity (19), the following inequality will prove useful:

\[
\frac{|\Phi(C_{in}) - \Phi(C_{in}(j,z))|}{t(C_{in})} \leq \max \left\{ \frac{|\Phi(C_{in}) - \Phi(C_{in}(j,z))|}{t(C_{in})}, \frac{|\Phi(C_{in}) - \Phi(C_{in}(j,z))|}{t(C_{in})} \right\}
\]

(18)

Following the above inequality, the fact that \(t(C_{in}(j,z))\) is bounded below by \(t(C_{in}) - 1\), quantity (19) is bounded above by

\[
\sum_{j:C_{in}(j) \in O_{\psi}, z \notin O_{\psi}} \frac{|\Phi(C_{in}) - \Phi(C_{in}(j,z))|}{t(C_{in})} 2^{-n} \\
+ \sum_{j:C_{in}(j) \notin O_{\psi}} \sum_{z \in O_{\psi}} \max\{\Phi(C_{in}), \Phi(C_{in}(j,z))\} 2^{-n} \\
\]

(19)

which in turn is bounded above by

\[
\sum_{j:C_{in}(j) \in O_{\psi}} \sum_{z \in O_{\psi} \cap \Phi_{pn}(y)} \frac{|\Phi(C_{in}) - \Phi(C_{in}(j,z))|}{t(C_{in})} 2^{-n} \\
+ \sum_{j:C_{in}(j) \notin O_{\psi}} \sum_{z \in O_{\psi}} \max\{\Phi(C_{in}), \Phi(C_{in}(j,z))\} 2^{-n} \\
\]

(20)

Following the fact that both \(\Phi(C_{in})\) and \(\Phi(C_{in}(j,z))\) are bounded above by \(t(C_{in})\) when \(j \in C_{out}\), and following the inequality \(|\Phi_{pn}(y)| \leq 2t_{(p,n)}\) and \(|O_{\psi} \cap C_n| \leq t(C_{in})\), the bound is as follows:

\[
\sum_{m:C_{in}(m) \in O_{\psi}, z \in O_{\psi} \cap \Phi_{pn}(y)} \sum_{j:C_{in}(j) \in O_{\psi}} \frac{|\phi_m(C_{in}) - \phi_m(C_{in}(j,z))|}{t(C_{in})} 2^{-m}
\]

(21)

which in turn is bounded above by \(i + 1 + 2^{-c_o n}\) following Lemma 7.

**Third term:** Next, we bound term (17). Using an approach that is similar in the bounding of term (16), term (17) is bounded above by

\[
\sum_{j:C_{in}(j) \notin O_{\psi}} \sum_{z \in O_{\psi}} \max\{\Phi(C_{in}), \Phi(C_{in}(j,z))\} 2^{-n} \\
+ \sum_{j:C_{in}(j) \notin O_{\psi}} \sum_{z \in O_{\psi}} \frac{|\Phi(C_{in}) - \Phi(C_{in}(j,z))|}{t(C_{in})} 2^{-n}.
\]

(22)
which in turn is bounded above by
\[
\sum_{j:C_{in}(j)\notin \mathcal{O}_T} \sum_{m:C_{out}(m)\in \mathcal{O}_T} \frac{|\phi_m(C_{in}) - \phi_m(C_{in}(j, z))|}{t(C_{in})} 2^{-n} + \sum_{j:C_{in}(j)\notin \mathcal{O}_T} \sum_{(j)\in \mathcal{O}_T} \frac{\Phi_{\text{max}}}{t^2(C_{in})} 2^{-n} \tag{22}
\]
implies that \( j \neq \text{int}(C_{out}(m)) \). These conditions are sufficient to satisfy the hypothesis of Claim 4.

We now prove Lemma 8. For \( k = 1, 2 \), define \( S_k = \cap_{m:C_{out}(m)\in \mathcal{O}_T} S_k(m) \). Let \( j \in [N] \) such that \( C_{in}(j) \in S_1 \) and let \( z \in S_2 \). First note that \( C_{in}(j) \) is not in \( \mathcal{O}_T \) and \( z \) is not in \( \mathcal{O}_T \). Hence, \( \mathcal{O}_T \cap C_{in} = \mathcal{O}_T \cap \mathcal{C}_{in}(j, z) \cap \mathcal{O}_T \) and thus \( t(C_{in}) = t(C_{in}(j, z)) \). In turn, the difference \( q'(C_{in}) - q'(C_{in}(j, z)) \) is equal to
\[
\sum_{m:C_{out}(m)\in \mathcal{O}_T} \phi_m(C_{in}) - \sum_{m:C_{out}(m)\in \mathcal{O}_T} \phi_m(C_{in}(j, z)) \tag{23}
\]
in which in turn is equal to 0 following Claim 6. Thus, \( \Delta'(j, z, C_{in}) = 0 \). It follows that \( V'(C_{in}) \) is equal to
\[
\sum_{j\in [N]: C_{in}(j)\in \mathcal{O}_T} |\Delta'(j, z, C_{in})|^2 + \sum_{j\in [N]: C_{in}(j)\in \mathcal{O}_T} \frac{\Delta'(j, z, C_{in})^2}{2^n} \tag{26}
\]
We now upper bound \( V'(C_{in}) \) by bounding equation (26). Since \( C_{in} \in \mathcal{T} \), codeword \( C_{in} \) is \([T, p]\) list decodable and we have that \( |S_2'| \) is bounded above by \( |\mathcal{O}_T \cap C_{in}| \). By a simple union bound, \( |S_2'| \) is bounded above by \( |\mathcal{O}_T \cap C_{in}| 2^{H(p)\cdot n + |\mathcal{O}_T|} \) and \( |S_1| \leq N \). By Lemma 5, \( \Delta'(j, z, C_{in}) \leq K \) for all \( j \in [N], z \in [0, 1]^n \) and thus equation (26) is upper bounded by \(|S_1'| + |S_1||S_2'| 2^{-n})K^2 \), which in turn, for any \( \alpha \in [0, 1] \), is upper bounded by
\[
|\mathcal{O}_T \cap C_{in}|(K^2+2^{(\rho - \alpha)n}) \left((|\mathcal{O}_T \cap C_{in}|2^{(H(p)-1+\alpha)n} + |\mathcal{O}_T|2^{-(1-\alpha)n}\right)K^2. \]

The value of \( a_T^{LB} \) follows by setting \( \alpha = \rho \) and applying the bound \( |\mathcal{O}_T \cap C_{in}| \leq t_U \).

F. Concentration of q

The following Lemma shows that if \( q'(C_{in}) = q(C_{in}) \) w.h.p. over \( \Omega(n, N) \), then \( q' \) concentrated implies that \( q \) is concentrated.

Lemma 9. For \( \lambda > 0 \),
\[
\mathbb{P}_{C_{in}}(q - \mathbb{E}_{C_{in}}[q] > \lambda) \leq \mathbb{P}_{C_{in}}(q' - \mathbb{E}_{C_{in}}[q'] > \lambda) + \mathbb{P}_{C_{in}}(q \neq q'). \tag{27}
\]

Proof of Lemma 9. Let \( \Omega \) denote the set of all \([n, N]\) codebooks. We have that \( \mathbb{P}_{C_{in}}(q - \mathbb{E}_{C_{in}}[q] > \lambda) \) is equal to
\[
\int_{C_{in} \in \Omega} 1(q - \mathbb{E}_{C_{in}}[q] > \lambda) d\mathbb{P}_{C_{in}} = \int_{q(C_{in}) = q'(C_{in})} 1(q - \mathbb{E}_{C_{in}}[q] > \lambda) d\mathbb{P}_{C_{in}} + \int_{q(C_{in}) \neq q'(C_{in})} 1(q - \mathbb{E}_{C_{in}}[q] > \lambda) d\mathbb{P}_{C_{in}} \leq \int_{q(C_{in}) = q'(C_{in})} 1(q' - \mathbb{E}_{C_{in}}[q'] > \lambda) d\mathbb{P}_{C_{in}} + \int_{q(C_{in}) \neq q'(C_{in})} d\mathbb{P}_{C_{in}} \leq \mathbb{P}_{C_{in}}(q' - \mathbb{E}_{C_{in}}[q'] > \lambda) + \mathbb{P}_{C_{in}}(q \neq q').\]
where inequality (a) follows from $E_{C_{in}}[q'] \leq E_{C_{in}}[q]$. \hfill \blacksquare

We now state and prove our concentration inequality for $q'$, which follows from a modified logarithmic Sobolev inequality [18, Theorem 2].

**Lemma 10.** Let $L \geq 1$ be an integer constant. Suppose that $|O_\psi| \geq 2(1-K)n^{2}\ln n$, $i \in [L]$ and suppose that the following Assumptions hold:

1) For large enough $n$ (depending only on $\delta$, $\epsilon_R$, there exists a global variation coefficient $\alpha_G \in (0, \infty)$ such that $V'(C_{in}) \leq \alpha_G$ for all $n, N$ codebooks $C_{in}$.
2) There exists a typical variation coefficient $\alpha_T \in (0, 1]$ such that $V'(C_{in}) \leq \alpha_T$ for all $C_{in} \in T$.
3) As a sequence in $n$, the ratio $\frac{\alpha_T}{\alpha_G}$ is $o(-\ln P_{in}(C_{in} \notin T))$.

Then for $\lambda \in (\sqrt{\alpha_T}, 1)$ and for large enough $n$ (depending only on $\lambda$, $\delta$, $\epsilon_R$ and $\lambda$),

$$P_{C_{in}}\left(q'_i(\bar{O}, \psi, e) - E_{C_{in}}[q'_i(\bar{O}, \psi, e)] > \lambda\right) \leq \exp\left\{-\frac{\lambda^2}{8\alpha_T}\right\}. \quad (28)$$

**Proof of Lemma 10** We begin by restating a modified logarithmic Sobolev inequality.

**Lemma 11 (18 Theorem 2).** For $\theta > 0$ and $\mu \in (0, 1/\theta)$,

$$\ln E_{C_{in}}[\exp\{\mu(q' - E_{C_{in}}[q'])\}] \leq \frac{\mu}{1 - \mu} \ln E_{C_{in}}[\exp\{\frac{V'}{\theta}\}] \quad (29)$$

The quantity $E_{C_{in}}[\exp\{\mu(q' - E_{C_{in}}[q'])\}]$ on the RHS of (29) is bounded above by $\exp\{\mu\theta/\theta\} + \exp\{\mu\theta/\theta\}P_{C_{in}}(C_{in} \notin T)$; this follows from the law of total probability (applied with the sample space partition $\{C_{in} \in T\}$, $\{C_{in} \notin T\}$) and Assumption 1 and Assumption 2. Set $\theta$ such that $\exp\{\mu\theta/\theta\}P_{C_{in}}(C_{in} \notin T) = \exp\{\mu\theta/\theta\}$. Then

$$\ln E_{C_{in}}[\exp\{\frac{V'}{\theta}\}] \leq \mu\theta/\theta + \ln(2) \quad \text{where it follows from} \quad (29)$$

that

$$E_{C_{in}}[\exp\{\mu(q' - E_{C_{in}}[q'])\}] \leq \exp\left\{\frac{\mu^2\theta}{1 - \mu} + \frac{\mu\theta\ln 2}{1 - \mu}\right\}. \quad (30)$$

By Markov’s inequality,

$$P_{C_{in}}(q' - E_{C_{in}}[q'] > \lambda) \leq \exp\left\{\frac{\mu^2\theta}{1 - \mu} + \frac{\mu\theta\ln 2}{1 - \mu} - \lambda\right\}. \quad (30)$$

To finish the proof, we choose some round numbers to simplify the RHS of (30). We choose $\mu$ to minimize the RHS of (30). To simplify this optimization, we can treat $\theta \approx 0$ as a constant independent of $\mu$, which gives us $\lambda = \lambda/\mu(2\alpha_T)$. It follows that $\mu\theta$ is equal to $\frac{\lambda^2}{4\alpha_T} - \ln P_{in}(C_{in} \notin T)$, which in turn is less than both $1/5$ and $\mu^2\alpha_T/5$ for large enough $n$ by Assumption 3. Furthermore, $\mu^2\alpha_T + \mu\theta\ln 2$ is less than $(6/5)\mu^2\alpha_T$. Substituting these bounds into (30) gives us the desired inequality.

Putting together Lemma 6 through Lemma 10, we get the following concentration inequality for $q$.

**Lemma 12.** Let $L \geq 1$ be an integer constant. Suppose that $|O_\psi| \geq 2(1-K)n^{2}\ln n$ and $i \in [L]$ for small enough $\epsilon_R > 0$ (depending only on $\delta$) and for large enough $n$ (depending only on $\delta$, $L$, and $\epsilon_R$),

$$P_{C_{in}}\left(q_i(\bar{O}, \psi, e) - E_{C_{in}}[q_i(\bar{O}, \psi, e)] > 1/n\right) \leq \exp\left\{-\frac{2\beta_1}{8\alpha_T}\right\}. \quad (31)$$

**Proof of Lemma 12** Recall that $\delta' \geq \delta$ is the unique constant such that $|O_\psi| = 2(1-K)n^{2}\ln n$.

We first bound the probability that $C_{in}$ is not typical. Note that $E_{C_{in}}[|O_\psi \cap C_{in}| = 2(1-K)n^{2}\ln n] = 2^{\delta'\ln n}$. It follows from Lemma 1 and Lemma 3 that

$$P_{C_{in}}(q' \neq q) \leq P_{C_{in}}(C_{in} \notin T) \leq 2^{-2\alpha_T}$$

and $P_{C_{in}}(|O_\psi \cap C_{in}| < t_L) \leq 2 \exp\left\{-\frac{t_L^2}{4}\right\}$ for large enough $n$ (depending only on $\delta$).

Next, we bound $K$ and $a_T^{LB}$ (defined by equation 12 and equation 25, respectively). Recall that $K$ is equal to $2^{-n/4}$.

Substitution of the typical parameters, $K$ is equal to $2^{-n/5} + 2$ for large enough $n$ (depending only on $\delta$).

Similarly, $a_T^{LB}$ is equal to $2^{-3\delta/13 + 6} + 2^{-5\delta/13 + 6\delta n + 6\epsilon_R}$ for large enough $n$ (depending only on $\delta$ and $\epsilon_R$).

Hence, for large enough $n$ (depending only on $\delta$ and $\epsilon_R$) we can choose any $a_T$ such that

$$a_T \geq a_T^{LB} = 2^{-3\delta/13 + 6} + 2^{-5\delta/13 + 6\delta n + 6\epsilon_R}. \quad (31)$$

Finally, we are ready to apply Lemma 10. We first check that Assumption 3 of Lemma 10 holds. We have that

$$\frac{a_T - a_T}{a_T^2(-\ln P_{C_{in}}(C_{in} \notin T))} < \frac{a_G}{a_T^2\frac{1}{2\alpha_T}} \quad (32)$$

for large enough $n$ (depending only on $\delta$ and $\epsilon_R$). Pick $\epsilon_R \in (0, (\lambda - 1/\delta))$ and set $a_T = 2^{-3\delta/13}$; this choice of $a_T$ satisfies equation (31). Following Lemma 6, $a_G$ is bounded above by $5L + 14$ for large enough $n$ (depending only on $\delta$, $\epsilon_R$, and $L$). In turn, using $\delta' \geq \delta$, the RHS of quantity (32) is bounded above by

$$\frac{5L + 14}{2^{-4\alpha_T}}$$

for large enough $n$ (depending on $\delta$, $\epsilon_R$, and $L$), and therefore, is o(1) and Assumption 3 holds.

To complete the proof, we apply Lemma 2 to bound the quantity $P_{C_{in}}(q - E_{C_{in}}[q] > 1/n)$ above by $P_{C_{in}}(q' - E_{C_{in}}[q'] > 1/n) + P_{C_{in}}(q' \neq q)$. Lemma 12 follows from applying Lemma 10 to bound $P_{C_{in}}(q' - E_{C_{in}}[q'] > 1/n)$ above by $\exp\{-2\delta n/30\}$ for large enough $n$ (depending only on $\delta$, $\epsilon_R$, and $L$), and by observing that $P_{C_{in}}(q' \neq q)$ is bounded above by $P_{C_{in}}(|O_\psi \cap C_{in}| < t_L) \leq \exp\{-2\delta n/30\}$. \hfill \blacksquare
Our goal is to show that for the chosen $\epsilon_R, \epsilon_p > 0$, there exists an $[n, N]$ codebook $C_{in}$ such that $P_e(C_{in})$ is going to 0 in the limit of the blocklength $n \to \infty$. By Condition 2, $\epsilon_R$ is small enough to satisfy the conditions of Lemma 12.

Next, we define the set of good $[n, N]$ codebooks and bound the probability that $C_{in}$ is not good. Let $L > 1/\epsilon_p$ be an integer constant and let $E_1$ be the event that $C_{in}$ is not in $[L, p]$ list decodable. Let $E_2$ be the event that some $q_i(\vec{O}, \psi, e, C_{in}) > C_{in}$ is not concentrated for some parameters, i.e., $q_i(\vec{O}, \psi, e, C_{in}) > C_{in}$ is not for some $e \in B_{pn}(0), i \in [L], \vec{O} \in \text{CPX}(r, cn^s)$ and $\psi \in \{0, 1\}^n$ such that $|\vec{O}| \geq 2(1-R)n^2n$. Let $E_3$ be the event that some small observation set is not typical, i.e., $|\vec{O}| \leq 2^{(\delta + \Delta \delta)n}$ for some $\vec{O} \in \text{CPX}(r, cn^s), \psi \in \{0, 1\}^n$ such that $|\vec{O}| < 2(1-R)n^{2n}$. Finally, let $G = (E_1 \cup E_2 \cup E_3)\in$ denote the set of $[n, N]$ codebooks. We say that an $[n, N]$ codebook $C_{in}$ is not good if $C_{in}$ is not in $G$.

We now bound the probability that $C_{in}$ is not good by bounding $P_{\text{in}}(E_1), P_{\text{in}}(E_2)$, and $P_{\text{in}}(E_3)$. Following Lemma 2, $P_{\text{in}}(E_1)$ is bounded above by $1/n$ for large enough $n$ (depending only on $\epsilon_p$). The adversary’s computational bound will help us to bound $P_{\text{in}}(E_2)$ and $P_{\text{in}}(E_3)$. Let $S$ denote the number of unique observation sets in $\text{CPX}(r, cn^s)$, i.e., $S = |\vec{O} \subseteq \{0, 1\}^n : \vec{O} \in \text{CPX}(r, cn^s), \vec{O} = \vec{O}_1$ for some $\psi \in \{0, 1\}^r|$. We can bound $S$ by counting the number of Boolean circuits with $cn^s$ logic gates.

Lemma 13. For large enough $n$ (depending only on $c$ and $s$), the number of functions in $\text{CPX}(r, cn^s)$ is bounded above by $2^{s^n+n}$, and thus $S = S(r, cn^s) \leq 2^{s^n+n}$. Proof is in Appendix 12.

Following Lemma 12 and a simple union bound, $P_{\text{in}}(E_2)$ is bounded above by $S2^n L^2 \exp\{-2^{s^n+n}/(8n^2)\}$ for large enough $n$ (depending only on $\delta, L, \epsilon_p, c$ and $s$). Following Lemma 3 and a simple union bound, $P_{\text{in}}(E_3)$ is bounded above by $S2^{s^n-n}$ for large enough $n$ (depending only on $\delta, \Delta \delta, c$ and $s$). The probability that $C_{in}$ drawn from distribution $\Omega[n, N]$ is not good is bounded above by $P_{\text{in}}(E_1) + P_{\text{in}}(E_2) + P_{\text{in}}(E_3)$, which in turn is going to zero in $n$.

We now show that for any $C_{in} \in G$, the probability of decoding error $P_e(C_{in})$ is going to 0. Let $C_{in} \in G$. The probability of decoding error upper bound (35) is equal to

$$\max_{\vec{O} \in \text{CPX}(r, cn^s)} \sum_{i=1}^{L} \text{E}[\max_{e \in B_{pn}(0)} q_i(\vec{O}, \psi, e, C_{in})].$$

(33)

In our analysis, we allow decoding to fail on all observation sets smaller than $T = 2^{(1-R)n^2n}$. Specifically, if $|\vec{O}| < T$, then we use the trivial bound $q_i(\vec{O}, \psi, e) \leq 1$. It follows that (33) is upper bounded by the maximum over $\vec{O} \in \text{CPX}(r, cn^s)$ of

$$\sum_{i=1}^{L} \text{E}[\max_{e \in B_{pn}(0)} q_i(\vec{O}, \psi, e, C_{in})] I(|\vec{O}| \geq T) + L \text{E}[\max_{e \in B_{pn}(0)} q_i(\vec{O}, \psi, e, C_{in}) I(|\vec{O}| < T)].$$

(34)

In equation (34) above, the term $L \text{E}[\max_{e \in B_{pn}(0)} q_i(\vec{O}, \psi, e, C_{in}) I(|\vec{O}| < T)]$ is equal to $L \sum_{\psi \in \{0, 1\}^n} I(|\vec{O}| < T) I(\vec{O} \in \text{CPX}(r, cn^s))\in$ going to 0 uniformly for all $\vec{O} \in \text{CPX}(r, cn^s)$ as $n \to \infty$.

To complete the proof, given that $C_{in} \in G$, we show that the first term of quantity (34), i.e.,

$$\sum_{i=1}^{L} \text{E}[\max_{e \in B_{pn}(0)} q_i(\vec{O}, \psi, e, C_{in}) I(|\vec{O}| \geq T)],$$

(36)

is $o(1)$. Following from the definition of $E_2$, (36) is bounded above by the maximum over $i \in [N], \vec{O} \in \text{CPX}(r, cn^s), \psi \in \{0, 1\}^n$, $e \in B_{pn}(0)$ of $L(|\vec{O}| \geq T) I(\vec{O} \in \text{CPX}(r, cn^s))$, which in turn is going to 0 following Lemma 4. Hence, for any codebook $C_{in}$ in $G$ (which is non-empty), quantity (34) and (36) is going to 0 uniformly for all $\vec{O} \in \text{CPX}(r, cn^s)$ as $n \to \infty$. This completes the proof.

V. CONCLUSION

In this work, we study the capacity of adversarial channels in which the adversary can observe the transmitted codeword via some computationally bounded process. We characterize the capacity for certain parameters under deterministic and average probability of error criterion.

APPENDIX A

A TALAGRAND-TYPE CONCENTRATION INEQUALITY

Let $g(\cdot)$ be a function mapping the set of $[n, N]$ codebooks to $(-\infty, \infty)$. For $b > 0$, $g$ is said to be $b$-Lipschitz if for any $[n, N]$ codebooks $C_{in}$ and $C_{in}'$ differing by at most 1 codeword, then $|g(C_{in}) - g(C_{in}')| \leq b$. An index set $J(\cdot) \subseteq [N]$ is said to be a certificate of $g$ if for any $[n, N]$ codebook $C_{in}$, $g(C_{in}) \geq |J(C_{in})|$ and $g(C_{in}) \geq g(C_{in})$ for any $C_{in}$ that agrees with $C_{in}$ on the codewords indexed in $J(C_{in})$. Lastly, for $c > 0$, $g$ is said to be $c$-certifiable if there exists a certificate $J(\cdot)$ such that $|J(C_{in})| \leq c g(C_{in})$ for all $[n, N]$ codebooks $C_{in}$.

Lemma 14 (20 Theorem 11.3). Let $M[g]$ denote a median of $g$. For any $t > 0$,

$$P(g - M[g] > t) \leq 2 \exp\{-\frac{-t^2}{4b^2 c[M[g] + t]}\}$$

and

$$P(g - M[g] < -t) \leq 2 \exp\{-\frac{-t^2}{4b^2 cM[g]}\}.$$
For $y \in \{0, 1\}^n$, define $g_y(C_{in}) = |C_{in} \cap B_{pn}(y)|$. Our goal is to show that $g_y$ is strongly concentrated around its expectation $E_{C_{in}}[g_y]$ (with respect to $C_{in}$). Note the following: $g_y$ is 1-Lipschitz and $J(C_{in}) = \{ k \in [N] : C_{in}(k) \in B_{pn}(y) \}$ is a certificate of $g_y(C_{in})$ where it follows that $g_y$ is 1-certifiable.

Since $g_y(C_{in})$ is a binomial random variable, the value floor($E_{C_{in}}[g_y]$) is a median. Set $M[g_y] = \text{floor}(E_{C_{in}}[g_y])$. Note that $R < 1 - H(p)$ implies that $E_{C_{in}}[g_y]$ (which is equal to $\sum_{i=1}^{2^{rn}} P_{C_{in}}(x_i \in B_{pn}(y)) \leq 2^{(R-1+H(p))n}$) is going to zero in $n$. It follows that for large enough $n$, $M[g_y] = 0$.

By Lemma 1 for $\ell > 0$ the probability that $g_y > \ell$ is at most $2^{-\log(\ell)}2^{\ell+1}$. In conclusion, $P_{C_{in}}(\exists y \in \{0, 1\}^n \text{ s.t. } g_y(C_{in}) > \ell) = P_{C_{in}}(\bigcup_{y \in \{0, 1\}^n}\{g_y(C_{in}) > \ell\}) < 2^n2^{-\log(\ell)\frac{2^{\ell+1}}{2^{\ell+1}}} + 1$.

Define $g(C_{in}) = \{A \cap C_{in}\}$. Note the following: $g(\cdot)$ is 1-Lipschitz and $J(C_{in}) = \{ m \in [M] : C_{in}(C_{out}(m)) \in A \}$ is a certificate of $g(C_{in})$ where it follows that $g(C_{in})$ is 1-certifiable.

Since $g(C_{in})$ is a binomial random variable, the expected value $E_{C_{in}}[g]$ is a median. Set $M[g] = E_{C_{in}}[g] = 2^{-(1-H(p))n} |A|$. The desired result follows from Lemma 1.

Let $W$ be the number of functions of the form $f_n : \{0, 1\}^n \rightarrow \{0, 1\}$ that can be computed by a Boolean circuit (of $n$ inputs and 1 output) of size $cn^8$. We first show that $W < 2^{2^{(s+1)n^8+1}}$.

Note that each gate can be 1 of 16 unique functions from $\{0, 1\}^2 \rightarrow \{0, 1\}$. Furthermore, for a given gate, the number of ways to choose 2 gate inputs from $n$ circuit inputs, $cn^8 - 1$ gate outputs, and 2 constant inputs $\{0, 1\}$ is bounded above by $(n + cn^8 + 1)^2$. It follows that $W$ is bounded above by $(16(n + cn^8 + 1)^2)^{cn^8}$ which in turn, for large enough $n$, is bounded above by $(n^{8+1}2^{cn^8})^{\log n} = 2^{2^{(s+1)n^8}1^{\log n}}$. Done.

We now prove Lemma 1. Any function in $F_{n,r}$ that is computable by a Boolean circuit (of $n$ inputs and $rn$ outputs) of size $cn^8$ can be computed by $rn$ Boolean circuits (of $n$ inputs and 1 output) each of size $cn^8$. Hence, the number of functions in $F_{n,r}$ that can be computed by a Boolean circuit (of $n$ inputs and $rn$ outputs) of size $cn^8$ is bounded above by $W^{rn}$. We finish the proof by observing that $W^{rn}$ is smaller than $2^{2^{s+2}}$ for large enough $n$.

REFERENCES

[1] C. E. Shannon, “A Mathematical Theory of Communication,” *Bell System Technical Journal*, vol. 27, no. 3, 1948.

[2] R. J. McEliece, E. R. Rodemich, H. Rumsey, and L. R. Welch, “New Upper Bounds on the Rate of a Code via the Delsarte—MacWilliams Inequalities,” *IEEE Transactions on Information Theory*, 1977.

[3] R. J. Lipton, “A new approach to information theory,” in *Annual Symposium on Theoretical Aspects of Computer Science*, 1994.

[4] V. Gurusswami and A. Smith, “Optimal rate code constructions for computationally simple channels,” *Journal of the ACM*, vol. 63, no. 4, 2016.

[5] A. D. Sarwate, “Coding against myopic adversaries,” in *2010 IEEE Information Theory Workshop*.

[6] B. K. Dey, S. Jaggi, and M. Langberg, “Sufficiently Myopic Adversaries Are Blind,” *IEEE Transactions on Information Theory*, 2019.

[7] A. J. Budkuley, B. K. Dey, S. Jaggi, M. Langberg, A. D. Sarwate, and C. Wang, “Symmetrizability for Myopic AVCs,” in *IEEE International Symposium on Information Theory*.

[8] Z. Chen, S. Jaggi, and M. Langberg, “A characterization of the capacity of online (causal) binary channels,” in *ACM symposium on Theory of Computing*, 2015, pp. 287–296.

[9] B. K. Dey, S. Jaggi, M. Langberg, and A. D. Sarwate, “A bit of delay is sufficient and stochastic encoding is necessary to overcome online adversarial erasures,” in *IEEE International Symposium on Information Theory - Proceedings*, vol. 2016-August, 2016.

[10] V. Suresh, E. Ruzomberka, and D. J. Love, “Stochastic-Adversarial Channels: Online Adversaries with Feedback Snooping,” in *IEEE International Symposium on Information Theory - Proceedings*, vol. 2021-July, 2021.

[11] I. Csiszár and P. Narayan, “Capacity and Decoding Rules for Classes of Arbitrarily Varying Channels,” *IEEE Transactions on Information Theory*, vol. 35, no. 4, 1989.

[12] I. Csiszár and P. Narayan, “The Capacity of the Arbitrarily Varying Channel Revisited: Positivity, Constraints,” *IEEE Transactions on Information Theory*, vol. 34, no. 2, 1988.

[13] M. Langberg, “Oblivious communication channels and their capacity,” *IEEE Transactions on Information Theory*, 2008.

[14] C. Wang, “On the capacity of the binary adversarial wiretap channel,” in *45th Annual Allerton Conference on Communication, Control, and Computing, Allerton 2016*, 2017.

[15] M. Sigman, *Introduction to the theory of computation*, 2nd ed. Boston: Thompson Course Technology, 2006.

[16] E. N. Gilbert, “A comparison of signaling alphabets,” *Bell System Technical Journal*, vol. 31, no. 3, p. 504–522, 1952.

[17] R. R. Varshamov, “Estimate of the number of signals in error correcting codes,” *Dokl. Akad. Nauk*, vol. 117, no. 739–741, 1957.

[18] S. Boucheron, G. Lugosi, and P. Massart, “Concentration inequalities using the entropy method,” *Annals of Probability*, vol. 31, no. 3, 2003.

[19] V. H. Vu, “Concentration of Non-Lipschitz Functions and Applications,” in *Random Structures and Algorithms*, 2002.

[20] D. P. Dubhashi and A. Panconesi, *Concentration of measure for the analysis of randomized algorithms*, 2009.