The Capacity of Anonymous Communications

Hua Sun

Abstract

We consider the communication scenario where $K$ transmitters are each connected to a common receiver with an orthogonal noiseless link. One of the transmitters has a message for the receiver, who is prohibited from learning anything in the information theoretic sense about which transmitter sends the message (transmitter anonymity is guaranteed). The capacity of anonymous communications is the maximum number of bits of desired information that can be anonymously communicated per bit of total communication. For this anonymous communication problem over a parallel channel with $K$ transmitters and 1 receiver, we show that the capacity is $1/K$, i.e., to communicate 1 bit anonymously, each transmitter must send a 1 bit signal. Further, it is required that each transmitter has at least 1 bit correlated randomness (that is independent of the messages) per message bit and the size of correlated randomness at all $K$ transmitters is at least $K - 1$ bits per message bit.

This paper will be presented in part at ISIT 2018. Hua Sun (email: hua.sun@unt.edu) is with the Department of Electrical Engineering at the University of North Texas.
1 Introduction

Traditional studies in information theoretic security and cryptography focus on efficient coding techniques for protecting the information contents. There is much recent interest in shifting the objective to hide user behaviors. For example, private information retrieval (PIR) aims to pursue communication efficient methods for hiding the identity of the desired message that the user wants to retrieve from a set of distributed replicated databases. The fundamental capacity limits of PIR and several of its variants are characterized recently in [1–3].

In this work, we consider the anonymous communication problem, where the goal is to hide the identity of the transmitters, receivers and the association between the two in a network. This problem of anonymous communications has been studied extensively in cryptography and computer science communities [4–6], where typically the objective is to provide scalable solutions over large networks while information theoretic optimality guarantees are not considered or treated in the approximate order sense.

We focus on an elemental model where $K$ transmitters want to communicate to a common receiver anonymously with interference-free noiseless parallel channels. Our goal is to identify the exact information theoretic limits on the rate and common randomness for anonymous communications. For example, consider the case where we have $K = 3$ transmitters. As each transmitter is connected to the receiver with a parallel channel, the received signal $Y$ is the collection of all transmitted signals, $X_1, X_2, X_3$ (see Figure 1).

![Network topology: transmitters are connected to a single receiver with parallel interference-free noiseless links.](image)

Figure 1: Network topology: transmitters are connected to a single receiver with parallel interference-free noiseless links.

One of the transmitters wishes to send a desired message to the receiver without being identified, i.e., the receiver decodes the message correctly, but has no knowledge about which transmitter sends the message. This anonymity constraint requires that no matter which transmitter wants to send the message, the received signal must be identically distributed and the decoding mapping cannot depend on the desired transmitter index. To accomplish the task of keeping the transmitter identity anonymous, we assume that the transmitters share some correlated random variables that are independent of the messages. In this case, we assume that Transmitter 1 holds $a$, Transmitter 2 holds $b$ and Transmitter 3 holds $a + b$, where $a, b$ are two i.i.d uniform random bits (that form the correlated random variables). Then a simple scalar linear coding scheme that guarantees transmitter anonymity is presented next. Suppose the desired transmitter index is $\theta \in \{1, 2, 3\}$.

---

1 Separate and perfect communication links are the least favorable channel conditions for anonymity because this assumption eliminates the possibility of hiding over direct interactions between the signals and noise.
The transmitted signals are

\[ X_1 = a + 1(\theta = 1)W_1 \] \hspace{1cm} (1)

\[ X_2 = b + 1(\theta = 2)W_2 \] \hspace{1cm} (2)

\[ X_3 = a + b + 1(\theta = 3)W_3 \] \hspace{1cm} (3)

where \(1(x)\) is the indicator function that takes value 1 if the event \(x\) is true and 0 otherwise. Each message is assumed to be 1 independent uniform bit as well.

![Diagram](image)

Figure 2: The anonymous coding scheme with \(K = 3\) transmitters. (a). \(\theta = 1\). (b). \(\theta = 2\). (c). \(\theta = 3\).

Note that no matter which message is sent, the receiver sees 3 uniform random bits and the decoding rule is always an addition.

Correctness is easy to see as for all cases, the randomness cancels with each other after the addition operation. Anonymity holds because regardless of the value of \(\theta\), the received signal consists of 3 uniform random bits and the decoding mapping is always an addition. As such, the receiver learns nothing about which transmitter is the source of the message. We see that in order to communicate 1 bit anonymously, each transmitter needs to send 1 bit out. It is not hard to see that this is information theoretically optimal as even if there is no anonymity constraint, each transmitter will send out the desired message bit. What is non-trivial is the requirement on the correlated randomness. In this context, we show that for all linear schemes, each transmitter must hold a correlated random variable whose size is at least the size of the message and the total amount of randomness available at all transmitters must be at least as large as the size of \(K - 1\) messages. Further, when the scheme is capacity achieving, both the individual and total randomness sizes are optimal information theoretically (i.e., for all non-linear schemes as well). A scheme of similar nature appears in a different context in \[7,8\], where coded randomness is not allowed and optimality on the communications and randomness is not considered.

Notation: For integers \(Z_1, Z_2, Z_1 \leq Z_2\), we use the compact notation \([Z_1 : Z_2] = \{Z_1, Z_1 + 1, \cdots, Z_2\}\). The notation \(X \sim Y\) is used to indicate that random variables \(X\) and \(Y\) are identically distributed.

### 2 Problem Statement

Consider a network with \(K\) transmitters and 1 receiver. Each transmitter is connected to the receiver with an orthogonal noiseless link. Each link can carry one symbol from a finite field \(\mathbb{F}_p\) per channel use for a prime \(p\).
Transmitter $k, k \in [1 : K]$ has a message $W_k$. The messages $W_1, \cdots, W_K$ are independent and are each comprised of $L$ i.i.d. uniform symbols from $\mathbb{F}_p$. In $p$-ary units,

$$H(W_1) = \cdots = H(W_K) = L,$$

(4)

$$H(W_1, \cdots, W_K) = H(W_1) + \cdots + H(W_K).$$

(5)

The transmitters wish to communicate with the receiver anonymously. The transmitters privately generate $\theta$ uniformly over $[1 : K]$ (without loss of generality) and wish to communicate $W_\theta$ to the receiver while keeping $\theta$ a secret to the receiver. Depending on $\theta$, there are $K$ strategies that the transmitters employ to privately communicate the desired message. For example, if $\theta = k$, then in order to communicate $W_k$, Transmitter $i$ sends a signal $X_i^{[k]}$ over $N$ channel uses. To fulfill the task of communicating anonymously, we assume that Transmitter $i$ holds a correlated random variable $Z_i$. The correlated random variables are generated offline, i.e., before the realizations of the messages are known, so that the correlated random variables are independent of the messages.

$$H(Z_1, \cdots, Z_K, W_1, \cdots, W_K) = H(Z_1, \cdots, Z_K) + H(W_1, \cdots, W_K)$$

(6)

The transmitted signal, $X_i^{[k]}$, is a function of the information available at the transmitter (i.e., the message and the correlated random variable),

$$H(X_i^{[k]}|W_i, Z_i) = 0$$

(7)

The received signal at the receiver is a collection of the $K$ transmitted signals.

$$Y^{[k]} = [X_1^{[k]}, \cdots, X_K^{[k]}]^T$$

(8)

From $Y^{[k]}$, the receiver decodes the desired message $W_k$ according to a decoding mapping $g$. Note that the receiver is not allowed to learn anything about the index of the desired transmitter, so the decoding rule does not depend on $k$. The decoding mapping $g$ is fixed and known at every node (including the transmitters).

$$W_k = g(Y^{[k]})$$

(9)

To ensure transmitter anonymity, the $K$ strategies must be indistinguishable (identically distributed) from the perspective of the receiver, i.e., the following anonymity constraint must be satisfied $\forall k \in [1 : K]$,

$$[\text{Anonymity}] \quad (Y^{[l]}, g) \sim (Y^{[k]}, g)$$

i.e., $(X_1^{[l]}, \cdots, X_K^{[l]}, g) \sim (X_1^{[k]}, \cdots, X_K^{[k]}, g)$

(10)

The anonymous communication rate characterizes how many symbols of desired information are communicated per symbol of total communication, and is defined as

$$R \triangleq \frac{L}{KN}$$

(11)

---

2It turns out that for our achievable scheme, the transmitters do not need to know the exact value of the desired transmitter index $\theta$. It suffices for each transmitter to know that whether he is the desired or not.

3The encoding and decoding functions are globally known (akin to codebooks). Note that the receiver might try arbitrary other operations, for which no guarantees are made (e.g., correctness). Further note that the received signals are identically distributed, so the operations by the receiver do not depend on the desired message index (revealing no information).
Note that by symmetry\footnote{\footnotetext[4]{Given any (asymmetric) achievable scheme that might employ a different number of channel uses for each transmitter, a symmetric scheme with the same rate (defined as the message size over the total number of channel uses by all transmitters) is obtained by repeating the original scheme \( K \) times, and in the \( i \)-th repetition shifting the transmitter indices cyclicly by \( i \).}} the number of channel uses for each transmitter does not depend on the transmitter indices. A rate \( R \) is said to be achievable if there exists an anonymous communication scheme of rate greater than or equal to \( R \), for which zero error decoding is guaranteed. The supremum of achievable rates is called the capacity \( C \).

The individual randomness size \( \rho \) measures the amount of correlated randomness at each transmitter relative to the message size (by symmetry, without loss of generality, we assume that each transmitter holds the same amount of correlated randomness, i.e., \( H(Z_1) = \cdots = H(Z_K) \)). The total randomness size \( \eta \) measures the total amount of correlated randomness at all transmitters relative to the message size.

\[
\rho = \frac{H(Z_i)}{L} \quad \text{(12)}
\]
\[
\eta = \frac{H(Z_1, \cdots, Z_K)}{L} \quad \text{(13)}
\]

## 3 Capacity of Anonymous Communications

Theorem 1 states our main result.

**Theorem 1** The capacity of anonymous communications over a parallel channel with \( K \) transmitters and 1 receiver is \( C = 1/K \). To achieve capacity, the minimum requirement on randomness size is \( \rho = 1 \) individually and \( \eta = K - 1 \) in total.

The achievability proof appears in Section 4, where we provide a scalar linear anonymous coding scheme. The converse proof on the rate appears in Section 5. The converse proof on the randomness appears in Section 6 for linear schemes and Section 7 for all possible schemes (i.e., the information theoretic converse).

When there is no anonymity constraint, the capacity is trivially 1 (only the desired transmitter sends its message) and no common randomness is needed. Therefore, in order to obtain anonymity among a set of \( K \) transmitters, the price for anonymity in communication cost is \( K \) times of that with no anonymity constraint and we further need \( K - 1 \) bits of common randomness overall and 1 bit per transmitter, to communicate 1 bit anonymously.

## 4 Proof of Theorem 1: Achievability

The achievable scheme with \( K \) transmitters is an immediate generalization of that when \( K = 3 \), presented in the introduction section. We show that to communicate 1 bit anonymously, each transmitter uses its channel once, so that the rate achieved is \( 1/K \).

We present the scheme over the binary field (any field will work in general). Denote \( a_1, \cdots, a_{K-1} \) as \( K - 1 \) i.i.d. uniform bits, that are independent of the messages. The correlated random variables are assigned as follows.

\[
Z_i = a_i, \quad i \in [1: K - 1]
\]
\[
Z_K = a_1 + \cdots + a_{K-1}
\]
The transmitted signals are

\[ X_i = a_i + 1(\theta = i)W_i, \ i \in [1 : K - 1] \]
\[ X_K = a_1 + \cdots + a_{K-1} + 1(\theta = K)W_K \]  

from which we can easily identify \( X_{i[k]} \forall i, k \in [1 : K] \).

The decoding mapping is the addition operation.

\[ g(Y) = X_1 + X_2 + \cdots + X_K \]

i.e., \( g(Y[k]) = X_{1[k]} + X_{2[k]} + \cdots + X_{K[k]} = W_k \)  

Correctness is easy to verify as the \( K \) correlated random variables lie in a \( K - 1 \) dimensional space (in fact, any \( K - 1 \) dimensional space will work) and the decoding mapping is along the null space of the correlated random variables. Anonymity is guaranteed because for all possible values of \( \theta \), the received signal is comprised of \( K \) uniform i.i.d. bits and the decoding mapping does not depend on \( \theta \). That is, when \( \theta = k, \forall k \in [1 : K] : \)

\[ H(Y[k]) = H(X_{1[k]}, \cdots, X_{K[k]}) \]

\[ = H(a_1, a_2, \cdots, a_{K-1}, W_k) \]

\[ = K \]

Remark (Coded Randomness): In our coding scheme, the common randomness variables are correlated in coded form at the transmitters. Combining with the converse, we know that coded randomness is necessary to minimize the randomness size (i.e., if we do not allow randomness to be mixed, then we must use more randomness).

Remark (Collusion): Our achievable scheme is resilient to user collusions (equivalently, prior knowledge to preclude a set of non-desired transmitters) in the following sense. Suppose each transmitter only knows he is desired or not, then any collusion of \( K - 2 \) non-desired transmitters with the receiver can not identify the desired transmitter index (i.e., the transmitters that are not in the colluding set are equally likely to be the desired).

Remark (Security): Our achievable scheme is perfectly secure in that the receiver obtains absolutely no information about all other messages beyond the desired one.

5 Proof of Theorem 1: Converse on Rate

We show that to transmit \( L \) symbols anonymously, each transmitter must use the channel at least \( N \geq L \) times. Then the rate bound \( R = \frac{L}{NK} \leq 1/K \) follows.

We first show that \( H(X_i) \geq L \), i.e., when Transmitter \( i \) is the desired transmitter, he must send a signal that contains at least as much information as that contained in his message, from the correctness constraint. Define \( W_i = (W_1, \cdots, W_{i-1}, W_{i+1}, \cdots, W_K) \).

\[ L \quad H(W_i) \]  
\[ H(W_i ; Y[i]) \]  
\[ I(W_i ; X[i], \cdots, X[K], Z[i], \cdots, Z_K, W_i) \]
Next, we show that \( H(X_i^k) \geq L, k \neq i \), i.e., when Transmitter \( i \) is not the desired transmitter, he must send a statistically equivalent signal so that the entropy is also not less than the message size, from the anonymity constraint.

Combining with the fact that \( H(X_i^k) \leq N, \forall k \), we arrive at the desired rate bound.

6 Proof of Theorem

6.1 Proof for scalar linear case when \( K = 3 \)

To illustrate the main idea in a simpler setting, we first consider the \( K = 3 \) setting and assume the scheme is scalar linear, i.e., each message and each correlated random variable is only 1 symbol. We show that each correlated random symbol must be uniformly random, \( H(Z_i) \geq L, i \in \{1, 2, 3\} \) and any two random symbols are independent, \( H(Z_i, Z_j) \geq 2L, i \neq j, i, j \in \{1, 2, 3\} \).

For a linear scheme, the transmitted signal is a linear combination of the message symbol and the correlated random variable, and the decoding mapping is also a linear combination of the received signal symbols (so the only operation allowed is taking linear combinations). Specifically, the transmitted signals are

\[
X_i^k = V_i^k W_i + U_i^k Z_i, \quad i, k \in \{1, 2, 3\}
\]

where \( V_i^k, U_i^k \) are scalars over \( \mathbb{F}_p \). The decoding coefficients are denoted as \( G_1, G_2, G_3 \in \mathbb{F}_p \) (note that the constants \( G_1, G_2, G_3 \) do not depend on the desired transmitter index \( k \)) and the decoding works as follows.

\[
W_k = G_1 X_1^k + G_2 X_2^k + G_3 X_3^k
= G_1 V_1^k W_1 + G_2 V_2^k W_2 + G_3 V_3^k W_3 + G_1 U_1^k Z_1 + G_2 U_2^k Z_2 + G_3 U_3^k Z_3
\]

As such, for any \( k \in \{1, 2, 3\} \), the undesired messages cannot appear. It follows from the equality \( 31 \) that

\[
G_1 V_1^{[1]} \neq 0, G_2 V_2^{[1]} = 0, G_3 V_3^{[1]} = 0
\]
\[ G_1V_1^{[2]} = 0, G_2V_2^{[2]} \neq 0, G_3V_3^{[2]} = 0 \]  
\[ G_1V_1^{[3]} = 0, G_2V_2^{[3]} = 0, G_3V_3^{[3]} \neq 0 \]  
\[ \Rightarrow G_1 \neq 0, G_2 \neq 0, G_3 \neq 0, V_1^{[1]} \neq 0, V_2^{[2]} \neq 0, V_3^{[3]} \neq 0, V_2^{[1]} = V_3^{[1]} = 0, V_1^{[2]} = V_2^{[2]} = 0, V_1^{[3]} = V_3^{[3]} = 0 \]  

Consider now \( X_1^{[2]} = U_1^{[2]}Z_1 \). From (28), we have

\[ \begin{align*}
L & \leq H(X_1^{[2]}) \\
& \leq H(U_1^{[2]}Z_1) \\
& = H(Z_1) \\
& \geq \rho L
\end{align*} \]  

where (38) follows from the observation that

\[ U_1^{[2]} \neq 0, \]  

as otherwise \( H(X_1^{[2]}) = 0 \), contradicting (28). Therefore, we have proved that the individual randomness size \( \rho \geq 1 \). Symmetrically, from (38) and (40), we have

\[ L \leq H(Z_2), L \leq H(Z_3), \]  
\[ U_i^{[k]} \neq 0, \ k \neq i \]  

Next, we consider \( \langle X_1^{[1]}, X_2^{[1]} \rangle = (V_1^{[1]}W_1 + U_1^{[1]}Z_1, U_2^{[1]}Z_2) \).

\[ \begin{align*}
H(X_1^{[1]}, X_2^{[1]}) & \geq H(U_2^{[1]}Z_2) + H(V_1^{[1]}W_1 + U_1^{[1]}Z_1 | U_2^{[1]}Z_2) \\
& \geq H(Z_2) + H(V_1^{[1]}W_1 + U_1^{[1]}Z_1 | Z_2, Z_1) \\
& \geq L + H(V_1^{[1]}W_1 | Z_2, Z_1) \\
& \geq L + H(V_1^{[1]}W_1)
\end{align*} \]

Then we consider \( H(X_1^{[3]}, X_2^{[3]}) \).

\[ \begin{align*}
\eta L & \geq H(Z_1, Z_2, Z_3) \\
& \geq H(Z_1, Z_2) \\
& \geq H(X_1^{[3]}, X_2^{[3]}) \\
& \geq H(X_1^{[1]}, X_2^{[1]}) \\
& \geq 2L
\end{align*} \]

Therefore we have proved that the sum randomness size \( \eta \geq 2 = K - 1 \).

Remark: From (37), we know that the correlated random variables must satisfy some linear equation, i.e., they must lie in a lower dimensional space (rank deficient) for successful decoding.
6.2 General proof for vector linear case with arbitrary $K$

We generalize the above proof to the vector linear case with arbitrary number of transmitters, $K$. We show that $H(Z_1) \geq L$ and $H(Z_1, \cdots, Z_{K-1}) \geq (K - 1)L$.

The vector linear scheme is represented as follows.

$$X_i^{[k]} = V_i^{[k]}W_i + U_i^{[k]}Z_i, \quad i, k \in [1 : K]$$  \hspace{1cm} (53)

where $V_i^{[k]}, U_i^{[k]}$ are $N \times L$ constant encoding matrices, over $\mathbb{F}_p$ (and are globally known). Note that there is no loss of generality in assuming that $Z_i$ contains $L$ symbols over $\mathbb{F}_p$, as we do not impose any statistical properties on the $L$ symbols (e.g., they are not necessarily independent). For any $i, k \in [1 : K]$,

$$W_k = \sum_{i=1}^{K} G_i X_i^{[k]}$$  \hspace{1cm} (54)

$$= \sum_{i=1}^{K} G_i V_i^{[k]}W_i + \sum_{i=1}^{K} G_i U_i^{[k]}Z_i$$  \hspace{1cm} (55)

The decoding mapping is specified by the constant filtering matrices $G_i$, which have dimension $L \times N$ over $\mathbb{F}_p$. Then we have

$$\text{rank}(G_i V_i^{[k]}) = L, \quad k \in [1 : K]$$  \hspace{1cm} (56)

$$G_i V_i^{[k]} = 0, \quad k \neq i, \quad i, k \in [1 : K]$$  \hspace{1cm} (57)

Following the proof presented in the previous section, we proceed to consider $G_1 X_1^{[2]}$.

$$L \quad H(W_1)$$  \hspace{1cm} (58)

$$= \text{H}(G_1 V_1^{[1]}W_1)$$  \hspace{1cm} (59)

$$H(G_1 V_1^{[1]}W_1|Z_1)$$  \hspace{1cm} (60)

$$H(G_1 X_1^{[1]}|Z_1)$$  \hspace{1cm} (61)

$$\leq H(G_1 X_1^{[1]})$$  \hspace{1cm} (62)

$$H(G_1 X_1^{[2]})$$  \hspace{1cm} (63)

$$H(G_1 U_1^{[2]}Z_1)$$  \hspace{1cm} (64)

$$\leq H(Z_1)$$  \hspace{1cm} (65)

$$\rho L$$  \hspace{1cm} (66)

Therefore, we have proved that the individual randomness size $\rho \geq 1$. As a byproduct, from (64), we obtain that

$$\text{rank}(G_1 U_1^{[2]}) = L$$  \hspace{1cm} (67)
as otherwise we have the contradiction that \( H(G_1 U_1^{[2]} Z_1) < L \). Symmetrically, from (67), we have
\[
\text{rank}(G_k U_k^{[i]}) = L, \ k \neq i
\] (68)

Next, we consider the total randomness size. We first prove a lemma.

**Lemma 1** For all \( i \in [1 : K - 1] \), we have
\[
H(G_1 X_1^{[i+1]}, G_2 X_2^{[i+1]}, \ldots, G_i X_i^{[i+1]}) \geq iL
\] (69)

**Proof:** The proof is based on induction. Note that the basis case where \( i = 1 \) is proved in (63). Suppose now (69) holds when \( i = j, j \in [1 : K - 2] \), i.e.,
\[
H(G_1 X_1^{[j+1]}, G_2 X_2^{[j+1]}, \ldots, G_j X_j^{[j+1]}) \geq jL
\] (70)

Now consider the case where \( i = j + 1 \).
\[
H(G_1 X_1^{[j+2]}, G_2 X_2^{[j+2]}, \ldots, G_{j+1} X_{j+1}^{[j+2]})
\]
\[
= H(G_1 X_1^{[j+1]}, G_2 X_2^{[j+1]}, \ldots, G_j X_j^{[j+1]})
\]
\[
+ H(G_{j+1} X_{j+1}^{[j+1]} | G_1 X_1^{[j+1]}, \ldots, G_j X_j^{[j+1]})
\] (71)
\[
\geq jL + H(G_{j+1} X_{j+1}^{[j+1]} | Z_1, \ldots, Z_j, Z_{j+1})
\] (72)
\[
\geq jL + H(G_{j+1} V_j^{[j+1]} | W_{j+1} | Z_1, \ldots, Z_{j+1})
\] (73)
\[
\geq jL + H(W_{j+1})
\] (74)
\[
= (j + 1)L
\] (75)

Since both the basis and the inductive steps have been performed, by mathematical induction, we have proved that (69) holds for all \( i \in [1 : K - 1] \). The proof for Lemma 1 is complete.

Finally, consider (69) and set \( i = K - 1 \). We have
\[
(K - 1)L \leq H(X_1^{[K]}, \ldots, X_{K-1}^{[K]})
\] (76)
\[
\leq H(Z_1, \ldots, Z_{K-1})
\] (77)
\[
\leq \eta L
\] (78)

Therefore we have proved that the sum randomness size \( \eta \geq K - 1 \), for any rate \( R = \frac{L}{NK} \).

7 Proof of Theorem 1: Information Theoretic Converse on Randomness for Capacity Achieving Schemes

We show that when the scheme is capacity achieving, i.e., the rate achieved is \( 1/K \), i.e., \( H(X_i^{[k]}) = N = L, \forall i, k \in [1 : K] \), then the randomness sizes \( \rho = 1 \) and \( \eta = K - 1 \) are both information theoretically optimal.
7.1 Proof for binary scalar case when $K = 3$

Before presenting the general proof for arbitrary $K$, we first consider the $K = 3$ case and assume that each message is one bit, to illustrate the idea. Then in this case, $L = 1$ and the field is $\mathbb{F}_2$. In this case, we need to show that $H(Z_i) \geq 1$ and $H(Z_1, Z_2, Z_3) \geq 2$.

First, for capacity achieving schemes, i.e.,

$$H(X_i^{[k]}) = 1, \forall i, k \in \{1, 2, 3\}$$

the received signal is uniformly random. The proof is deferred to Lemma 2 for the general case. That is, for any $k$,

$$X_1^{[k]} \, X_2^{[k]} \, X_3^{[k]} \text{ is uniformly distributed.}$$

Next, consider $X_1^{[2]} \, X_2^{[2]} \, X_3^{[2]} \, W_2$. Note that

$$H(W_2|X_1^{[2]}, X_3^{[2]}) = H(W_2) = L$$

$$H(X_2^{[2]}|X_1^{[2]}, X_3^{[2]}) = L$$

$$H(W_2|X_1^{[2]}, X_2^{[2]}, X_3^{[2]}) = 0$$

Then we have the observation that for any realization of $X_1^{[2]} \, X_3^{[2]} \, W_2$ has a one-to-one mapping to $X_2^{[2]}$, i.e.,

$$H(X_2^{[2]}|W_2, X_1^{[2]}, X_3^{[2]}) = 0$$

Repeating the argument for $W_1$ and $W_3$, we have

$$H(X_1^{[1]}|W_1, X_2^{[1]}, X_3^{[1]}) = 0$$

$$H(X_3^{[3]}|W_3, X_1^{[3]}, X_2^{[3]}) = 0$$

From the anonymity constraint (10) and the correctness constraint (9), we know that

$$(X_1^{[1]}, X_2^{[1]}, X_3^{[1]}, W_1) \sim (X_1^{[k]}, X_2^{[k]}, X_3^{[k]}, g, W_k)$$

We now consider the individual randomness size. Combining (85) and (88), we have

$$H(X_2^{[1]}|W_1, X_1^{[1]}, X_3^{[1]}) = 0$$

Then

$$I(X_2^{[1]}; W_2) \leq I(X_2^{[1]}, W_1, X_1^{[1]}; W_2)$$

$$= I(W_1, X_1^{[1]}, X_3^{[1]}; W_2) + I(X_2^{[1]}|W_2, W_1, X_1^{[1]}, X_3^{[1]})$$

$$\leq I(W_1, X_1^{[1]}, Z_1, X_3^{[1]}; W_3, Z_3; W_2) + 0$$

$$\leq 0$$
and

\[ \begin{align*}
1 & \quad H(X_2^{[1]}) \\
2 & \quad I(X_2^{[1]}; W_2, Z_2) \\
3 & \quad I(X_2^{[1]}; Z_2 | W_2) \\
\leq & \quad H(Z_2)
\end{align*} \]

(93) (94) (95) (96) (97)

Therefore the individual randomness size satisfies that \( \rho \geq 1 \).

We proceed next to consider the sum randomness size. Combining (85), (86), (87) and (88), we have obtained the structure of the decoding mapping, i.e., for any 3-tuple of the received signal, if 2 elements are fixed, the remaining element has a one-to-one mapping with the desired message. For example, when \( Y[k] = (0, 0, 0) \), suppose that \( W_k = g(Y[k]) = g(0, 0, 0) = w \), \( w \in \{0, 1\} \), then \( g(0, 0, 1) = g(h, 0, 0) = g(1, 0, 0) = 1 - w \). Proceeding along this line, the decoding mapping is uniquely identified as follows.

| \( Y[k] \) | \( W_k \) |
|---|---|
| (0, 0, 0) | \( w \) |
| (0, 0, 1) | \( 1 - w \) |
| (0, 1, 0) | \( 1 - w \) |
| (0, 1, 1) | \( w \) |
| (1, 0, 0) | \( 1 - w \) |
| (1, 0, 1) | \( w \) |
| (1, 1, 0) | \( w \) |
| (1, 1, 1) | \( 1 - w \) |

(99)

We are now ready to show that

\[ H(X_2^{[1]}, X_3^{[1]} | Z_1, Z_2, Z_3) = 0. \]

(100)

Consider an arbitrary realization of \((W_1, Z_1, Z_2, Z_3) = (w_1, z_1, z_2, z_3)\), drawn according to the correct joint distribution \((W_1\) is independent of \(Z_1, Z_2, Z_3)\). Then \( X_1^{[1]} \) is a constant (denoted as \( x_1 \)) as \( X_1^{[1]} \) is a function of \( W_1 \) and \( Z_1 \). We now show that \( X_2^{[1]}\) and \( X_3^{[1]} \) are now constants as well. Note that the only variables that are random now are \( W_2, W_3 \). Suppose \( X_2^{[1]} \) is still random, depending on the value of \( W_2 \). Then consider two realizations of \( X_2^{[1]} \) (denoted as \( x_2, x_2', x_2 \neq x_2' \)) and the received signal realizations

\[ \begin{align*}
y_1 & = (x_1, x_2, X_3^{[1]}) \\
y_2 & = (x_1, x_2', X_3^{[1]})
\end{align*} \]

(101) (102)

From the decoding mapping table, we know that \( g(y_1) \neq g(y_2) \). However, from the correctness constraint, we know that \( g(y_1) = g(y_2) = w_1 \). Therefore, we arrive at the contradiction and \( X_2^{[1]}, X_3^{[1]} \) are deterministic functions of the correlated random variables. Then we have

\[ \eta \quad H(Z_1, Z_2, Z_3) \]

(103)
Therefore the sum randomness size $\eta \geq 2$ and the proof is complete.

### 7.2 Proof for Arbitrary $K$

We follow the steps of the proof for $K = 3$ binary case and show $H(Z_i) \geq L, H(Z_1, Z_2, \cdots, Z_K) \geq (K - 1)L$.

First, we present a lemma, which says that the received signals are uniformly random, when the scheme is capacity achieving.

**Lemma 2**

\[ H(X_i^k) = N = L, \forall i, k \in [1 : K] \] (107)

\[ \Rightarrow \ H(X_1^k, \cdots, X_K^k) = KL, \forall k \in [1 : K] \] (108)

**Proof:** Note that (107) implies that $H(X_1^k, \cdots, X_K^k) \leq KL$. It suffices to prove only the other direction. Define $X_i^k = (X_i^{[1]}, \cdots, X_i^{[i-1]}, X_i^{[i]}, \cdots, X_i^{[K]}).

\[
H(X_1^k, \cdots, X_K^k) = \sum_{i=1}^{K} H(X_i^k | X_1^k, \cdots, X_i^{[i-1]}) \geq \sum_{i=1}^{K} H(X_i^k | X_i^k) \geq \sum_{i=1}^{K} I(W_i; X_i^k | X_i^k) \geq \sum_{i=1}^{K} H(W_i | X_i^k) \geq \sum_{i=1}^{K} H(W_i | X_i^k, W_i, Z_1, \cdots, Z_K) \geq KL
\] (115)

Next, note that

\[ H(W_i | X_i^k) \stackrel{[113]}{=} L, \] (116)
\( H(X_i^{[i]} | X_i^{[i]}) \leq L, \) \hspace{1cm} (117)

\( H(W_i | X_i^{[i]}, X_i^{[i]}) \leq 0 \) \hspace{1cm} (118)

Note that \( W_i \) is independent of \( X_i^{[i]} \). Then we have the observation that for any realization of \( X_i^{[i]} \), 
\( W_i \) has a one-to-one mapping to \( X_i^{[i]} \), i.e.,
\( H(X_i^{[i]} | W_i, X_i^{[i]}) = 0 \) \hspace{1cm} (119)

From the anonymity constraint (10) and the correctness constraint (9), we know that for any \( i \in [1 : K] \),
\( (X_i^{[i]}, X_i^{[i]}, g, W_i) \sim (X_i^{[1]}, X_i^{[1]}, g, W_i) \) \hspace{1cm} (120)

Combining (119) and (120), we have
\( H(X_i^{[1]} | W_1, X_i^{[1]}) = 0 \) \hspace{1cm} (121)

Then for \( i \neq 1 \),
\[
\begin{align*}
I(X_i^{[1]}; W_i) &\leq I(X_i^{[1]}, W_1, X_i^{[1]}; W_i) \\
&= I(W_1, X_i^{[1]}; W_i) + I(X_i^{[1]}; W_i | W_1, X_i^{[1]}) \\
&\overset{(123)}{\leq} I(W_1, X_i^{[1]}, Z_i; W_i) + 0 \\
&\overset{(124)}{=} 0
\end{align*}
\]

where \( Z_i = (Z_1, \ldots, Z_{i-1}, Z_{i+1}, \ldots, Z_K) \).

For the individual randomness size, we have
\[
L \overset{(107)}{=} H(X_i^{[1]}) \overset{(7)}{=} I(X_i^{[1]}; W_i, Z_i) \overset{(125)}{=} I(X_i^{[1]}; Z_i | W_i) \overset{(128)}{\leq} H(Z_i) \overset{(12)}{=} \rho L 
\]

Therefore \( \rho \geq 1 \).

For the sum randomness size, as (119) holds for all \( i \in [1 : K] \) and from (120), we know that if any \( K - 1 \) elements of the received signal are determined, the remaining element has a one-to-one mapping with the desired message, which means that

For 2 received signal tuples that differ in 1 element,
\( i.e., \ y_1 = (x_1, \ldots, x_k, \ldots, x_K), \)
\( y_2 = (x_1, \ldots, x_k', \ldots, x_K), \)
we have \( g(y_1) \neq g(y_2). \) \hspace{1cm} (131)

Then we claim that \( X_2^{[1]}, \ldots, X_K^{[1]} \) are functions of \( Z_1, \ldots, Z_K \). This result is stated in the following lemma.
Lemma 3

\[ H(X_2^{[1]}, \ldots, X_K^{[1]}|Z_1, \ldots, Z_K) = 0 \]  

Proof: Consider one arbitrary realization of \( W_1, Z_1, \ldots, Z_K \), denoted as \( (W_1, Z_1, \ldots, Z_K) = (w_1, z_1, \ldots, z_K) \). As \( W_1, Z_1 \) are fixed, then \( X_1^{[1]} \) is a constant, denoted as \( x_1 \). We show that \( X_2^{[1]}, \ldots, X_K^{[1]} \) are constants now. To set up the proof by contradiction, suppose there exists one \( X_k^{[1]} \) that can take multiple values. Denote two such values as \( x_k, x'_k, x_k \neq x'_k \). The other \( X_i^{[1]}, i \neq k \) are assumed to be constants and denoted as \( x_i \). Note that for fixed \( z_2, \ldots, z_k, X_2^{[1]}, \ldots, X_K^{[1]} \) are conditionally independent as now the randomness only comes from the messages \( W_2, \ldots, W_K \) and the messages are independent. We now have two different received signal tuples

\[ y_1 = (x_1, \ldots, x_k, \ldots, x_K) \]  
\[ y_2 = (x_1, \ldots, x'_k, \ldots, x_K) \]

From (131), we know that \( g(y_1) \neq g(y_2) \). However, this contradicts with the fact that \( g(y_1) = g(y_2) = w_1 \). Therefore we have arrived at the contradiction and \( X_2^{[1]}, \ldots, X_K^{[1]} \) are functions of \( Z_1, \ldots, Z_K, W_1 \). Further \( X_2^{[1]}, \ldots, X_K^{[1]} \) are independent of \( W_1 \) and we have proved the lemma.

From Lemma 3 we have

\[ \eta L \geq H(Z_1, \ldots, Z_K) \]  
\[ \geq H(X_2^{[1]}, \ldots, X_K^{[1]}|Z_1, \ldots, Z_K) \]  
\[ \geq H(X_2^{[1]}, \ldots, X_K^{[1]}) \]  
\[ \geq (K - 1) L \]

Therefore the desired sum randomness size bound follows and the proof is complete.

Remark: The above proof relies on the assumption that the scheme is capacity achieving. Otherwise, Lemma 3 and Lemma 5 may not hold.

Remark: The individual randomness size bound holds without the constraint that the achieved rate is equal to the capacity, i.e., we have \( \rho \geq 1 \) for any positive rate (the total randomness size bound, however, hinges on the assumption of capacity achieving schemes). A sketch of proof idea is as follows (the above proof is more informative in that the combinatorial structure of the decoding mapping is revealed). We first note that the transmitted signal from Transmitter \( i, i \neq 1 \) is independent of \( W_1 \), i.e., \( I(X_i^{[1]}; W_1) = 0 \). Next, from the anonymity constraint, the same relation on the mutual information must hold when \( W_i \) is desired, i.e., \( I(X_i^{[1]}; W_i) = 0 \), meaning that the transmitted signal from Transmitter \( i \) does not contain any information about \( W_i \). To guarantee this, the randomness needed must be at least as large as the message size.

Remark: A more general condition where the bound on the total randomness size holds unconditionally for arbitrary positive rates is when we require the transmitted signal to be deterministic functions of the correlated random variable when he is not desired, i.e., \( H(X_i^{[k]}|Z_i) = 0, i \neq k \) (in other words, the messages do not play a role when they are not desired. As the messages are independent among themselves and of the correlated random variables, it will be interesting if they help to reduce total randomness size). Lemma 3 proves that this deterministic condition holds for capacity achieving schemes. After we assume this deterministic condition to be satisfied, the proof is the same as that presented above after Lemma 3.
8 Conclusion

We consider the problem of anonymous communications from an information theory perspective. We have characterized the capacity of anonymous communications over a parallel channel with \(K\) transmitters and 1 receiver, to be \(C = 1/K\). Further, the minimum randomness sizes required are \(\rho = 1\) per transmitter and \(\eta = K - 1\) for all transmitters. This work represents a step towards using information theoretic tools to understand the fundamental limits of anonymous network communications.

References

[1] H. Sun and S. A. Jafar, “The Capacity of Private Information Retrieval,” *IEEE Transactions on Information Theory*, vol. 63, no. 7, pp. 4075–4088, 2017.

[2] ——, “The Capacity of Robust Private Information Retrieval with Colluding Databases,” *IEEE Transactions on Information Theory*, vol. 64, no. 4, pp. 2361–2370, 2018.

[3] K. Banawan and S. Ulukus, “The Capacity of Private Information Retrieval from Coded Databases,” *IEEE Transactions on Information Theory*, vol. 64, no. 3, pp. 1945–1956, 2018.

[4] K. Peng, *Anonymous communication networks: Protecting privacy on the web*. CRC Press, 2014.

[5] J. Ren and J. Wu, “Survey on anonymous communications in computer networks,” *Computer Communications*, vol. 33, no. 4, pp. 420–431, 2010.

[6] G. Danezis and C. Diaz, “A survey of anonymous communication channels,” Tech. Rep., 2008.

[7] D. Chaum, “The dining cryptographers problem: Unconditional sender and recipient untraceability,” *Journal of cryptology*, vol. 1, no. 1, pp. 65–75, 1988.

[8] S. Dolev and R. Ostrobsky, “Xor-trees for efficient anonymous multicast and reception,” *ACM Transactions on Information and System Security (TISSEC)*, vol. 3, no. 2, pp. 63–84, 2000.