Universal Multiparty Data Exchange

Himanshu Tyagi †
Shun Watanabe ‡

Abstract—Multiple parties observing correlated data seek to recover each other’s data and attain omniscience. To that end, they communicate interactively over a noiseless broadcast channel: Each bit transmitted over this channel is received by all the parties. We give a universal interactive protocol for omniscience which requires communication of rate only $O(n^{-1/2} \sqrt{\log n})$ more than the optimal rate for every independent and identically distributed (in time) sequence of data.

I. INTRODUCTION

An $m$ party omniscience protocol is an interactive communication protocol that enables $m$ parties to recover each other’s data. The communication is error-free and is in a broadcast mode wherein the transmission of each party is received by all the other parties. Such protocols were first considered in [4] in a two-party setup, where bounds for the number of bits communicated on average and in the worst-case were derived for the case when no error is allowed. The $m$ party version, and the omniscience terminology, was proposed in [3] where the collective observations of the parties was assumed to be an independent and identically distributed (IID) sequence. It was shown in [3] that a noninteractive protocol based on sending random hash bits of appropriate rates is sum-rate optimal. Recently, a single-shot protocol was proposed for the two-party case in [6] which used interactive communication and, when applied to IID observations, was of optimal length even up to the second-order asymptotic term.

A common feature of these prior works is that the protocol relies on the knowledge of the underlying distribution. We present a universal omniscience protocol, i.e., it does not rely on the knowledge of the underlying IID distribution, and is of optimal rate. In fact, we show that our protocol has an excess rate less than $O(n^{-1/2} \sqrt{\log n})$ over the optimal rate, on average as well as in the worst-case. Note that even when the distribution is known, for $m = 2$, [6] showed that a worst-case excess rate of $O(n^{-1/2})$ is necessary. While no converse bounds are available for the average rate, our results are the best-known for this case as well.

It was shown in [8] that interaction enables an ACK−NACK based universal variable-length coding scheme for the Slepian-Wolf problem, where only party 1 needs to send its data to party 2. Our protocol, too, is interactive in a similar spirit, but it relies on carefully increasing the rate of communication for each party. Note that while for $m = 2$ a simple extension of the protocol in [8] works for the data exchange problem as well, this is not the case when $m > 2$. For $m > 2$, the order in which the parties communicate must be carefully chosen. We give a very simple criterion for choosing this order of communication and show that the resulting protocol is universally rate-optimal. A key feature of our protocol is its recursive structure wherein whenever a subset of parties attains local omniscience, the rate vector appears as if the parties were together to begin with. Furthermore, any subset $A$ attaining local omniscience does so by using only the communication from parties in $A$ and using the minimum possible rate of communication. We analyze our protocol for individual sequences, and the rate optimality result follows upon taking limits. We omit the proofs due to lack of space. Complete proofs, a more elaborate discussion and an application of our protocol to the secret key agreement can be found in the full version [7].

The remainder of this paper is organized as follows: The next section contains the formal description of the omniscience problem. We give our protocol in Section III. The section is divided into three parts with the first two parts containing a simpler description of our protocol under ideal assumptions, illustrated using examples, and the final part containing the full description of our protocol and main results.

II. OMNISCIENCE OR MULTIPARTY DATA EXCHANGE

Parties in a set $\mathcal{M} = \{1, \ldots, m\}$ observe an IID sequence $X^n_{\mathcal{M}} = (X^n_{\mathcal{M}1}, \ldots, X^n_{\mathcal{M}m})$, with the $i$th party observing $\{X^n_{\mathcal{Mi}}\}_{i=1}^n$ and $X^n_{\mathcal{M}} = (X^n_{\mathcal{Mi}} \in X^n_{\mathcal{Mi}}, i \in \mathcal{M}) \sim P_{X^n_{\mathcal{M}}}$ denoting the collective data at the $t$th time instance. The parties have access to shared public randomness $U$ such that $U$ is independent jointly of $X^n_{\mathcal{M}}$. Thus, the $i$th party observes $(X^n_{\mathcal{Mi}}, U)$.

The parties communicate with each other interactively using a $2$-tree-protocol; see [5] for a definition of tree-protocols. The (worst-case) length $\|\pi\|_{\text{w}}$ of a protocol $\pi$ is the maximum number of bits that are transmitted in any execution of the protocol and equals the depth of the protocol-tree. Also, the average length $\|\pi\|_{\text{av}}$ is given by the expected value of the number of bits transmitted in an execution of the protocol $\pi$.

In the omniscience problem, the parties communicate interactively to recover each other’s data. A protocol $\pi$ constitutes an $\epsilon$-omniscience protocol if the $i$th party can output an estimate $\hat{X}_i = \hat{X}_i(X^n_{\mathcal{Mi}}, U, \Pi) \in X^n_{\mathcal{Mi}}$ such that

$$\mathbb{P}(\hat{X}_i = X^n_{\mathcal{Mi}}, i \in \mathcal{M}) \geq 1 - \epsilon.$$  

Definition 1 (Communication for omniscience). Given IID observations with a common distribution $P_{X_{\mathcal{M}}}$, as above, for $0 \leq \epsilon < 1$, a rate $R \geq 0$ is an $\epsilon$-achievable omniscience rate if there exists an $\epsilon$-omniscience protocol $\pi$ with length

---

1†Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560012, India. Email: htyagi@ece.iisc.ernet.in.
2‡Department of Computer and Information Sciences, Tokyo University of Agriculture and Technology, Tokyo 184-8588, Japan. Email: shunwata@cc.tuat.ac.jp.

†HT is supported by DRDO, India under grant DRDO00649. SW is supported in part by the JSPS KAKENHI under grant 16H06091.
less than \( nR \) for all \( n \) sufficiently large. The infimum over all 
\( \epsilon \)-achievable omniscience rates is denoted by \( R_{\epsilon}(P_{X|M}) \). The 
minimum rate of communication for omniscience \( R(P_{X|M}) \) is 
given by \( R(P_{X|M}) = \lim_{\epsilon \to 0} R_{\epsilon}(P_{X|M}) \).

The minimum average rate of communication for omniscience \( R^w(P_{X|M}) \) is 
defined similarly by replacing length \( |\pi| \) with average length \( \|\pi\|_w \).

A simple characterization of \( R(P_{X|M}) \) as a linear program 
was derived in [3]. Here we recall an alternative characterization 
in terms of partitions of \( \mathcal{M} \). This form was obtained 
as a lower bound for \( R(P_{X|M}) \) in [3], which was shown to 
be tight for \( m = 3 \). Later, it was shown in [2] that, in fact, 
this lower bound is tight for all \( m \). Our universal protocol 
for omniscience directly achieves this rate, thereby providing an 
alternative proof for the tightness of this lower bound.

**Theorem 1** ([3], [2]). Given a distribution \( P_{X|M} \),

\[
R(P_{X|M}) = R^w(P_{X|M}) = \max_{\sigma \in \Sigma(M)} \mathbb{H}_\sigma(P_{X|M}),
\]

(1)

where \( \Sigma(M) \) is the set of partitions \( \sigma = \{\sigma_1, \ldots, \sigma_k\} \), \( k \geq 2 \),
of \( \mathcal{M} \) and \( \mathbb{H}_\sigma(P_{X|M}) \) is 
\[
|\sigma|^{-1} \sum_{i=1}^{\|\sigma\|} H(X_M|X_{\sigma_i}).
\]

**Remark 1** (Rate-region of communication for omniscience).
The right-side of (1) is equal to the minimum sum 
\[
\sum_{i=1}^{\|\sigma\|} R_i \text{ of rates } R = (R_1, \ldots, R_m) \text{ such that } \sum_{i \in B} R_i \geq H(X_B|X_{B'}), \quad \forall B \subset M.
\]

Following [3], the collection of all rate vectors \( R \) satisfying 
the constraints above, termed the CR region, will be 
denoted by \( R_{\text{CR}}(\mathcal{M}|P_{X|M}) \), and the minimum sum rate by 
\( R_{\text{CR}}(\mathcal{M}|P_{X|M}) \). When the distribution \( P_{X|M} \) is clear from the 
context, we shall omit it from the notation.

We shall call the finest partition \( \sigma \) that maximizes \( \mathbb{H}_\sigma \) 
in (1) (see [1]) the dominant partition. The finest partition 
\( \sigma_f(M) = \{\{i\}, i \in M\} \) plays a particularly important role in 
our protocol. Note that when the finest partition is dominant, 
the optimal rate assignment is uniquely given by the solution 
\[
R^*(M) = (R^*_1, \ldots, R^*_m)
\]

of

\[
\sum_{i \in M \setminus \{j\}} R_i = H(X_M|X_j), \quad j = 1, \ldots, m; \quad (2)
\]

\( R^*(A) \) is defined similarly with \( A \) replacing \( M \).

**III. A UNIVERSAL PROTOCOL FOR OMNISCIENCE**

We shall describe a universal protocol for omniscience, which, when \( X_M \) is observed, will transmit communication 
of rate no more than \( R_{\text{CR}}(\mathcal{M}|P_{X|M}) \). Since the empirical 
distribution approaches the true distribution in the limit as \( n \) 
goes to \( \infty \), and \( R_{\text{CR}}(\mathcal{M}|P_{X|M}) \) is continuous in \( P_{X|M} \), 
we achieve the optimal rate universally. To present the main idea 
behind our protocol, we first describe it under ideal conditions.

**A. Simplified protocol under ideal assumptions**

We make two assumptions: (a) **Continuous rate assumption**: Rates can be continuously \(^3\) increased as a function of time;

(b) **Ideal decoder assumption**: We assume the availibility of an error-free, ideal decoder \( \text{DEC}_{\text{id}} \) which correctly decodes 
a sequence once sufficient communication is sent and 
declares a NACK otherwise.\(^4\) A standard universal decoder used 
in source coding is the minimum entropy decoder which, 
given side-information \( y \) and an \( nR \)-bit random hash of \( X^n \), 
searches for the unique sequence \( x \) such that the joint type 
\( P_{X^nY} = P_{XY} \) satisfies \( H(X|Y) \leq R \) and the hash of \( x \) 
matches the received hash bits. The decoder that we prescribe 
in the next section works on a similar principle and it searches 
for any possible subsequence \( x_A \) of data that it can decode 
with the current rate. To avoid the additional complications 
due to decoding error, we first assume the availability of an 
ideal decoder \( \text{DEC}_{\text{id}} \) which enables omniscience for all parties 
j \( j \in A \) as soon as the rate received from the parties in \( A \) is 
sufficient. That is, the ideal decoder guarantees that each party 
i \( i \in A \) can recover the correct sequence \( x_A \) if the rates of 
communication \( R = (R_i, i \in A) \) satisfy \( R \in R_{\text{CR}}(A|P_{X_A}) \). Furthermore, if \( R \notin R_{\text{CR}}(A|P_{X_A}) \), the ideal decoder does not 
mistakenly output a wrong sequence \( x_A^* \) but declares a NACK 
instead.

**Protocol 1:** Ideal decoder \( \text{DEC}_{\text{id}}(j, \sigma, R) \)

**Input:** An index \( 1 \leq j \leq m \), a partition \( \sigma \in \Sigma(M) \), a 
rate vector \( R = (R_1, \ldots, R_m) \).

**Output:** An ACK message \( \text{ACK}(A) \) or a NACK message

1) For \( \sigma_i \) such that \( j \in \sigma_i \), find the maximal set \( A \subset M \)

such that \( \sigma_i \subset A \), \( (R_i, i \in A) \in R_{\text{CR}}(A|P_{X_A}) \) and 
reveal \( x_A \) to party \( j \).

2) **if such an \( A \) was found in Step 1 then**

   | return \( \text{ACK}(A) \).

   **else**

   \( \text{NACK} \).

With this ideal decoder at our disposal, under the continuous 
rates assumption, finding a universal protocol is tantamount to 
finding a policy for increasing the rates \( (R_1, \ldots, R_m) \) such that 
when the rate vector enters \( R_{\text{CR}}(\mathcal{M}|P_{X|M}) \) for the first time, 
the sum rate is \( R_{\text{CR}}(\mathcal{M}|P_{X|M}) \). Note that initially the marginal 
types \( P_{X_A} \) are available to each party and can be transmitted 
using \( O(\log n) \) bits, since there are only polynomially many 
types. Also, if a subset \( A \) attains local omniscience in the 
middle of the protocol, any \( j \in A \) upon recovering \( x_A \) can 
transmit \( P_{X_A} \) in \( O(\log n) \) bits to all the parties, who in turn 
can use it to compute \( H(P_{X_A}) \).

As an illustration, consider the simple case when \( m = 2 \). 
Parties first share \( P_{X_1} \) and \( P_{X_2} \); suppose \( H(P_{X_1}) \geq H(P_{X_2}) \). 
Then, party 1 starts communicating and increases its rate \( R_1 \) at 
slope 1. When the rate \( R_1 \) reaches \( H(P_{X_1}) - H(P_{X_2}) \), party 
2 starts communicating at slope 1 as well. Throughout the 
protocol, each party is trying to decode the observation of 
the other by using the ideal decoder \( \text{DEC}_{\text{id}} \), and they keep on 
communicating as long as the ideal decoders output NACKs.

\(^3\)Clearly, this does not hold in practice since the rate of communication 
corresponds to the number of bits transmitted per \( n \) and can only be increased 
in discrete steps.

\(^4\)In our ideal protocol, we do not account for the rate needed to send NACKs. 
In the actual protocol, described below, each NACK symbol counts for a bit of 
communication.
The parties will decode each other as soon as \((R_1, R_2)\) enters \(R_{CO} \{1, 2\} | P_{x_1, x_2}\), i.e., when
\[
R_1 \geq H(\mathcal{X}_1 | \mathcal{X}_2) \quad \text{and} \quad R_2 \geq H(\mathcal{X}_2 | \mathcal{X}_1),
\]
where \((\mathcal{X}_1, \mathcal{X}_2) \sim P_{x_1, x_2}\). Note that once both the parties start communicating, the difference \(R_1 - R_2\) is maintained as \(H(\mathcal{X}_1) - H(\mathcal{X}_2)\). Thus, when \((R_1, R_2)\) enters \(R_{CO} \{1, 2\}\) above, \((R_1, R_2)\) must coincide with the solutions \(\mathbf{R}^* \{1, 2\}\) of \(2\). The red line in Figure 1 illustrates this evolution of rates. Note that it is also possible to proceed along the blue line for the \(m = 2\) case. However, its extension to a general \(m\) is not clear.

Our protocol extends the simple protocol above to a general \(m\). Specifically, we first arrange parties in decreasing order of the entropy of the empirical distribution of their local observations, which are shared in \(O(\log n)\)-bits. Assuming \(H(P_{x_i}) \geq H(P_{x_2}) \geq \ldots \geq H(P_{x_m})\), party 1 starts communicating, and the \(i\)th party starts communicating when \(R_1 \geq H(P_{x_i}) - H(P_{x_{i+1}}), \) a condition termed the balanced difference condition. When a subset \(A\) attains local omniscience, our protocol changes mode and works with a modified model treating the parties in \(A\) as collocated. To that end, it makes two changes: (i) the rate-slope\(^5\) for each party \(i \in A\) is decreased to \(1/|A|\), thereby ensuring that collectively parties in \(A\) increase the rate of communication \(R_A\) at slope 1; and (ii) a new party now starts communicating when the balanced difference condition holds for the modified model. Note that since parties in \(A\) have recovered \(x_A\), any one party \(i \in A\) can compute the type \(P_{x_A}\), and transmit it using \(O(\log n)\) bits.

Our key observation is that at this point the rates appear as if collocated. To that end, it makes two changes: (i) the rate-slope\(^5\) for each party \(i \in A\) is decreased to \(1/|A|\), thereby ensuring that collectively parties in \(A\) increase the rate of communication \(R_A\) at slope 1; and (ii) a new party now starts communicating when the balanced difference condition holds for the modified model. Note that since parties in \(A\) have recovered \(x_A\), any one party \(i \in A\) can compute the type \(P_{x_A}\), and transmit it using \(O(\log n)\) bits.

Our observation is that at this point the rates appear as if collocated. To that end, it makes two changes: (i) the rate-slope\(^5\) for each party \(i \in A\) is decreased to \(1/|A|\), thereby ensuring that collectively parties in \(A\) increase the rate of communication \(R_A\) at slope 1; and (ii) a new party now starts communicating when the balanced difference condition holds for the modified model. Note that since parties in \(A\) have recovered \(x_A\), any one party \(i \in A\) can compute the type \(P_{x_A}\), and transmit it using \(O(\log n)\) bits.

Example 1. For \(m = 3\), let \(X_1 \sim \text{Ber}(1/2)\), \(X_2 \sim \text{Ber}(q)\), and \(X_3 = X_1 \oplus X_2\). In this case, \(R_{CO} \{1, 2, 3\}\) is given by rate vectors satisfying \(R_1 + R_2 \geq 1, R_2 + R_3 \geq h(q),\) and \(R_1 + R_3 \geq h(q)\). When \(1/2 < h(q) \leq 1\), the finest partition is dominant, and \(R_{CO} \{1, 2, 3\}\) is achieved by the unique rate assignment \(\mathbf{R}^* = (1/2, 1/2, 2h(q) - 1)/2\). In our protocol, parties 1 and 2 communicate first and increase their rates at slope 1 until \(R_1 = R_2 = H(X_1) - H(X_3) = H(X_2) - H(X_3) = 1 - h(q)\). At this point, party 3 starts communicating and all the parties increase their rates at slope 1. Owing to the initial lead of \(R_1\) and \(R_2\) over \(R_3\), all the parties reach \(\mathbf{R}^*\) simultaneously.\(^7\)

When \(\mathbb{H}_\sigma\) is maximized by a partition \(\sigma\) other than the finest partition \(\sigma_f(\mathcal{M})\), as our protocol proceeds, the parties in parts of \(\sigma\) attain local omniscience, along the way, before all the parties attain omniscience. Consider the following example.

Example 2. Let \(W_1, W_2 \sim \text{Ber}(1/2)\) and \(V_1, V_2 \sim \text{Ber}(q)\) for some \(0 < q < 1/2\), and let \(X_1 = (W_1, W_2), X_2 = (W_1 \oplus V_1, W_2),\) and \(X_3 = W_2 \oplus V_2\). In this case, \(\mathbb{H}_{12|3} = 1 + 3h(q)\) is dominant, and our protocol proceeds as follows: Parties 1 and 2 start increase their rates at slope 1. When their rates reach \(h(q)\), they attain local omniscience and the protocol changes its mode. At this point they start increasing their rates at slope 1/2 and continue doing so until \(R_1 + R_2\) reaches \(H(X_1, X_2) - H(X_3) = 1 + h(q)\), namely the balanced difference condition for \(12|3\) holds. Now, party 3 starts communicating at slope 1. When all the parties reach \(((1 + 2h(q))/2, (1 + 2h(q))/2, h(q))\), they attain omniscience.

In Example 2, when the mode of the protocol changes, all the communicating parties attain omniscience and start behaving as one. In fact, the recursive structure of our protocol holds even when only a subset of communicating parties attains omniscience, as our final example with \(m = 4\) illustrates.

\(^5\) Rate-slope refers to the derivative of the communication rate as a function of blocklength (or time).

\(^6\) Note that \(\mathbb{H}_\sigma(P_{x_M})\) corresponds to \(\mathbb{H}\) for the finest partition with each part of \(\sigma\) treated as one party.

\(^7\) One crucial observation that drives our protocol is that \(R_1^* - R_j^* = H(X_1) - H(X_j)\), which holds even for a general \(m\).
Example 3. Let $W_1, W_2, W_3 \sim \text{Ber}(1/2)$ and $V_1, V_2 \sim \text{Ber}(q)$ for some $0 < q < 1/2$, and let $X_1 = (W_1, W_2)$, $X_2 = (W_1 \oplus V_1, W_2)$, $X_3 = W_2 \oplus V_2$, and $X_4 = W_3$. Note that the observations of subsequence $\{1, 2, 3\}$ are exactly as in Example 2. In this case, $\mathbb{H}_{123|4}) = 3 + 2h(q)$ is dominant, and our protocol proceeds as in Figure 3. At $t_1$, parties 1 and 2 attain local omniscience and change the slopes of $R_1$ and $R_2$ to 1/2. At $t_2$, parties 3 and 4 start communicating. At $t_3$, parties in $\{1, 2, 3\}$ attain local omniscience and change their slope to 1/3. Note that up to $t_3$, the evolution of $(R_1, R_2, R_3)$ is exactly as in Example 2. Also, at $t_3$ the rate difference $(R_1 + R_2 + R_3 − R_4)$ equals $H(X_1, X_2, X_3) − H(X_4) = 1 + 2h(q)$. Thus, after $t_3$ the rate pair $(R_1 + R_2 + R_3, R_4)$ behaves as if the parties in $\{1, 2, 3\}$ were collocated to begin with. Finally, all parties attain omniscience at $t_4$.

C. Full description of the protocol and its performance

Moving now to the real world, rates must be increased in discrete increments and a positive decoding error probability must be tolerated. To that end, the parties incrementally transmit independent hash bits, $n\Delta$ at a time. The ideal decoder of the previous section is replaced with a typical decoder DEC$(j, \sigma, R)$ which searches for the maximal set $A$ such that there exists a unique sequence $x_A$ that contains the current rate vector in its CO region and is consistent with the local observation and the received hash values. In fact, instead of working with the original CO region $\mathcal{R}_0$, we use the more restrictive region $\mathcal{R}_0^\Delta(A)$ consisting of vectors $(R_i, i \in A)$ such that

$$R_B \geq H(X_B \mid X_{A \setminus B}) + |B|\Delta, \quad \forall B \subseteq A.$$  

The complete decoder is described in Protocol 2.

Note that the decoder declares (ACK, $A$) if it can find a unique maximal set $A$ and a unique sequence $x_A$, declares NACK if it finds no such set or sequence, or an ERR otherwise. In fact, an error may occur even when it is not detected, i.e., when ERR is not transmitted. However, we can identify an event $\mathcal{E}$ of small probability such that under $\mathcal{E}$ the real decoder DEC behaves exactly like DEC$_{1d}$, but with $\mathcal{R}_0(A)$ replaced with $\mathcal{R}_0^\Delta(A)$. Therefore, omniscience can be achieved in a similar manner as the ideal protocol of the previous section.

Protocol 2: DEC$(j, \sigma, R)$

**Input:** An index $1 \leq j \leq m$, a partition $\sigma \in \Sigma(M)$, a rate vector $R = (R_1, \ldots, R_m)$.  

**Output:** A NACK message, an ACK message (ACK, $A$), or an error message ERR.  

1) For $\sigma_i$ such that $j \neq \sigma_i$, find the maximal set $A \subseteq M$ such that $\sigma_i \not\subseteq A$ and there exists a joint type $P_{X_A}$ for $X_A$ satisfying the following:

(i) $P_{X_A} = P_X$, 

(ii) $(R_i, i \in A) \in \mathcal{R}_{\sigma_i}^{\Delta}(A \mid P_{X_A})$, and 

(iii) there is a unique $x_A$ of joint type $P_{X_A}$ such that the hashes of $x_A$ match all the previously received hashes from parties in $A \setminus \{j\}$.

2) if there is a unique maximal $A$ found in Step 1 then

| return (ACK, $A$).

else if there is no sequence found in Step 1(iii) then

| return NACK.

else

| Return ERR.

The main component of our universal protocol is the one step omniscience protocol OMN described in Protocol 3, which uses DEC for decoding. Protocol OMN proceeds very much like the ideal protocol except with a modified balanced difference condition: A new party $i$ starts communicating when $R_1 \geq H(P_{X_i}) + \alpha \Delta$, where $\alpha \in \mathbb{N}$ is an increasing threshold parameter which is updated as the protocol proceeds. Throughout the protocol, a rate $R_i = -1$ indicates that the $i$th party is not yet transmitting and only the parties with $R_i \geq 0$ communicate. The decoder tries to attain omniscience only among the communicating parties.

The ideal protocol of the previous section works due to its recursive structure whereby when a subset $A$ attains local omniscience, the rate vector appears as if the parties in $A$ have been collocated from the start. Moreover, the first subset to attain local omniscience does so by using a communication of rate $\mathbb{H}_{\sigma_i}(A)$. The OMN protocol almost has both these features and conserves a certain “validity” property of rate vectors, which we define next.

Definition 2. For $\alpha \in \mathbb{N}$, $\sigma \in \Sigma(M)$ with $|\sigma| = k$ and $H = (H_{\sigma_1}, \ldots, H_{\sigma_k})$, a rate vector $(R_1, \ldots, R_m)$ is $(\sigma, H, \alpha)$-valid if, for $s = \max\{i : R_{\sigma_i} \geq 0\}$, the following conditions hold:

(i) For $1 \leq i, j \leq s$, $R_{\sigma_i} - R_{\sigma_j} \leq H_{\sigma_i} - H_{\sigma_j} + \alpha \Delta$;

(ii) $R_{\sigma_1} < H_{\sigma_1} - H_{\sigma_{s+1}} + \alpha \Delta$;

(iii) $\forall 1 \leq i \leq k$ s.t. $|\sigma_i| \geq 2$, $(R_j, j \in \sigma_i) \in \mathcal{R}_{\sigma_i}^{\Delta}(\sigma_i)$;

(iv) for all $A \subseteq \{1, \ldots, k\}$ with $|A| \geq 2$, $(R_j : j \in \sigma_i, i \in A) \not\in \mathcal{R}_{\sigma_i}^{\Delta}\left(\bigcup_{i \in A} \sigma_i\right)$.

Theorem 2. For $\alpha \in \mathbb{N}$, $\sigma \in \Sigma(M)$ with $|\sigma| = k$ and $H = (H_{\sigma_1}, \ldots, H_{\sigma_k})$ with $H_{\sigma_1} \geq H_{\sigma_2} \geq \ldots, H_{\sigma_k}$, let $R_{\text{in}} = (R_{\text{in}, \sigma_1}, \ldots, R_{\text{in}, \sigma_k})$ be $(\sigma, H, \alpha)$-valid. Then, if OMN$(\sigma, \alpha, H, R_{\text{in}})$ is executed and error $\mathcal{E}$ does not occur, the final rates $R_{\text{OUT}}$ and the omniscience family $\mathcal{O}$ satisfy the following:

1) Every $A \in \mathcal{O}$ comprises parts of $\sigma$ and the sum-rate
Protocol 3: OMN($\sigma, \alpha, H, R$)

Input: A partition $\sigma \in \Sigma(M)$ with $|\sigma| = k$, an $\alpha \in \mathbb{N}$, an entropy estimate vector $H = (H_{\sigma_i} : 1 \leq i \leq k)$, a rate vector $R = (R_1, \ldots, R_m)$; we assume that $H$ is sorted, i.e., $H_{\sigma_1} \geq H_{\sigma_2} \geq \cdots \geq H_{\sigma_k}$.

Output: A rate vector $R^{\text{out}}$, a family of subsets $O$ that have attained omniscience.

1) Initialize $s := \max\{ j : R_j \geq 0 \}$.
2) All parties $j$ such that $j \in \sigma_i$ for some $1 \leq i \leq k$ send $\lceil n \Delta / |\sigma| \rceil$ random hash bits. Update $R_j \rightarrow R_j + \Delta / |\sigma|$. 
3) If there exists $i > s$ such that $R_{\sigma_i} \geq H_{\sigma_i} - H_{\sigma_i} + \Delta$ then set $R_j = 0$ for all $j \in \sigma_i$, and set $s = \max\{ j : R_j \geq 0 \}$.
4) For all $j$ such that $j \in \sigma_i$ for some $1 \leq i \leq s$, execute $\text{DEC}(j, \sigma, R)$, which outputs $\text{ACK}(j, \sigma, A_j)$ or $\text{ERR}$. 
5) If all parties send a $\text{NACK}$ then return to Step 2.
   - else if No party declares an $\text{ERR}$ and some parties declare an $\text{ACK}$, then
     - Identify the omniscience family $O = \{ B \subset M : \text{all } j \in B \text{ returned } (\text{ACK}, B) \}$, if $O$ is nonempty then
       - Set $R^{\text{out}} = R$, and return $(R, O)$.
     - else
       - Declare an error.
   - else
     - Declare an error.

$R^{\text{out}}$ satisfies

$$R^{\text{out}}_A - \mathbb{H}_{\sigma_{[1]} \cdots \sigma_{[s]}}(P_{X_A}) \leq c_m \Delta,$$

where $c_m$ is a constant depending only on $m$.

2) Let $\sigma^{\text{out}} \in \Sigma(M)$ be the partition obtained by combining the parts in $\sigma$ that belong to the same $A$ in $O$. Let $H^{\text{out}}_{\sigma^{\text{out}}}$ denote the entropy of the type of $X_{\sigma^{\text{out}}}$. Then, with $H^{\text{out}} = (H^{\text{out}}_{\sigma^{\text{out}}}, 1 \leq i \leq |\sigma^{\text{out}}|)$, $R^{\text{out}}$ is $(\sigma^{\text{out}}, H^{\text{out}}, c'_m \alpha)$-valid, where $c'_m$ is a constant depending only on $m$.

We are now in a position to describe our universal protocol for omniscience. We begin by calling OMN with $\sigma = \sigma_f(M)$, $\alpha = 1$, the sorted entropy estimates $H$ computed from marginal empirical distributions $P_{X_A}$, and the rate vector $R = (0, -1, \ldots, -1)$ indicating that party 1 starts communicating and everyone else remains quiet. Note that $R$ is $(\sigma, H, 1)$-valid. A new party $i$ starts communicating when $R_i \geq H_i - H_i + \Delta$. If no error occurs, OMN will terminate when a subset $A$ attains omniscience. In view of Theorem 2, at this point $R_A$ should be close to $\mathbb{H}_{\sigma_f(A)}(P_{X_A})$ and the rates will be $(\sigma^{\text{out}}, H^{\text{out}}, c'_m \alpha)$-valid. Thus, we are in a similar situation as the first call to OMN except that $\alpha$ must be replaced by $c'_m \alpha$ and the parties in a single part of $\sigma^{\text{out}}$ are behaving as if they are collocated. The protocol proceeds by calling OMN again with these updated parameters. Note that under $E^c$, any party $j \in A$ for $A \in O$ can correctly compute $P_{X_A}$ and transmit it using $O(\log n)$ bits. Proceeding recursively in this manner, the protocol stops when parties in $M$ attain omniscience, which by Theorem 2 can only happen when the sum-rate $R_M$ is close to $\mathbb{H}_{\sigma}(P_{X_M})$ for some partition $\sigma$ of $M$. Thus, omniscience will be attained in communication of rate roughly less than $R_{\text{CO}}(M|P_{X_M})$. We formally state our overall protocol in Protocol 4.

Protocol 4: A universal omniscience protocol

1) Initialize $\sigma = \sigma_f(M)$, $R = (0, -1, -1, \ldots, -1)$, $k = |\sigma|$, $\alpha = 1$.
2) While $k > 1$ do
   - (i) For $1 \leq i \leq k$, a party $j \in \sigma_i$ computes $P_{X_{\sigma_i}}$ and broadcasts it. Each party computes $H_{\sigma_i} = H(P_{X_{\sigma_i}}), 1 \leq i \leq k$.
   - (ii) Let $H$ be the sorted version of $(H_{\sigma_1}, 1 \leq i \leq k)$, i.e., assume $H_{\sigma_1} \geq H_{\sigma_2} \geq \cdots \geq H_{\sigma_k}$. Call OMN$(\sigma, \alpha, H, R)$.
   - if There is no error declared then
     - let $(R^{\text{out}}, O)$ be its output.
   - else
     - Terminate.
   - (iii) Let $\sigma^{\text{out}} = \{ \sigma_i : \sigma_i \in s.t. \sigma_i \not\subset A \forall A \in O \} \cup \{ A : A \in O \}$.
     - Update $R = R^{\text{out}}$, $\sigma = \sigma^{\text{out}}$, $k = |\sigma^{\text{out}}|$, and $\alpha \rightarrow c'_m \alpha$.

We close with the following result claiming the universal rate optimality of our protocol for every IID distribution. Proof is a simple consequence of Theorem 2. Note that while Protocol 4 is a variable length protocol, its fixed length variant can be obtained simply by aborting the protocol once the total number of bits communicated crosses $nR$.

Theorem 3. For $\Delta = \frac{1}{\sqrt{n}}$ and every distribution $P_{X_M}$, Protocol 4 has average length $|\pi|_{|\sigma|}$ less than $nR_{\text{CO}}(M) + O(\sqrt{n \log n})$ and probability of error vanishing to 0 in the limit $n \rightarrow \infty$. Furthermore, for a fixed $R > 0$, the fixed length variant of Protocol 4 has probability of error vanishing to 0 as $n \rightarrow \infty$ for all distributions $P_{X_M}$ with $R > R_{\text{CO}}(M) + O(\sqrt{n^{-1} \log n})$.

References