Decentralized Detection of Nomadic Transmitter via Helping Agents

Amichai Sanderovich*,Shlomo Shamai (Shitz)*,Yossef Steinberg*and Gerhard Kramer†
*Technion, Haifa, Israel
†Bell Laboratories Lucent Technologies, NJ

***Abstract*— The problem of a wireless terminal sending information to a remote destination via agents with reliable connections is investigated. Such a setting typifies, nomadic users communicating with access points to a wireless network, where each access point (agent) is equipped with a prescribed reliable connection bandwidth. The agents are assumed to be ignorant of the code-book employed by the nomadic user, and hence they are not able to decode messages. We focus here on a decentralized quantization based approach, and provide outer and inner bounds for the reliable transmission rate when the channel between the user to the agents is a general broadcast channel. For a Gaussian channel model, the best reliable rate is determined when the transmitter uses typical Gaussian codewords.**

I. INTRODUCTION

Information theory view of networks and especially wireless networks is in the focus of an extensive research activity. This interest is partly due to many recent results about the multiple antenna channel, which demonstrate significant improvements, especially for the fading channels.

Many papers propose and analyze ad-hoc wireless network in information theoretic terms. Among these, reported comparisons between relaying and dirty paper coding are [2] and coding schemes which achieve $O(n)$ transport capacity [1]. The relaying technique, or as is sometimes called multi-hop, makes use of several intermediate wireless nodes to assist the communication between two nodes. An information theoretic framework for the relay channel was given by El Gamal and Cover in [3] for a single relay node and extended by [4] to several relaying nodes. Relaying can be coarsely divided into compress-and-forward and decode-and-forward, depending on whether the relays decode the transmitted message or just forward the processed received signal to the destination. Relaying schemes can take advantage of their common knowledge for the sake of forwarding to the final destination. This cooperation is commonly used, for example in [5],[6]. Cooperation between receiving nodes in a degraded broadcast channel is described in [7]. We conclude that an upper bound derived by [8] suggests that as the number of users in an ad-hoc network is going to infinity, the total rate per user tends to zero. This bound motivates the use of networks that are not solely ad-hoc, but are composed of base stations or access points as well.

The problems of conveying a source which is observed by remote agents to a single destination are built around similar settings, where the source is modelled as an i.i.d. random variable. Different aspects of these problems are analyzed in information theoretic frameworks such as distributed source coding, CEO [9] and sensor networks. The setting of wireless network with base stations and/or access points is closely related to these problems, as is evident in this presentation. A small list of papers that are relevant to this end are [10],[11] for distributed reconstruction of the sources. Distributed lossy reconstruction of sources, as opposed to central processing [12] is still essentially unsolved. The Gaussian CEO problem [13],[14],[15] which was recently solved by the entropy power inequality. Multi-terminal lattice approaches are described in [16]. Relating these rate-distortion problems to the network scenario is the subset of [17],[18],[19]. The use of other measures instead of the distortion is addressed in [20]. The dissertation of Schein [21] focuses on the characteristics of the problem of communicating via two agents, and several achievable rates are demonstrated.

Here we consider the problem of reliable communication of a nomadic transmitter through non-decoding agents which are connected via lossless links to the final destination. The agents use a distorted version (via their respective channels) of the transmitted message, and are able to transmit a predetermined number of bits to the destination without any errors. The destination is reached only via the agents which serve as access points. The rest of the paper is organized as follows: in section II the setting of the problem is given. An achievable rate and an upper bound are presented in sections III and IV respectively. The Gaussian channel is presented as an example in section V which also includes a complete characterization of the rate-region for the case where the agents are unaware of the code used. We use capital letters for random variables, capital letter with subscript X_i denotes the i-th element in a random vector and capital letter with superscript X^n denotes the vector $(X_1,\ldots,X_n)$. When using the notation X_k^m it refers to the vector $(X_k,\ldots,X_m)$. A calligraphic letter $\mathcal{X}$ denotes the signal space of the random variable X or a set $\mathcal{T}$.

II. PROBLEM SETTINGS

We consider the problem of a single transmission through T agents, playing the role of decentralized processors, as is seen on figure 1. For the purpose of stating a converse to the case where the agents do not know the codebook, we artificially introduce random code for the transmitter S. Such random coding is also used in [22] for a miss-match scenario, while the advantages of random codebooks were demonstrated in [23].

The following properties hold (unless clearly stated otherwise) for the scheme described in this presentation:

1) Define $\mathcal{C}$ as the ensemble of all $N_{\mathcal{C}} = |\mathcal{X}|^{k2^{kR}}$ codebooks with rate R, codeword length of k and input channel alphabet of $\mathcal{X}$. Let F denote a key which is an index into a code from $\mathcal{C}$, so $1 \le F \le N_{\mathcal{C}}$.
2) In an initialization stage, the transmitter S randomly selects the key F. It then sends to the channel the signal
$$X^k = \phi_{S,F}(M) : I_{2^{Rk}} \to \mathcal{X}^k \quad (1)$$
during k channel uses, where $\phi_{S,F}$ represents the coding with code F and M is the message to be sent $M = [1, 2^{kR}]$. The key F is chosen from $N_{\mathcal{C}}$ according to the probability
$$P_F(F) = \prod_{M=1}^{2^{kR}} P_{X^k}(\phi_{S,F}(M)) \quad (2)$$
and $P_{X^k}(X^k) = \prod_{i=1}^{k} P_X(X_i)$, for some single letter probability $P_X(X)$.
3) T agents $A1, \ldots, AT$, receive the k outputs of a memoryless broadcast channel
$$P_{Y_1^k,\ldots,Y_T^k|X^k}(Y_1^k, \ldots, Y_T^k|X^k) = \prod_{i=1}^{k} P_{Y_1,\ldots,Y_T|X}(Y_{1,i}, \ldots, Y_{T,i}|X_i), \quad (3)$$
where $Y_t \in \mathcal{Y}_t$.
4) The agents are not informed about the the key F but know $\mathcal{C}$ and therefore $P_X(X)$. All agents encode every $n \le k$ channel outputs (where $m = k/n$ is an integer) with T encoding functions:
$$0 < t \le T,\ 0 < j \le m : \quad V_{t,j} = \phi_{At}(Y_{t,(j-1)n}^{jn}) : \mathcal{Y}_t^n \to I_{C_t n}. \quad (4)$$
Where C_t is the capacity in bits per channel use of a lossless link which connects the final destination D to the agent At.
5) The final destination D knows all the encoding functions in the system, the code ensemble $\mathcal{C}$ and most important, the chosen key F. Denote $\mathcal{T} \triangleq \{1, \ldots, T\}$. So D can decode the message M from the set $V_{\mathcal{T}}^m$ of $m = k/n$ length T vectors $V_{\mathcal{T}}^m \triangleq (V_{\mathcal{T},1}, \ldots, V_{\mathcal{T},m})$ which are sent to the destination from the T agents:
$$\hat{M} = \phi_{D,F}(V_{\mathcal{T}}^m) : I_{2^{m\sum nC_t}} \to I_{2^{kR}}. \quad (5)$$

Notice that with the knowledge of F, X^k is uniformly distributed over 2^{kR} codewords, while without the key, X^k is distributed according to $\prod_{i=1}^{k} P_X(X_i)$. We use the two following simple lemmas in the sequel:

Lemma 1: Without the key F, the received vector X^k is distributed according to $P_{X^k}(X^k) = \prod_{i=1}^{k} P_X(X_i)$, and due to the memoryless property, Y_t^k are also distributed as $P_{Y_t^k}(Y_t^k) = \prod_{i=1}^{k} \sum_X P_{Y_t|X}(Y_{t,i}|X)P_X(X)$. This lemma is proved in appendix I.

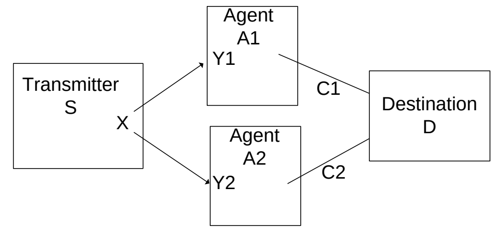

Fig. 1. Scheme of a system with two agents between the transmitter and the destination

Lemma 2: The chosen code book F is identical to a standard single codebook which is constructed by randomly and independently selecting codewords according to the probability law $P_{X^k}(X^k) = \prod_{i=1}^{k} P_X(X_i)$.

Proof: It is easy to see from the probability laws and the fact that all possible codebooks are considered, that the resulting codebooks are chosen with the same probability. ■

These settings correspond to the situation where the final destination decodes the message from the transmitter via simple agents who are not able to decode the transmitted message and use short compression schemes of the received signals. This enables to relate to the problem with two separated problems: first we would like to build agents that will convey the received signals to the destination so the reliable transmission rate is maximized, and secondly we would like to design a transmitter that will maximize the total rate.

III. An achievable rate

The following theorem is proven in appendix III.

Theorem 1: if the codebook ensemble $\mathcal{C}$ contains codes within the rate:
$$R < \max I(X; U_{\mathcal{T}}) \quad (6)$$
under the constraints
$$\forall \mathcal{S} \subseteq \mathcal{T} : \sum_{t \in \mathcal{S}} C_t > I(U_{\mathcal{S}}; Y_{\mathcal{S}}|U_{\mathcal{S}^C}) \quad (7)$$
the transmitted message can be decoded correctly in the destination using a suitable encoder, decoder and agent encoders. The maximization in (6) is over $P_{X,U_{\mathcal{T}},Y_{\mathcal{T}}}(X, U_{\mathcal{T}}, Y_{\mathcal{T}})$ such that:
$$P_{X,U_{\mathcal{T}},Y_{\mathcal{T}}}(X, U_{\mathcal{T}}, Y_{\mathcal{T}}) = P_X(X)P_{Y_{\mathcal{T}}|X}(Y_{\mathcal{T}}|X) \prod_{t=1}^{T} P_{U_t|Y_t}(U_t|Y_t). \quad (8)$$
The following Markov relations hold as a consequence of (8):
$$U_t - Y_t - \{X, U_{\mathcal{T}\setminus t}, Y_{\mathcal{T}\setminus t}\}. \quad (9)$$

Since this achievable region is attained through the use of compression which is independent of the message index and the codebook used by the transmitter, the proof is valid for agents which are ignorant of the code used by the nomadic transmitter.

IV. An outer bound

We start by stating the maximum rate R so error free decoding is possible at the destination, when the agents' encoding functions are given. Using Fano's inequality, an error free decoding at the destination is possible only if:

$$H(M|V_{\mathcal{T}}^m, F) \leq k\epsilon_k, \tag{10}$$

where $k\epsilon_k \to 0$ as $k \to \infty$.
Define $\mathbf{X}_j \triangleq (X_{(j-1)n+1}, \ldots, X_{jn})$, now we have:

$$\begin{aligned}
kR &= H(M) = I(M; V_{\mathcal{T}}^m, F) + H(M|V_{\mathcal{T}}^m, F) && (11)\\
&\leq H(V_{\mathcal{T}}^m) + H(F|V_{\mathcal{T}}^m) - H(F|M) + \\
&\quad -H(V_{\mathcal{T}}^m|M, F) + k\epsilon_k && (12)\\
&= I(V_{\mathcal{T}}^m; M, F) - I(F; V_{\mathcal{T}}^m) + k\epsilon_k && (13)\\
&\leq I(V_{\mathcal{T}}^m; M, F) + k\epsilon_k && (14)\\
&= I(\mathbf{X}^k(M, F); V_{\mathcal{T}}^m) + k\epsilon_k && (15)\\
&\leq \sum_{j=1}^m H(V_{\mathcal{T},j}) - H(V_{\mathcal{T},j}|\mathbf{X}_j) + n\epsilon_k && (16)\\
&\leq m \max_{P(X)} I(V_{\mathcal{T}}; X^n) + k\epsilon_k, && (17)
\end{aligned}$$

where (13) is since F is independent with M so $H(F|M) = H(F)$ and (16) is due to properties 3 and 4. From (17) we conclude that the transmission rate R is upper bounded by

$$R \leq \max_{P(X)} \frac{1}{n} I(V_{\mathcal{T}}; X^n). \tag{18}$$

Turning to the agents, we would like to upper bound the maximization of

$$\frac{1}{n} I(V_{\mathcal{T}}; X^n), \tag{19}$$

over the encoding functions of the agents without the key F. By defining $U_{\mathcal{S},i} \triangleq (V_{\mathcal{S}}, Y_{\mathcal{T}}^{i-1}, X^{i-1})$ we get to the following theorem, which is proved in appendix II and by theorem 2 of [13]:

Theorem 2: The maximum achievable rate when the agents do not know the codebook key F is

$$R \leq \max I(X; U_{\mathcal{T}}), \tag{20}$$

and $U_{\mathcal{T}}$ must fulfill the constraints:

$$\forall \mathcal{S} \subseteq \mathcal{T}: \sum_{t \in \mathcal{S}} C_t \geq I(U_{\mathcal{S}}; Y_{\mathcal{T}}|U_{\mathcal{S}^C}). \tag{21}$$

and considering lemma 1, also fulfill the following Markov relations:

$$U_t - Y_t - \{X, Y_{\mathcal{T}\setminus t}\}. \tag{22}$$

The difference between theorems 1 and 2 besides the difference of (7) and (21), is that for the former

$$P_{U_t|Y_{\mathcal{T}}, U_{\mathcal{T}\setminus t}}(U_t|Y_{\mathcal{T}}, U_{\mathcal{T}\setminus t}) = P_{U_t|Y_t}(U_t|Y_t) \tag{23}$$

and for the latter

$$P_{U_t|Y_{\mathcal{T}}, U_{\mathcal{T}\setminus t}}(U_t|Y_{\mathcal{T}}, U_{\mathcal{T}\setminus t}) = P_{U_t|Y_t, U_{\mathcal{T}\setminus t}}(U_t|Y_t, U_{\mathcal{T}\setminus t}). \tag{24}$$

V. The Gaussian channel

In this section we explore the Gaussian channel. Using the latest results on the Gaussian CEO rate-distortion problem [15], a converse for the maximum achievable rate is shown for the former case.
We use the results of section III with continuous alphabets, where the extension relies on standard arguments.

A. The capacity

For the Gaussian channel, assume that $Y_t = X + n_t$ where X, n_t are independent Gaussian random variables with $\mathrm{E}X^2 = P_X$ and $\mathrm{E}n_t^2 = P_{n_t}$. As specified since we deal with non-decoding agents, with "typical" codebooks, X is i.i.d. according to $P_X(X)$.

Theorem 3: The capacity for this case is

$$R \leq \max_{r_t \geq 0} \min_{\mathcal{S} \subseteq \mathcal{T}} \sum_{t \in \mathcal{S}} C_t - r_t + \frac{1}{2} \log_2 \left(1 + P_X \sum_{t \in \mathcal{S}^C} \frac{1 - 2^{-2r_t}}{P_{n_t}} \right). \tag{25}$$

This is proved by showing that the region of minimum required links $\underline{C} \triangleq (C_1, \ldots, C_T)$ which are sufficient for communication with rate R is equal to the region of minimum links $\overline{C} \triangleq (C_1, \ldots, C_T)$ which are necessary for communicating with rate R. An adaptation of [15] to a communication problem instead of the quadratic distortion is used to this end. The altered proof is briefly scatched here.
First, the sum-rate under the constraints

$$\sum_{t=1}^m C_t \leq a_m, \; m = 1, \ldots, T-1 \tag{26}$$

is shown to be

$$C_{\min}(a_1, \ldots, a_{T-1}) = R + \min_{r_t \geq 0} \sum_{t=1}^T r_t \tag{27}$$

where the minimum is over the space:

$$\begin{cases}
\forall\ 1 \leq t \leq T: \; r_t \geq 0 \\
\forall\ 1 \leq m \leq T-1: \\
\sum_{t=1}^m r_t + R - \frac{1}{2} \log_2 \left(1 + P_X \sum_{t=m+1}^T \frac{1-2^{-2r_t}}{P_{n_t}} \right) \leq a_m \\
R = \frac{1}{2} \log_2 \left(1 + P_X \sum_{t=1}^T \frac{1-2^{-2r_t}}{P_{n_t}} \right).
\end{cases} \tag{28}$$

The direct part of this sum rate is subtantiated by the contra-polymatroid form of the achievable region which is evident in [15] or [13]. The converse part is shown in appendix IV using [15] via the entropy power inequality.
Next it is shown that the above sum-rate identity leads to identical rate-regions ($\underline{C}(R) \leq \overline{C}(R)$). This is since both regions are convex [15] and since for all non-negative vectors $(\alpha_1, \ldots, \alpha_T)$,

$$\min_{(C_1, \ldots, C_T) \in \overline{C}(R)} \sum_{t=1}^T \alpha_t C_t \geq \min_{(C_1, \ldots, C_T) \in \underline{C}(R)} \sum_{t=1}^T \alpha_t C_t. \tag{29}$$

Identity (29) is proved in [15] as follows. Assume that $\alpha_1 \geq \cdots \geq \alpha_T$, then

$$\begin{aligned}
&\min_{(C_1,\ldots,C_T)\in\overline{C}(R)} \sum_{t=1}^{T} \alpha_t C_t \\
&\geq \min_{(C_1,\ldots,C_{T-1})\in\overline{C}(R)} \alpha_T C_{\min}\left(C_1,\ldots,\sum_{t=1}^{T-1} C_t\right) \\
&\qquad\qquad + \sum_{m=1}^{T-1}(\alpha_m-\alpha_{m+1})\sum_{t=1}^{m} C_t \\
&\geq \min_{(r_1,\ldots,r_T)} \alpha_T(R+\sum_{t=1}^{T} r_t) + \sum_{m=1}^{T-1}(\alpha_m-\alpha_{m+1})\times \\
&\quad \left(\sum_{t=1}^{m} r_t + R - \frac{1}{2}\log_2\left(1+P_X\sum_{t=m+1}^{T}\frac{1-2^{-2r_t}}{P_{n_t}}\right)\right), \qquad (30)
\end{aligned}$$

where $\{r_t\}$ must satisfy the last equality of (28). Now the minimizers $\{r_t\}$ can be used in the direct part of (27), completing the argument.

B. *Example of two equivalent agents*

Next we consider the case where there are only two users with $P_{n_1} = P_{n_2}$ and with $C_1 = C_2 = C$. Recall that a simple uppr bound for this case can be derived from the cut-set bound, which appears in [24]. This bound for our case reduces to:

$$R \leq \min\left[\frac{1}{2}\log_2\left(1+2\frac{P_X}{P_n}\right), 2C\right]. \qquad (31)$$

We denote the term $\frac{1}{2}\log_2\left(1+2\frac{P_X}{P_n}\right)$ in (31) as the MIMO upper bound, since it is identical to the maximum achievable rate when $C \to \infty$.

The achievable rate of (6) can be calculated by maximizing R over R, r_1, r_2 such that following set of inequalities is valid:

$$\begin{cases}
R \leq C - r_1 + \frac{1}{2}\log_2\left(1+P_X\frac{1-2^{-2r_2}}{P_n}\right) \\
R \leq C - r_2 + \frac{1}{2}\log_2\left(1+P_X\frac{1-2^{-2r_1}}{P_n}\right) \\
R \leq 2C - r_1 - r_2 \\
R \leq \frac{1}{2}\log_2\left(1+P_X\frac{2-2^{-2r_1}-2^{-2r_2}}{P_n}\right)
\end{cases} \qquad (32)$$

Solving for the case where the two lower inequalities are active results with the rate of:

$$R = \frac{1}{2}\log_2\left(1+2S\left(1-\frac{\sqrt{S^2+2^{4C}(1+2S)}-S}{2^{4C}}\right)\right). \qquad (33)$$

where $S = \frac{P_X}{P_n}$. This solution fulfilles the inequalities (32) and thus is the capacity for the problem. Figure 2 demonstrates this achievable rate for several C values as a function of the signal to noise ratio S. It is noticed that for the lower values of S, the achievable rate is near optimal for a system with fixed users, and that at most 0.7 bit per channel use is lost to maintain the scheme robustness.

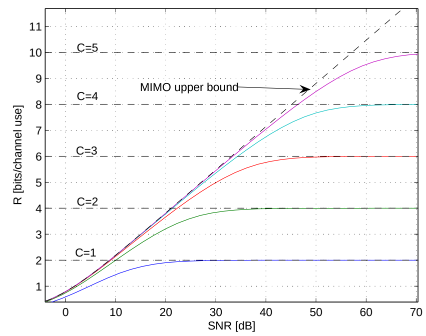

Fig. 2. The achievable rates of systems with two equivalent agents versus the signal to noise ratio in dB where 5 different values of C were considered. The dashed lines represent the cut-set bounds where the flat lines are $2C$ and the MIMO upper bound is the upper-left dashed line.

VI. Conclusion

The case of communication of nomadic transmitter via separated non-cooperative agents is presented. The agents are assumed to be ignorant about the codebook used, and the scheme is robust of the codebooks used by the transmitter. A direct coding theorem based on decentralized quantization and an outer bound are presented. Considering the Gaussian channel, a converse is proved by the entropy power inequality.

Acknowledgment

The research was supported by the EU 6th framework program via the NEWCOM network of excellence. The authors are indebt to Mr. Michael Peleg for helpful discussions.

Appendix I
Proof of Lemma 1

In this section we show that when the key F in unknown the transmission X^k is a memoryless random process, distributed according to: $P_{X_{t_1},\ldots,X_{t_L}}(X_{t_1},\ldots,X_{t_L}) = \prod_{t=t_1}^{t_L} P_X(X_t)$ for all $t_{\mathcal{L}} \triangleq t_1 < \cdots < t_L$ and all $X_{t_{\mathcal{L}}} \triangleq X_{t_1},\ldots,X_{t_L}$. We have that

$$P_{X_{t_1},\ldots,X_{t_L}}(X_{t_1},\ldots,X_{t_L}) = \sum_{F,M:(\phi_{S,F}(M))_{t_{\mathcal{L}}}=X_{t_{\mathcal{L}}}} P_{F,M}(F,M). \qquad (34)$$

Now since we consider all possible codebooks, we consider also all the possible $2^{kR}!$ permutations of every codebook, so the above is equal to:

$$\begin{aligned}
&P_{X_{t_1},\ldots,X_{t_L}}(X_{t_1},\ldots,X_{t_L}) = \\
&\sum_{M=1}^{2^{kR}} \sum_{F:(\phi_{S,F}(M))_{t_{\mathcal{L}}}=X_{t_{\mathcal{L}}}} P_{F|M}(F|M)P_M(M) = \\
&\qquad \sum_{F:(\phi_{S,F}(M=1))_{t_{\mathcal{L}}}=X_{t_{\mathcal{L}}}} P_{F|M}(F|M=1). \qquad (35)
\end{aligned}$$

Now we calculate $P(F|M=1)$ and sum over all the possible $|\mathcal{X}|^{(2^{kR}-1)k+k-L}$ codebooks F, that is over all possible $X_{t_{\mathcal{L}^C}}, X^k(M=2), \ldots, X^k(M=2^{kR})$. We get that:

$$P_{X_{t_{\mathcal{L}}}}(X_{t_{\mathcal{L}}}) = \prod_{t=t_1}^{t_L} P_X(X_t) \sum_{X_{t_{\mathcal{L}^C}}, X^k(M=2),\ldots,X^k(M=2^{kR})}$$

$$\prod_{j\in t_{\mathcal{L}^C}} P_X(X_j) \prod_{i=[1,k2^{kR}-k])} P_X(X_i) = \prod_{t=t_1}^{t_L} P_X(X_t). \quad (36)$$

Concluding the proof.

Appendix II
Proof of the outer bound of Theorem (2)

Turning to the agents, we would like to upper bound the maximization of

$$\frac{1}{n} I(V_{\mathcal{T}}; X^n), \quad (37)$$

over the encoding functions of the agents, when the agent do not know F:

$$\begin{aligned}
I(V_{\mathcal{T}}; X^n) &= H(X^n) - H(X^n|V_{\mathcal{T}}) && (38)\\
&= \sum_{i=1}^n H(X_i) - H(X_i|V_{\mathcal{T}}, X^{i-1}) && (39)\\
&\leq \sum_{i=1}^n H(X_i) - H(X_i|V_{\mathcal{T}}, Y_{\mathcal{T}}^{i-1}, X^{i-1}) && (40)\\
&= \sum_{i=1}^n H(X_i) - H(X_i|U_{\mathcal{T},i}) && (41)\\
&\leq n \max I(X; U_{\mathcal{T}}). && (42)
\end{aligned}$$

When $U_{\mathcal{S}} \triangleq (V_{\mathcal{S}}, Y_{\mathcal{T}}^{i-1}, X^{i-1})$ for any $\mathcal{S} \subseteq \mathcal{T}$.

The random variables $U_{\mathcal{S}}$ must also fulfill the following constraints:

$$\begin{aligned}
& \sum_{i=1}^n I(U_{\mathcal{S},i}; Y_{\mathcal{T},i}|U_{\mathcal{S}^C,i}) && (43)\\
= & \sum_{i=1}^n I(V_{\mathcal{S}}, Y_{\mathcal{T}}^{i-1}, X^{i-1}; Y_{\mathcal{T},i}|V_{\mathcal{S}^C}, Y_{\mathcal{T}}^{i-1}, X^{i-1}) && (44)\\
= & \sum_{i=1}^n I(V_{\mathcal{S}}; Y_{\mathcal{T},i}|V_{\mathcal{S}^C}, Y_{\mathcal{T}}^{i-1}, X^{i-1}) && (45)\\
\leq & \sum_{i=1}^n I(V_{\mathcal{S}}; Y_{\mathcal{T},i}|V_{\mathcal{S}^C}, Y_{\mathcal{T}}^{i-1}, X^{i-1}) + && (46)\\
& I(V_{\mathcal{S}}; Y_{\mathcal{T},i}, X_i|V_{\mathcal{S}^C}, Y_{\mathcal{T}}^{i}, X^{i-1}) && (47)\\
= & \sum_{i=1}^n H(V_{\mathcal{S}}|V_{\mathcal{S}^C}, Y_{\mathcal{T}}^{i-1}, X^{i-1}) + && (48)\\
& -H(V_{\mathcal{S}}|V_{\mathcal{S}^C}, Y_{\mathcal{T}}^{i}, X^{i-1}) + H(V_{\mathcal{S}}|V_{\mathcal{S}^C}, Y_{\mathcal{T}}^{i}, X^{i-1}) + && (49)\\
& -H(V_{\mathcal{S}}|V_{\mathcal{S}^C}, Y_{\mathcal{T}}^{i}, X^{i}) && (50)\\
= & \sum_{i=1}^n I(V_{\mathcal{S}}; Y_{\mathcal{T},i}, X_i|V_{\mathcal{S}^C}, Y_{\mathcal{T}}^{i-1}, X^{i-1}) && (51)\\
= & \sum_{i=1}^n H(Y_{\mathcal{T},i}, X_i|V_{\mathcal{S}^C}, Y_{\mathcal{T}}^{i-1}, X^{i-1}) + && (52)\\
& -H(Y_{\mathcal{T},i}, X_i|V_{\mathcal{T}}, Y_{\mathcal{T}}^{i-1}, X^{i-1}) && (53)\\
= & H(Y_{\mathcal{T}}^n, X^n|V_{\mathcal{S}^C}) - H(Y_{\mathcal{T}}^n, X^n|V_{\mathcal{T}}) && (54)\\
= & H(V_{\mathcal{S}}|V_{\mathcal{S}^C}) - H(V_{\mathcal{S}}|V_{\mathcal{S}^C}, Y_{\mathcal{T}}^n, X^n) && (55)\\
= & H(V_{\mathcal{S}}|V_{\mathcal{S}^C}) \leq n \sum_{t\in\mathcal{S}} C_t, && (56)
\end{aligned}$$

Combining (56) and (42) and following the proof of *Lemma* 7 in [13] we prove theorem 2.

Notice the following equality which is due to Lemma 1, and the memoryless channel:

$$H(X_i, Y_{\mathcal{T}\setminus t,i}|Y_{t,i}) = H(X_i, Y_{\mathcal{T}\setminus t,i}|Y_{t,i}, Y_{t,i+1}^n, X^{i-1}, Y_{\mathcal{T}}^{i-1}). \quad (57)$$

This identity implies that:

$$\begin{aligned}
0 &= I(X_i, Y_{\mathcal{T}\setminus t,i}; Y_{t,i+1}^n, X^{i-1}, Y_{\mathcal{T}}^{i-1}|Y_{t,i})\\
&\geq I(V_t, X^{i-1}, Y_{\mathcal{T}}^{i-1}; X_i, Y_{\mathcal{T}\setminus t,i}|Y_{t,i}), \quad (58)
\end{aligned}$$

Due to data processing inequality. Thus $U_{\mathcal{T}}$ in theorem 2 also fulfill the Markov relations (22).

Appendix III
Proof for the achievable rate

For the proof we use ideas from [25], which presents an achievable rate region for compress and forward technique for the multiple relays problem. The difference being that the agents benefit from a fixed non-interfering links to the destination, thus multiple access communication and the interferences of simultaneously transmitting relays are avoided. Notice that the construction of the compression does not assume knowledge of the codebook at the agents. The network is composed of T agents $t \in \mathcal{T} = \{1,\ldots,T\}$, a source transmitter and a final destination. Apposed to [25], we do not need the *block Markov encoding* technique here.

The transmitter is sending $X^n(M)$ where $M \in [1, 2^{nR}]$. The t-agent compresses the received signal Y_t^n into z_t, where $z_t \in [1, 2^{n\hat{R}_t}]$. Since the compressed signals $\{z_t\}$ are dependent with each other and with $M_{\mathcal{T}}$, bandwidth from the agents to D can be saved by using the Slepian-Wolf (SW) lossless distributed source coding. Each agent uses C_t to send the compression information. So by using SW the agents send $\{S_t \in [1, 2^{nC_t}]\}$. The final destination uses $S_1, \ldots, S_T$ to decode $z_1, \ldots, z_T$. By knowing $z_1, \ldots, z_T$ it decodes $M \in [1, 2^{nR}]$. The detailed proof goes as follows: we first describe the code construction, in the transmitter and in the agents. Next the processing at transmitter, agents and the decoding at the final destination are given. The conditions (7) result from the described construction so that when $n \to \infty$ the error probability is arbitrary small.

A. Code construction:

For compress and forward transmission:
For every $t = (1, \ldots, T)$:

- Randomly generate $2^{n[\hat{R}_t - C_t]}$ vectors U_t^n according to $P_{U_t^n}(U_t^n) = \prod_i P_{U_t}(U_{t,i})$. Label these $U_t^n(z_t)$.
- Repeat the last step 2^{nC_t} times so that $z_t \in [1, 2^{n\hat{R}_t}]$. For each repetition label the resulting set of z_t by S_{s_t} where $s_t \in [1, 2^{nC_t}]$.
- The message of the t-agent to the destination is therefore s_t.

For every $M \in [1, 2^{nR}]$:

- Randomly choose X^n with $P_{X^n}(X^n) = \prod_i P_X(X_i)$. Label these $X^n(M)$, $M = [1, 2^{nR}]$.

B. Encoding:

Let M be the message to be sent, the source terminal then sends $X^n(M)$.

C. Processing at the agents:

1) Compression: The t agent chooses any of the z_t such that

$$\left(U_t^n(z_t), Y_t^n\right) \in \mathbf{T}_\varepsilon^{t,1}. \tag{59}$$

The probability of an independently generated $U_t^n(\hat{z}_t)$ not to be in $\mathbf{T}_\varepsilon^{t,1}$ is bounded from above by

$$1 - \frac{2^{n[H(U_t,Y_t)-\epsilon]}}{2^{n[H(U_t)+\epsilon]}2^{n[H(Y_t)+\epsilon]}} = 2^{-n[I(U_t;Y_t)+3\epsilon]}. \tag{60}$$

So there is no such z_t with probability P which is upper bounded by

$$P \leq \left(1 - 2^{-n[I(U_t;Y_t)+3\epsilon]}\right)^{2^{n\hat{R}_t}} \tag{61}$$

It is easy to see that this probability goes to zero for sufficiently large n as long as

$$\hat{R}_t > I(U_t;Y_t). \tag{62}$$

After deciding on z_t the agent transmits s_t which corresponds to $z_t \in S_{s_t}$.

D. Decoding (at the destination):

The destination finds the set of indices $\hat{z}_\mathcal{T} \triangleq \{\hat{z}_1, \ldots, \hat{z}_T\}$ of the decoded indices of the compressed vectors which satisfies

$$\begin{cases} \left(U_1^n(\hat{z}_1), \ldots, U_T^n(\hat{z}_T)\right) \in \mathbf{T}_\epsilon^3 \\ \hat{z}_\mathcal{T} \in (S_{s_1}, \ldots, S_{s_T}). \end{cases} \tag{63}$$

Error analysis:
Assume that $\hat{z}_\mathcal{S} \neq z_\mathcal{S}$ and $\hat{z}_{\mathcal{S}^C} = z_{\mathcal{S}^C}$ for some $S \subseteq \mathcal{T}$. Such vector is distributed according to $P_{U_{\mathcal{S}^C}}(U_{\mathcal{S}^C}) \prod_\mathcal{S} P_{U_t}(U_t)$. Thus the probability that such vector belongs to $\mathbf{T}_\epsilon^3$ is upper bounded by

$$\frac{2^{n[H(U_\mathcal{T})+\epsilon]}}{2^{n[H(U_{\mathcal{S}^C})-\epsilon])}2^{n[\sum_{t\in\mathcal{S}} H(U_t)-\epsilon]}} = 2^{n[H(U_\mathcal{T})-H(U_{\mathcal{S}^C})-\sum_{t\in\mathcal{S}} H(U_t)+(|\mathcal{S}|+2)\epsilon]}. \tag{64}$$

Since there are at most $2^{n\sum_\mathcal{S}[\hat{R}_t - C_t]}$ such vectors in the set $(S_{s_1}, \ldots, S_{s_T})$, the probability of such error is upper bounded by:

$$2^{n[\sum_\mathcal{S}\left[\hat{R}_t - C_t - H(U_t)\right] + H(U_\mathcal{T}) - H(U_{\mathcal{S}^C}) + (|\mathcal{S}|+2)\epsilon]}. \tag{65}$$

Which means that as long as

$$\sum_{t\in\mathcal{S}} C_t > \sum_{t\in\mathcal{S}} \left[\hat{R}_t - H(U_t)\right] + H(U_\mathcal{S}|U_{\mathcal{S}^C}) \tag{66}$$

for all $\mathcal{S} \subseteq \mathcal{T}$, the destination will be able to reliably decode $z_\mathcal{T}$ for sufficiently large n.
Now the destination decides that M was sent if

$$\left(X^n(M), U^n(\hat{z}_\mathcal{T})\right) \in \mathbf{T}_\epsilon^4. \tag{67}$$

The probability that there exists $\hat{M} \neq M$ that satisfies (67) is upper bounded by

$$2^{-n[I(X;U_\mathcal{T})-3\epsilon]}. \tag{68}$$

Now summing over $2^{nR} - 1$ and upper bounding, we find that reliable detection of M is possible if

$$R < I(X;U_\mathcal{T}). \tag{69}$$

Taking (62) and (66) and noticing that $\{U_t\}_{t\in\mathcal{T}}$ are independent given (Y_t) we can write the constraints as:

$$\forall \mathcal{S} \subseteq \mathcal{T}: \sum_{t\in\mathcal{S}} C_t > I(U_\mathcal{S}; Y_\mathcal{S}|U_{\mathcal{S}^C}). \tag{70}$$

Which proves (7). The achievable rate (6) is through (69).

Appendix IV
Direct and converse for (27)

A simple compression scheme where:

$$U_t = Y_t + W_t, \tag{71}$$

and W_t is a Gaussian i.i.d. random variable independent of Y_t (no connection with W in the previous appendices) is used for the direct part.
This scheme leads to

$$U_t = X + d_t, \tag{72}$$

and

$$D_t \triangleq \mathrm{E}d_t^2 = P_{n_t} + P_{W_t}. \tag{73}$$

Define $r_t \triangleq I(Y_t;U_t|X)$, which can be written as

$$\begin{aligned} r_t = &\quad H(U_t|X) - H(U_t|Y_t,X) \\ = &\quad \frac{1}{2}\log_2(2\pi e D_t) - \frac{1}{2}\log_2(2\pi e P_{W_t}) \\ = &\quad -\frac{1}{2}\log_2\left(1 - \frac{P_{n_t}}{D_t}\right), \end{aligned} \tag{74}$$

and D_t in terms of r_t

$$\frac{1}{D_t} = \frac{1 - 2^{-2r_t}}{P_{n_t}}. \tag{75}$$

The terms $\{r_t\}$ can take any positive value, when $\{D_t\}$ are determined accordingly (this space $\{\mathbb{R}^+\}^T$ is limited as seen in the next lines, by the SW compression). The last equality can be used to express the maximum mutual information (through maximal ratio combining) in terms of $\{r_t\}$ between X and some subset $U_{\mathcal{S}}$

$$I(X;U_{\mathcal{S}}) = \frac{1}{2}\log_2\left(1+\frac{P_X}{(\sum_{t\in\mathcal{S}} D_t)^{-1}}\right) \quad (76)$$

$$= \frac{1}{2}\log_2\left(1+P_X\sum_{t\in\mathcal{S}} D_t\right) \quad (77)$$

$$= \frac{1}{2}\log_2\left(1+P_X\sum_{t\in\mathcal{S}}\frac{1-2^{-2r_t}}{P_{n_t}}\right). \quad (78)$$

Using (7) as a SW compression (since 9 is fulfilled) we get for $\{r_t\}$ and R in the above simple memoryless compression

$$\sum_{t\in\mathcal{S}} C_t \geq I(U_{\mathcal{S}};Y_{\mathcal{S}}|U_{\mathcal{S}^C}) \quad (79)$$

$$= I(X,Y_{\mathcal{S}};U_{\mathcal{S}}|U_{\mathcal{S}^C}) \quad (80)$$

$$= I(X;U_{\mathcal{S}}|U_{\mathcal{S}^C}) + I(Y_{\mathcal{S}};U_{\mathcal{S}}|X) \quad (81)$$

$$= I(X;U_{\mathcal{T}}) - I(X;U_{\mathcal{S}^C}) + \sum_{t\in\mathcal{S}} I(Y_t;U_t|X) \quad (82)$$

$$= I(X;U_{\mathcal{T}}) - I(X;U_{\mathcal{S}^C}) + \sum_{t\in\mathcal{S}} r_t, \quad (83)$$

where $r_t \triangleq I(Y_t;U_t|X)$ and (82) is true when using memoryless compression, as described in (71).

With (78) these inequalities become

$$\sum_{t\in\mathcal{S}} C_t \geq \frac{1}{2}\log_2\left(1+P_X\sum_{t\in\mathcal{T}}\frac{1-2^{-2r_t}}{P_{n_t}}\right) - \frac{1}{2}\log_2\left(1+P_X\sum_{t\in\mathcal{S}^C}\frac{1-2^{-2r_t}}{P_{n_t}}\right) + \sum_{t\in\mathcal{S}} r_t. \quad (84)$$

Which define a contra-polymatroid. The sum-rate of a polymatroid is known to be equal over all the $T!$ polymatroid vertices [15]. By using the constraints (26) for the contra-polymatroid, a permutation $1,\ldots,T-1$ for the vertex and that $R = \frac{1}{2}\log_2\left(1+P_X\sum_{t\in\mathcal{T}}\frac{1-2^{-2r_t}}{P_{n_t}}\right)$, the direct part of the sum rate is proven [15].

The converse part is showed by using Oohama [14] rather than through the results such as those of section IV. We have that

$$\overline{C}_{\min} \geq \sum_{t=1}^{T} C_t \geq \frac{1}{n} I(Y_{\mathcal{T}}^n;V_{\mathcal{T}}) \quad (85)$$

$$= \frac{1}{n} I(Y_{\mathcal{T}}^n,X^n;V_{\mathcal{T}}) \quad (86)$$

$$= \frac{1}{n} I(X^n;V_{\mathcal{T}}) + \frac{1}{n} I(Y_{\mathcal{T}}^n;V_{\mathcal{T}}|X^n) \quad (87)$$

$$\geq R + \frac{1}{n} I(Y_{\mathcal{T}}^n;V_{\mathcal{T}}|X^n) \quad (88)$$

where (88) is due to (18). Now also by imposing the conditions (26) we can get

$$a_m \geq \quad (89)$$

$$\sum_{t=1}^{m} C_t \geq \frac{1}{n} I(Y_{\mathcal{T}}^n;V_{1,\ldots,m}|V_{m+1,\ldots,T}) \quad (90)$$

$$= \frac{1}{n} I(Y_{\mathcal{T}}^n;V_{\mathcal{T}}) - \frac{1}{n} I(Y_{\mathcal{T}}^n;V_{m+1,\ldots,T}) \quad (91)$$

$$= \frac{1}{n} I(Y_{\mathcal{T}}^n,X^n;V_{\mathcal{T}}) - \frac{1}{n} I(Y_{\mathcal{T}}^n,X^n;V_{m+1,\ldots,T}) \quad (92)$$

$$= \frac{1}{n} I(X^n;V_{\mathcal{T}}) + \frac{1}{n} I(Y_{\mathcal{T}}^n;V_{\mathcal{T}}|X^n) + -\frac{1}{n} I(X^n;V_{m+1,\ldots,T}) + -\frac{1}{n} I(Y_{m+1,\ldots,T}^n;V_{m+1,\ldots,T}|X^n) \quad (93)$$

$$= \frac{1}{n} I(X^n;V_{\mathcal{T}}) + \sum_{t=1}^{T} r_t - \frac{1}{n} I(X^n;V_{\mathcal{T}}) + - \sum_{t=m+1}^{T} r_t \quad (94)$$

$$= \frac{1}{n} I(X^n;V_{\mathcal{T}}) - \frac{1}{n} I(X^n;V_{m+1,\ldots,T}) + \sum_{t=1}^{m} r_t \quad (95)$$

Where (92) and (94) is since V_t is a deterministic function of Y_t^n. These lead, by redefining $r_t \triangleq \frac{1}{n} I(Y_t^n;V_t|X^n)$ and since the rate R is upper bounded by (18) to (27) with the constraints

$$a_m \geq \sum_{t=1}^{m} r_t + R - \frac{1}{n} I(X^n;V_{m+1,\ldots,T}) \quad (96)$$

$$R \leq \frac{1}{n} I(X^n;V_{\mathcal{T}}). \quad (97)$$

Now using the entropy power inequality, as in [14] it is seen that:

$$2^{\frac{2}{n} I(X^n;V_{\mathcal{S}})} \leq 1 + P_X \sum_{t\in\mathcal{S}} \frac{1-2^{-2r_t}}{P_{n_t}}. \quad (98)$$

So that we get to the constraints (28), where the last equality of (28) is since the constraint $R \leq \frac{1}{n} I(X^n;V_{\mathcal{T}}) \leq \frac{1}{2}\log_2\left(1+P_X\sum_{t\in\mathcal{T}}\frac{1-2^{-2r_t}}{P_{n_t}}\right)$ is maximized with equality.

References

[1] C. Ng and A. Goldsmith, “Transmitter cooperation in ad-hoc wireless networks: Does dirty-paper coding beat relaying?” in *the proceedings of IEEE INFORMATION THEORY WORKSHOP (ITW2004)*, San Antonio, Texas, October.

[2] P. Gupta and P. R. Kumar, “Towards an information theory of large networks: an achievable rate region,” *IEEE Trans. Inform. Theory*, vol. 49, no. 8, pp. 1877–1894, Aug. 2003.

[3] T. M. Cover and A. A. El-Gamal, “Capacity theorems for the relay channel,” *IEEE Trans. Inform. Theory*, vol. 25, no. 5, pp. 572–584, Jan 1979.

[4] G. Kramer, M. Gastpar, and P. Gupta, “Information-theoretic multi-hopping for relay networks,” in *International Zürich seminar on communication*, Swizerland, Feb. 2004.

[5] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec 2004.

[6] E. Erkip and M. Yuksel, “Diversity gains and clustering in wireless relaying,” in *Proc IEEE Int Symp Info Theory ISIT2004*, Chicago, IL, June 2004.

[7] R. Dabora and S. D. Servetto, “Broadcast channels with cooperating receivers: a downlink for sensor reachback problem,” in *Proc IEEE Int Symp Info Theory ISIT2004*, Chicago,IL, June.

[8] O. Lévêque and E. Telatar, “Information theoretic upper bounds on the capacity of large extended ad hoc wireless networks,” *IEEE Trans. Inform. Theory*, Mar. 2005.

[9] T. Berger, Z. Zhang, and H. Viswanathan, “The CEO problem,” *IEEE Trans. Inform. Theory*, vol. 42, no. 3, pp. 887–902, May 1996.

[10] A. El-Gamal and T. Cover, “Achievable rates for multiple descriptions,” *IEEE Trans. Inform. Theory*, vol. 6, pp. 8151–8157, November 1982.

[11] D. Slepian and J. K. Wolf, “Noiseless coding of correlated information sources,” *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 471–480, July 1973.

[12] A. D. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” *IEEE Trans. Inform. Theory*, vol. 22, no. 1, pp. 1–10, Jan 1976.

[13] J. Chen, X. Zhang, T. Berger, and S. B. Wicker, “An upper bound on the sum rate distortion function and its corresponding rate allocation schemes for the CEO problem,” *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 977–987, Aug. 2004.

[14] Y. Oohama, “The rate-distortion function for the quadrati Gaussian CEO problem,” *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 1057–1070, May 1998.

[15] V. Prabhakaran, D. Tse, and K. Ramchandran, “Rate region of the quadratic Gaussian CEO problem,” in *Proc IEEE Int Symp Info Theory ISIT2004*, Chicago, IL, June 2004, p. 119.

[16] R. Zamir, S. Shamai, and U. Erez, “Nested linear/lattice codes for structured multiterminal binning,” *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1250–1276, June 2002.

[17] Y. Steinberg and N. Merhav, “On successive refinement for the Wyner-Ziv problem,” *IEEE Trans. Inform. Theory*, vol. 50, no. 8, pp. 1636–1654, Aug. 2004.

[18] J. Barros and S. D. Servetto, “A note on cooperative multiterminal source coding,” in *the Proceedings of the 38th Annual Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, Mar. 2004.

[19] A. D. Murugan, P. K. Gopala, and H. El-Gamal, “Correlated sources over wireless channels: cooperative source-channel coding,” *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 988–999, Aug. 2004.

[20] A. Orlitsky and R. Roche, “Coding for computing,” *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 903–917, Mar. 2001.

[21] B. E. Schein, “Distributed coordination in network information theory,” Ph.D. dissertation, MIT, October 2001.

[22] S. Shamai and A. Lapidoth, “Fading channels: How perfect need ”perfect side information” be?” *IEEE Trans. Inform. Theory*, vol. 48, no. 5, pp. 1118–1134, May 2002.

[23] I. G. Stiglitz, “Coding for a class of unknown channels,” *IEEE Trans. Inform. Theory*, vol. IT-12, no. 2, pp. 189–195, April 1966.

[24] T. M. Cover and J. A. Thomas, *Elements of Information theory*. John Wiley & Sons, Inc., 1991.

[25] G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” Bell Labs, Lucent, Murray Hill, NJ,” Technical Memorandum, Feb. 2004.