On Wireless Sensor Networks with Arbitrarily Correlated Sources

By

Seyyed Hassan Mahboubi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2008

Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Seyyed Hassan Mahboubi

Abstract

An achievable rate region for general wireless sensor networks is proposed. A general multi-source, multi-relay, multi-destination wireless sensor network with arbitrarily correlated sources is considered. Each node can sense some real phenomena and send its readings to one or more sinks (data gathering nodes) via some relays. It also can relay some correlated or independent readings of other nodes, simultaneously. In this problem the source and channel coding separation is not optimal and the information which each reading has about others nodes is destroyed in separation. Thus, a joint source channel coding scheme can be used. The problem consists of relay channels and multiple access channels with arbitrarily correlated sources. The proposed scheme is based on regular block Markov encoding/backward decoding and code division multiple-access (CDMA) and the result is a combination of multi-relay and multiple-access with correlated sources.

Acknowledgement

I wish to take this opportunity to express my appreciation for numerous individuals who made my years at Waterloo a rewarding experience.

I am very grateful to my thesis supervisor, Prof. Amir K. Khandani, for giving me the opportunity and freedom to pursue my research under his supervision. He has been instrumental in bringing me to Waterloo and has provided support during my stay here. I wish to thank him for his constant encouragement and patience. His profound knowledge and scientific curiosity have set high standards and are a constant source of inspiration. During my interaction with him as his student and advisee, I benefited tremendously from his numerous qualities of a mentor and collaborator. His clarity of thought and keen insights have greatly influenced my research. His advice, both scholarly and non-academic, and most of all his friendship, leave me greatly indebted.

I have been fortunate to associate with many remarkable people during the time spent at Waterloo. My gratitude goes to numerous friends and colleagues at the University of Waterloo especially in the Coding and Signal Transmission (CST) laboratory for providing such a creative, fun and intellectually stimulating environment. There are numerous members of the CST lab who contribute to the excellent working environment and creative atmosphere. In fact, there are too many to list in the acknowledgements but I am nevertheless truly grateful. However, in particular, I would like to thank Alireza Bayesteh, Seyyed-Ali Ahmadzadeh and also my friend Sahinaz Safari for many lively discussions ranging over a rather wide variety of topics. I will always remember their support and friendship.

Last, and most importantly, I am forever indebted to my parents, Seyyed Naser Mahboubi and Maryam Chinikar, and my sister,Parisa Mahboubi and my grandparents. Their invaluable and relentless support, encouragement, and love are without doubt the most important reasons for my success.and at the end I should appreciate remarkable support and endless love from my best friend ever ,Golmehr Sistani,who brought happiness to my life.

Dedication

This thesis is dedicated to my wonderful parents, my beautiful sister and my beloved grandparents.

Table of Contents
List of Figures

List of Figures

Chapter 1:

Introduction to Wireless Sensor Networks

1.1-Sensor Networks:

Recent technological improvements have made the deployment of small, inexpensive, low power, distributed devices, which are capable of local processing and wireless communication, a reality. Such nodes are called as sensor nodes. Each sensor node is capable of only a limited amount of processing. But when coordinated with the information from a large number of other nodes, they have the ability to measure a given physical environment in great detail. A sensor network can be described as a collection of sensor nodes which co-ordinate to perform some specific action. Unlike traditional networks, sensor networks depend on dense deployment and co-ordination to carry out their tasks.

Previously, sensor networks consisted of small number of sensor nodes that were wired to a central processing station. However, nowadays, the focus is more on wireless, distributed, sensing nodes. But, why distributed, wireless sensing [1]? The reason is the exact location of a particular phenomenon is unknown, distributed sensing allows for closer placement to the phenomenon than a single sensor would permit. Also in many cases, multiple sensor nodes are required to overcome environmental obstacles like obstructions, line of sight constraints etc. In most cases, the environment to be monitored does not have an existing infrastructure for either energy or communication. It becomes imperative for sensor nodes to survive on small, finite sources of energy and communicate through a wireless communication channel.

Another requirement for sensor networks would be distributed processing capability. This is necessary since communication is a major consumer of energy. A centralized system would mean that some of the sensors would need to communicate over long distances that lead to even more energy depletion. Hence, it would be a good idea to process locally as much information as possible in order to minimize the total number of bits transmitted.

1.2-Applications of sensor networks:

Sensor networks have a variety of applications. Examples include environmental monitoring – which involves monitoring air soil and water, condition based maintenance, habitat monitoring (determining the plant and animal species population and behavior), seismic detection, military surveillance, inventory tracking, smart spaces etc. In fact, due to the pervasive nature of micro-sensors, sensor networks have the potential to revolutionize the very way we understand and construct complex physical system [2].

1.3. The Sensor Reachback Problem:

Wireless sensor networks made up of mostly unreliable devices equipped with limited sensing, processing and transmission capabilities, have recently sparked a fair amount of interest in communications problems involving multiple correlated sources and large-scale wireless networks [3]. It is envisioned that an important class of applications for such networks involves a dense deployment of a large number of sensors over a fixed area. In these areas a physical process unfolds the task of these sensors is then to collect measurements, encode them, and relay them to some data collection point where this data is to be analyzed, and possibly acted upon. There are several aspects that make this communications problem interesting:

- ***Correlated Observations:*** If we have a large number of nodes sensing a physical process within a confined area, it is reasonable to assume that their measurements are correlated. This correlation may be exploited for efficient encoding/decoding.
- ***Cooperation among Nodes:*** Before transmitting data to the remote receiver, the sensor nodes may establish a *conference* to exchange information over the wireless medium and increase their efficiency or flexibility through cooperation.
- ***Channel Interference:*** If multiple sensor nodes use the wireless medium at the same time (either for conferencing or reachback), their signals will necessarily interfere with each other. Consequently, reliable communication in a reachback network requires a set of rules that control (or exploit) the interference in the wireless medium.

In order to capture some of these key aspects, while still being able to provide complete results, we make some modeling assumptions, discussed next.

1.3.1- Source Model:

It is common to assume that the sources are memoryless, and thus consider only the spatial correlation of the observed samples and not their temporal dependence (since the latter dependencies could be dealt with by simple extensions of results to the case of ergodic sources). Furthermore, each sensor node v_i observes only one component U_i and must transmit enough information to enable the sink node v_0 to reconstruct the whole vector $U_1U_2......U_M$. This assumption is the most natural one to make for scenarios in which data is required at a remote location for fusion and further processing.however the data capture process is distributed, with sensors able to gather *local* measurements only, and deeply embedded in the environment.

A conceptually different approach would be to assume that all sensor nodes get to observe independently corrupted noisy versions of one and the same source of information U, and it is this source (and not the noisy measurements) that needs to be estimated at a remote location. This approach seems better suited for applications involving non-homogeneous sensors, in which each one of the sensors gets to observe different characteristics of the same source (e.g., multispectral imaging), and therefore leads to a conceptually very different type of sensing applications from those of interest in this work. Such an approach leads to the so called *CEO problem* studied by Berger, Zhang and Viswanathan in [4].

1.3.2- Independent Channels:

Common motivation to consider a network of independent DMCs(Discrete Memoryless Channels) is two-fold. From a pure information-theoretic point of view independent channels are interesting . Moreover, a corollary of said coding theorem does provide a conclusive answer for a special case of the multiple access channel with correlated sources. A problem for which no general converse is known. From a more practical point of view, the assumption of independent channels is valid for any network that controls interference by means of a reservation-based medium-access control protocol (e.g., TDMA). This option seems perfectly reasonable for sensor networking scenarios in which sensors collect data over extended periods of time, and must then transmit their accumulated measurements simultaneously. In this Case, a key assumption in the design of standard random access techniques for multiaccess communication breaks down—the fact those individual nodes will transmit with low probability [5, Chapter 4].

As a result, classical random access would result in too many collisions and hence low throughput. Alternatively, instead of *mitigating* interference, a medium access control (MAC) protocol could attempt to *exploit* it, in the form of using cooperation among nodes to generate waveforms that add up constructively at the receiver (cf. [6], [7], [8]). Providing an information-theoretic analysis of such cooperation mechanisms would be very desirable. It entails dealing with correlated sources and a general multiple access channel, dealing with correlated sources and an array of independent channels constitutes a reasonable first step towards that goal, and is also interesting in its own right, since it provides the ultimate performance limits for an important class of sensor networking problems.

1.3.3- Perfect Reconstruction at the Receiver:

In the formulation of the sensor reachback problem, the far receiver is interested in reconstructing the entire field of sensor measurements with arbitrarily small probability of error. This formulation leads to a natural *capacity* problem, in the classical sense of Shannon Theorem. Alternatively, one could relax the condition of perfect reconstruction, and tolerate some distortion in their construction of the field of measurements at the far receiver. Leading to the so called *Multiterminal Source Coding* problem studied by Berger [9]. This condition could be further relaxed, to require a faithful reproduction of the *image* of some function f of the sources, leading to a problem studied extensively by Csiszar, Körner and Marton [10], [11].

1.4- An Information Theoretic View of Architectural Issues:

For large-scale, complex systems of the type of interest in this work, the complexity of basic questions of design and performance analysis appears daunting:

• How should nodes cooperate to relay messages to the data collector node v_0? Should they decode received messages, re-encode them, and forward to other nodes? Should they map channel outputs to channel inputs without attempting to decode? Should they do something else?

• How should redundancy among the sources be exploited? Should we compress the information as much as possible? Should we leave some of that redundancy to combat noise in the channels? Is there a source/channel separation theorem in these networks?

• How do we measure performance of these networks, what are appropriate cost metrics? How do we design networks that are efficient under an appropriate cost metric?

In [12], a number of examples are identified in which the existence of a simple architecture has played an enabling role in the proliferation of technology: the von Neuman computer architecture, separation of source and channel coding in communications, separation of plant and controller in control systems, and the OSI layered architecture model. So what all these questions boil down to is an issue similar to those considered in [12]:

- what are appropriate abstractions of the network, similar to the IP protocol stack for the Internet, based on which we can break the design task into independent reusable components, optimize the design of these components, and obtain an *efficient* system as a result?

In this work, we show how information theory is indeed capable of providing very meaningful answers to this problem. Information theory, in one of its applications, deals with the analysis of performance of communication systems. So, to some it may seem the natural theory to turn to for guidance in the task of searching for suitable network architecture. However, to others it may seem unnatural to do so. It is well known that information theory and communication networks have not had fruitful interactions in the past, as explained by Ephremides and Hajek [13]. In the presence of these mixed indicators, we take the stand that indeed information theory has a great deal to offer in the task at hand.

For this setup, Shannon established that reliable communication of a source over a noisy channel is possible if and only if the entropy rate of the source is less than the capacity of the channel [14, Ch. 8.13]. This result, known as the source/channel separation theorem, has a double significance. On one hand, it provides an exact single-letter characterization of conditions under which reliable communication is possible. On the other hand, and of particular interest to the task at hand for us, it is a statement about the *architecture* of an optimal communication system: the encoder/decoder design task can be split into the design and optimization of two independent components. So it is inspired by Shannon's teachings for point-to-point systems that we ask in this work, and answer in the affirmative, the question of whether it is possible or not to derive similar useful architectural guidelines for the class of networks under consideration.

1.5- Related Work and Our Work:

The problem of communicating distributed correlated sources over a network of point-to-point links is closely related to several classical problems in network information theory. To set the stage for the main contributions of this thesis, we now review related previous work.

1.5.1- Distributed Correlated Sources and Multiple Accesse:

The concept of separate encoding of correlated sources was studied by Slepian and Wolf in their seminal paper [15], where they proved that two correlated sources (U_1U_2) drawn *i.i.d.* ~ $p(u_1u_2)$ can be compressed at rates (R_1, R_2) if and only if

$$\mathrm{R}_1 \geq \mathrm{H}(\mathrm{U}_1 \mid \mathrm{U}_2)$$

$$\mathrm{R}_2 \geq \mathrm{H}(\mathrm{U}_2 \mid \mathrm{U}_1)$$

$$\mathrm{R}_1 + \mathrm{R}_2 \geq \mathrm{H}(\mathrm{U}_1\mathrm{U}_2).$$

Assume now that (U_1U_2) are to be transmitted with arbitrarily small probability of error to a joint receiver over a multiple access channel with transition probability $\mathrm{p}(\mathrm{y} \mid \mathrm{x}_1\mathrm{x}_2)$. Knowing that the capacity of the multiple access channel with independent sources is given by the convex hull of the set of points (R1,R2) satisfying [14, Ch. 14.3]

$$R_1 < I(X_1;Y|X_2)$$
$$R_2 < I(X_2;Y|X_1)$$
$$R_1 + R_2 < I(X_1X_2;Y),$$

it is not difficult to prove that Slepian-Wolf source coding of (U_1U_2) followed by separate channel coding yields the following *sufficient* conditions for reliable communication

$$H(U_1|U_2) < I(X_1;Y|X_2)$$
$$H(U_2|U_1) < I(X_2;Y|X_1)$$
$$H(U_1U_2) < I(X_1X_2;Y).$$

These conditions, which basically state that the Slepian-Wolf region and the capacity region of the multiple access channel have a non-empty intersection, are sufficient but not necessary for reliable communication, as shown by Cover, El Gamal, and Salehi with a simple counterexample in [16]. In that same paper, the authors introduce a class of *correlated* joint source/channel codes, which enables them to increase the region of achievable rates to(for some $p(u_1u_2x_1x_2y) = p(u_1u_2)\cdot p(x_1|u_1)\cdot p(x_2|u_2)\cdot p(y|x_1x_2)$)

$$H(U_1|U_2) < I(X_1;Y|X_2U_2)$$
$$H(U_2|U_1) < I(X_2;Y|X_1U_1)$$
$$H(U_1U_2) < I(X_1X_2;Y),$$

Also in [16], the authors generalize this set of sufficient conditions to sources (U1U2) with a common part $W = f(U_1) = g(U_2)$, but they were not able to prove a converse, i.e., they were not able to show that their region is indeed the capacity region of the multiple access channel with correlated sources. Later, it was shown with a carefully constructed example by Dueck in [17] that indeed the region defined by eqns. (1)-(3) is not tight. Related problems were considered by Slepian and Wolf [18], and Ahlswede and Han [19]. To this date however, the general problem still remains open.

Assuming independent sources, Willems investigated a cooperative scenario, in which encoders exchange messages over *conference* links of limited capacity prior to transmission over the multiple access channel [20].In this case, the capacity region is given by

$$R_1 < I(X_1;Y|X_2Z) + C_{12}$$
$$R_2 < I(X_2;Y|X_1Z) + C_{21}$$

$$R_1 + R_2 < \min\{ I(X_1X_2;Y|Z) + C_{12} + C_{21}, I(X_1X_2;Y)\},$$

for some auxiliary random variable Z such that $|Z| = \min(|X_1|\cdot|X_2|+2, |Y|+3)$, and for a joint distribution $p(zx_1x_2y_1y_2) = p(z)\cdot p(x_1|z)\cdot p(x_2|z)\cdot p(y|x_1x_2)$.

1.5.2- Correlated Sources and Networks of DMCs:

Recently, an early paper was brought to our attention, in which Han considers the transmission of correlated sources to a common sink over a network of independent channels [21]. Although the problem setup is less general than ours, in that (a) each source block and each transmitted codeword participate only once in the encoding process and (b) the intermediate nodes are assumed to decode the data before passing it on.

.

1.5.3 Network Coding:

Another closely related problem is the well known *network coding* problem, introduced by Ahlswede, Cai, Li and Yeung [22]. In that work, the authors establish the need for applying coding operations at intermediate nodes to achieve the max-flow/min-cut bound of a general multicast network. A converse proof for this problem was provided by Borade [23]. Linear codes were proposed by Li, Yeung and Cai in [24], and Koetter and M´edard in [25].

Effros, M´edard et al. have developed a comprehensive study on separate and joint design of linear source, channel and network codes for networks with correlated sources under the assumption that all operations are defined over a common finite field [26]. For this particular case, optimality of separate linear source and channel coding was observed in the one-receiver instance, but the result of [26] does not prove that it holds for general networks and channels with arbitrary input and output alphabets. Error exponents for multicasting of correlated sources over a network of noiseless channels were given by Ho, M´edard et al. in [27], and networks with undirected links were considered by Li and Li in [28]. Another problem in which network flow techniques have been found useful is that of finding the maximum stable throughput in certain networks. In this problem, posed by Gupta and Kumar in [29], it is sought to determine the maximum rate at which nodes can inject bits into a network, while keeping the system stable. This problem was reformulated by Peraki and Servetto as a multicommodity flow problem, for which tight bounds were obtained using elementary counting techniques [30], [31].

Xie and Kumar [32] propose an achievable rate region for multi-source, multi-relay, multi-destination networks. Actually, they propose a channel coding scheme for independent source date. As long as the source and channel coding separation theorem is not valid in such problem [16], their scheme can not exploit the high dependency among sensors readings. Cover, *et*. *al*. [16] propose an achievable rate region for multiple-access channels with arbitrarily correlated sources. Their region is larger than the region achieved by separate source channel coding scheme but it is not optimal in general [17]. Their problem is only a multiple-access channel with two sources and without any relay. There are also some other works in multi-user information theory with correlated sources [33]-[38].

In real wireless sensor networks, we deal with large networks with correlated sources. In this thesis, the idea of [32] and [16] is used to find an achievable rate region for large sensor networks with correlated data.

Chapter 2:

Achievability

2.1-Model and Preliminaries:

In this section, we describe the notations, information theoretical preliminaries and model of systems.

2.1.1.Notations:

We denote random variables with capital letter X and its realization with lower case letter x. a discrete random variable X takes values in a finite discrete set X and ||X|| is the cardinality of X. Vectors are denoted by boldface letters, e.g. $\mathbf{X}^N$, and its realization is denoted by x^{N}. X_n and x_n are the *n*'th element of $\mathbf{X}^N$ and x^{N}, respectively. $p_X(x)$ And $p_X^N(\mathrm{x}^N)$ denote the probability distribution of X on X and X^{N} on X^{N}. For brevity, we may omit the subscript X when it is obvious from the context. We denote the set of ε-weakly typical sequences and the set of ε-strongly typical sequences w.r.t. distribution $p_X(x)$ by $A_\varepsilon^N(X)$ and $A_\varepsilon^{*(N)}(X)$.

2.1.2-Typical sequences:

Our proofs are based on typical sequences properties. Here, we review some properties of typical sequences which are used in this article. The details can be found in [14] and [39]. Let $(X_1, X_2, ..., X_n)$ denote a finite collection of discrete random variables with some fixed joint distribution, $p(x_1, x_2, ..., x_n)$, $(x_1, x_2, ..., x_n) \in X_1 \times X_2 \times ... \times X_n$. Let *S* denote an ordered subset of these random variables and consider *N* independent copies of *S*. Thus,

$$p(\mathrm{s}^N) = \prod_{n=1}^{N} p(s_n), \quad \mathrm{s}^N \in S^N .$$

Definition 1:

The set $A_\varepsilon^N(X)$ of *ε*-weakly typical *n*-sequences or *ε*-typical *n*-sequences $(\mathrm{x}_1^N, \mathrm{x}_2^N, ..., \mathrm{x}_k^N)$ is defined by

$$A_\varepsilon^N(X_1, X_2, ..., X_k) = A_\varepsilon^N = \left\{ (\mathrm{x}_1^N, \mathrm{x}_2^N, ..., \mathrm{x}_k^N) : \left| -\frac{1}{N} \log p(\mathrm{s}^N) - H(S) \right| < \varepsilon, \forall S \subseteq \{X_1, X_2, ..., X_k\} \right\}$$

For any $\varepsilon > 0$, there exists an integer *N* such that $A_\varepsilon^N(S)$ satisfies

1) $\Pr\left(A_\varepsilon^N(S)\right) \geq 1 - \varepsilon, \forall S \subseteq \{X_1, X_2, ..., X_k\}$

2) $\mathrm{s}^N \in A_\varepsilon^N(S) \Rightarrow \left| -\frac{1}{N} \log p(\mathrm{s}^N) - H(S) \right| < \varepsilon$

3) $(1-\varepsilon) 2^{N(H(S)-\varepsilon)} \leq \left\| A_\varepsilon^N(S) \right\| \leq 2^{N(H(S)+\varepsilon)}$

4) Let S_1 and S_2 be two subsets of $\{X_1, X_2, ..., X_k\}$. If $(\mathrm{s}_1^N, \mathrm{s}_2^N) \in A_\varepsilon^N(S_1, S_2)$, then

$$(1-\varepsilon) 2^{-N(H(S_1|S_2)+2\varepsilon)} \leq p(\mathrm{s}_1^N \mid \mathrm{s}_2^N) \leq 2^{-N(H(S_1|S_2)-2\varepsilon)} .$$

5) Let S_1, S_2 and S_3 be three subsets of $\{X_1, X_2, ..., X_k\}$. If $(\mathrm{s}_1^N, \mathrm{s}_2^N, \mathrm{s}_3^N) \in A_\varepsilon^N(S_1, S_2, S_3)$, and if

$$p(\mathrm{s}_1^N, \mathrm{s}_2^N, \mathrm{s}_3^N) = \prod_{n=1}^{N} p(s_{1n} \mid s_{3n}) p(s_{2n} \mid s_{3n}) p(s_{3n})$$

Then

$$(1-\varepsilon)2^{-N\left(I(S_1;S_2|S_3)+6\varepsilon\right)} \le p\left\{(\mathbf{S}_1^N,\mathbf{S}_2^N,\mathbf{S}_3^N)\in A_\varepsilon^N(S_1,S_2,S_3)\right\} \le 2^{-N\left(I(S_1;S_2|S_3)-6\varepsilon\right)}.$$

***Definition 2*:**

Let S_1 and S_2 be two subsets of $\{X_1,X_2,...,X_k\}$. For any $\varepsilon>0$, define $A_\varepsilon^N\left(S_1\,|\,\mathrm{s}_2^N\right)$ to be the set of s_1^N sequences that are jointly ε-typical with a particular s_2^N sequence. If $\mathrm{s}_2^N\in A_\varepsilon^N(S_2)$, then for sufficiently large N, we have

$$\left\|A_\varepsilon^N(S_1\,|\,\mathrm{s}_2^N)\right\| \le 2^{N\left(H(S_1|S_2)+2\varepsilon\right)}.$$

Remark 1:

These results also apply to continuous random variables with minor modifications. Replace the entropy H(X) of a discrete random variable X with differential entropy h(X) of a continuous random variable X. Instead of cardinality of the typical set $\left\|A_\varepsilon^N(S)\right\|$ in the discrete case, the volume of the typical set, $Vol\left(A_\varepsilon^N(S)\right)$, in the continuous case is used.

***Lemma 1*[16]:**

Let (Z_1,Z_2,Z_3,Z_4,Z_5) be random variables with joint distribution with $p(z_1,z_2,z_3,z_4,z_5)$. Fix $(\mathrm{z}_1^N,\mathrm{z}_2^N)\in A_\varepsilon^N$, and let $\mathrm{Z}_3^N,\mathrm{Z}_4^N,\mathrm{Z}_5^N$ be drown according to

$$p(\mathbf{Z}_3^N=\mathrm{z}_3^N,\mathbf{Z}_4^N=\mathrm{z}_4^N,\mathbf{Z}_5^N=\mathrm{z}_5^N\,|\,\mathrm{z}_1^N,\mathrm{z}_2^N)=\prod_{n=1}^N p(z_{3n}\,|\,z_{1n},z_{2n})p(z_{4n}\,|\,z_{3n},z_{2n})p(z_{5n}\,|\,z_{3n},z_{1n})$$

In other words $\mathbf{Z}_3^N$ depends only on $\mathbf{Z}_1^N$, $\mathbf{Z}_2^N$; $\mathbf{Z}_4^N$ depends only on $\mathbf{Z}_3^N$, $\mathbf{Z}_2^N$; and $\mathbf{Z}_5^N$ depends only on $\mathbf{Z}_3^N$, $\mathbf{Z}_1^N$. Then

$$\Pr\left\{(\mathrm{z}_1^N,\mathrm{z}_2^N,\mathbf{Z}_3^N,\mathbf{Z}_4^N,\mathbf{Z}_5^N)\in A_\varepsilon^n\right\} \le 2^{-N\left(I(Z_1;Z_4|Z_2,Z_3)+I(Z_5;Z_2,Z_4|Z_1,Z_3)-8\varepsilon\right)}$$.greater than n

Proof*:** In ***appendix I.

2.1.3-Model of the system:

In the general case, we consider a multi-source, multi-destination, multi-relay network with *n* nodes N= {1, 2,…, *n*}. Let M= {1, 2,…, *m*} denote the set of correlated sources and each of them $k \in \mathcal{M}$ corresponds to a source node $s^{(k)} \in \mathcal{N}$ and a set of destination nodes $\mathcal{D}^{(k)} \subset \mathcal{N}$. For each $k \in \mathcal{M}$, the source node $s^{(k)}$ should send the source *k* to all the nodes in $\mathcal{D}^{(k)}$. The number *m* can be greater than *n*, since multiple sources having different destinations can originate from the same node.

For each source $k \in \mathcal{M}$, there is a multi-relay route $\mathcal{N}^{(k)} \subseteq \mathcal{N}$ which is an ordered set of nodes starting with $s^{(k)}$. For any $i, j \subset \mathcal{N}^{(k)}$, the order is defined by $i \prec j(k)$ if node *i* is upstream of node *j* along the route. Since all the nodes in the multi-relay route will obtain the source information, the multi-cast is fulfilled as long as the route is chosen such that $\mathcal{D}^{(k)} \subset \mathcal{N}^{(k)}$.

Consider a discrete memoryless network channel model described by

$$\left(\mathcal{X}_1 \times \ldots \times \mathcal{X}_L, p(y_1,\ldots,y_L \mid x_1,\ldots,x_L), \mathcal{Y}_1 \times \ldots \times \mathcal{Y}_L\right)$$

where $\mathcal{X}_i$ and $\mathcal{Y}_i$, i=1,…,n are finite input and output alphabets respectively, and $p(y_1,\ldots,y_L \mid x_1,\ldots,x_L)$ is a probability distribution on $\mathcal{Y}_1 \times \ldots \times \mathcal{Y}_L$ for each $(x_1,\ldots,x_L)$. At any time *t*=1,2,…, each node $i \in \mathcal{N}$ sends $x_{it} \in \mathcal{X}_i$ into the channel and receives $y_{it} \in \mathcal{Y}_i$ from the channel. The distribution of the outputs $(y_{1t},\ldots,y_{nt})$ depends only on the inputs at the time *t* via

$$p\left(y_{1t},\ldots,y_{nt} \mid x_{1t},\ldots,x_{nt}\right).$$

Like [32], we choose a multi-relay route $\mathcal{N}^{(k)} \subseteq \mathcal{N}$ for each source $k \in \mathcal{M}$, such that $\mathcal{D}^{(k)} \subset \mathcal{N}^{(k)}$. Along each route, we use the scheme of regular block Markov encoding/backward decoding. According to [32] these *m* routs can be unified into a joint backward decoding scheme, if it is possible to assign a nonnegative integer to each node in the network, such that along any multi-relay route excluding the source node, the integers are strictly increasing.

2.2-Main Results:

Consider the simplest multi-source, multi-relay, multi-destination network with five nodes which is depicted in Fig 1 where two source nodes want to send two sources W_1 and W_2 which are correlated according to $p(w_1, w_2)$ to the same destination node 5 via two relays node 3 and node 4. There are two routs $1 \rightarrow 3 \rightarrow 5$ and $2 \rightarrow 4 \rightarrow 5$ in the network, and relays help source nodes with decode and forward strategy. The source nodes 1 and 2 transmit x_1 and x_2 based on their data (readings). The relay nodes 3 and 4 transmit x_3 and x_4 based only on the signals they have already received:

$$x_{3n} = f_{3,n}(y_{31}, \dots, y_{3,t-1}),$$
$$x_{4n} = f_{4,n}(y_{41}, \dots, y_{4,t-1}),$$

A rate pair (R_1, R_2) of nodes 1 and 2 is achievable if both the messages can be decoded at the destination node 5 with an arbitrarily small probability of error. According to Xie and Kumar [1, theorem 2.1], for this network the following rate region is achievable:

$$\begin{cases} R_1 < I(X_1; Y_3 \mid X_3) \\ R_2 < I(X_2, Y_4 \mid X_4) \\ R_1 < I(X_1, X_3; Y_5 \mid X_2, X_4) \\ R_2 < I(X_2, X_4; Y_5 \mid X_1, X_3) \\ R_1 + R_2 < I(X_1, X_2, X_3, X_4; Y_5) \end{cases} \qquad (1)$$

For some $p(x_1, x_3)p(x_2, x_4)$. It follows easily that W_1 and W_2 can be sent over the above channel if for some

$p(x_1, x_2, x_3, x_4) = p(x_1, x_3)p(x_2, x_4)$,

$$\begin{cases} H(W_1) < I(X_1; Y_3 \mid X_3) \\ H(W_2) < I(X_2, Y_4 \mid X_4) \\ H(W_1) < I(X_1, X_3; Y_5 \mid X_2, X_4) \\ H(W_2) < I(X_2, X_4; Y_5 \mid X_1, X_3) \\ H(W_1) + H(W_2) < I(X_1, X_2, X_3, X_4; Y_5) \end{cases} \quad . \ (2)$$

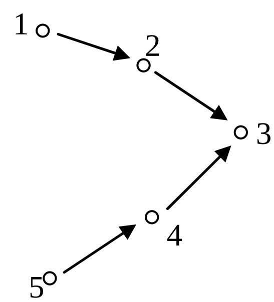

Figure 1. A simple two source, two relay, one destination network.

This achievable rate region can be increased by decreasing the left side of (2) and by increasing the right side of (2). The left side can be decreased using Slepin-Wolf coding [15] and the right side can be increased by allowing X_1 and X_2 to depend on W_1 and W_2.

***Theorem 1*:**

A source $(W_1^N, W_2^N) \sim \prod_n p(w_{1n}, w_{2n})$ can be sent with arbitrarily small probability of error over the above channel if there exist $p(x_1 \mid w_1), p(x_2 \mid w_2)$, such that

$$\begin{cases} H(W_1) < I(X_1; Y_3 \mid X_3) \\ H(W_2) < I(X_2, Y_4 \mid X_4) \\ H(W_1 \mid W_2) < I(X_1, X_3; Y_5 \mid X_2, X_4, W_2) \\ H(W_2 \mid W_1) < I(X_2, X_4; Y_5 \mid X_1, X_3, W_1) \\ H(W_1, W_2) < I(X_1, X_2, X_3, X_4; Y_5) \end{cases} \quad (3)$$

Where $p(w_1, w_2, x_1, x_2, y) = p(w_1, w_2) p(x_1 \mid w_1) p(x_2 \mid w_2) p(y \mid x_1, x_2)$.

This result can be generalized to the case of two sources (W_1, W_2) with common part $W_c = f(W_1) = g(W_2)$. The common part W_c can be defined as follows:

***Definition 3*[40]:**

The common part W_c of two random variables W_1 and W_2 is defined by finding the maximum integer k such that there exist functions f and g

$$f : W_1 \to \{1, 2, ..., k\}$$
$$g : W_2 \to \{1, 2, ..., k\}$$

with $P\{f(W_1 = i)\} > 0$, $P\{g(W_2 = i)\} > 0$, i=1,2,…,k, such that $f(W_1) = g(W_2)$ with probability one and then defining $W_c = f(W_1) = g(W_2)$.

***Theorem 2*:**

A source $(W_1^N, W_2^N) \sim \prod_n p(w_{1n}, w_{2n})$ can be sent with arbitrarily small probability of error over the above channel if there exist $p(s)$, $p(x_1 \mid s, w_1), p(x_2 \mid s, w_2)$, such that

$$\begin{cases} H(W_1) < I(X_1; Y_3 \mid X_3) \\ H(W_2) < I(X_2, Y_4 \mid X_4) \\ H(W_1 \mid W_2) < I(X_1, X_3; Y_5 \mid X_2, X_4, W_2, S) \\ H(W_2 \mid W_1) < I(X_2, X_4; Y_5 \mid X_1, X_3, W_1, S) \\ H(W_1, W_2 \mid W_c) < I(X_1, X_2, X_3, X_4; Y_5 \mid S, W_c) \\ H(W_1, W_2) < I(X_1, X_2, X_3, X_4; Y_5) \end{cases} \quad (4)$$

where $p(s, w_1, w_2, x_1, x_2, y) = p(s)p(w_1, w_2)p(x_1 \mid s, w_1)p(x_2 \mid s, w_2)p(y \mid x_1, x_2)$.

These two theorems show that in the above problem the source and channel coding separation is not optimal in general and one can achieve a larger rate region using joint source channel coding.

Now, consider the general wireless sensor network of section II.C with condition (II.C.1). For any source $k \in \mathcal{M}$, introduce an auxiliary random variable $U_i^{(k)}$ with cardinality equal to $\|\mathcal{X}_i\|$ for each node $i \in \mathcal{N}^{(k)}$ which stands for the information of source *k* in node *i*, like [32]. For any node $i \in \mathcal{N}$, denote by $\mathcal{M}_i := \{k : i \in \mathcal{N}^{(k)}\}$ the set of all sources with the multi-relay route passing through node *i*. For any $\mathcal{S} \subseteq \mathcal{M}_i$ we define:

$$\mathcal{U}^{(k)} = \{U_j^{(k)} : j \in \mathcal{N}^{(k)}\},$$

$$\mathcal{U}_{i-}^{(k)} = \{U_j^{(k)} : j \prec i(k), j \in \mathcal{N}^{(k)}\},$$
$$\mathcal{U}_{i+}^{(k)} = \{U_j^{(k)} : i \prec j(k), j \in \mathcal{N}^{(k)}\},$$

$$\mathcal{U}_i^{(\mathfrak{J})} = \left\{ U_i^{(k)} : k \in \mathfrak{J} \right\},$$
$$\mathcal{U}^{(\mathfrak{J})} = \bigcup_{k \in \mathfrak{J}} \mathcal{U}^{(k)},$$
$$\mathcal{U}_{i-}^{(\mathfrak{J})} = \bigcup_{k \in \mathfrak{J}} \mathcal{U}_{i-}^{k},$$
$$\mathcal{U}_{i+}^{(\mathfrak{J})} = \bigcup_{k \in \mathfrak{J}} \mathcal{U}_{i+}^{k},$$

$$\mathfrak{J}^C = \mathcal{M}_i \setminus \mathfrak{J} .$$

Then we have the following theorem along the m multi-relay routs by a join backward decoding scheme:

Theorem 3**:**

For the system a source

$$(W_1^N, W_2^N, ..., W_m^n) \sim \prod_n p(w_{1n}, w_{2n}, ..., w_{mn})$$

Can be sent with arbitrarily small probability of error over a wireless sensor network

$$\left(\mathcal{X}_1 \times ... \times \mathcal{X}_L, p(y_1, ..., y_L \mid x_1, ..., x_L), \mathcal{Y}_1 \times ... \times \mathcal{Y}_L \right),$$

With allowed codes

$$x_i(u_i^{(k)}, k \in \mathcal{M}_i), \quad i \in \mathcal{N},$$

If there exist probability mass functions

$$\prod_{k \in \mathcal{M}} p(u_j^{(k)}, j \in \mathcal{N}^{(k)}), \; p(x_i \mid u_i^{(k)}, k \in \mathcal{M}_i), \quad i \in \mathcal{N},$$

Such that for any node $i \in \mathcal{N}$ and any $\mathfrak{J} \subseteq \mathcal{M}_i$,

$$H(W^{(\mathfrak{J})} \mid W^{(\mathfrak{J}^C)}) \leq I(\mathcal{U}_{i-}^{(\mathfrak{J})}; Y_i \mid \mathcal{U}_i^{(\mathfrak{J})}, \mathcal{U}_{i+}^{(\mathfrak{J})}, \mathcal{U}^{(\mathfrak{J}^C)}, W^{(\mathfrak{J}^C)}), \quad (5)$$

Where

$$p(w_1,...,w_m,x_1,...,x_L,y_1,...,y_L) = p(w_1,...,w_m).\prod_i p(x_i \mid \mathscr{M}_i).p(y_1,...,y_L \mid x_1,...,x_L)\,.$$

Example 1:

consider a simple wireless sensor network topology depicted in Fig 2 where three source nodes s_1, s_2, and s_3, desire to send three correlated sources W_1, W_2, and W_3, respectively, to the same destination d via the relays r_1, r_2, and r_3. By the theorem 3, the sources can be sent with arbitrarily small probability of error over the wireless sensor network, if there exists some product distribution

$$p(u_{s_1}^{(1)},u_{r_1}^{(1)},u_d^{(1)})\,p(u_{s_2}^{(2)},u_{r_2}^{(2)},u_{r_3}^{(2)},u_d^{(2)})\,p(u_{s_3}^{(3)},u_{r_3}^{(3)},u_d^{(3)})$$

And some functions

$$x_{s_1}(u_{s_1}^{(1)}),\ x_{s_2}(u_{s_2}^{(2)}),\ x_{s_3}(u_{s_3}^{(3)}),$$

$$x_{r_1}(u_{r_1}^{(1)}),\ x_{r_2}(u_{r_2}^{(2)}),$$

$$x_{r_3}(u_{r_3}^{(2)},u_{r_3}^{(3)}),$$

$$x_d(u_d^{(1)},u_d^{(2)},u_d^{(3)}),$$

Such that for node r_1,

$$H(W_1) < I(U_{s_1}^{(1)};Y_{r_1} \mid U_{r_1}^{(1)},U_d^{(1)});$$

For node r_2,

$$H(W_2) < I(U_{s_2}^{(2)};Y_{r_2} \mid U_{r_2}^{(2)},U_{r_3}^{(2)},U_d^{(2)});$$

For node r_3,

$$\begin{cases} H(W_2 \mid W_3) < I(U_{s_2}^{(2)},U_{r_2}^{(2)};Y_{r_3} \mid U_{r_3}^{(2)},U_d^{(2)},U_{s_3}^{(3)},U_{r_3}^{(3)},U_d^{(3)},W_3) \\ H(W_3 \mid W_2) < I(U_{s_3}^{(3)};Y_{r_3} \mid U_{r_3}^{(3)},U_d^{(3)},U_{s_2}^{(2)},U_{r_2}^{(2)},,U_{r_3}^{(2)},U_d^{(2)},W_2); \\ H(W_2,W_3) < I(U_{s_2}^{(2)},U_{r_2}^{(2)},U_{s_3}^{(3)};Y_{r_3} \mid U_{r_3}^{(2)},U_d^{(2)},U_{r_3}^{(3)},U_d^{(3)}) \end{cases}$$

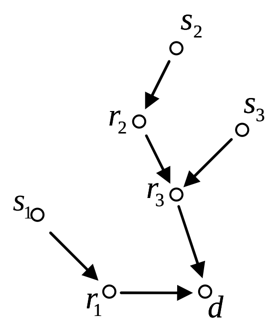

Figure 2. A multi-source, multi-relay, single-destination network.

And for node *d*,

$$\begin{cases} H(W_1 \mid W_2, W_3) < I(U_{s_1}^{(1)}, U_{r_1}^{(1)}; Y_d \mid U_d^{(1)}, U_{s_2}^{(2)}, U_{r_2}^{(2)}, U_{r_3}^{(2)}, U_d^{(2)}, U_{s_3}^{(3)}, U_{r_3}^{(3)}, U_d^{(3)}, W_2, W_3) \\ H(W_2 \mid W_1, W_3) < I(U_{s_2}^{(2)}, U_{r_2}^{(2)}, U_{r_3}^{(2)}; Y_d \mid U_d^{(2)}, U_{s_1}^{(1)}, U_{r_1}^{(1)}, U_d^{(1)}, U_{s_3}^{(3)}, U_{r_3}^{(3)}, U_d^{(3)}, W_1, W_3) \\ H(W_3 \mid W_1, W_2) < I(U_{s_3}^{(3)}, U_{r_3}^{(3)}; Y_d \mid U_d^{(3)}, U_{s_1}^{(1)}, U_{r_1}^{(1)}, U_d^{(1)}, U_{s_2}^{(2)}, U_{r_2}^{(2)}, U_{r_3}^{(2)}, U_d^{(2)}, W_1, W_2) \\ H(W_1, W_2 \mid W_3) < I(U_{s_1}^{(1)}, U_{r_1}^{(1)}, U_{s_2}^{(2)}, U_{r_2}^{(2)}, U_{r_3}^{(2)}; Y_d \mid U_d^{(1)}, U_d^{(2)}, U_{s_3}^{(3)}, U_{r_3}^{(3)}, U_d^{(3)}, W_3) \\ H(W_1, W_3 \mid W_2) < I(U_{s_1}^{(1)}, U_{r_1}^{(1)}, U_{s_3}^{(3)}, U_{r_3}^{(3)}; Y_d \mid U_d^{(1)}, U_d^{(3)}, U_{s_2}^{(2)}, U_{r_2}^{(2)}, U_{r_3}^{(2)}, U_d^{(2)}, W_2) \\ H(W_1, W_2, W_3) < I(U_{s_1}^{(1)}, U_{r_1}^{(1)}, U_{s_2}^{(2)}, U_{r_2}^{(2)}, U_{r_3}^{(2)}, U_{s_3}^{(3)}, U_{r_3}^{(3)}; Y_d \mid U_d^{(1)}, U_d^{(2)}, U_d^{(3)}) \end{cases};$$

This region is larger than the region that can be achieved by separate source and channel coding in the scheme of [32].

2.3-Nothing but Proofs:

2.3-1.Proof of theorem 1:

We consider any fixed $p(w_1, w_2, x_1, x_2, x_3, x_4)$. Consider the following marginal and conditional probability mass functions obtained from $p(w_1, w_2, x_1, x_2, x_3, x_4)$ and $p(y_3, y_4, y_5 \mid x_1, x_2, x_3, x_4)$:

$p(w_1, w_2)$, $p(x_3 \mid w_1)$, $p(x_1 \mid x_3, w_1)$, $p(x_4 \mid w_2)$, $p(x_2 \mid x_4, w_2)$, $p(y_3 \mid x_1, x_3)$, $p(y_4 \mid x_2, x_4)$ and $p(y_5 \mid x_1, x_2, x_3, x_4)$.

Consider B blocks of transmission, each of N transmission slots. Sequences $\mathrm{w}_1^N(b) \in \{1, 2, ..., 2^{NH(W_1)}\}$ and $\mathrm{w}_2^N(b) \in \{1, 2, ..., 2^{NH(W_2)}\}$, $b = 1, 2, ..., B-1$ will be sent in nB transmission slots. Thus, the rate $\frac{NR_1(B-1)}{nB}$ and $\frac{NR_2(B-1)}{nB}$ will be achieved and as $B \to \infty$ they are arbitrarily close to R_1 and R_2 for any N.

2.3.1.1-Random code construction:

At the encoder, regular encoding is used.

1) Fix $p(x_3 \mid w_1)$ and $p(x_4 \mid w_2)$; for each $\mathrm{w}_1^N \in \mathcal{W}_1^o$ generate one x_3^N sequence according to $\prod_{n=1}^{N} p(x_{3n} \mid w_{1n})$ and for each $\mathrm{w}_2^N \in \mathcal{W}_2^o$ generate one x_4^N sequence according to $\prod_{n=1}^{N} p(x_{4n} \mid w_{2n})$. Call these sequences $\mathrm{x}_3^N(\mathrm{w}_1^N)$ and $\mathrm{x}_4^N(\mathrm{w}_2^N)$, respectively.

2) Fix $p(x_1 \mid x_3, w_1)$ and $p(x_2 \mid x_4, w_2)$; for each $\mathrm{w}_1^N \in \mathcal{W}_1^o$ and x_3^N generate one x_1^N sequence according to $\prod_{n=1}^{N} p(x_{1n} \mid x_{3n}, w_{1n})$ and for each $\mathrm{w}_2^N \in \mathcal{W}_2^o$ and x_4^N generate one x_2^N sequence according to $\prod_{n=1}^{N} p(x_{2n} \mid x_{4n}, w_{2n})$. Call these sequences $\mathrm{x}_1^N(\mathrm{w}_1^{N(1)} \mid \mathrm{w}_1^{N(3)})$ and $\mathrm{x}_2^N(\mathrm{w}_2^{N(2)} \mid \mathrm{w}_2^{N(4)})$, respectively. Then all codebooks are revealed to all parties.

***2.3.2-Encoding*:**

At the beginning of each block $b = 1, 2, \ldots, B-1$, node 3 has an estimate $\hat{\mathrm{w}}_1^{N(3)}(b-1)$ of $\hat{\mathrm{w}}_1^{N(1)}(b-1) = \mathrm{w}_1^N(b-1)$, and sends $\mathrm{x}_3^N(b) = \mathrm{x}_3^N\left(\hat{\mathrm{w}}_1^{N(3)}(b-1)\right)$. In the same block, node 1 sends $\mathrm{x}_1^N(b) = \mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(1)}(b) \mid \hat{\mathrm{w}}_1^{N(1)}(b-1)\right)$. We encode the source *W₂* with the same algorithm and $\mathrm{x}_4^N(b) = \mathrm{x}_4^N\left(\hat{\mathrm{w}}_2^{N(4)}(b-1)\right)$ and $\mathrm{x}_2^N(b) = \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(2)}(b) \mid \hat{\mathrm{w}}_2^{N(2)}(b-1)\right)$ will be sent over the channel. For the synchronization of all the nodes at the initial time, we set $\hat{\mathrm{w}}_i^{N(j)}(b_1) = \mathrm{w}_i^N(b_1) = 1$ for every $b_1 \le 0$ or $b_1 \ge B, i \in \{0,1\}, j \in \{1,2,3,4,5\}$.

Every node $k \in \{2,3,4,5\}$ will receive $\mathbf{Y}_k^N(b) = \mathbf{Y}_k^N\left(\mathbf{X}_1^N(b), \mathbf{X}_2^N(b), \mathbf{X}_3^N(b), \mathbf{X}_4^N(b)\right)$ with probability

$$\Pr\left\{\mathbf{Y}_k^N(b) \mid \mathbf{X}_1^N(b), \mathbf{X}_2^N(b), \mathbf{X}_3^N(b), \mathbf{X}_4^N(b)\right\} = \prod_{n=1}^{N} p(y_{kn}(b) \mid x_{1n}(b), x_{2n}(b), x_{3n}(b), x_{4n}(b))$$

***2.3.3-Decoding*:**

At the end of each block $b = 1, 2, ..., B-1$, decoding at nodes 3 and 4 happen simultaneously, but independently. Node 3 declares $\hat{\mathrm{w}}_1^{N(3)}(b) = \mathrm{w}_1^N$ if w_1^N is the unique value in $\left\{1, 2, ..., 2^{NH(W_1)}\right\}$ such that in the block *b*

$$\left\{\mathrm{x}_1^N\left(\mathrm{w}_1^N \mid \hat{\mathrm{w}}_1^{N(3)}(b-1)\right), \mathrm{x}_3^N\left(\hat{\mathrm{w}}_1^{N(3)}(b-1)\right), \mathrm{y}_3^N(b)\right\} \in A_\varepsilon^N(X_1, X_3, Y_3)$$

Otherwise, if no unique w_1^N as above exist, an error will declared with $\hat{\mathrm{w}}_1^{N(3)}(b) = 0$. Node 4 decodes $\hat{\mathrm{w}}_2^{N(4)}(b) = \mathrm{w}_2^N$ in the same way.

The destination node 5 does not commence decoding until the end of block *B* (backward decoding). First, it decodes $\mathrm{w}_1^N(b-1)$ and $\mathrm{w}_2^N(b-1)$, based on $\mathrm{y}_5^N(B)$. It declares $\left(\mathrm{w}_1^N(b-1), \mathrm{w}_2^N(b-1)\right) = \left(\mathrm{w}_1^N, \mathrm{w}_2^N\right)$ if $\left(\mathrm{w}_1^N, \mathrm{w}_2^N\right)$ is the unique pare in $\left\{1, 2, ..., 2^{NH(W_1)}\right\}$ and $\left\{1, 2, ..., 2^{NH(W_2)}\right\}$ such that in the block *B*

$$\begin{aligned}\{&\hat{\mathrm{w}}_1^{N(5)}(B-1), \hat{\mathrm{w}}_2^{N(5)}(B-1), \mathrm{x}_1^N\left(1 \mid \hat{\mathrm{w}}_1^{N(5)}(B-1)\right), \mathrm{x}_2^N\left(1 \mid \hat{\mathrm{w}}_2^{N(5)}(B-1)\right), \\ &\mathrm{x}_3^N\left(\hat{\mathrm{w}}_1^{N(5)}(B-1)\right), \mathrm{x}_4^N\left(\hat{\mathrm{w}}_2^{N(5)}(B-1)\right), \mathrm{y}_5^N(B)\} \\ &\in A_\varepsilon^N(W_1, W_2, X_1, X_2, X_3, X_4, Y_5)\end{aligned}$$

according to

$$p(w_1, w_2)\, p(x_1, x_3 \mid w_1)\, p(x_2, x_4 \mid w_2)\, p(y_5 \mid x_1, x_2, x_3, x_4)$$

Then recursively for $b = B-1, ..., 2$, node 5, which knows $\left(\hat{\mathrm{w}}_1^{N(5)}(b), \hat{\mathrm{w}}_2^{N(5)}(b)\right)$, decodes $\left(\hat{\mathrm{w}}_1^{N(5)}(b-1), \hat{\mathrm{w}}_2^{N(5)}(b-1)\right)$ according to

$$\begin{aligned}\{&\hat{\mathrm{w}}_1^{N(5)}(b-1), \hat{\mathrm{w}}_2^{N(5)}(b-1), \mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \hat{\mathrm{w}}_1^{N(5)}(b-1)\right), \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \hat{\mathrm{w}}_2^{N(5)}(b-1)\right), \\ &\mathrm{x}_3^N\left(\hat{\mathrm{w}}_1^{N(5)}(b-1)\right), \mathrm{x}_4^N\left(\hat{\mathrm{w}}_2^{N(5)}(b-1)\right), \mathrm{y}_5^N(b)\} \\ &\in A_\varepsilon^N(W_1, W_2, X_1, X_2, X_3, X_4, Y_5)\end{aligned}$$

***2.3.4-Probability of error*:**

Error can be occurred in the relays or the destination.

1) At the relays: Consider node 1 sends $\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(1)}(b)\,|\,\hat{\mathrm{w}}_1^{N(3)}(b-1)\right)$ at block *b*. Thus,

$$\Pr\left\{\left\{\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(1)}(b)\,|\,\hat{\mathrm{w}}_1^{N(3)}(b-1)\right),\mathrm{x}_3^N\left(\hat{\mathrm{w}}_1^{N(3)}(b-1)\right),\mathrm{y}_3^N(b)\right\}\notin A_\varepsilon^N(X_1,X_3,Y_3)\right\}<\varepsilon$$

According to typicality properties. Also,

$$\begin{aligned}&\Pr\left\{\exists \mathrm{w}_1^N\neq\hat{\mathrm{w}}_1^{N(1)}(b):\left\{\mathrm{x}_1^N\left(\mathrm{w}_1^N\,|\,\hat{\mathrm{w}}_1^{N(3)}(b-1)\right),\mathrm{x}_3^N\left(\hat{\mathrm{w}}_1^{N(3)}(b-1)\right),\mathrm{y}_3^N(b)\right\}\in A_\varepsilon^N(X_1,X_3,Y_3)\right\}\\&<\sum_{i=1}^{2^{NH(W_1)}}\Pr\left\{\left\{\mathrm{x}_1^N\left(\mathrm{w}_1^N\,|\,\hat{\mathrm{w}}_1^{N(3)}(b-1)\right),\mathrm{x}_3^N\left(\hat{\mathrm{w}}_1^{N(3)}(b-1)\right),\mathrm{y}_3^N(b)\right\}\in A_\varepsilon^N(X_1,X_3,Y_3)\right\}\\&<2^{NH(W_1)}.2^{-N\left(I(X_1,Y_3|X_3)-6\varepsilon\right)}\end{aligned}$$

According to typicality properties of section II.B. Thus,

$P_{e_3}<\varepsilon+2^{N(H(W_1)-I(X_1;Y_3|X_3)+6\varepsilon)}$, where P_{e_3} is the probability of the decoding error at node 3.

Similarly, we have

$P_{e_4}<\varepsilon+2^{N(H(W_2)-I(X_2;Y_4|X_4)+6\varepsilon)}$, where P_{e_4} is the probability of the decoding error at node 4.

2) At the destination: Consider node 1 sends $\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(1)}(b)\,|\,\hat{\mathrm{w}}_1^{N(3)}(b-1)\right)$ and node 2 sends $\mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(2)}(b)\,|\,\hat{\mathrm{w}}_2^{N(4)}(b-1)\right)$ at block *b* and the decoder already knows $\hat{\mathrm{w}}_1^{N(5)}(b)$ and $\hat{\mathrm{w}}_2^{N(5)}(b)$. Thus,

$$\Pr\left\{\begin{array}{r}\{\hat{\mathrm{w}}_1^{N(1)}(b-1),\hat{\mathrm{w}}_2^{N(2)}(b-1),\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b)\,|\,\hat{\mathrm{w}}_1^{N(1)}(b-1)\right),\mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b)\,|\,\hat{\mathrm{w}}_2^{N(2)}(b-1)\right),\\ \mathrm{x}_3^N\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right),\mathrm{x}_4^N\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\right),\mathrm{y}_5^N(b)\}\\ \notin A_\varepsilon^N(W_1,W_2,X_1,X_2,X_3,X_4,Y_5)\end{array}\right\}<\varepsilon$$

According to typicality properties. Also,

$$\Pr\left\{\begin{array}{l}\exists(\mathrm{w}_1^N,\mathrm{w}_2^N)\neq\left(\hat{\mathrm{w}}_1^{N(1)}(b-1),\hat{\mathrm{w}}_2^{N(2)}(b-1)\right)\ : \\ \{\mathrm{w}_1^N,\mathrm{w}_2^N,\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b)\,|\,\mathrm{w}_1^N\right),\mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b)\,|\,\mathrm{w}_2^N\right), \\ \qquad\qquad \mathrm{x}_3^N\left(\mathrm{w}_1^N\right),\mathrm{x}_4^N\left(\mathrm{w}_2^N\right),\mathrm{y}_5^N(b)\} \\ \qquad\qquad\qquad \in A_\varepsilon^N(W_1,W_2,X_1,X_2,X_3,X_4,Y_5)\end{array}\right\}$$

$$< \sum_{\left(\hat{\mathrm{w}}_1^{N(5)},\hat{\mathrm{w}}_2^{N(5)}\right)\in A_\varepsilon^N} p\left(\hat{\mathrm{w}}_1^{N(5)},\hat{\mathrm{w}}_2^{N(5)}\right). \sum_{\substack{\mathrm{w}_1^N\neq\hat{\mathrm{w}}_1^{N(5)}\\ \mathrm{w}_2^N=\hat{\mathrm{w}}_2^{N(5)}}} \Pr\left\{\begin{array}{r}\{\mathrm{w}_1^N,\mathrm{w}_2^N,\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b)\,|\,\mathrm{w}_1^N\right),\mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b)\,|\,\mathrm{w}_2^N\right), \\ \mathrm{x}_3^N\left(\mathrm{w}_1^N\right),\mathrm{x}_4^N\left(\mathrm{w}_2^N\right),\mathrm{y}_5^N(b)\} \\ \in A_\varepsilon^N(W_1,W_2,X_1,X_2,X_3,X_4,Y_5)\end{array}\right\}$$

$$+ \sum_{\left(\hat{\mathrm{w}}_1^{N(5)},\hat{\mathrm{w}}_2^{N(5)}\right)\in A_\varepsilon^N} p\left(\hat{\mathrm{w}}_1^{N(5)},\hat{\mathrm{w}}_2^{N(5)}\right). \sum_{\substack{\mathrm{w}_1^N=\hat{\mathrm{w}}_1^{N(5)}\\ \mathrm{w}_2^N\neq\hat{\mathrm{w}}_2^{N(5)}}} \Pr\{.\}$$

$$+ \sum_{\left(\hat{\mathrm{w}}_1^{N(5)},\hat{\mathrm{w}}_2^{N(5)}\right)\in A_\varepsilon^N} p\left(\hat{\mathrm{w}}_1^{N(5)},\hat{\mathrm{w}}_2^{N(5)}\right). \sum_{\substack{\mathrm{w}_1^N\neq\hat{\mathrm{w}}_1^{N(5)}\\ \mathrm{w}_2^N\neq\hat{\mathrm{w}}_2^{N(5)}}} \Pr\{.\}$$

$$\begin{array}{l}\leq 2^{N\left(H(W_1|W_2)+\varepsilon\right)}.2^{-N\left(I\left(X_1,X_3;Y_5|X_2,X_4,W_2\right)-6\varepsilon\right)} \\ +2^{N\left(H(W_2|W_1)+\varepsilon\right)}.2^{-N\left(I\left(X_2,X_4;Y_5|X_1,X_3,W_1\right)-6\varepsilon\right)} \\ +2^{NH(W_1,W_2)}.2^{-N\left(I\left(X_1,X_2,X_3,X_4;Y_5\right)-\varepsilon\right)}\end{array}.$$

Thus,

$$\begin{array}{l}P_{e_5}\leq\varepsilon+2^{N\left(H(W_1|W_2)-I\left(X_1,X_3;Y_5|X_2,X_4,W_2\right)+5\varepsilon\right)} \\ +2^{N\left(H(W_2|W_1)-I\left(X_2,X_4;Y_5|X_1,X_3,W_1\right)+5\varepsilon\right)} \\ +2^{N\left(H(W_1,W_2)-I\left(X_1,X_2,X_3,X_4;Y_5\right)+\varepsilon\right)}\end{array}$$

where P_{e_5} is the probability of the decoding error at node 5.

Thus, we have

$$\begin{array}{l}P_e\leq P_{e_3}+P_{e_4}+P_{e_5}\leq \\ 3\varepsilon+2^{N(H(W_1)-I(X_1;Y_3|X_3)+6\varepsilon)} \\ +2^{N(H(W_2)-I(X_2;Y_4|X_4)+6\varepsilon)} \\ +2^{N\left(H(W_1|W_2)-I\left(X_1,X_3;Y_5|X_2,X_4,W_2\right)+5\varepsilon\right)} \\ +2^{N\left(H(W_2|W_1)-I\left(X_2,X_4;Y_5|X_1,X_3,W_1\right)+5\varepsilon\right)} \\ +2^{N\left(H(W_1,W_2)-I\left(X_1,X_2,X_3,X_4;Y_5\right)+\varepsilon\right)}\end{array}$$

Where P_e is the probability of the decoding error.

Consequently, $P_e \to 0$ if

$$\begin{cases} H(W_1) < I(X_1;Y_3 \mid X_3) \\ H(W_2) < I(X_2,Y_4 \mid X_4) \\ H(W_1 \mid W_2) < I(X_1,X_3;Y_5 \mid X_2,X_4,W_2) \\ H(W_2 \mid W_1) < I(X_2,X_4;Y_5 \mid X_1,X_3,W_1) \\ H(W_1,W_2) < I(X_1,X_2,X_3,X_4;Y_5) \end{cases}$$

And this completes the proof of theorem 1.

2.4-Proof of theorem 2:

2.4.1-Random code construction:

1) Fix $p(r)$; for each $\mathrm{w}_c^N \in \mathcal{W}_c^o$ generate one r^N sequence according to $\prod_{n=1}^{N} p(r_n)$ and index them by $\mathrm{r}^N(\mathrm{w}_c^N)$, $\mathrm{w}_c^N \in \mathcal{W}_c^o$.

2) For each $\mathrm{w}_1^N \in \mathcal{W}_1^o$ find the corresponding $\mathrm{w}_c^N = f(\mathrm{w}_1^N) = \left(f(\mathrm{w}_{11}),\dots,f(\mathrm{w}_{1n})\right)$. Fix $p(x_3 \mid r, w_1)$ and $p(x_4 \mid r, w_2)$; independently generate one x_3^N sequence according to $\prod_{n=1}^{N} p(x_{3n} \mid w_{1n}, r_n(w_c))$ and index it by $\mathrm{x}_3^N\left(\mathrm{w}_1^N \mid \mathrm{r}^N\left(f\left(\mathrm{w}_1^N\right)\right)\right)$ or $\mathrm{x}_3^N\left(\mathrm{w}_1^N \mid \mathrm{r}^N\right)$, $\mathrm{w}_1^N \in \mathcal{W}_1^o$, $\mathrm{r}^N \in \mathcal{R}$ and $\mathrm{r}^N = \mathrm{r}^N\left(f\left(\mathrm{w}_1^N\right)\right)$ as generated in 1). We do the same procedure for each $\mathrm{w}_2^N \in \mathcal{W}_2^o$ sequences using $\prod_{n=1}^{N} p(x_{4n} \mid w_{2n}, r_n(w_c))$. We index them by $\mathrm{x}_4^N\left(\mathrm{w}_2^N \mid \mathrm{r}^N\left(g\left(\mathrm{w}_2^N\right)\right)\right)$ or $\mathrm{x}_4^N\left(\mathrm{w}_2^N \mid \mathrm{r}^N\right)$, $\mathrm{w}_2^N \in \mathcal{W}_2^o$, $\mathrm{r}^N \in \mathcal{R}$ and $\mathrm{r}^N = \mathrm{r}^N\left(g\left(\mathrm{w}_2^N\right)\right)$.

3) Fix $p(x_1 \mid x_3, w_1, r)$ and $p(x_2 \mid x_4, w_2, r)$; for each $\mathrm{w}_1^N \in \mathcal{W}_1^o$ and x_3^N generate one x_1^N sequence according to $\prod_{n=1}^{N} p(x_{1n} \mid x_{3n}, w_{1n}, r_n(w_c))$ and for each $\mathrm{w}_2^N \in \mathcal{W}_2^o$ and x_4^N

generate one x_2^N sequence according to $\prod_{n=1}^{N} p(x_{2n} \mid x_{4n}, w_{2n}, r_n(w_c))$. Call these sequences $\mathrm{x}_1^N(\mathrm{w}_1^{N(1)} \mid \mathrm{w}_1^{N(3)}, \mathrm{r}^{N(1)})$ and $\mathrm{x}_2^N(\mathrm{w}_2^{N(2)} \mid \mathrm{w}_2^{N(4)}, \mathrm{r}^{N(2)})$, respectively. Then all codebooks are revealed to all parties.

***2.4.2-Encoding*:**

At the beginning of each block $b = 1, 2, ..., B-1$, node 3 has an estimate $\hat{\mathrm{w}}_1^{N(3)}(b-1)$ of $\hat{\mathrm{w}}_1^{N(1)}(b-1) = \mathrm{w}_1^N(b-1)$, it also finds $\hat{\mathrm{r}}^{N(3)}(b-1) = f\left(\hat{\mathrm{w}}_1^{N(3)}(b-1)\right)$ and sends $\mathrm{x}_3^N(b) = \mathrm{x}_3^N\left(\hat{\mathrm{w}}_1^{N(3)}(b-1) \mid \hat{\mathrm{r}}^{N(3)}(b-1)\right)$. In the same block, node 1 sends $\mathrm{x}_1^N(b) = \mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(1)}(b) \mid \hat{\mathrm{w}}_1^{N(1)}(b-1), \hat{\mathrm{r}}^{N(1)}(b)\right)$, where $\hat{\mathrm{r}}^{N(1)}(b-1) = f(\hat{\mathrm{w}}_1^{N(1)}(b-1))$. We encode the source W_2 with the same algorithm and $\mathrm{x}_4^N(b) = \mathrm{x}_4^N\left(\hat{\mathrm{w}}_2^{N(4)}(b-1) \mid \hat{\mathrm{r}}^{N(4)}(b-1)\right)$ and $\mathrm{x}_2^N(b) = \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(2)}(b) \mid \hat{\mathrm{w}}_2^{N(2)}(b-1), \hat{\mathrm{r}}^{N(2)}(b)\right)$ will be sent over the channel, where $\hat{\mathrm{r}}^{N(i)}(b-1) = g(\hat{\mathrm{w}}_2^{N(i)}(b-1)),\ i = 2, 4$. For the synchronization of all the nodes at the initial time, we set $\hat{\mathrm{w}}_i^{N(j)}(b_1) = \mathrm{w}_i^N(b_1) = 1$ for every $b_1 \le 0$ or $b_1 \ge B, i \in \{0,1\}, j \in \{1,2,3,4,5\}$.

Every node $k \in \{2,3,4,5\}$ will receive $\mathbf{Y}_k^N(b) = \mathbf{Y}_k^N\left(\mathbf{X}_1^N(b), \mathbf{X}_2^N(b), \mathbf{X}_3^N(b), \mathbf{X}_4^N(b)\right)$ with probability

$$\Pr\left\{\mathbf{Y}_k^N(b) \mid \mathbf{X}_1^N(b), \mathbf{X}_2^N(b), \mathbf{X}_3^N(b), \mathbf{X}_4^N(b)\right\} = \prod_{n=1}^{N} p(y_{kn}(b) \mid x_{1n}(b), x_{2n}(b), x_{3n}(b), x_{4n}(b))$$

***2.4.3-Decoding*:**

At the end of each block $b = 1, 2, ..., B-1$, decoding at nodes 3 and 4 happen simultaneously, but independently. Node 3 declares $\hat{\mathrm{w}}_1^{N(3)}(b) = \mathrm{w}_1^N$ if w_1^N is the unique value in $\left\{1, 2, ..., 2^{NH(W_1)}\right\}$ such that in the block b

$$\left\{ \mathrm{x}_1^N \left(\mathrm{w}_1^N \mid \hat{\mathrm{w}}_1^{N(3)}(b-1), \mathrm{r}^N \right), \mathrm{x}_3^N \left(\hat{\mathrm{w}}_1^{N(3)}(b-1) \mid \mathrm{r}^N \left(f\left(\hat{\mathrm{w}}_1^{N(3)}(b-1) \right) \right) \right), \mathrm{y}_3^N(b) \right\} \in A_\varepsilon^N(X_1, X_3, Y_3)$$

Where $\mathrm{r}^N = \mathrm{r}^N(\mathrm{w}_c^N)$ and $\mathrm{w}_c^N = f(\mathrm{w}_1^N)$. Otherwise, if no unique w_1^N as above exist, an error will declared with $\hat{\mathrm{w}}_1^{N(3)}(b) = 0$. Node 4 decodes $\hat{\mathrm{w}}_2^{N(4)}(b) = \mathrm{w}_2^N$ in the same way.

The destination node 5 does not commence decoding until the end of block B (backward decoding). Consider $\mathrm{w}_c^N = f(\mathrm{w}_1^N) = g(\mathrm{w}_2^N)$ and $\mathrm{r}^N = \mathrm{r}^N(\mathrm{w}_c^N)$. First, it decodes $\mathrm{w}_1^N(b-1)$ and $\mathrm{w}_2^N(b-1)$, based on $\mathrm{y}_5^N(B)$. It declares $\left(\mathrm{w}_1^N(b-1), \mathrm{w}_2^N(b-1)\right) = \left(\mathrm{w}_1^N, \mathrm{w}_2^N\right)$ if $\left(\mathrm{w}_1^N, \mathrm{w}_2^N\right)$ is the unique pare in $\left\{1, 2, ..., 2^{NH(W_1)}\right\}$ and $\left\{1, 2, ..., 2^{NH(W_2)}\right\}$ such that in the block B

$$\begin{aligned} \{&\mathrm{w}_1^N, \mathrm{w}_2^N, \mathrm{w}_c^N, \\ &\mathrm{x}_1^N\left(1 \mid \mathrm{w}_1^N, 1\right), \mathrm{x}_2^N\left(1 \mid \mathrm{w}_2^N, 1\right), \\ &\mathrm{r}^N, \mathrm{x}_3^N\left(\mathrm{w}_1^N \mid \mathrm{r}^N\right), \mathrm{x}_4^N\left(\mathrm{w}_2^N \mid \mathrm{r}^N\right), \\ &\mathrm{y}_5^N(B)\} \in A_\varepsilon^N(W_1, W_2, X_1, X_2, X_3, X_4, Y_5) \end{aligned}$$

According to

$$p(r)p(w_1, w_2)p(x_1, x_3 \mid w_1, r)p(x_2, x_4 \mid w_2, r)p(y_5 \mid x_1, x_2, x_3, x_4)$$

Then recursively for $b = B-1, ..., 2$, node 5, which knows $\left(\hat{\mathrm{w}}_1^{N(5)}(b), \hat{\mathrm{w}}_2^{N(5)}(b)\right)$, decodes $\left(\hat{\mathrm{w}}_1^{N(5)}(b-1), \hat{\mathrm{w}}_2^{N(5)}(b-1)\right)$ according to

$$\begin{aligned} \{&\mathrm{w}_1^N, \mathrm{w}_2^N, \mathrm{w}_c^N, \\ &\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \mathrm{w}_1^N, \mathrm{r}^{N(5)}(b)\right), \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \mathrm{w}_2^N, \mathrm{r}^{N(5)}(b)\right), \\ &\mathrm{r}^N, \mathrm{x}_3^N\left(\mathrm{w}_1^N \mid \mathrm{r}^N\right), \mathrm{x}_4^N\left(\mathrm{w}_2^N \mid \mathrm{r}^N\right), \\ &\mathrm{y}_5^N(b)\} \in A_\varepsilon^N(W_1, W_2, X_1, X_2, X_3, X_4, Y_5), \end{aligned}$$

2.4.4- Probability of error:

Error can be occurred in the relays or the destination.

1) At the relays: Consider node 1 sends $\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(1)}(b)\,|\,\hat{\mathrm{w}}_1^{N(3)}(b-1)\right)$ at block *b*. Thus,

$$\Pr\left\{\begin{array}{c}\{\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(1)}(b)\,|\,\hat{\mathrm{w}}_1^{N(3)}(b-1),\mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(1)}(b)\right)\right)\right),\mathrm{x}_3^N\left(\hat{\mathrm{w}}_1^{N(3)}(b-1)\,|\,\mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(3)}(b-1)\right)\right)\right)\\ ,\mathrm{y}_3^N(b)\}\notin A_\varepsilon^N(X_1,X_3,Y_3)\end{array}\right\}<\varepsilon$$

According to typicality properties. Also,

$$\Pr\left\{\begin{array}{l}\exists \mathrm{w}_1^N\neq\hat{\mathrm{w}}_1^{N(1)}(b):\{\mathrm{x}_1^N\left(\mathrm{w}_1^N\,|\,\hat{\mathrm{w}}_1^{N(3)}(b-1),\mathrm{r}^N\right),\\ \mathrm{x}_3^N\left(\hat{\mathrm{w}}_1^{N(3)}(b-1)\,|\,\mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(3)}(b-1)\right)\right)\right),\mathrm{y}_3^N(b)\}\in A_\varepsilon^N(X_1,X_3,Y_3)\end{array}\right\}$$
$$<2^{NH(W_1)}.2^{-N\left(I(X_1,Y_3|X_3)-6\varepsilon\right)}$$

According to typicality properties of section II.B. Thus,

$$P_{e_3}<\varepsilon+2^{N(H(W_1)-I(X_1;Y_3|X_3)+6\varepsilon)},$$

Where P_{e_3} is the probability of the decoding error at node 3.

Similarly, we have

$$P_{e_4}<\varepsilon+2^{N(H(W_2)-I(X_2;Y_4|X_4)+6\varepsilon)},$$

Where P_{e_4} is the probability of the decoding error at node 4.

2) At the destination: Consider node 1 sends $\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(1)}(b)\,|\,\hat{\mathrm{w}}_1^{N(3)}(b-1)\right)$ and node 2 sends $\mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(2)}(b)\,|\,\hat{\mathrm{w}}_2^{N(4)}(b-1)\right)$ at block *b* and the decoder already knows $\hat{\mathrm{w}}_1^{N(5)}(b)$ and $\hat{\mathrm{w}}_2^{N(5)}(b)$. Thus,

$$\Pr\left\{\begin{array}{l}\{\hat{\mathrm{w}}_1^{N(1)}(b-1),\hat{\mathrm{w}}_2^{N(2)}(b-1),f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right),\mathrm{r}^{N(1)}(b-1),\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b)\,|\,\hat{\mathrm{w}}_1^{N(1)}(b-1),\mathrm{r}^{N(1)}(b)\right),\\ \quad\mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b)\,|\,\hat{\mathrm{w}}_2^{N(2)}(b-1),\mathrm{r}^{N(1)}(b)\right),\mathrm{x}_3^N\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\,|\,\mathrm{r}^{N(1)}(b-1)\right),\\ \qquad\qquad\mathrm{x}_4^N\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\,|\,\mathrm{r}^{N(1)}(b-1)\right),\mathrm{y}_5^N(b)\}\\ \qquad\qquad\qquad\notin A_\varepsilon^N(W_1,W_2,X_1,X_2,X_3,X_4,Y_5)\end{array}\right\}<\varepsilon$$

According to typicality properties. Also, we can have the following error events:

$$E1:\ \exists \mathrm{w}_1^N \neq \hat{\mathrm{w}}_1^{N(1)}(b-1):$$
$$\{\mathrm{w}_1^N, \hat{\mathrm{w}}_2^{N(2)}(b-1), g\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\right), \mathrm{r}^N\left(g\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\right)\right),$$
$$\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \mathrm{w}_1^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \hat{\mathrm{w}}_2^{N(2)}(b-1), \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right),$$
$$\mathrm{x}_3^N\left(\mathrm{w}_1^N \mid \mathrm{r}^N\left(g\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\right)\right)\right), \mathrm{x}_4^N\left(\hat{\mathrm{w}}_2^{N(2)}(b-1) \mid \mathrm{r}^N\left(g\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\right)\right)\right), \mathrm{y}_5^N(b)\} \in A_\varepsilon^N$$

$$E2:\ \exists \mathrm{w}_2^N \neq \hat{\mathrm{w}}_2^{N(2)}(b-1)\ :$$
$$\{\hat{\mathrm{w}}_1^{N(1)}(b-1), \mathrm{w}_2^N, f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right), \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right)\right),$$
$$\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \hat{\mathrm{w}}_1^{N(1)}(b-1), \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \mathrm{w}_2^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right),$$
$$\mathrm{x}_3^N\left(\hat{\mathrm{w}}_1^{N(1)}(b-1) \mid \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right)\right)\right), \mathrm{x}_4^N\left(\mathrm{w}_2^N \mid \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right)\right)\right), \mathrm{y}_5^N(b)\} \in A_\varepsilon^N$$

$$E3:\ \exists(\mathrm{w}_1^N, \mathrm{w}_2^N) \neq \left(\hat{\mathrm{w}}_1^{N(1)}(b-1), \hat{\mathrm{w}}_2^{N(2)}(b-1)\right)\ :$$
$$\{\mathrm{w}_1^N, \mathrm{w}_2^N, f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right), \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right)\right),$$
$$\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \mathrm{w}_1^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \mathrm{w}_2^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right),$$
$$\mathrm{x}_3^N\left(\mathrm{w}_1^N \mid \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right)\right)\right), \mathrm{x}_4^N\left(\mathrm{w}_2^N \mid \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right)\right)\right), \mathrm{y}_5^N(b)\} \in A_\varepsilon^N$$

$$E4:\ \exists\begin{pmatrix} \mathrm{w}_1^N, \mathrm{w}_2^N, \\ \mathrm{w}_c^N, r^N(\mathrm{w}_c^N) \end{pmatrix} \neq \begin{pmatrix} \hat{\mathrm{w}}_1^{N(1)}(b-1), \hat{\mathrm{w}}_2^{N(2)}(b-1), \\ \hat{\mathrm{w}}_c^{N(1)}(b-1), \mathrm{r}^N(\hat{\mathrm{w}}_c^{N(1)}(b-1)) \end{pmatrix}\ :$$
$$\{\mathrm{w}_1^N, \mathrm{w}_2^N, \mathrm{w}_c^N, \mathrm{r}^N\left(\mathrm{w}_c^N\right),$$
$$\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \mathrm{w}_1^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \mathrm{w}_2^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right),$$
$$\mathrm{x}_3^N\left(\mathrm{w}_1^N \mid \mathrm{r}^N\left(\mathrm{w}_c^N\right)\right), \mathrm{x}_4^N\left(\mathrm{w}_2^N \mid \mathrm{r}^N\left(\mathrm{w}_c^N\right)\right), \mathrm{y}_5^N(b)\} \in A_\varepsilon^N$$

$$E5:\ \exists\left(\mathrm{w}_1^N, \mathrm{w}_2^N, \mathrm{w}_c^N\right) \neq \left(\hat{\mathrm{w}}_1^{N(1)}(b-1), \hat{\mathrm{w}}_2^{N(2)}(b-1), \hat{\mathrm{w}}_c^{N(1)}(b-1)\right)\ :$$
$$\{\mathrm{w}_1^N, \mathrm{w}_2^N, \mathrm{w}_c^N, \mathrm{r}^N(\hat{\mathrm{w}}_c^{N(1)}(b-1)),$$
$$\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \mathrm{w}_1^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \mathrm{w}_2^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right),$$
$$\mathrm{x}_3^N\left(\mathrm{w}_1^N \mid \mathrm{r}^N(\hat{\mathrm{w}}_c^{N(1)}(b-1)\right), \mathrm{x}_4^N\left(\mathrm{w}_2^N \mid \mathrm{r}^N(\hat{\mathrm{w}}_c^{N(1)}(b-1)\right), \mathrm{y}_5^N(b)\} \in A_\varepsilon^N$$

Lemma 2**:**

$$\Pr\left\{\bigcup_{i=1}^{5} E_i\right\} \le \sum_{i=1}^{5} \Pr\left\{E_i\right\} \le$$
$$2^{N(H(W_1|W_2)-I(X_1,X_3;Y|X_2,X_4,W_2,R)+10\in)} +$$
$$2^{N(H(W_2|W_1)-I(X_2,X_4;Y|X_1,X_3,W_1,R)+10\in)} +$$
$$2^{N(H(W_1,W_2|W_c)-I(X_1,X_2,X_3,X_4;Y|W_c,R)+10\in)} +$$
$$2^{N(H(W_1,W_2)-I(X_1,X_2,X_3,X_4;Y)+12\in+1/N)}$$

Proof**:** In the ***Appendix II***.

Thus,

$$p_e \le 3\varepsilon +$$
$$2^{N\left(H(W_1)-I\left(X_1;Y_3|X_3\right)+\varepsilon\right)} +$$
$$2^{N\left(H(W_2)-I\left(X_2;Y_4|X_4\right)+\varepsilon\right)} +$$
$$2^{N(H(W_1|W_2)-I(X_1,X_3;Y|X_2,X_4,W_2,R)+10\in)} +$$
$$2^{N(H(W_2|W_1)-I(X_2,X_4;Y|X_1,X_3,W_1,R)+10\in)} +$$
$$2^{N(H(W_1,W_2|W_c)-I(X_1,X_2,X_3,X_4;Y|W_c,R)+10\in)} +$$
$$2^{N(H(W_1,W_2)-I(X_1,X_2,X_3,X_4;Y)+12\in+1/N)}$$

Where P_e is the probability of the decoding error.

Consequently, $P_e \to 0$ if

$$\begin{cases} H(W_1) < I(X_1;Y_3 \mid X_3) \\ H(W_2) < I(X_2,Y_4 \mid X_4) \\ H(W_1 \mid W_2) < I(X_1,X_3;Y_5 \mid X_2,X_4,W_2,S) \\ H(W_2 \mid W_1) < I(X_2,X_4;Y_5 \mid X_1,X_3,W_1,S) \\ H(W_1,W_2 \mid W_c) < I(X_1,X_2,X_3,X_4;Y_5 \mid S,W_c) \\ H(W_1,W_2) < I(X_1,X_2,X_3,X_4;Y_5) \end{cases}$$

And this completes the proof of theorem 2.

2.5-Proof of theorem 3:

The proof is similar to [41] and [32]. We just give the outline of the proof here. Consider the multi-relay scheme of [41]. Still we use the backward decoding at relays. Each node $i \in \{2,...,L\}$ decodes at the end of every B^{i-2} blocks according to [32]. For each node $i \in \{2,...,L\}$ the upstream nodes $1,2,...,i-1$ appear to be cooperating in sending information to it, while the inputs by the downstream nodes are predictable. Thus, if

$$H(W) \le I(X_1,...,X_i;Y_i \mid X_{i+1},...,X_L)$$

we can send the source W over this channel with arbitrarily small probability of error.

According to [32], the decoding time of each node $i \in \{2,...,L\}$ can be scheduled at the end of every B^{j_i} block, as long as j_i, $i \in \{2,...,L\}$ satisfies $0 \le j_2 < j_3 < ... < j_n$. This allows the merging of multiple multi-relay routs. According to (II.C.1), one can assign an integer j_i to each node $i \in \mathscr{N}$ such that these integers are strictly increasing by i. Hence, each node $i \in \mathscr{N}$ decodes its received block at the end of every B^{j_i} block. It decodes the received blocks of different routs at the same time and using backward decoding scheme. Thus, here we have a combination of backward decoding and CDMA with correlated sources like section IV.A. Hence, if

$$H(W^{(\mathfrak{F})} \mid W^{(\mathfrak{F}^C)}) \le I(\mathcal{U}_{i-}^{(\mathfrak{F})};Y_i \mid \mathcal{U}_i^{(\mathfrak{F})},\mathcal{U}_{i+}^{(\mathfrak{F})},\mathcal{U}^{(\mathfrak{F}^C)},W^{(\mathfrak{F}^C)}),$$

The information can be sent over the channel with arbitrarily small probability of error.

Chapter3:

Conclusion and Future Work

3.1-Concusion:

In this thesis we have proven some conditions which some correlated sources need to be sent from some transmitters to some receivers via some relays in a large wireless network reliably (theorem 3). Wireless sensor networks are an important application of such networks. It is shown that the optimality of source-channel separation theorem is not valid in general. The conditions contain a larger region than that of Xie and Kumar's [32] with separate source-channel coding. The codebooks are constructed according to their corresponding sources to utilize the dependency among the sources at the decoder. A combination of backward decoding and CDMA is used at the decoder.

For a simpler channel with 2 source nodes, 2 relays and 1 destination, the sufficient conditions for reliable transmission when sources contain common information have been found (theorem 2).

3.2-Future Work:

After having established coding theorems for the problem of network information flow with correlated sources, a natural question that arises: what if, in a given scenario, $R = 0$? In that case, the best we can hope for is to reconstruct an *approximation* to the original source message, and the answer is given by rate-distortion theory [42]. The rate-distortion formulation of our problem in the case of non-cooperating encoders is equivalent to the well known (and still open) *Multiterminal Source Coding* problem [9]. Although you can find some special case answers but in general scenario we have still a lot to go forward.

APPENDIX I

Proof of Lemma 1:

We have

$$\Pr\left\{\left(z_1^N, z_2^N, Z_3^N, Z_4^N, Z_5^N\right) \in A_\varepsilon^N\right\} = \sum_{\substack{\left(z_3^N, z_4^N, z_5^N\right): \\ \left(z_1^N, z_2^N, z_3^N, z_4^N, z_5^N\right) \in A_\varepsilon^N}} \Pr\left\{\left(Z_3^N, Z_4^N, Z_5^N\right) = \left(z_3^N, z_4^N, z_5^N\right) \mid z_1^N, z_2^N\right\}.$$

But from the assumption of the Lemma 1,

$$\Pr\left\{\left(Z_3^N, Z_4^N, Z_5^N\right) = \left(z_3^N, z_4^N, z_5^N\right) \mid z_1^N, z_2^N\right\} = \Pr\left\{Z_3^N = z_3^N \mid z_1^N, z_2^N\right\}.\Pr\left\{Z_4^N = z_4^N \mid z_3^N, z_2^N\right\}.\Pr\left\{Z_5^N = z_5^N \mid z_3^N, z_1^N\right\}.$$

Now, from the typicality properties we have

$$\Pr\left\{Z_3^N = z_3^N \mid z_1^N, z_2^N\right\} \le 2^{-N\left(H(Z_3|Z_1,Z_2)+2\varepsilon\right)},$$
$$\Pr\left\{Z_4^N = z_4^N \mid z_3^N, z_2^N\right\} \le 2^{-N\left(H(Z_4|Z_3,Z_2)+2\varepsilon\right)},$$
$$\Pr\left\{Z_5^N = z_5^N \mid z_3^N, z_1^N\right\} \le 2^{-N\left(H(Z_5|Z_3,Z_1)+2\varepsilon\right)},$$

Thus we have

$$\Pr\left\{\left(z_1^N, z_2^N, Z_3^N, Z_4^N, Z_5^N\right) \in A_\varepsilon^N\right\} \le 2^{N\left(H(Z_3,Z_4,Z_5|Z_1,Z_2)-2\varepsilon\right)}. 2^{-N\left(H(Z_3|Z_1,Z_2)+2\varepsilon\right)}.2^{-N\left(H(Z_4|Z_3,Z_2)+2\varepsilon\right)}.2^{-N\left(H(Z_5|Z_3,Z_1)+2\varepsilon\right)}.$$

Substituting

$$H\left(Z_3, Z_4, Z_5 \mid Z_1, Z_2\right) = H\left(Z_3 \mid Z_1, Z_2\right) + H\left(Z_4 \mid Z_1, Z_2, Z_3\right) + H\left(Z_5 \mid Z_1, Z_2, Z_3, Z_4\right)$$

Into the previous inequality we have

$$\Pr\left\{\left(z_1^N, z_2^N, Z_3^N, Z_4^N, Z_5^N\right) \in A_\varepsilon^N\right\} \le 2^{-N\left(I(Z_4;Z_1|Z_2,Z_3)+I(Z_5;Z_2,Z_4|Z_1,Z_3)-8\varepsilon\right)}.$$

This completes the proof.

APPENDIX II

***Proof of Lemma 2*:**

Consider node 1 sends $\mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(1)}(b)\,|\,\hat{\mathrm{w}}_1^{N(3)}(b-1)\right)$ and node 2 sends $\mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(2)}(b)\,|\,\hat{\mathrm{w}}_2^{N(4)}(b-1)\right)$ at block b and the decoder already knows $\hat{\mathrm{w}}_1^{N(5)}(b)$ and $\hat{\mathrm{w}}_2^{N(5)}(b)$. This is denoted by the event B. Also consider $\hat{\mathrm{w}}_c^{N(1)} = f(\hat{\mathrm{w}}_1^{N(1)})$

$$\Pr\left\{E_1\,|\,\mathscr{B}\right\} = \sum_{\substack{\mathrm{w}_1^N \neq \hat{\mathrm{w}}_1^{N(1)}(b-1) \\ \left(\mathrm{w}_1^N,\hat{\mathrm{w}}_2^{N(2)}(b-1), f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right)\right)\in A_\varepsilon^N}} \Pr\left\{\begin{array}{l} \{\mathrm{w}_1^N, \hat{\mathrm{w}}_2^{N(2)}(b-1), g\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\right), \mathrm{r}^N\left(g\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\right)\right), \\ \mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b)\,|\,\mathrm{w}_1^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b)\,|\,\hat{\mathrm{w}}_2^{N(2)}(b-1), \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_3^N\left(\mathrm{w}_1^N\,|\,\mathrm{r}^N\left(g\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\right)\right)\right), \\ \mathrm{x}_4^N\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\,|\,\mathrm{r}^N\left(g\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\right)\right)\right), \mathrm{y}_5^N(b)\} \in A_\varepsilon^N \end{array}\right\}$$

but by substituting

$$\mathbf{Z}_1^N = \left(\mathbf{W}_2^N, \hat{\mathrm{w}}_c^{N(1)}\right),\ \mathrm{z}_2^N = \mathrm{w}_1^N,\ \mathbf{Z}_3^N = \mathbf{R}^N\left(\hat{\mathrm{w}}_c^{N(1)}\right),$$
$$\mathbf{Z}_4^N = \mathbf{X}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b)\,|\,\mathrm{w}_1^N, \mathrm{r}^N(\mathrm{w}_1^N)\right),\ \mathbf{Z}_5^N = \left(\mathbf{X}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b)\,|\,\mathrm{w}_2^N, \mathrm{r}^N(\mathrm{w}_2^N)\right), \mathbf{Y}_5^N\right)$$

into the lemma 1 we have

$$\Pr\left\{\begin{array}{l} \{\mathrm{w}_1^N, \hat{\mathrm{w}}_2^{N(2)}(b-1), g\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\right), \mathrm{r}^N\left(g\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\right)\right), \\ \mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b)\,|\,\mathrm{w}_1^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b)\,|\,\hat{\mathrm{w}}_2^{N(2)}(b-1), \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_3^N\left(\mathrm{w}_1^N\,|\,\mathrm{r}^N\left(g\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\right)\right)\right), \\ \mathrm{x}_4^N\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\,|\,\mathrm{r}^N\left(g\left(\hat{\mathrm{w}}_2^{N(2)}(b-1)\right)\right)\right), \mathrm{y}_5^N(b)\} \in A_\varepsilon^N \end{array}\right\} \leq 2^{-N\left(I\left(X_1,X_3;Y_5|X_2,X_4,W_2,R\right)-8\varepsilon\right)}$$

Thus,

$$\Pr\left\{E_1\,|\,\mathscr{B}\right\} \leq 2^{N\left(H(W_1|W_2)-I\left(X_1,X_3;Y_5|X_2,X_4,W_2,R\right)-10\varepsilon\right)}.$$

Similarly

$$\Pr\left\{E_2\,|\,\mathscr{B}\right\} \leq 2^{N\left(H(W_2|W_1)-I\left(X_2,X_4;Y_5|X_1,X_3,W_1,R\right)+10\varepsilon\right)}.$$

$$\Pr\left\{E_3 \mid \mathscr{B}\right\} = \sum_{\substack{\mathrm{w}_1^N \neq \hat{\mathrm{w}}_1^{N(1)}(b-1), \mathrm{w}_2^N \neq \hat{\mathrm{w}}_2^{N(2)}(b-1) \\ \left(\mathrm{w}_1^N, \mathrm{w}_2^N, \hat{\mathrm{w}}_c^{N(1)}\right) \in A_\varepsilon^N}} \Pr\left\{\begin{array}{l} \{\mathrm{w}_1^N, \mathrm{w}_2^N, f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right), \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right)\right), \\ \mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \mathrm{w}_1^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \mathrm{w}_2^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_3^N\left(\mathrm{w}_1^N \mid \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right)\right)\right), \\ \mathrm{x}_4^N\left(\mathrm{w}_2^N \mid \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right)\right)\right), \mathrm{y}_5^N(b)\} \in A_\varepsilon^N \end{array}\right\}$$

but by substituting

$$\mathrm{z}_1^N = \hat{\mathrm{w}}_c^{N(1)},\ \mathrm{z}_2^N = \left(\mathrm{w}_1^N, \mathrm{w}_2^N\right),\ \mathbf{Z}_3^N = \mathbf{R}^N\left(\hat{\mathrm{w}}_c^{N(1)}\right),$$

$$\mathbf{Z}_4^N = \left(\mathbf{X}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \mathrm{w}_1^N, \mathrm{r}^N(\mathrm{w}_c^N)\right), \mathbf{X}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \mathrm{w}_2^N, \mathrm{r}^N(\mathrm{w}_c^N)\right)\right),\ \mathbf{Z}_5^N = \mathbf{Y}_5^N$$

into the lemma 1 we have

$$\Pr\left\{\begin{array}{l} \{\mathrm{w}_1^N, \mathrm{w}_2^N, f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right), \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right)\right), \\ \mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \mathrm{w}_1^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \mathrm{w}_2^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_3^N\left(\mathrm{w}_1^N \mid \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right)\right)\right), \\ \mathrm{x}_4^N\left(\mathrm{w}_2^N \mid \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(1)}(b-1)\right)\right)\right), \mathrm{y}_5^N(b)\} \in A_\varepsilon^N \end{array}\right\} \leq 2^{-N\left(I\left(X_1, X_2 . X_3, X_4; Y_5 \mid W, S\right) - 8\varepsilon\right)}.$$

Thus,

$$\Pr\left\{E_3 \mid \mathscr{B}\right\} \leq 2^{N\left(H\left(W_1, W_2 \mid W_c\right) - I\left(X_1, X_2, X_3, X_4; Y_5 \mid W_c, R\right) + 10\varepsilon\right)}.$$

$$\Pr\left\{E_4 \mid \mathscr{B}\right\} = \sum_{\substack{\mathrm{w}_1^N \neq \hat{\mathrm{w}}_1^{N(1)}(b-1), \mathrm{w}_2^N \neq \hat{\mathrm{w}}_2^{N(2)}(b-1) \\ \left(\mathrm{w}_1^N, \mathrm{w}_2^N, \mathrm{w}_c^N\right) \in A_\varepsilon^N, \mathrm{w}_c^N \neq \hat{\mathrm{w}}_c^{N(1)}}} \Pr\left\{\begin{array}{l} R(\mathrm{w}_c^N) \neq R(\hat{\mathrm{w}}_c^{N(1)}) \,\&\, \{\mathrm{w}_1^N, \mathrm{w}_2^N, \mathrm{w}_c^N, \mathrm{r}^N\left(\mathrm{w}_c^N\right), \\ \mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \mathrm{w}_1^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \mathrm{w}_2^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_3^N\left(\mathrm{w}_1^N \mid \mathrm{r}^N\left(\mathrm{w}_c^N\right)\right), \mathrm{x}_4^N\left(\mathrm{w}_2^N \mid \mathrm{r}^N\left(\mathrm{w}_c^N\right)\right), \mathrm{y}_5^N(b)\} \in A_\varepsilon^N \end{array}\right\}$$

But by substituting

$$\mathrm{z}_1^N = \Phi,\ \mathrm{z}_2^N = \left(\mathrm{w}_1^N, \mathrm{w}_2^N, \mathrm{w}_c^N, \mathrm{r}^{\mathrm{N}}\left(\mathrm{w}_c^N\right)\right),\ \mathbf{Z}_3^N = \Phi,$$

$$\mathbf{Z}_4^N = \left(\mathbf{X}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \mathrm{w}_1^N, \mathrm{r}^N(\mathrm{w}_c^N)\right), \mathbf{X}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \mathrm{w}_2^N, \mathrm{r}^N(\mathrm{w}_c^N)\right)\right),\ \mathbf{Z}_5^N = \mathbf{Y}_5^N$$

Into the lemma 1 we have

$$\Pr\left\{\begin{array}{l} R(\mathrm{w}_c^N) \neq R(\hat{\mathrm{w}}_c^{N(1)}) \,\&\, \{\mathrm{w}_1^N, \mathrm{w}_2^N, \mathrm{w}_c^N, \mathrm{r}^N\left(\mathrm{w}_c^N\right), \\ \mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \mathrm{w}_1^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \mathrm{w}_2^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_3^N\left(\mathrm{w}_1^N \mid \mathrm{r}^N\left(\mathrm{w}_c^N\right)\right), \mathrm{x}_4^N\left(\mathrm{w}_2^N \mid \mathrm{r}^N\left(\mathrm{w}_c^N\right)\right), \mathrm{y}_5^N(b)\} \in A_\varepsilon^N \end{array}\right\} \leq 2^{-N\left(I(X_1, X_2, X_3, X_4; Y_5) - 8\varepsilon\right)}$$

Thus,

$$\Pr\left\{E_4 \mid \mathscr{B}\right\} \leq 2^{N\left(H(W_1, W_2) - I(X_1, X_2, X_3, X_4; Y_5) + 10\varepsilon\right)}.$$

$$\Pr\left\{E_5 \mid \mathscr{B}\right\} = \sum_{\substack{\mathrm{w}_1^N \neq \hat{\mathrm{w}}_1^{N(1)}(b-1), \mathrm{w}_2^N \neq \hat{\mathrm{w}}_2^{N(2)}(b-1) \\ \left(\mathrm{w}_1^N, \mathrm{w}_2^N, \mathrm{w}_c^N\right) \in A_\varepsilon^N, \mathrm{w}_c^N \neq \hat{\mathrm{w}}_c^{N(1)}}} \Pr\left\{\begin{array}{l} R(\mathrm{w}_c^N) = R(\hat{\mathrm{w}}_c^{N(1)}) \,\&\, \{\mathrm{w}_1^N, \mathrm{w}_2^N, \mathrm{w}_c^N, \mathrm{r}^N(\hat{\mathrm{w}}_c^{N(1)}(b-1)), \\ \mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \mathrm{w}_1^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \mathrm{w}_2^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_3^N\left(\mathrm{w}_1^N \mid \mathrm{r}^N(\hat{\mathrm{w}}_c^{N(1)}(b-1)\right), \\ \mathrm{x}_4^N\left(\mathrm{w}_2^N \mid \mathrm{r}^N(\hat{\mathrm{w}}_c^{N(1)}(b-1)\right), \mathrm{y}_5^N(b)\} \in A_\varepsilon^N \end{array}\right\}$$

But by substituting

$$\mathrm{z}_1^N = \mathrm{r}^{\mathrm{N}}\left(\mathrm{w}_c^N\right),\ \mathrm{z}_2^N = \left(\mathrm{w}_1^N, \mathrm{w}_2^N, \mathrm{w}_c^N, \mathrm{r}^{\mathrm{N}}\left(\mathrm{w}_c^N\right)\right),\ \mathbf{Z}_3^N = \Phi,$$
$$\mathbf{Z}_4^N = \left(\mathbf{X}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \mathrm{w}_1^N, \mathrm{r}^N(\mathrm{w}_c^N)\right), \mathbf{X}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \mathrm{w}_2^N, \mathrm{r}^N(\mathrm{w}_c^N)\right)\right),\ \mathbf{Z}_5^N = \mathbf{Y}_5^N$$

Into the lemma 1 we have

$$\Pr\left\{\begin{array}{l} R(\mathrm{w}_c^N) = R(\hat{\mathrm{w}}_c^{N(1)}) \,\&\, \{\mathrm{w}_1^N, \mathrm{w}_2^N, \mathrm{w}_c^N, \mathrm{r}^N(\hat{\mathrm{w}}_c^{N(1)}(b-1)), \\ \mathrm{x}_1^N\left(\hat{\mathrm{w}}_1^{N(5)}(b) \mid \mathrm{w}_1^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_2^N\left(\hat{\mathrm{w}}_2^{N(5)}(b) \mid \mathrm{w}_2^N, \mathrm{r}^N\left(f\left(\hat{\mathrm{w}}_1^{N(5)}(b)\right)\right)\right), \\ \mathrm{x}_3^N\left(\mathrm{w}_1^N \mid \mathrm{r}^N(\hat{\mathrm{w}}_c^{N(1)}(b-1)\right), \\ \mathrm{x}_4^N\left(\mathrm{w}_2^N \mid \mathrm{r}^N(\hat{\mathrm{w}}_c^{N(1)}(b-1)\right), \mathrm{y}_5^N(b)\} \in A_\varepsilon^N \end{array}\right\} \leq 2^{-N\left(I(X_1, X_2, X_3, X_4; Y_5 \mid R) - 8\varepsilon\right)}$$

Thus,

$$\Pr\left\{E_5 \mid \mathscr{B}\right\} \leq 2^{N\left(H(W_1, W_2) - I(X_1, X_2, X_3, X_4; Y_5 \mid R) - H(R) + 12\varepsilon\right)} \leq 2^{N\left(H(W_1, W_2) - I(X_1, X_2, X_3, X_4; Y_5) + 12\varepsilon\right)}.$$

Thus,

$$
\begin{aligned}
&\Pr\left\{\bigcup_{i=1}^{5} E_i\right\} \le \sum_{i=1}^{5} \Pr\left\{E_i\right\} \le \\
&2^{N(H(W_1|W_2)-I(X_1,X_3;Y|X_2,X_4,W_2,R)+10\in)} + \\
&2^{N(H(W_2|W_1)-I(X_2,X_4;Y|X_1,X_3,W_1,R)+10\in)} + \\
&2^{N(H(W_1,W_2|W_c)-I(X_1,X_2,X_3,X_4;Y|W_c,R)+10\in)} + \\
&2^{N(H(W_1,W_2)-I(X_1,X_2,X_3,X_4;Y)+12\in+1/N)}
\end{aligned}
\quad .
$$

This completes the proof of Lemma 2.

REFERENCES

[1] D. Estrin, L. Girod, G. Pottie, M. Srivastava, "Instrumenting the world with wireless sensor networks,"*In Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP 2001),*Salt Lake City, Utah, May 2001.

[2] D. Estrin, R. Govindan, J. Heidemann and S. Kumar, "Next Century Challenges: Scalable Coordination in Sensor Networks," *In Proceedings of the Fifth Annual International Conference on Mobile Computing and Networks (MobiCOM '99)*, August 1999, Seattle, Washington.

[3] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. A Survey on Sensor Networks. *IEEE Communications Mag.*, 40(8):102–114, 2002.

[4] T. Berger, Z. Zhang, and H. Viswanathan. The CEO Problem. *IEEE Trans. Inform. Theory*, 42(3):887–902, 1996.

[5] D. Bertsekas and R. Gallager. *Data Networks (2nd ed)*. Prentice Hall, 1992.

[6] A. Hu and S. D. Servetto. dFSK: *Distributed* Frequency Shift Keying Modulation in Dense Sensor Networks. In *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, 2004.

[7] A. Hu and S. D. Servetto. Algorithmic Aspects of the Time Synchronization Problem in Large-Scale Sensor Networks. *ACM/Kluwer Mobile Networks and Applications*, 10:491–503, 2005. Special issue with selected (and revised) papers from
ACM WSNA 2003.

[8] A. Hu and S. D. Servetto. On the Scalability of Cooperative Time Synchronization in Pulse-Connected Networks. IEEE Trans. Inform. Theory, to appear. Available from http://cn.ece.cornell.edu/.

[9] T. Berger. *The Information Theory Approach to Communications (G. Longo, ed.)*, chapter Multiterminal Source Coding.Springer-Verlag, 1978.

[10] I. Csisz´ar and J. K¨orner. Towards a General Theory of Source Networks. *IEEE Trans. Inform. Theory*, 26(2):155–166, 1980.

[11] J. K¨orner and K. Marton. How to Encode the Modulo-Two Sum of Binary Sources. *IEEE Trans. Inform. Theory*, 25(2):219–221,1979.October 2, 2005. DRAFT 33

[12] V. Kawadia and P. R. Kumar. A Cautionary Perspective on Cross-Layer Design. IEEE Wireless Comm. Mag., 2004.

[13] A. Ephremides and B. Hajek. Information Theory and Communication Networks: An Unconsummated Union. *IEEE Trans. Inform. Theory*, 44(6):2416–2434, 1998.

[14] T. M. Cover and J. Thomas. *Elements of Information Theory*. John Wiley and Sons, Inc., 1991.

[15] D. Slepian and J. K. Wolf. Noiseless Coding of Correlated Information Sources. *IEEE Trans. Inform. Theory*, IT-19(4):471–480, 1973.

[16] T. M. Cover, A. A. El Gamal, and M. Salehi. Multiple Access Channels with Arbitrarily Correlated Sources. *IEEE Trans.Inform. Theory*, IT-26(6):648–657, 1980.

[17] G. Dueck. A Note on the Multiple Access Channel with Correlated Sources. *IEEE Trans. Inform. Theory*, IT-27(2):232–235,1981.

[18] D. Slepian and J. K. Wolf. A Coding Theorem for Multiple Access Channels with Correlated Sources. *Bell Syst. Tech. J.*,52(7):1037–1076, 1973.

[19] R. Ahlswede and T. S. Han. On Source Coding with Side Information via a Multiple-Access Channel, and Related Problems in Multi-User Information Theory. *IEEE Trans. Inform. Theory*, 29(3):396–411, 1983.

[20] F. M. J. Willems. The Discrete Memoryless Multiple Access Channel with Partially Cooperating Encoders. *IEEE Trans. Inform. Theory*, 29(3):441–445, 1983.

[21] T. S. Han. Slepian-Wolf-Cover Theorem for a Network of Channels. *Inform. Contr.*, 47(1):67–83, 1980.

[22] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung. Network Information Flow. *IEEE Trans. Inform. Theory*, 46(4):1204–1216, 2000.

[23] S. Borade. Network Information Flow: Limits and Achievability. In *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Lausanne, Switzerland, 2002.

[24] S.-Y. R. Li, R. W. Yeung, and N. Cai. Linear Network Coding. *IEEE Trans. Inform. Theory*, 49(2):371–381, 2003.

[25] R. Koetter and M. M´edard. An Algebraic Approach to Network Coding. *IEEE/ACM Trans. Networking*, 11(5):782–795, 2003.

[26] M. Effros, M. M´edard, T. Ho, S. Ray, D. Karger, and R. Koetter. Linear Network Codes: A Unified Framework for Source,Channel, and Network Coding. In *Discrete Mathematics and Theoretical Computer Science (DIMACS) series on Network Information Theory*, Piscataway, NJ, 2003.

[27] T. Ho, M. M´edard, M. Effros, and R. Koetter. Network Coding for Correlated Sources. In *Proc. 38th Annual Conf. Inform.Sciences Syst. (CISS)*, Princeton, NJ, March 2004.

[28] Z. Li and B. Li. Network Coding in Undirected Networks. In *Proc. 38th Annual Conf. Inform. Sciences Syst. (CISS)*, Princeton, NJ, March 2004.

[29] P. Gupta and P. R. Kumar. The Capacity of Wireless Networks. *IEEE Trans. Inform. Theory*, 46(2):388–404, 2000.

[30] C. Peraki and S. D. Servetto. On the Maximum Stable Throughput Problem in Random Networks with Directional Antennas. In *Proc. ACM MobiHoc*, Annapolis, MD, 2003.

[31] C. Peraki and S. D. Servetto. Capacity, Stability and Flows in Large-Scale Random Networks. In *Proc. IEEE Inform. Theory Workshop (ITW)*, San Antonio, TX, 2004.

[32] L-L. Xie, P. R. Kumar, "Multi-source, multi-destination, multi-relay wireless networks," *Submitted to IEEE Trans. on Information Theory*, August 2006.

[33] K. De Bruyn, V. V. Prelov, E. C. Vander, "Reliable transmission of two correlated sources over an asymmetric multiple-access channel," *IEEE Trans. on Information Theory*, Vol. IT-33, No. 5, pp. 716-718, Sep. 1987.

[34] T. S. Han, M. H. M. Costa, "Broadcast channels with arbitrarily correlated sources," *IEEE Trans. on Information Theory*, Vol. IT-33, No. 5, pp. 641-650, Sep. 1987.

[35] M. Salehi, "Random multiple access channels with correlated sources,"

[36] M. Salehi, "multiple access channels with correlated sources - coding subject to a fidelity criterion,"

[37] M. Salehi, E. Kurtas, "Interference channels with correlated sources,"

[38] W. Kang, S. Ulukus, "An outer bound for multiple access channels with correlated sources,"

[39] H-F Chong, M. Motani, and H. K. Grag, "Lower bounds to the capacity region of the relay channel,"

[40] H. Witsenhausen, "On sequences of pairs of dependent random variables," *SIAM J. Math*., Vol. 28, pp. 100-113, Jan. 1975.

[41] L-L. Xie and P. R. Kumar, "An achievable rate for the multiple-level relay channel," *IEEE Trans. on Information Theory*, Vol. 51, No. 4, pp. 1348-1358, Apr.

[42] T. Berger. *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Prentice-Hall, Inc., 1971.