Network Coding Theorem for Dynamic Communication Networks

Søren Riis and Maximilien Gadouleau
School of Electronic Engineering and Computer Science
Queen Mary, University of London
Emails: (soren.riis,mgadouleau)@eecs.qmul.ac.uk

***Abstract*—A symbolic approach to communication networks, where the topology of the underlying network is contained in a set of formal terms, was recently introduced. The so-called min-cut of a term set represents its number of degrees of freedom. For any assignment of function symbols, its dispersion measures the amount of information sent to the destinations and its Rényi entropies measure the amount of information that can be inferred about the input from the outputs. It was proved that the maximum dispersion and the maximum Rényi entropy of order less than one asymptotically reach the min-cut of the term set. In this paper, we first generalize the term set framework for multi-user communications and state a multi-user version of the dispersion (and Rényi entropy) theorem. We then model dynamic networks as a collection of term sets and the possible demands of users via a utility function. We apply the multi-user theorem to derive a general principle for many-to-many cast communications in dynamic multi-user networks. In general, we show that if each user's demand can be satisfied locally, then all the demands can be satisfied globally.**

I. Introduction

A multi-user communication problem is given by a network with prescribed sets of sources and of destinations, where the destinations request messages sent by the sources. The general problem of determining whether all the demands of the destinations can be satisfied at the same time (i.e. whether the problem is solvable) is crucial, and as such has been widely studied [1], [2] via techniques from network coding [3], [4].

In [5], a communication problem with a single user was represented via a set of *terms*, which are concepts from logic. A term is built on variables representing the messages sent by the sources, and on function symbols representing coding functions at the intermediate nodes. A term thus formally represents all the possible operations undergone by messages from the sources to the user. The *min-cut* of a term set can be viewed as the number of degrees of freedom of the set and hence represents the information bottlenecks on the network. One measure of information introduced in [5] is the *dispersion*, which is (the logarithm of) the number of possible outputs, once the coding functions have been assigned. Another measure is the one-to-one dispersion, which is (the logarithm of) the number of outputs with exactly one input, i.e. for which perfect decoding is possible. The maximum dispersion and one-to-one dispersion are asymptotically given by the min-cut of the term set. A similar result was also proved for the *Rényi entropy*, which characterizes the amount of information that can be inferred about a uniformly distributed input. The main results derived in [5] are reviewed in Section II.

The first main contribution of this paper is the multi-user max-flow min-cut theorem, derived in Section III. It is obtained by assigning a term set to a multi-user communication problem. This result shows that maximum dispersion equal to the min-cut received by each user can be asymptotically attained simultaneously. In other words, if a dispersion can be achieved locally, i.e. while disregarding the other users, it can be achieved globally, i.e. when the other users have to be accommodated as well. This also holds for the one-to-one dispersion and the Rényi entropy. This result is then applied to multi-user communication problems such as satellite communication (the butterfly network) and data storage.

Then, the framework based on term sets is extended in Section IV to simulate dynamic networks whose topologies may change over time. We view a dynamic network as possible "worlds", i.e. states in which the network is, and we allow the users to have requirements on the dispersion that change over time. A dynamic network can thus be modeled as one main term set, viewed as the union of the term sets for all users, possible worlds, and time-slots. The second main contribution of this paper is the multi-user theorem for dynamic networks, which proves that if the demand (over all worlds and all time-slots) of each user can be satisfied locally, then they can all be satisfied globally. In that sense, this can be viewed as a network coding theorem for dynamic multi-user networks. This theorem is expressed in terms of dispersion, one-to-one dispersion, or Rényi entropy demands.

A surprising corollary of our results is the inefficiency of *clairvoyance*, informally defined as the case where the coding functions "know" in which world the network is. In that case, we show that the min-cuts of the term sets corresponding to clairvoyance are equal to that of the case without clairvoyance. Therefore, if a utility demand can be satisfied with clairvoyance, then it can also be satisfied without clairvoyance. However, in the case where users not only have dispersion demands, but actually request specific messages, then clairvoyance may greatly outperform "natural" coding.

II. Communication networks based on term sets

A. Term sets

Let $X = \{x_1, x_2, \dots, x_k\}$ be a set of *variables* and consider a set of *function symbols* $\{f_1, f_2, \dots, f_l\}$ with respective ari-

978-1-61284-140-3/11/$26.00 ©2011 IEEE

ties (numbers of arguments) $d_1, d_2, \ldots, d_l$. A *term* is defined to be an object obtained from applying function symbols to variables recursively.

We say that u is a subterm of t if the term u appears in the definition of t. Furthermore, u is a *direct subterm* of t if $t = f_j(v_1, \ldots, u, \ldots, v_{d_j})$, and f_j is referred to as the *principal function* of t.

We shall consider finite term sets, typically referred to as $\Gamma = \{t_1, t_2, \ldots, t_r\}$. We denote the set of variables that occur in terms in Γ as $\Gamma_{\rm var}$ and the collection of subterms of one or more terms in Γ as $\Gamma_{\rm sub}$.

We now define a term-cut, which can be viewed as replacing some subterms in the definition of a term by variables.

Definition 1 (Term-cut): A set of subterms $s_1, s_2, \ldots, s_\rho \in \Gamma_{\rm sub}$ provides a *term-cut* of size ρ for Γ if all the terms can be expressed syntactically by applying function symbols to $s_1, s_2, \ldots, s_\rho$.

The minimum size of a term-cut for Γ is referred to as the *min-cut* of Γ, which can be viewed as the number of degrees of freedom of the term set. Clearly the min-cut is no more than the number of variables k and the number of terms r.

These concepts can be graphically explained as follows.

Definition 2 (The graph G_Γ): For a given term set Γ, the directed graph $G_\Gamma = (V, E, S, T)$ is defined to have vertex set $V = \Gamma_{\rm sub}$, edge set $E = \{(u, v) : u \,\text{direct subterm of}\, v\}$, source set $S = \Gamma_{\rm var}$, and target set $T = \Gamma$.

A set of subterms is a term-cut for Γ if and only if it is a vertex cut that separates $S = \Gamma_{\rm var}$ from $T = \Gamma$ in the directed graph G_Γ. Notice that $S \cap T$ is non-empty if Γ contains one or more terms that are variables.

B. Dispersion

So far, we have treated function symbols as abstract entities. We now study the case when these function symbols are assigned explicit values.

Definition 3 (Interpretation): Let A be a finite set with $|A| \geq 2$, referred to as the alphabet. An *interpretation* for Γ over A is an assignment of the function symbols $\psi = \{\bar{f}_1, \bar{f}_2, \ldots, \bar{f}_l\}$, where $\bar{f}_i : A^{d_i} \to A$ for all $1 \leq i \leq l$.

Once all the function symbols f_i are assigned coding functions $\bar{f}_i$, then by composition each term $t_j \in \Gamma$ is assigned a function $\bar{t}_j : A^k \to A$. In order to simplify notations, we shall write functions by the way they map a tuple $\mathbf{a} = (a_1, a_2, \ldots, a_k) \in A^k$. We shall abuse notations and also denote the *induced mapping* of the interpretation as $\psi : A^k \to A^r$, defined as

$$\psi(\mathbf{a}) = \big(\bar{t}_1(\mathbf{a}), \bar{t}_2(\mathbf{a}), \ldots, \bar{t}_r(\mathbf{a})\big).$$

Note that the definition of the induced mapping depends on the ordering of terms in Γ. However, our performance measures for interpretations and induced mappings will not depend on a particular ordering.

We are interested in how ψ disperses its outputs, and how much information about the inputs can be obtained from the outputs. For any $\mathbf{b} \in A^r$, we denote the pre-image of $\mathbf{b}$ as ${\rm pre}(\mathbf{b}) = \{\mathbf{a} \in A^k : \psi(\mathbf{a}) = \mathbf{b}\}$. The image and the one-to-one image of ψ are respectively defined as ${\rm image}(\psi) := \{\mathbf{b} \in A^r : |{\rm pre}(\mathbf{b})| \geq 1\}$ and ${\rm one}(\psi) := \{\mathbf{b} \in A^r : |{\rm pre}(\mathbf{b})| = 1\}$. The Γ-*dispersion* and *one-to-one* Γ-*dispersion* of an interpretation ψ for Γ over A are respectively defined as

$$\gamma(\psi) := \log_{|A|} |{\rm image}(\psi)|, \quad \gamma_{\rm one}(\psi) := \log_{|A|} |{\rm one}(\psi)|.$$

These can be viewed as the analogue of the value of a flow for information transfer on networks based on logic.

If the input $\bar{\mathbf{a}} \in A^k$ is uniformly distributed, then $\psi(\bar{\mathbf{a}})$ is a random variable with values in A^r, where for all $\mathbf{b} \in A^r$

$$p_{\mathbf{b}} = \mathbb{P}\{\psi(\bar{\mathbf{a}}) = \mathbf{b}\} = |A|^{-k} |{\rm pre}(\mathbf{b})|.$$

The normalized Rényi entropy $H_\alpha(\psi)$ of order α of the random variable $\psi(\bar{\mathbf{a}})$, is given for $0 < \alpha < \infty$, $\alpha \neq 1$ by [6]

$$\begin{aligned} H_\alpha(\psi) &:= \frac{1}{1-\alpha} \log_{|A|} \sum_{\mathbf{b} \in A^r} p_{\mathbf{b}}^\alpha \\ &= \frac{\alpha}{\alpha - 1} \left(k - \frac{1}{\alpha} \log_{|A|} \sum_{\mathbf{b} \in A^r} |{\rm pre}(\mathbf{b})|^\alpha \right). \end{aligned}$$

Three further special cases need close attention. First, when $\alpha = 0$, the so-called *Hartley entropy* is the dispersion of ψ:

$$H_0(\psi) := \log_{|A|} |\{\mathbf{b} \in A^r : p_{\mathbf{b}} > 0\}| = \gamma(\psi).$$

Second, the *Shannon entropy*, obtained when $\alpha = 1$, is the expected amount of information inferred from the term set about the input messages:

$$H_1(\psi) := -\sum_{\mathbf{b} \in A^r} p_{\mathbf{b}} \log_{|A|} p_{\mathbf{b}} = \mathbb{E}\left\{k - \log_{|A|} |{\rm pre}(\mathbf{b})|\right\}.$$

Third, when $\alpha = \infty$, the *min-entropy* quantifies the amount of information that can be inferred from any output:

$$H_\infty(\psi) := -\log_{|A|} \max_{\mathbf{b} \in A^r} p_{\mathbf{b}} = \min_{\mathbf{b} \in A^r} \left\{k - \log_{|A|} |{\rm pre}(\mathbf{b})|\right\}.$$

C. Single-user max-flow min-cut theorems

We refer to the maximal Γ-dispersion over all interpretations for Γ over A as the dispersion of Γ over A, which we denote by $\gamma(\Gamma, |A|)$ as this quantity clearly depends on A via its cardinality only. We similarly denote $\gamma_{\rm one}(\Gamma, |A|)$.

Theorem 1 (Max-flow min-cut theorem for dispersion): Let Γ be a term set with min-cut of ρ, then for any alphabet A, $\gamma_{\rm one}(\Gamma, |A|) \leq \gamma(\Gamma, |A|) \leq \rho$. Conversely,

$$\lim_{|A| \to \infty} \gamma_{\rm one}(\Gamma, |A|) = \lim_{|A| \to \infty} \gamma(\Gamma, |A|) = \rho.$$

Similarly to the dispersion, we denote the maximum Rényi entropy over all interpretations for Γ over A as $H_\alpha(\Gamma, |A|)$. The max-flow min-cut theorem for the dispersion indicates that the Rényi entropy for $\alpha = 0$ tends to the min-cut of the term set. This can be generalized to the Rényi entropy for $\alpha < 1$.

Theorem 2 (Max-flow min-cut theorem for the Rényi entropy): Let Γ be a term set with min-cut of ρ, then

$$\lim_{|A| \to \infty} H_\alpha(\Gamma, |A|) = \rho \quad \text{for all}\, 0 \leq \alpha < 1.$$

Conversely, for any $\alpha > 1$, there exists a term set Γ with min-cut ρ for which $\lim_{|A| \to \infty} H_\alpha(\Gamma, |A|) < \rho$.

III. MULTI-USER MAX-FLOW MIN-CUT THEOREM

We now consider multi-user communications. To each user (receiver) $1 \leq j \leq m$, we associate a term set Γ_j. Each choice of coding functions determines a dispersion of $(\Gamma_1, \Gamma_2, \ldots, \Gamma_m)$ which is an array of dispersions. Before detailing how the term sets correspond to the multi-user communication problem, let us derive the max-flow min-cut theorem for a collection of term sets.

Theorem 3 (Multi-user max-flow min-cut theorem): Let $\Gamma_1, \Gamma_2, \ldots, \Gamma_m$ be term sets with respective min-cuts $\rho_1, \rho_2, \ldots, \rho_m$. Then for any $\epsilon > 0$, there exists n_0 such that for all A with $|A| \geq n_0$, $\gamma_{\rm one}(\Gamma_j, \rho_j) \geq \rho_j - \epsilon$ for all $1 \leq j \leq m$. Also, for any $0 \leq \alpha < 1$, there exists n such that $H_\alpha(\Gamma_j, |A|) \geq \rho_j - \epsilon$ for all $|A| \geq n$.

The proof is based on diversifying variables. Let $X = \{x_1, x_2, \ldots, x_k\}$ be the set of variables on which all terms are built. For any $1 \leq j \leq m$, let $X^j = \{x_1^j, x_2^j, \ldots, x_k^j\}$ be a new set of variables, and let Γ^j be the term set obtained from Γ_j by replacing the variable x_i by x_i^j for all $1 \leq i \leq k$. It is clear that any interpretation for Γ_j can be viewed as an interpretation for Γ^j, and that its Γ_j-dispersion (also one-to-one dispersion and Rényi entropy) is equal to its Γ^j-dispersion.

Consider now

$$\Gamma := \bigcup_{j=1}^{m} \Gamma^j,$$

where the union is disjoint and with min-cut equal to $\rho = \sum_{j=1}^m \rho_j$. In fact the graphs G_{Γ^j} are components of G_Γ, which shows that for any interpretation ψ for Γ, we have $\gamma(\psi) = \sum_{j=1}^m \gamma(\psi^j)$, where ψ^j is the corresponding interpretation for Γ^j. Therefore, the theorem is proved by choosing an interpretation ψ for Γ such that $\gamma(\psi) \geq \rho - \epsilon$, hence for any j, we have $\gamma(\psi^j) \geq \rho_j - \epsilon$.

A. Multi-user communication problems and term sets

We now apply the multi-user max-flow min-cut theorem to communication problems. A many-to-many cast is defined as follows.

Definition 4: A multi-user communication problem instance (also referred to as a *many-to-many cast*) is a tuple (V, E, S, T, A), where

- $G = (V, E)$ is an acyclic directed graph, where the vertices $v_1, v_2, \ldots, v_{|V|}$ of V are sorted such that $(v_i, v_j) \in E$ only if $i < j$.
- $S = \{s_1, s_2, \ldots, s_k\} \subseteq V$ is the set of sources, which are nodes with in-degree 0. Without loss of generality, $s_i = v_i$ for $1 \leq i \leq k$.
- $T = \{r_1, r_2, \ldots, r_m\} \subseteq V$ is the set of receivers, which are nodes with out-degree 0. Without loss, $r_j = v_{|V|-m+j}$ for $1 \leq j \leq m$.
- A is an alphabet of size $|A| \geq 2$. Each source s_i sends a distinct message $a_i \in A$, and each receiver r_j requests $\{a_1, a_2, \ldots, a_k\}$.

Each vertex $v_k \in V \backslash S$ can manipulate the data it receives and transmit a function f_k of its inputs onto all its out-edges.

A communication problem can be equivalently defined with terms. Each receiver obtains a term built on the variables sent by the sources, where the function symbols represent the operations made by the intermediate nodes in $V \backslash (S \cup T)$. More formally, to each source s_i we associate the variable x_i and we denote $X = \{x_1, x_2, \ldots, x_k\}$. Each receiver r_j requests all the variables in X. We then associate the function symbol f_l to all intermediate nodes $v_l \in V \backslash (S \cup T)$ and each vertex v_l is recursively assigned the term $u_l = f_l(u_{l_1}, u_{l_2}, \ldots, u_{l_d})$, where $\{v_{l_1}, v_{l_2}, \ldots, v_{l_d}\}$ is the in-neighborhood of v_l. We denote the in-neighborhood of the receiver r_j as $\{v_{j,1}, v_{j,2}, \ldots, v_{j,p_j}\}$. Note that using this notation, it is possible that $v_{j,i} = v_{j',i'}$ for distinct j, j' and i, i'. However, this will not affect our definitions below. We finally associate the term $t_{j,i}$ to the vertex $v_{j,i}$.

A solution for the many-to-many cast instance is a choice of the functions at the intermediate nodes such that all the receivers' demands can be satisfied at the same time. Using terms, receiver j is satisfied if and only if it can recover X from the term set $\Gamma_j = \{t_{j,i}\}_{i=1}^{p_j}$ obtained from its in-neighborhood. This implies that the term set Γ_j must have dispersion of k. In order to take into account the fact that all demands have to be satisfied at the same time, we use the term set Γ defined by diversifying variables. By our construction and the remarks on Γ made above, we obtain the following result.

Proposition 1: A many-to-many cast instance is solvable over A if and only if

$$H_\infty(\Gamma_j, |A|) = \gamma_{\rm one}(\Gamma_j, |A|) = \gamma(\Gamma_j, |A|) = k$$

for all $1 \leq j \leq m$ or equivalently,

$$H_\infty(\Gamma, |A|) = \gamma_{\rm one}(\Gamma, |A|) = \gamma(\Gamma, |A|) = km.$$

The multi-user max-flow min-cut theorem then shows that if the min-cut between the sources and each receiver is equal to k, then the multi-user instance is asymptotically solvable. Note that the equivalence in Proposition 1 together with Theorem 3 indicate that if each user can be asymptotically satisfied individually, then all the users' demands can asymptotically be satisfied simultaneously.

B. Examples

The following examples are consequences of Proposition 1.

Example 1 (Satellite communication): In [3], Yeung and Zhang made a simple observation that lay the foundation for network coding. The authors considered a situation where two users communicate via a satellite. User X wants to send a message $x \in A$ to Y, while Y at the same time wants to send a message $y \in A$ to X.

The satellite communication problem is equivalent to the communication problem in Figure 1a), referred to as the butterfly network. In Figure 1a), the function symbols are affected to the vertices, accordingly to Definition 4; however the problem is sometimes represented in the literature by Figure 1b), where the function symbols are affected to the edges. A solution over an alphabet A is a function $\bar{f} : A^2 \to A$ with the property

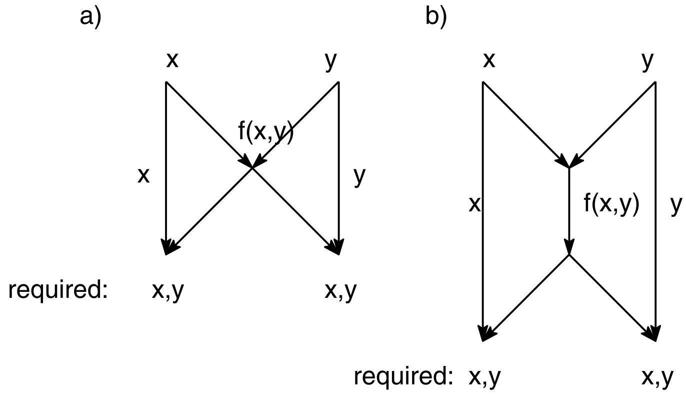

Fig. 1. Butterfly network.

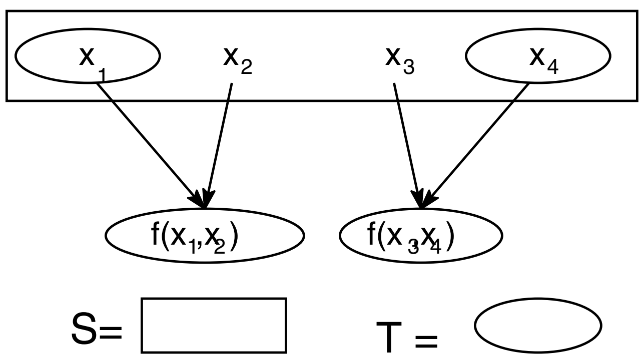

Fig. 2. Graph G_Γ for the butterfly network.

that there exist decoding functions $\bar{d}_1, \bar{d}_2 : A^2 \to A$ such that $b = \bar{d}_1\big(b, \bar{f}(a,b)\big)$ and $a = \bar{d}_2\big(a, \bar{f}(a,b)\big)$ for any $(a,b) \in A^2$.

By Proposition 1, the satellite communication problem is equivalent to the following problem on terms.

Problem 1 (Equivalent to satellite communication): Construct a function $\bar{f} : A^2 \to A$ with Γ-dispersion equal to 4, where

$$\Gamma := \{x_1, f(x_1,x_2), f(x_3,x_4), x_4\}.$$

This term set corresponds to the graph with vertex set

$$V = \Gamma_{\mathrm{sub}} = \{x_1, x_2, x_3, x_4, f(x_1,x_2), f(x_3,x_4)\},$$

source set $S = \{x_1, x_2, x_3, x_4\}$ and target set $T = \Gamma$. The edge set $E \subseteq V \times V$ is given by

$$E = \big\{\big(x_1, f(x_1,x_2)\big), \big(x_2, f(x_1,x_2)\big), \big(x_3, f(x_3,x_4)\big), \big(x_4, f(x_3,x_4)\big)\big\}.$$

The graph G_Γ is displayed in Figure 2.

Problem 1 can be viewed as a communication problem between a single source that transmits a message $(a_1,a_2,a_3,a_4) \in A^4$ and a receiver who receives a message of the form $\big(a_1, \bar{f}(a_1,a_2), \bar{f}(a_3,a_4), a_4\big) \in A^4$.

Example 2 (Distributed storage): Assume that we want to store two messages $x, y \in A$ at four locations. The messages x and y are stored at two of the locations. At the two remaining locations two messages $f(x,y) \in A$ and $g(x,y) \in A$ are stored. The problem is to select the coding functions $f : A^2 \to A$ and $g : A^2 \to A$ such that it is always is possible to reconstruct x and y from accessing only two of the four locations.

This type of problem has already been studied in the literature [7] as part of network coding, as well as an application of error correcting codes. The actual problem can be shown to be equivalent to the existence of two orthogonal Latin squares of order $|A|$ [8]. This problem was first posed by Euler around 1780 and was eventually completely solved in 1960, where it was shown in [9] that there exist orthogonal Latin squares of any order except of order 2 and order 6. The distributed storage problem is equivalent to the following problem.

Problem 2 (Equivalent to the distributed storage problem): Construct two functions $\bar{f}, \bar{g} : A^2 \to A$ with Γ-dispersion equal to 10, where

$$\Gamma := \big\{x_1, f(x_1,y_1), y_2, f(x_2,y_2), x_3, g(x_3,y_3), y_4, g(x_4,y_4), f(x_5,y_5), g(x_5,y_5)\big\}.$$

Proof: The storage problem is a many-to-many cast instance with two sources sending x and y and six receivers, corresponding to the six possible ways of accessing the stored data, obtaining $\{x,y\}$, $\{x, f(x,y)\}$, $\{y, f(x,y)\}$, $\{x, g(x,y)\}$, $\{y, g(x,y)\}$, and $\{f(x,y), g(x,y)\}$ respectively. The demands of the first receiver are trivially satisfied, so only the last five need to be considered. Applying the transformations above and Proposition 1, we obtain the desired term set. ■

Problem 2 can be viewed as a communication problem between a single source that transmits a message $\mathbf{a} \in A^{10}$ and a receiver who receives a message $\psi(\mathbf{a}) \in A^{10}$.

We can give a network coding interpretation of the diversified term set. When the term set is diversified, then we can select a different solution for each receiver independently. Therefore, we have full dispersion in the diversified case if and only if each request can be satisfied individually.

Example 3: Consider the term set $\Gamma = \{x_1, f(x_1,x_2), x_4, f(x_3,x_4)\}$ corresponding to the satellite communication problem. The diversified term set can be written as $\Gamma^{\mathrm{div}} = \{x_1, f(x_1,x_2), x_4, g(x_3,x_4)\}$. Since the coding functions f and g are not required to be identical, it is easier to find coding functions that achieve the maximal dispersion of 4 in the case of Γ^{div}. One possible choice of coding functions that achieve dispersion of 4 is to let $f(x_1,x_2) = x_2$ and let $g(x_3,x_4) = x_3$. However, this choice does not correspond to any real-life situation for the original satellite communication problem. In fact, diversifying the term set is equivalent to considering the case where the satellite has access to two independent channels to the stations on which it can send a different message (one constructed by the function f, the other by the function g).

The max-flow min-cut theorem for the dispersion then shows that if all demands can be satisfied individually, then for all large enough alphabets, all demands can be “nearly” satisfied at once. Theorem 3 quantifies this statement in terms of a small loss in dispersion, one-to-one dispersion, or Rényi entropy. However, this small loss may be critical when network coding is considered. Indeed let Γ be the term set associated to a given multi-user communication problem, and let ψ be the induced function of an interpretation for Γ. Then, the one-to-one pre-image of ψ may contain very few points $(\mathbf{a}, \mathbf{a}, \dots, \mathbf{a}) \in A^{km}$, where $\mathbf{a} \in A^k$. However, only these points make sense for the multi-user communication, as the variable x_i^j is merely an artificial variable representing x_i for all j. Consider the butterfly network for example. Then all the points in the one-to-one pre-image of dynamic routing satisfy $x_1 \neq x_3$, which does not correspond to any real-life situation for the satellite communication problem.

IV. Dynamic networks

A. Theorem for multi-user dynamic networks

In the analysis of dynamic communication networks it is natural to take into account that networks change over time. Potential network changes including link failures, point failures and noisy channels can be modeled by ideas vaguely akin to Kripke's *possible world semantics* from logic and philosophy. The idea is to consider a collection of *possible worlds* which each could become *the actual world* as time progresses.

A world is not only a representation of the network at a given time, but an expansion over a number of time slots. This is a generalization of the butterfly network, which can be viewed as an expansion over time of the satellite communication problem. As the model is discrete there are only a finite a number of possible worlds. We can think of each node in a world as a network node at a certain time slot. A point in each world thus represents a node at a given time slot, and as such there is a link from each node to its successor in time. This link is not a communication link, but represents the data transformation at the node in the two time slots. Communications might be instantaneous i.e. connect different nodes in the same time slot. The resulting network is acyclic. Formally we define

Definition 5 (Dynamic network): Let U, W, and T be finite sets. A *dynamic network* is a collection

$$\tilde{\Gamma} := \{\Gamma_{u,w,t} : u \in U, w \in W, t \in T\}$$

of term sets indexed by a user/receiver $u \in U$, a world $w \in W$, and a time-slot $t \in T$.

Assume all coding functions that occur in terms in $\tilde{\Gamma}$ have been given interpretations. Then we can associate a dispersion to each set term set $\Gamma_{u,w,t}$. For each set $\Gamma_{u,w,t}$ we associate a variable $\gamma_{u,w,t}$ which for each choice of coding functions denotes the dispersion of the term set $\Gamma_{u,w,t}$. Typically the same function symbol might occur in terms sets in multiple worlds. A coding function might be a good choice for some of the worlds, while it might be a bad choice for other worlds.

Some worlds may be more likely than others. We tackle this issue by considering the general case where each user $u \in U$ is assigned a *utility function*, defined as a real-valued, non-decreasing, and continuous function F_u in the variables $\gamma_u = \{\gamma_{u,w,t} : w \in W, t \in T\}$. Each user u has a *utility demand* which is the real number $\mathrm{dem}_u \in \mathbb{R}$, where u requires that the utility of the received dispersion strictly exceeds this value i.e. that $F_u(\gamma_u) > \mathrm{dem}_u$. We say that the demand

$$D_u \equiv F_u(\gamma_u) > \mathrm{dem}_u$$

is satisfied *locally* if it can be satisfied when all other user demands are disregarded. Conversely, we say that $D_1, D_2, \dots, D_m$ are satisfied *globally* if they can all be satisfied by the same interpretation.

Theorem 4 (Dynamic multi-user theorem): In a dynamic many-to-many cast where the demand D_u of each receiver $1 \leq u \leq m$ can be satisfied locally, the demands $D_1, D_2, \dots, D_m$ can be satisfied globally. The same holds for Rényi entropy demands with $\alpha < 1$ and one-to-one dispersion demands.

Proof: Suppose that each $F_u(\gamma_u) > \mathrm{dem}_u$ with $1 \leq u \leq m$ can be achieved locally. If we select $\delta > 0$ such that $\delta < \min_u\{F_u(\gamma_u) - \mathrm{dem}_u\}$, then in fact $F_u(\gamma_u) > \mathrm{dem}_u + \delta$ with $1 \leq u \leq m$ can be achieved locally. Assume that $F_u(\gamma_u) > \mathrm{dem}_u + \delta$ is achieved (locally) by the dispersions $\gamma_u = \{\gamma_{u,w,t} : w \in W, t \in T\}$. Since F_u is continuous, there exists $\epsilon > 0$ such that if the dispersion γ'_u of $\{\Gamma_{u,w,t} : w \in W, t \in T\}$ has $|\gamma'_{u,w,t} - \gamma_{u,w,t}| < \epsilon$, then $|F_u(\gamma'_u) - F_u(\gamma_u)| < \delta$. According to the multi-user max-flow min-cut theorem, for each $\epsilon > 0$ there exists an interpretation (over a sufficiently large alphabet) which globally achieves the dispersions $\gamma_{u,w,t} - \epsilon$, $u \in U, w \in W, t \in T$. Thus there exist coding functions such that $F_u(\gamma_u) > \mathrm{dem}_u$ for $1 \leq u \leq m$. ■

Example 4 below shows how the utility function can cover a broad family of performance measures.

Example 4: Let $\tilde{\Gamma}$ consist of term sets $\Gamma_{u,w,t}$ where $u \in U, w \in W$, and $t \in T$. Assume world w occurs with probability p_w and the user u has assigned a weight ω_t proportional to the utility of the dispersion achieved in time slot t. Then the utility for each user $u \in U$ can be defined as

$$F_u(\gamma_u) = \sum_{w \in W, t \in T} p_w \omega_t \gamma_{u,w,t}$$

which is a continuous function in the variables γ_u. Asymptotically, the maximal utility achievable for user u is given by $\sum_{w \in W, t \in T} p_w \omega_t \rho_{u,w,t}$, where $\rho_{u,w,t}$ is the min-cut of $\Gamma_{u,w,t}$.

The example above can be viewed as an asymptotic generalization of the network coding theorem, which only considers one time slot and one possible world. A proper generalization of that result can be obtained by considering linear coding functions over a a diversified term set for a multi-user communication problem.

B. Clairvoyance and term equations

In a dynamic network, various unpredictable network changes (e.g. link failures) might happen in various of the possible worlds. A realistic choice of coding functions cannot look into the future and take into account which link might fail during transmission. Nonetheless, we define clairvoyance as the case where each node "knows" in which of the possible worlds the network is. We can define the formally as follows:

Definition 6 (Clairvoyant coding): The clairvoyant version $\tilde{\Gamma}^{\text{clair}}$ of $\tilde{\Gamma}$ is defined as the collection of term sets where function symbols have been diversified so function symbols in different worlds are distinct, e.g. each function symbol the occurs in term sets with index $w \in W$ is assigned an (additional) index w. An assignment of functions symbols to $\tilde{\Gamma}^{\text{clair}}$ is said to consists of *clairvoyant coding functions* for $\tilde{\Gamma}$.

As another application of the multi-user dispersion theorem we obtain that clairvoyance does not improve the performance of a network in terms of dispersion. The proof of Proposition 2 below is similar to that of Theorem 1.

Proposition 2 (Clairvoyance does not increase dispersion): If in a dynamic many-to-many network the users demands $D_1, D_2, \ldots, D_r$ can be satisfied in $\tilde{\Gamma}^{\text{clair}}$, then they can be satisfied in $\tilde{\Gamma}$.

It should be noticed that this result for the dispersion and the Rényi entropy is very much in the spirit of diversity coding and random linear network coding and is thus not surprising. Indeed, diversity coding intuitively deals with link failures and noisy channels by mixing the inputs and transmitting a large number of independent messages. Clairvoyance is thus rendered useless. However, our result about the one-to-one dispersion is remarkable, for a high one-to-one dispersion involves a controlled non-linear mixing, which contradicts the philosophy of diversity coding.

We finish this section by revealing that a model based on term equations can take into account the fact that each user not only requires a high dispersion, but also a certain number of specific messages. If a user receives the terms $\{t_1, t_2, \ldots, t_r\}$, we associate the decoding functions $\bar{d}_1, \bar{d}_2, \ldots, \bar{d}_s$ where we require that $\bar{d}_i(\psi(\mathbf{a})) = a_i$ for all $1 \leq i \leq s$. This can be more succinctly expressed in the term equation $\tau_i = x_i$, where τ is the term defined as $\tau_i = d_i(t_1, t_2, \ldots, t_r)$.

In the case of message demands, then clairvoyance can clearly make a difference, as seen in Example 5 below.

Example 5: Consider the dynamic network depicted in Figure 3, where one of two links is always contaminated with pure noise. Remark that this is not the butterfly network in Figure 1. Without using clairvoyance, the message demands of the destinations can be expressed as term equations

$$\begin{aligned} d_1(\text{noise}_1, f(x,y)) &= x \\ d_2(y, f(x,y)) &= y \\ d_3(x, f(x,y)) &= x \\ d_4(\text{noise}_2, f(x,y)) &= y. \end{aligned}$$

Here we assumed that the decoding functions can distinguish the messages x and y from noise, which is why we can apply different decoding functions to each of the four potential decoding situations. It is clear that the message demands of both receivers cannot be satisfied globally without clairvoyance. On the other hand, using clairvoyance, the problem is turned into

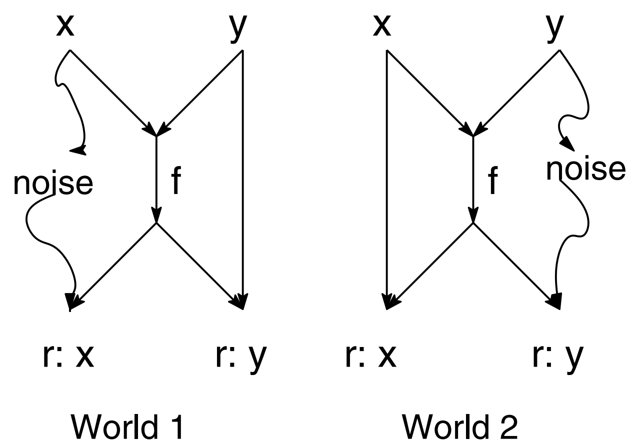

Fig. 3. Dynamic network with message demands.

the following set of term equations:

$$\begin{aligned} d_1(\text{noise}_1, f_1(x,y)) &= x \\ d_2(y, f_1(x,y)) &= y \\ d_3(x, f_2(x,y)) &= x \\ d_4(\text{noise}_2, f_2(x,y)) &= y. \end{aligned}$$

Thus, letting $\bar{f}_1(x,y) = x$ for world 1 and $\bar{f}_2(x,y) = y$ for world 2 solves the communication problem by an appropriate choice of decoding functions $\bar{d}_1, \bar{d}_2, \bar{d}_3$, and $\bar{d}_4$.

In brief, clairvoyance does not help in many-to-many casting where the receivers have dispersion demands. But not surprisingly in general–when different users have different message demands–clairvoyance can greatly increase the performance of the network.

References

[1] R. Dougherty, C. Freiling, and K. Zeger, "Linearity and solvability in multicast networks," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2243–2256, October 2004.

[2] ——, "Networks, matroids, and non-Shannon information inequalities," *IEEE Transactions on Information Theory*, vol. 53, no. 6, pp. 1949–1969, June 2007.

[3] R. Yeung and Z. Zhang, "Distributed source coding for satellite communications," *IEEE Trans. Info. Theory*, vol. 45, no. 3, pp. 1111–1120, May 1999.

[4] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Info. Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.

[5] S. Riis and M. Gadouleau, "Max-flow min-cut theorems for multi-user communication networks," *submitted*, avaible at http://arxiv.org/abs/1012.5224.

[6] A. Rényi, "On measures of information and entropy," in *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, 1961, p. 547561.

[7] A. Dimakis, P. Godfrey, Y. Wu, M. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," *IEEE Trans. Info. Theory*, vol. 56, no. 9, pp. 4539–4551, September 2010.

[8] S. Riis and R. Ahlswede, "Problems in network coding and error-correcting codes," in *Proc. NetCod*, Riva del Garda, Italy, April 2005.

[9] R. Bose, S. Shrikhande, and E. Parker, "Further results on the construction of mutually orthogonal latin squares and the falsity of Euler's conjecture," *Canadian Journal of Mathematics*, vol. 12, pp. 189–203, 1960.