5-2012

Selfish Distributed Compression Over Networks: Correlation Induces Anarchy

Aditya Ramamoorthy
Iowa State University, adityar@iastate.edu

Vwani Roychowdhury
University of California, Los Angeles

Sudhir Kumar Singh
University of California, Los Angeles

Selfish Distributed Compression over Networks

Aditya Ramamoorthy*, Vwani Roychowdhury†, Sudhir Kumar Singh†

* Department of Electrical and Computer Engineering, Iowa State University, Ames, Iowa 50011.
Email: adityar@iastate.edu
†Department of Electrical Engineering, University of California, Los Angeles, CA 90095.
Emails:{vwani,suds}@ee.ucla.edu

***Abstract*—We consider the min-cost multicast problem (under network coding) with multiple correlated sources where each terminal wants to losslessly reconstruct all the sources. This can be considered as the network generalization of the classical distributed source coding (Slepian-Wolf) problem. We study the inefficiency brought forth by the selfish behavior of the terminals in this scenario by modeling it as a noncooperative game among the terminals. The solution concept that we adopt for this game is the popular local Nash equilibrium (Waldrop equilibrium) adapted for the scenario with multiple sources. The degradation in performance due to the lack of regulation is measured by the *Price of Anarchy* (POA), which is defined as the ratio between the cost of the worst possible Waldrop equilibrium and the socially optimum cost. Our main result is that in contrast with the case of independent sources, the presence of source correlations can significantly increase the price of anarchy. Towards establishing this result we make several contributions. We characterize the socially optimal flow and rate allocation in terms of four intuitive conditions. This result is a key technical contribution of this paper and is of independent interest as well. Next, we show that the Waldrop equilibrium is a socially optimal solution for a different set of (related) cost functions. Using this, we construct explicit examples that demonstrate that the POA > 1 and determine near-tight upper bounds on the POA as well. The main techniques in our analysis are Lagrangian duality theory and the usage of the supermodularity of conditional entropy. Finally, all the techniques and results in this paper will naturally extend to a large class of network information flow problems where the Slepian-Wolf polytope is replaced by any contra-polymatroid (or more generally polymatroid-like set), leading to a nice class of succinct multi-player games and allow the investigation of other practical and meaningful scenarios beyond network coding as well.**

I. Introduction

In large scale networks such as the Internet, the agents involved in producing and transmitting information often exhibit selfish behavior e.g. if a packet needs to traverse the network of various ISP's, each ISP will behave in a greedy manner and ensure that the packet spends the minimum time on its network. While this minimizes the ISP's cost it may not be the best strategy from a overall network cost perspective. Selfish routing, that deals with the question of network performance under a lack of regulation has been studied extensively (see [15], [20]) and has developed as an area of intense research activity. However, by and large most of these studies have considered the network traffic injected into the network at various sources to be independent.

From an information theoretic perspective there is no need to consider the sources involved in the transmission to be independent. In this paper we initiate the study of network optimization issues related to the transmission of correlated sources over a network when the agents involved are selfish. In particular, we concentrate on the problem of multicasting correlated sources over a network to different terminals, where each terminal is interested in losslessly reconstructing all the sources. We assume that the network is capable of network coding. Under this scenario, a generalization of the classical Slepian-Wolf theorem of distributed source coding [8] holds for arbitrary networks. In particular when the network performs random linear network coding each terminal can recover the sources under appropriate conditions on the Slepian-Wolf region and the capacity region of the terminals with respect to the sources, thereby allowing distributed source coding over networks. The selfish agents in our set-up are the terminals who pay for the resources. Each terminal aims to minimize her own cost while ensuring that she can satisfy her demands. It is important to note that this is a generalization of the problem of minimum cost selfish multicast of independent sources considered by Bhadra et al. [3].

Our Results: In this work, we model the scenario as a noncooperative game amongst the selfish terminals who request rates from sources and flows over network paths such that their individual cost is minimized (i.e. with no regard for social welfare) while allowing for reconstruction of all the sources. We investigate properties of the socially optimal solution and define appropriate solution concepts (Nash equilibrium and Waldrop equilibrium) for this game and investigate properties of the flow-rates at equilibrium. We briefly describe our contributions below.

i) *Characterization of social-optimality conditions.* The problem of computing the socially optimal cost is a convex program. We present a precise characterization of the optimality conditions of this convex program in terms of four intuitive conditions, using Lagrangian duality theory and by judiciously exploiting the super-modularity of conditional entropy.

ii) *Demonstrating the equivalence of flow-rates at equilibrium with social-optimal solutions for alternative instances.* We consider certain meaningful market models that split resource costs amongst the different terminals and show that the flows and rates under the game-theoretic equilibriums are in fact socially optimal solutions for a different set of cost functions. This characterization allows us to quantify the degradation caused by the lack of regulation. The measure of performance degradation due to such loss in regulation that we adopt is

the *Price of Anarchy* (POA), which is defined as the ratio between the cost of the worst possible equilibrium and the socially optimum cost [9], [17], [21], [20].
iii) *Showing that source correlation induces anarchy.* The main result of this paper is that the presence of source correlations can significantly increase the POA under reasonable cost-splitting mechanisms. This is in stark contrast to the case of multicast with independent sources, where for a large class of cost functions, cost-splitting mechanisms can be designed that ensure that the price of anarchy is one. We construct explicit examples where the POA is greater than one and also obtain an upper bound on the POA which is near tight.

Finally, we expect that the techniques developed in this paper will be applicable to a large class of network information flow problems with correlated sources where the Slepian-Wolf polytope is replaced by polymatroid-like objects. These include multi-terminal source coding with high resolution [23] and the CEO problem [18].

Background and Related Work: Distributed source coding (or distributed compression) (see [5], Ch. 14 for an overview) considers the problem of compressing multiple discrete memoryless sources that are observing correlated random variables. The landmark result of Slepian and Wolf [22] characterizes the feasible rate region for the recovery of the sources. However, the problem of Slepian and Wolf considers a direct link between the sources and the terminal. More generally one would expect that the sources communicate with the terminal over a network. Different aspects of the Slepian-Wolf problem over networks have been considered in ([2], [6], [19]). Network coding (first introduced in the seminal work of Ahlswede et al. [1]) for correlated sources was considered by Ho et al. [8]. They considered a network with a set of sources and a set of terminals and showed that as long as the minimum cuts between all non-empty subsets of sources and a particular terminal were sufficiently large (essentially as long as the Slepian-Wolf region of the sources has an intersection with the capacity region of a given terminal), random linear network coding over the network followed by appropriate decoding at the terminals achieves the Slepian-Wolf bounds.

The problem of minimum cost multicast under network coding has been addressed in the work of [13], [12]. The multicast problem has also been examined by considering selfish agents [3], [10], [11]. Our work is closest in spirit to the analysis of Bhadra et al. [3] that considers selfish terminals. In this scenario, for a large class of edge cost functions, they develop a pricing mechanism for allocating the edge costs among the different terminals and show that it leads to a globally optimal solution to the original optimization problem, i.e. the price of anarchy is one. Their POA analysis is similar to that in the case of selfish routing [21], [20]. Our model is more general and our results do not generalize from theirs in a straightforward manner. In particular, we need to judiciously exploit several non-trivial properties of the Slepian-Wolf polytope in our analysis.

II. The Model

Consider a directed graph $G = (SUTUV, E)$. There is a set of source nodes S that may be correlated and a set of sinks T that are the terminals (i.e. receivers). Each source node observes a discrete memoryless source X_i. The Slepian-Wolf region of the sources is assumed to be known and is denoted $\mathcal{R}_{\mathcal{SW}}$. For notational simplicity, let $N_S = |S|, N_T = |T|$, $S = \{1, 2, \ldots, N_S\}$, and $T = \{t_1, t_2, \ldots, t_{N_T}\}$. The set of paths from source s to terminal t is denoted by $\mathcal{P}_{s,t}$. Further, define $\mathcal{P}_t = \cup_{s\in S}\mathcal{P}_{s,t}$ i.e. the set of all possible paths going to terminal t, and $\mathcal{P} = \cup_{t\in T}\mathcal{P}_t$, the set of all possible paths. A *flow* is an assignment of non-negative reals to each path $P \in \mathcal{P}$. The flow on P is denoted f_P. A *rate* is a function $R : S \times T \longrightarrow \mathcal{R}^+$, i.e. the rate requested by the terminal t from the source s is $R_{s,t}$. We will refer to a flow and rate pair (f, R) as *flow-rate*. Also, let us denote the rate vector for terminal t by $\boldsymbol{R}_t$ and the vector of requested rates at source s by $\boldsymbol{\rho}_s$ i.e. $\boldsymbol{R}_t = (R_{1,t}, R_{2,t}, \ldots, R_{N_S,t})$ and $\boldsymbol{\rho}_s = (R_{s,t_1}, R_{s,t_2}, \ldots, R_{s,t_{N_T}})$.

Associated with each edge $e \in E$ is a cost c_e, which takes as argument a scalar variable z_e that depends on the flows to various terminals passing through e. Similarly, let d_s be the cost function corresponding to the source s, which takes as argument a scalar variable y_s that depends on the rates that various terminals request from s. These functions c_e's and d_s's are assumed to be *convex, positive, differentiable and monotonically increasing*. Further, the functions $\int \frac{c_e(x)}{x}\ dx$ are also convex, positive, differentiable and monotonically increasing. In particular, these conditions are satisfied by functions like $x^a, a > 1$ and $xe^{bx}, b > 0$ among others.

The network connection we are interested in supporting is one where each terminal can reconstruct all the sources. i.e. we need to jointly allocate rates and flows for each terminal so that it can reconstruct the sources. We now present a formal description of the optimization problem under consideration.

Min-Cost Multicast with Multiple Sources: Let us call the quadruple $(G, c, d, \mathcal{R}_{\mathcal{SW}})$ an *instance*. The problem of minimizing the total cost for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$ can be formulated as

$$\begin{array}{ll} \text{minimize} & C(f,R) = \sum_{e\in E} c_e(z_e) + \sum_{s\in S} d_s(y_s) \\ \text{subject to} & f_P \geq 0 \quad \forall P \in \mathcal{P} \\ (NIF-CP) & \sum_{P\in\mathcal{P}_{s,t}} f_P \geq R_{s,t} \quad \forall s \in S, \forall t \in T \quad (1) \\ & \boldsymbol{R}_t \in \mathcal{R}_{\mathcal{SW}} \quad \forall t \in T \end{array}$$

where $z_e, \forall e \in E$ is a function of $x_{e,t_1}, x_{e,t_2}, \ldots, x_{e,t_{N_T}}$, that we shall denote $z_e(x_{e,t_1}, x_{e,t_2}, \ldots, x_{e,t_{N_T}})$ with $x_{e,t} = \sum_{P\in\mathcal{P}_t:e\in P} f_P \quad \forall e \in E, \quad \forall t \in T$, and $y_s, \forall s \in S$ is a function of $\boldsymbol{\rho}_s$ that we will denote $y_s(\boldsymbol{\rho}_s)$.

The formulation above is similar to the one presented in [3]. However since we consider source correlations as well, their formulation is a specific case of our formulation. Since network coding allows the sharing of edges, the penalty at an edge is only the maximum and not the sum i.e. z_e is the maximum flow (among the different terminals) across the edge e. Similarly, the penalty at the sources for higher resolution quantization is also driven by the maximum level requested by each terminal i.e. y_s is also maximum. In this work, for differentiability requirements the maximum function will be approximated as L_p norm with $p \to \infty$. Nevertheless, most of our analysis is done where z_e and y_s are non-decreasing functions partially differentiable with respect to their arguments, such that $c_e(z_e)$ and $d_s(y_s)$ are convex, positive, differentiable and monotonically increasing. Note that in the formulation above, the objective function is convex and all constraints are linear which implies that this is a convex optimization problem.

The constraint (1) above models the fact that the total flow from the source s to a terminal t needs to be at least $R_{s,t}$. Finally, the rate point of each terminal $\boldsymbol{R}_t$ needs to be within the Slepian-Wolf polytope. A flow-rate (f, R) satisfying all the conditions in the above optimization problem (i.e. *(NIF-CP)*) will be called a *feasible* flow-rate for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$ and the cost $C(f, R)$ will be referred to as *social cost* corresponding to this flow-rate. Also, we will call a solution (f^*, R^*) of the above problem as an *OPT* flow-rate for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$.

Consider a feasible flow-rate (f, R) for the above optimization problem. It can be seen that the value of the flow from $A \subseteq S$ to a terminal $t \in T$ is $\sum_{P \in \cup_{s \in A} \mathcal{P}_{s,t}} f_P \geq \sum_{s \in A} R_{s,t}$. Since $\mathbf{R}_t \in \mathcal{R}_{\mathcal{SW}}$ the result of [8] shows that random linear network coding followed by appropriate decoding at the terminals can recover the sources with high probability. Conversely the result of [2] shows the necessity of the existence of such a flow.

Terminals' Incentives and the Distributed Compression Game: The above formulation for social cost minimization for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$ disregards the fact that the agents who pay for the costs incurred at the edges and the sources may not be cooperative and may have incentives for strategic manipulation. In this work we consider the scenario where the terminals pay for the network resources they are being provided. The terminals are noncooperative and will behave selfishly trying to minimize their own respective costs without regard to the social cost, while ensuring that they can reconstruct all the sources. We have the following assumptions.

(i) Let (f, R) denote a feasible flow rate for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$. The network operates via random linear network coding over the subgraph of G induced by the corresponding $\{z_e\}$ for $e \in E$. The terminals are capable of performing appropriate decoding to recover the sources.

(ii) Each terminal $t \in T$ can request for any specific set of flows on the paths $P \in \mathcal{P}_t$ and rates $\mathbf{R}_t$ as long as such a request allows reconstruction of the sources at t. There is a mechanism in the network by means of which this request is accommodated i.e. the subgraph over which random linear network coding is performed is adjusted appropriately.

In this work we wish to characterize flow-rates that represent an equilibrium among selfish terminals while they act strategically to minimize their own costs. Furthermore, we shall systematically study the loss that occurs due to the mismatch between the social goals and terminal's selfish goals.

Towards this end, we now formally model the game originating from the selfish behavior of the terminals. We model this game as a *normal formal game*[16], i.e. a static one shot strategic game of complete information, which we refer to as *Distributed Compression Game(DCG)*.

A normal form game, denoted $(\mathcal{N}, \{A_i\}_{i \in \mathcal{N}}, \{\succeq_i\}_{i \in \mathcal{N}})$, consists of the set of *players* $\mathcal{N}$, the tuple of *set of strategies* A_i for each player $i \in \mathcal{N}$, and the tuple of *preference relations* $\succeq_i$ for each player $i \in \mathcal{N}$ on the set $\mathcal{A} = \times_{i \in \mathcal{N}} A_i$. For $a, b \in \mathcal{A}$, $a \succeq_i b$ means that the player i prefers the tuple of strategies a to the tuple of strategies b. In the context of *Distributed Compression Game*, given an instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$, these parameters are defined as follows.

1) The Distributed Compression Game:

Players: $\mathcal{N} = T$, i.e. the terminals are the players. This is because, as mentioned above, the terminals are the users and they are the ones who pay for the network resources they are being provided.

Strategies: The strategy set of a player $t \in T$ consists of tuples $(f_t, \boldsymbol{R}_t)$ where

- f_t is the vector of flows on paths going to t, i.e. the vector of values f_P for all $P \in \mathcal{P}_t$, and recall that $\boldsymbol{R}_t$ denotes the rate vector for terminal t;
- $f_P \geq 0 \quad \forall P \in \mathcal{P}_t, \sum_{P \in \mathcal{P}_{s,t}} f_P \geq R_{s,t} \quad \forall s \in S$ and $\boldsymbol{R}_t \in \mathcal{R}_{\mathcal{SW}}$.

Therefore,

$$A_t = \left\{ (f_t, \boldsymbol{R}_t) : \begin{array}{l} f_P \geq 0 \quad \forall P \in \mathcal{P}_t, \\ \sum_{P \in \mathcal{P}_{s,t}} f_P \geq R_{s,t} \quad \forall s \in S, \\ \boldsymbol{R}_t \in \mathcal{R}_{\mathcal{SW}} \end{array} \right\}. \quad (2)$$

Note that a feasible flow-rate (f, R) for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$ is an element of the set $\mathcal{A} = \times_{t \in T} A_t$ defined for the same instance.

Preference Relations: To specify the preference relation of terminal $t \in T$, we need to know how much does she pay given a feasible flow-rate (f, R) i.e. what fractions of the costs at various edges and sources are being paid by t? To this end, we need market models, i.e. mechanisms for splitting the costs among various terminals.

Edge Costs: At a flow f, the cost of an edge $e \in E$ is $c_e(z_e)$. It is split among the terminals $t \in T$, each paying a fraction of this cost. Let us say that the fraction paid by the player t is $\Psi_{e,t}(\boldsymbol{x}_e)$ i.e. the player t pays $c_e(z_e)\Psi_{e,t}(\boldsymbol{x}_e)$ for the edge e where $\boldsymbol{x}_e$ denotes the vector $(x_{e,t_1}, x_{e,t_2}, \ldots, x_{e,t_{N_T}})$. Of course, $\sum_{t \in T} \Psi_{e,t}(\boldsymbol{x}_e) = 1$ to ensure that the total cost is borne by someone or the other. The total cost borne by t across all the edges is $\sum_{e \in E} c_e(z_e)\Psi_{e,t}(\boldsymbol{x}_e)$, denoted $C_E^{(t)}(f)$.

Source Costs: At a rate R, the cost for the source s is $d_s(y_s)$, which is split among the terminals $t \in T$, such that t pays a fraction $\Phi_{s,t}(\boldsymbol{\rho}_s)$ i.e. the player t pays $d_s(y_s)\Phi_{s,t}(\boldsymbol{\rho}_s)$ for

the source s. Of course, $\sum_{t\in T}\Phi_{s,t}(\boldsymbol{\rho}_s)=1$. Therefore, the total cost borne by t for all sources, denoted $C_S^{(t)}(R)$, is $\sum_{s\in S} d_s(y_s)\Phi_{s,t}(\boldsymbol{\rho}_s)$.
Thus, with the *edge-cost-splitting mechanism* Ψ and the *source-cost-splitting mechanism* Φ, the total cost incurred by the player $t\in T$ at flow-rate (f,R) denoted $C^{(t)}(f,R)$ is

$$\begin{aligned} C^{(t)}(f,R) &= C_E^{(t)}(f)+C_S^{(t)}(R) \\ &= \sum_{e\in E} c_e(z_e)\Psi_{e,t}(\boldsymbol{x}_e)+\sum_{s\in S} d_s(y_s)\Phi_{s,t}(\boldsymbol{\rho}_s). \end{aligned}$$

Now, each terminal t would like to minimize its own cost i.e. the function $C^{(t)}(f,R)$ and therefore the preference relations $\{\succeq_t\}$ are as follows. For two flow-rates $(f,R)\in\mathcal{A}$ and $(\tilde{f},\tilde{R})\in\mathcal{A}$, $(f,R)\succeq_t(\tilde{f},\tilde{R})$ if and only if $C^{(t)}(f,R)\le C^{(t)}(\tilde{f},\tilde{R})$. Also, $(f,R)\succ_t(\tilde{f},\tilde{R})$ iff $C^{(t)}(f,R)<C^{(t)}(\tilde{f},\tilde{R})$.

We will call $(G,c,d,\mathcal{R}_{\mathcal{SW}},\Psi,\Phi)$ as an instance of the Distributed Compression Game.

2) Solution Concepts for the Distributed Compression Game: We now outline the possible solution concepts in our scenario. These are essentially dictated by the level of sophistication of the terminals. Sophistication refers to the amount of information and computational resources available to a terminal. In this work we shall work with two different solution concepts that we now discuss.

a) *Nash Equilibrium.* The solution concept of Nash equlibrium requires the complete information setting and requires each terminal to compute her best response to any given tuple of strategies of the other players. For notational simplicity, let f_{-t} be the vector of flows on paths not going to terminal t i.e. the vector of values f_P for all $P\in\mathcal{P}-\mathcal{P}_t$, therefore $f=(f_{-t},f_t)$. Similarly, $\boldsymbol{R}_{-t}$ is the vector of rates corresponding to all players other than t, therefore $R=(\boldsymbol{R}_{-t},\boldsymbol{R}_t)$. In our setting, the best response problem of a terminal t is to minimize her cost function $C^{(t)}(f_{-t},f_t,\boldsymbol{R}_{-t},\boldsymbol{R}_t)$ over $(f_t,\boldsymbol{R}_t)\in A_t$ given any $(f_{-t},\boldsymbol{R}_{-t})$. Therefore a Nash flow-rate is defined as follows.

Definition 1: **(Nash flow-rate)** A flow-rate (f,R) feasible for the instance $(G,c,d,\mathcal{R}_{\mathcal{SW}})$ is at Nash equilibrium, or is a Nash flow-rate for instance $(G,c,d,\mathcal{R}_{\mathcal{SW}},\Psi,\Phi)$, if $\forall t\in T$,

$$C^{(t)}(f,R)\le C^{(t)}(f_{-t},\tilde{f}_t,\boldsymbol{R}_{-t},\tilde{\boldsymbol{R}}_t)\quad\forall(\tilde{f}_t,\tilde{\boldsymbol{R}}_t)\in A_t.$$

We note that computing the best response will in general require a given terminal to know flow assignments on all possible paths and rate vectors for all the terminals. Moreover, convexity of the objective function in $NIF-CP$ (i.e. social cost $C(f,R)$) does not imply convexity of $C^{(t)}(f_{-t},f_t,\boldsymbol{R}_{-t},\boldsymbol{R}_t)$ in the variables $(f_t,\boldsymbol{R}_t)\in A_t$ in general. Therefore the computational requirements at the terminals may be large. Consequently Nash equilibrium does not seem to be an appropriate solution concept for the Distributed Compression Game when we look through the algorithmic lens.

b) *Waldrop Equilibrium.* From a practical standpoint, a terminal may only have partial knowledge of the system and may be computationally constrained. A solution concept more appropriate under such situations is that of local Nash equilibrium or Waldrop equilibrium that is widely adopted in selfish routing and transportation literature[20], [14], [7]. We note that this solution concept has also been utilized in [3]. We first present the precise definition of the Waldrop equilibrium in our case and then provide an intuitive justification. Towards this end, we need to define the marginal cost of a path.

Definition 2: **(Marginal Cost of a Path)** For a $P\in\mathcal{P}_t$ its marginal cost is $C_P(f):=\sum_{e\in P}\frac{c_e(z_e)\Psi_{e,t}(\boldsymbol{x}_e)}{x_{e,t}}$.
Therefore, for the terminal t, the total cost for the edges, $C_E^{(t)}$, can be equivalently written as $C_E^{(t)}(f)=\sum_{P\in\mathcal{P}_t}C_P(f)f_P$.

Definition 3: **(Waldrop flow-rate)** A flow-rate (f,R) feasible for the instance $(G,c,d,\mathcal{R}_{\mathcal{SW}})$ is at local Nash equilibrium, or is a Waldrop flow-rate for instance $(G,c,d,\mathcal{R}_{\mathcal{SW}},\Psi,\Phi)$, if it satisfies the following conditions.
(1) $\forall t\in T,\ \ \forall s\in S$, we have $\sum_{P\in\mathcal{P}_{s,t}}f_P=R_{s,t}$.
(2) $\forall t\in T$, we have $\sum_{s\in S}R_{s,t}=H(X_S)$.
(3) $\forall t\in T,\ \forall s\in S,\ P,Q\in\mathcal{P}_{s,t}$ with $f_P>0$, $C_P(f)\le C_Q(f)$.
(4) For $t\in T$, let $j\in S$ participates in **all tight** rate inequalities involving $i\in S$ (i.e. if $A\subseteq S$, such that $i\in A$ and $\sum_{l\in A}R_{l,t}=H(X_A|X_{-A})$[1], then $j\in A$) and let $P\in\mathcal{P}_{i,t},Q\in\mathcal{P}_{j,t}$ with $f_P>0$ then we have

$$C_P(f)+\frac{\partial C_S^{(t)}(R)}{\partial R_{i,t}}\le C_Q(f)+\frac{\partial C_S^{(t)}(R)}{\partial R_{j,t}}.$$

Intuitively, conditions (1) and (2) require that each terminal requests as little rate and flow as possible. Condition (3) ensures that an infitesimally small change in flow allocations from path P (where $f_P>0$) to path Q where $P,Q\in\mathcal{P}_{s,t}$, will increase the sum cost along paths in $\mathcal{P}_t$. Now, consider an infitesimally small change in flow allocation from $P\in\mathcal{P}_{i,t}$ (where $f_P>0$) to $Q\in\mathcal{P}_{j,t}$. This also requires a corresponding change in the rates requested from sources i and j by terminal t. Under certain constraints on the source j, Condition (4) ensures that the overall effect of this change will serve to increase terminal t's cost. The conditions on the source j are well-motivated in light of the characterization of Nash flow-rate at the end of section III in the case when the best response problem of every terminal is convex.

We remark that a Nash flow-rate may not always be a Waldrop flow-rate and vice versa. When sources are independent, condition (2) implies that $R_{s,t}=H(X_s)$ for all $s\in S,t\in T$ and it is not required to check the condition (4). Also we can recover condition (3) by setting $i=j$ in condition (4). They are stated separately for the sake of clarity.

As we discussed earlier, the solution concept based on Waldrop equilibrium seems more suitable to our scenario and consequently we define the price of anarchy [9], [17], [20] in terms of Waldrop flow-rate instead of Nash flow-rate.

Definition 4: **Price of Anarchy(POA)***:* Let $\mathcal{C}$ be a class of edge cost functions, $\mathcal{D}$ be a class of source cost functions, $\mathcal{G}$ be a class of networks/graphs, Ψ be an edge cost splitting

[1]We use $H(X_A|X_{-A})$ and $H(X_A|X_{A^c})$ interchangeably in the text to denote the joint entropy of the sources in set A given the remaining sources.

mechanism, Φ be a source cost splitting mechanism, and $\mathcal{M}$ be a set of Slepian-Wolf polytopes. We will refer to $(\mathcal{G}, \mathcal{C}, \mathcal{D}, \Psi, \Phi, \mathcal{M})$ as a *scenario*. The price of anarchy for the scenario $(\mathcal{G}, \mathcal{C}, \mathcal{D}, \Psi, \Phi, \mathcal{M})$, denoted $\rho(\mathcal{G}, \mathcal{C}, \mathcal{D}, \Psi, \Phi, \mathcal{M})$, is defined as

$$\rho(\mathcal{G}, \mathcal{C}, \mathcal{D}, \Psi, \Phi, \mathcal{M}) = \max_{G\in\mathcal{G}, c\in\mathcal{C}, d\in\mathcal{D}, \mathcal{R}_{\mathcal{SW}}\in\mathcal{M}} \left(\frac{1}{C_{OPT}(G, c, d, \mathcal{R}_{\mathcal{SW}})} \max_{(f,R)\text{ is a Waldrop flow-rate for }(G,c,d,\mathcal{R}_{\mathcal{SW}},\Psi,\Phi)} C(f, R) \right),$$

where $C_{OPT}(G, c, d, \mathcal{R}_{\mathcal{SW}})$ refers to the optimal cost of $NIF-CP$ for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$. Let us denote the set of Slepian-Wolf polytopes corresponding to the case where there are no source correlations (i.e. $H(X_A|X_{-A}) = H(X_A)$ for all $A \in S$) by $\mathcal{M}_{ind}$ (subscript ind denotes - *independent*) and the set of Slepian-Wolf polytopes corresponding to the case where sources are correlated (i.e. there exists $A \subseteq S$ with $H(X_A|X_{-A}) < H(X_A)$) by $\mathcal{M}_c$. Also, we use $\mathcal{G}_{all}$ to denote the class of all graphs where every $t \in T$ is connected to every $s \in S$, and $\mathcal{G}_{dsw}$ (subscript dsw denotes - *direct Slepian-Wolf*) to denote the class of complete bipartite graphs between the set of sources and the set of terminals. Note that $\mathcal{G}_{dsw}$ corresponds to the case where every terminals is directly connected to every source by an edge and no network coding is required. A question we will be most concerned with in this paper is whether $\rho(\mathcal{G}, \mathcal{C}, \mathcal{D}, \Psi, \Phi, \mathcal{M}_c) > \rho(\mathcal{G}, \mathcal{C}, \mathcal{D}, \Psi, \Phi, \mathcal{M}_{ind})$, and in particular whether $\rho(\mathcal{G}, \mathcal{C}, \mathcal{D}, \Psi, \Phi, \mathcal{M}_c) > 1$ but $\rho(\mathcal{G}, \mathcal{C}, \mathcal{D}, \Psi, \Phi, \mathcal{M}_{ind}) = 1$ for meaningful classes of cost functions $\mathcal{C}, \mathcal{D}$ and reasonable splitting mechanisms Ψ and Φ i.e. does correlation induce anarchy?

III. Characterizing the Optimal Flows and Rates

In this section, we investigate the properties of an *OPT* flow-rate via Lagrangian duality theory [4]. Since the optimization problem *(NIF-CP)* is convex and the constraints are such that the strong duality holds, the *Karush-Kuhn-Tucker(KKT)* conditions exactly characterize optimality [4]. Therefore, we start out by writing the Lagrangian dual of *NIF-CP*,

$$\begin{aligned} L = &\sum_{e\in E} c_e(z_e) + \sum_{s\in S} d_s(y_s) - \sum_{P\in\mathcal{P}} \mu_P f_P \\ &+ \sum_{s\in S}\sum_{t\in T} \lambda_{s,t}(R_{s,t} - \sum_{P\in\mathcal{P}_{s,t}} f_P) \\ &+ \sum_{t\in T}\left[\sum_{A\subseteq S} \nu_{A,t}\left(H(X_A|X_{A^c}) - \sum_{i\in A} R_{i,t}\right)\right] \end{aligned}$$

where $\mu_P \geq 0, \lambda_{s,t} \geq 0$ and $\nu_{A,t} \geq 0$ are the dual variables (i.e. Lagrange multipliers). For notational simplicity, let us denote the partial derivative of z_e with respect to $x_{e,t}$, $\frac{\partial z_e}{\partial x_{e,t}}$ by $z'_{e,t}$. Note that the partial derivative of $x_{e,t}$ w.r.t. to f_P is 1 for a $P \in \mathcal{P}_t$. Similarly, we denote the partial derivative of y_s with respect to $R_{s,t}$, $\frac{\partial y_s}{\partial R_{s,t}}$ by $y'_{s,t}$. The *KKT* conditions are then given by the following equations that hold $\forall\, s \in S, t \in T$,

$$\frac{\partial L}{\partial f_P} = \sum_{e\in P} c'_e(z_e) z'_{e,t}(\boldsymbol{x}_e) - \mu_P - \lambda_{s,t} = 0 \ \forall P \in \mathcal{P}_{s,t}, \quad (3)$$

$$\frac{\partial L}{\partial R_{s,t}} = d'_s(y_s) y'_{s,t}(\boldsymbol{\rho}_s) + \lambda_{s,t} - \sum_{A\subseteq S: s\in A} \nu_{A,t} = 0 \quad (4)$$

along with the feasibility of the flow-rate (f, R) and the complementary slackness conditions $\mu_P f_P = 0$ for all $P \in \mathcal{P}$, $\lambda_{s,t}(R_{s,t} - \sum_{P\in\mathcal{P}_{s,t}} f_P) = 0$ for all $s \in S, t \in T$, and $\nu_{A,t}\left(H(X_A|X_{A^c}) - \sum_{i\in A} R_{i,t}\right) = 0$ for all $A \subseteq S, t \in T$.

Let us now interpret the KKT conditions at the *OPT flow-rate* (f^*, R^*). Suppose that $f^*_P > 0$ for $P \in \mathcal{P}_{s,t}$. Then due to complementary slackness, we have $\mu^*_P = 0$ and consequently from equation (3) we get $\sum_{e\in P} c'_e(z^*_e) z'_{e,t}(\boldsymbol{x}^*_e) = \lambda^*_{s,t}$ i.e. if there exists another path $Q \in \mathcal{P}_{s,t}$ such that $f^*_Q > 0$ then $\sum_{e\in P} c'_e(z^*_e) z'_{e,t}(\boldsymbol{x}^*_e) = \sum_{e\in Q} c'_e(z^*_e) z'_{e,t}(\boldsymbol{x}^*_e)$. On further investigation we can obtain four necessary and sufficient conditions for optimality. Due to lack of space we will omit the proof of the necessity of these four conditions under strict convexity of the various cost functions and instead concentrate on the proof of sufficiency of these conditions for optimality.

Theorem 5: A feasible flow-rate (f, R) for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$, which satisfies the following four conditions is an OPT flow-rate for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$. Also, there is always an OPT flow-rate that satisfies these four conditions. Further, when the edge cost functions c_e for all $e \in E$ and the source cost functions d_s for all $s \in S$ are strictly convex, that is when the optimization problem *(NIF-CP)* is strictly convex, these conditions are also necessary for optimality.

1) $\forall t \in T, \quad \forall s \in S$, we have $\sum_{P\in\mathcal{P}_{s,t}} f_P = R_{s,t}$.
2) $\forall t \in T$, we have $\sum_{s\in S} R_{s,t} = H(X_S)$.
3) $\forall t \in T, \ \forall s \in S, \ P, Q \in \mathcal{P}_{s,t}$ with $f_P > 0$, $\sum_{e\in P} c'_e(z_e) z'_{e,t}(\boldsymbol{x}_e) \leq \sum_{e\in Q} c'_e(z_e) z'_{e,t}(\boldsymbol{x}_e)$.
4) For $t \in T$, suppose that there exist $i, j \in S$ that satisfy the following property. If $A \subseteq S$, such that $i \in A$ and $\sum_{l\in A} R_{l,t} = H(X_A|X_{-A})$, then $j \in A$. For such i and j let $P \in \mathcal{P}_{i,t}, Q \in \mathcal{P}_{j,t}$ with $f_P > 0$. Then
$$\begin{aligned} &\sum_{e\in P} c'_e(z_e) z'_{e,t}(\boldsymbol{x}_e) + d'_i(y_i) y'_{i,t}(\boldsymbol{\rho}_i) \\ &\leq \sum_{e\in Q} c'_e(z_e) z'_{e,t}(\boldsymbol{x}_e) + d'_j(y_j) y'_{j,t}(\boldsymbol{\rho}_j). \end{aligned}$$

Proof: We prove that the above four conditions imply optimality of (f, R). Our assumptions guarantee that the optimization problem (*NIF-CP*) for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$ is convex and since all the feasibility constraints are linear, strong duality holds [4]. This implies that the KKT conditions are necessary and sufficient for optimality. We show that a feasible flow-rate (f, R) with the above four properties satisfies the KKT conditions for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$ for a suitable choice of the dual variables given below.

Choosing $\lambda_{i,t}$'s: $\lambda_{i,t} := \min_{P\in\mathcal{P}_{i,t}} \sum_{e\in P} c'_e(z_e) z'_{e,t}(\boldsymbol{x}_e)$. Note that, using **Condition 3**, for $i \in S$, if there exist a $P_i \in \mathcal{P}_{i,t}$ such that $f_{P_i} > 0$ then we have $\lambda_{i,t} = \sum_{e \in P_i} c_e^{'}(z_e) z_{e,t}^{'}(\boldsymbol{x}_e)$.

Choosing μ_P's: For $P \in P_{i,t}$ take $\mu_P := \sum_{e \in P} c_e^{'}(z_e) z_{e,t}^{'}(\boldsymbol{x}_e) - \lambda_{i,t}$.

Choosing $\nu_{A,t}$'s: Let $h_{i,t} := d_i^{'}(y_i) y_{i,t}^{'}(\boldsymbol{\rho}_i) + \lambda_{i,t}$. Let π denote a permutation such that $0 \le h_{\pi(1),t} \le h_{\pi(2),t} \le \dots h_{\pi(N_S),t}$. Now take

$$\nu_{A,t} = \begin{cases} h_{\pi(1),t} \text{ if } A = \{\pi(1), \pi(2), \dots, \pi(N_S)\} \\ h_{\pi(i),t} - h_{\pi(i-1),t} \text{ if } A = \{\pi(i), \dots, \pi(N_S)\} \\ \qquad\qquad\qquad \text{and } 2 \le i \le N_S \\ 0 \quad \text{otherwise.} \end{cases}$$

Now, with the above choice of dual variables we will check all the KKT conditions one by one.

Dual Feasibility:

• $\lambda_{i,t} \ge 0$ as c_e and z_e are non-decreasing functions i.e. $c_e^{'}(z_e) \ge 0$ and $z_{e,t}^{'}(\boldsymbol{x}_e) \ge 0$.

• $\mu_P \ge 0$ by the definition because $\lambda_{i,t} \le \sum_{e \in P} c_e^{'}(z_e) z_{e,t}^{'}(\boldsymbol{x}_e)\ \forall P \in P_{i,t}$.

• $\nu_{A,t} \ge 0$ by definition.

KKT Conditions as per equation 3:

$$\frac{\partial L}{\partial f_P} = \sum_{e \in P} c_e^{'}(z_e) z_{e,t}^{'}(\boldsymbol{x}_e) - \lambda_{i,t} - \mu_P$$

$$= \sum_{e \in P} c_e^{'}(z_e) z_{e,t}^{'}(\boldsymbol{x}_e) - \lambda_{i,t} - \left(\sum_{e \in P} c_e^{'}(z_e) z_{e,t}^{'}(\boldsymbol{x}_e) - \lambda_{i,t} \right) = 0$$

KKT Conditions as per equation 4:

$$\frac{\partial L}{\partial R_{\pi(i),t}} = d_{\pi(i)}^{'}(y_{\pi(i)}) y_{\pi(i),t}^{'}(\boldsymbol{\rho}_{\pi(i)}) + \lambda_{\pi(i),t} - \sum_{A \subseteq S : \pi(i) \in A} \nu_{A,t}$$
$$= h_{\pi(i),t} - \sum_{A \subseteq S : \pi(i) \in A} \nu_{A,t}$$
$$= h_{\pi(i),t} - \sum_{j \in \{1,2,\dots,i\}} \nu_{\{\pi(j), \pi(j+1), \dots, \pi(N_S)\},t}$$
$$= h_{\pi(i),t} - [h_{\pi(1),t} + (h_{\pi(2),t} - h_{\pi(1),t})$$
$$+ (h_{\pi(3),t} - h_{\pi(2),t}) + \dots + (h_{\pi(i),t} - h_{\pi(i-1),t})]$$
$$= h_{\pi(i),t} - h_{\pi(i),t} = 0.$$

Complementary Slackness Conditions:

• $\mu_P f_P = 0$ for all $P \in \mathcal{P}$. Let $P \in \mathcal{P}_{i,t}$ and $f_P > 0$ then using **Condition** 3 and definition of $\lambda_{i,t}$ we get $\sum_{e \in P} c_e^{'}(z_e) z_{e,t}^{'}(\boldsymbol{x}_e) = \lambda_{i,t}$ and therefore, $\mu_P = \sum_{e \in P} c_e^{'}(z_e) z_{e,t}^{'}(\boldsymbol{x}_e) - \lambda_{i,t} = 0$.

• $\lambda_{s,t}(R_{s,t} - \sum_{P \in \mathcal{P}_{s,t}} f_P) = 0$ for all $s \in S, t \in T$. This follows from the **Condition** 1.

• $\nu_{A,t}\left(H(X_A|X_{A^c}) - \sum_{i \in A} R_{i,t}\right) = 0$ for all $A \subseteq S, t \in T$. Note that $\nu_{A,t} = 0$ except for $A = \{\pi(i), \pi(i+1), \dots, \pi(N_S)\}$, for $i = 1, 2, \dots, N_S$. Therefore the only condition that needs to be checked is that if $\sum_{j=i}^{N_S} R_{\pi(j),t} > H(X_{\pi(i)}, X_{\pi(i+1)}, \dots, X_{\pi(N_S)} | X_{\pi(i-1)}, \dots, X_{\pi(1)})$, then $h_{\pi(i),t} - h_{\pi(i-1),t} = 0$.

Towards this end let $j \in \{\pi(i), \pi(i+1), \dots, \pi(N_S)\}$, and let A_j be the minimum cardinality set such that $j \in A_j$ and $\sum_{l \in A_j} R_{l,t} = H(X_{A_j}|X_{-A_j})$ i.e.

$$A_j = \arg \min_{A \subseteq S : j \in A, \sum_{l \in A} R_{l,t} = H(X_A|X_{-A})} |A|.$$

Such a set A_j always exists because from **Condition** 2 we have $\sum_{l=1}^{N_S} R_{l,t} = H(X_1, \dots, X_{N_S})$ and therefore the set $\{A \subseteq S : j \in A, \sum_{l \in A} R_{l,t} = H(X_A|X_{-A})\}$ is not empty.

We claim that there exists a $j^* \in \{\pi(i), \pi(i+1), \dots, \pi(N_S)\}$ such that $A_{j^*} \cap \{\pi(1), \pi(2), \dots, \pi(i-1)\}$ is not empty. If this is not true then clearly we have $\cup_{j=\pi(i)}^{\pi(N_S)} A_j = \{\pi(i), \pi(i+1), \dots, \pi(N_S)\}$ and using the supermodularity property of conditional entropy (ref. **Lemma 10**), we obtain $\sum_{j=\pi(i)}^{\pi(N_S)} R_{j,t} = H(X_{\pi(i)}, X_{\pi(i+1)}, \dots, X_{\pi(N_S)} | X_{\pi(i-1)}, \dots, X_{\pi(1)})$ which is a contradiction, therefore we must have such a $j^* \in \{\pi(i), \pi(i+1), \dots, \pi(N_S)\}$ such that $A_{j^*} \cap \{\pi(1), \pi(2), \dots, \pi(i-1)\}$ is not empty.

Next, we show that there exists a source $k \in \{\pi(1), \pi(2), \dots, \pi(i-1)\}$ such that if $j^* \in A$ and $\sum_{l \in A} R_{l,t} = H(X_A|X_{-A})$, then $k \in A$. Towards this end suppose that there exist subsets S_1 and S_2 of such that $j^* \in S_1 \cap S_2$ and $\sum_{l \in S_1} R_{l,t} = H(X_{S_1}|X_{-S_1})$ and $\sum_{l \in S_2} R_{l,t} = H(X_{S_2}|X_{-S_2})$, then using the supermodularity property of conditional entropy we can show that rate inequality involving $S_1 \cap S_2$ is also tight (**Lemma 10**) i.e. $\sum_{l \in S_1 \cap S_2} R_{l,t} = H(X_{S_1 \cap S_2}|X_{-(S_1 \cap S_2)})$. This implies that A_{j^*}, being of minimum cardinality, is the intersection of all sets that have j^* as a member on which the rate inequality is tight i.e.

$$A_{j^*} = \bigcap_{A \subseteq S} \{A : j^* \in A, \sum_{l \in A} R_{l,t} = H(X_A|X_{-A})\}.$$

Moreover note that A_{j^*} is not a singleton set since $A_{j^*} \cap \{\pi(1), \pi(2), \dots, \pi(i-1)\} \ne \phi$. Therefore there exists a $k \in A_{j^*}$ such that $k \ne j^*$. By our above arguments this implies that if $A \subseteq S$ is such that $j^* \in A$ and $\sum_{l \in A} R_{l,t} = H(X_A|X_{-A})$ then $k \in A$.

Clearly, $R_{j^*,t} > H(X_{j^*}|X_{-j^*})$ as k does not participate in this rate inequality. Therefore, $R_{j^*,t} > 0$ which implies that there exists a $P \in \mathcal{P}_{j^*,t}$ with $f_P > 0$, therefore using **Condition** 3 and the definition of $\lambda_{j^*,t}$ we have $\sum_{e \in P} c_e^{'}(z_e) z_{e,t}^{'}(\boldsymbol{x}_e) = \lambda_{j^*,t}$. Also, by the definition of $\lambda_{k,t}$ there is a $Q \in \mathcal{P}_{k,t}$ such that $\sum_{e \in Q} c_e^{'}(z_e) z_{e,t}^{'}(\boldsymbol{x}_e) = \lambda_{k,t}$.

Now using **Condition** 4, we get

$$\sum_{e \in P} c_e^{'}(z_e) z_{e,t}^{'}(\boldsymbol{x}_e) + d_{j^*}^{'}(y_{j^*}) y_{j^*,t}^{'}(\boldsymbol{\rho}_{j^*})$$
$$\le \sum_{e \in Q} c_e^{'}(z_e) z_{e,t}^{'}(\boldsymbol{x}_e) + d_k^{'}(y_k) y_{k,t}^{'}(\boldsymbol{\rho}_k) \quad \forall Q \in \mathcal{P}_{k,t}$$

which implies that

$$\lambda_{j^*,t} + d_{j^*}^{'}(y_{j^*}) y_{j^*,t}^{'}(\boldsymbol{\rho}_{j^*}) \le \lambda_{k,t} + d_k^{'}(y_k) y_{k,t}^{'}(\boldsymbol{\rho}_k)$$

and therefore we get $h_{j^*,t} \le h_{k,t}$. Now note that $k \in \{\pi(1), \pi(2), \dots, \pi(i-1)\}$ while $j^* \in \{\pi(i), \dots, \pi(N_S)\}$. This implies in turn that $h_{\pi(i),t} \le h_{j^*,t} \le h_{k,t}$. But, we know that $h_{k,t} \le h_{\pi(i-1),t}$ i.e. $h_{\pi(i),t} - h_{\pi(i-1),t} \le 0$ but we already have $h_{\pi(i),t} - h_{\pi(i-1),t} \ge 0$ and hence $h_{\pi(i),t} - h_{\pi(i-1),t} = 0$.

Corollary 6: If the sources are independent (i.e. $\mathcal{R}_{\mathcal{SW}} \in \mathcal{M}_{ind}$), there is a feasible flow-rate for instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$

that is an OPT flow-rate for both the instances $(G, c, d, \mathcal{R}_{\mathcal{SW}})$ and $(G, \tilde{c}, \tilde{d}, \mathcal{R}_{\mathcal{SW}})$, where $\tilde{c}_e(x) = \alpha c_e(x)$ for constant $\alpha > 0$, and $\tilde{d}_s$ is any convex, differentiable, positive and non-decreasing function. Further, this OPT flow-rate satisfies the four conditions in Theorem 5 for both the instances $(G, c, d, \mathcal{R}_{\mathcal{SW}})$ and $(G, \tilde{c}, \tilde{d}, \mathcal{R}_{\mathcal{SW}})$.

We omit the proof due to lack of space. The idea is that when the sources are independent, Condition (2) in Theorem 5 implies that $R_{s,t} = H(X_s)$ for all $s \in S, t \in T$, and therefore, there is no pair (i, j) such that j participates in all tight rate inequalities involving i and consequently it is not required to check Condition (4).

We conclude this section with an important note that whenever the best response problem of each terminal is convex, using an approach essentially similar to the proof of Theorem 5, it can be shown that the four conditions in the definitions of Waldrop flow-rate but with $C_P(f)$ replaced by $\frac{\partial C_E^{(t)}(f)}{\partial f_P}$ characterizes the Nash flow-rate. Further, under similar convexity conditions, we can also show that a Nash flow-rate always exists for the *Distributed Compression Game*. The proofs are omitted due to space constraints.

IV. Waldrop Flow-Rate and the Price of Anarchy

In this section, we investigate the inefficiency brought forth by the selfish behavior of terminals. First, we will show that the Waldrop equilibrium is a socially optimal solution for a different set of (related) cost functions. Using this, we will construct explicit examples that demonstrate that the POA > 1 and determine near-tight upper bounds on the POA as well. We start out with the characterization of Waldrop flow-rate.

Theorem 7: Let $z_e(\boldsymbol{x}_e) = \left(\sum_{t\in T} x_{e,t}^n\right)^{\frac{1}{n}}, \Psi_{e,t}(\boldsymbol{x}_e) = \frac{x_{e,t}^n}{\left(\sum_{j\in T} x_{e,j}^n\right)}$ and $\Phi_{s,t}(\boldsymbol{\rho}_s) = \frac{1}{N_T}$. A Waldrop flow-rate for $(G, c, d, \mathcal{R}_{\mathcal{SW}}, \Psi, \Phi)$ is an OPT flow-rate for $(G, \tilde{c}, d, \mathcal{R}_{\mathcal{SW}})$, where $\tilde{c}_e(x) = N_T \int \frac{c_e(x)}{x} dx$. Further, when the edge cost functions c_e for all $e \in E$ and the source cost functions d_s for all $s \in S$ are strictly convex, an OPT flow-rate for $(G, c, d, \mathcal{R}_{\mathcal{SW}})$ is also a Waldrop flow-rate for $(G, \hat{c}, d, \mathcal{R}_{\mathcal{SW}}, \Psi, \Phi)$, where $\hat{c}_e(x) = \frac{1}{N_T} x c_e'(x)$.

Proof: We will show that the definition of a Waldrop flow-rate for instance $(G, c, d, \mathcal{R}_{\mathcal{SW}}, \Psi, \Phi)$ exactly corresponds to the four conditions for the instance $(G, \tilde{c}, d, \mathcal{R}_{\mathcal{SW}})$ in Theorem 5.

We have, $z_{e,t}'(\boldsymbol{x}_e) = \frac{1}{n}\left(\sum_{j\in T} x_{e,j}^n\right)^{\frac{1}{n}-1} n x_{e,t}^{n-1} = \frac{z_e}{x_{e,t}} \frac{x_{e,t}^n}{\sum_{j\in T} x_{e,j}^n}$. Therefore,

$$C_P(f) = \sum_{e\in P} c_e(z_e) \frac{x_{e,t}^{n-1}}{\left(\sum_{j\in T} x_{e,j}^n\right)}$$
$$= \sum_{e\in P} c_e(z_e) \frac{z_{e,t}'(x_e)}{z_e} = \frac{1}{N_T} \sum_{e\in P} \tilde{c}_e'(z_e) z_{e,t}'(x_e)$$

where the last equality follows from the fact that $\tilde{c}_e(x) = N_T \int \frac{c_e(x)}{x} dx \implies \tilde{c}_e'(x) = N_T \frac{c_e(x)}{x}$. Also, $C_S^{(t)}(R) = \frac{1}{N_T} \sum_{i\in S} d_i(y_i), \implies \frac{\partial C_S^{(t)}(R)}{\partial R_{i,t}} = \frac{1}{N_T} d_i'(y_i) y_{i,t}'(\boldsymbol{\rho}_i)$. Therefore, $C_P(f) + \frac{\partial C_S^{(t)}(R)}{\partial R_{i,t}} = \frac{1}{N_T}\left[\sum_{e\in P} \tilde{c}_e'(z_e) z_{e,t}'(\boldsymbol{x}_e) + d_i'(y_i) y_{i,t}'(\boldsymbol{\rho}_i)\right]$. The result follows from the equivalence of conditions coming from Definition 3 and Theorem 5. ■

In contrast with the result of [3] that holds for a single source with the edge cost splitting mechanism used above, from Theorem 7, we can note that for most reasonable cost splitting mechanisms, the POA will not equal one for all monomial edge cost functions. We construct explicit examples for POA > 1 in the Figure 1. The example in Figure 1(a) is near tight as will be evident from an upper bound on POA derived in Theorem 9.

It is interesting to note that in the case when sources are independent, in the Waldrop or OPT solutions, the rates requested at various sources will equal their respective lower bounds (i.e. their entropy). Therefore, the cost term corresponding to the sources will be fixed, and one only needs to find flows that minimize the edge costs. In this situation, it is not hard to see that the POA will again equal one for *all* monomial edge cost functions. i.e. *it is the correlation among the sources that is responsible for bringing more anarchy*. We formalize this below.

Let $\mathcal{C}_k = \{c : c_e(x) = a_e x^k,\ a_e > 0, \forall e \in E\}$ be the set of edge cost functions where all edge cost functions are monomial of the same degree k possibly with different coefficients, and $\mathcal{C}_{mon} = \cup_{k\geq 1}\mathcal{C}_k$. Similarly, $\mathcal{D}_k = \{d : d_i(y) = b_i x^k,\ b_i > 0, \forall s \in S\}$. Also, let $D_{convex} = \{d : d_i \text{ is convex } \forall i \in S\}$.

Corollary 8: **Correlation Induces Anarchy:** Let $z_e(\boldsymbol{x}_e) = \left(\sum_{t\in T} x_{e,t}^n\right)^{\frac{1}{n}}$, $\Psi_{e,t}(\boldsymbol{x}_e) = \frac{x_{e,t}^n}{\left(\sum_{j\in T} x_{e,j}^n\right)}$, $y_s(\boldsymbol{\rho}_s) = \left(\sum_{t\in T} R_{s,t}^m\right)^{\frac{1}{m}}$, and $\Phi_{s,t}(\boldsymbol{\rho}_s) = \frac{1}{N_T}$, then we have

1) $\rho(\mathcal{G}_{all}, \mathcal{C}_{mon}, \mathcal{D}_{convex}, \Psi, \Phi, \mathcal{M}_{ind}) = 1$.
2) $\rho(\mathcal{G}_{all}, \mathcal{C}_{N_T}, \mathcal{D}_{convex}, \Psi, \Phi, \mathcal{M}_c) = 1$.
3) $\rho(\mathcal{G}_{all}, \mathcal{C}_{mon}, \mathcal{D}_{convex}, \Psi, \Phi, \mathcal{M}_c) > 1$ for large values of m and n. In fact, $\rho(\mathcal{G}_{all}, \mathcal{C}_1, \mathcal{D}_2, \Psi, \Phi, \mathcal{M}_c) > \frac{1+N_T}{5}$.
4) $\rho(\mathcal{G}_{dsw}, \mathcal{C}_{mon}, \mathcal{D}_{convex}, \Psi, \Phi, \mathcal{M}_c) > 1$ for large values of m and n.

Proof: Let $c \in \mathcal{C}_{mon}$ i.e. $c_e(x) = a_e x^k$ for $a_e > 0$ for all $e \in E$, therefore, $\int \frac{c_e(x)}{x} dx = \int a_e x^{k-1}\, dx = a_e \frac{1}{k} x^k = \frac{1}{k} c_e(x)$. Also, $d \in \mathcal{D}_{convex}$. Now, since the sources are independent (i.e. $\mathcal{R}_{\mathcal{SW}} \in \mathcal{M}_{ind}$), from Theorem 7 and Corollary 6 it follows that a Waldrop flow-rate for instance $(G, c, d, \mathcal{R}_{\mathcal{SW}}, \Psi, \Phi)$ is also an OPT flow-rate for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$ which implies that $\rho(\mathcal{G}_{all}, \mathcal{C}_{mon}, \mathcal{D}_{convex}, \Psi, \Phi, \mathcal{M}_{ind}) = 1$.

Even if the sources are correlated, when we have $k = N_T$, we have $N_T \int \frac{c_e(x)}{x} dx = c_e(x)$ and using Theorem 7, a Waldrop flow-rate for instance $(G, c, d, \mathcal{R}_{\mathcal{SW}}, \Psi, \Phi)$ is also an OPT flow-rate for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$ which implies that $\rho(\mathcal{G}_{all}, \mathcal{C}_{N_T}, \mathcal{D}_{convex}, \Psi, \Phi, \mathcal{M}_c) = 1$.

We prove $\rho(\mathcal{G}_{all}, \mathcal{C}_1, \mathcal{D}_2, \Psi, \Phi, \mathcal{M}_c) > \frac{1+N_T}{5}$ and consequently $\rho(\mathcal{G}_{all}, \mathcal{C}_{mon}, \mathcal{D}_{convex}, \Psi, \Phi, \mathcal{M}_c) > 1$, by explicitly constructing an example as provided in Figure 1(a). All sources are identical with entropy h, therefore, $\mathcal{R}_{\mathcal{SW}} \in \mathcal{M}_c$.

Further, $d_s(y) = C_1 y^2$ for all $s \in S$, therefore, $d \in \mathcal{D}_2$. All edge cost functions are $c_e(x) = x$ except for the edge (u, v) for which $c_e(x) = C_2\ x$. Therefore, $c \in \mathcal{C}_1$. Let us consider the following flow-rate (f, R)

$$R_{1,t} = h\ \forall t \in T,\ R_{s,t} = 0\ \forall s \in S - \{1\}, t \in T$$
$$f_{(1,t)} = h\ \forall t \in T, f_P = 0\ \forall P \in \mathcal{P}_t - \{(1,t)\}, t \in T.$$

Clearly, (f, R) is feasible for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$. We claim that (f, R) is a Waldrop flow-rate for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}}, \Psi, \Phi)$ when $\frac{2C_1 h}{N_T} \leq 1 + C_2$. To see this, first note that (f, R) satisfies the Conditions (1) and (2) in the definition of Waldrop flow-rate (Definition 3) for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}}, \Psi, \Phi)$. To check the conditions (3) and (4) in the definition, we compute the marginal cost for the various paths and the differential cost for various sources. Note that $\Psi_{e,t}(\boldsymbol{x}_e) = \frac{1}{N_T}$ whenever $x_{e,t} = x$ for all $t \in T$ for some $x > 0$ and by continuity this is true even if $x = 0$. Therefore,

$$C_{(1,t)}(f) = \sum_{e \in \{(1,t)\}} \frac{c_e(z_e)\Psi_{e,t}(\boldsymbol{x}_e)}{x_{e,t}} = \frac{h\ .\ 1}{h} = 1,$$
$$C_{(1,u,v,t)}(f) = \sum_{e \in \{(1,u),(u,v),(v,t)\}} \frac{c_e(z_e)\Psi_{e,t}(\boldsymbol{x}_e)}{x_{e,t}}$$
$$= \lim_{x \longrightarrow 0} \left[\frac{x\ .\ (1/N_T)}{x} + \frac{C_2 x\ .\ (1/N_T)}{x} + \frac{x\ .\ 1}{x} \right] = 1 + \frac{1 + C_2}{N_T},$$
$$C_{(s,u,v,t)}(f) = 1 + \frac{1 + C_2}{N_T},\ s \in S - \{1\}.$$

Clearly, the condition (3) is satisfied as $C_{(1,t)}(f) < C_{(1,u,v,t)}(f)$. Also,

$$\frac{\partial C_S^{(t)}(R)}{\partial R_{i,t}} = \frac{1}{N_T} d_i^{'}(y_i) y_{i,t}^{'}(\boldsymbol{\rho}_i) = \frac{1}{N_T} 2C_1 y_i y_{i,t}^{'}(\boldsymbol{\rho}_i)$$
$$= \frac{2C_1}{N_T} y_i^2 \frac{R_{i,t}^{m-1}}{\sum_{j \in T} R_{i,j}^m} = \frac{2C_1}{N_T} \left(\sum_{j \in T} R_{i,j}^m \right)^{2/m} \frac{R_{i,t}^{m-1}}{\sum_{j \in T} R_{i,j}^m}.$$
$$\therefore\ \frac{\partial C_S^{(t)}(R)}{\partial R_{1,t}} = \frac{2C_1}{N_T} \left(N_T h^m\right)^{2/m} \frac{h^{m-1}}{N_T h^m} = \frac{2C_1 h}{N_T^2}$$
$$\text{as } m \to \infty, \text{ and } \frac{\partial C_S^{(t)}(R)}{\partial R_{s,t}} \geq 0, \forall s \in S - \{1\}.$$

Therefore, when $\frac{2C_1 h}{N_T} \leq 1 + C_2$, we get $C_{(1,t)}(f) + \frac{\partial C_S^{(t)}(R)}{\partial R_{1,t}} \leq C_{(s,u,v,t)}(f) + \frac{\partial C_S^{(t)}(R)}{\partial R_{s,t}}$ for all $s \in S - \{1\}$ which implies that the condition (4) is also satisfied. Thus, (f, R) is indeed a Waldrop flow-rate for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}}, \Psi, \Phi)$. Further after simple calculation we get, $C(f, R) = N_T h + C_1 h^2$, as $m \to \infty$. Now let us consider another flow-rate (f^*, R^*)

$$R_{s,t}^* = \frac{h}{N_S} \quad \forall s \in S, t \in T$$
$$f_{(1,t)}^* = 0\ \forall t \in T, \text{and } f_{(s,u,v,t)}^* = \frac{h}{N_S} \quad \forall s \in S, t \in T.$$

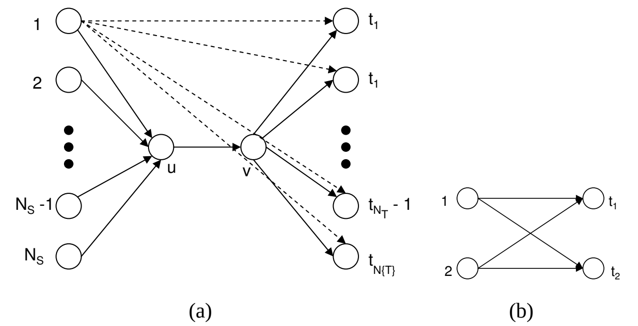

Fig. 1. (a) Example of a network where POA is linear in N_T. (b) Classical Slepian-Wolf network with appropriate costs also has POA > 1.

Clearly, (f^*, R^*) is feasible for the instance $(G, c, d, \mathcal{R}_{\mathcal{SW}})$. Further, we get $C(f^*, R^*) = h(1 + C_2 + N_T) + \frac{C_1 h^2}{N_S}$ as $m \to \infty, n \to \infty$.

Thus, when $\frac{1+C_2}{C_1} < h\ (1 - \frac{1}{N_S})$, we have $C(f^*, R^*) < C(f, R)$. As $OPT(G, c, d, \mathcal{R}_{\mathcal{SW}}) \leq C(f^*, R^*)$, this implies that the POA is greater than one. In particular, $\rho(\mathcal{G}_{all}, \mathcal{C}_1, \mathcal{D}_2, \Psi, \Phi, \mathcal{M}_c) > \frac{C_1 + \frac{N_T}{h}}{\frac{1+C_2+N_T}{h} + \frac{C_1}{N_S}}$. Now, take $h = 1, N_S = N_T > 4, 1 + C_2 = 3N_T, C_1 = N_T^2$, and note that $\frac{2C_1 h}{N_T} = 2N_T < 3N_T = 1 + C_2$, as well as, $\frac{1+C_2}{C_1} = \frac{3}{N_T} < (1 - \frac{1}{N_T}) = (1 - \frac{1}{N_S})$ as $N_T > 4$. Therefore, we get $\rho(\mathcal{G}_{all}, \mathcal{C}_1, \mathcal{D}_2, \Psi, \Phi, \mathcal{M}_c) > \frac{1+N_T}{5}$. This is near tight as evident from Theorem 9. To establish (4), we will prove a stronger result, $\rho(\mathcal{G}_{dsw}, \mathcal{C}_3, \mathcal{D}_3, \Psi, \Phi, \mathcal{M}_c) > 1$, by constructing an example as described below. As shown in Figure 1(b) there are two sources and two terminals which are directly connected to each source. Both sources are identical with entropy 1, $d_1(y) = C_1 y^3, d_2(y) = C_2 y^3$ with $C_1, C_2 > 0, C_1 \neq C_2$ and $c_e(x) = x^3$ for all edges. We now outline the argument that shows that the POA > 1.

First, observe that the instance is symmetric with respect to terminals and all cost functions are strictly convex. Therefore the OPT flow rate for the instance, denoted (f^*, R^*) is such that $R_{s,t_1}^* = R_{s,t_2}^*$ for $s = 1, 2$. Next, by the characterization of Waldrop as per Theorem 7, the Waldrop flow-rate, denoted (f, R) is an OPT flow-rate for $\tilde{c}_e(x) = \frac{2}{3}x^3$ with the source cost functions remaining the same. This new instance with $\tilde{c}_e(x) = \frac{2}{3}x^3$ is also symmetric with respect to the terminals and the cost functions remain strictly convex. Therefore we conclude that for the Waldrop flow-rate as well $R_{s,t_1} = R_{s,t_2}$ for $s = 1, 2$. Let $R_{1,t_1} = R_{1,t_2} = h$ and $R_{1,t_1}^* = R_{1,t_2}^* = h^*$. Using the properties of Waldrop flow-rate and OPT flow rate as per Condition (2) in Theorem 5, we have $R_{2,t_1} = R_{2,t_2} = 1 - h$ and $R_{2,t_1}^* = R_{1,t_2}^* = 1 - h^*$. We argue below that $h \neq h^*$. Consequently, by uniqueness of the OPT flow-rate (due to strict convexity of the objective function) we will have $C(f, R) > C(f^*, R^*)$ implying $\rho(\mathcal{G}_{dsw}, \mathcal{C}_3, \mathcal{D}_3, \Psi, \Phi, \mathcal{M}_c) > 1$. We have, for $t = t_1, t_2$, $\frac{\partial C_S^{(t)}(R)}{\partial R_{1,t}} = \frac{1}{N_T} d_1^{'}(y_1) y_{1,t}^{'}(\boldsymbol{\rho}_1) = \frac{3}{2} C_1 y_1^2 y_1 \frac{R_{1,t}^{m-1}}{\sum_{j=1}^{2} R_{1,j}^m} = \frac{3}{4} C_1 h^2$ as $m \to \infty$.

Similarly, $\frac{\partial C_S^{(t)}(R)}{\partial R_{2,t}} = \frac{3}{4} C_2 (1 - h)^2$. By the definition of Waldrop flow-rate, we have $f_{(1,t)} = h, f_{(2,t)} = (1 - h)$. Thus,

$C_{(1,t)}(f) = h^2, C_{(2,t)}(f) = (1-h)^2$. Further, $\frac{\partial C_S^{(t)}(R)}{\partial R_{1,t}} + C_{(1,t)}(f) = \frac{\partial C_S^{(t)}(R)}{\partial R_{2,t}} + C_{(2,t)}(f)$ implies that $\frac{3}{4}C_1 h^2 + h^2 = \frac{3}{4}C_2(1-h)^2 + (1-h)^2$. Therefore, $\frac{h}{1-h} = \sqrt{\frac{\frac{3}{4}C_2+1}{\frac{3}{4}C_1+1}}$. Now, from Theorem 7, (f^*, R^*) is a Waldrop flow-rate for the instance where everything remains the same except for the edge cost functions which are now $\frac{3}{2}x^3$ instead of x^3 and performing the similar calculations as above for (f, R), we obtain $\frac{h^*}{1-h^*} = \sqrt{\frac{\frac{3}{4}C_2+\frac{3}{2}}{\frac{3}{4}C_1+\frac{3}{2}}}$. Clearly, since $C_1 \neq C_2$, we get $h \neq h^*$. In particular, take $C_1 = 4, C_2 = 8$, then $h = 0.5695$ and $h^* = 0.5635$. Thus, $C(f, R) = 1.9061, C(f^*, R^*) = 1.9052$ implying that $POA \geq 1.004 > 1$, in this example. ■

We now state an upper bound that is nearly attained by the example of Figure 1(a).

Theorem 9: Let $z_e(\boldsymbol{x}_e) = \left(\sum_{t \in T} x_{e,t}^n\right)^{\frac{1}{n}}, \Psi_{e,t}(\boldsymbol{x}_e) = \frac{x_{e,t}^n}{\left(\sum_{j \in T} x_{e,j}^n\right)}$ and $\Phi_{s,t}(\boldsymbol{\rho}_s) = \frac{1}{N_T}$. Then, $\rho(\mathcal{G}_{all}, \mathcal{C}_k, \mathcal{D}_{convex}, \Psi, \Phi, \mathcal{M}_c) \leq \max\{\frac{N_T}{k}, \frac{k}{N_T}\}$.

Proof: Omitted due to lack of space.

Now we consider another splitting mechanism Φ that looks more like the edge cost splitting mechanism Ψ. Specifically, take $y_s(\boldsymbol{\rho}_s) = \left(\sum_{t \in T}(R_{s,t})^m\right)^{\frac{1}{m}}$ and $\Phi_{i,t}(\boldsymbol{\rho}_s) = \frac{(R_{i,t})^m}{\sum_{j \in T}(R_{j,t})^m}$. Arguing in a manner similar to Corollary 8(1) (rates need to equal their corresponding entropies) we obtain $\rho(\mathcal{G}_{all}, \mathcal{C}_{mon}, \mathcal{D}_{convex}, \Psi, \Phi, \mathcal{M}_{nc}) = 1$. Now, we will argue that with $y_s(\boldsymbol{\rho}_s) = \left(\sum_{t \in T}(R_{s,t})^m\right)^{\frac{1}{m}}$ and $\Phi_{i,t}(\boldsymbol{\rho}_s) = \frac{(R_{i,t})^m}{\sum_{j \in T}(R_{j,t})^m}$ we have $\rho(\mathcal{G}_{dsw}, \mathcal{C}_{mon}, \mathcal{D}_{convex}, \Psi, \Phi, \mathcal{M}_c) > 1$ for large values of m and n. Let us consider the same example as in Figure 1(b) but with the new source cost splitting mechanism. The previously calculated OPT flow-rate for this instance (f^*, R^*) is given by $R^*_{1,t} = f^*_{(1,t)} = h^*$ and $R^*_{2,t} = f^*_{(2,t)} = 1 - h^*$. We will argue that this is not a Waldrop flow-rate and since the OPT flow-rate is unique (by strict convexity) we will obtain $POA > 1$. After some simple calculations we get

$$\frac{\partial C_S^{(t)}(R)}{\partial R_{i,t}} = d_i^{'}(y_i)\frac{y_i}{R_{i,t}}\Phi_{i,t}^2(\boldsymbol{\rho}_i) + m\frac{d_i(y_i)}{R_{i,t}}\Phi_{i,t}(\boldsymbol{\rho}_i)\left(1 - \Phi_{i,t}(\boldsymbol{\rho}_i)\right).$$

Therefore, $\frac{\partial C_S^{(t)}(R^*)}{\partial R_{1,t}} = (m+3)(N_T)^{3/m}\frac{C_1}{4}(h^*)^2$ and $\frac{\partial C_S^{(t)}(R^*)}{\partial R_{2,t}} = (m+3)(N_T)^{3/m}\frac{C_2}{4}(1-h^*)^2$. Also, $C_{(1,t)}(f^*) = (h^*)^2$ and $C_{(1,t)}(f^*) = (1-h^*)^2$. Note that $N_T = 2$ in this example. Now, with $C_1 = 4, C_2 = 8$, we have $h^* = 0.5635$ and therefore $\frac{C_{(1,t)}(f^*) + \frac{\partial C_S^{(t)}(R^*)}{\partial R_{1,t}}}{C_{(2,t)}(f^*) + \frac{\partial C_S^{(t)}(R^*)}{\partial R_{2,t}}} = \frac{(h^*)^2 + (m+3)(N_T)^{3/m}\frac{C_1}{4}(h^*)^2}{(1-h^*)^2 + (m+3)(N_T)^{3/m}\frac{C_2}{4}(1-h^*)^2} = \frac{(m+3)(N_T)^{3/m}+1}{2(m+3)(N_T)^{3/m}+1}\frac{0.5635^2}{(1-0.5635)^2} = \frac{1}{2}\frac{0.5635^2}{(1-0.5635)^2} = 0.8333 \neq 1$ as $m \to \infty$.

REFERENCES

[1] R. Ahlswede, N. Cai, S.-Y. Li, and R. W. Yeung. Network Information Flow. *IEEE Trans. on Info. Th.*, 46, no. 4:1204–1216, 2000.

[2] J. Barros and S. Servetto. Network Information Flow with Correlated Sources. *IEEE Trans. on Info. Th.*, 52:155–170, Jan. 2006.

[3] S. Bhadra, S. Shakkottai, and P. Gupta. Min-Cost Selfish Multicast With Network Coding. *IEEE Trans. on Info. Th.*, 52, no. 11:5077–5087, 2006.

[4] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[5] T. Cover and J. Thomas. *Elements of Information Theory*. Wiley Series, 1991.

[6] R. Cristescu, B. Beferull-Lozano, and M. Vetterli. Networked Slepian-Wolf: theory, algorithms, and scaling laws. *IEEE Trans. on Info. Th.*, 51, no. 12:4057–4073, 2005.

[7] S. C. Dafermos and F. T. Sparrow. The traffic assignment problem for a general network. *J. Res. Nat. Bureau of Standards, Series B*, 73B(2):91–118, 1969.

[8] T. Ho, M. Médard, M. Effros, and R. Koetter. Network Coding for Correlated Sources. In *Conf. on Information Sciences and Systems*, 2004.

[9] E. Koutsoupias and C. H. Papadimitriou. Worst-case equilibria. In *STACS*, pages 404–413, 1999.

[10] Z. Li. Min-Cost Multicast of Selfish Information Flows. In *Proc. of IEEE INFOCOM*, 2007.

[11] Z. Li. Cross-Monotonic Multicast. In *Proc. of IEEE INFOCOM*, 2008.

[12] Z. Li and B. Li. Efficient and Distributed Computation of Maximum Multicast Rates. In *Proc. of IEEE INFOCOM*, 2005.

[13] D. S. Lun, N. Ratnakar, M. Médard, R. Koetter, D. R. Karger, T. Ho, E. Ahmed, and F. Zhao. Minimum-Cost Multicast over Coded Packet Networks. *IEEE Trans. on Info. Th.*, 52:2608–2623, June 2006.

[14] C. B. M. M Beckman and C. B. Winsten. *Studies in the Economics of Transportation*. Yale University Press, New Haven,CT, 1956.

[15] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, New York, NY, USA, 2007.

[16] M. J. Osborne and A. Rubinstein. *A Course in Game Theory*. The MIT Press, Cambridge, MA, USA, 1994.

[17] C. H. Papadimitriou. Algorithms, games, and the internet. In *STOC '01: Proceedings of the thirty-third annual ACM symposium on Theory of computing*, pages 749–753, New York, NY, USA, 2001. ACM.

[18] V. Prabhakaran, D. Tse, and K. Ramchandran. Rate Region of the quadratic Gaussian CEO problem. In *IEEE Intl. Symposium on Info. Th.*, pages 119–119, 2004.

[19] A. Ramamoorthy. Minimum cost distributed source coding over a network. In *IEEE Intl. Symposium on Info. Th.*, pages 1761–1765, 2007.

[20] T. Roughgarden. *Selfish Routing and the Price of Anarchy*. The MIT Press, 2005.

[21] T. Roughgarden and Éva Tardos. How bad is selfish routing? *J. ACM*, 49(2):236–259, 2002.

[22] D. Slepian and J. Wolf. Noiseless coding of correlated information sources. *IEEE Trans. on Info. Th.*, 19:471–480, Jul. 1973.

[23] R. Zamir and T. Berger. Multiterminal source coding with high resolution. *IEEE Trans. on Info. Th.*, 45, no. 1:106–117, 1999.

APPENDIX

Lemma 10: Let $\boldsymbol{R}_t \in \mathcal{R}_{\mathcal{SW}}$ i.e. $\sum_{l \in A} R_{l,t} \geq H(X_A|X_{-A})$ for all $A \subseteq S$. If $S_1, S_2 \in S$ satisfy $\sum_{l \in S_1} R_{l,t} = H(X_{S_1}|X_{-S_1})$ and $\sum_{l \in S_2} R_{l,t} = H(X_{S_2}|X_{-S_2})$ then we have $\sum_{l \in S_1 \cap S_2} R_{l,t} = H(X_{S_1 \cap S_2}|X_{-(S_1 \cap S_2)})$ and $\sum_{l \in S_1 \cup S_2} R_{l,t} = H(X_{S_1 \cup S_2}|X_{-(S_1 \cup S_2)})$.

Proof: We have, $\sum_{l \in S_1 \cap S_2} R_{l,t} + \sum_{l \in S_1 \cup S_2} R_{l,t} = \sum_{l \in S_1} R_{l,t} + \sum_{l \in S_2} R_{l,t} = H(X_{S_1}|X_{-S_1}) + H(X_{S_2}|X_{-S_2}) \leq H(X_{S_1 \cap S_2}|X_{-(S_1 \cap S_2)}) + H(X_{S_1 \cup S_2}|X_{-(S_1 \cup S_2)})$ where in the second step we have used the supermodularity property of conditional entropy. Now we are also given that $\sum_{l \in S_1 \cap S_2} R_{l,t} \geq H(X_{S_1 \cap S_2}|X_{-(S_1 \cap S_2)})$ and $\sum_{l \in S_1 \cup S_2} R_{l,t} \geq H(X_{S_1 \cup S_2}|X_{-(S_1 \cup S_2)})$. Therefore we can conclude that $\sum_{l \in S_1 \cup S_2} R_{l,t} \leq H(X_{S_1 \cup S_2}|X_{-(S_1 \cup S_2)})$ and $\sum_{l \in S_1 \cap S_2} R_{l,t} \leq H(X_{S_1 \cap S_2}|X_{-(S_1 \cap S_2)})$ and consequently that $\sum_{l \in S_1 \cup S_2} R_{l,t} = H(X_{S_1 \cup S_2}|X_{-(S_1 \cup S_2)})$ and $\sum_{l \in S_1 \cap S_2} R_{l,t} = H(X_{S_1 \cap S_2}|X_{-(S_1 \cap S_2)})$. ■