Rate-optimal schemes for Peer-to-Peer live streaming

Laurent Massoulié and Andy Twigg
Thomson, Corporate Research

Abstract

In this paper we consider the problem of sending data in real time from information sources to sets of receivers, using peer-to-peer communications. We consider several models of communication resources, and for each model we identify schemes that achieve successful diffusion of information at optimal rates.

For edge-capacitated networks, we show optimality of the so-called "random-useful" packet forwarding algorithm. As a byproduct, we obtain a novel proof of a famous theorem of Edmonds, characterising the broadcast capacity of a capacitated graph.

For node-capacitated networks, assuming a complete communication graph, we show optimality of the so-called "most-deprived" neighbour selection scheme combined with random useful packet selection. We then show that optimality is preserved when each peer can exchange data with a limited number of neighbours, when neighbourhoods are dynamically adapted according to a particular scheme.

Finally, we consider the case of multiple information sources, each creating distinct information to be disseminated to a specific set of receivers. In this context, we prove optimality of the so-called "bundled most-deprived neighbour random useful packet" selection.

I. INTRODUCTION

There are nowadays many operational peer-to-peer systems supporting *live streaming*, i.e. real-time dissemination of audio-video data injected at a source towards a set of receivers (e.g. PPLive, TVants, Sopcast, and CoolStreaming, to name a few). The detailed algorithms implemented in those systems are in general not publicly available, a notable exception being CoolStreaming (see e.g.[14]).

However, some degree of reverse engineering is possible [12], and it is commonly believed that the above-mentioned systems implement schemes of the following kind. Data generated at the source is partitioned into *windows* of contiguous data packets. Each peer then aims to download all packets from a window, before it starts downloading packets from the subsequent window. To download the packets from a given window, a peer implements schemes similar to those used in BitTorrent [2] for downloading a file.

The schemes used for determining packet exchanges result from a combination of peer selection strategy (who to send to) and packet selection strategy (what to send). For instance in BitTorrent, peers serve preferentially those from whom they received more data in the recent

The results in Sections 2-4 of this paper were announced in [7]. Early versions of their proofs appeared in Technical Report [8]. The results in Section 5 appear here for the first time.

past (this is the so-called Tit-for-Tat mechanism, discussed in detail in [6]). The decision of what to send is made by the receiver, who selects the rarest packet, rarity being evaluated locally across neighbour peers.

The aim of the present paper is to introduce mathematical models and give proofs that simple strategies, similar to those just described, are capable of achieving real-time dissemination of content to receivers at optimal rates. By optimal rate, we mean that dissemination at larger rates is infeasible with the available communication resources.

The organisation of the paper is as follows. In Section II, we introduce the model of edge-capacitated networks. In this context, the optimal diffusion, or broadcast rate, has been characterised by Edmonds [5] as the minimum over receivers of the minimum of the capacity of cuts between the data source and the corresponding receiver (also known as the min-min-cut capacity). Our main result, Theorem 1, states that under random-useful packet forwarding, the system is ergodic provided the data injection rate is strictly below the min-min-cut capacity. The theorem of Edmonds [5] is retrieved as a corollary of our result.

The proof of Theorem 1 is given in Section III.

In Section IV, we introduce the model of node-capacitated networks, and the most-deprived neighbour selection strategy. Assuming a complete communication graph (i.e., everyone can send to anyone), we prove that the system is ergodic under the most-deprived neighbour random-useful packet selection rules, when the injection rate is strictly below the maximal sustainable injection rate.

In Section V, we relax the assumption of a complete communication graph. We show that rate optimality is preserved for communication graphs with small neighbourhoods that are continuously adapted over time. We then extend the results of Section IV to the case of multiple information sources, each creating distinct information to be disseminated to a specific set of receivers. In this context, we prove optimality of the so-called bundled most-deprived neighbour random useful packet selection.

Conclusions are drawn in Section VI.

Preliminary versions of the proofs in Sections II–IV appeared previously in the technical report [8], while the corresponding results have been announced in [7]. Both the results and the proofs in Section V appear here for the first time.

II. Edge capacitated networks

A. System model

A directed, edge-capacitated graph $G = (V, E)$ is given. A distinguished source node s wants to send data to all other nodes, i.e. broadcast information over the graph G. Data transfers consist of packet replication from some node u to a node v such that $(u, v) \in E$. The *injection rate* at the source is denoted by λ, and is by definition the rate at which the source gets new packets.

The scheduling strategy we consider in this section is the following. When any node u has some packets that one of its neighbours v has not received yet, or is not currently receiving, and when there is no current transmission from node u to node v, then u picks uniformly at random one of the packets that it could usefully send to v, and sends it to v. Such a transfer proceeds at speed c_{uv}, where c_{uv} is the capacity of link $(uv) \in E$.

Once injected at the source, a packet p can be in a number of different states. It can be replicated at all nodes in the system, hence successfully broadcast. Alternatively, it can be *idle*, that is not actively transferred, and replicated at nodes u in some set $S \subset V$. In this paper the symbol $\subset$ is used to denote strict subsets; the symbol $\subseteq$ is used to denote non strict subsets. The subsets over which packets can be replicated is not arbitrary: it must contain a spanning tree rooted at s, and hence in particular it must contain s. We shall denote by $\mathcal{S}$ the collection of strict subsets of V that contain the source node s.

Alternatively, it can be replicated at some nodes $u \in S$, for some subset $S \in \mathcal{S}$, but also actively transferred along some edges $e \in F$, for some subset $F \subseteq E$.

We shall adopt the following description of the system state: for all $S \in \mathcal{S}$, X_S denotes the number of idle packets, that are replicated exactly at the nodes $u \in S$. In addition, an unordered list of subgraphs $A = \{G_1 = (W_1, F_1), \dots, G_m = (W_m, F_m)\}$ is maintained, describing the "active packets": W_i is the set of nodes at which the i-th active packet is currently replicated; F_i is the set of edges along which the i-th active packet is currently transferred.

We shall assume that at any given time, at most one packet is transferred along a given edge. Thus, the total number of active packets is at most $|E|$. We shall further assume that the following constraints are met: for an active packet with description (W, F), for each $(u, v) \in F$, then $u \in W$, $v \notin W$, and there is no other edge $e \in F$ that points towards v.

The physical meaning of this assumption is the following. No packet is sent towards a node that already has it, and no packet is sent simultaneously from several nodes to the same destination node.

One practical implementation that ensures this property consists in letting each receiver node inform its neighbour nodes of which packets it has not received yet, and omitting from this list those packets currently being transferred.

Introduce the notation:

$$X_{+u-v} = \sum_{S \in \mathcal{S}: u \in S, v \notin S} X_S.$$

This counts the number of idle packets that are present at node u and absent at node v.

Let also X^a_{+u-v} denote the number of active packets that could possibly be forwarded along edge (u, v), given the above constraints. That is to say, let

$$X^a_{+u-v} = \sum_{(W,F) \in A} \mathbf{1}_{u \in W} \mathbf{1}_{v \notin W} \mathbf{1}_{\forall u' \in V, (u',v) \notin F}.$$

The following **activity condition** will be enforced at all times: for any edge (u, v), either there is an active packet that is actively transferred along edge (u, v), or:

$$X_{+u-v} = 0 \text{ and } X^a_{+u-v} = 0.$$

In words, if there is no ongoing transfer along some edge (u, v), then necessarily no packet present in the system could be transferred along this edge.

We now describe the transitions that the system state can experience under the proposed random useful scheduling strategy.

Primary transitions: The first type of primary transitions is due to a fresh packet arrival at the source. After such a transition, the state variable $X_{\{s\}}$ is updated to $X_{\{s\}} + 1$.

The second type of primary transitions is due to completion of transfer of an active packet along some edge. Let this packet be represented by (W, F), and let $e = (u, v) \in F$ be the edge along which replication has just completed.

Then two cases may occur. If $e = (u, v)$ was the only edge in F, then the packet under consideration, characterised by (W, F), is removed from the collection of active packets, and the number of idle packets replicated at $S = W \cup \{v\}$ is increased by 1. If instead $F \setminus \{e\}$ is not empty, then (W, F) is replaced by $(W \cup \{v\}, F \setminus \{e\})$ in the list of active packets.
Secondary transitions: These happen subsequently to primary transitions, to ensure that the activity condition is met. If, after a primary transition, there is an edge (u, v) for which the activity condition is not met, this means that this edge is not actively used, while the number of packets $X_{+u-v} + X^a_{+u-v}$ which could potentially be transferred along that edge is positive. In this case, one of these $X_{+u-v} + X^a_{+u-v}$ packets will be selected uniformly at random, and start being replicated along edge (u, v).

More precisely, for each $S \in \mathcal{S}$ such that $u \in S$, $v \notin S$, with probability

$$\frac{X_S}{X_{+u-v} + X^a_{+u-v}},$$

the following state updates are made:

$$\begin{aligned} &X_S \leftarrow X_S - 1, \\ &A \leftarrow A \cup (S, (u, v)). \end{aligned}$$

For each active packet (W, F) such that $u \in W$, $v \notin W$, and for all $u' \in V$, $(u', v) \notin F$, then with probability $1/(X_{+u-v} + X^a_{+u-v})$, the active set A is updated as follows:

$$A \leftarrow A \setminus (W, F) \cup (W, F \cup (u, v)).$$

Note that all these transition probabilities sum to 1, as required. Moreover, these capture the uniform selection of useful packets that could be sent along edge (u, v) that we assume throughout.

A Markovian special case: A general version of the model would assume that the time intervals between fresh packet arrivals at the source are i.i.d. random variables, and that packet transfer times along a given edge are also i.i.d. random variables. Under these assumptions, the model we just described is a Markov process, provided we augment the state space to keep track of the residual times till (i) arrival of the next fresh packet, and (ii) completion of transmission along a given edge. Of particular interest is the case where these i.i.d. random variables are in fact deterministic.

The general i.i.d. case is beyond the scope of the present work. In this article, we focus on the special case where the i.i.d. random variables involved are Exponential random variables, where the mean inter-packet arrival at the source equals λ^{-1}, and the mean packet transfer time along edge (u, v) is c_{uv}^{-1}. In this particular case, the evolution of the state variables described above is Markovian, without the adjunction of residual time variables. In the sequel we focus on this particular setup.

B. Edge capacities: main result

We shall denote by $\lambda^*(G)$ the min-min-cut of graph G with source node s. That is,

$$\lambda^*(G) = \min_{u \in V} \min_{S \subset V : s \in S, u \notin S} \sum_{v \in S} \sum_{w \notin S} c_{vw}. \tag{1}$$

The main result in the present context is the following

Theorem 1: The Markov process $((X_S)_{S \in \mathcal{S}}, A)$ corresponding to random useful packet forwarding is ergodic under the condition

$$\lambda < \lambda^*(G). \tag{2}$$

The proof of this result will be given in Section III . It relies on the so-called “fluid limits” approach, introduced and popularised by [10] and [4]. Informally, the approach consists in first establishing that trajectories of the original Markov process, after joint rescaling of both time and space, evolve according to some simpler, “fluid” dynamics, and then to prove that trajectories of the fluid dynamics converge to zero in finite time.

We define the “spanning tree packing number”, $\pi(G)$ of graph G with distinguished source node s as the solution of the following optimization problem:

$$\begin{aligned} &\text{Maximize} && \sum_{T \in \mathcal{T}} \lambda_T && && (3)\\ &\text{over} && \lambda_T \geq 0, && T \in \mathcal{T} && (4)\\ &\text{subject to} && \sum_{T \in \mathcal{T} : (i,j) \in T} \lambda_T \leq c_{ij}, && (ij) \in E(G), && (5) \end{aligned}$$

where $\mathcal{T}$ denotes the collection of spanning trees of G, rooted at s.

A direct consequence of Theorem 1 is the following

Corollary 1: (Edmonds, 1972 [5]) For any oriented graph G with edge capacities c_{ij}, $(ij) \in E(G)$, and source node s, the spanning tree packing number $\pi(G)$ is equal to the min-min-cut number, $\lambda^*(G)$.

Proof: To any $i \in V(G)$ and any collection of non-negative numbers λ_T satisfying inequalities (5), one can associate a flow from s to i with total capacity $\sum_{T \in \mathcal{T}} \lambda_T$. Thus, necessarily $\pi(G)$ is at most the maximum flow between s and i, for any $i \in V(G)$. The celebrated max-flow min-cut theorem states that this maximum flow coincides with the minimum cut capacity between s and i. This establishes that $\pi(G)$ is at most $\lambda^*(G)$.

We now establish the converse inequality using Theorem 1. Let $\epsilon > 0$ be arbitrary, and consider the injection rate $\lambda = \lambda^*(G) - \epsilon$. Theorem 1 guarantees that the Markov process keeping track of the number of packets in any possible state is ergodic. As a consequence, there exists an equilibrium distribution for the time it takes a packet to be successfully broadcast, and a steady state distribution for the spanning tree along which a packet is effectively broadcast.

Let q be the discrete probability distribution over the collection $\mathcal{T}$ characterising the tree along which a packet is broadcast in equilibrium. Let also F_T denote the cumulative distribution function of the time to broadcast a packet in equilibrium, conditionally on the fact that it is broadcast along tree T, for all $T \in \mathcal{T}$.

Let $\delta > 0$ and $M > 0$ be some fixed positive numbers. Let τ be some time index, that we shall let increase to infinity. Let $N_T(\tau)$ denote the number of packets that have been injected at the source during $[0, \tau]$, and which have been successfully broadcast along tree T by time $(1+\delta)\tau$. Clearly, provided $M \leq \delta\tau$, $N_T(\tau)$ is larger than the number $N'_T(\tau)$ of packets injected at the source during $[0, \tau]$, that have been successfully broadcast with a broadcast time no larger than M.

By the ergodic theorem, the following holds:

$$\lim_{\tau\to\infty} \frac{1}{\tau} N'_T(\tau) = \lambda\, q(T) F_T(M), \quad \text{almost surely.}$$

In turn, this implies that:

$$\liminf_{\tau\to\infty} \frac{1}{\tau} N_T(\tau) \geq \lambda q(T) F_T(M).$$

Let $N_{ij}((1+\delta)\tau)$ denote the number of packet transmissions along edge (ij) during $[0, (1+\delta)\tau]$. Again by the ergodic theorem, and the fact that the time for a packet transmission along edge (ij) is exponentially distributed with mean $1/c_{ij}$, it holds that

$$\limsup_{\tau\to\infty} \frac{1}{(1+\delta)\tau} N_{ij}((1+\delta)\tau) \leq c_{ij}, \quad \text{almost surely.}$$

On the other hand, for any edge $(ij) \in E$, the process dynamics are such that necessarily:

$$\forall \tau \geq 0, \qquad \sum_{T:(ij)\in T} N_T(\tau) \leq N_{ij}((1+\delta)\tau), \quad \text{almost surely.}$$

Dividing this last inequality by τ, and letting τ tend to infinity, the two previous inequalities entail that

$$\sum_{T:(ij)\in T} \lambda q(T) F_T(M) \leq (1+\delta) c_{ij}, \quad (ij) \in E(G).$$

Let successively M tend to infinity and δ tend to zero to obtain:

$$\sum_{T:(ij)\in T} \lambda q(T) \leq c_{ij}, \quad (ij) \in E(G).$$

Thus, the numbers $\lambda_T := \lambda q(T)$ satisfy the constraints (5). Since they sum to λ, this ensures that $\lambda \leq \pi(G)$. Since this is true for any positive $\epsilon = \lambda^*(G) - \lambda$, the desired inequality $\lambda^*(G) \leq \pi(G)$ follows.

III. Proof of Theorem 1

The proof consists of three main parts. We first characterize the "fluid trajectories" that are valid limits of the process trajectories after joint rescaling of both time and space. We then establish that, under the stability condition $\lambda < \lambda^*(G)$, these must converge to zero in finite time. We finally deduce the ergodicity result by applying a suitable version of Foster's criterion.

A. Fluid dynamics: characterization and convergence

Let us introduce the following definition.

Definition 1: The real-valued non-negative functions $t \to y_S(t)$, $S \in \mathcal{S}$, are called fluid trajectories of the above Markov process if they satisfy the following conditions.

For all $S \in \mathcal{S}$, all $u \in S$, all $v \notin S$, there exist non-negative functions $t \to \phi_{S,(uv)}(t)$ such that

$$\begin{aligned} y_{\{s\}}(t) &= y_s(0) + \lambda t - \textstyle\sum_{v \in V \setminus \{s\}} \phi_{\{s\},(sv)}(t) \\ S \neq \{s\}: \; y_S(t) &= y_S(0) + \textstyle\sum_{u \in S} \sum_{v \in S \setminus \{u\}} \phi_{S \setminus \{v\},(uv)}(t) \\ & \quad - \textstyle\sum_{u \in S} \sum_{v \notin S} \phi_{S,(uv)}(t), \end{aligned} \tag{6}$$

and that are non-decreasing, Lipschitz continuous with Lipschitz constants c_{uv}. In addition, for all $(u, v) \in E$, it holds that:

$$\sum_{S \in \mathcal{S}: u \in S, v \notin S} \phi_{S,(uv)} \text{ is } c_{uv}\text{-Lipschitz.}$$

Moreover at almost every point t, the function $\phi_{S,(uv)}$ is differentiable, and the following holds:

$$y_{+u-v}(t) > 0 \Rightarrow \frac{d}{dt}\phi_{S,(uv)}(t) = c_{uv} \frac{y_S(t)}{y_{+u-v}(t)}, \tag{7}$$

where we have used the notation

$$y_{+u-v}(t) := \sum_{S' \in \mathcal{S}: u \in S', v \notin S'} y_{S'}(t). \tag{8}$$

$\diamondsuit$

The following notation will be used in the sequel. For any $y \in \mathbb{R}_+^{\mathcal{S}}$, $S(y)$ denotes the set of all fluid trajectories of the system with initial condition y. Thus it is a subset of $\mathcal{C}([0, +\infty), \mathbb{R}_+^{\mathcal{S}})$, that is the space of continuous, $\mathbb{R}_+^{\mathcal{S}}$-valued functions on $[0, +\infty)$. Note that at this stage, neither existence nor uniqueness of fluid trajectories has been established.

The following result shows in what sense such fluid trajectories describe the dynamics of the original Markov process after spatial and temporal rescaling. It implies as a corollary that the set $S(y)$ is nonempty, for any $y \in \mathbb{R}_+^{\mathcal{S}}$. However no claim of uniqueness of fluid trajectories is made.

Theorem 2: Consider a sequence of initial conditions $(X^N(0), A^N(0))$, $N > 0$, such that for a sequence of positive numbers $(z_N)_{N>0}$, $\lim_{N \to \infty} z_N = +\infty$, and the limit

$$\lim_{N \to \infty} \frac{1}{z_N} X^N(0) = x(0)$$

exists in $\mathbb{R}_+^{\mathcal{S}}$. Introduce the rescaled process

$$Y_S^N(t) := \frac{1}{z_N} X_S^N(z_N t), \quad S \in \mathcal{S},$$

where $X_S^N(t)$ represents the (S-coordinate of the) state of the Markov process with initial conditions $(X^N(0), A^N(0))$ at time t. Then for all $T > 0$, all $\epsilon > 0$, the following convergence takes place:

$$\lim_{N \to \infty} \mathbf{P} \left(\inf_{f \in S(x(0))} \sup_{t \in [0,T]} ||Y^N(t) - f(t)|| \geq \epsilon \right) = 0. \tag{9}$$

In words, the restriction of the rescaled process Y^N to any compact interval $[0,T]$ converges in probability to the set $S(x(0))$ of fluid trajectories with initial condition $x(0)$, where convergence of processes is for the uniform norm.

Proof: It will be more convenient to work with the state variables $\tilde{X}_S$, which count the total number of packets, active or idle, present at nodes $u \in S$. That is:

$$\tilde{X}_S = X_S + \sum_{(W,F)\in A} \mathbf{1}_{W=S}.$$

We shall thus consider the rescaled processes

$$\begin{aligned}\tilde{Y}_S^N(t) &:= \tfrac{1}{z_N}\tilde{X}_S^N(z_N t)\\ &= Y_S^N(t) + \tfrac{1}{z_N}\textstyle\sum_{W\in A^N} \mathbf{1}_{W=S}.\end{aligned}$$

Since they differ from Y_S^N by at most $|E|/z_N$, the processes agree in the limit $N\to\infty$.

Let P_{uv}, $(u,v)\in E$, be independent unit rate Poisson processes. The Poisson process P_{uv} will be used to determine the instants at which packet transfers along edge (u,v) complete. Introduce the notation:

$$\Phi_{S,(uv)}^N(t) = P_{uv}\left(c_{uv}\int_0^t \sum_{(W,F)\in A^N(s-)} \mathbf{1}_{W=S,(u,v)\in F} ds\right).$$

This process keeps track of the number of completions of packet transfers along edge (u,v), for packets that were previously present at node set S.

We thus have the following, for all $S\in\mathcal{S}$, $S\neq\{s\}$:

$$\tilde{X}_S^N(t) = \tilde{X}_S^N(0) + \sum_{u\in S, v\in S\setminus\{u\}} \Phi^N_{S\setminus\{v\},(uv)}(z_N t) - \sum_{u\in S, v\notin S} \Phi^N_{S,(uv)}(z_N t).$$

We use another unit rate Poisson process P_0 to count fresh arrivals at the source, and write:

$$\tilde{X}^N_{\{s\}}(t) = \tilde{X}^N_{\{s\}}(0) + P_0(\lambda t) - \sum_{v\neq s} \Phi^N_{\{s\},(sv)}(z_N t).$$

We now show that for any (deterministic) subsequence of the original sequence, there exists a further (deterministic) subsequence $f(N)$ for which the following property holds.

Almost surely, the sequence of rescaled processes $t\to \frac{1}{z_{f(N)}}\Phi^{f(N)}_{S,(uv)}(z_{f(N)}t)$ is tight (for the topology of uniform convergence), and any collection of functions $\phi_{S,(uv)}$ that are accumulation points of this sequence define a fluid trajectory as per the previous definition.

$$\sup_{t\in[0,T]}\left|\frac{1}{z_N}\Phi^N_{S,(uv)}(z_N t) - c_{uv}\int_0^t \sum_{(W,F)\in A^N(z_N s-)} \mathbf{1}_{W=S,(u,v)\in F} ds\right| \le \sup_{t\in[0,c_{uv}T]}\left|\frac{1}{z_N}P_{uv}(z_N t) - t\right|. \tag{10}$$

The following lemma, which is a classical result on the maximal deviation of a Poisson process from its mean is now needed:

Lemma 1: Let Ξ be a unit rate Poisson process. Then for all $T > 0$, $N > 0$, and all $\epsilon > 0$, it holds that

$$\mathbf{P}(sup_{0 \le t \le T} |\Xi(Nt) - Nt| \ge \epsilon NT) \le e^{-NTh(\epsilon)} + e^{-NTh(-\epsilon)}, \tag{11}$$

where

$$h(\lambda) := (1+\lambda)\log(1+\lambda) - \lambda \tag{12}$$

is the Cramér transform of a unit mean, centered Poisson random variable. In the above formula, it is understood that $h(-\lambda) = +\infty$ if $\lambda > 1$.
Define the subsequence $f(N)$, together with a sequence $\epsilon(N)$ as follows:

$$\begin{cases} f(N) & = \inf\{k > f(N-1) : z_k \ge N\}, \\ \epsilon(N) & = N^{-1/4}. \end{cases}$$

Define the event $\mathcal{A}_N$ as

$$\mathcal{A}_N = \cup_{S \in \mathcal{S}, u \in S, v \notin S} \left\{ \sup_{t \in [0,T]} \left| \frac{1}{z_{f(N)}} \Phi^{f(N)}_{S,(uv)}(z_{f(N)} t) - c_{uv} \int_0^t \sum_{(W,F) \in A^{f(N)}(z_{f(N)} s-)} \mathbf{1}_{W=S,(u,v) \in F} ds \right| \ge \epsilon(N) \right\}.$$

It is readily seen, using the above Lemma and the inequality (10), that the following holds:

$$\sum_{N>0} \mathbf{P}(\mathcal{A}_N) < +\infty.$$

Thus, by Borel-Cantelli's lemma, with probability 1 only finitely many events $\mathcal{A}_N$ occur. In the sequel, to lighten notation we write N instead of $f(N)$.

To establish the claimed convergence of the rescaled processes $\frac{1}{N}\Phi^N_{S,(uv)}(Nt)$ to Lipschitz-continuous, non-decreasing functions $\phi_{S,(uv)}$ along subsequences, it is therefore sufficient to establish that such convergence holds for the functions:

$$t \to c_{uv} \int_0^t \sum_{(W,F) \in A^N(Ns-)} \mathbf{1}_{W=S,(u,v) \in F} ds. \tag{13}$$

To this end, we use the following lemma, taken from Ye et al. [13]:

Lemma 2: (Lemma 6.3, Ye et al. [13]) Suppose that a sequence of functions $f_k : [0,T] \to \mathbb{R}$ has the following properties:

(i) $\{f_k(0)\}_{k \ge 0}$ is bounded;

(ii) there is a constant $M > 0$, and a sequence of positive numbers σ_k, with $\sigma_k \to 0$ as $k \to \infty$, such that

$$|f_k(t) - f_k(s)| \le M(t-s) + \sigma_k, \quad k \ge 0,\ s,t \in [0,T].$$

Then the sequence admits a subsequence that converges uniformly on $[0,T]$ to a Lipschitz continuous function $f : [0,T] \to \mathbb{R}$ with Lipschitz constant M.
Clearly the conditions of the Lemma are met for the functions (13), with as a Lipschitz constant $M = c_{uv}$. Moreover, any limiting function must be non-decreasing since the functions (13) are all non-decreasing.

Note now that for $t < t'$,

$$\sum_{S \in \mathcal{S}: u \in S, v \notin S} c_{uv} \int_t^{t'} \sum_{(W,F) \in A^N(Ns-)} \mathbf{1}_{W=S,(u,v)\in F} ds \leq c_{uv}(t'-t).$$

This readily implies that for any given $(u,v) \in E$, the limiting functions $\phi_{S,(uv)}$ summed over $S \in \mathcal{S}$ such that $u \in S$ and $v \notin S$ are c_{uv}-Lipschitz.

It now remains to establish the last property in the definition of fluid trajectories, that is: at almost every t, the function $\phi_{S,(uv)}(t)$ is differentiable, and provided $y_{+u-v}(t) > 0$, then:

$$\frac{d}{dt}\phi_{S,(uv)}(t) = c_{uv}\frac{y_S(t)}{y_{+u-v}(t)}\cdot$$

By Rademacher's theorem, a Lipschitz-continuous function is differentiable almost everywhere. Let thus t be a point where $\phi_{S,(uv)}(t)$ is differentiable. Consider first the case where $y_S(t) > 0$. Fix some $h > 0$. We want to evaluate the following quantity:

$$\frac{1}{h}c_{uv}\int_t^{t+h} \sum_{(W,F)\in A^N(Ns-)} \mathbf{1}_{W=S,(u,v)\in F} ds.$$

Note that on the interval $\tau \in [t, t+h]$, $N^{-1}X_S^N(N\tau)$ equals $y_S(t)+0(h)+\epsilon_N$, where $\epsilon_N \to 0$ as $N \to \infty$, by convergence of the rescaled trajectories, and by Lipschitz continuity of the limiting trajectories.

Thus, after each completion of a transfer along edge (u,v) during the interval $[Nt, N(t+h)]$, the probability that the next packet selected for transmission along edge (u,v) is a previously idle packet, replicated at nodes $w \in S$ is asymptotic to $y_S(t)/y_{+u-v}(t) + 0(h)$. Furthermore, once such a transfer is started, the probability that the packet under consideration is elected for transmission along another edge converges to zero as $N \to \infty$, since there are close to $Ny_S(t)$ other idle packets that could alternatively have been selected for such a transmission. Together these arguments ensure that

$$\lim_{N\to\infty}\frac{1}{h}c_{uv}\int_t^{t+h} \sum_{(W,F)\in A^N(Ns-)} \mathbf{1}_{W=S,(u,v)\in F} ds = c_{uv}\frac{y_S(t)}{y_{+u-v}(t)} + O(h)\cdot$$

However, the left-hand side of this expression also reads

$$\frac{1}{h}\left(\phi_{S,(uv)}(t+h) - \phi_{S,(uv)}(t)\right),$$

and thus the derivative of $\phi_{S,(uv)}$ at t must equal $c_{uv}\frac{y_S(t)}{y_{+u-v}(t)}$ as announced.

Finally, consider the case where $y_S(t) = 0$, and choose a particular t at which all S' with $u \in S'$, $v \notin S'$ are such that $\phi_{S',(uv)}(t)$ are differentiable. We know that almost everywhere, the sum of these derivatives can not exceed c_{uv}, because it is a Lipschitz constant for the sum of these functions. However, the sum of the derivatives for those S' such that $y_{S'}(t) > 0$ equals c_{uv}, therefore the derivatives for those S such that $y_S(t) = 0$ must equal zero. ■

B. Fluid dynamics: stability

In the present section, we establish that any fluid trajectories as per Definition 1 satisfy a suitable stability property:

Theorem 3: Assume that Condition (2) holds. Let $(y_S)_{S\in\mathcal{S}}$ denote fluid trajectories as per Definition 1. For all $S\subset V$, define:

$$y_{\subseteq S}=\sum_{S'\in\mathcal{S},S'\subseteq S} y_{S'}.$$

Then there exist positive parameters $\beta_1,\ldots,\beta_{|V|-1}$, and $\epsilon>0$ such that the function

$$L(\{y_S\}_{S\in\mathcal{S}}):=\sup_{S\subset V}\beta_{|S|}y_{\subseteq S}$$

verifies:

$$L(y(t))\leq \max\left(0,L(y(0))-\epsilon t\right). \tag{14}$$

Denote by K the total number of nodes, that is $K=|V|$. The proof will rely on the following lemma:

Lemma 3: Let $\alpha>0$ be fixed. For given $\delta, A>0$, define:

$$\begin{array}{ll}\epsilon_{K-1}=\delta; & \\ \epsilon_{K-1-i}=\delta A(1+A)^{i-1}, & i=1\ldots,K-2,\\ \beta_{K-1}=1; & \\ \beta_{K-i}=\prod_{j=K-i+1}^{K-1}\left(\frac{1}{1-\epsilon_j}\right), & i=2,\ldots,K-1.\end{array} \tag{15}$$

Then A and δ can be chosen so that the following properties hold for any $(y_S)_{S\in\mathcal{S}}\in\mathbb{R}_+^{\mathcal{S}}$, $(y_S)_{S\in\mathcal{S}}\neq 0$. For all $S\subset V$, all $u\in S$, $v\notin S$:

$$y_{+u-v}<\epsilon_{|S|}y_{\subseteq S}\Rightarrow \beta_{|S|-1}y_{\subseteq S\setminus\{u\}}>\beta_{|S|}y_{\subseteq S}. \tag{16}$$

Moreover for all $S\subset V$ such that, for all $u\in S$, all $v\notin S$, $y_{+u-v}\geq\epsilon_{|S|}y_{\subseteq S}$, assuming there exist $u\in S$ and $v\notin S$ such that for some $S'\not\subseteq S$: $u\in S'$, $v\notin S'$ and $y_{S'}>\alpha y_{+u-v}$, then it holds that:

$$\beta_{|S\cup S'|}y_{\subseteq S\cup S'}>\beta_{|S|}y_{\subseteq S}. \tag{17}$$

Proof: (of Lemma 3) Let us first establish sufficient conditions on the parameters ϵ_i, β_i for the conclusions of the Lemma to hold. Consider the first requirement (16), and let thus S be such that for some $u\in S$ and $v\notin S$, one has

$$y_{+u-v}<\epsilon_{|S|}y_{\subseteq S}.$$

Write now:

$$\begin{aligned}y_{\subseteq S}&=y_{\subseteq S\setminus\{u\}}+\textstyle\sum_{S'\in\mathcal{S}:u\in S',S'\subseteq S}y_{S'}\\ &\leq y_{\subseteq S\setminus\{u\}}+y_{+u-v}\\ &< y_{\subseteq S\setminus\{u\}}+\epsilon_{|S|}y_{\subseteq S}.\end{aligned}$$

It thus follows that

$$y_{\subseteq S\setminus\{u\}}>(1-\epsilon_{|S|})y_{\subseteq S}.$$

Thus the desired conclusion (16) will follow provided:

$$\beta_{i-1}(1-\epsilon_i) \geq \beta_i, \quad i=2,\ldots,K-1. \tag{18}$$

Clearly, this condition will be satisfied with the particular choice of coefficients β_i as in (15), provided the ϵ_i lie in the interval $(0,1)$, which will be ensured by taking $\delta > 0$ sufficiently small.

Let us now turn to Condition (17). Let thus $S \subset V$ be such that for all $u \in S$ and $v \notin S$, $\epsilon_{|S|} y_{\subseteq S} \leq y_{+u-v}$. Assume moreover the existence of $u \in S$, $v \notin S$, and $S' \not\subseteq S$ such that $u \in S'$, $v \notin S'$, and satisfying in addition:

$$y_{S'} > \alpha y_{+u-v}.$$

Then necessarily, one has:

$$y_{S'} > \alpha \epsilon_{|S|} y_{\subseteq S}.$$

The left-hand side of Condition (17) then verifies:

$$\begin{aligned} \beta_{|S\cup S'|} y_{\subseteq S\cup S'} &\geq \beta_{|S\cup S'|}\left(y_{S'} + y_{\subseteq S}\right) \\ &> \beta_{|S\cup S'|}(1+\alpha\epsilon_{|S|}) y_{\subseteq S}. \end{aligned}$$

Therefore, (17) will hold provided

$$\beta_{|S\cup S'|}(1+\alpha\epsilon_{|S|}) \geq \beta_{|S|}.$$

For sufficiently small $\delta > 0$, the coefficients ϵ_i as in (15) will be strictly less than 1, and hence the coefficients β_i as in (15) will be decreasing with i. Thus, the above condition will be satisfied provided:

$$\beta_{K-1}(1+\alpha\epsilon_i) \geq \beta_i, \quad i=1,\ldots,K-2. \tag{19}$$

For $i = K-2$, this condition reads $1+\alpha\epsilon_{K-2} \geq 1/(1-\epsilon_{K-1})$. Recalling from (15) that $\epsilon_{K-1} = \delta$, the right-hand side reads $1+\delta+o(\delta)$, while the left-hand side reads $1+\alpha\delta A$. Thus, this particular condition is met provided $A\alpha > 1$, and $\delta > 0$ is small enough.

Let us now consider $i \in \{1,\ldots,K-3\}$. Note that the right-hand side of (19) is equivalent to, for small $\delta > 0$:

$$\begin{aligned} \beta_i &= \prod_{j=i+1}^{K-1}\left(\frac{1}{1-\epsilon_j}\right) \\ &= 1 + \sum_{j=i+1}^{K-1} \epsilon_j + o(\delta) \\ &= 1 + \delta + \sum_{j=i+1}^{K-2} \delta A (1+A)^{K-2-j} + o(\delta) \\ &= 1 + \delta + \delta A \sum_{j=0}^{K-3-i} (1+A)^j + o(\delta) \\ &= 1 + \delta + \delta A \frac{(1+A)^{K-3-i+1}-1}{A} + o(\delta) \\ &= 1 + \delta(1+A)^{K-2-i} + o(\delta). \end{aligned}$$

On the other hand, the left-hand side of (19) equals $1+\alpha\delta A(1+A)^{K-2-j}$. Thus, provided $\alpha A > 1$, and $\delta > 0$ is small enough, the announced properties hold. ■

Proof: (of Theorem 3) Consider the particular parameters β_i, ϵ_i as in Lemma 3. Clearly, for a vector $y \in \mathbb{R}_+^{\mathcal{S}}$ that is non-zero, any set $S^* \subset V$ achieving the maximum in $\max_{S\subset V} \beta_{|S|} y_{\subseteq S}$ is such that $y_{\subseteq S^*} > 0$. Moreover, for all $u \in S^*$, $v \notin S^*$, one must have:

$$y_{+u-v} \geq \epsilon_{|S^*|} y_{\subseteq S^*} > 0,$$

for otherwise optimality of the set S^* would be contradicted by Condition (16). In addition, for all $u \in S^*$, $v \notin S^*$, and all $S' \not\subseteq S^*$ such that $u \in S'$, $v \notin S'$, necessarily $y_{S'} \leq \alpha y_{+u-v}$, for otherwise optimality of S^* would be contradicted by (17).

One thus has the following evaluation:

$$\begin{array}{ll} \frac{d}{dt} y_{\subseteq S^*} & = \sum_{S\subseteq S^*} \frac{d}{dt} y_S \\ & = \lambda - \sum_{u\in S^*, v\notin S^*} \sum_{S\subseteq S^*, u\in S} \frac{d}{dt} \phi_{S,(uv)} \\ & = \lambda - \sum_{u\in S^*, v\notin S^*} c_{uv} \left[1 - \sum_{S'\not\subseteq S^*, u\in S', v\notin S'} \frac{y_{S'}}{y_{+u-v}}\right] \\ & \leq \lambda - \sum_{u\in S^*, v\notin S^*} c_{uv} + \sum_{u\in S^*, v\notin S^*} c_{uv} \sum_{S'\not\subseteq S^*, u\in S', v\notin S'} \alpha \\ & \leq \lambda - \sum_{u\in S^*, v\notin S^*} c_{uv} + \max_{e\in E} c_e |E| 2^K \alpha. \end{array}$$

In the above, we have used the expression (7) for the derivative of the functions $\phi_{S,e}$, and the bound of α on the ratio $y_{S'}/y_{+u-v}$ previously established.

Furthermore, the conditions (18) and (19) used in the proof of Lemma 3 can be shown to imply the following. For a set S such that $\beta_{|S|} y_{\subseteq S} \geq (1-r)\beta_{|S^*|} y_{\subseteq S^*}$, where $r > 0$ is some small positive constant, necessarily for all $u \in S$, $v \notin S$,

$$y_{+u-v} \geq \left(1 - \frac{1-\epsilon_{|S|}}{1-r}\right) y_{\subseteq S}.$$

In addition, for $u \in S$, $v \notin S$ and $S' \not\subseteq S$ such that $u \in S'$, $v \notin S'$, then one has:

$$y_{S'} \leq \left(\frac{1+\alpha\epsilon_{|S|}}{1-r} - 1\right) \frac{1}{1 - \frac{1-\epsilon_{|S|}}{1-r}} y_{+u-v} = (\alpha + O(r))\, y_{+u-v}.$$

Thus, for such S, one has the similar evaluation

$$\frac{d}{dt} y_{\subseteq S} \leq \lambda - \sum_{u\in S, v\notin S} c_{uv} + \max_{e\in E} c_e |E| 2^K \alpha\, (1 + O(r)). \tag{20}$$

Note that the choice of $\alpha > 0$ in Lemma 3 was arbitrary. For definiteness, set

$$\alpha = \frac{1}{2} \frac{\min_{S\subset V} \sum_{u\in S, v\notin S} c_{uv} - \lambda}{|E| 2^K \max_{e\in E} c_e}$$

This is positive, under the stability condition (2). Then from the above evaluation (20), it follows that necessarily, almost everywhere the Lipschitz continuous function $L(y(t))$ must satisfy:

$$\frac{d}{dt} L(y(t)) \leq -\epsilon \mathbf{1}_{y(t)\neq 0},$$

where

$$\epsilon := \frac{1}{2} \left(\min_{S\subset V} \sum_{u\in S, v\notin S} c_{uv} - \lambda \right).$$

The result of Theorem 3 follows. ■

C. Proof of Theorem 1

The proof of Theorem 1 will require to combine Theorems 2, 3 and the following ergodicity criterion, which is a direct consequence of Theorem 8.13, p.224 in Robert [9]:

Theorem 4: Let $Z(t)$ be a Markov jump process on a countable state space $\mathcal{Z}$. Assume there exists a function $L:\mathcal{Z}\to\mathbb{R}_+$ and constants M, ϵ, $\tau>0$ such that for all $z\in\mathcal{Z}$:

$$L(z)>M\Rightarrow\frac{1}{L(z)}\mathbf{E}_zL\left(Z(L(z)\tau)\right)\leq 1-\epsilon. \tag{21}$$

If in addition the set $\{z:L(z)\leq M\}$ is finite, and $\mathbf{E}_zL(Z(1))<+\infty$ for all $z\in\mathcal{Z}$, then the process $Z(t)$ is ergodic.

Let us show how this result applies in the present context. Here we have $Z(t)=(X(t),A(t))$, and our candidate Lyapunov function takes as argument the X-component only, and reads

$$L(Z)=\sup_{S\subset V}\beta_{|S|}X_{\subseteq S}.$$

Let us set $\tau=1$, where ϵ is as in Theorem 3, and establish that (21) holds by contradiction. Assuming it fails, there must exist a sequence of initial conditions $Z^N(0)$ such that $L(Z^N(0))\to\infty$, and such that

$$\lim_{N\to\infty}\frac{1}{L(Z^N(0))}\mathbf{E}L\left(Z^N(L(Z^N(0))\tau)\right)>1-\epsilon. \tag{22}$$

However, by Theorem 1, any accumulation point of the sequence

$$\frac{1}{L(Z^N(0))}X^N(L(Z^N(0))\tau)$$

must be equal to $y(\tau)$ for some fluid trajectory y issued from an initial condition $y(0)$ such that $L(y(0))=1$. Furthermore, this family of random vectors is uniformly integrable: indeed, writing

$$\frac{1}{L(Z^N(0))}X_S^N(L(Z^N(0))\tau)\leq\frac{X_S^N(0)}{\beta_{|S|}X_S^N(0)}+\frac{1}{L(Z^N(0))}\sum_{e\in E}P_e(L(Z^N(0))c_e\tau),$$

where the P_e are the Poisson processes previously introduced, uniform integrability can be readily checked. Since the function L grows not faster than linearly, the family of random variables

$$\frac{1}{L(Z^N(0))}L\left(X^N(L(Z^N(0))\tau)\right)$$

is also uniformly integrable. Since the function L is continuous, accumulation points of this sequence must be of the form $L(y(\tau))$, for some fluid trajectory y issued from an initial condition $y(0)$ such that $y(0)=1$. By Theorem 3, all such accumulation points are less than, or equal to $1-\epsilon$. This together with uniform integrability ensures that

$$\limsup_{N\to\infty}\frac{1}{L(Z^N(0))}\mathbf{E}L\left(Z^N(L(Z^N(0))\tau)\right)\leq 1-\epsilon,$$

which contradicts (22). The proof is concluded by verifying the other assumptions of Theorem 4, i.e. that $\{z:L(z)\leq M\}$ is finite for sufficiently large M. This holds trivially, because for any X-component the number of potential A-components is bounded (say by $|E|$ times the number of subgraphs of G).

Finally, one must check that $\mathbf{E}_z L(Z(1)) < +\infty$ for all z; this is easily verified, once more by bounding $X_S(1)$ by its initial value plus increments of Poisson processes.

IV. Node-capacitated networks

A. Model and Algorithm

Neighbour selection: Here, the system is also described by a a graph $G = (V, E)$. However, the capacities are now associated with nodes rather than with edges. We shall denote by c_u the capacity of node u, and assume that each node devotes its capacity to one of its "most deprived neighbours". By this, the following is meant. For each of its neighbours v, node u evaluates the number Z_{+u-v} of packets that it could usefully forward to node v. Using the same notation as before, this reads:

$$Z_{+u-v} = X_{+u-v} + X^a_{+u-v}.$$

It then elects one neighbour v for which the corresponding quantity Z_{+u-v} is maximal. Ties can be broken either at random, or in a systematic manner. Once the target neighbour v is chosen, then one of the Z_{+u-v} packets held by u and useful to v is chosen, and forwarded from u to v, at rate c_u.

Packet selection: We now describe how packets are elected for transmission once a node's capacity becomes available. For non-source nodes u, who have chosen to transmit to some most deprived neighbour v, then the packet to be transmitted is selected at random among all the possible Z_{+u-v} possible choices.

For the source node s, having chosen to transmit to some most deprived neighbour v, the following strategy is used: if the source has a packet that it has not sent to anyone before (a *fresh* packet), that is if $X_{\{s\}} > 0$, then one such fresh packet is forwarded to node v; if no such fresh packet is available, then the packet to be forwarded is selected uniformly at random from the Z_{+s-v} possible choices.

As in the edge capacitated case, the state space consists in the collection of variables X_S, for all $S \in \mathcal{S}$, and the collection of active packet states $A = ((W_1, F_1), \ldots, (W_m, F_m))$. The constraints on these active packet states are different though: we now assume that each node forwards a packet to only one of its neighbours at a given time. Thus for each node u, there is at most one edge (u, w) appearing in the sets F_i, $i = 1, \ldots, m$. Otherwise the same constraints apply: for a given active packet (W, F), and each edge $(u, v) \in F$, necessarily, $u \in W$ and $v \notin W$; also, there is no other edge (u', v) pointing towards v in F.

We shall assume that packet transmissions are not preempted, even if a neighbour of some node u becomes more deprived than the neighbour v to which node u is currently transmitting.

As in the edge-capacitated case, we assume Exponentially distributed inter-event timers. Specifically, the time for transmission of a packet from some node u is exponentially distributed with mean $1/c_u$, and fresh packets arrive at the source node s at the instants of a Poisson process with rate λ.

B. Fluid limits

We first define the candidate fluid trajectories for the system under consideration:

Definition 2: The real-valued, non-negative functions $(y_S)_{S\in\mathcal{S}}$ are called fluid trajectories of the node-capacitated system if the following properties hold.

For all $S \in \mathcal{S}$, $u \in S$, $v \notin S$ such that $(u,v) \in E$, there exist non-decreasing, Lipschitz-continuous functions $\phi_{S,(uv)}$ with Lipschitz constant c_u, such that Equations (6) hold. Furthermore, using notation

$$y_{+u-v} := \sum_{S\in\mathcal{S}:u\in S, v\notin S} y_S,$$

for all $S \in \mathcal{S}$, $u \in S$, the functions $\{\phi_{S,(uv)}\}_{v\notin S,(uv)\in E}$ are differentiable at almost every t, and if $\sum_{v:(u,v)\in E} y_{+u-v}(t) > 0$, their derivatives satisfy:

$$\frac{d}{dt}\phi_{S,(uv)}(t) = 0 \text{ if } y_{+u-v}(t) < \max_{v':(u,v')\in E}\left(y_{+u-v'}(t)\right), \tag{23}$$

$$\sum_{v:(uv)\in E}\ \sum_{S:u\in S, v\notin S} \frac{d}{dt}\phi_{S,(uv)}(t) = c_u. \tag{24}$$

If $u \neq s$, that is for a non-source node, one also has, for all v such that $(uv) \in E$ and assuming the condition

$$\sum_{S:u\in S, v\notin S} \frac{d}{dt}\phi_{S,(uv)}(t) > 0$$

holds, the following equation:

$$\forall S/u \in S, v \notin S,\ \frac{d}{dt}\phi_{S,(uv)}(t) = \frac{y_S(t)}{\sum_{S':u\in S', v\notin S'} y_{S'}(t)} \sum_{S':u\in S', v\notin S'} \frac{d}{dt}\phi_{S',(uv)}(t). \tag{25}$$

For the source node s, one has the following:

$$y_{\{s\}} > 0 \Rightarrow \sum_{v\neq s} \frac{d}{dt}\phi_{\{s\},(sv)}(t) = c_s. \tag{26}$$

In the case where $y_{\{s\}} = 0$, one then has for all v such that $(sv) \in E$, assuming the condition

$$\sum_{S\in\mathcal{S}:S\neq\{s\}, v\notin S} \frac{d}{dt}\phi_{S,(sv)}(t) > 0$$

holds, the following:

$$\forall S \in \mathcal{S}/S \neq \{s\}, v \notin S:\ \frac{d}{dt}\phi_{S,(sv)}(t) = \frac{y_S(t)}{\sum_{S'\in\mathcal{S}:S'\neq\{s\}, v\notin S'} y_{S'}(t)} \sum_{S'\in\mathcal{S}:S'\neq\{s\}, v\notin S'} \frac{d}{dt}\phi_{S',(sv)}(t). \tag{27}$$

$\diamondsuit$

We now establish the following

Theorem 5: The statement of Theorem 2 holds true with $(X^N(t), A^N(t))$ denoting the state of the process corresponding to the node-capacitated system, and with $S(x)$ denoting the set of fluid trajectories defined in Definition 2. That is, rescaled trajectories converge in probability to the set of fluid trajectories.

Proof: Introduce the functions

$$\Phi^N_{S,(uv)}(t) := P_u\left(c_u \int_0^t \sum_{(W,F)\in A^N(s-)} \mathbf{1}_{W=S,(u,v)\in F} ds\right),$$

where P_u are independent, unit rate Poisson processes. The existence of functions $\phi_{S,(uv)}$ that are non-increasing and Lipschitz continuous with Lipschitz constant c_u, and such that for functions y_S given by (6), the claimed convergence in probability holds, is established exactly as in the proof of Theorem 2, and hence the detailed argument is omitted.

It only remains to establish properties (23–27) of the derivatives $\frac{d}{dt}\phi_{S,(uv)}(t)$. Fix thus $h > 0$, and consider the quantity

$$\frac{1}{h}\left(\frac{1}{N}\Phi^N_{S,(uv)}(N(t+h)) - \frac{1}{N}\Phi^N_{S,(uv)}(Nt)\right). \tag{28}$$

Assume that the node u is such that the limiting processes (y) satisfy

$$\sum_{v'\neq u} y_{+u-v'}(t) > 0. \tag{29}$$

Then, provided $y_{+u-v}(t) < \max_{v'\neq v} y_{+u-v}(t)$, by Lipschitz continuity of the limiting trajectories, the same inequality holds throughout the interval $[t, t+h]$. Thus, by convergence of the rescaled trajectories to the fluid limits, for large enough N, neighbour v is never selected for transmission by node u over the whole interval $[Nt, N(t+h)]$. It then follows that the term (28) converges to 0 as $N \to \infty$. This establishes (23).

Note next that, when (29) holds, for large enough N one has the following equality:

$$\sum_{v\neq u, S\in\mathcal{S}:u\in S, v\notin S} \frac{1}{h}\left(\frac{1}{N}\Phi^N_{S,(uv)}(N(t+h)) - \frac{1}{N}\Phi^N_{S,(uv)}(Nt)\right) = \frac{1}{N}\left(P_u(N(t+h)) - P_u(Nt)\right).$$

This is because node u's capacity is always used when there are packets that node u can usefully transmit. This identity guarantees that

$$\lim_{N\to\infty} \sum_{v\neq u, S\in\mathcal{S}:u\in S, v\notin S} \frac{1}{h}\left(\frac{1}{N}\Phi^N_{S,(uv)}(N(t+h)) - \frac{1}{N}\Phi^N_{S,(uv)}(Nt)\right) = c_u,$$

from which (24) follows.

Assume now that for non-source node u, node v is such that

$$\sum_{S:u\in S, v\notin S} \frac{d}{dt}\phi_{S,(uv)}(t) > 0.$$

Then for all S such that $u \in S$, $v \notin S$, of all the instants during the interval $[Nt, N(t+h)]$ at which node u chooses to send a packet to node v, a fraction $y_S(t)/y_{+u-v}(t) + 0(h) + 0(1/N)$ of these choices is towards an idle packet previously replicated at all nodes in S.

Furthermore, once transfer of such previously idle packets has started, such a packet is elected for transmission by some other node with probability $0(1/N)$. This thus shows that

$$\lim_{N\to\infty} \frac{1}{N}\left[\Phi^N_{S,(uv)}(N(t+h)) - \Phi^N_{S,(uv)}(Nt)\right] = \left(\frac{y_S(t)}{y_{+u-v}(t)} + 0(h)\right) \times \cdots \times \sum_{S':u\in S',v\notin S'} \left[\phi_{S',(uv)}(t+h) - \phi_{S',(uv)}(t)\right].$$

Dividing by h and letting h tend to zero establishes (25).

Equation (26) follows by similar arguments, relying on the fact that the source node s forwards fresh packets, whenever there are some available. Equation (27) is also established by similar arguments, now relying on the fact that the source, when sending non-fresh packets, selects such packets uniformly at random. ■

C. Stability for the complete graph

The main result we shall establish is in the case of the complete graph, that is all edges (u,v), $u \neq v$, are present in E. We then have the following

Theorem 6: Assume that the graph $G = (V,E)$ is complete, and that the injection rate λ verifies:

$$\lambda < \min\left(c_s, \frac{\sum_{u\in V} c_u}{K-1}\right), \tag{30}$$

where $K = |V|$. Then the Markov process keeping track of the system state under "random useful to most deprived neighbour" scheduling strategy is ergodic.

The proof of Theorem 6 parallels exactly that of Theorem 1, relying on a combination of Theorem 4 with Theorem 5 (taking the role played by Theorem 2 in the proof of Theorem 1) and of Theorem 7 below (taking the role played by Theorem 3 in the proof of Theorem 1). We shall not reproduce the whole argument, but shall instead only detail the proof of the following result on stability of fluid trajectories:

Theorem 7: For any $y = (y_S)_{S\in\mathcal{S}} \in \mathbb{R}_+^{\mathcal{S}}$, define the *workload* function $w(y)$ as:

$$w(y) = \sum_{S\in\mathcal{S}} y_S\left(K - |S|\right), \tag{31}$$

where $K = |V|$. Under the assumption (30), when the graph G is complete, any fluid trajectory y as per Definition 2 is such that, for some $\epsilon > 0$,

$$w(y(t)) \leq \max(0, w(y(0)) - \epsilon t). \tag{32}$$

Proof: To establish (32), it suffices to show that, for all fluid trajectory y, at a point t where $y(t)$ is differentiable and $y(t) \neq 0$, one has

$$\frac{d}{dt} w(y(t)) \leq -\epsilon.$$

This is true because the function $t \to w(y(t))$ is Lipschitz-continuous, which follows from Lipschitz continuity of the individual functions $t \to y_S(t)$.

We distinguish two cases. First, consider the case where at t, for all $u \in V$, one has

$$\sum_{v \neq u} y_{+u-v}(t) > 0. \tag{33}$$

Write then, using (6):

$$\begin{aligned} \frac{d}{dt}w(y(t)) &= \sum_{S\in\mathcal{S}}(K-|S|)\frac{d}{dt}y_S(t) \\ &= \lambda(K-1) - \sum_{S\in\mathcal{S}}\sum_{u\in S, v\notin S}\frac{d}{dt}\phi_{S,(uv)}(t) \\ &= \lambda(K-1) - \sum_{u\in V}\sum_{v:(uv)\in E}\sum_{S:u\in S, v\notin S}\frac{d}{dt}\phi_{S,(uv)}(t) \\ &= \lambda(K-1) - \sum_{u\in V} c_u, \end{aligned}$$

where the last equality follows from (24), which is applicable in view of Assumption (33). Thus in the present case, under Assumption (30), the time derivative $(d/dt)w(y(t))$ decreases at a constant speed as desired.

Consider now the case where for a non-empty set S^*, all $u \in S^*$ are such that

$$\sum_{v \neq u} y_{+u-v}(t) = 0.$$

Equivalently, for all $S \in \mathcal{S}$ such that $u \in S$, one has $y_S(t) = 0$. It readily follows that for any node $u \in V$, the set of most deprived neighbours consists precisely of those nodes $v \in S^*$.

Distinguish now according to whether $y_{\{s\}}(t) = 0$ or not. In the first case where $y_{\{s\}}(t) = 0$, necessarily there must exist $T \in \mathcal{S}$, $T \neq \{s\}$ for which $y_T(t) > 0$, by the assumption that $y(t) \neq 0$. Note now that, by non-negativity of the function $t \to y_{\{s\}}(t)$, one must necessarily have:

$$\frac{d}{dt}y_{\{s\}}(t) = 0, \tag{34}$$

and by the same argument, for all S such that $S \cap S^* \neq \emptyset$, one also has

$$\frac{d}{dt}y_S(t) = 0. \tag{35}$$

On the other hand, it follows from Equation (23) that the left-hand side of (34) also reads

$$\lambda - \sum_{v\in S^*}\frac{d}{dt}\phi_{\{s\},(sv)}(t).$$

It thus follows from (24) that

$$\sum_{S\in\mathcal{S}, S\neq\{s\}}\sum_{v\in S^*}\frac{d}{dt}\phi_{S,(sv)}(t) = c_s - \lambda > 0.$$

Using (34–35), write then

$$\begin{aligned} \frac{d}{dt}w(y(t)) &= \sum_{S\in\mathcal{S}:S\neq\{s\}, S\cap S^*=\emptyset}(K-|S|)\frac{d}{dt}y_S(t) \\ &= -\sum_{S\in\mathcal{S}:S\neq\{s\}, S\cap S^*=\emptyset}(K-|S|)\sum_{u\in S, v\in S^*}\frac{d}{dt}\phi_{S,(uv)}(t) \\ &\leq -\sum_{S\in\mathcal{S}:S\cap S^*=\emptyset}(K-|S|)\sum_{v\in S^*}\frac{d}{dt}\phi_{S,(sv)}(t) \\ &= -(c_s - \lambda). \end{aligned}$$

In the above, we have used the fact that the most deprived nodes are those in S^*, and hence by (23), for all S such that $S \cap S^* = \emptyset$, all $u \in S$, $v \in S \setminus \{u\}$, necessarily

$$\frac{d}{dt}\phi_{S\setminus\{v\},(uv)}(t) = 0,$$

for the capacity of node u is fully targeted towards nodes in S^*.

The last case to consider is when $y_{\{s\}}(t) > 0$. Then in view of (26),

$$\frac{d}{dt}y_{\{s\}}(t) = \lambda - c_s.$$

This entails that

$$\begin{aligned} \tfrac{d}{dt}w(y(t)) &= -(K-1)(c_s-\lambda) + \textstyle\sum_{S\in\mathcal{S}:S\neq\{s\},S\cap S^*=\emptyset}(K-|S|)\tfrac{d}{dt}y_S(t) \\ &= -(K-1)(c_s-\lambda) - \textstyle\sum_{S\in\mathcal{S}:S\neq\{s\},S\cap S^*=\emptyset}(K-|S|)\sum_{u\in S,v\in S^*}\tfrac{d}{dt}\phi_{S,(uv)}(t) \\ &\leq -(c_s-\lambda)(K-1). \end{aligned}$$

Thus, it follows that (32) holds, with $\epsilon = \min(c_s - \lambda, \sum_{u\in V} c_u - (K-1)\lambda)$. ■

V. Limited neighbourhoods and multiple commodities

We now extend the results of the previous section to limited neighbourhoods and multiple commodities.

A. Limited neighbourhoods

Given a set V of nodes u with associated capacities c_u, we assume that each node has at any given time a finite set of neighbours $\mathcal{N}(u)$, that it can send to. We adapt the most deprived neighbour selection rule to this context, by requiring that each node u sends to the most deprived node v from its limited neighbourhood $\mathcal{N}(u)$.

In addition, we assume that for any $v \in V \setminus \{u\}$, node u contacts node v at the instants of some Poisson process S_{uv} with intensity σ_{uv}. It then updates its neighbourhood as follows. It first adds v to it, and then removes a least deprived peer from the resulting set (breaking ties at random). Thus, the neighbourhood is eventually modified, but remains of constant size, say d_u.

In this context, we have the following:

Theorem 8: Consider the most deprived random useful selection mechanism with adaptive neighbourhoods as previously defined. Assume that the sampling rates s_{uv} and the neighbourhood sizes d_u are positive. Then the resulting Markov process is ergodic under the condition (30) on the injection rate λ.

Proof: The result follows from the fact that the Markov process keeping track of the states of packets, as well as the composition of the dynamic neighbourhoods, admits exactly the same fluid limits as the process with full neighbourhoods considered in the previous section. Ergodicity then follows exactly as in the proof of Theorem 6.

We shall only show that candidate fluid trajectories must satisfy (23); we omit the detailed arguments for the other equations (24-27), since these consist in very similar adaptations of their proofs for the complete graph scenario.

For fixed $h > 0$, consider the quantity

$$\frac{1}{h}\left(\frac{1}{N}\Phi^N_{S,(uv)}(N(t+h)) - \frac{1}{N}\Phi^N_{S,(uv)}(Nt)\right). \tag{36}$$

Assume that node u is such that the limiting processes (y) satisfy

$$\sum_{v' \neq u} y_{+u-v'}(y) > 0.$$

Then, provided $y_{+u-v}(t) < \max_{v' \neq u} y_{+u-v'}(t)$, by Lipschitz continuity of the limiting trajectories, the same inequality holds throughout the interval $[t, t+h]$. Let w be some arbitrary node that achieves the maximum in $y_{+u-v'}(t)$ over v'. Let now $\epsilon \in (0, h)$ be some fixed, arbitrary number. Then necessarily, there exists a subsequence (still denoted by N) under which the sampling process S_{uw} is such that

$$S_{uw}(N(t+\epsilon)) - S_{uw}(Nt) > 0$$

for large enough N. For large enough N, no transfer from u to v can take place during $[N(t+\epsilon), N(t+h)]$. Indeed, as previously explained, v is strictly less deprived than w over this interval; however, since w has been considered for inclusion in the neighbourhood of u during $[Nt, N(t+\epsilon)]$, either it has been added, or an even more deprived node was present in the neighbourhood at time $N(t+\epsilon)$. Thus, throughout the remaining interval $[N(t+\epsilon), N(t+h)]$, node u always has a neighbour that is more deprived than v, and never sends to v throughout this interval. Thus the quantity (36) is asymptotically no larger than $c_u\epsilon$. Since ϵ was arbitrary, then it must converge to zero. This establishes (23). ■

Remark 1: The mechanism we have considered for sampling new candidate neighbours consisted in node u contacting any node v at the instants of a Poisson process S_{uv}. The ergodicity result does not depend on the corresponding rate σ_{uv}. It is not hard to see that the ergodicity result does not depend either on the statistical details of the sampling mechanism.

For instance, it would still hold under the following neighbour sampling mechanism. Assume a connected, undirected graph $G_S = (V, E_S)$ is given, and further assume that each node $u \in V$ can communicate directly, at all times, with the nodes v such that $(u, v) \in E_S$. Graph G_S is used only for sampling purposes, and this in the following manner. At the instants of a Poisson process S_u, node u picks uniformly at random one node –say v– in the set $\mathcal{N}(u)$, and then picks uniformly at random a neighbour of v according to the neighbourhood structure of the sampling graph G_S, say node w. This is the node that is considered for inclusion in the neighbourhood $\mathcal{N}(u)$.

B. Multiple commodities

We now assume that there are multiple commodities, each with a dedicated source node. We let $\mathcal{I} \subset V$ denote the collection of such source nodes. For each source node $i \in \mathcal{I}$, we let $V_i \subset V$ denote the set of receivers of the corresponding commodity, including the source node i.

Ordinary nodes can be receivers of several commodities. However, we do not allow source nodes to be receivers of any commodities, except the ones they are sources of. Formally, this reads

$$\forall i, j \in \mathcal{I},\ i \neq j \Rightarrow i \notin V_j.$$

We shall also make use of the notation $I(v)$ to represent the set of source nodes i that v is a receiver of, i.e. $v \in V_i$. By convention, for source nodes $i \in \mathcal{I}$, we let $I(i) = \{i\}$.

In this setup, we denote by λ_i the packet creation rate at source node i and c_u the capacity of any node $u \in V$. We still assume exponentially distributed random variables for the inter-times between both packet creations at any source node, and packet transmissions from any node.

In the sequel, we shall assume that nodes are only willing to relay content that they are themselves interested in. This constraint can be thought of as modeling a form of selfishness of users, who will not participate in the delivery of data they do not need. This can be related to the so-called Tit-for-tat mechanism implemented in BitTorrent, with two important distinctions. Our constraint is imposed at the level of commodities, while BitTorrent's Tit-for-tat is used to determine which node to send to, in a group of users interested in the same content. Also, we do not put constraints on reciprocation, i.e. users are not prevented from sending data to others who do not provide them service in return.

Under these assumptions, the natural necessary conditions on the injection rates λ_i of commodities $i \in \mathcal{I}$ for feasibility of broadcast are the following:

$$\lambda_i \leq c_i, \quad i \in \mathcal{I}. \tag{37}$$

$$\sum_{i \in J} (|V_i| - 1)\lambda_i \leq \sum_{u \in \cup_{i \in J} V_i} c_u, \quad J \subseteq \mathcal{I}. \tag{38}$$

Indeed, the first inequality (37) states that each source node i must have a capacity larger than the rate λ_i at which it receives fresh data. Inequality (38) states that for each set of sources J, the total capacity required to forward the corresponding commodities to their receivers is no larger than the total capacity of users that can take part in their transmissions.

Let us now describe the scheduling policy we shall consider in the present setup. To determine who to send to, any node u evaluates the overall deprivation of potential receivers v as the total number of packets that it can send to v, and that v has neither received, nor is currently receiving. This evaluation is done over all commodities that both nodes are receivers of, i.e. over packets generated by all sources $i \in I(u) \cap I(v)$.

Node u then sends to its most deprived neighbour, where deprivation is measured as just decribed. Ties between equally deprived nodes are broken uniformly at random. The decision of which packet to send is done as follows. If u is a source node, it sends a fresh packet if it has any. In any other event, node u chooses which packet to send uniformly at random among the set of "useful packets", i.e. packets originating from sources $i \in I(u) \cap I(v)$. We refer to these rules as the "bundled most deprived - random useful" strategy, where bundling is over commodities.

In this context, we have the following

Theorem 9: Assume that conditions (37–38) hold, with strict inequalities. Assume further a complete communication graph, i.e. selection of the most deprived neighbour is made from the whole collection of nodes V. Then the Markov process describing the state of the system is ergodic.

Remark 2: The proof will again rely on fluid limit techniques. In fact, the fluid limits will be the same if we assume as in the previous subsection dynamic, restricted neighbourhoods instead of a complete communication graph. Thus, the above theorem remains true under the relaxed assumption on the neighbourhoods.

We now provide the main lines of the proof. We omit the parts that can readily be filled in from the previous proofs, and only describe the fluid trajectories in the present setup, the Lyapunov function and the proof that it decreases along these fluid trajectories.

C. Fluid Trajectories

In the multicommodity setup, we let $\mathcal{S}^i$ denote the collection of subsets of V_i that include node i, and $y_S^i(t)$ denote the quantity of packets generated by source i, and currently replicated at nodes $u \in S$, at time t. The cumulative number of packets originated from source i, which node u started to forward to node v while they were replicated at nodes in set S over the time interval $[0,t]$ is denoted by $\phi^i_{S,(uv)}(t)$.

We now describe the fluid trajectories of the multicommodity system.

Definition 3: The real-valued, non-negative functions $(y^i_S)_{i\in\mathcal{I},S\in\mathcal{S}^i}$ are called fluid trajectories of the multi-commmodities node-capacitated system if the following properties hold.

For all $i \in \mathcal{I}$, $S \in \mathcal{S}^i$, $u \in S$, $v \notin S$ such that $(u,v) \in V_i$, there exist non-decreasing, Lipschitz-continuous functions $\phi^i_{S,(uv)}$ with Lipschitz constant c_u, such that Equations (6) hold. Furthermore, using notation

$$y^i_{+u-v} := \sum_{S\in\mathcal{S}:u\in S,v\notin S} y^i_S,$$

for all $i \in \mathcal{I}$, $S \in \mathcal{S}^i$, $u \in S$, the functions $\{\phi^i_{S,(uv)}\}_{v\in V_i\setminus S}$ are differentiable at almost every t, and if $\sum_{j\in\mathcal{I}}\sum_{v\in V} y^j_{+u-v}(t) > 0$, their derivatives satisfy:

$$\frac{d}{dt}\phi^i_{S,(uv)}(t) = 0 \text{ if } \sum_{j\in I(u)\cap I(v)} y^j_{+u-v}(t) < \max_{v'\in V}\left(\sum_{j\in I(u)\cap I(v')} y^j_{+u-v'}(t)\right), \tag{39}$$

$$\sum_{i\in\mathcal{I}}\sum_{v\in V}\sum_{S:u\in S,v\notin S}\frac{d}{dt}\phi^i_{S,(uv)}(t) = c_u. \tag{40}$$

If $u \notin \mathcal{I}$, that is for a non-source node, one also has, for all $v \in V$, and assuming the condition

$$\sum_{i\in\mathcal{I}}\sum_{S:u\in S,v\notin S}\frac{d}{dt}\phi^i_{S,(uv)}(t) > 0$$

holds, for all $i \in I(u) \cap I(v)$ and all $S \in \mathcal{S}^i$ such that $u \in S$, $v \notin S$, the following equation:

$$\frac{d}{dt}\phi^i_{S,(uv)}(t) = \frac{y^i_S(t)}{\sum_{j\in I(u)\cap I(v)} y^j_{+u-v}(t)}\sum_{j\in I(u)\cap I(v)}\sum_{S'\in\mathcal{S}^j:u\in S',v\notin S'}\frac{d}{dt}\phi^j_{S',(uv)}(t)\cdot \tag{41}$$

For a source node $i \in \mathcal{I}$, one has the following:

$$y^i_{\{i\}} > 0 \Rightarrow \sum_{v \in V_i \setminus \{i\}} \frac{d}{dt}\phi^i_{\{i\},(iv)}(t) = c_i. \tag{42}$$

In the case where $y^i_{\{i\}} = 0$, one then has for all $v \in V_i \setminus \{i\}$, assuming the condition

$$\sum_{S \in \mathcal{S}^i : S \neq \{i\}, v \notin S} \frac{d}{dt}\phi^i_{S,(iv)}(t) > 0$$

holds, for all $S \in \mathcal{S}^i$ such that $S \neq \{i\}$ and $v \notin S$ the following:

$$\frac{d}{dt}\phi^i_{S,(iv)}(t) = \frac{y^i_S(t)}{\sum_{S' \in \mathcal{S}^i : S' \neq \{i\}, v \notin S'} y^i_{S'}(t)} \sum_{S' \in \mathcal{S}^i : S' \neq \{i\}, v \notin S'} \frac{d}{dt}\phi^i_{S',(iv)}(t). \tag{43}$$

$\Diamond$

D. Lyapunov Stability

The Lyapunov function we consider is the natural workload function $w(t) = \sum_{i \in \mathcal{I}} w^i(t)$, where

$$w^i(t) = \sum_{S \in \mathcal{S}^i} (|V_i| - |S|) y^i_S(t).$$

Let us evaluate the time derivative $\frac{d}{dt}w^i(t)$ at a point t where the workload function w^i is differentiable. To this end, we let $J(i)$ denote the set of nodes in V_i that are fully deprived of commodity i, i.e.

$$J(i) = \{u \in V_i : \sum_{v \in V_i \setminus \{u\}} y^i_{+u-v}(t) = 0\}.$$

We also introduce the notations

$$f^i_{S,(uv)} = \frac{d}{dt}\phi^i_{S,(uv)}(t)$$

and

$$c^i_u := \sum_{S \in \mathcal{S}^i : u \in S} \sum_{v \in V_i \setminus S} f^i_{S,(uv)}.$$

The latter quantity c^i_u can be interpreted as the rate at which node u forwards data from commodity i. We assume for now that $w^i(t) > 0$. The case where $w^i(t) = 0$ is easily dealt with: the derivative of the function must then be zero, since the function is non-negative.

We then have the following

$$\begin{array}{rcl} \frac{d}{dt}w^i(t) & = & \sum_{S \in \mathcal{S}^i} (|V_i| - |S|) \frac{d}{dt} y^i_S(t) \\ & = & (|V_i| - 1)\lambda_i - \sum_{v \in V_i \setminus \{i\}} f^i_{\{i\},(iv)} \\ & & + \sum_{S \in \mathcal{S}^i : S \neq \{i\}, S \cap J(i) = \emptyset} (|V_i| - |S|) \quad \Big[\sum_{u \in S, v \in S \setminus \{u\}} f^i_{S \setminus \{v\},(uv)} \\ & & \qquad - \sum_{u \in S, v \in V_i \setminus S} f^i_{S,(uv)} \Big] \end{array}$$

We further distinguish several cases.

Case 1: $J(i)=\emptyset$. Then the above sums telescope, and one obtains after recombinations:

$$\frac{d}{dt}w^i(t)=(|V_i|-1)\lambda_i-\sum_{u\in V_i}c^i_u. \tag{44}$$

Case 2: $J(i)\neq\emptyset$. Note that for any node $u\in V_i$, all terms $f^i_{S,(uv)}$ appear with a coefficient $-(|V_i|-|S|)\mathbf{1}_{S\cap J(i)=\emptyset}+(|V_i-|S|-1)\mathbf{1}_{(S\cup\{v\})\cap J(I)=\emptyset}$, which is non-positive, and at most -1 when $S\cap J(i)=\emptyset$. Hence, using (39), we obtain

$$\begin{array}{rcl}\frac{d}{dt}w^i(t) & \leq & (|V_i|-1)\lambda_i \\ & & -\sum_{S\in\mathcal{S}^i,S\cap J(i)=\emptyset}\sum_{v\in V_i\setminus S}f^i_{S,(iv)}\left[(|V_i|-|S|)-\mathbf{1}_{v\notin J(i)}(|V_i|-|S|-1)\right] \\ & & -\sum_{v\in V_i\setminus J(i),v\neq i}c^i_v.\end{array}$$

Note now that, because of Condition (39), $f^i_{S,(iv)}$ can only be positive when $v\in J(i)$. Therefore we obtain that

$$\begin{aligned}\frac{d}{dt}w^i(t) \quad & \leq \quad (|V_i|-1)\lambda_i-\sum_{v\in J(i)}f^i_{\{i\},(iv)}(|V_i|-1) && (45)\\ & \quad -\sum_{S\in\mathcal{S}^i,S\neq\{i\},S\cap J(i)=\emptyset}\sum_{v\in J(i)}f^i_{S,(iv)}(|V_i|-|S|) && (46)\\ & \quad -\sum_{v\in V_i\setminus J(i),v\neq i}c^i_v. && (47)\end{aligned}$$

Also, by Condition (40), necessarily

$$\sum_{S\in\mathcal{S}^i,S\cap J(i)=\emptyset}\sum_{v\in J(i)}f^i_{S,(iv)}=c_i.$$

Now, distinguish further acording to whether $y^i_{\{i\}}(t)=0$ or not. In the first case, necessarily the right-hand side in (45), being the derivative $\frac{d}{dt}y^i\{i\}(t)$, must equal zero, while the term (46) must necessarily equal $c_i-\lambda_i$. In the case where $y^i_{\{i\}}(t)>0$, in view of (42), necessarily the derivative $\frac{d}{dt}w^i(t)$ is no larger than $-(c_i-\lambda_i)(|V_i|-1)$.

It therefore holds that, when $J(i)\neq\emptyset$, necessarily

$$\frac{d}{dt}w^i(t)\leq-(c_i-\lambda_i)-\sum_{v\in V_i\setminus J(i),v\neq i}c^i_v. \tag{48}$$

Let us denote by $\mathcal{I}'$ the set of those sources $i\in\mathcal{I}$ for which $J(i)\neq\emptyset$ and $w^i(t)>0$, and by $\mathcal{I}''$ the set of those sources $i\in\mathcal{I}$ for which $J(i)=\emptyset$.

Combined together, (44) and (48) yield

$$\begin{array}{rcl}\sum_{i\in\mathcal{I}}\frac{d}{dt}w^i(t) & \leq & -\sum_{i\in\mathcal{I}'}(c_i-\lambda_i) \\ & & +\sum_{i\in\mathcal{I}''}(|V_i|-1)\lambda_i-\sum_{u\in\cup_{i\in\mathcal{I}''}V_i}c_u \\ & & +\sum_{u\in\cup_{i\in\mathcal{I}''}V_i}\sum_{i\in I(u)}c^i_u\mathbf{1}_{u\in J(i)},\end{array}$$

where we have used (40). We now argue that for any node u, if it is a receiver of a commodity $i \in \mathcal{I}''$ and belongs to $J(i')$ for another commodity i', then necessarily, $c_u^{i'} = 0$. This follows indeed from (41). It therefore follows that

$$\begin{array}{rcl} \sum_{i\in\mathcal{I}} \frac{d}{dt} w^i(t) & \leq & -\sum_{i\in\mathcal{I}'}(c_i-\lambda_i) \\ & & +\sum_{i\in\mathcal{I}''}(|V_i|-1)\lambda_i - \sum_{u\in\cup_{i\in\mathcal{I}''}V_i} c_u. \end{array} \tag{49}$$

Thus the workload function $\sum_{i\in\mathcal{I}} w^i(t)$ verifies

$$w(t) \leq \max(0, w(0)-\epsilon t)$$

where

$$\epsilon = \min\left\{\min i \in \mathcal{I}\left[c_i-\lambda_i\right], \min_{J\subseteq\mathcal{I}, J\neq\emptyset}\left[\sum_{u\in\cup_{i\in J}V_i} c_u - \sum_{i\in J}(|V_i|-1)\lambda_i\right]\right\}$$

is strictly positive under the assumptions of Theorem 9.

Ergodicity then follows exactly by the same arguments as in the proofs of Theorems 1 and 6.

VI. Conclusion

We have identified distributed scheduling strategies for live streaming, and proven a rate optimality property for several network capacity models.

Many open problems remain, concerning the performance achievable using simple, "unstructured" peer-to-peer mechanisms such as those considered in this paper. In particular, we do not provide any guarantees on the delays with which packets reach receivers. Very few results are available on the delay performance of such distributed schemes (notable exceptions being [11], [1]), and it would be interesting to obtain such results for the schemes we just presented.

Another issue of interest concerns the performance of such schemes in the presence of "relay nodes". There, the use or not of network coding will dramatically affect the theoretical rate at which data can be streamed. But in any case, with or without network coding, it is an open question whether simple "epidemic" schemes of the kind we discussed can achieve optimal diffusion rates.

References

[1] T. Bonald, L. Massoulié, F. Mathieu, D. Perino and A. Twigg. Epidemic Live-Streaming: Optimal Performance Trade-Offs. Preprint, 2007.

[2] B. Cohen, Incentives build robustness in BitTorrent. Technical Report, bittorrent.org, 2003.

[3] M. Castro, P. Druschel, A.-M. Kermarrec, A. Nandi, A. Rowstron and A. Singh, Splistream: High-bandwidth multicast in cooperative environments. In Symposium on Operating Systems Principles (SOSP 2003), 2003.

[4] J.G. Dai, On positive Harris recurrence of multiclass queuing networks: a unified approach via fluid limit models, Ann. of Applied Probability 5, 49-77, 1995.

[5] J. Edmonds, Edge-disjoint branchings, in *Combinatorial Algorithms*, R. Rustin, Ed., pp. 21–31. Algorithmics Press, 1972.

[6] A. Legout, N. Liogkas, E. Kohler and L. Zhang. Clustering and Sharign Incentives in BitTorrent Systems. In Proc. ACM Sigmetrics'2007, 2007.

[7] L. Massoulié, A. Twigg, C. Gkantsidis and P. Rodriguez. Randomized decentralized broadcasting algorithms. In IEEE INFOCOM 07, 2007.

[8] L. Massoulié, A. Twigg, C. Gkantsidis and P. Rodriguez. Provably optimal decentralized broadcast algorithms. Tech. Report MSR-TR-2006-105, Microsoft Research, Jul. 2006.

[9] P. Robert, *Stochastic Networks and Queues*, Springer, 2003.

[10] A.N. Rybko and A.L. Stolyar, Ergodicity of stochastic processes describing the operations of open queueing networks. Problemy Pederachi Informatsii, 28, 2–26, 1992.

[11] S. Sanghavi, B. Hajek and L. Massoulié, Gossiping with Multiple Messages. IEEE Trans. Information Theory, Vol. 53 (12), pp.4640–4654, 2007.

[12] T. Silverston and O. Fourmeaux, P2P IPTV measurement: A case study of TVants. In Proc. of the 2nd Conference on Future Networking Technologies (CoNEXT 06), 2006.

[13] H.Q. Ye, J.H. Ou and X.M. Yuan, Stability of Data Networks: Stationary and Bursty Models. Operations Research, 53(1), 107-125, 2005.

[14] X. Zhang, J. Liu, B. Li amnd T.-S. Yum, “DONet/CoolStreaming: A Data-driven Overlay Network for Live Media Streaming”, in Proc. IEEE INFOCOM’05, Miami, 2005.