Fluid Models of Population and Download Progress in P2P Networks

Andrés Ferragut and Fernando Paganini, *Fellow, IEEE*

***Abstract*—This paper studies partial differential equation (PDE) models for the dynamics of peer-to-peer (P2P) file-sharing networks. Using as independent variables time and a fluid measure of residual work, our PDE model tracks the population profile of the P2P swarm, allowing for general file-size distributions. Focusing on the processor sharing discipline, which we validate as an accurate model of homogeneous P2P networks, we provide a series of analytical results invoking tools of feedback control theory. We establish local stability of the equilibrium, analyze variability around this equilibrium, and compute transient response times, all of which are shown to match tightly with simulation results for a full packet-level implementation of the BitTorrent protocol. We also extend our model to heterogeneous bandwidth scenarios, and to the case of peers contributing to the system after they finish download.**

***Index Terms*—Communication networks, peer-to-peer computing, stability.**

I. Introduction

IN RECENT years, peer-to-peer (P2P) file sharing systems such as BitTorrent [4] have become widespread, representing an important portion of total Internet traffic. The power of P2P as a means of content distribution lies on the fact that downloading peers contribute their upload bandwidth, so the supply of capacity scales with demand.

This feature of self-scaling presents challenges when it comes to modeling the evolution of the system over time. Peer population dynamics is driven by the download rate the network is able to provide, which in turn is determined by the population size: understanding this feedback loop is key to establishing that the P2P network reaches a stable, scalable operation. Another modeling challenge is to describe the resource allocation between peers: given the piece-exchange mechanisms (e.g. BitTorrent's tit-for-tat rules), is the global result egalitarian or influenced by the amount of content possessed? Finally, populations do not completely characterize the dynamics, because job-sizes do not follow a memoryless distribution. For instance, a swarm of peers interested in the same content (e.g. a file) has a fixed, deterministic download job size. Dynamic models must therefore track download progress in addition to population.

In this paper we address these issues by constructing a fluid model for a P2P file sharing system for general file sizes, which may include a variety of resource allocation disciplines. Based on a Partial Differential Equation (PDE), the model keeps track of populations and download progress, and can be used to analyze both steady-state download profiles and dynamic questions of stability, variability and transient response. This fluid model is appropriate to describe the average behavior of the system for large population sizes.

We begin in Section II by reviewing the structure of a P2P file sharing system and some of the previous works on its dynamics. In Section III we present our PDE model, focusing on a *processor sharing* model for bandwidth allocation, which we validate as accurate for the case of peers with homogeneous network access parameters. For our dynamic studies we consider here the case of a fixed number of server peers (*seeders* in BitTorrent terminology) in the system, while downloading *leechers* follow an arrival/departure process. We characterize the equilibrium download profile, and establish its local stability invoking tools of feedback control theory; we also use such methods to evaluate the variability of populations around equilibrium, and the transient times to empty the system when there are no arrivals. All these mathematical derivations are shown to match very closely with packet level simulations performed in `ns2`, based on a complete implementation of the BitTorrent protocol, and are shown to improve on previously existing models.

In Section IV the model is generalized to several classes of peers with heterogeneous upload bandwidth constraints. Under a plausible model for bandwidth allocation among these classes, we show that equilibrium and stability results extend and again validate the results by packet-level simulations. In Section V we extend our model to incorporate seeder variability by modeling peers staying in the system after completing download. We extend our stability and variability results to this case and validate them by packet simulation. Conclusions and lines of future work are presented in Section VI. Partial results leading up to this paper were presented in [9], [22].

II. Background and related work

In a P2P system, *content* is disseminated by subdividing it into small *pieces* or *chunks*, and enabling peers to exchange such units bidirectionally. Thus every peer present is a server; those who are also clients are referred to as *leechers*. Peers present in the system only to altruistically distribute content are referred as *seeders*, and possess an entire copy of the content. Typically there is a common file or set of files of fixed size in which all peers are interested (e.g. a swarm around a single torrent). In this case, the job size each client demands from the system is deterministic in nature. More generally, the content may be a *bundle* of different files as proposed in [20].

The authors are with the Mathematics Applied to Telecommunications and Energy Research Group, Universidad ORT Uruguay, Montevideo, Uruguay (e-mail: ferragut@ort.edu.uy).

Research supported by AFOSR-US under Grant FA9550-12-1-0398.

Here, the system stores a large directory and different peers could have interest in downloading only a portion of it, which introduces variability in the job size distribution peers demand from the network.

Understanding the dynamics of peer populations has been a topic of active research. In [28], a Markov chain model for the populations of leechers and seeders is proposed, assuming exponential transition times. Despite this simplifying assumption, the model proves hard to solve explicitly and numerical studies were given. A fluid version, with some additional extensions, was studied in [24]; this ordinary differential equation (ODE) model leads to explicit results for the equilibrium of the system and its stability, and is a direct predecessor to the PDE generalization to be pursued in this paper. The case of heterogeneous peer rates was considered in [3]. The main limitation of these models is that they summarize the state of the system in the total number of leechers, giving no information on download progress.

At the other extreme are Markov models which consider a population variable for each subset of possessed pieces [13], [18]; the number of state variables is thus huge, exponential in the number of chunks (which is commonly in the hundreds). For such an enormous state dimension it is harder to justify the use of fluid limits, since in practice most subsets will have empty population.

A successful application of such Markov models has been to identify, via Foster-Lyapunov techniques, certain population instabilities that may arise due to the scarcity of a certain piece [11], [32]. These occur when the rate of new peer arrivals exceeds the ability of seeders to inject new chunks. However, these models assume leecher encounters are random, and also the requested piece is random. Therefore, as one peer approaches the end of the file, the probability of finding the right peer to exchange is small, introducing a bottleneck that leads to instability in some cases. While interesting, real world implementations of BitTorrent include several mechanisms to direct piece exchanges in a more efficient way. In particular, BitTorrent systems spread information on chunk availability ("*have*" messages) so peers can target their requests to peers that have valuable pieces, and "*rarest-first*" policies of piece selection are used to combat chunk rarity. In this regard, [19] models the piece availability mechanisms of BitTorrent and concludes that, in an homogeneous scenario, each piece takes roughly the same time to download, and piece availability is uniform. From an empirical perspective, [14] performs extensive real-world studies using BitTorrent, which agree with this conclusion.

In this paper we adopt an intermediate approach in which remaining workload is used to index a population state, but without taking into account the identity of individual chunks. This will enable us to go beyond memoryless distributions for job sizes, and thus improve the accuracy of predictions, yet avoid the high complexity of full state models. A similar approach was successfully applied in [23] in the context of bandwidth-sharing networks. A related line of work to model download progress in P2P is the use of multiple tandem queues, as in [5], [15], [16], [19]. In particular, [7] shows that modeling two download stages by ODEs provides an improvement in predictive power with respect to the single ODE in [24]; this can be seen as a first discrete step in endowing job sizes with a distribution of smaller variance than that of an M/M queue. Our PDE model is the natural conclusion in this progression, covering all job type distributions, including the deterministic case with no variance in job size.

Another aspect of P2P that has received attention is the resource allocation between leechers. A high-level discussion of this design space is given in [8], see also [1], [2], [30]. In the case of BitTorrent, this allocation is indirectly determined by the tit-for-tat reciprocity mechanism introduced in [4] to avoid free-riding. An alternative proportional reciprocity mechanism, more amenable to analytic studies, was studied in [29] and implemented in [17]; theoretical results on resulting allocation under these strategies are given in [27]. A conclusion of these studies is that for peers of *homogeneous* upload capacity, an egalitarian download distribution results. We will incorporate this assumption in the next section.

III. Fluid model for a P2P system

We introduce here our fluid model for the population of leechers in a P2P system, using as state descriptor a function valued state that keeps track of the download progress. As discussed above, the main motivation to introduce download progress in the state is to incorporate the possibility that job-sizes come from a general distribution. In fact, if the content consists of a single file, then the job size is the same for every peer (deterministic). If the content is a *bundle* of files, peers may be interested in different parts of it, which introduces randomness in the demand.

As in [28], consider a P2P system where fresh peers arrive as a Poisson process of intensity λ. Instead of assuming a distribution of *service time*, it is natural here to model the *job size* requirement:[1] each peer n requires some amount of content σ_n, which we assume random with complementary cumulative distribution function (CCDF) $H(\sigma) := P(\sigma_n > \sigma)$. We assume that H is normalized such that $E[\sigma_n] = \int_0^\infty H(\sigma)d\sigma = 1$, which amounts to a choice of units. Examples of $H(\sigma)$ are the exponential case $H(\sigma) = e^{-\sigma}$, and the deterministic size case $H(\sigma) = \mathbf{1}_{[0,1)}(\sigma)$ where all peers demand the same amount of content. We denote by $x(t)$ the total amount of leechers present in the system at time t.

As a first step we shall assume that the total number of seeders in the system is fixed and equal to y_0, whereas leechers completing their download leave the system immediately. This idealizes a rather common scenario in practice, where torrents often rely mainly on a set of permanent, highly committed seeders (publishers). In Section V we discuss how to include seeder variability. We also assume that all participating peers contribute the same amount of upload capacity into the system, which we denote by μ (in files/second). In Section IV we discuss how to generalize the model to heterogeneous upload capacities.

For generally distributed job sizes, to characterize the system state the total population $x(t)$ does not suffice [25]. Instead, the state must keep track of the residual workload of

[1] Both are not interchangeable when service capacity per job varies in time.

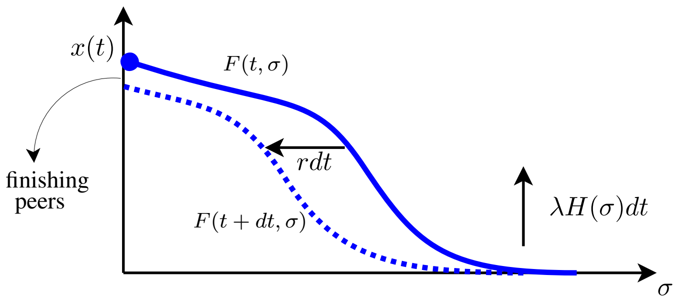

Fig. 1. Fluid state evolution.

each present peer. Let $F(t,\sigma)$ be a function that counts the number of leechers that at time t have residual workload larger than σ. When the scale of the system is large, we can treat $F(t,\sigma)$ as a fluid (real) variable. If leechers are served at rate r (possibly dependent on the population size), the evolution equation for $F(t,\sigma)$ is:

$$F(t+dt,\sigma) = F(t,\sigma+rdt) + \lambda H(\sigma)dt.$$

This evolution equation is better explained through Fig. 1. The first term is due to downloading: after a short time dt, the population profile $F(t,\cdot)$ is transported towards zero at rate r. The second term $\lambda H(\sigma)dt$ accounts for the fresh arrivals above level σ, corresponding to the thinned Poisson arrival process with probability $H(\sigma)$.

Subtracting $F(t,\sigma)$, dividing by dt and letting $dt \to 0$ we arrive at the following PDE state equation:

$$\frac{\partial F}{\partial t} = \lambda H(\sigma) + r\frac{\partial F}{\partial \sigma}.$$

A derivation of the above dynamics as a fluid limit of stochastic processes is outside the scope of this paper. We refer the reader to [10], [23] for a discussion on the connection between stochastic fluid limits and PDE models. Here, we shall focus on the fluid state descriptor and its predictions of the system behavior, and validate its conclusions through simulation experiments.

To completely specify our model, we must choose a suitable rate function r. In principle, $r = r(F, y, \sigma)$ may depend on the current profile, the number of seeders and the download stage (see [9]), and would be affected by the piece-exchange details outside our model. Nevertheless, we will show evidence in Section III-D that the following macroscopic assumptions are approximately satisfied in the case of BitTorrent:

- A1 The P2P sharing mechanism is *efficient*, in the sense that the total upload bandwidth $\bar{R}_{up} = \mu(x+y_0)$ is used for sharing.
- A2 This available bandwidth is distributed evenly among all downloading peers independently of their download progress, i.e. the system behaves as a Processor Sharing queue with state-dependent rates.

Under these assumptions, the downloading rate of a given leecher when the population size is x is given by:

$$r = \mu \cdot \left(\frac{x+y_0}{x}\right);$$

and therefore, the complete dynamics for the system are captured by the following PDE:

$$\frac{\partial F}{\partial t} = \lambda H(\sigma) + \mu \cdot \left(\frac{x+y_0}{x}\right)\frac{\partial F}{\partial \sigma}, \tag{1}$$

with $x(t) := F(t,0)$, the total amount of leechers in the system, and the boundary condition $F(t,\infty) = 0$ for all t. Recall that μ is the constant upload bandwidth contributed by peers in the system, in files per second.

As an example, consider the case of exponentially distributed job sizes, $H(\sigma) = e^{-\sigma}$. In that case, we can seek a solution by separation of variables, i.e. $F(t,\sigma) = x(t)e^{-\sigma}$. Substituting in (1), we have

$$\dot{x}(t)e^{-\sigma} = \lambda e^{-\sigma} - \mu\frac{x(t)+y_0}{x(t)}x(t)e^{-\sigma};$$

after canceling terms, we see that F is a solution provided $x(t)$ follows:

$$\dot{x}(t) = \lambda - \mu(x(t)+y_0). \tag{2}$$

This is a version of the ODE dynamics proposed by [24], specialized to the case of a fixed seeder population. This indicates that the ODE model carries with it the assumption of exponentially distributed jobs.

A practically more interesting case is when all peers demand the same content, so job sizes are deterministic and $H(\sigma) = \mathbf{1}_{[0,1)}(\sigma)$. In this case, since no mass arrives beyond $\sigma = 1$ we can restrict the dynamics to $\sigma \in [0,1]$ leading to:

$$\frac{\partial F}{\partial t} = \lambda + \mu \cdot \left(\frac{x+y_0}{x}\right)\frac{\partial F}{\partial \sigma}, \tag{3}$$

for $0 \leqslant \sigma \leqslant 1$ with the boundary condition $F(t,1) = 0$ at all times. As we shall see, (3) will provide a better description of the system in this case than the ODE dynamics (2).

A. Equilibrium and stability

Let us analyze the equilibrium of the dynamics (1). Denote by * the values of the system at equilibrium. Setting $\partial F/\partial t = 0$ and integrating in the positive real line we have:

$$\lambda \int_0^\infty H(\sigma)d\sigma + \mu\left(\frac{x^*+y_0}{x^*}\right)\int_0^\infty \frac{\partial F^*}{\partial \sigma}d\sigma = 0. \tag{4}$$

Since we normalized the job size, $\int_0^\infty H(\sigma)d\sigma = 1$, and due to the hypotheses on F, $\int_0^\infty \frac{\partial F^*}{\partial \sigma}d\sigma = -F^*(0) = -x^*$. We conclude that in equilibrium the number of leechers should satisfy:

$$\lambda - \mu(x^*+y_0) = 0. \tag{5}$$

We now distinguish two cases, as a function of $\rho := \frac{\lambda}{\mu}$, which can be interpreted as the system load, since $1/\mu$ is the time to upload a copy of the file. If $\rho < y_0$, seeders alone can cope with the load, in which case we have a *seeder sustained* system. In this case, no equilibrium with positive x^* exists and the solution of the PDE approaches 0 as $t \to \infty$.

The most important case is when $\rho > y_0$: here the seeders alone cannot cope with the load and the effect of P2P is to stabilize the system through leecher contribution around a nonzero population. We have a *globally sustained* system, and the number of leechers in equilibrium is:

$$x^* = \rho - y_0.$$

Substituting this value in the equilibrium condition, and using the boundary condition $F^*(\infty) = 0$ we arrive at:

$$F^*(\sigma) = \int_\sigma^\infty \frac{\partial F^*}{\partial \sigma} d\sigma = (\rho - y_0) \int_\sigma^\infty H(s) ds.$$

The last term can be interpreted by recalling the concept of *residual lifetime distribution* from renewal theory: Given a stationary renewal process on the real line, with inter-event distribution with CCDF H and mean 1, the distance to the first event starting from any given point is distributed with CCDF

$$\bar{H}(\sigma) = \int_\sigma^\infty H(s) ds.$$

With this notation the equilibrium becomes simply:

$$F^*(\sigma) = (\rho - y_0)\bar{H}(\sigma). \tag{6}$$

Therefore the fluid equilibrium predicts that the system will have $x^* = \rho - y_0$ leechers present, and with remaining workloads distributed as the residual lifetimes[2] associated to H. This is consistent with known properties of processor sharing systems in equilibrium [10]. If we specialize (6) to the case of deterministic job sizes, it is easy to see that $\bar{H}(\sigma) = 1 - \sigma$ for $\sigma \in [0, 1]$. Therefore, in equilibrium the peers show a uniform download progress.

We now analyze local stability of the equilibrium (6), introducing incremental variables δx, f as $x = x^* + \delta x$ and $F = F^* + f$, and a perturbation signal $n(t)$ in the arrivals as an input to the system. With this notation, (1) becomes:

$$\frac{\partial(F^* + f)}{\partial t} = (\lambda + n(t))H(\sigma) + (r^* + \delta_r)\frac{\partial(F^* + f)}{\partial \sigma}$$

By canceling the equilibrium terms and discarding higher order terms, the linearization becomes:

$$\frac{\partial f}{\partial t} = H(\sigma)n + r^* \frac{\partial f}{\partial \sigma} + \frac{\partial F^*}{\partial \sigma}\delta r.$$

Noting that $\delta r = -\mu \frac{y_0}{x^{*2}} \delta x$ and $\frac{\partial F^*}{\partial \sigma} = -x^* H(\sigma)$ we have

$$\frac{\partial f}{\partial t} = r^* \frac{\partial f}{\partial \sigma} + \mu \frac{y_0}{\rho - y_0} H(\sigma)\delta x + H(\sigma)n.$$

Denoting now by $\tau = \frac{1}{r^*} = \frac{1}{\mu}\frac{\rho - y_0}{\rho}$ the equilibrium download time, the linear dynamics becomes

$$\frac{\partial f}{\partial t} = \frac{1}{\tau}\frac{\partial f}{\partial \sigma} + \frac{1}{\tau} H(\sigma) \underbrace{\left(\frac{y_0}{\rho}\delta x + \tau n\right)}_{u}. \tag{7}$$

Identifying the feedback signal u from the previous equation, we are led to the structure of Fig. 2, in which P is the infinite dimensional system

$$P : \begin{cases} \dfrac{\partial f}{\partial t} = \dfrac{1}{\tau}\dfrac{\partial f}{\partial \sigma} + \dfrac{1}{\tau} H(\sigma) u, \\ \delta x = f(t, 0). \end{cases} \tag{8}$$

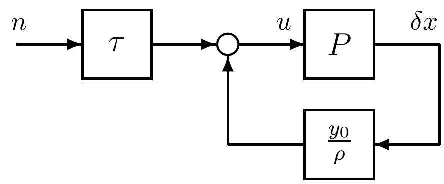

Fig. 2. Linearized dynamics as a feedback loop, with injected noise.

We analyze the stability of the feedback loop using Laplace transforms, which is non-trivial due to infinite dimensionality.

Theorem 1: The transfer function of subsystem P is

$$\hat{P}(s) = \int_0^\infty e^{-s\tau\eta} H(\eta) d\eta, \tag{9}$$

and the input-output transfer function is:

$$\hat{Q}(s) = \frac{\tau \hat{P}(s)}{1 - \frac{y_0}{\rho}\hat{P}(s)}, \tag{10}$$

which is stable (analytic in $\mathrm{Re}(s) \geqslant 0$).

Proof: To derive the transfer function of the subsystem P, define $\hat{f}(s, \sigma)$ the Laplace transform in the time variable of $f(t, \sigma)$. Transforming (8) with zero initial conditions yields:

$$s\hat{f}(s, \sigma) = \frac{1}{\tau}\frac{\partial \hat{f}}{\partial \sigma} + \frac{1}{\tau}\hat{u}(s)H(\sigma).$$

This is now an ordinary differential equation in σ. It can be readily verified that the solution of this equation with the boundary condition $\hat{f}(s, \infty) = 0$ can be represented as:

$$\hat{f}(s, \sigma) = \hat{u}(s) \int_0^\infty e^{-s\tau\eta} H(\sigma + \eta) d\eta,$$

and recalling that $\hat{\delta x}(s) = \hat{f}(s, 0)$, taking $\sigma = 0$ in the above equation we get the desired result for $\hat{P}(s)$. The expression for $\hat{Q}(s)$ now follows by calculating the closed loop gain of the feedback loop.

For stability, we will use the small-gain theorem (cf. [31]) which applies to transfer functions seen as operators, in particular the $\mathcal{H}_\infty$ (L_2-induced) norm of P satisfies:

$$\|\hat{P}(s)\|_{\mathcal{H}_\infty} = \sup_{\omega \in \mathbb{R}} |\hat{P}(j\omega)| = \sup_{\omega \in \mathbb{R}} \left| \int_0^\infty e^{-j\omega\tau\eta} H(\eta) d\eta \right| \leqslant \int_0^\infty H(\eta) d\eta = 1.$$

where we recall that job sizes have mean 1. Since we are in the case $\rho > y_0$, the loop gain is

$$\left\| \hat{P}(s) \frac{y_0}{\rho} \right\|_{\mathcal{H}_\infty} \leqslant \frac{y_0}{\rho} < 1;$$

the small-gain theorem implies that the closed loop transfer function $\hat{Q}(s)$ in (10) is analytic in the closed right half-plane. The system is thus locally stable. ■

[2]Despite the use of this standard terminology, we emphasize that the independent variable here is workload, not time.

B. Variability analysis

The feedback loop analysis performed above enables us to also characterize the variability around the equilibrium value. We consider the case of deterministic job sizes, with $H(\sigma) = \mathbf{1}_{[0,1)}(\sigma)$, where the only source of variability corresponds to peer arrivals. In this case, the noise signal $n(t)$ above represents variability in the Poisson arrival process, which at time t can be modeled as $\int_0^t (\lambda + n(\tau))d\tau$. Here $n(\tau)$ is a stationary process of power spectral density $S_n(\omega) \equiv \lambda$.

Using this representation we can characterize the steady state variance of δx. As a stationary process, the output variations δx will have power spectral density $S_x(\omega) = \lambda|\hat{Q}(j\omega)|^2$, where $\hat{Q}$ is the input-output transfer function of the loop (10). It turns out that:

$$\|\hat{Q}\|_2^2 = \int_{-\infty}^{\infty} |\hat{Q}(j\omega)|^2 \frac{d\omega}{2\pi} = \frac{1}{\mu},$$

and therefore the steady state variance of δx is given by:

$$\int_{-\infty}^{\infty} S_x(\omega) \frac{d\omega}{2\pi} = \frac{\lambda}{\mu} = \rho.$$

C. Transient analysis

We now analyze the predictive power of the PDE model in a transient situation, away from equilibrium. An interesting such scenario is when the arrivals are "turned off", and thus the system will empty in a finite time after all leechers complete their download. We wish to evaluate this completion time.

The relevant model is (1) without the arrivals term:

$$\frac{\partial F}{\partial t} = \mu \left(\frac{x + y_0}{x} \right) \frac{\partial F}{\partial \sigma}. \tag{11}$$

As for the initial condition, we assume that $F(0, \sigma) = \varphi(\sigma)$, a strictly decreasing differentiable function of σ, with $x_0 := \varphi(0)$, the initial total number of leechers, and $\varphi(\infty) = 0$.

Proposition 1: The time needed to empty a processor-sharing P2P system with y_0 servers and starting from an initial condition $\varphi(\sigma)$ is given by:

$$T = \frac{1}{\mu} \int_0^{\infty} \frac{\varphi(\sigma)}{\varphi(\sigma) + y_0} d\sigma. \tag{12}$$

Proof: For this simplified system with no input, the PDE model (11) with initial condition $\varphi(\sigma)$ can be written in the integral form

$$F(t, \sigma) = \varphi \left(\sigma + \int_0^t \mu \left(1 + \frac{y_0}{x(\tau)} \right) d\tau \right),$$

valid for $\{t : x(t) > 0\}$. That this is a solution can be readily verified by substituting in (11). Evaluating the preceding expression in $\sigma = 0$ gives the following integral equation for the number of leechers $x(t)$:

$$x(t) = F(t, 0) = \varphi \left(\int_0^t \mu \left(1 + \frac{y_0}{x(\tau)} \right) d\tau \right).$$

This is valid while $x(t) > 0$, and since φ is strictly decreasing in this range we can solve for

$$\varphi^{-1}(x(t)) = \int_0^t \mu \left(1 + \frac{y_0}{x(\tau)} \right) d\tau.$$

Differentiating in t we get the following autonomous differential equation for x:

$$(\varphi^{-1})'(x)\dot{x} = \mu \left(1 + \frac{y_0}{x} \right), \quad x(0) = x_0.$$

Applying separation of variables and integrating in $[0, T]$ gives

$$\frac{1}{\mu} \int_{x_0}^{x(T)} \frac{x}{x + y_0} (\varphi^{-1})'(x) dx = T.$$

When $x(T)$ tends to zero we obtain the expression for the completion time:

$$T = \frac{1}{\mu} \int_{x_0}^{0} \frac{x}{x + y_0} (\varphi^{-1})'(x) dx.$$

Finally, the change of variables $\sigma = \varphi^{-1}(x)$ in the above integral leads to (12). ■

It is worth specializing the above result to the case of deterministic workloads of size 1, where (12) becomes:

$$T = \frac{1}{\mu} \int_0^1 \frac{\varphi(\sigma)}{\varphi(\sigma) + y_0} d\sigma. \tag{13}$$

and we assumed $\varphi(\sigma) = 0$ for $\sigma > 1$. Noting that $\varphi(\sigma) \leqslant x_0$ $\forall \sigma$ and the function $\frac{\xi}{\xi + y_0}$ is increasing in $\xi > 0$, we have the following:

Corollary 1: For deterministic job sizes, the completion time T satisfies

$$T \leqslant \frac{1}{\mu} \frac{x_0}{x_0 + y_0}. \tag{14}$$

In fact, the equality in the above expression is achieved when the initial condition φ approaches the function $x_0 \mathbf{1}_{[0,1)}(\sigma)$, i.e. when all the leechers start empty.

Note that in particular, T is bounded above by $1/\mu$, i.e. the time for completion is finite, and is at most $1/\mu$, the time to upload a copy of the file. This uniform bound holds regardless of the initial number of leechers! This again evidences the scalability of P2P file exchange mechanisms: when the demand is large, the supply scales to keep up with it.

Let us compare these results with the prediction of ODE models which do not track download progress. In particular, the model (2) with arrivals turned off becomes:

$$\dot{x} = -\mu (x + y_0), \quad x(0) = x_0.$$

With analogous (simpler) calculations, the completion time for this model can be readily calculated as:

$$T' = \int_0^{x_0} \frac{1}{\mu} \frac{1}{x + y_0} dx = \frac{1}{\mu} \log \left(1 + \frac{x_0}{y_0} \right).$$

In particular the ODE model predicts $T' \to \infty$ as $x_0 \to \infty$, albeit logarithmically. We show in our simulations below that this is pessimistic, and the time we found in (14) gives closer predictions.

The analysis presented here is only one way of considering the transient behavior, where the uplink bandwidth bottleneck plays the crucial role, and thus the scalability of P2P is exhibited. A complementary view of transients is that of epidemic models, where download times are neglected and propagation is controlled by the evolving peer connectivity, see for instance [21]. Also, in the case of bundles of multiple content files, [20], [26] analyze the possibility of a specific content becoming unavailable due to peer departures.

D. Simulations with fixed seeders, homogeneous leechers

In this section we offer simulation experiments that validate the accuracy of the models presented so far. All simulations were performed using the network simulator `ns2` with the BitTorrent library [6], which closely mimics the behavior of the BitTorrent protocol, including individual chunk availability, tit-for-tat rules and transport layer connections [4]. The file size of interest is of 100 MB (deterministic) and is subdivided in 400 small size chunks. Each peer may open connections up to 40 other peers in the network, thus not every peer interacts with every other. The physical uplink bandwidth of clients is 256kbps; however a fraction of this is spent in protocol overheads (TCP headers, BitTorrent control messages, etc.), empirically we found in our simulations that the useful bandwidth seen from the application layer is about 90% of the total. This amounts to choosing:

$$\mu = 0.9 \times \frac{256kbps}{8 \times 100MB} = 2.8 \times 10^{-4} s^{-1}.$$

This means it takes $\approx 1h$ for a single peer to upload a full copy of the file.

We begin by validating the assumptions that the bandwidth sharing attained by BitTorrent is efficient, and can be well approximated by a processor sharing discipline. To do so, we simulated the system with a fixed number of seeders $y_0 = 25$, and leechers arriving empty at a rate $\lambda = 1.6$ arrivals/min. Leechers that finish download depart immediately.

To measure the efficiency, we measure the total download rate $R_{down}(t)$ by counting the total bits downloaded at intervals of $\Delta t = 120s$, and compare its behavior with $\bar{R}_{up}(t) = \mu(x(t) + y_0)$, with $x(t)$ the current leecher population. In Fig. 3 we plot the evolution of the ratio $R_{down}/\bar{R}_{up}$, which is close to 1 at all times, meaning the system fully utilizes the available bandwidth, despite a varying population.

To evaluate the processor sharing hypothesis, we use the Jain index [12], i.e. we compute the individual download rates $(r_1(t), \ldots, r_x(t))$ of the peers present in the system at time t and define:

$$J(t) = \frac{(\sum r_i(t))^2}{x(t) \sum r_i(t)^2}.$$

Recall that $J(t) \leqslant 1$, with equality if and only if all download rates are equal.

The second graph in Fig. 3 shows the evolution of the Jain index, which stays most of the time with $J \geqslant 0.9$. Complete fairness ($J = 1$) is not easy to attain, since the underlying protocol incorporates complex mechanisms for peer discovery and reciprocity (tit-for-tat rules and optimistic unchoking), and works in a completely decentralized fashion.

To validate the uniformity of download rate across download stages, we performed another experiment where we fix the total population. The setting is as before, but instead of a random arrival process of leechers, each departing peer is replaced by a new fresh leecher. In this case $x(t) = x_0 = 75$ leechers and $y_0 = 25$ seeders. This enables us to fix the population and thus the total available rate in the system, isolating the piece-related effects from population variability.

In Fig. 4 we plot the received rate measured at different download stages and compare it to the processor sharing allocation $r = \mu\left(\frac{x_0+y_0}{x_0}\right)$. We observe that there is a short period at the beginning of download where the received rate is lower, but immediately the system approaches the egalitarian allocation. A microscopic explanation could be the following: as soon as a leecher owns a modest amount of pieces, in a diverse swarm it will find others interested in them and therefore occupy its limited number of unchoke slots. In a tit-for-tat environment, this implies the leecher's "bargaining power" becomes comparable to other peers with the same access rate, regardless of the portion of file they possess; thus obtained rates tend to equalize over the download. As for the last piece, a slight but not significant lowering of the rate is observed towards the end. Simulations across several population sizes show that, provided the number of seeders is not too small, the system behaves in this fashion. We refer the reader to [5] for a discussion of the case when the number of seeders becomes too small. We believe that these results validate approximating the resource allocation by processor sharing for homogeneous peers in the case of large swarms, as those described by the fluid limit.

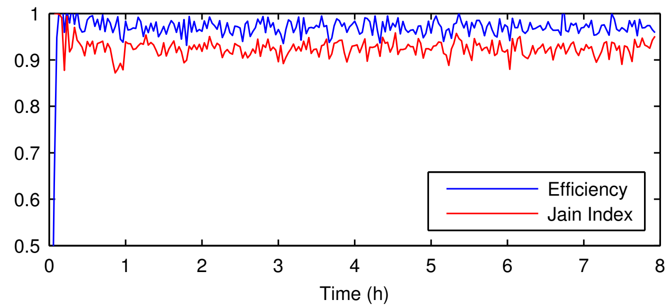

Fig. 3. Evolution of a system with fixed seeders. Efficiency (download/upload rate ratio) and Jain index of instantaneous download rates.

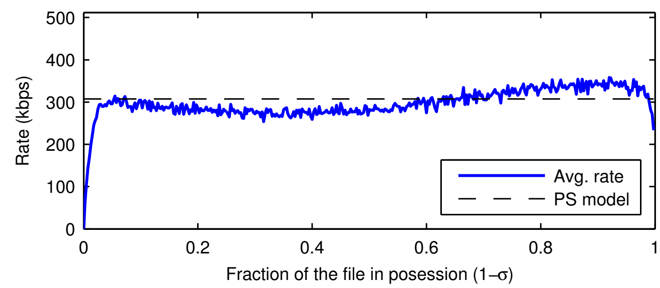

Fig. 4. Received rate r rate across download stages for a system with fixed total population.

We now compare the fluid model (3) with the simulation experiment. Again, we use $\lambda = 1.6$ arrivals per minute, $1/\mu \approx 1h$ and $y_0 = 25$. To capture the transient and steady-state behavior, we take $16h$ of simulation time. The load of this system is then $\rho = \frac{\lambda}{\mu} = 100$, and the equilibrium value for the leecher population is $x^* = \rho - y_0 = 75$. In Fig. 5 we plot the results compared against the trajectory predicted by the model for the same parameters. The PDE model tracks the evolution of the system during the transient, and correctly

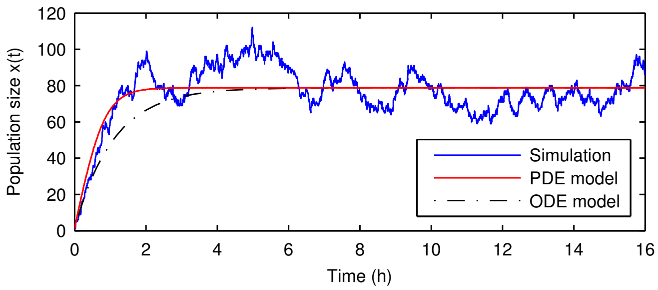

Fig. 5. Evolution of a system with fixed seeders: simulation and fluid model.

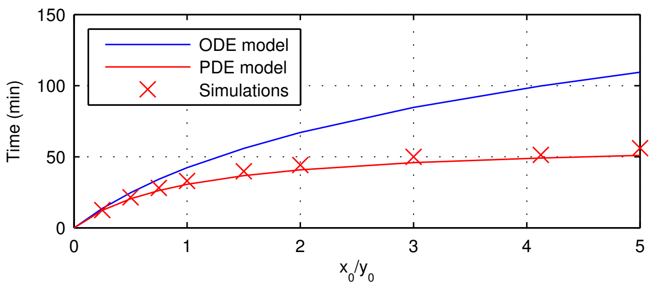

Fig. 6. Time to finish service in a P2P system with varying initial leechers.

tracks the average of the steady-state population. Moreover, from the analysis of III-B the steady state variance of $x(t)$ should verify $E[(\delta x)^2] = \rho$, corresponding in this case to a standard deviation of 10, whereas the measured standard deviation in the simulation is ≈ 10.67. For comparison, the ODE trajectory resulting from equation (2) is also shown. We observe that, although the equilibrium population predicted is the same, the transient is better captured by the PDE dynamics.

Finally we compare the time to empty the system, analyzed in III-C. We start the P2P system with x_0 leechers with no initial content, and y_0 seeders, which remain during the entire simulation. Finishing peers leave the system immediately. The time to finish service of the initial leechers is given by equation (14) which in particular in this case is bounded by $1/\mu \approx$ 60min. In Fig. 6 we plot the results for several initial values of the ratio between leechers and seeders x_0/y_0. The time to finish download predicted by our model is compared with simulation results, showing good fit. We also plot the time T' predicted by the simpler ODE model, which is pessimistic.

This experiment emphasizes again the scalability of P2P file-sharing, and the fact that to fully represent this feature an ODE model does not suffice; information on download progress plays a crucial role.

IV. Uplink bandwidth diversity

The model analyzed in Section III assumes an homogeneous uplink bandwidth across all leechers participating in the system. This is a strong assumption, since typically peers have different access links and may also commit different amounts of bandwidth to the P2P network. We now describe how to generalize the preceding model to the case of heterogeneous uplink bandwidths. For ease of exposition, we will only focus here on the case where all peers want to download the same content of fixed size. However, we note that the results extend to the case of general job sizes.

Consider a P2P system where leechers are classified in n groups according to their upload bandwidths, which we denote by $\{\mu_i\}$, measured in files per second. Leechers of class i arrive at a rate λ_i, and request the download of the file. Assume as before that there is a fixed number of seeders y_0, each one with uplink bandwidth μ_0.[3]

Let as before $F_i(t,\sigma)$, $i = 1,\dots,n$ denote the amount of leechers of class i that at time t have a pending download of at least σ, and define $x_i(t) := F_i(t,0)$ the total population of class i. In the case of deterministic job sizes, we can restrict ourselves to $\sigma \in [0,1]$ and the corresponding multi-class PDE model for the system is now:

$$\frac{\partial F_i}{\partial t} = \lambda_i + r_i(F,y_0)\frac{\partial F_i}{\partial \sigma}, \quad \sigma \in [0,1],\ i = 1,\dots,n.$$

To complete our model, we now choose r_i, the rate of service each class receives from the system, which comes from two contributions. On one hand, the seeders' portion, which is split evenly among all participating peers, in a processor-sharing fashion. On the other hand, each peer receives data from fellow leechers. As analyzed theoretically in [8] and by extensive simulations in [14], the tit-for-tat mechanism of BitTorrent tends to cluster leechers in groups of similar bandwidths, thus each will receive from other leechers roughly the same amount of bandwidth it provides to the network. This leads to the following model for the individual rates:

$$r_i = \frac{\mu_0 y_0}{\sum_j x_j} + \mu_i. \tag{15}$$

Plugging this rate into the PDE model gives the dynamics

$$\frac{\partial F_i}{\partial t} = \lambda_i + \left(\frac{\mu_0 y_0}{\sum_{j=1}^n x_j} + \mu_i\right)\frac{\partial F_i}{\partial \sigma}, \quad \sigma \in [0,1]. \tag{16}$$

A. Equilibrium and stability

We now analyze the equilibrium and stability of the dynamics (16). As in Section III, we do so in the *globally sustained* case, i.e. when the seeders alone cannot cope with the demand, which for heterogeneous rates translates to:

$$\sum_i \lambda_i > \mu_0 y_0. \tag{17}$$

Imposing equilibrium in (16) yields:

$$\lambda_i + \left(\frac{\mu_0 y_0}{\sum_{j=1}^n x_j^*} + \mu_i\right)\frac{\partial F_i^*}{\partial \sigma} \equiv 0,$$

which together with the boundary condition $F_i^*(1) = 0$ implies that $F_i^*(\sigma) = x_i^*(1-\sigma)$, where x^* satisfies:

$$\lambda_i = \left(\frac{\mu_0 y_0}{\sum_{j=1}^n x_j^*} + \mu_i\right) x_i^* \quad i = 1,\dots,n. \tag{18}$$

[3] Heterogeneous seeders could be included with essentially no change, this is avoided for simplicity.

Define

$$\alpha = \frac{\mu_0 y_0}{\sum_{j=1}^{n} x_j^*} > 0, \tag{19}$$

representing the equilibrium seeder bandwidth available per leecher, and solve (18) for

$$x_i^* = \frac{\lambda_i}{\alpha + \mu_i}. \tag{20}$$

Summing (18) over i and substituting for x_i^* leads to an equilibrium condition for α:

$$\sum_{i=1}^{n} \lambda_i \frac{\alpha}{\alpha + \mu_i} = \mu_0 y_0. \tag{21}$$

Denote the left-hand side of (21) as $g(\alpha)$. It is strictly increasing in $\alpha \geqslant 0$, $g(0) = 0$ and $g(+\infty) = \sum_i \lambda_i > \mu_0 y_0$ by condition (17). Therefore, there exists a unique root $\alpha > 0$ to (21). Substituting in (20), there is a unique point x^* with $x_i^* > 0$ that satisfies the conditions for equilibrium. It is straightforward to show conversely that satisfying (20), (21) is also sufficient, so the dynamics has a unique equilibrium.

We now analyze the local stability of this equilibrium through linearization. Denote by δx_i, δr_i the incremental scalar variables, and $f_i(t, \sigma) = F_i(t, \sigma) - F_i^*(\sigma)$; we have $f_i(t, 1) \equiv 0$ and $f_i(t, 0) = \delta x_i$. The linearization of (16) is:

$$\frac{\partial f_i}{\partial t} = r_i^* \frac{\partial f_i}{\partial \sigma} - x_i^* \delta r_i \tag{22}$$

where the incremental rate is given by

$$\delta r_i = -\frac{\mu_0 y_0}{(\sum_j x_j^*)^2} \sum_j \delta x_j = -\frac{\alpha}{\sum_j x_j^*} \sum_j \delta x_j,$$

with α from (19). The second term in (22) is now expressed as $-x_i^* \delta r_i = r_i^* \kappa_i \sum_j \delta x_j$, where

$$\kappa_i := \frac{\alpha}{r_i^*} \frac{x_i^*}{\sum_j x_j^*}.$$

Note $r_i^* = \alpha + \mu_i > \alpha$, so $\sum_i \kappa_i < 1$. Let us further denote

$$u_i := \kappa_i \sum_j \delta x_j. \tag{23}$$

It will also be convenient to define $\tau_i = (r_i^*)^{-1} = \frac{x_i^*}{\lambda_i}$ (equilibrium download time per peer). It follows that the linearized dynamics is the feedback interconnection of:

- A set of parallel blocks P_i, with input u_i and output δx_i, characterized by the infinite-dimensional dynamics

$$\frac{\partial f_i}{\partial t}(t, \sigma) = \frac{1}{\tau_i} \frac{\partial f_i}{\partial \sigma}(t, \sigma) + \frac{1}{\tau_i} u_i(t), \tag{24a}$$

$$\delta x_i(t) = f_i(t, 0), \tag{24b}$$

$$0 \equiv f_i(t, 1). \tag{24c}$$

- The static mapping (23), represented in matrix form by

$$u = K \mathbf{1}\mathbf{1}^T \delta x, \tag{25}$$

  where K is the diagonal matrix diag(κ_i) and $\mathbf{1}$ a column vector of ones.

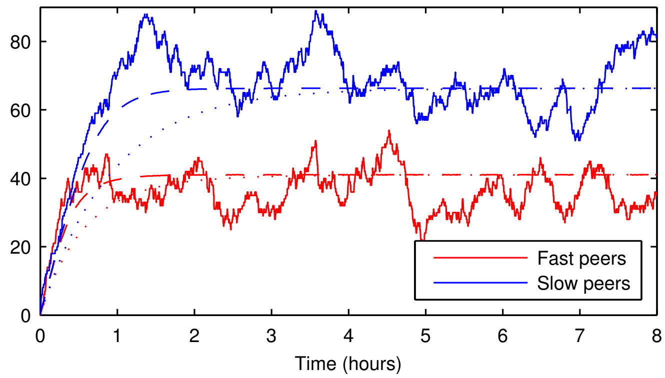

Fig. 7. Two-class experiment. Peer populations for simulated BitTorrent, PDE model (solid) and ODE model (dashed).

The block (24) is of the form (8), for deterministic job sizes, hence its transfer function follows from (9), giving:

$$\hat{P}_i(s) = \frac{1 - e^{-\tau_i s}}{\tau_i s},$$

In particular, it satisfies $\|\hat{P}_i(s)\|_\infty = \sup_{\omega \in \mathbb{R}} |\hat{P}_i(j\omega)| = 1$. In the present case, the loop transfer function $K\mathbf{1}\mathbf{1}^T$diag$(P_i(s))$ is of rank one, so its input-output stability is equivalent to that of the scalar loop gain

$$L(s) = \mathbf{1}^T \text{diag}(P_i(s)) K \mathbf{1} = \sum_i \kappa_i \hat{P}_i(s).$$

It follows that

$$|L(j\omega)| \leq \sum_i \kappa_i |\hat{P}_i(j\omega)| \leq \sum_i \kappa_i < 1,$$

hence input-output stability holds by a small-gain argument.

B. Simulations with fixed seeders, heterogeneous leechers

To illustrate and validate our analysis we report a packet-level simulation of the BitTorrent algorithm. There are $y_0 = 20$ seeders, the file size is $100MB$ and seeders have $512kbps$ of uplink bandwidth, which accounting for protocol inefficiencies gives $\mu_0 = 5.3 \times 10^{-4}$ files/sec (a copy is uploaded in $\approx 30min$). Leechers of two classes arrive with $\lambda_1 = \lambda_2 = 1.5$ peers/min, with $\mu_1 = \mu_0, \mu_2 = \frac{1}{2}\mu_0$.

Fig. 7 shows the results of an 8-hour run, with an empty initial condition. For comparison we include the traces of the numerical solution of (16). We also include as a reference an ODE version of the PDE model, including the rates in (15). Both models show convergence to an equilibrium roughly consistent with the experimental populations, with the PDE version giving tighter transient predictions.

V. Endogenously generated seeders

In this Section we extend our previous model to the case of peers finishing download and staying in the system to contribute, as first proposed in the models of [24], [28]. As in these references, we return here to the single class case with homogeneous leecher upload bandwidths.

Assume as before that peers arrive into the system with rate λ, demanding job sizes of CCDF $H(\sigma)$. When a leecher finishes its service, it becomes a seeder, staying in the network and contributing its upload bandwidth for some time. Let $y(t)$ the number of seeders present in the system at time t.

Under assumptions A1-A2, the service rate of leechers is $r = \mu \cdot \left(\frac{x+y}{x}\right)$, and the dynamics of the leecher population can be captured by (1) as before, but now with variable $y(t)$:

$$\frac{\partial F}{\partial t} = \lambda H(\sigma) + \underbrace{\mu \cdot \left(\frac{x+y}{x}\right)}_{r} \frac{\partial F}{\partial \sigma}, \tag{26a}$$

where again $x(t) = F(t, 0)$. To model the seeder dynamics, assume that each seeder departs the network at rate γ, then we have:

$$\dot{y} = -r \left.\frac{\partial F}{\partial \sigma}\right|_{\sigma=0} - \gamma y, \tag{26b}$$

where the first term accounts for download termination.

The P2P system is thus represented by the combined dynamics (26a), (26b) in the state variables $F(t, \sigma)$ and $y(t)$. Alternatively, one can replace $y(t)$ by the variable:

$$z(t) = x(t) + y(t) = F(t, 0) + y(t),$$

the total population of the network. By evaluating (26a) in $\sigma = 0$ and adding it to (26b) we deduce the following dynamic equation for z:

$$\dot{z} = \lambda - \gamma z + \gamma x. \tag{26c}$$

A. Equilibrium and stability

Denoting as before by * the equilibrium values, and imposing the equilibrium condition in (26c) we get:

$$\lambda - \gamma z^* + \gamma x^* = 0,$$

so in any equilibrium of (26) the number of seeders $y^* = z^* - x^*$ should be $y^* = \frac{\lambda}{\gamma}$. Plugging the equilibrium values in (26a) and integrating in the positive real line we get the following condition for x^*:

$$\lambda - \mu \left(\frac{x^* + y^*}{x^*}\right) x^* = 0. \tag{27}$$

If $\gamma > \mu$, there is an unique positive solution to the above equation with $y^* = \lambda/\gamma$, given by:

$$x^* = \lambda \left(\frac{1}{\mu} - \frac{1}{\gamma}\right)$$

and in this case the equilibrium in (26a) yields:

$$F^*(\sigma) = x^* \bar{H}(\sigma)$$

where $\bar{H}$ is the residual lifetime distribution associated to H. Note also that the equilibrium rate satisfies $r^* = \lambda/x^*$, which yields the following expression for the mean download time:

$$\tau = \frac{1}{r^*} = \frac{1}{\mu} - \frac{1}{\gamma}. \tag{28}$$

The condition $\gamma > \mu$ (or $\mu^{-1} > \gamma^{-1}$) can be interpreted as follows: the mean time γ^{-1} a peer spends as a seeder is insufficient to upload a full copy of the file (which takes μ^{-1}); the extra time τ of upload must be provided when the peer is a leecher; so we are again in a *globally sustained* case as in Section III. If instead $\gamma \leqslant \mu$ no equilibrium with positive x^* exists, consistent with the fact that in this case, the average number of seeders can cope with the load, and the fluid model converges to $F^* \equiv 0$.

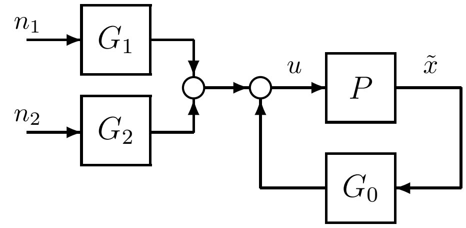

Fig. 8. Linearized dynamics with noise.

In analogous fashion to Section III, we wish to analyze the local stability of the globally sustained case. Once again, we use incremental variables $F = F^* + f$, $x = x^* + \delta x$, $z = z^* + \delta z$, and carry out similar operations as those leading to (7), yielding the linearization:

$$\frac{\partial f}{\partial t} = \frac{1}{\tau}\frac{\partial f}{\partial \sigma} + \frac{1}{\tau} H(\sigma) \underbrace{(-\mu\tau\delta z + \delta x + \tau n_1)}_{u}, \tag{29}$$

where n_1 is a perturbation introduced in the arrival rate. Since there is randomness in the seeder departures, it makes sense to introduce a second noise source n_2 on the right-hand side of (26b). When translated to the total population variable as in (26c), we end up with an incremental equation that involves both noise sources:

$$\dot{\delta z} = -\gamma \delta z + \gamma \delta x + n_1 + n_2. \tag{30}$$

As before we can express the dynamics as a feedback loop. One of the components is the same infinite dimensional system P of (8), reproduced here for convenience:

$$P: \begin{cases} \dfrac{\partial f}{\partial t} = \dfrac{1}{\tau}\dfrac{\partial f}{\partial \sigma} + \dfrac{1}{\tau} H(\sigma) u, \\ \delta x = f(t, 0). \end{cases}$$

As for the feedback term, in this case we have the first order dynamics

$$\begin{aligned} \dot{\delta z} &= -\gamma \delta z + \gamma \delta x + n_1 + n_2; \\ u &= -\mu\tau\delta z + \delta x + \tau n_1; \end{aligned}$$

which can also be expressed in transfer function form

$$\hat{u}(s) = \hat{G}_0(s)\hat{\delta x}(s) + \hat{G}_1(s)\hat{n}_1(s) + \hat{G}_2(s)\hat{n}_2(s);$$

$$\hat{G}_0(s) = \frac{s+\mu}{s+\gamma},\ \hat{G}_1(s) = \frac{\tau(s+\gamma-\mu)}{s+\gamma},\ \hat{G}_2(s) = -\frac{\mu\tau}{s+\gamma}.$$

The feedback dynamics in transfer function form is depicted in the block diagram of Fig. 8.

Theorem 2: The closed loop transfer function $\hat{P}(s)\hat{G}_0(s)$ from Fig. 8 has the small-gain property $\|\hat{P}\hat{G}_0\|_{\mathcal{H}_\infty} < 1$, and therefore the closed loop transfer function is stable (analytic in $\mathrm{Re}(s) \geqslant 0$).

Proof: Recall from Theorem 1 that P has transfer function $\hat{P}(s) = \int_0^\infty e^{-s\tau\eta} H(\eta) d\eta$, which satisfies $\|\hat{P}(s)\|_{\mathcal{H}^\infty} = \sup_\omega |\hat{P}(j\omega)| \leqslant 1$. Moreover, it is readily checked that this supremum is attained at frequency $\omega = 0$.

Analyzing now the transfer function $\hat{G}_0$, we see that it corresponds to a lead-lag system with the zero at μ and the pole at $\gamma > \mu$. The modulus of the frequency response of this system is thus increasing and we conclude that $\|\hat{G}_0(s)\|_{\mathcal{H}^\infty} = \sup_\omega |\hat{G}_0(j\omega)| = 1$, achieved as $\omega \to \infty$.

The feedback loop gain therefore satisfies $\|\hat{P}\hat{G}_0\|_{\mathcal{H}^\infty} < 1$, since both terms have norm $\leqslant 1$ attained at opposite ends of the spectrum. By the small-gain theorem, the feedback loop is stable. Furthermore, since $\hat{G}_1$ and $\hat{G}_2$ are both stable transfer functions, the overall relationship

$$\hat{\delta x}(s) = \frac{\hat{P}(s)}{1 - \hat{G}_0(s)\hat{P}(s)} \left[\hat{G}_1(s)\hat{n_1}(s) + \hat{G}_2(s)\hat{n_2}(s) \right]$$

is input-output stable. ■

B. Variability analysis

As before, we are interested in evaluating the variability of populations around their equilibrium values, using the above linear models and frequency domain computation.

Assume that the only sources of variability are the arrival and departure noises n_1, n_2. As in the previous Section, variability of the Poisson(λ) arrivals will be modeled by taking n_1 to be white noise of power spectral density λ. Note that locally around equilibrium, departures also follow a Poisson process of intensity $\gamma y^* = \lambda$, and thus n_2 will have the same fluid noise representation, independent of n_1.

The resulting power spectral density for the output process δx is given by

$$S_x(\omega) = \lambda \frac{|\hat{P}(j\omega)|^2 \cdot (|\hat{G}_1(j\omega)|^2 + |\hat{G}_2(j\omega)|^2)}{|1 - \hat{P}(j\omega)\hat{G}_0(j\omega)|^2}. \quad (31)$$

Integrating over frequency we have the population variance $E[\delta x^2] = \int_{-\infty}^{\infty} S_x(\omega) \frac{d\omega}{2\pi}$. Similar calculations can be used to the evaluate the variance around equilibrium for the number of seeders $E[\delta y^2]$.

In Fig. 9 we plot the power spectral density in (31) for two cases: the solid curve corresponds to deterministic job sizes, where we have

$$\hat{P}(s) = \frac{1 - e^{-\tau s}}{\tau s},$$

and the dashed curve is for exponential job sizes, where $\hat{P}(s) = \frac{1}{\tau s + 1}$. Note that in this case the model becomes finite-dimensional since it can be reduced to an ODE through separation of variables. The parameters are $\gamma = 4\mu$ in normalized units ($\mu = 1$). Both models coincide at low frequency, and the infinite dimensional model predictably adds many high order modes. Around the cutoff frequency, with most impact on the variance, both models differ substantially, the PDE model predicting more variability.

When job sizes are deterministic, the only sources of variability are clearly the arrival and departure processes, in particular the transition from leechers to seeders is endogenously determined by the transport equation. In this case we expect the above frequency domain integral to provide a good estimate of variance; below, we will validate its accuracy in simulation. For job-sizes of general random distribution, this additional variability must be accounted for; in particular it does not suffice to express arrival noise as variations in the Poisson rate affecting the mean distribution $H(\sigma)$. Dealing properly with this case remains open for future research.

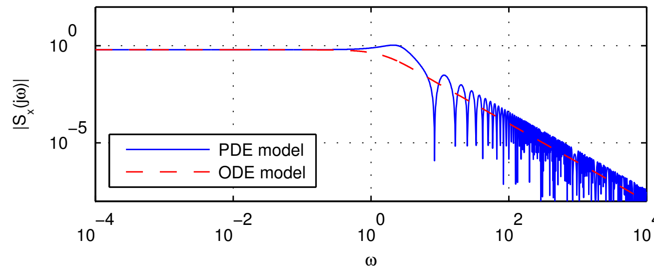

Fig. 9. Bode plot of the power spectral density S_x; deterministic and exponential case.

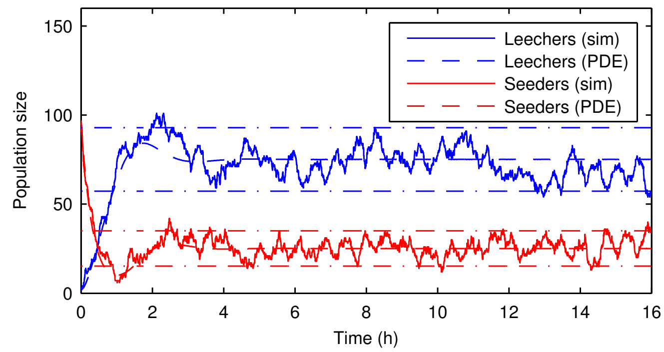

Fig. 10. Evolution of the number of leechers and seeders in a globally sustained P2P system.

C. Simulations with variable seeders

We now simulate the scenario where peers that finish download remain in the system as seeders, departing after an exponentially distributed time with parameter γ. The system model is then (26), and we choose $\gamma = 4\mu$, so seeders depart on average after 15 min. Since $\gamma > \mu$, the system is globally sustained, and the equilibrium values for the population are $x^* = 75$, $y^* = 25$. We simulate 16 hs. of system evolution to capture transient and steady-state behavior. At time $t = 0$, the system starts with no leechers and 100 initial seeders.

Results are shown in Fig. 10, for both leecher and seeder populations. Again, the fluid model correctly tracks the transient and steady-state average behavior. To validate our predictions of variability, we numerically computed the frequency integrals for the variance of x and y, with noise terms of power λ. Using these variances, we plot the corresponding 95% confidence intervals for the system population around equilibrium, showing good agreement with the simulation.

VI. Conclusions

In this paper, we analyzed PDE models for population evolution in a P2P file exchange network, capturing download progress when the file of interest is of general distribution. When the resource allocation policy is processor-sharing, we derive results for local stability and variability around the equilibrium in the cases of fixed and variable number of seeders, and perform transient analysis. Simulations show good agreement of the model with the behavior of BitTorrent under homogeneous access parameters. We also extended some of the results to the case of heterogeneous populations under a suitable model of bandwidth allocation. Further analysis of such generalized resource allocation remains open for future work.

References

[1] C. Aperjis, R. Johari, and M. J. Freedman, "Bilateral and Multilateral Exchanges for Peer-Assisted Content Distribution," *IEEE/ACM Trans. on Networking*, vol. 19, no. 5, pp. 1290–1303, 2011.

[2] A. Chow, L. Golubchik, and V. Misra, "BitTorrent: An extensible heterogeneous model," in *Proc. of IEEE INFOCOM*, 2009, pp. 585–593.

[3] F. Clévenot, P. Nain, and K. Ross, "Multiclass P2P Networks: Static Resource Allocation for Service Differentiation and Bandwidth Diversity," in *Proc. of IFIP Performance, Juan-les-Pins, France*, 2005.

[4] B. Cohen, "Incentives build robustness in BitTorrent," in *Proc. 1st Workshop on the Econ. of Peer-2-Peer Systems, Berkeley, CA*, 2003.

[5] E. de Souza e Silva, R. M. Leao, D. S. Menasché, and A. Rocha, "On the interplay between content popularity and performance in P2P systems," in *Quantitative Evaluation of Systems*. Springer, 2013, pp. 3–21.

[6] K. Eger, T. Hossfeld, A. Binzenhofer, and G. Kunzmann, "Efficient simulation of large scale P2P networks: packet-level vs. flow-level simulations," in *Proc. of UPGRADE-CN '07, Monterey Bay, USA*, 2007.

[7] B. Fan, D.-M. Chiu, and J. C. S. Lui, "Stochastic differential equation approach to model BitTorrent-like P2P systems," in *Proc. of ICC 2006*, 2006, pp. 915–920.

[8] B. Fan, J. C. Lui, and D.-M. Chiu, "The Design Trade-Offs of BitTorrent-Like File Sharing Protocols," *IEEE/ACM Trans. on Networking*, vol. 17, no. 2, pp. 365–376, 2009.

[9] A. Ferragut, F. Kozynski, and F. Paganini, "Dynamics of content propagation in BitTorrent-like P2P file exchange systems," in *Proc. of 50th IEEE Conf. on Decision and Control, Orlando, FL*, 2011.

[10] A. Ferragut and F. Paganini, "Content dynamics in P2P networks from queueing and fluid perspectives," in *Proc. of 24th International Teletraffic Congress, Krakow, Poland*, 2012.

[11] B. Hajek and J. Zhu, "The missing piece syndrome in peer-to-peer communication," *Stochastic Systems*, vol. 1, no. 2, pp. 246–273, 2011.

[12] R. Jain, D.-M. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer system," *DEC Research Report TR-301.*, 1984.

[13] G. Kesidis, T. Konstantopoulos, and P. Sousi, "Modeling file-sharing with BitTorrent-like incentives," in *Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Honolulu, HI*, 2007.

[14] A. Legout, N. Liogkas, E. Kohler, and L. Zhang, "Clustering and sharing incentives in BitTorrent systems," in *ACM SIGMETRICS Performance Evaluation Review*, vol. 35, no. 1, 2007, pp. 301–312.

[15] K. Leibnitz, T. Hossfeld, N. Wakamiya, and M. Murata, "Modeling of Epidemic Diffusion in Peer-to-Peer File-Sharing Networks," *Lecture Notes on Computer Science*, vol. 3853, pp. 322–329, 2006.

[16] L. Leskela, P. Robert, and F. Simatos, "Interacting branching processes and linear file-sharing networks," *Adv. in Applied Probability*, vol. 42, no. 3, pp. 834–854, 2010.

[17] D. Levin, K. LaCurts, N. Spring, and B. Bhattacharjee, "BitTorrent is an Auction: Analyzing and Improving BitTorrent's Incentives," *ACM SIGCOMM Computer Comm. Review*, vol. 38, pp. 243–254, 2008.

[18] L. Massoulié and M. Vojnović, "Coupon replication systems," *IEEE/ ACM Transactions on Networking*, vol. 16, no. 3, pp. 603–616, 2008.

[19] D. Menasché, A. Rocha, E. de Souza, R. Leao, D. Towsley, and A. Venkataramani, "Estimating self-sustainability in peer-to-peer swarming systems," *Performance Evaluation*, vol. 67, no. 11, pp. 1243–1258, 2010.

[20] D. Menasché, A. Rocha, B. Li, D. Towsley, and A. Venkataramani, "Content availability and bundling in swarming systems," in *Proc. of the 5th Intl. Conf. on Emerging Networking Experiments and Technologies*, 2009, pp. 121–132.

[21] R. Núnez-Queija and B. Prabhu, "Scaling laws for file dissemination in P2P networks with random contacts," in *Proc. of the 16th International Workshop on Quality of Service (IWQoS) 2008.* IEEE, 2008, pp. 75–79.

[22] F. Paganini and A. Ferragut, "PDE models for population and residual work applied to peer-to-peer networks," in *Proc. of 46th Annual Conference on Information Sciences and Systems, Princeton, NJ*, 2012.

[23] F. Paganini, K. Tang, A. Ferragut, and L. Andrew, "Stability of networks under general file size distribution and alpha fair bandwidth allocation," *IEEE Trans. on Automatic Control*, vol. 57, no. 3, pp. 579–591, 2012.

[24] D. Qiu and R. Srikant, "Modeling and performance analysis of BitTorrent-like peer-to-peer networks," *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 4, pp. 367–378, 2004.

[25] P. Robert, *Stochastic Networks and Queues*, ser. Stochastic Modelling and Applied Probability Series. New York: Springer-Verlag, 2003.

[26] S. Wong, E. Altman, and M. Ibrahim, "P2P networks: The interplay between legislation and information technology," *INRIA Research Report RR-6889*, 2009.

[27] F. Wu and L. Zhang, "Proportional response dynamics leads to market equilibrium," in *Proc. of the 39th Annual ACM Symposium on Theory of Computing, San Diego, CA*, 2007.

[28] X. Yang and G. De Veciana, "Service capacity of peer-to-peer networks," in *Proc. of IEEE Infocom, Hong Kong*, 2004.

[29] ——, "Fairness, incentives and performance in peer-to-peer networks," *Performance Evaluation*, vol. 63, no. 3, pp. 175–194, 2006.

[30] B. Q. Zhao, J. C. Lui, and D. Chiu, "A mathematical framework for analyzing adaptive incentive protocols in p2p networks," *IEEE/ACM Trans. on Networking*, vol. 20, no. 2, pp. 367–380, 2012.

[31] K. Zhou, J. C. Doyle, and K. Glover, *Robust and optimal control*. NJ: Prentice-Hall, 1996.

[32] J. Zhu and B. Hajek, "Stability of a peer-to-peer communication system," *IEEE Trans. on Information Theory*, vol. 58, no. 7, pp. 4693–4713, 2012.

Andres Ferragut obtained his PhD degree in Electrical Engineering (Telecommunications) from Universidad de la Republica, Uruguay (2011). He has held teaching and research positions at the Mathematics and Electrical Engineering Depts. at Univ. de la Republica (2000 to 2009) and since 2007 he is affiliated with the Mathematics Applied to Telecommunications and Energy (MATE) research group, Universidad ORT Uruguay. He also spent two years (2004-2006) of graduate work at Télécom Paris, France. In 2012, Dr. Ferragut was awarded the annual Uruguayan Engineering Academy prize for the best doctoral thesis in Electrical Engineering. His research interests are stochastic processes and queueing theory applied to the mathematical modeling of networks.

Fernando Paganini (M'90–SM'05–F'14) received his Ingeniero Electricista and Licenciado en Matematica degrees from Universidad de la Republica, Montevideo, Uruguay, in 1990, and his M.S. and PhD degrees in Electrical Engineering from the California Institute of Technology, Pasadena, in 1992 and 1996 respectively. His PhD thesis received the 1996 Wilts Prize and the 1996 Clauser Prize at Caltech. From 1996 to 1997 he was a postdoctoral associate at MIT. Between 1997 and 2005 he was on the faculty the Electrical Engineering Department at UCLA, reaching the rank of Associate Professor. Since 2005 he is Professor of Electrical and Telecommunications Engineering at Universidad ORT Uruguay. Dr. Paganini has received the 1995 O. Hugo Schuck Best Paper Award, the 1999 NSF CAREER Award, the 1999 Packard Fellowship, the 2004 George S. Axelby Best Paper Award, and the 2010 Elsevier Scopus Prize. He is a member of the Uruguayan National Academy of Sciences and a Fellow of the IEEE. His research interests are control and networks.