

# Large deviations of the stationary measure of networks under proportional fair allocations

M. Jonckheere, S. López

July 26, 2012

## Abstract

We address a conjecture introduced by Massoulié (2007), concerning the large deviations of the stationary measure of bandwidth-sharing networks functioning under the Proportional fair allocation. For Markovian networks, we prove that Proportional fair and an associated reversible allocation are geometrically ergodic and have the same large deviations characteristics using Lyapunov functions and martingale arguments. For monotone networks, we give a more direct proof of the same result relying on stochastic comparisons that hold for general service requirement distribution. These results comfort the intuition that Proportional fairness is ‘close’ to allocations of service being insensitive to the service time requirement.

## 1 Introduction

Bandwidth-sharing networks describe the evolution of the number of flows (or calls) in a communication network where different classes of traffic compete for the bandwidth.

They have become a standard modeling tool over the past decades for modeling communication networks [4, 11] and have been used in particular to represent the flow level dynamics of a wide range of wireline and wireless networks [2], generalizing henceforth more traditional voice traffic models [7].

In queuing theory, these models boil down to a particular class of processor sharing networks with state-dependent service rates. Assuming that class- $i$  flows arrive subject to a Poisson process of intensity  $\lambda_i$  and require exponentially distributed service times of mean  $\mu_i$  (the arrival processes of

all classes being mutually independent), and in the absence of internal routing, the stochastic process  $X = (X_1, \dots, X_N)$  describing the number of flows (or calls) in progress in the network is a multi-dimensional birth and death process with transition rates:

$$\begin{aligned} q(x, x - e_i) &= \mu_i \phi_i(x), \\ q(x, x + e_i) &= \lambda_i, \end{aligned}$$

where  $x = (x_1, \dots, x_N)$  is the number of flows in each class of traffic. The service rates of the  $N$  traffic classes  $\phi = (\phi_i(\cdot))_{i=1}^N$  encodes the particularities of the network resulting from the specific topology, technology, radio conditions, interference and multi-diversity effects and the packet protocols and congestion control mechanisms in use. In a wireline network, the vector  $\phi(x)$  is usually assumed to belong for all  $x$  to a polyhedron describing the capacities constraints of each link that are used by the different routes while for wireless networks,  $\phi(x)$  generally belongs to a more complicated closed convex set containing 0 corresponding to the achievable rates. In many situations like TDMA or CDMA data networks, the capacity set can be assumed to be convex but this is not necessarily the case for decentralized schemes like for the 802.11 WLAN. However, it has been shown that the capacity sets in that case are log-convex (see [9]).

Some specific bandwidth allocations have received a lot of attention in recent years. These include the max-min fair allocation and the proportional fair allocation (PF) that maximize a log-utility function. More generally, Mo and Walrand introduce the following family of utility functions:

$$U^\alpha(x, \eta) = \sum_{i=1}^N x_i \frac{\eta_i^{1-\alpha}}{1-\alpha}, \quad \alpha \in (1, \infty),$$

$$U^1(x, \eta) = \sum_{i=1}^N x_i \log(\eta_i), \quad \alpha = 1,$$

including as special cases the proportional fair allocation and the max-min fair allocation for  $\alpha$  tending  $\infty$ . On the other hand, the balanced fairness allocation (BF) was defined in [3] as the allocation ensuring the reversibility of the Markov process  $X$  and maximizing the probability of the network to be empty (among the ‘reversible’ allocations).

More generally, the selection of a specific bandwidth allocation is motivated by several properties of the resulting process  $X$ . Among those properties, the following ones are of particular interest:

1. *maximal stability*: one expects the bandwidth sharing mechanism to stabilize the system whenever it can be stabilized,
2. *decentralized protocol*: the bandwidth allocation can be implemented in the network using decentralized schemes,
3. *robustness*: when changing the traffic conditions<sup>1</sup> and in particular the service time distribution (but keeping the mean flow size fixed), one could expect the stationary measure of  $X$  to remain the same, in which case the system and the bandwidth allocation are said to be insensitive,
4. *relaxed robustness* : the large deviations characteristics do not depend on the service time distribution, except for its mean and coincide with the large deviations of the most efficient insensitive allocation.

The PF allocation satisfies properties (1) and (2) but fails to satisfy property (3) on general topologies. On the other hand, the BF allocation satisfies (1) and (3) but it is not known whether it satisfies (2) [2]. On the one hand, it was shown in [10] that an appropriate modification of the proportional fair allocation, called modified proportional fair allocation, (mPF) coinciding asymptotically (point-wise) with PF, has the same large deviations characteristics as BF. On the other hand, it has been recently proven that an insensitive allocation being maximal stable, is asymptotically equivalent (point-wise) to PF [15]. However, it remained an open problem to prove that the large deviation characteristics of the stationary measure of the PF allocation itself coincide with those of mPF, and BF as it was conjectured in [10].

Our contribution is the following. We first recall the principles of the Freidlin and Wentzell theory for birth and death processes on  $\mathbb{Z}^N$  with rates being Lipschitz and with bounded logarithms. As underlined for instance in [14], it is very demanding to extend these results to processes on state spaces with boundaries since the technical conditions of the classical theory

---

<sup>1</sup>When the size distributions of the flows are not exponentially distributed, the process  $X$  is not Markov by itself anymore and the dynamics have to be defined using the residual service time.

are never fulfilled. (Remark that boundness of the logarithm is never verified for birth and death processes living in the orthant, while the Lipschitz assumption is not verified in our case for several network topologies).

A main contribution of the present article consists in overcoming these difficulties for the specific processes we are studying. We first show that for Poisson arrivals and exponentially distributed flow sizes, the stationary distribution  $\pi^{PF}$  of the number of flows associated with PF and the stationary distribution  $\pi^{mPF}$  associated with the mPF allocation have the same large deviations characteristics. More precisely:

**Theorem 1.** *For all  $x \in \mathbb{Z}_+^N$ :*

$$\frac{1}{n} \log \left( \frac{\pi^{PF}(nx)}{\pi^{mPF}(nx)} \right) \leq O(n^{-\frac{1}{2}+\epsilon}), \quad \forall \epsilon > 0.$$

*In the particular case that  $x \in \mathbb{N}^N$  the bound can be improved to  $O(n^{-1})$ .*

This is achieved by first proving the geometric ergodicity of both the mPF allocation and the PF allocation. For that purpose, we exhibit appropriate Lyapunov functions, relying on some structural results of PF described in [10]. This then allows us to use simple martingale arguments.

Finally, for monotone networks and generally distributed flow size, we give a more direct proof establishing that the large deviations characteristics are actually insensitive to the service time distribution. This shows that the proportional fair allocations indeed satisfy properties (1) and (2) and (4) at least on monotone topologies.

The rest of the paper is organized as follows. In Section 2, we introduce the allocations functions used in the sequel as well as several of their properties. In Section 3, we show that the rate function of the proportional fair allocation coincides for Markovian dynamics to the rate function of Balanced fairness. In Section 4, we prove property (4) for monotone networks using stochastic comparisons.

## 2 Properties of bandwidth sharing allocations

**Notations** We define here a few notations that we need in the sequel.  $\mathbb{R}_+$  denotes the set of non-negative real numbers. The bandwidth allocation vector is denoted  $\phi(x) \equiv (\phi_i(x))_{i=1\dots N}$ . For any vector  $v$  in  $\mathbb{R}^N$ , and function

$f : \mathbb{R} \rightarrow \mathbb{R}$ , we denote  $f(v)$  the vector  $(f(v_i))_{i=1\dots N}$ . Similarly, we also use the notation  $a^v$  for  $\prod_{i=1}^N a_i^{v_i}$ . The usual scalar product in  $\mathbb{R}^N$  between  $u$  and  $v$  is denoted  $\langle u, v \rangle$ . We use  $\|\cdot\|_p$  to denote the  $l_p$ -norm, but we reserve  $|\cdot|$  to denote the  $l_1$ -norm:  $|v| = \sum_{i=1}^N |v_i|$ . For  $x, y \in \mathbb{R}^N$ , we also use the notation  $x \leq y$  to denote the partial order  $x_i \leq y_i$  for all  $i = 1 \dots N$ . Finally, given a set  $\mathcal{C}$  and function  $f : \mathbb{R} \rightarrow \mathbb{R}$ , we shall denote  $f(\mathcal{C})$  the set  $\{f(\eta) : \eta \in \mathcal{C}\}$ .

## 2.1 Gradient allocations

In the characterization of the large deviations of the stationary regime, we shall rely on two properties of the allocation function  $\phi$  playing a crucial role and being intrinsically related: being a gradient allocation or a discrete gradient allocation.

**Definition 1.** *A bandwidth allocation is called gradient if there exists a function  $P : \mathbb{R}^N \rightarrow \mathbb{R}$  (that we call a continuous potential) such that*

$$\log(\phi(x)) = -\nabla P(x), \quad \forall x \in \mathbb{R}_+^N \setminus \{0\}.$$

Recall that the proportional fair allocation is defined by the optimization problem associated to a capacity set  $\mathcal{C}$  as follows:

$$\phi^{PF}(x) = \arg \max_{\eta \in \mathcal{C}} U^1(\eta, x) = \arg \max_{\eta \in \mathcal{C}} \langle x, \log(\eta) \rangle.$$

Following Massoulié [10], observe that the proportional fair bandwidth allocation is gradient. Indeed, let  $\delta_{\mathcal{A}}^*$  the support function of a bounded convex set  $\mathcal{A}$  i.e.:

$$\delta_{\mathcal{A}}^*(x) = \max_{\eta \in \mathcal{A}} \langle x, \eta \rangle.$$

**Proposition 1.** *Assume that the set  $\mathcal{C}$  is log-convex, i.e., the set  $\log(\mathcal{C})$  is convex, then*

$$\log(\phi^{PF}(x)) = \nabla \delta_{\log(\mathcal{C})}^*(x), \quad \forall x \in \mathbb{R}_+^N \setminus \{0\}.$$

**Proof.**

The function  $\delta^*$  is sub-differentiable because it is convex and finite (see [13]) for all  $x \in \mathbb{R}_*^N$ . The unicity of the sub-gradient comes from the strict concavity of the log function and implies the differentiability. □

In the sequel, we always assume that  $\mathcal{C}$  is convex and contains the set  $\{\eta : \sum \eta_i \leq c\}$  for some  $c > 0$ . Of course, this is not a restriction for applications. We denote by  $P^{PF} \equiv \delta_{\log(\mathcal{C})}^*$ .

## 2.2 Reversible allocations

**Definition 2.** A bandwidth allocation is called reversible or discrete gradient, if there exists a function  $\tilde{P} : \mathbb{R}^N \rightarrow \mathbb{R}$  (that we call a discrete potential) such that for all  $x \in \mathbb{N}^N$ ,  $i = 1, \dots, N$ :

$$\log(\phi_i(x)) = -\mathbf{D}\tilde{P}(x) \equiv \tilde{P}(x) - \tilde{P}(x - e_i).$$

In the latter case, the stationary measure of the process is easily described:

**Proposition 2.** The process  $\tilde{X}$  associated with a reversible allocation with discrete potential  $\tilde{P}$  is reversible (in the usual sense) and its stationary measure (when it exists) is:

$$\pi(x) = C\lambda^x \exp(\tilde{P}(x)).$$

Assume further that  $\frac{1}{n}\tilde{P}(nx) \rightarrow \gamma(x)$  as  $n \rightarrow \infty$ . Then:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \pi(nx) = -(\gamma(x) - \sum_{i=1}^N x_i \log(\lambda_i)).$$

A particular role in what follows shall be played by two reversible allocations ‘close’ to PF: the modified PF allocation (mPF) and the BF allocation which we now define mathematically.

**Definition 3** (Modified proportional fair and Balanced fair allocations).

$$\log(\phi^{mPF}(x)) = -\mathbf{D}\tilde{P}^{PF}(x),$$

with  $P^{PF}$  the continuous potential associated with PF, and

$$\log(\phi^{BF}(x)) = -\mathbf{D}\tilde{P}(x),$$

where the potential  $\tilde{P}$  is recursively defined by:

$$\tilde{P}(x) = 0,$$

$$\tilde{P}(x) = \max\{a > 0 : \tilde{P}(x - e_i) - a \in \log(C)\}.$$

We have the following large deviations results for the BF allocation:

**Proposition 3.** (Massoulié, 2007)

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \pi^{BF}(nx) = P^{mPF}(x).$$

The main challenge of the subsequent analysis is to provide similar results for the Proportional fair allocation itself and hence verify that both allocations share the same large deviations characteristics. The main difficulty of this program is that the stationary measure of a network under the proportional fair allocation does not have a closed form in general (except in the very particular cases of symmetric hypergrids).

### 2.3 Stability

We state two stability results, one for the original stochastic system, and one for the deterministic analogue. Both results are well known for Markovian bandwidth sharing networks in [1]. Massoulié extended the stability results to general service time distributions in [10].

**Theorem 2** ([10]). *The stability set of the process  $X$  associated with the PF allocation is contained in, and contains the interior of the set  $\mathcal{S} = \bar{\mathcal{C}}$ .*

**Remark 2.1.** *For Markov processes, a direct proof follows from taking  $x \mapsto \sum_{i=1}^N \frac{x_i^2}{\lambda_i}$  as a Lyapunov function.*

Define now the deterministic dynamical system  $(DS)_{x_0}$  by

$$\dot{x}_i = (\lambda_i - \phi_i(x))1_{x_i > 0}, \quad (1)$$

$$x(0) = x_0. \quad (2)$$

It was proved [8] (see also [10]) that the fluid limits of the process  $X$  are solutions of this dynamical system at all regular points.

### 2.4 Examples

We illustrate the above description on a few toy examples.

**Example 1** (Single link). *Consider first a single link shared by  $N$  classes. This corresponds to choosing  $\mathcal{C} = \{\sum \frac{n_i}{\mu_i} \leq 1\}$ , where the  $\{\mu_i\}$  correspond to the mean flow sizes. In this case, the proportional fair allocation coincides with the balanced fairness allocation. Its service rates are given by:*

$$\phi_i(x) = \mu_i \frac{x_i}{|x|}.$$

*Note however that the continuous and the discrete potential functions associated with the allocation  $\phi$  do not coincide:*

$$\delta_{\log(\mathcal{C})}^*(x) = \sum_{i=1}^N \frac{x_i}{|x|} \log \left( \frac{x_i}{|x|} \right),$$

while the discrete potential is given by:

$$P(x) = \log \left( \binom{|x|}{x_1, \dots, x_N} \right).$$

One can however prove using the Stirling formula that

$$\frac{1}{n}(P(nx) - \delta_{\log(c)}^*(nx)) \rightarrow 0, \text{ as } n \rightarrow \infty.$$

An extension of these formulae can be obtained for hypergrids topologies [4].

**Example 2** (Tree network). A 2-level tree network with  $N$  routes is defined by the following polyhedron for the capacity set:

$$\mathcal{P} = \{ \eta \in \mathbb{R}_+^N : \eta_i \leq c_i, \sum_i \eta_i \leq 1 \},$$

for some constants  $\{c_i\}$ . An expression of the stationary measure for the balanced fairness allocation can be found in [4]. In general (i.e., when the tree is not degenerate), the proportional fair allocation does not coincide with the balanced fair allocation and does not raise a reversible process. Hence its stationary distribution can be obtained only numerically. For the case  $N = 2$ , and  $\max(c_1, c_2) \leq 1$ , the proportional fair allocation is:

$$\begin{aligned} \phi_1(x) &= 1 - \phi_2(x), \\ \phi_2(x) &= \min(c_2, \frac{x_2}{x_1 + x_2}). \end{aligned}$$

An important property of the proportional fair allocation in that case is its monotonicity (see Section 4 for a definition) [5]:

$$\frac{\phi_i^{PF}(x)}{x_i} \geq \frac{\phi_i^{PF}(y)}{y_i}, \quad x_j \leq y_j, \quad \forall j.$$

**Example 3** (Wireless network). In [9], the rate region  $\mathcal{C}$  of a 2-station network functioning under the 802.11e protocol is studied. The rate region is the set of achievable throughput vectors at the fine time scale (packet level). Their findings show that the rate region, (which exact expression depends in a complicated manner of the probabilities of transmitting) is generally not convex but is however log-convex.

### 3 Large deviations for Markovian dynamics

#### 3.1 The Freidlin-Wentzell theory for birth and death processes with smooth transitions

In this Section, we remind the concepts introduced by Freidlin and Wentzell [6] allowing to get a grasp on the sample-paths large deviations of stochastic differential equations with small noise. These results were often applied to diffusions processes but can equally be applied (as indicated in [6]) to some jump processes. We show here how these classical results can be used to prove a large deviation principle for birth and death processes with “gradient” rates on  $\mathbb{Z}^N$ .

Let  $Y^n$  be a multi-dimensional birth and death process with transition rates:

$$\begin{aligned} q\left(\frac{x}{n}, \frac{x}{n} - \frac{e_i}{n}\right) &= n\phi_i\left(\frac{x}{n}\right), \\ q\left(\frac{x}{n}, \frac{x}{n} + \frac{e_i}{n}\right) &= n\lambda_i. \end{aligned}$$

We suppose additionally that the death rates are 0-homogeneous (i.e.  $\phi(az) = \phi(z)$ ,  $\forall a > 0$ ).

Define the logarithmic moment-generating of the increment of the process for  $z \in \mathbb{R}^N/n, y \in \mathbb{R}^N$  by

$$H_n(z, y) = \frac{d}{dt} E^z[\exp(\langle y, Y^n(t) - z \rangle)]. \quad (3)$$

Using the structure of the generator and the 0-homogeneity of the rates:

$$H_n(z, y) = n \left( \sum_i \lambda_i (e^{y_i/n} - 1) + \phi_i(z) (e^{-y_i/n} - 1) \right). \quad (4)$$

We thus obtain that  $H_n(z, y/n) = nH(z, y/n)$  with

$$H(z, y) = \langle e^y - 1, \lambda \rangle + \langle e^{-y} - 1, \phi(z) \rangle. \quad (5)$$

Define now the Fenchel-Legendre  $L$  transform of  $H$ :

$$L(x, y) = \sup_{\theta \in \mathbb{R}^N} \langle y, \theta \rangle - H(x, \theta), \quad (6)$$

and define the action functional  $S : \mathcal{C}([0, T]) \rightarrow \mathbb{R}$  as follows:

$$S_T(r) = \int_0^T L(r_s, \dot{r}_s) ds. \quad (7)$$

Now define the quasi-potential  $V$  by:

$$V(x) = \min_{r, T, r(0)=0, r(T)=x} S_T(r, \dot{r}). \quad (8)$$

Assume that the function  $L$  is such that:

$$\sup_{|x-\tilde{x}|} \frac{|L(x, y) - L(\tilde{x}, y)|}{1 + L(x, y)} \rightarrow 0.$$

The original results of Freidlin and Wentzell are stated for Lipschitz-continuous transitions (extended to  $\mathbb{R}^N$ ):

**Theorem 3** (Freidlin and Wentzell). *Assume that for each  $i$ , the function  $\log \phi_i(\cdot)$  is bounded and Lipschitz continuous. Let  $\pi^n$  the stationary measure of the birth and death processes  $X^n$ . Then, a large deviations principle holds for the family of probabilities  $\pi^n(\cdot)$  with rate function  $V(\cdot)$ .*

### 3.1.1 Further characterization of the potential

In the case of general multi-dimensional birth and death processes, it is difficult to solve the variational problem from which the potential  $V$  is defined. However, assuming that the rates are gradient, as defined in the previous Section, greatly simplifies the expression of the potential.

**Proposition 4.** *Assume that the rates (allocation)  $\phi$  is gradient with continuous potential  $P$ . Then the quasi-potential of the birth and death process is equal to  $R(x) = -\langle \log(\lambda), x \rangle + P(x)$ .*

**Proof.**

We first show that:

$$V(x) = \max_{x_s: x_0=0, x_T=x} \int_0^T \langle \dot{x}_s, \log \left( \frac{\lambda}{\phi(x_s)} \right) \rangle ds.$$

The first simplification comes from the time homogeneity of the process. Using the Euler-Lagrange principle, combined with the time homogeneity (which implies that  $\nabla_x L = 0$ ) the minimizing path satisfies the Beltrami identity:

$$L - \langle y, \nabla_y L \rangle = 0.$$

Now, observe that if  $\theta_y = \arg \max \langle y, \theta \rangle - H(x, \theta)$ , we obtain that  $\nabla_y L = \theta_y$  and  $L - \langle y, \nabla_y L \rangle = H(x, \theta_y) = 0$ . Solving the last equation leads to:

$$\theta_x = \frac{1}{2} \log\left(\frac{\phi(x)}{\lambda}\right),$$

which allows us to conclude since

$$L(x, \dot{x}) = \langle \dot{x}, \nabla_y L(x, \dot{x}) \rangle = \langle \dot{x}, \theta_x \rangle.$$

The expression for the quasi-potential then follows from integration.  $\square$

We insist on the facts that this theory allows to prove the large deviation rate of the stationary measure of birth and death processes only for smooth allocations and that it becomes a very technical issue to weaken this assumption. Also it does not provide a rate of convergence in general.

We address these issues in the next Section: we prove directly the large deviation principle using martingale arguments by taking advantage of the geometric ergodicity of the process.

### 3.2 Large-deviations results for general gradient allocations

In this Section, we call  $X$  the process associated with the PF allocation, denoted  $\phi$  and corresponding to the **continuous gradient** of the potential  $P^{PF}$ . We further denote by  $\tilde{X}$  the process associated to the allocation  $\tilde{\phi}$  itself corresponding to the **discrete gradient** of  $P^{PF}$ , i.e.:

$$\log(\phi(x)) = \nabla P^{PF}(x), \quad (9)$$

$$\log(\tilde{\phi}(x)) = -\mathbf{D}P^{PF}(x) = (P^{PF}(x) - P^{PF}(x - e_i))_{i=1\dots N}. \quad (10)$$

#### Structural properties of the PF allocation

From the structural representation of PF, we know that  $\tilde{\phi}$  is a perturbation of the original rates  $\phi$  in the sense that (see Lemma 9 in [10]):

$$\left| \log(\phi_i(x)) - \log(\tilde{\phi}_i(x)) \right| = \left| \log(\phi_i(x)) - (P^{PF}(x) - P^{PF}(x - e_i)) \right| \leq \frac{1}{x_i}, \quad \forall i, x. \quad (11)$$

## Geometric ergodicity

We first consider the dynamics corresponding to the Proportional fairness allocation.

**Proposition 5** (Geometric ergodicity of  $X$ ). *Suppose  $\lambda \in \text{int}(\mathcal{C})$ , then there exists a constant  $K$  such that for  $|x| > K$ :*

$$\Delta G(x) \leq -\gamma G(x).$$

Hence:

$$\left| P_t^0(x) - \pi^{PF}(x) \right| \leq K_1 e^{-K_2 t},$$

for some constants  $K_1, K_2 > 0$ , and  $\pi^{PF}$  the stationary distribution of  $X$ .

### Proof.

The proof has two steps. We first construct a Lyapunov function with bounded drift. We then use this Lyapunov function to construct a new one verifying a geometric drift inequality.

**First step:** let  $F(x) = (\sum_{i=1}^N \frac{x_i^2}{\lambda_i})^{1/2}$ . Observe that  $F$  is a norm in  $\mathbb{R}^N$  (hence is positive, 1-homogeneous, and diverges to infinity when  $|x| \rightarrow \infty$ ). Furthermore, it is  $C^2$  for all  $x \neq 0$  and:

$$\nabla F(x) = \frac{x_i}{\lambda_i F(x)}.$$

Hence, there exists  $K > 0$  such that for  $|x| > K$

$$\frac{1}{|x|} \sup_{z=x, x \pm e_i} \left| \frac{\partial^2 F(z)}{\partial^2 x_i} \right| \leq \epsilon.$$

Using that  $F$  is 1-homogeneous, this leads for  $|x| > K$  to:

$$\begin{aligned} \Delta F(x) &\equiv \sum_{i=1}^N \lambda_i (F(x + e_i) - F(x)) + \phi_i(x) (F(x - e_i) - F(x)), \\ &\leq \langle \lambda - \phi(x), \nabla F(x) \rangle + \epsilon, \\ &= \langle \lambda - \phi(x), \frac{x}{F(x)\lambda} \rangle + \epsilon, \\ &= \frac{|x|}{F(x)} \langle \lambda - \phi(x), \frac{x}{|x|\lambda} \rangle + \epsilon. \end{aligned}$$

Using the definition of Proportional fairness (as the allocation maximizing  $U^1(\cdot, x)$  and the strict concavity of the log function), for all  $\eta \in \mathcal{C}$ ,

$$\sum_{i=1}^N \frac{x_i}{\eta_i} (\eta_i - \phi_i(x)) = \left\langle \left( \frac{\partial U(x, \eta)}{\eta_i} \right), \eta - \phi \right\rangle < 0,$$

which implies that there exists  $\gamma$  such that:

$$\left\langle \lambda - \phi(x), \frac{x}{|x|} \frac{1}{\lambda} \right\rangle \leq -\gamma,$$

which in turn implies (together with the fact that  $F$  is norm-like) that there exists  $\tilde{\gamma}$  such that:

$$\Delta F(x) \leq -\tilde{\gamma}, \quad \forall x, |x| > K.$$

Remark also that

$$|\Delta F(x)| \leq C, \quad \forall x.$$

**Second step:** we now calculate the drift of  $G(x) = \exp(\delta F(x))$ :

$$\begin{aligned} \Delta G(x) &= \sum_{i=1}^N \lambda_i (e^{\delta F(x+e_i)} - e^{F(x)}) + \phi_i(x) (e^{\delta F(x-e_i)} - e^{\delta F(x)}), \\ &= G(x) \sum_{i=1}^N \lambda_i (e^{\delta(F(x+e_i)-F(x))} - 1) + \phi_i(x) (e^{\delta(F(x-e_i)-F(x))} - 1), \\ &\leq G(x) (\delta \Delta F(x) + C_1 \delta^2 \sum_{i=1}^N \exp(c|\Delta F(x)|) + O(\delta^3)), \\ &\leq G(x) (-C_2 \delta + C_3 \delta^2), \\ &\leq -\gamma G(x). \end{aligned}$$

This implies that (see [12]):

$$|P^0(X(t) = x) - \pi^{PF}(x)| \leq K_3 \exp^{-K_4 t}.$$

□

We now deal with the dynamics of the modified proportional fair allocation.

**Proposition 6** (Geometric ergodicity of  $\tilde{X}$ ). *Suppose  $\lambda \in \text{int}(\mathcal{C})$ , then there exists a constant  $\tilde{K}$  such that for  $|x| > \tilde{K}$ :*

$$\Delta G(x) \leq -\gamma G(x).$$

Hence,

$$\left| P_t^0(x) - \pi^{PF}(x) \right| \leq K_3 e^{-K_4 t},$$

for some constants  $K_3, K_4 > 0$ ,  $\pi^{mPF}$  being the stationary distribution of  $\tilde{X}$ .

**Proof.**

We proceed as previously. The only difference consists in proving that the perturbations of the rates are small enough to be negligible in the drift calculations. Let  $F(x) = (\sum_{i=1}^N \frac{x_i^2}{\lambda_i})^{1/2}$ . For  $|x| > \tilde{K}$ , using the bounds on the difference of  $\phi$  and  $\tilde{\phi}$ , we obtain that:

$$|\phi_i(x) - \tilde{\phi}_i(x)| \leq c_0 |\log(\phi_i(x)) - \log(\tilde{\phi}_i(x))| \leq \frac{c_0}{x_i}.$$

Hence:

$$\begin{aligned} \Delta F(x) &= \sum_{i=1}^N \lambda_i (F(x + e_i) - F(x)) + \tilde{\phi}_i(x) (F(x - e_i) - F(x)), \\ &\leq \langle \lambda - \tilde{\phi}(x), \nabla F(x) \rangle + \epsilon, \\ &\leq \langle \lambda - \phi(x), \frac{x}{F(x)\lambda} \rangle + \langle \phi(x) - \tilde{\phi}(x), \frac{x}{F(x)\lambda} \rangle + \epsilon, \\ &\leq -\tilde{\gamma} + \langle \phi(x) - \tilde{\phi}(x), \frac{x}{F(x)\lambda} \rangle, \\ &\leq -\tilde{\gamma} + \frac{c}{|x|}. \end{aligned}$$

which gives (together with the fact that  $F$  is norm-like) that there exists  $\hat{\gamma}$  such that:

$$\tilde{\Delta} F(x) \leq -\hat{\gamma}, \quad \forall x, |x| > \tilde{K}.$$

Remark also that

$$|\tilde{\Delta} F(x)| \leq C.$$

We can conclude following exactly as in the previous proof.  $\square$

Before proving the main result of this section, we establish a useful lemma using the geometric ergodicity of the processes.

**Lemma 3.1.** *There exists  $C > 0$  such that for  $t_n = Cn$ , we have*

$$\log \frac{\mathbb{P}^0(X_{t_n} \geq nx)}{\pi^{PF}(\{nx\}^\uparrow)} \leq C_1 \exp\{-C_2 n\}$$

for  $C_1, C_2 > 0$ , where  $\pi^{PF}(\{k\}^\uparrow) = \sum_{j \geq k} \pi^{PF}(j)$ . The analogue inequality is true for process  $\tilde{X}$ .

**Proof.**

We need to bound from below the stationary probabilities of the process. This is easy since the rates are bounded. Let us define a process with the same arrival rates and the service rates equal to  $\bar{\phi}$ , the maximum of  $\phi_i$  for all coordinates. It is clear that

$$\pi^{PF}(\{nx\}^\uparrow) \geq K_1 \left(\frac{\lambda}{\bar{\phi}}\right)^n,$$

for some constant  $K_1 > 0$ .

We have

$$\begin{aligned} \log \frac{\mathbb{P}^0(X_{t_n} \geq nx)}{\pi^{PF}(\{nx\}^\uparrow)} &\leq \left| 1 - \frac{\mathbb{P}^0(X_{t_n} \geq nx)}{\pi^{PF}(\{nx\}^\uparrow)} \right| \\ &= \left| \frac{\mathbb{P}^0(X_{t_n} \geq nx) - \pi^{PF}(\{nx\}^\uparrow)}{\pi^{PF}(\{nx\}^\uparrow)} \right| \\ &\leq \frac{K_2 e^{-K_3 t_n}}{\pi^{PF}(\{nx\}^\uparrow)} \\ &\leq \frac{K_2 e^{-K_3 t_n}}{K_1 \left(\frac{\lambda}{\bar{\phi}}\right)^n} = \frac{K_2}{K_1} \exp\{-K_3 t_n + \log\left(\frac{\bar{\phi}}{\lambda}\right)n\} \end{aligned}$$

where first inequality comes by  $\log(x) \leq |1 - x|$  for all  $x > 0$ , and the second one by the ergodicity of  $X$ . We take  $C > 0$  such that  $K_3 C > \log\left(\frac{\bar{\phi}}{\lambda}\right)$ . Note that the same argument functions for process  $\tilde{X}$  taking its respective transition rates and stationary distribution.  $\square$

### Change of measure and control of the martingale

To relate the distribution of  $X$  and  $\tilde{X}$ , we recall the Proposition B.6 513 of Schwartz and Weiss [14]:

**Proposition 7.** *Let  $Y_t^i$  for  $i = 1, 2$  two multidimensional birth-death process in  $\mathbb{Z}_+^N$  with step directions  $e_j$ , bounded step rates  $\{q_j^i(x)\}_{x \in \mathbb{Z}_+^N}$  and law  $P^i$ .*

Assume that for all  $x$  and  $j$ ,  $q_j^1(x) = 0$  if and only if  $q_j^2(x) = 0$ . Then we can relate the distributions of the processes by  $dP^2 = M_t dP^1$  where

$$M_t = \exp \left( \int_0^t \sum_{j=1}^N \log \left( \frac{q_j^2(X_{s-})}{q_j^1(X_{s-})} \right) dN_s^j - \int_0^t \sum_{j=1}^N (q_j^2(X_s) - q_j^1(X_s)) ds \right).$$

Here  $\{N^j\}$  denote a family of counting processes describing the jumps of the processes in the  $j$ -th direction and  $M_t$  is a càdlàg martingale.

In our case,  $\phi_i(x)$ ,  $\tilde{\phi}_i(x)$  are both positive whenever  $x_i > 0$ , and 0 otherwise. All the rates of  $X$  and  $\tilde{X}$  are bounded so we meet the conditions of the previous Proposition and we can write that  $d\tilde{P}(\omega) = M_t dP(\omega)$ , with

$$\begin{aligned} M_t &= \exp \left( \int_0^t \sum_{j=1}^N 1_{\{X_{s-} > 0\}} \log \left( \frac{\tilde{\phi}_j(X_{s-})}{\phi_j(X_{s-})} \right) dN_s^j \right. \\ &\quad \left. - \int_0^t 1_{\{X_s > 0\}} (\tilde{\phi}_j(X_s) - \phi_j(X_s)) ds \right), \\ &= \prod_{j=1}^N M_t^j, \end{aligned}$$

and each of the  $M^j$  is itself a martingale.

**Remark 3.1.** An important observation for the following is that the counting processes  $N^j$  can be seen as thinning (according to  $X_s$ ) of some Poisson processes  $\tilde{N}^j$  which are all independent.

We denote by  $M_{s,t}$  the last expression with the integral running from  $s$  to  $t$ , so  $M_t = M_{0,t}$ .

The change of measure formula is then:

$$E[1_{\{\tilde{X}_t = nx\}}] = E[M_t 1_{\{X_t = nx\}}]. \quad (12)$$

We can now prove our main result.

### Proof of Theorem 1:

We consider first the case where  $x_i > 0$  for all  $i = 1 \dots, N$ .

Define the sequence of stopping times:

$$\begin{aligned} \tau_1 &= 0, \\ \tau_i &= \inf \{t > \tau_{i-1}, \tilde{X}_t \geq nx\}, \text{ for } i \text{ even, } i \geq 2, \\ \tau_i &= \inf \{t > \tau_{i-1}, \tilde{X}_t < nx\}, \text{ for } i \text{ odd, } i \geq 3 \end{aligned}$$

Observe that if  $\tilde{X}_t \geq nx$ , then necessarily, there exists  $k$  even (a.s. finite) such that  $\tau_k \leq t \leq \tau_{k+1}$ . Using the Markov property and the martingale property, we then have

$$\begin{aligned}
E^0(M_{0,t}1_{\{\tilde{X}_t \geq nx\}}) &= E^0 \sum_{k \text{ even}} 1_{\tau_k \leq t < \tau_{k+1}} M_t 1_{\{\tilde{X}_t \geq nx\}} \\
&= \sum_{k \text{ even}} E^0 1_{\tau_k \leq t < \tau_{k+1}} M_{0,\tau_k} M_{\tau_k,t} 1_{\{\tilde{X}_t \geq nx\}}, \\
&= \sum_{k \text{ even}} E^0 (E(1_{\tau_k \leq t < \tau_{k+1}} M_{0,\tau_k} M_{\tau_k,t} 1_{\{\tilde{X}_t \geq nx\}} | \tilde{\mathcal{F}}_{\tau_k})) \\
&= \sum_{k \text{ even}} E^0 (M_{0,\tau_k} E(1_{\tau_k \leq t < \tau_{k+1}} M_{\tau_k,t} 1_{\{\tilde{X}_t \geq nx\}} | \tilde{\mathcal{X}}_{\tau_k})),
\end{aligned}$$

with  $\{\tilde{\mathcal{F}}_k\}$  being the natural filtration of process  $\tilde{X}$ .

We define

$$g_k = \sup_y E(1_{\tau_k \leq t < \tau_{k+1}} M_{\tau_k,t} 1_{\{\tilde{X}_t \geq nx\}} | \tilde{X}_{\tau_k} = y),$$

so

$$E^0(M_t 1_{\{\tilde{X}_t \geq nx\}}) \leq \sum_{k \text{ even}} g_k E^0(M_{0,\tau_k}) = \sum_{k \text{ even}} g_k.$$

Realize that on  $\{\tilde{X}_t \geq nx\} \cap \{\tau_k \leq t < \tau_{k+1}\}$  ( $k$  even) we have  $\{\tilde{X}_s \geq nx : s \in [\tau_k, t]\}$ . Recall that  $x$  is such that  $x_i > 0$  for all  $i$ . Using the assumption on  $\phi$  and  $\tilde{\phi}$ , we get that asymptotically in  $n$ :

$$\begin{aligned}
E^0 \left( 1_{\tau_k \leq t < \tau_{k+1}} 1_{\{\tilde{X}_t \geq nx\}} \exp \left( \int_{\tau_k}^t \sum_{j=1}^N 1_{\{X_{s-} > 0\}} \log \left( \frac{\tilde{\phi}_j(X_{s-})}{\phi_j(X_{s-})} \right) dN_s^j \right. \right. \\
\left. \left. - \int_{\tau_k}^t \sum_{j=1}^N 1_{\{X_{s-} > 0\}} (\tilde{\phi}_j(X_s) - \phi_j(X_s)) ds \right) \right) \\
\leq E^0 \left( 1_{\tau_k \leq t < \tau_{k+1}} \exp \left\{ \sum_{j=1}^N \frac{C_1}{n} (N_t^j + t) \right\} 1_{\{\tilde{X}_t \geq nx\}} \right), \\
\leq E^0 \left( 1_{\tau_k \leq t < \tau_{k+1}} \exp \left\{ \frac{C_1 N}{n} (\bar{N}_t + t) \right\} 1_{\{\tilde{X}_t \geq nx\}} \right)
\end{aligned}$$

where  $\bar{N}_t$  is a Poisson process with parameter  $\bar{\lambda}$  equal as the maximum of the parameters of the Poisson processes  $\{N_t^j\}_{j=1}^N$ .

Summing these inequalities and using Hölder's inequality

$$E^0(M_t 1_{\{\tilde{X}_t \geq nx\}}) \leq \exp\left\{\frac{C_1 N t}{n}\right\} \mathbb{P}^0(\tilde{X}_t \geq nx)^{\frac{p-1}{p}} \exp\{\bar{\lambda} t [e^{\frac{C_1 p N}{n}} - 1]\}^{\frac{1}{p}}.$$

We now choose  $t_n = C_2 n$  such the result of Lemma 3.1 holds, and the sequence  $p_n = n$  to obtain

$$\begin{aligned} \frac{1}{n} \log \mathbb{P}^0(X_{t_n} \geq nx) &\leq \frac{C_1 N C_2}{n} + \frac{\bar{\lambda} C_2 (\exp(C_1 N) - 1)}{n} \\ &\quad + \frac{n-1}{n^2} \log \mathbb{P}^0(\tilde{X}_{t_n} \geq nx), \\ \frac{1}{n} \log(\pi^{PF}(\{nx\}^\uparrow)) + \frac{C_3}{n} \exp(-C_4 n) &\leq O(n^{-1}) + \left(\frac{n-1}{n}\right) \frac{1}{n} \log(\pi^{mPF}(\{nx\}^\uparrow)) \\ &\quad + \frac{\tilde{C}_3(n-1)}{n^2} \exp(-\tilde{C}_4 n). \end{aligned}$$

So we have

$$\frac{1}{n} \log(\pi^{PF}(\{nx\}^\uparrow)) \leq \frac{1}{n} \log(\pi^{mPF}(\{nx\}^\uparrow)) + O(n^{-1}).$$

We also have the converse of last inequality using the same arguments interchanging  $X$  and  $\tilde{X}$ .

Assume now that  $x_i > 0$  for  $i \in \mathcal{U}$ , while  $x_i = 0$  for  $i \in \mathcal{S}$ , where  $\mathcal{U}, \mathcal{S} \subseteq \{1, \dots, N\}$ . If classes of  $\mathcal{U}$  are independent of classes  $\mathcal{S}$ , the result is obvious. Hence suppose that there exists a coupling between classes of  $\mathcal{U}$  and  $\mathcal{S}$ . We shall make use of the following fact:

**Lemma 3.2.** *Given the definition of Proportional fairness, if class  $i \in \mathcal{S}$  and class  $j \in \mathcal{U}$  are coupled, then there exist  $K > 0$  such that if  $y_i \leq n\epsilon$  and  $y_j \geq n$  for all  $j \in \mathcal{U}$ , then:*

$$|\phi_i(y) - \tilde{\phi}_i(y)| \leq K\epsilon,$$

**Proof.**

If at least two classes are coupled, there exists a Lagrange multiplier  $\alpha > 0$  and some positive constants  $c_i > 0$  and  $c_j$  (with at least one  $j$  such that  $c_j > 0$ ) such that

$$\frac{\partial U^1(y, \eta_i)}{\partial \eta_i} = \frac{x_i}{\eta_i} = \alpha c_i,$$

$$\frac{\partial U^1(y, \eta_i)}{\partial \eta_j} = \frac{x_j}{\eta_j} = \alpha c_j.$$

Combined with the fact that there exists  $c > 0$  such that  $0 < c \leq \sum_{j=1}^N \eta_j \leq C$ , we obtain that

$$\tilde{c} \frac{y_i}{y_i + \sum_{j \in \mathcal{U}_i} y_j} \leq \phi_i(y) \leq \tilde{C} \frac{y_i}{y_i + \sum_{j \in \mathcal{U}_i} y_j},$$

where  $\mathcal{U}_i \subset \mathcal{U}$ . Hence if  $y_i \leq n\epsilon$  and  $y_j \geq n$  for all  $j \in \mathcal{U}$

$$\phi_i(y) \leq \tilde{C} \frac{\epsilon n}{1+n} \leq \tilde{K} \epsilon.$$

Using the control on the modified proportional fair allocation for  $x_i \geq 1$ :

$$\tilde{\phi}_i(y) \leq \exp\left(\frac{1}{y_i}\right) \phi_i(y) \leq e \phi_i(y) \leq e \tilde{K} \epsilon.$$

□

Now we do a finer classification of indexes in  $\mathcal{S}$ . Let  $\{\epsilon_n\}$  be a sequence such that  $\epsilon_n$  goes to 0 as  $n$  grows, and  $\{t_n = C_2 n\}$ . For each  $n$ , we define the sets

$$\bar{\mathcal{S}}_n = \{i \in \mathcal{S} : X_i(t_n) \geq n\epsilon_n\}, \quad \underline{\mathcal{S}}_n = \{i \in \mathcal{S} : X_i(t_n) < n\epsilon_n\}.$$

We look at the set of events:

$$\mathcal{A} = \{X_i(t_n) \geq nx_i, i \in \mathcal{U}_R, X_i(t_n) \geq n\epsilon_n, i \in \bar{\mathcal{S}}_n, X_i(t_n) \leq n\epsilon_n, i \in \underline{\mathcal{S}}_n\}.$$

Recall that

$$\left| \log \left( \frac{\phi_i(y)}{\tilde{\phi}_i(y)} \right) \right| \leq 1,$$

and remark that on  $A \in \mathcal{A}$ , using the previous lemma, the process counting the number of downwards jumps in direction  $j \in \mathcal{S}$  which are not common for  $X$  and  $\tilde{X}$  is dominated by a Poisson process  $N^{\epsilon_n}$  of intensity  $C_5 \epsilon_n$  independent of  $N^j$ ,  $j \in \mathcal{S}$ .

Hence:

$$\begin{aligned} & E^0 1_{\tau_k \leq t < \tau_{k+1}} 1_A M_{\tau_k, t} \\ & \leq E^0 \left( \exp \left( \sum_{j \in \mathcal{U}} \frac{C_1}{n} (N_t^j + t) + \sum_{j \in \bar{\mathcal{S}}_n} \frac{C_1}{n\epsilon_n} (N_t^j + t) + \sum_{j \in \underline{\mathcal{S}}_n} [C_5 N_t^{\epsilon_n} + (C_6 \epsilon_n) t] \right) 1_{\tau_k \leq t < \tau_{k+1}} 1_A \right). \end{aligned}$$

and,

$$\begin{aligned}
E^0(M_t 1_A) &\leq E^0\left(\exp\left(\sum_{j \in \mathcal{U}} \frac{C_1}{n}(N_t^j + t) + \sum_{j \in \bar{\mathcal{S}}_n} \frac{C_1}{n\epsilon_n}(N_t^j + t) + \sum_{j \in \underline{\mathcal{S}}_n} [C_5 N_t^{\epsilon_n} + (C_6 \epsilon_n)t]\right) 1_A\right) \\
&\leq E^0\left(\exp\left(\frac{\#\mathcal{U}C_1}{n}(\bar{N}_t + t) + \frac{\#\bar{\mathcal{S}}_n C_1}{n\epsilon_n}(\bar{N}_t + t) + \#\underline{\mathcal{S}}_n [C_5 N_t^{\epsilon_n} + C_6 \epsilon_n t]\right) 1_A\right) \\
&\leq E^0\left(\exp\left\{\frac{C_1 p N}{n}(\bar{N}_t + t)\right\}\right)^{\frac{\#\mathcal{U}}{pN}} E^0\left(\exp\left\{\frac{C_1 p N}{n\epsilon_n}(\bar{N}_t + t)\right\}\right)^{\frac{\#\bar{\mathcal{S}}_n}{pN}} \\
&E^0\left(\exp\{pN[C_5 N_t^{\epsilon_n} + C_6 \epsilon_n t]\}\right)^{\frac{\#\underline{\mathcal{S}}_n}{pN}} \mathbb{P}^0(\tilde{X}_t \geq nx)^{\frac{p-1}{p}}, \quad (13)
\end{aligned}$$

where we used Hölder's inequality for  $1 = \frac{\#\mathcal{U}}{pN} + \frac{\#\bar{\mathcal{S}}_n}{pN} + \frac{\#\underline{\mathcal{S}}_n}{pN} + \frac{p-1}{p}$ .

We now bound the speed of convergence, in the large deviations scale, of each of the three first factors on the right-hand side of inequality (13). Choosing  $t_n = C_2 n$ ,  $p_n = C_7 \log n$ ,  $\epsilon_n = n^{-\frac{1}{2}}$ , we have:

$$\begin{aligned}
\frac{1}{n} \log E^0\left(\exp\left\{\frac{C_1 N p_n}{n}(\bar{N}_{t_n} + t_n)\right\}\right)^{\frac{\#\mathcal{U}}{N p_n}} &= \frac{1}{n} \log\left(\exp\left\{\frac{\#\mathcal{U} C_1 t_n}{n}\right\} \exp\{\bar{\lambda} t_n [e^{\frac{C_1 N p_n}{n}} - 1]\}\right)^{\frac{\#\mathcal{U}}{N p_n}} \\
&= \frac{\#\mathcal{U} C_1 C_2}{n} + \frac{\#\mathcal{U} \bar{\lambda} t_n}{N n p_n} [e^{\frac{C_1 N p_n}{n}} - 1] \\
&= \frac{\#\mathcal{U} C_1 C_2}{n} + \frac{\#\mathcal{U} \bar{\lambda} C_2}{N C_7 \log n} [e^{\frac{C_1 C_7 N \log n}{n}} - 1] \\
&= O(n^{-1}).
\end{aligned}$$

We do similarly for the second factor:

$$\begin{aligned}
\frac{1}{n} \log E^0\left(\exp\left\{\frac{C_1 N p_n}{\epsilon_n n}(\bar{N}_{t_n} + t_n)\right\}\right)^{\frac{\#\bar{\mathcal{S}}_n}{N p_n}} &= \frac{1}{n} \log\left(\exp\left\{\frac{\#\bar{\mathcal{S}}_n C_1 t_n}{\epsilon_n n}\right\} \exp\{\bar{\lambda} t_n [e^{\frac{C_1 N p_n}{\epsilon_n n}} - 1]\}\right)^{\frac{\#\bar{\mathcal{S}}_n}{N p_n}} \\
&= \frac{\#\bar{\mathcal{S}}_n C_1 C_2}{\epsilon_n n} + \frac{\#\bar{\mathcal{S}}_n \bar{\lambda} t_n}{N n p_n} [e^{\frac{C_1 N p_n}{\epsilon_n n}} - 1] \\
&= \frac{\#\bar{\mathcal{S}}_n C_1 C_2}{\epsilon_n n} + \frac{\#\bar{\mathcal{S}}_n \bar{\lambda} C_2}{N C_7 \log n} [e^{\frac{C_1 C_7 N \log n}{\epsilon_n n}} - 1] \\
&= O((n\epsilon_n)^{-1}) = O(n^{-\frac{1}{2}}).
\end{aligned}$$

Finally, for the third factor we have:

$$\begin{aligned}
\frac{1}{n} \log E^0 \left( \exp\{Np_n[C_5 N_{t_n}^{\epsilon_n} + C_6 \epsilon_n t_n]\} \right)^{\frac{\#\underline{S}_n}{Np_n}} &= \frac{1}{n} \log \left( \exp\{C_6 \#\underline{S}_n \epsilon_n t_n\} \exp\{\epsilon_n t_n [e^{C_5 Np_n} - 1]\} \right)^{\frac{\#\underline{S}_n}{Np_n}} \\
&= \frac{C_6 \epsilon_n \#\underline{S}_n t_n}{n} + \frac{\#\underline{S}_n \epsilon_n t_n}{Np_n n} [e^{C_5 Np_n} - 1] \\
&= C_6 C_2 \#\underline{S}_n \epsilon_n + \frac{C_2 \#\underline{S}_n \epsilon_n}{C_7 N \log n} [e^{C_5 C_7 N \log n} - 1] \\
&= C_6 C_2 \#\underline{S}_n \epsilon_n + \frac{C_2 \#\underline{S}_n \epsilon_n [n^{C_5 C_7 N} - 1]}{C_7 N \log n} \\
&= O(\epsilon_n) + O\left(\frac{\epsilon_n n^\epsilon}{\log n}\right) = O(n^{-\frac{1}{2} + \epsilon}).
\end{aligned}$$

where we chose  $C_7 > 0$  small enough such that  $C_7 C_5 N < \epsilon$ , for any given  $\epsilon > 0$ .

Then along the same lines as in the case  $x$  has all entries positive, it follows

$$\frac{1}{n} \log(\pi^{PF}(\{nx\}^\uparrow)) \leq \frac{1}{n} \log(\pi^{mPF}(\{nx\}^\uparrow)) + O(n^{-\frac{1}{2} + \epsilon}) \quad \forall \epsilon > 0.$$

□

## 4 Large deviations for monotone networks with general service time distributions

In this Section, we consider a processor sharing network (i.e. a set of processor sharing nodes) with a proportionally fair bandwidth allocation. This means that a flow is served at node  $i$  with speed  $\frac{\phi_i^{PF}(X_t)}{X_i(t)}$ . We need the notion of strong monotonicity introduced for instance in [5].

**Definition 4.**  $\phi$  is strongly decreasing if the function  $\psi_i$  defined by  $\psi_i(x) = \frac{\phi_i(x)}{x_i}$ , if  $x_i > 0$  and 0 otherwise, is decreasing in  $x_j$  for all  $j \neq i$ .

We assume in this Section that the network is monotone, i.e. that  $\phi$  is strongly increasing. This is verified on all tree topologies for instance and for many wireless networks instances. In the proof of the following theorem, we use the monotony of the network to obtain stochastic comparisons. We recall the following Proposition:

**Proposition 8.** *Let  $X_t$  and  $\tilde{X}_t$  two processes associated with the allocations  $\phi$  and  $\tilde{\phi}$ . Suppose that  $\phi$  is strongly decreasing, and that  $\frac{\phi_i(x)}{x_i} \geq \frac{\tilde{\phi}_i(x)}{x_i}$  for each  $x$ . Then for all  $t$  and for all service time distribution:*

$$X_t \prec_{st} \tilde{X}_t.$$

**Proof.**

Remark that the processes  $X_t$  and  $\tilde{X}_t$  are not Markov in general. We can however easily construct a sample-path comparison between the processes. Starting with ordered initial configuration of flows  $x \prec y$ , the rates at which each flow is served in the network verify:

$$\frac{\phi_i(x)}{x_i} \geq \frac{\phi_i(y)}{y_i} \geq \frac{\tilde{\phi}_i(y)}{y_i}.$$

Hence, a coupling can be constructed such that  $X_t \prec \tilde{X}_t$  almost surely.  $\square$

Recall that we define  $P^{PF}$  as  $P^{PF}(x) = \max_{\eta \in \log(\mathcal{C})} \langle \eta, x \rangle$ , and that  $\log(\phi_i(x)) = \nabla P^{PF}(x)$ .

Now define the discrete potential functions  $\overline{\Psi}$  and  $\underline{\Psi}$  by the recursive formula:

$$\begin{aligned} \overline{\Psi}(x) &= \max_{i=1 \dots N} (\overline{\Psi}(x - e_i) - \log(\phi_i(x))), \\ \underline{\Psi}(x) &= \min_{i=1 \dots N} \{\underline{\Psi}(x - e_i) - \log(\phi_i(x))\}. \end{aligned}$$

We call supPF and infPF the reversible allocations associated with the discrete potentials  $\overline{\Psi}$  and  $\underline{\Psi}$  and define the rate function  $R(x) = \langle \log(\lambda), x \rangle - P(x)$ . We can now state the main result of this Section:

**Theorem 4.** *Suppose the network monotone, i.e.  $\phi$  is strongly monotone. If  $\lambda$  is in the interior of the capacity set  $\mathcal{C}$ , then the allocations Proportional fairness, supPF and infPF as well as Balanced fairness are stable and admit the same large deviation characteristics with rate function  $R$ .*

**Proof.**

Given the definitions of the discrete potentials  $\overline{\Psi}$  and  $\underline{\Psi}$ , observe that there exists two paths  $\overline{\mathcal{P}}(x)$  and  $\underline{\mathcal{P}}(x)$  from 0 to  $x$  such that:

$$\begin{aligned}\overline{\Psi}(x) &= - \sum_{k:(i_k, z_k) \in \overline{\mathcal{P}}(x)} \log(\phi_{i_k}(z_k)), \\ \underline{\Psi}(x) &= - \sum_{k:(i_k, z_k) \in \underline{\mathcal{P}}(x)} \log(\phi_{i_k}(z_k)),\end{aligned}$$

where with a slight abuse of notations the indexes  $i(z)$  correspond to the indexes defined by the specific paths  $\overline{\mathcal{P}}(x)$  and  $\underline{\mathcal{P}}(x)$ .

Let  $\overline{X}$  and  $\underline{X}$  the process associated with the balance function  $\overline{\Psi}$  and  $\underline{\Psi}$ . Recall now that  $P$  is the potential associated with the proportional fair allocation, and since we can define this potential up to an additive constant, let us choose it such that  $P(0) = 0$ . So, we can write that for any path  $\mathcal{P}(x)$  going from 0 to  $x$ ,  $P(x) = P(x) - P(0) = \sum_{k:(i_k, z_k) \in \mathcal{P}(x)} \mathbf{D}_{i_k} P(z_k)$ . Choosing  $\mathcal{P}(x) = \overline{\mathcal{P}}(x)$ :

$$\begin{aligned}|\overline{\Psi}(nx) - P(nx)| &= \left| \sum_{k:(i_k, z_k) \in \overline{\mathcal{P}}(nx)} \mathbf{D}_{i_k} P(z_k) - \log(\phi_{i_k}(z_k)) \right|, \\ &\leq \sum_{k:(i_k, z_k) \in \overline{\mathcal{P}}(nx)} |\mathbf{D}_{i_k} P(z_k) - \log(\phi_{i_k}(z_k))|. \quad (14)\end{aligned}$$

Recalling that for  $x_i \geq 1$ :

$$|P(x) - P(x - e_i) - \log(\phi_i(x))| \leq \frac{1}{x_i},$$

we can bound (14) by:

$$\begin{aligned}|\overline{\Psi}(nx) - P(nx)| &\leq \sum_{(i_k, z_k) \in \overline{\mathcal{P}}(nx)} \frac{1}{z_{i_k}}, \\ &\leq C \sum_{i=1}^n \frac{1}{i} \leq C' \log(n). \quad (15)\end{aligned}$$

Define the invariant measures (not necessarily stationary)  $\overline{\pi}$  and  $\underline{\pi}$  of the processes  $\overline{X} \prec X \prec \underline{X}$ . Since we have constructed these processes from a discrete potential, we have that (see Proposition 2):

$$\overline{\pi}(x) = \lambda^x \exp(-\overline{\Psi}(x)),$$

$$\underline{\pi}(x) = \lambda^x \exp(-\underline{\Psi}(x)).$$

Now observe that:

$$\begin{aligned} \frac{1}{n} \log(\overline{\pi}(nx)) &= \langle \log(\lambda), x \rangle - \frac{1}{n} \overline{\Psi}(nx), \\ \frac{1}{n} \log(\underline{\pi}(nx)) &= \langle \log(\lambda), x \rangle - \frac{1}{n} \underline{\Psi}(nx). \end{aligned}$$

Using (15) we obtain that

$$\begin{aligned} \frac{1}{n} \log(\overline{\pi}(nx)) &= \frac{1}{n} \langle \log(\lambda), nx \rangle - \frac{1}{n} P(nx) + O(n^{-1}), \\ \frac{1}{n} \log(\underline{\pi}(nx)) &= \frac{1}{n} \langle \log(\lambda), nx \rangle - \frac{1}{n} P(nx) + O(n^{-1}), \end{aligned}$$

and hence, by the 1-homogeneity of  $R$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log(\overline{\pi}(nx)) = \lim_{n \rightarrow \infty} \frac{1}{n} \log(\underline{\pi}(nx)) = R(x).$$

Assume  $\lambda$  is in the interior of the capacity set  $\mathcal{C}$ . Using the definition of  $P$ , this implies that there exists  $a > 0$  such that

$$R(x) = \langle \log(\lambda), x \rangle - P(x) = \langle \log(\lambda) - \log(\phi(x)), x \rangle \leq -a|x|,$$

which implies that the invariant measure  $\overline{\pi}$  and  $\underline{\pi}$  are summable. This hence proves that the stationary distributions of both processes  $\overline{X}$  and  $\underline{X}$  are well defined for Markovian dynamics while the reversibility condition (i.e. the balance property of the service rates) implies the insensitivity of the stationary distribution to the service time distribution [3, 16]. Hence, for any service time distribution, the networks with allocation infPF and supPF admit the stationary distribution  $C_1 \overline{\pi}$  and  $C_2 \underline{\pi}$  where  $C_1$  and  $C_2$  are normalizing constants. Now using the assumption of monotonicity of the network, we obtain that:

$$P(|\overline{X}| \geq n) \geq P(|X| \geq n) \geq P(|\underline{X}| \geq n), \quad (16)$$

and by Cramér's theorem we conclude:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log P(|X| \geq n) = \lim_{n \rightarrow \infty} \frac{1}{n} \log P(|X| = n).$$

It has further been proven in [10] that mPF and BF have  $R(x)$  as rate functions.

□

## 5 Conclusion

We proved that the stationary measure of the number of flows in progress in a bandwidth sharing network functioning under the proportional fair allocation shares the same large deviations characteristics with the stationary measure of the number of flows in progress of the same network under the balanced fair allocation. This formalizes the idea that in long excursions proportional fair allocation behaves similarly to the most efficient insensitive allocation.

## References

- [1] T. Bonald, S. Borst, N. Hegde, M. Jonckheere, and A. Proutiere. Flow-level performance and capacity of wireless networks with user mobility. *To appear in Queueing Syst. Theory Appl.*
- [2] T. Bonald, L. Massoulié, A. Proutière, and J. Virtamo. A queueing analysis of max-min fairness, proportional fairness and balanced fairness. *Queueing Syst. Theory Appl.*, 53(1-2):65–84, 2006.
- [3] T. Bonald and A. Proutière. Insensitivity in processor-sharing networks. *Perform. Eval.*, 49(1-4):193–209, 2002.
- [4] T. Bonald and A. Proutière. Insensitive bandwidth sharing in data networks. *Queueing Syst. Theory Appl.*, 44(1):69–100, 2003.
- [5] T. Bonald and A. Proutière. On stochastic bounds for monotonic processor sharing networks. *Queueing Syst. Theory Appl.*, 47(1/2):81–106, 2004.
- [6] M. Freidlin and A. Wentzell. *Random perturbations of dynamical systems*. Springer, 1984.
- [7] F. Kelly. *Reversibility and Stochastic Networks*. Wiley, 1979.
- [8] F. P. Kelly and R. J. Williams. Fluid model for a network operating under a fair bandwidth-sharing policy. *Annals of Applied Probability*, 14(3):1055–1083, 2004.
- [9] D. J. Leith, V. G. Subramanian, and K. R. Duffy. Log-convexity of rate region in 802.11e wlangs. *IEEE Communications Letters*, 14(1):57–59, 2010.

- [10] L. Massoulié. Structural properties of proportional fairness: Stability and insensitivity. *Ann. Appl. Probab.*, 17(3):809–839, 2007.
- [11] L. Massoulié and J. Roberts. Bandwidth sharing: objectives and algorithms. *IEEE/ACM Trans. Netw.*, 10(3):320–328, 2002.
- [12] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer-Verlag, 1993.
- [13] T. Rockafellar. *Convex analysis*, volume 28. Princeton landmarks in mathematics, 1970.
- [14] A. Schwartz and A. Weiss. *Large Deviations for performance analysis. Queues, communication and computing*. Chapman & Hall, 1995.
- [15] N. Walton. Insensitive, maximum stable allocations converge to proportional fairness. *Queueing Systems*, 68(1):51–60, 2011.
- [16] S. Zachary. A note on insensitivity in stochastic networks. *Journal of Applied Probability*, 44(1):238–248, 2007.