Stability of spatial queueing systems

Charles Bordenave

HAL Id: inria-00070695
https://inria.hal.science/inria-00070695

Submitted on 19 May 2006

R INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Stability of spatial queueing systems

Charles Bordenave*

N° 5305

THÈME 1

_R apport
de recherche_

Stability of spatial queueing systems

Charles Bordenave*

Thème 1 — Réseaux et systèmes
Projets TREC

Abstract: In this report, we analyze a queueing system characterized by a space-time arrival process of customers served by a countable set of servers. Customers arrive at some points in space and the server stations have space-dependent processing rates. The workload is seen as a Radon measure and the server stations can adapt their power allocation to the current workload. We derive the stability region of the queuing system in the usual stationary ergodic framework. From the analysis of this stability region, we derive optimal partitions of space among server stations. Some specific subclasses of policies are also studied. Wireless communication networks provides a natural field of application for this model.

Key-words: Stability Region, Space-time point process, Wireless networks, Dynamic scheduling.

* INRIA/ENS, 45 rue d'Ulm, 75005 Paris ; e-mail: charles.bordenave@ens.fr

Stabilité des files d'attente spatiales

Résumé : Dans ce rapport, nous analysons une file d'attente caractérisée par un processus spacio-temporel d'arrivée de clients servis par un ensemble dénombrable de serveurs. Les clients arrivent sur des points de l'espace et les serveurs ont des vitesses de traitement qui dépendent de l'espace. La quantité de travail est vue comme une mesure de Radon et les serveurs peuvent adapter leur allocation de puissance avec la quantité de travail en cours. Nous établissons la region de stabilité de la file d'attente dans le cadre stationnaire ergodique usuel. L'analyse de la région de stabilité permet de définir des partitions de l'espace entre serveurs optimales. Certaines sous-classes de politiques sont également étudiées. Les réseaux sans fils est le champ d'application naturel de ce model.

Mots-clés : région de stabilité, processus ponctuels spaciaux, réseaux sans fils, politique adaptative.

1 Introduction : Model description

In this paper, we analyze a space-time arrival process of customers served by a countable set of servers. This model is motivated by large scale wireless communications networks but could suit to other types of infinite queuing systems. This model is along the line of other allocation problems studied by Tassiulas and Ephremides [11] and Bambos et al. [8], [1], [3]. Applications of this work to wireless networks are done in [5].

We consider a system where some jobs arrive exogenously, the jobs are located in the space. Some server stations serve the incoming jobs. Each server station can process the jobs at a rate depending of the localization of the job and a random environment variable.

All the random variables, we introduce in this section are defined on a common probability space $(\Omega, \mathcal{F}, P)$. This space is endowed with a measurable flow $\{\theta_t\}, t \in \mathbb{R}$. We suppose that (P, θ_t) is *ergodic*.

1.1 Customers arrival point process

In our system, the customers (or jobs) are seen as points of a marked point process A. Namely A is a spatial marked point process on $\mathbb{R} \times \mathbb{R}^d$ with marks on $\mathbb{R}_+$. We use the notation : $A = \sum_n \delta_{\{T_n, X_n, \sigma_n\}}$ to represent the point process. The nth job arrives at time $T_n \in \mathbb{R}$ located at $X_n \in \mathbb{R}^d$ and it requires a service time of $\sigma_n \in \mathbb{R}_+$. We suppose that A is compatible with the flow θ that is : if $A(\omega) = \sum_n \delta_{\{T_n, X_n, \sigma_n\}}$, $A \circ \theta_t(\omega) = A(\theta_t \omega) = \sum_n \delta_{\{T_n - t, X_n, \sigma_n\}}$. This implies that the point process A is stationary along the first coordinate. The intensity measure of A is defined for a Borel set $B \subset \mathbb{R}^{d+1}$ as :

$$E(A(B)) = E\Big(\sum_n \mathbb{1}((T_n, X_n) \in B) \Big).$$

It is supposed to be finite for all compacts K. Thus the intensity is a *Radon measure* of the form : $\lambda(dx)dt$ (see for example Lemma A2.7.11., p.634 in Daley and Vere-Jones [7]). By definition for a compact $K \subset \mathbb{R}^{d+1}$:

$$E(A(K)) = \int_K \lambda(dx)dt < +\infty. \tag{1}$$

Note that λ is not necessarily a finite measure on $\mathbb{R}^d$. Therefore between time t and $t' > t$, there arrives an infinite number of jobs on the space. In particular, the indexation of customers cannot be time monotonous. However, on a compact set $K \subset \mathbb{R}^d$, only a finite number of jobs arrives almost-surely (a.s.).

As usual, we will often drop the "ω" in the expression of random variables to simplify notations.

Define $P_A^{t,x}(\bullet)$ as the Palm probability of the point process A at $(t, x) \in \mathbb{R}^{d+1}$ (refer to [7], chapter 12). Since A is compatible with θ_t, we have :

$$P_A^{t,x}(\bullet) = P_A^{0,x}(\theta_{-t} \bullet).$$

Under $P_A^{t,x}$, let $\sigma_{t,x}$ be the required service of a customer arriving at time t in x. It is consistent since there is a customer arrived at t in x $P_A^{t,x}$-a.s.. We will suppose that $0 < E^{0,x}(\sigma_{0,x}) < +\infty$, $\lambda(dx)$-a.e.. Remind that $E^{0,x}(\sigma_{0,x})$ can be understood as the mean number of service time requirement of a typical customer arriving in x.

1.2   Server stations adapative policy

We have a *countable* set of server stations denoted by $\mathcal{J}$. The servers can provide service to all points in space and they have a different processing rate for each point in space ; the server j serves a customer located at x at a rate of $r_j(x)$. We suppose that $x \mapsto r_j(x)$ is a bounded positive measurable function and :

$$\lim_{|x| \to +\infty} r_j(x) = 0, \tag{2}$$

where $|\bullet|$ denotes the Euclidean norm.

   The server stations are in a random environment and their processing powers varies over time. At time t, the total processing power available is $\epsilon_j(t) \in \mathbb{R}_+$. As all the other random variables we consider, we suppose that the driving process $\epsilon_j(t)$ is compatible with the shift, that is to say : $\epsilon_j(t) = \epsilon_j(0) \circ \theta_t$.

   We suppose :

$$p_j = E(\epsilon_j(0)) < +\infty.$$

   The workload at time t is the set of all the jobs waiting to be processed. It is denoted by W_t and it is an atomic measure on $\mathbb{R}^d$: for a Borel set B, $\int_B W_t(dx)$ represents the total service time remaining at time t for all customers located in B.

   The server stations divide their processing power between the jobs required according to a policy scheme. This power allocation depends on the current workload. We suppose that our queuing system cannot handle an infinite amount of service in finite time for a given location x : more precisely :

$$\text{for all bounded borel set B,} \quad \sum_j p_j \sup_{x \in B} r_j(x) < +\infty. \tag{3}$$

   We can define a *policy* properly.

Definition 1. *Let $\mathcal{M}$ be the set of Radon measures on $\mathbb{R}^d$, endowed with the vague topology (refer to [7]). A policy is a measurable mapping :*

$$\pi : \begin{cases} \mathcal{M} \times \Omega & \to & \mathcal{M}^{\mathcal{J}} \\ (m, \omega) & \mapsto & \{\pi_j(m, \omega)\}, j \in \mathcal{J}, \end{cases}$$

satisfying, for all $\omega \in \Omega$:

$$\int_{\mathbb{R}^d} \pi_j(m, \omega)(dx) \le \epsilon_j(0)(\omega) \tag{4}$$

and

$$\pi_j(m, \omega) \text{ is absolutely continuous with respect to } m. \tag{5}$$

The policy enforced at time t is : $\pi(W_t, \theta_t\omega)$.

Equation (4) means that the server stations cannot allocate more than their total processing power. If the total workload on a Borel set is zero, it is useless to dedicate some processing power to this set. This is the meaning of Condition (5).

$\pi_j(W_t)(B)$ is the processing power the server station j dedicates to jobs located in B at time t. Therefore, the server station j achieves at time t an instantaneous service rate for a job located in x of :

$$r_j(x)\pi(W_t)(\{x\}).$$

The policies we are considering in our model are stochastic and stationary : if W is stationary, that is if $W_t = W \circ \theta_t$, for a given random measure W, $\pi(W_t, \theta_t\omega) = \pi(W, \omega) \circ \theta_t$ and thus π is stationary. The study of non-stationary policies is performed in [11] in an other framework ; in this article the authors show that the non-stationary policies have the same performance as stationary policies.

In the next sections, we will define some interesting classes of policies.

Here is a simple example of a deterministic policy defined for atomic measures :

$$\pi_j^+(m) = \begin{cases} 0 & \text{if} & m \text{ is the measure zero} \\ \epsilon_j \delta_{x_j^+} & \text{otherwise} . \end{cases}$$

where, $x_j^+ = \arg\max\left\{x : r_j(x)\mathbb{1}(m(\{x\}) > 0)\right\}$. If multiple choices of x are possible, choose the first in the lexicographic order.

With this policy, the server stations serve the user with the best processing rate first. Note in particular that this policy is *work conserving* : if W_t is not the 0 measure, the server is active.

1.3 Evolution Equation

The dynamic of our queuing system is given by the following integral equation, for all Borel sets B and $t' > t$,

$$W_{t'}(B) = W_t(B) + \int_t^{t'} \int_B \sigma_{s,x} A(ds \times dx) - \sum_j \int_t^{t'} \int_B r_j(x)\pi_j(W_s)(dx)ds, \tag{6}$$

where by definition : $\int_t^{t'} \int_B \sigma_{s,x} A(ds \times dx) = \sum_n \sigma_n \mathbb{1}(T_n \in [t, t'[)\mathbb{1}(X_n \in B)$. If W_0 is an atomic measure then from Condition (5), $W_t \geq 0$ for all $t \geq 0$ and it is also an atomic measure (a.s.).

Using Borovkov's terminology (see [6], chap. 4), Equation (6) defines a *stochastic recursive process* : we can write $W_{t+h} = f_h(W_t, A_t^{t+h})$, where A_t^{t+h} denotes the trajectory of the arrival point process A between t and $t+h$.

Under a given policy π, for a Radon measure m, define W_t^m as the workload at time t when $W_0 = m$.

1.4   Some Examples

Here come a few examples which illustrate the model.

Example 1.   All the jobs arrive on a countable set of points : $\{x_i\}, i \in \mathbb{N}$, with no accumulation points. This points are buffers and Assumption (1) reads :

$$\lambda_i = E(A([0,1] \times \{x_i\})) < +\infty.$$

The server stations serve users arriving at x_i with a processing rate $r_j(x_i)$.

In this example, the system reduces to a multi-class job traffic with processing rates depending on the class and the server station. Stability results on this type of systems have been presented in [8].

Example 2.   In a wireless communication scenario, $d = 2$ and the server stations are base antennas and the customers are mobile users which want to receive some data from the network. Server station j is located at $Y_j \in \mathbb{R}^2$. The processing rate can be written : $r_j(x) = f(x)L(x, Y_j)$. $L(x, y)$ is an attenuation function on the channel between x and y. A natural assumption is to assume that $L(x, y)$ depends only on $|y - x|$. If customer arrivals are uniform on the plane, then the intensity measure $\lambda(dx)$ is proportional to the Lebesgue measure.

Example 3.   When $\epsilon_j \in \{0, 1\}$, the model has a random connectivity. The server stations are switched on or switched off. At time t, if $\epsilon_j(t) = 1$ (resp. 0) the base station j is switched on (resp. switched off).

1.5   Stability Region

In this paper we analyze the stability of the queuing system described above. The definition of stability we use relies on the existence of a stationary regime.

Definition 2.   *A policy π is* stable *if there exists a finite stationary workload $\{M_t = M \circ \theta_t\}, t \in \mathbb{R}$, for an atomic random measure M satisfying Equation (6).*

The queuing system is said to be stable if there exists a stable policy.

This definition is called *stochastic stability* in [3]. In the next section, we will adopt a stronger definition of stability.

The parameter of the queuing system is the arrival marked point process A. Let $\mathcal{N}$ be the set of possible point processes A (namely the point processes of finite intensity and with a service time with finite expectation under the Palm measure). We define :

$$\mathcal{F} = \Big\{ \quad f = (x \mapsto f_j(x), j \in \mathcal{J}) \text{ such that,}$$
$$\forall j, \, f_j \text{ is non negative, measurable and } \lambda(dx)\text{-a.e. } \sum_j f_j(x) = 1 \Big\}.$$

$$\mathcal{N}^s = \Big\{ A \in \mathcal{N} : \exists f \in \mathcal{F} \text{ such that : } \forall j, \int_{\mathbb{R}^d} \frac{E_A^{0,x}(\sigma_{0,x})f_j(x)}{r_j(x)} \lambda(dx) < p_j \Big\}. \tag{7}$$

$$\bar{\mathcal{N}}^s = \Big\{ A \in \mathcal{N} : \exists f \in \mathcal{F} \text{ such that : } \forall j, \int_{\mathbb{R}^d} \frac{E_A^{0,x}(\sigma_{0,x})f_j(x)}{r_j(x)} \lambda(dx) \le p_j \Big\}. \tag{8}$$

These sets are continuous generalizations of the stability sets derived in [8].

Note that $\int_{\mathbb{R}^d} \frac{E_A^{0,x}(\sigma_{0,x})f_j(x)}{r_j(x)} \lambda(dx)$ is a traffic load : $E_A^{0,x}(\sigma_{0,x})\lambda(dx)$ is the mean number of service requirements per unit of surface and $r_j(x)$ is the processing rate at x for server station j.

We can now state the following stability theorem :

Theorem 1. *For the queuing model described above,*

- *if $A \in \mathcal{N}^s$, then there exists a stable policy,*

- *if there is a stable policy then $A \in \bar{\mathcal{N}}^s$.*

Note that, as in Loynes' Theorem, the stability region depends on the distribution of the point process A only through its means. To fix ideas, suppose that there is only one server station and that all jobs arrive at the same place, say 0. Then the theorem says that the stability region is given by $E_A^{0,0}(\sigma_{0,0})\lambda(0) < p_1 r_1(0)$. This result is the usual condition $\rho < 1$ for G/G/1 queues.

In the proof of the theorem, we establish that to a given stable policy π corresponds a set of functions $(f_j)_{j \in \mathcal{J}}$ in $\mathcal{F}$ such that :

$$\int_{\mathbb{R}^d} \frac{E_A^{0,x}(\sigma_{0,x})f_j(x)}{r_j(x)} \lambda(dx) \le p_j.$$

$f_j(x)$ is the proportion of service carried by the server j for users in x in the stationary regime and for the stable policy π.

The converse mapping is also available, for a set of functions $(f_j)_{j \in \mathcal{J}}$ in $\mathcal{F}$ such that the above inequality is satisfied with a strict inequality, there exists a stable policy. In fact, as it will be seen in the proof of Theorem 1 this last assertion is only true for a dense subset of functions of $\mathcal{F}$.

A proof of Theorem 1 is given in Section 2.

In Section 3, we give a stronger stability results on our queuing systems which follows from classical coupling arguments.

In Section 4, we consider a subclass of policies which preserves a kind of monotony for the workload measure.

Section 5 is dedicated to the stability region defined by Equation (7) when the spatial intensity measure in absolutely continuous with respect to the Lebesgue measure.

The last results of this paper concern a subclass of policy for which the stability region is easy to compute. This work is done in Section 6.

2 Proof of Theorem 1

2.1 Necessary Condition

This technical lemma is needed in what follows.

Lemma 1. *Let Z be a non-negative a.s. finite random variable, such that for a given t, $Z - Z \circ \theta_t \in L^1(P)$, then $E(Z - Z \circ \theta_t) = 0$.*

For the proof, see for example Lemma 2.2.1 of Baccelli and Bremaud [2].

Suppose exists a stable policy π. There exists a stationary workload measure $W_t = W \circ \theta_t$ a.s. finite. Let B be a bounded Borel set, from Equation (6), we have :

$$W \circ \theta_t(B) = W(B) + \int_0^t \int_B \sigma_{s,x} A(ds \times dx) - \sum_j \int_t^{t+h} \int_B r_j(x) \pi_j(W \circ \theta_s))(dx) ds.$$

The Campbell Formula for marked point processes, using (1.1), implies :

$$
\begin{aligned}
E(\int_0^t \int_B \sigma_{s,x} A(ds \times dx)) &= \int_0^t \int_B \int_{\mathbb{R}_-} \sigma_{s,x} P_A^{s,x}(d\sigma) \lambda(dx) ds \\
&= t \int_B E_A^{0,x}(\sigma_{0,x}) \lambda(dx).
\end{aligned}
$$

Since $\pi_j(B)$ is bounded by ϵ_j and we can define $\bar{\pi}_j(B) = E(\pi_j(W)(B))$. $\bar{\pi}_j$ is a Radon measure. Moreover, since $\int_{\mathbb{R}^d} \pi_j(W)(dx) \leq \epsilon_j$, we have : $\int_{\mathbb{R}^d} \bar{\pi}_j(dx) \leq p_j$. Using Condition (3), we deduce :

$$E(|W \circ \theta_t(B) - W(B)|) \leq t \int_B E_A^{0,x}(\sigma_{0,x}) \lambda(dx) + t \sum_j p_j \sup_{x \in B} r_j(x) < +\infty.$$

Thus $W \circ \theta_t(B) - W(B) \in L^1(P)$ and we can apply Lemma 1 to conclude that :

$$0 = \int_B E_A^{0,x}(\sigma_{0,x}) \lambda(dx) - \sum_j \int_B r_j(x) \bar{\pi}_j(dx). \tag{9}$$

Since Equation (9) holds for all bounded Borel sets, the Radon measures : $\sum_j r_j(x)\bar{\pi}_j(dx)$ and $E_A^{0,x}(\sigma_{0,x})\lambda(dx)$ are equal. In particular the measure $r_j(x)\bar{\pi}_j(dx)$ is absolutely continuous with respect to $E_A^{0,x}(\sigma_{0,x})\lambda(dx)$. We define the measurable function $\tilde{\pi}_j(x)$ as the Radon-Nicodym derivative of $\bar{\pi}_j(dx)$ with respect to $\lambda(dx)$. Let $f_j(x) = \frac{r_j(x)\tilde{\pi}_j(x)}{E_A^{0,x}(\sigma_{0,x})}$. We deduce from (9) that for all Borel B :

$$\int_B E_A^{0,x}(\sigma_{0,x})\lambda(dx) = \sum_j \int_B f_j(x)E_A^{0,x}(\sigma_{0,x})\lambda(dx).$$

Thus $\lambda(dx)$-a.e. : $\sum_j f_j(x) = 1$.

We have proved that $f \in \mathcal{F}$. $\int_{\mathbb{R}^d} \bar{\pi}_j(dx) \leq p_j$ reads $\int_{\mathbb{R}^d} \frac{E_A^{0,x}(\sigma_{0,x})f_j(x)}{r_j(x)}\lambda(dx) \leq p_j$.

The second assertion of Theorem 1 follows.

2.2 Sufficient Condition

Suppose, $A \in \mathcal{N}^s$. There exists $(f_j), j \in \mathcal{J}$, in $\mathcal{F}$ such that : for all j, $\int_{\mathbb{R}^d} \frac{E_A^{0,x}(\sigma_{0,x})f_j(x)}{r_j(x)}\lambda(dx) = \rho_j < 1$. We can suppose that f has the properties given in Proposition 6 (in Appendix). For a given policy π and an initial atomic workload m, we define the sets :

$$\mathcal{A}_j^m(t) = \left\{ x : r_j(x) \int_0^t \pi_j(W_s^m)(\{x\})ds \leq f_j(x)\Big(m(\{x\}) + \int_{[0,t)\times\{x\}} \sigma_{s,x} A(ds \times dx)\Big) \right\}.$$

The fact that x is in $\mathcal{A}_j^m(t)$ means that server j has contributed for less than $f_j(x)$ to fulfill the service requirements of customers located in x.

Consider the following non-stationary policy : for all $j \in \mathcal{J}$, for $t \geq 0$:

$$\pi_j : \begin{cases} \Omega \times \mathbb{R} & \to & \mathcal{M} \\ (\omega, t) & \mapsto & \begin{cases} \epsilon_j(t)\delta_{x_j^*} & \text{if} & \mathcal{A}_j^{W_0}(t) \neq \emptyset \\ 0 & \text{if} & \mathcal{A}_j^{W_0}(t) = \emptyset. \end{cases} \end{cases} \tag{10}$$

where,

$$x_j^* = \arg\max\left\{ x : \mathbb{1}(x \in \mathcal{A}_j^{W_0}(t))\mathbb{1}(W_t(\{x\}) > 0))r_j(x) \right\}$$

if multiple choices of x_j^* are possible, choose the first in the lexicographic order.

The existence of this policy follows from Condition (2).

The policy π divides the workload among all servers with proportions f_j and processes the jobs at the faster available rate. This policy is not stationary and is defined for an initial workload at time 0. In other words, $\pi_j(t)$ is the policy enforced at time t when at time 0, the workload was equal to m : with the usual notation, $\pi_j(t) = \pi_j(W_t^m)$.

Note also that if at time T, $\mathcal{A}_j^m(T) = \emptyset$, then the server j will only serve after time T jobs arrived after time T and for $t > 0$:

$$\mathcal{A}_j^m(t + T) = \mathcal{A}_j^0(t) \circ \theta_T. \tag{11}$$

Let B be a bounded Borel set, we have :

$$
\begin{aligned}
W_t(B) &= \int_0^t \int_B \sigma_{s,x} A(ds \times dx) - \sum_j \int_0^t \int_B r_j(x)\pi_j(s)(dx)ds \\
&= \sum_j \int_0^t \int_B f_j(x)\sigma_{s,x} A(ds \times dx) - \sum_j \int_0^t \int_B r_j(x)\pi_j(s)(dx)ds \\
&= \sum_j \int_B r_j(x)\tilde{W}_t^j(dx),
\end{aligned}
$$

where :

$$\tilde{W}_t^j = \int_0^t \int_B \frac{f_j(x)\sigma_{s,x}}{r_j(x)} A(ds \times dx) - \int_0^t \int_B \pi_j(s)(dx)ds. \tag{12}$$

Definition 3. *We define a total order $\prec_j$ on $\mathbb{R}^d$: $x \succ_j y$ if $r_j(x) > r_j(y)$ or $r_j(x) = r_j(y)$ and x is in the lexicographic order smaller than y.*
A Borel set B is a j-max set if : $\{x : x \succeq_j y \text{ for all } y \in B\} = B$.

Lemma 2. *For a j-max set B :*

$$
\begin{aligned}
\tilde{W}_t^j(B) &= \Big(\max \quad \big(\int_B \frac{f_j(x)W_0(dx)}{r_j(x)} + \int_0^t \int_B \frac{f_j(x)\sigma_{s,x}}{r_j(x)} A(ds \times dx) - \int_0^t \epsilon_j(s)ds, \\
&\qquad \sup_{0 \le h \le t} \int_0^h \int_B \frac{f_j(x)\sigma_{s,x}}{r_j(x)} A(ds \times dx) - \int_0^h \epsilon_j(s)ds \big) \Big)^+ .
\end{aligned}
$$

Proof. The policy π divides the global queuing system with the set of servers $\mathcal{J}$ into distinct queues, one for each server in $\mathcal{J}$ with an incoming workload equal to :

$$\sum_j \int_{-t}^0 \int_B f_j(x)\sigma_{s,x} A(ds \times dx).$$

$\tilde{W}_t^j(B)$ is the rescaled workload : the server j serves the user at x with a unit processing rate. If B is a j-max set, from the definition of policy π_j, the server j dedicates all of its processing power to B if the workload in B is not 0. The customers in $\mathbb{R}^d \backslash B$ are served if there is no customer in B. Thus the statement of the lemma follows from the usual formula for the G/G/1 queue. $\qquad\square$

Let $M_t = W_t^0 \circ \theta_{-t}$ be the Loynes sequence under policy π and $\tilde{M}_t^j = \tilde{W}_t^0 \circ \theta_{-t}$.

$$M_t(B) = \sum_j \int_B r_j(x) \tilde{M}_t^j(dx)$$

. Note from Lemma 2 that $\tilde{M}_t^j(B)$ is a non-decreasing sequence for a *j-max* set.

Lemma 3. *As t tends to infinity, $\tilde{M}_t^j$ couples a.s. with $\tilde{M}_\infty^j$, a finite random measure.*

Proof. Let B be a j-max set, from the Birkhoff's Theorem a.s. :

$$\lim_{t \to +\infty} \frac{1}{t} \int_{-t}^0 \int_B \frac{f_j(x)\sigma_{s,x}}{r_j(x)} A(ds \times dx) = \int_B \frac{f_j(x)E_A^{0,x}(\sigma_{0,x})}{r_j(x)} \lambda(dx) < 1.$$

Therefore : $\tilde{M}_t^j(B) = (\sup_{0 \le h \le t} \int_{-h}^0 \int_B \frac{f_j(x)\sigma_{s,x}}{r_j(x)} A(ds \times dx) - h)^+$ is a bounded increasing sequence and it couples with $\tilde{M}_\infty^j(B) = (\sup_{h \in \mathbb{R}_+} \int_{-h}^0 \int_B \frac{f_j(x)\sigma_{s,x}}{r_j(x)} A(ds \times dx) - h)^+$.

It remains to prove the result for any Borel set. $\mathbb{R}^d$ is a *j-max* set and we have for any set B : $\tilde{M}_t^j(B) \le \tilde{M}_t^j(\mathbb{R}^d) \le \tilde{M}_\infty^j(\mathbb{R}^d)$. Therefore a subsequence $\tilde{M}_{t_n}^j(B)$ converges to $\tilde{M}_\infty^j(B)$ and $\tilde{M}_\infty^j$ is well defined and it is a finite random measure.

Since $\tilde{W}_t^j(\mathbb{R}^d)$ is a G/G/1 queue, we can define T the first time after the coupling time with $\tilde{M}_\infty^j \circ \theta_t(\mathbb{R}^d)$ such that $\tilde{W}_t^j(\mathbb{R}^d) = 0$. T is a.s. finite. In view of Equation (11), for $t \ge T$, $\tilde{W}_t^j = \tilde{M}_\infty^j \circ \theta_t$. Thus : $\tilde{M}_t^j = \tilde{W}_t^j \circ \theta_{-t} = \tilde{M}_\infty^j$ and $\tilde{M}_t^j$ couples with $\tilde{M}_\infty^j$. $\square$

Let B be a bounded set. For our specific choice of f, given by Proposition 6 :

$$M_t(B) = \sum_{j \in \mathcal{J}_B} \int_B r_j(x) \tilde{M}_t^j(dx), \tag{13}$$

Since $|\mathcal{J}_B|$ is finite, we deduce from Lemma 3 that $M_t(B)$ couples a.s. with $M_\infty(B) = \sum_{j \in \mathcal{J}_B} \int_B r_j(x) \tilde{M}_\infty^j(dx)$.

We thus have proved the existence of the limit :

$$\lim_{t \to \infty} M_t = M_\infty = \sum_j \int_B r_j(x) \tilde{M}_\infty^j(dx).$$

To conclude the proof, it remains to prove that M_∞ is a stationary solution of Equation (6) for a policy π'.

Along the line of Lemma 3, we can also prove that the process $\{\tilde{M}_{t+s}^j \circ \theta_s\}_{0 \le s \le h}$ couples with $\{\tilde{M}_\infty^j \circ \theta_s\}_{0 \le s \le h}$ for any positive h.

From Equation (12), for a bounded set B :

$$
\begin{aligned}
\tilde{M}_{t+h}^j \circ \theta_h(B) &= \int_{-t-h}^0 \int_B \frac{f_j(x)\sigma_{s,x}}{r_j(x)} A(ds \times dx) - \int_0^{t+h} \int_B \pi_j(\theta_{-t}, s)(dx) ds \\
&= \tilde{M}_t^j(B) + \int_0^h \int_B \frac{f_j(x)\sigma_{s,x}}{r_j(x)} A(ds \times dx) - \int_0^h \int_B \pi_j(\theta_{-t}, t+s)(dx) ds.
\end{aligned}
$$

If Z is the maximum of coupling time of $\{\tilde{M}^j_{t+s} \circ \theta_s(B)\}_{0 \le s \le h}$, for $t \ge Z$ we deduce :

$$\int_0^h \int_B \pi_j(\theta_{-t}, t+s)(dx)ds = \tilde{M}^j_\infty \circ \theta_h(B) - \tilde{M}^j_\infty(B) + \int_0^h \int_B \frac{f_j(x)\sigma_{s,x}}{r_j(x)} A(ds \times dx).$$

For all h, the right hand side of the last equation does not depend on t, thus $\pi_j(\theta_{-t}, t+s)$ couples a.e. with a measure $\pi'_j(s) = \lim_t \pi_j(\theta_{-t}, t+s)$. Let t_0 such that the coupling occurs, we have $\pi'_j(s) = \lim_t \pi_j(\theta_{-t+s-t_0} \circ \theta_s, t + t_0 - s + s) = \pi'_j(t_0) \circ \theta_{s-t_0}$.

In consequence, π' is a stationary policy and M_∞ is a stationary solution of Equation (6).

Theorem 1 is proved.

Remark 1. *In the particular case described in Example 1, a simpler proof is available. We have :* $\sum_i \frac{\lambda_i E_A^{0,x_i}(\sigma_{0,x_i}) f_j(x_i)}{r_j(x_i)} = \rho_j < 1$. *Consider the following deterministic policy defined for atomic measure with atoms on $\{x_i\}, i \in \mathbb{N}$:*

$$\pi_j(m)(\{x_i\}) = \epsilon_j \mathbb{1}(m(\{x_i\}) \ne 0) \frac{f_j(x_i)\lambda_i E_A^{0,x_i}(\sigma_{0,x_i})}{r_j(x_i)\rho_j}.$$

When computing $M_t(\{x_i\})$, it appears that this policy is stable.

3   Convergence toward a Stationary Solution

When there exists a stationary regime in a queuing system it is an important issue to know if for any initial condition the workload converges in some sense to the stationary regime. The following proposition gives a positive answer for the policy defined in the proof of Theorem 1.

Proposition 1. *If the policy scheme defined by Equation (10) is enforced, for any finite initial workload at time $t = 0$, for all bounded Borel set B, $\{W_{t+T}(B)\}, t \in \mathbb{R}_+$, converges in variation toward $\{M \circ \theta_t(B)\}, t \in \mathbb{R}_+$, as T tends to infinity.*

Note that the workload measure does not converge in variation, convergence happens only on bounded sets. The proposition states that the workload converges in variation for the vague topology.

Proof. The proof relies on the following fact : if a stochastic process $\{X_t\}$ couples with $\{Y \circ \theta_t\}$ then $\{X_{t+T}\}, t \in \mathbb{R}_+$, converges in variation toward $\{Y \circ \theta_t\}, t \in \mathbb{R}_+$, as T tends to infinity (see Lindvall [9] or Theorem 2.4.1. of [2]).

From Lemma 2, $\tilde{W}^j_t(B)$ is G/G/1 queue, therefore the coupling of $\tilde{W}^j_t(B)$ for any initial condition follows from Property 2.4.1. in [2].

For a general Borel set B, it suffices to notice as already done that $\tilde{W}^j_t(B) \le \tilde{W}^j_t(\mathbb{R}^d)$. The same arguments used in the proof of Theorem 1 works to show that $\tilde{W}^j_t(B)$ couples for any initial condition.

If B is a bounded set, $W_t(B) = \sum_{j \in \mathcal{J}_B} \int_B r_j(x)\tilde{M}^j_t(dx)$. Since $|\mathcal{J}_B|$ is finite, the coupling also occurs. $\qquad\qquad\square$

4 Monotonous Policies

The policy we have defined to prove the sufficiency part of Theorem 1 is not of particular interest. In particular, it requires the knowledge of the mappings $f_j(x)$. Along the line of the work done in [11], [1], [3], it would be very appealing to find some stable policy which does not rely on the knowledge of the parameters of the system.

4.1 Loynes' Sequence and Loynes' Variable

We are now going to use the key ideas of the proof of Loynes for general Polish space (see paragraph 2.5.2. of [2]). The space $\mathcal{M}$ is a Polish space for the vague topology (refer to [7]).

We define a partial order $\preceq$ on $\mathcal{M}$, the set of Radon measures on $\mathbb{R}^d$. Let m and m' be two Radon measures, $m \preceq m'$ if for all bounded Borel set B in $\mathbb{R}^d$, $m(B) \le m'(B)$.

Lemma 4. *Let $(m_n, n \in \mathbb{N})$ be a monotonous sequence in $\mathcal{M}$ for $\preceq$, such that for all bounded Borel sets B, $m_n(B)$ is bounded. Then, $(m_n, n \in \mathbb{N})$ converges in $\mathcal{M}$ for the vague topology.*

Proof. Let f be a bounded continuous function vanishing outside a bounded set. Let $f_+ = \max(0, f)$ and $f_- = \min(0, f)$, it is straightforward to check that $(m_n(f_+), n \in \mathbb{N})$ and $(m_n(f_-), n \in \mathbb{N})$ are bounded monotonous sequences. Thus $(m_n(f), n \in \mathbb{N})$ converges. □

The Loynes sequence $(M_t, t \in \mathbb{R}_+)$ is defined as the workload found at time 0 supposing that the workload was zero at time $-t$, that is :

$$M_t = W_t^0 \circ \theta_{-t}.$$

The Loynes sequence is of particular interest for the class of monotonous policies.

A policy is said to be *monotonous* if $m' \preceq m$ implies for all $t \in \mathbb{R}_+$, $W_t^{m'} \preceq W_t^m$. Monotonous policies are quite natural in our queuing setting.

Lemma 5. *Suppose π is a monotonous policy then $\{M_t\}, t \in \mathbb{R}_+$, is non-increasing sequence (for $\preceq$).*

Proof. Let $h \in \mathbb{R}_+$, we have : $0 = W_0^0 \circ \theta_{-t} \preceq W_h^0 \circ \theta_{-t-h}$, from the monotonicity of $m \mapsto W_t^m$, we deduce that : $M_t = W_t^0 \circ \theta_{-t} \preceq W_t^{W_h^0 \circ \theta_{-t-h}} \circ \theta_{-t} = M_{t+h}$. □

If $(M_t), t \in \mathbb{R}_-$, is an non-decreasing sequence and is bounded by a random Radon measure Z (for $\preceq$), using Lemma 4, M_t converges a.s. in $\mathcal{M}$ and we can then define the Loynes variable as :

$$M_\infty = \lim_{t \to +\infty} M_t.$$

4.2   Monotonous Left Continuous Policies

In this paragraph, we gives some conditions to ensure that a stationary solution exists. To ensure that the Loynes' variable is a stationary solution we need some continuity properties of π_j.

Definition 4. *Let f be a measurable mapping from $\mathcal{M}$ to $\mathcal{M}$. We say that f is left continuous (resp. right continuous) if for all non-decreasing (resp. non-increasing) converging sequence $(m_n), n \in \mathbb{N}$, of $\mathcal{M}$, $\lim_n f(m_n) = f(\lim_n m_n)$.*

It is consistent to define some continuity properties with converging sequences since $\mathcal{M}$ is a complete metric space. Note also that a mapping can be left and right continuous without being continuous. This is due to the fact that $\preceq$ is only a partial order.

Right continuous policies cannot be relevant. When a new customer arrives, the policy can change drastically. In particular remark that a work-conserving policy cannot be right continuous.

We define the *discontinuity set* of a mapping $h : disc(f)$ as $\{x : h(x) \text{ not continuous at } x\}$.

Proposition 2. *Suppose π_j is left continuous and $\lambda(disc(r_j)) = 0$ for all $j \in \mathcal{J}$. When M_∞ is finite, it is a stationary solution of Equation (6).*

Proof. By definition, $W_s^{M_t} = W_{t+s}^0 \circ \theta_{-t} = W_{t+s}^0 \circ \theta_{-s-t} \circ \theta_s = M_{t+s} \circ \theta_s$. Therefore, from Equation (6), we have, for a Borel set B and $t \in \mathbb{R}^+$:

$$
\begin{aligned}
M_{t+h} \circ \theta_h(B) &= W_h^{M_t}(B) \\
&= M_t(B) + \int_0^h \int_B \sigma_{s,x} A(ds, dx) \\
&\quad - \sum_j \int_0^h \int_B r_j(x) \pi_j(M_{t+s} \circ \theta_s)(dx) ds. \qquad (14)
\end{aligned}
$$

When $(t_k), k \in \mathbb{N}$, be an increasing sequence converging toward $+\infty$. $(M_{t_k+s} \circ \theta_s), k \in \mathbb{N}$, is a non-decreasing sequence converging toward $M_\infty \circ \theta_s$. Since π_j is left continuous, we have $\lim_{k \to +\infty} \pi_j(M_{t_k+s} \circ \theta_s) = \pi_j(M_\infty \circ \theta_s)$ for the vague convergence.

For a fixed ω, the arrival point process is $A(\omega) = \sum_n \delta_{T_n(\omega), X_n(\omega), \sigma_n(\omega)}$. We define $C(\omega) = \{X_n(\omega), n \in \mathbb{N}\}$ and let B be a bounded Borel set such that $C(\omega) \cap \partial B = \emptyset$. ∂B avoids a countable set of points in $\mathbb{R}^d$.

From Equation (6) and Condition (5), $M_\infty \circ \theta_s$ and $\pi_j(M_\infty) \circ \theta_s$ are atomic measures and they are supported by $\sum_k \delta_{X_k} \mathbb{1}(T_k < s)$. Thus, for B as above, for all s in $\mathbb{R}$, we have $\pi_j(M_\infty \circ \theta_s)(\partial B) = 0$. Moreover, since $\lambda(disc(r_j)) = 0$, P-almost surely $\pi(M_\infty \circ \theta_s)(disc(r_j)) = 0$. From Lemma 6 (in Appendix), we deduce P-almost surely :

$$
\lim_{k \to +\infty} \int_B r_j(x) \pi_j(M_{t_k+s} \circ \theta_s)(dx) = \int_B r_j(x) \pi_j(M_\infty \circ \theta_s)(dx).
$$

Now, from Hypothesis (3), $\sum_j \int_0^h \int_B r_j(x)\pi_j(M_{t+s} \circ \theta_s)(dx) \leq \sum_j h \sup_{x\in B} r_j(x) < +\infty$. The dominated convergence theorem gives, P-almost surely :

$$M_\infty \circ \theta_h(B) = M_\infty(B) + \int_0^h \int_B \sigma_{s,x} A(ds, dx) - \sum_j \int_0^h \int_B r_j(x)\pi_j(M_\infty \circ \theta_s)(dx).$$

From Lemma 7 (in Appendix), this last equation is indeed satisfied for all Borel sets. Therefore, M_∞ is a stationary solution. $\qquad\square$

The hypothesis can be made more precise. Consider $E = \{\omega \in \Omega : \exists T, \forall t > T\ M_t(\omega) = M_\infty(\omega)\}$. On E, M_t converges in variation (or couples).

We have already noticed that : $M_{t+s}(\theta_s\omega) = W_s^{M_t}(\omega)$. Thus, if $M_t(\omega) = M_\infty(\omega)$, for $t > T$, we deduce that $M_t(\theta_s\omega) = W_s^{M_{t-s}}(\omega) = M_\infty(\theta_s\omega)$ for $t > T + s$. Thus E is a θ_s-invariant event and from ergodicity, $P(E) \in \{0, 1\}$.

If $P(E) = 1$, M_t couples with M_∞ a.s.. It follows that the hypothesis on the continuity of π_j and r_j are not needed. We deduce the corollary :

Corollary 1. *For a given policy π, if $M_t = W_t^Z \circ \theta_{-t}$ couples with M_∞ then M_∞ is a stationary workload solution of Equation (6).*

Example 4. Cone policies. An interesting class of policies has emerged in papers, see [11], [1], [3]. Let $\alpha > 0$, we define for an atomic policy m :

$$A_j(m) = \arg\max\left\{x : m(\{x\})^\alpha r_j(x)\right\}$$

$$\pi_j(m) = \begin{cases} C(m)\epsilon_j \sum_{x\in A_j(m)} r_j(x)^{-\frac{\alpha+1}{\alpha}}\delta_x & \text{if} \quad m \neq 0 \\ 0 & \text{if} \quad m = 0, \end{cases}$$

where $C(m)$ is the constant such that $\int_{\mathbb{R}^d} \pi_j(m)(dx) = \epsilon_j$ if $m \neq 0$.

Notice that $\pi(cm) = \pi(m)$ for $c > 0$. For finite workload measures, it can be shown that this policy is monotonous and left-continuous. However, it is not clear whether or not this policy is stable when $A \in \mathcal{N}^s$.

5 Optimal Spatial Allocation

In this section, we suppose that the spatial arrival intensity $\lambda(dx)$ is absolutely continuous with respect to the Lebesgue measure : it can be written as $\lambda(x)dx$.

We have seen in Theorem 1 that the stability of system relies on :

$$\rho = \inf_{f\in\mathcal{F}} \sup_{j\in\mathcal{J}} \frac{1}{p_j} \int_{\mathbb{R}^d} \frac{E_A^{0,x}(\sigma_{0,x})f_j(x)}{r_j(x)}\lambda(x)dx. \tag{15}$$

If $\rho < 1$, the system is stable, if $\rho > 1$, the system is unstable. In this section, we analyse this optimization problem. To this end, we define :

$$\rho_j(f) = \frac{1}{p_j} \int_{\mathbb{R}^d} \frac{E_A^{0,x}(\sigma_{0,x})f_j(x)}{r_j(x)}\lambda(x)dx \quad \text{and} \quad \rho(f) = \sup_j \rho_j(f).$$

$\mathcal{F}$ is a convex closed set and $f \rightarrow \rho(f)$ is a convex function, thus the minimum of Equation (15) is reached. In this section, we take interest to the optimal subset of $\mathcal{F}$ defined as :

$$\mathcal{F}^* = \{f \in \mathcal{F} : \rho(f) = \rho\}.$$

The extremal points of the convex set $\mathcal{F}$ are the measurable functions such that $f_j(x) = \mathbb{1}(x \in V_j)$, for a Borel set V_j. This class of function is called *tessellation*. A tessellation is a partition of the space : each point $x \in \mathbb{R}^d$ is affiliated to a unique base station j.

We know that there is a mapping from an $(f_j)_j$ to a stable policy π. Therefore, in Equation (15) we are looking for a policy scheme which maximizes the intensity of arrival. The policy scheme which corresponds to a tessellation is a cellular type policy : a customer is served only by one base station.

Proposition 3. *If ρ is finite there is a f such that :*

$$\forall j, \quad \rho_j(f) = \rho.$$

If there is a finite number of server stations, all $f \in \mathcal{F}^*$ satisfy the above equation.

Proof. Let $f \in \mathcal{F}$ and suppose for example, $\rho_1(f) < \rho_2(f)$, since, $\rho_2(f) > 0$, f_2 is not a.e. equal to 0. Thus, there exists a measurable non negative function $x \mapsto \epsilon(x)$ such that $f_2^\epsilon(x) = f_2(x) - \epsilon(x) \geq 0$, $f_1^\epsilon(x) = f_1(x) + \epsilon(x) \leq 1$ and $\rho_2(\epsilon) > 0$. Let $f_j^\epsilon(x) = f_j(x)$, for $j \notin \{1,2\}$. $f^\epsilon \in \mathcal{F}$ and we have $\rho_1(f^\epsilon) = \rho_1(f) + \rho_1(\epsilon)$ and $\rho_2(f^\epsilon) = \rho_2(f) - \rho_2(\epsilon)$. Thus for $\rho_1(\epsilon)$ small enough, $\sup_{j \in \{1,2\}} \rho_j(f^\epsilon) < \sup_{j \in \{1,2\}} \rho_j(f)$ and $\rho(f^\epsilon) \leq \rho(f)$.

Suppose now that $f \in \mathcal{F}^*$, then f^ϵ is also in $\mathcal{F}^*$. By iterating the construction above for all j, j', such that $\rho_{j'}(f) < \rho_j(f)$, the proposition follows. □

Proposition 3 as an intuitive meaning : for an optimal spatial allocation, the traffic load is the same on each server station. Along the same line, we can prove a more surprising result.

Definition 5. *The processing rates are said to be* singular *if there exist j and k in $\mathcal{J}$, a constant $C > 0$ and a Borel set A of positive Lebesgue measure such that :*

$$\forall x \in A, \; r_j(x) = Cr_k(x).$$

Proposition 4. *Suppose ρ is finite. If the processing rates are not singular, then there is an $f \in \mathcal{F}^*$ which is a tessellation.*

If there is finite number of base stations, all $f \in \mathcal{F}^*$ are tessellations.

This proposition gives a counter-intuitive result : the server stations do not need to share the jobs to reach the stability region. However, note that all the difficulty is to find the optimal tessellation which reaches the optimum. This result is not very surprising from the point of view of convex optimization : it only asserts that the extremum is reached at an extremal point.

Definition 5 is purely technical and does not relies on a natural assumption on the processing rates. In the wireless scenario (Example 1.4) and $r_j(x) = f(x)l(|Y_j - x|)$, if $Y_j \neq Y_k$ for all $j, k \in \mathcal{J}$ and l is a strictly convex mapping then the processing rates are non-singular.

Proof. We consider the $f \in \mathcal{F}^*$ given by proposition 3. Let $E = f_1(]0,1[)^{-1} \cap f_2(]0,1[)^{-1}$. In this proof, μ will denote the Lebesgue measure. We want to show that $\mu(E) = 0$.

Suppose instead that $\mu(E) > 0$, we can suppose without loss of generality that $\mu(E) < +\infty$. Let A, B be disjoint compact sets of positive Lebesgue measure included in E, these sets exist in view of Theorem 2.14 in Rudin [10] (Riesz Representation Theorem). We consider the mapping $\phi(x) = \mathbb{1}(x \in A) - \nu\mathbb{1}(x \in B), \nu > 0$.

Let $f_1^\epsilon(x) = f_1(x) + \epsilon\phi(x)$, $f_2^\epsilon(x) = f_2(x) - \epsilon\phi(x)$ and $f_i^\epsilon(x) = f_i(x)$ for $i \notin \{1,2\}$. If $\epsilon > 0$ is small enough, f^ϵ and $f^{-\epsilon}$ are in $\mathcal{F}$ and for $i \in \{1,2\}$:

$$\rho_i(f^{\pm\epsilon}) = \rho_i(f) \pm \epsilon\rho_i(\phi) = \rho \pm \epsilon\rho_i(\mathbb{1}_A) \mp \nu\epsilon\rho_i(\mathbb{1}_B).$$

Since $f \in \mathcal{F}^*$, $\max(\rho_1(f^{\pm\epsilon}), \rho_2(f^{\pm\epsilon})) \geq \rho$ and we deduce that $sign(\rho_1(\phi)) = sign(\rho_2(\phi))$, where $sign$ is the sign function ($sign(0) = 0$, and for $x \neq 0$, $sign(x) = \frac{x}{|x|}$). It follows that for all real ν, $\rho_1(\mathbb{1}_A) - \nu\rho_1(\mathbb{1}_B)$ and $\rho_2(\mathbb{1}_A) - \nu\rho_2(\mathbb{1}_B)$ have the same sign. Therefore the vector $(\rho_1(\mathbb{1}_A), \rho_1(\mathbb{1}_B))$ and $(\rho_2(\mathbb{1}_A), \rho_2(\mathbb{1}_B))$ are colinear : exists $C_{A,B}$ such that : $\rho_1(\mathbb{1}_A) = C_{A,B}\rho_2(\mathbb{1}_A)$: $C_{A,B}$ cannot depend on B and by symmetry does not depend neither on A. Thus, exists $C > 0$ such that :

$$\rho_1(\mathbb{1}_A) = C\rho_2(\mathbb{1}_A).$$

This last equality has been proved for any compact set included in E. From Theorem 2.14 in [10], it can be extended to any Borel set included in E. Thus, for all A included E, such that $\mu(A) > 0$, $\frac{1}{\mu(A)}\int_A \frac{E_A^{0,x}(\sigma_{0,x})}{p_1 r_1(x)} - C\frac{E_A^{0,x}(\sigma_{0,x})}{p_2 r_2(x)})dx = 0$. We can apply Theorem 1.40 of [10] and conclude that a.e. in E :

$$C' r_1(x) = r_2(x).$$

This contradicts our hypothesis on the processing rates. Therefore $\mu(E) = 0$. and we have proved that there exists an f in $\mathcal{F}^*$ such that a.e. $f_j(x) = \mathbb{1}(x \in V_j)$ and $V_j \cap V_k$ is a set of measure 0. Since for two functions equal a.e., $\rho_j(f)$ has the same value, $(\mathbb{1}(V_j))_{j\in\mathbb{N}}$ is a tessellation in $\mathcal{F}^*$. □

6 Cellular Policies

Definition 6. *Let* $\{V_j\}, j \in \mathcal{J}$, *be a tessellation, a* cellular policy *with cells* $\{V_j\}_j$ *is a policy scheme satisfying for all* j :

$$\forall m \in \mathcal{M}, \quad \pi_j(m)(\mathbb{R}^d \backslash V_j) = 0.$$

We have seen in Proposition 4, that cellular policies reach the stability region under some assumptions. In this section, we focus on this subclass of policies.

For a cellular policy, the server stations are not sharing the jobs which are divided among them. We say that a cellular policy is *work-conserving* if $m(V_j) > 0$ implies $\pi_j(m)(V_j) = 1$.

Proposition 5. *Let* $\{V_j\}, j \in \mathcal{J}$, *be a tessellation with bounded sets,* any work conserving cellular policy *with cells* $\{V_j\}_j$ *is stable if* :

$$\forall j, \quad \int_{V_j} \frac{E_A^{0,x}(\sigma_{0,x})}{r_j(x)} \lambda(dx) < p_j.$$

If there is a j *such that* : $\int_{V_j} \frac{E_A^{0,x}(\sigma_{0,x})}{r_j(x)} \lambda(dx) > p_j$ *then* any cellular policy *with cells* $\{V_j\}_j$ *is unstable.*

This proposition is along the line of the result on single server queue which asserts that the stability does not depend on the discipline, provided it is work-conserving.

Since there is no interactions between server stations when a cellular policy is enforced, to prove the proposition, it is sufficient to prove the following result :

Suppose that the intensity measure $\lambda(dx)$ *is finite and that there is a unique server. Then if* $A \in \mathcal{N}^s$, *any work conserving policy is stable, if* $A \notin \bar{\mathcal{N}}^s$ *any working policy is unstable.*

This result on multi-class queues is mainly known. A proof is given in Appendix.

7 Conclusion

In this paper, we have derived some basic results on a spatial queuing system. An important aspect of this model is that the workload is an atomic measure with an infinite total mass and not a vector in $\mathbb{R}_+^n$. Another feature is the infinite number of server stations but this was more easy to handle.

With stronger hypothesis on the arrival point process A, it could be possible to derive more precise results on the stability and on the law of the stationary workload measure.

As already pointed out, Theorem 1 does not provide an interesting stationary policy. It would be interesting to find other types of policies which are stable when the arrival point process is in $\mathcal{N}^s$. In this scope, deriving the stability region of cone policies described in example 4.2 is a challenging work.

Appendix : Property of the Stability Set

In this section, we find some properties on the set $\mathcal{N}^s$ as it is defined in Section 1.5.

To simplify notations, $\nu_j(dx)$ will denote the measure $\frac{E_A^{0,x}(\sigma_{0,x})}{r_j(x)}\lambda(dx)$ and $\mathcal{J}$ is set to be $\mathbb{N}$.

For $x \in \mathbb{R}^d$, let $\mathcal{J}_x = \{j : f_j(x) > 0\}$ and for a set B, $\mathcal{J}_B = \cup_{x \in B} \mathcal{J}_x$.

Proposition 6. *If $\mathcal{N}^s$ is not empty, there is a $f \in \mathcal{F}$ such that for all j, $x \mapsto f_j(x)$ is continuous and for all bounded set B, $|\mathcal{J}_B|$ is finite.*

Proof. Let $f \in \mathcal{F}$ such that $\forall j$, $\rho_j = \int_{\mathbb{R}^d} f_j(x)\nu_j(dx) < 1$. Let G be a open bounded set.

For all j, from Lusin's Theorem (see for example [10]) there exists a sequence of continuous functions on G, $g_j^n(\bullet)$, such that $\nu_j(dx)$-a.e., $0 \le g_j^n(x) \le f_j(x)$ and $\lim_n \int_G g_j^n(x)\nu_j(dx) = \int_G f_j(x)\nu_j(dx)$.

We have $\sum_j g_j^n(x) \le 1 = \sum_j f_j(x)$. Let $\epsilon^n(x) = 1 - \sum_{j \neq 1} g_j^n(x)$. We define

$$f_j^n(x) = \left\{ \begin{array}{ll} g_j^n(x) + \mathbb{1}_{j=1}\epsilon^n(x) & \text{if} \quad x \in G, \\ f_j(x) & \text{if} \quad x \notin G, \end{array} \right.$$

We check that f^n is in $\mathcal{F}$ and by the dominated convergence theorem, for n large enough : $\forall j$, $\int_{\mathbb{R}^d} f_j^n(x)\nu_j(dx) < 1$.

By iterating this construction for a set of open sets covering the $\mathbb{R}^d$, we deduce that there exists and $f \in \mathcal{F}$, such that $\forall j$, $\rho_j = \int_{\mathbb{R}^d} f_j(x)\nu_j(dx) < 1$ and f_j continuous.

Now we turn to the second part of the proposition. Let K be a compact subset, f_j as above and $\max(\rho_j, \frac{1}{2}) < \rho_j' < 1$. We define $g_j(x) = \frac{f_j(x)}{\rho_j'}$.

Then : $\int_{\mathbb{R}^d} g_j(x)\nu_j(dx) = \frac{\rho_j}{\rho_j'} < 1$ and $\sum_j g_j(x) > 1$. For all x, there exists j_x such that $\sum_{j=1}^{j_x} g_j(x) > 1$. By continuity, since K is compact, $j_K = \sup_{x \in K} j_x$ is finite.

It is immediate to check that $\tilde{f}_j(x) = \frac{g_j(x)\mathbf{1}_{j \le j_x}}{\sum_{j \le j_x} g_j(x)}$ has all the required properties. $\square$

Appendix : Vague Convergence in $\mathcal{M}$

In this paragraph, we give some technical results used in Section 4.

The following lemma is an adaptation of Theorem 5.2 of Billingsley [4] to the vague topology.

Lemma 6. *Let m_n be a sequence in $\mathcal{M}$ converging for the vague topology toward m. Let h be a bounded measurable function which $m(disc(h)) = 0$ and B a bounded Borel set of $\mathbb{R}^d$ with $m(\partial B) = 0$, then : $\lim_n \int_B h(x)m_n(dx) = \int_B h(x)m(dx)$*

Lemma 7. *Let C be a countable set of points in $\mathbb{R}^d$ and let $\mathcal{C}$ the set of bounded Borel sets of $\mathbb{R}^d$ with $C \cap \partial B = \emptyset$. Then :*

 - $\mathcal{C}$ is an algebra and the σ-algebra generated by $\mathcal{C}$, $\sigma(\mathcal{C})$ is the Borel σ-algebra $\mathcal{B}$.

- *From a measure m defined on $\mathcal{C}$, there is a unique extension to m on $\mathcal{B}$.*

Proof. From the relations : $\partial(A \cap B) \subset \partial A \cap \partial B$, $\partial(A \cup B) \subset \partial A \cup \partial B$ and $\partial(A^c) = \partial A$. It follows that $\mathcal{C}$ is an algebra on the set of sets of $\mathbb{R}^d$.

$\mathcal{B}$ is the σ-algebra of the algebra generated by the open rectangles of $\mathbb{R}^d$. To prove the first assertion of our lemma, it suffices to prove that the rectangle $]0, 1[^d$ can be written as $\cup_{n\in\mathbb{N}} B_n$, where $B_1 \subset B_2... \subset B_n \in \mathcal{C}$. To this end, consider the rectangle $R_\epsilon =]\epsilon, 1 - \epsilon[^d$, with $0 < \epsilon < \frac{1}{2}$. If $\epsilon \neq \epsilon'$, $\partial R_\epsilon \cap \partial R_{\epsilon'} = \emptyset$. Since C is countable there can only be a countable set of ϵ such that $C \cap \partial R_\epsilon \neq \emptyset$. In particular, there exists an increasing sequence ϵ_n such that $C \cap \partial R_{\epsilon_n} = \emptyset$. We have proved the first statement of the lemma.

The second assertion is the Caratheodory Extension Theorem.   □

Appendix : Spatial queuing system with one server

When there is only one server in the system, Theorem 1 can be made more precise. The system reduces to a multi-class queue. The condition $\lambda \in \mathcal{N}^s$ can be restated as :

$$\int_{\mathbb{R}^d} \frac{E_A^{0,x}(\sigma_{0,x})}{r(x)} \lambda(dx) < p,$$

where r is the processing rate for the server providing service for a user located at x and p is the expectation of the available processing power.

We have the following proposition :

Proposition 7. *Suppose that the intensity measure $\lambda(dx)$ is finite and that there is a unique server. Then if $A \in \mathcal{N}^s$, any work conserving policy is stable.*

Proof. We will only give a sketch of the proof in three steps. We will suppose that $\epsilon(t) = 1$ (hence $p = 1$). The proof in the general case follows with obvious modifications.

Step 1. Since λ is a finite Radon measure, $N = \sum_n \delta_{T_n, \frac{\sigma_n}{r(X_n)}}$ is a simple marked point process on $\mathbb{R}$ with finite intensity $E(N[0, 1]) = \int_{\mathbb{R}^d} \lambda(dx)$.

For a given work conserving policy π, define $Y_t = \int_{\mathbb{R}^d} \frac{W_t(dx)}{r(x)}$. From Equation (6), we deduce that for $t \in [T_n, T_{n+1})$:

$$Y_t = \left(Y(T_n-) + \frac{\sigma_n}{r(X_n)} + (t - T_n)\right)^+.$$

Therefore, Y_t does not depend on the policy and Y_t is the usual workload for $G/G/1$ queue (in particular $W_t(\mathbb{R}^d)$ is equal to zero at the same time for any work conserving policy). The workload for this queue is equal to $\int_{\mathbb{R}^d} \lambda(dx) E_N^0(\frac{\sigma_0}{r(X_0)}) = \int_{\mathbb{R}^d} \frac{E_A^{0,x}(\sigma_{0,x})}{r(x)} \lambda(dx) < 1$.

Similarly, by denoting $W_n = W_{T_n-}$, it appears that $(W_n), n \geq 0$, is generated by a stochastic recurrence, see [6], [2] (indeed we may write $W_{n+1} = h(W_n, \xi_n)$, for a suitable measurable mapping h and a driving sequence $(\xi_n), n \in \mathbb{N}$, compatible with the flow θ).

Step 2. Let $\mathcal{M}_f$ be the set of atomic measures with a finite set of atoms on $\mathbb{R}^d$. If $W_0 \in \mathcal{M}_f$ then a.s., $W_t \in \mathcal{M}_f$ for $t \geq 0$.

Now, we define the following policy on $\mathcal{M}_f$:

$$
\pi^- : \left\{ \begin{array}{ccc} \mathcal{M}_f & \to & \mathcal{M} \\ m & \mapsto & \left\{ \begin{array}{ccc} \delta_{x^*} & \text{if} & m \neq 0 \\ 0 & \text{if} & m = 0. \end{array} \right. \end{array} \right.
$$

where $x^* = \arg\max \left\{ x : r(x)\mathbb{1}(m(\{x\}) > 0) \right\}$. If multiple choices of x^* are possible, choose the first in the lexicographic order.

π^- is the work conserving policy which dedicates all the processing power to the slowest customer. For π^-, $m \to W_t^m$ is non-increasing : that is, $m' \succeq m$ implies for all $t \geq 0$, $W_t^{m'} \succeq W_t^m$. To prove this last assertion, remark that for $t \in [0, T_1)$ and t small enough :

$$
W_t^{m'} = m' - tr(x^*(m'))\delta_{x^*(m')}
$$

and

$$
W_t^m = m - tr(x^*(m))\delta_{x^*(m)}.
$$

If $m(\{x^*(m')\}) > 0$ then $W_t^{m'} - W_t^m = m' - m \succeq 0$. If $m(\{x^*(m')\}) = 0$ then since $m'(\{x^*(m)\}) > 0$, we have still $W_t^{m'} - W_t^m \succeq 0$.

Let M_t^- be the Loynes'sequence for policy π^- (a.s., $M_t^- \in \mathcal{M}_f$). Let $0 \leq s \leq t$, since $0 = W_0^0 \circ \theta_{-s} \preceq W_{t-s}^0 \circ \theta_{-t}$, from the monotonicity of $m \mapsto W_s^m$, we deduce that : $M_s^- = W_s^0 \circ \theta_{-s} \preceq W_s^{W_{t-s}^0 \circ \theta_{-t}} \circ \theta_{-s} = M_t^-$.

Thus $(M_t^-), t \in \mathbb{R}^+$, is a non-decreasing sequence and for all Borel set B, $M_t^-(B) \leq M_t^-(\mathbb{R}^d)$. $M_t^-(\mathbb{R}^d)$ is the Loynes' sequence for the usual $G/G/1$ queue (from Step 1). We deduce that M_t^- converges a.s. toward the Loynes' variable M_∞^- (from Lemma 4).

Step 3. Consider now any work conserving policy π, we define similarly the Loynes' variable M_t for policy π. As noticed above, π^- is the slowest policy and it follows :

$$
M_t(\mathbb{R}^d) \leq M_t^-(\mathbb{R}^d) \leq M_\infty^-(\mathbb{R}^d).
$$

Consider the event $A = \{M_{T_n}^- = 0\}$. This event is a renovating event for M_{T_n} and since the workload of the $G/G/1$ queue is strictly less than 1, $P_N^0(A) \geq P_N^0(M_\infty^-(\mathbb{R}^d) = 0) > 0$. From Theorem 2.5.3 and Property 2.5.5 of [2], we deduce that M_t converges to a stationary solution M_∞ and that M_t couples with M_∞ (in the strong backward sense). $\square$

Restating Property 2.4.1 of [2], we can also prove that from any finite initial condition m, W_t^m couples with $M_\infty \circ \theta_t$ (as t tends toward $+\infty$).

References

[1] ARMONY, M. AND BAMBOS, N. (2003). Queueing dynamics and maximal throughput scheduling in switched processing systems. *Queueing Syst.* **44,** 209–252.

[2] BACCELLI, F. AND BREMAUD, P. (2003). *Elements of Queuing Theory* second ed. Applications of Mathematics. Springer, Berlin.

[3] BAMBOS, N. AND MICHAILIDIS, G. (2004). Queueing and scheduling in random environments. *Adv. in Appl. Probab* **36,** 293–317.

[4] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley Series in Probability and Mathematical Statistics. Wiley.

[5] BORDENAVE, C. (2004). Stability properties of data flows on a cdma network in macrodiversity. *INRIA research report* RR–5257.

[6] BOROVKOV, A. (1994). *Ergodicity and Stability of Stochastic Processes.* Wiley Series in Probability and statistics. Wiley.

[7] DALEY, D. AND VERE-JONES, D. (1988). *An introduction to the Theory of Point Processes.* Springer Series in Statistics. Springer-Verlag, New-York.

[8] K. WASSERMAN, G. MICHAILIDIS, N. B. (2001). Differntiated processors' scheduling on heterogeneous task flows. *Technical Report SU-Netlab-2001-12/1* Engineering Library, Stanford University.

[9] LINDVALL, T. (1992). *Lectures on the Coupling Method.* Wiley.

[10] RUDIN, W. (1974). *Real and complex Analysis* second edition ed. McGraw-Hill Series in Higher Mathematics. McGraw-HillBook Co.

[11] TASSIULAS, L. AND EPRHEMIDES, A. (1992). Stability properties of constrained queueing systems and scheduling policies for maximum troughput in multihop radio networks. *IEEE Transactions on Automatic Control* **37,** 1936–1948.