Stability of the simultaneous processor sharing model

Cian John Reynolds

Submitted for the degree of
Doctor of Philosophy
on completion of research in the
Department of Actuarial Mathematics and Statistics,
School of Mathematical and Computer Sciences,
Heriot-Watt University

June 2007

Abstract

We study the phenomenon of *entrainment* in processor sharing networks, whereby, while individual network resources have sufficient capacity to meet demand, the requirement for simultaneous availability of resources means that a network may nevertheless be unstable. We show that instability occurs through a poor control strategy, and that, for a variety of network topologies, only small modifications to control strategies are required in order to ensure stability. For control strategies which possess a natural monotonicity property, we give some new results for the classification of the corresponding Markov processes, which lead to conditions both for stability and for instability. Finally, we study the effect of variation of call size distribution on stationary distributions and stability.

To Eileen & John and May & Paddy.

In memory of Martin Lennon.

Acknowledgements

I am most grateful to my two supervisors Stan Zachary and Jennie Hansen. The former for sharing with me his natural intuition for mathematics, during our chats, and for looking after me with respect to non-academic matters and the latter for sharing with me her natural ability for mathematical precision and for patiently teaching me (or reminding me of?!) the basics, especially while she frequently proof-read this thesis. I am also grateful to Terence Chan and Iain MacPhee both of whom took the time to examine the thesis and also gave me some useful comments during the viva voce.

The School of Mathematical and Computer Sciences at Heriot-Watt University generously provided me with funding and office space for the duration of my studies. I would also like to express my gratitude to the former *European Next Generation Internet Network of Excellence* (EuroNGI) for their funding to attend various workshops and summer schools across the European Union. EuroNGI also provided the funding for me to become a Research Associate at Heriot-Watt University for two months for which I am indebted to Serguei Foss for his efforts in securing this funding. I would also like to say thanks to my brother Dáire for providing me with money when all my other sources of funding ran out.

Two of my former teachers deserve to be mentioned for their enthusiasm. Dermot Cook demonstrated to me the usefulness of mathematics and convinced me to study the subject at university level. The late John T. Lewis instilled in me a passion for probability theory through an undergraduate course and by expertly advising me on my undergraduate dissertation.

I would also like to thank all of my office mates (Catherine, Katie, Erik, James, Tom, Ke Li, Mark, Seva and Mike) at Heriot-Watt for the useful chats, friendship and for enduring my constant company for three and a half years. For the same reasons I am appreciative of all the staff and Ph.D. students at the School of Mathematical and Computer Sciences. I am especially grateful to James Kirkby for helping me to write R code when I needed to run simulations.

On a personal level I am thankful to my parents Kathryn and Christopher and to all of my family and friends for their support. I would finally like to thank my girlfriend Kimberley for her encouragement and patience throughout my studies.

Contents

Chapter 1

Introduction

To begin the thesis we define the simultaneous processor sharing model and describe the issue of stability for this model. We report some of the previous work on stability for the simultaneous processor sharing model. Finally, we present an outline for the remaining chapters of the thesis.

1.1 Simultaneous processor sharing model

This thesis is concerned with models for communications networks such as file transfer applications on the Internet. We are interested in models in which "flows" or "calls" have simultaneous capacity requirements from a number of resources, each of which can share capacity between all calls present. It may then happen that while each resource in the network, considered in isolation, has sufficient capacity to service the demand placed on it, the control of the network is such that the requirement for *simultaneous* availability of capacity ensures that over time demand cannot be met, and that the network is unstable, i.e. the number of calls present in the network increases to infinity.

The canonical model for such networks is called the *simultaneous processor sharing model* and is defined as follows. Let $\mathcal{J}$ denote a finite set of *resources* and denote each resource in $\mathcal{J}$ by $j \in \mathcal{J}$. Let $\mathcal{R}$ denote a finite set of *routes* or *call types* and denote each route in $\mathcal{R}$ by $r \in \mathcal{R}$. Let $R = |\mathcal{R}|$ and $J = |\mathcal{J}|$. If a resource j is utilised by a route r we say that $j \in r$. For the sets $\mathcal{J}$ and $\mathcal{R}$ define the *incidence matrix* $A = (A_{jr})_{j\in\mathcal{J},r\in\mathcal{R}}$ by

$$A_{jr} = \begin{cases} 1 & \text{if } j \in r \\ 0 & \text{otherwise .} \end{cases} \tag{1.1}$$

For each $j \in \mathcal{J}$, let c_j denote the *capacity* of resource j where, $0 < c_j \leq \infty$ and let $\boldsymbol{c} = (c_j)_{j\in\mathcal{J}}$ be called the *capacity parameter*. For each $r \in \mathcal{R}$ we have Poisson arrivals which have *input rate*, $\nu_r \geq 0$. The volume of work associated with each call is assumed to be exponentially distributed with mean $\mu_r^{-1} > 0$, and each call size is independent over all other call sizes. Define the vectors $\boldsymbol{\nu} = (\nu_r)_{r\in\mathcal{R}}$ and $\boldsymbol{\mu} = (\mu_r)_{r\in\mathcal{R}}$. For convenience we write $\kappa_r := \nu_r/\mu_r$, $r \in \mathcal{R}$. The quantity κ_r may be interpreted as the rate at which "work" of type r arrives at the system. Let $\boldsymbol{\kappa} = (\kappa_r)_{r\in\mathcal{R}}$ be called the *input parameter*. Many, but not all, important results depend on the parameters ν_r and μ_r only through the corresponding value of κ_r. The *network topology* is specified by the fixed value of the incidence matrix A. We shall call the given fixed values of the parameters $\boldsymbol{\nu}$, $\boldsymbol{\mu}$, $\boldsymbol{c}$ and the matrix $A = (A_{jr})_{j\in\mathcal{J},r\in R}$ the *network* $\mathcal{N}$, with

$$\mathcal{N} = (\boldsymbol{c}, \boldsymbol{\nu}, \boldsymbol{\mu}, A, \mathcal{J}, \mathcal{R}). \tag{1.2}$$

We also define the dimension of $\mathcal{N}$ by

$$\dim \mathcal{N} = R. \tag{1.3}$$

For any time $t \geq 0$, let $\boldsymbol{n}(t) = (n_r(t))_{r\in\mathcal{R}}$, where $n_r(t)$ is the number of *calls* using route r at time t. At any time $t \geq 0$, work of type r is processed at total rate $b_r(\boldsymbol{n}(t))$. Our choice of the function $\boldsymbol{b}$ defined by $\boldsymbol{b}(\boldsymbol{n}) = (b_r(\boldsymbol{n}))_{r\in\mathcal{R}}$ is called the *control strategy* or *bandwidth allocation* for the model. The set of *feasible* control strategies $\boldsymbol{b}$ is defined by the following capacity constraints

$$\sum_{r\in\mathcal{R}} A_{jr} b_r(\boldsymbol{n}) \leq c_j \text{ for all } j \in \mathcal{J}. \tag{1.4}$$

Hence the only rôle that the capacities, $\{c_j : j \in \mathcal{J}\}$, have in this model is as a constraint on the control strategy $\boldsymbol{b}$. Given a control strategy $\boldsymbol{b}$ satisfying condition

(1.4), $\boldsymbol{n}(\cdot) = (n_r(\cdot))_{r\in\mathcal{R}}$ is an irreducible Markov process with state space $\mathbb{Z}_+^R$ (hence the dimension of the process is R), with allowable transitions

$$\boldsymbol{n} \to \begin{cases} \boldsymbol{n} + \boldsymbol{e}_r & \text{at rate} \quad \nu_r, \\ \boldsymbol{n} - \boldsymbol{e}_r & \text{at rate} \quad \mu_r b_r(\boldsymbol{n}) \end{cases} \tag{1.5}$$

for each $r \in \mathcal{R}$, where $\boldsymbol{e}_r = (e_{rs})_{s\in\mathcal{R}}$ is the R-dimensional unit vector with $e_{rr} = 1$ and $e_{rs} = 0$ for all $s \neq r$. We require for each $r \in \mathcal{R}$ that $b_r(\boldsymbol{n}) = 0$ whenever $n_r = 0$. This guarantees that for any $r \in \mathcal{R}$, $n_r(\cdot)$ cannot become negative and ensures that the state space for the process is $\mathbb{Z}_+^R$.

The control strategy $\boldsymbol{b}$ is said to be *stable* if the Markov process $\boldsymbol{n}(\cdot)$ is positive recurrent. Otherwise the control strategy $\boldsymbol{b}$ is said to be unstable. Informally, we can also consider a network to be stable if the number of calls does not increase to infinity. In this thesis we are interested in characterising stable control strategies.

We will be primarily interested in the stability of control strategies which, in addition to satisfying (1.4), are also *Pareto efficient*. A control strategy $\boldsymbol{b}$ is Pareto efficient if, for all $\boldsymbol{n} \in \mathbb{Z}_+^R$ and for all $r \in \mathcal{R}$ such that $n_r > 0$, there exists $j \in \mathcal{J}$ with $A_{jr} = 1$ such that the resource j is *saturated* with respect to $\boldsymbol{b}$, that is

$$\sum_{s\in\mathcal{R}} A_{js} b_s(\boldsymbol{n}) = c_j. \tag{1.6}$$

It should be noted that if a control strategy $\boldsymbol{b}$ is Pareto efficient then it is impossible to increase the bandwidth allocated to a call type without decreasing the bandwidth allocation to some other call type in the network. Otherwise, for at least one resource, (1.4) would be violated and the control strategy would not be feasible. In section 2.3 we show that a necessary and sufficient condition for the existence of a stable control strategy is the constraint

$$\sum_{r\in\mathcal{R}} A_{jr}\kappa_r < c_j \quad \text{for all} \quad j \in \mathcal{J} \tag{1.7}$$

on the input parameter $\boldsymbol{\kappa}$. In fact, we also show in Chapter 2 that the constraint (1.7) is also sufficient for the existence of a stable Pareto efficient control strategy.

Before discussing the simultaneous processor sharing model further and commenting on previous work for our model, it is instructive to consider how our model differs from the well-studied models of *loss networks*. There has been extensive research on loss networks and they are another important model for certain types of communication networks (see for example: Kelly [19]; Ross [30]; and Zachary and Ziedins [36]). A brief definition of loss networks with fixed routing (see also Zachary and Ziedins [36]) is as follows. Let $\mathcal{J}$ be a finite set of resources and let $\mathcal{R}$ be a finite set of call types. Each call type $r \in \mathcal{R}$ has a Poisson input rate $\nu_r \geq 0$ and each call (if accepted by the network) remains for a *holding time* which is exponentially distributed with mean μ_r^{-1} and which a holding time is independent over all other holding times in the network. For any time $t \geq 0$, let $\boldsymbol{n}(t) = (n_r(t))_{r\in\mathcal{R}}$, where $n_r(t)$ is the number of calls of type r in service at time t. The process $\boldsymbol{n}(\cdot)$ is an irreducible Markov process with some state space $\mathcal{M} \subset \mathbb{Z}_+^R$, where $\mathcal{M}$ is defined by a set of resource constraints

$$\sum_{r\in\mathcal{R}} A_{jr} n_r \leq c_j, \quad j \in \mathcal{J} \tag{1.8}$$

where $A_{jr}, c_j \in \mathbb{Z}_+$ for all $j \in \mathcal{J}$ and $r \in \mathcal{R}$, and where c_j denotes the total *capacity* of resource j. We can think of a call of type r as having a simultaneous requirement, for the duration of its holding time, for A_{jr} units of the total capacity c_j for resource $j \in \mathcal{J}$. If a call of type r attempts to enter the network and there is no spare capacity available for calls of type r then the call is *blocked*. Otherwise if there is free capacity available (i.e. constraint (1.8) can be satisfied) then the call is *accepted* into the network.

Both the simultaneous processor sharing model and loss model have calls arriving to certain routes which require simultaneous capacity from one or more resources. However, in a loss network calls are processed (allocated bandwidth) at a fixed rate and while the capacity of any resource is fully utilised, no further calls which use that resource are accepted into network and hence the input rate, ν_r, becomes zero. Hence in a loss network the input rates depend on the state of the Markov process. This is not the case in the simultaneous processor sharing model as every call is accepted into the network, no matter how many calls are currently in the system. Hence the input rates for the simultaneous processor sharing model are constant. For the simultaneous processor sharing model, we typically think of the total capacity allocated to calls of any type as being divided between them (though the dynamics of our Markov process are not sensitive to this assumption). This property is called *processor sharing*. Further, in the simultaneous processor sharing model, capacity is allocated at a flexible rate subject to the control strategy being feasible.

We now consider again the problem of stability in the simultaneous processor sharing network. It might be thought that condition (1.7) would also be sufficient to ensure that any Pareto efficient control strategy $\boldsymbol{b}$ is stable, but this is not the case, except when $J = 1$. Condition (1.7) fails to be sufficient for the stability of Pareto efficient control strategies due to the phenomenon called *entrainment*, whereby the capacity required by calls of a given type is indeed available at each resource but at different times. A consequence of this is that there is not sufficient capacity simultaneously available in the network. The following example shows how a Pareto efficient control strategy may not be stable even when condition (1.7) is satisfied.

Example 1.1.1 Suppose that $\mathcal{R} = \{1,2\}$, $\mathcal{J} = \{1,2\}$ and that the incidence matrix $A = (A_{jr})_{j\in\mathcal{J},r\in\mathcal{R}}$ is given by

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \tag{1.9}$$

Thus the capacity available for calls of type 1 is constrained by resource 1 only, whereas the capacity available for calls of type 2 is constrained by both resources 1 and 2. For simplicity of exposition suppose, that $\mu_1 = \mu_2 = 1$ and suppose also that $c_2 \le c_1$. We assume the condition (1.7) holds, which here becomes

$$\kappa_1 + \kappa_2 < c_1, \qquad \kappa_2 < c_2. \tag{1.10}$$

In figure 1.1.1 — as with all similar figures throughout the thesis which represent a network $\mathcal{N}$ — the lines represent the routes $r \in \mathcal{R}$, with κ_r labelled, and the boxes represent the resources $j \in \mathcal{J}$, with c_j labelled. A line r intersects a box j if and only $A_{jr} = 1$ in the network $\mathcal{N}$.

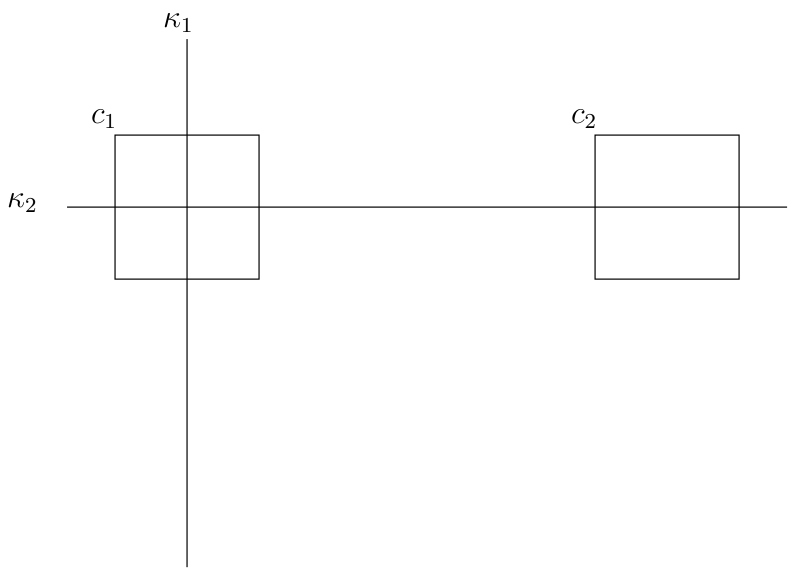

Figure 1.1 The network of Example 1.1.1.

Consider the Pareto efficient control strategy in which complete priority is given to calls of type 1 (i.e. $b_1(\boldsymbol{n}) = c_1$ whenever $n_1 > 0$.). Under this control, the long-run fraction of time in which the network is empty of calls of type 1—and so resource 1 is available for use by calls of type 2—is given by $1 - \kappa_1/c_1$. Since, when resource 1 is available, calls of type 2 are processed at rate c_2, standard arguments for the stability of a single-server queue show that the control strategy is stable if and only if

$$\kappa_2 < c_2(1 - \kappa_1/c_1). \tag{1.11}$$

This is a condition which is generally more restrictive than (1.10) above. When it is violated we have the phenomenon of entrainment referred to above; i.e. by (1.10), each resource in the network, considered in isolation, has sufficient capacity, but the given Pareto efficient control strategy $\boldsymbol{b}$ is nevertheless unstable. More formally, we have a Markov process in the positive quadrant with homogeneous transition rates. The fact that this Markov process is not positive recurrent follows from results for the necessary and sufficient conditions for positive recurrence given in, for example, Fayolle *et al* [12] and Borovkov [7].

Of course there is no issue of instability in loss models, since calls are *blocked* — in the sense that they do not enter the system — when their acceptance would violate the network capacity constraints (1.8). Indeed the interesting problem in a loss model is to evaluate the probability of a call being blocked after it attempts to enter the system. These are called *blocking probabilities.*

The example of entrainment given above shows that the naïve conjecture, that all Pareto efficient control strategies would be stable as long as condition (1.7) holds, is false for the simultaneous processor sharing model. In fact, because of the phenomenon of entrainment it is necessary to show that, under condition (1.7), at least one stable control strategy even exists (see section 2.3). It is also critical to understand what properties of a control strategy $\boldsymbol{b}$ cause entrainment and hence instability of the Markov process $\boldsymbol{n}(\cdot)$. In this thesis we attempt to determine conditions which result in stable control strategies. We further attempt to characterise which properties of control strategies make for stability. It will be shown that, for some simple topologies, any arbitrary Pareto efficient control strategy $\boldsymbol{b}$ can be made stable with only minor adjustments to the control strategy near some boundary of the state space $\mathbb{Z}_+^R$. Wasted capacity (and hence entrainment) usually arises when the Markov process is near the boundary of the state space, i.e. values of $\boldsymbol{n}$ such that $n_r = 0$ for at least one $r \in \mathcal{R}$. However we show in this thesis that for some network topologies this is not the only cause of entrainment and that entrainment can sometimes be caused by other forms of poor control strategy (see section 4.5). It will also be shown that the network topology can have an impact on entrainment.

1.2 Application to the Internet

As previously stated, the simultaneous processor sharing model is a natural model for file transfers on the Internet. The users connecting to the Internet are represented by calls in our terminology while each different type of user is represented by a different call type. The computers that allocate bandwidth to Internet users correspond to the resources, while the control strategy describes the allocation of bandwidth. The capacities of the resources can be viewed as constraints on the total available bandwidth. For example, consider again the Example 1.1.1.

Here calls of type 1 could correspond to all corporate users in a particular city while calls of type 2 could represent all residential users in the same city. This assumption is reasonable as in any city it is often the case that the bandwidth allocated to corporate Internet users is only constrained by the capacity of the city's main "bandwidth hub" while the bandwidth allocated to residential users would be further constrained by the modems or otherwise of each household.

The following table shows some units that can be used to measure the various quantities introduced for the simultaneous processor sharing model.

Quantity	unit
c_j	bits/second
ν_r	connections/second
μ_r^{-1}	bits
κ_r	bits/second
n_r	connections
$b_r(\boldsymbol{n})$	bits/second

Table 1.1 Units for the simultaneous processor sharing model.

When a control strategy for the simultaneous processor sharing network is unstable this means that the total number of calls will increase to infinity. In Internet networks this is, of course, impossible. Instead instability corresponds to downloading times increasing to such an extent that some users will experience disconnections from the network. Therefore for Internet Service Providers the problem of stability relates to the quality of service that they can provide.

1.3 Previous work

Most of the previous work on the problem of stability for the simultaneous processor sharing model and in particular, on the phenomenon of entrainment, has been in the context of *fair-sharing* control strategies (see for example: Bonald and Massoulié [3]; de Veciana *et al* [11]; Kelly *et al* [20]; Kelly and Williams [21]; and Roberts and Massoulié [29]). The notion of fair-sharing characterises how the capacity is allocated at saturated resources. It has been shown that fair-sharing control strategies are characterised by the solution to certain optimisation problems as outlined below. The first such control strategies to be studied are the so called max-min fair-sharing control strategies defined by Bertsekas and Gallager [2]. A control strategy $\boldsymbol{b}$ is said to be (weighted) *max-min fair* if the bandwidth allocated to any call in the network cannot be increased without decreasing the bandwidth to another call in the network which already has a smaller or equal bandwidth allocation. It is easy to see that (weighted) max-min fair control strategies are "fair" by considering the following algorithm which has been shown to yield these control strategies.

Suppose $\mathcal{N}$ is a network. For each $r \in \mathcal{R}$ let w_r be a positive *weight*. Then the weighted max-min fair control strategy $\boldsymbol{b}$ with weights $\{w_r : r \in \mathcal{R}\}$ is computed as follows. For each state $\boldsymbol{n} \in \mathbb{Z}_+^R$, $\boldsymbol{b}(\boldsymbol{n})$ is determined by maximising the function

$$f(\boldsymbol{b}) = \min_{r \in \mathcal{R}} \frac{b_r(\boldsymbol{n})}{w_r n_r} \tag{1.12}$$

subject to the capacity constraints $\sum_{r \in \mathcal{R}} A_{jr} b_r(\boldsymbol{n}) \leq c_j$ for all $j \in \mathcal{J}$ and to the restrictions $b_r(\boldsymbol{n}) \geq 0$ (and $b_r(\boldsymbol{n}) = 0$ if $n_r = 0$) for every $r \in \mathcal{R}$. It has been shown by Gafni and Bertsekas [14] (and outlined by de Veciana *et al* [11]) that for any $\boldsymbol{n} \in \mathbb{Z}_+^R$ the solution to the optimisation problem (1.12) is solved by the following iterative algorithm.

For each $j \in \mathcal{J}$, we define the function $c_j^{(1)}$ on $\mathbb{Z}_+^R$ by

$$c_j^{(1)}(\boldsymbol{n}) = \frac{c_j}{\sum_{r \in \mathcal{R}} A_{jr} w_r n_r}, \qquad \boldsymbol{n} \in \mathbb{Z}_+^R. \tag{1.13}$$

We also define the function $c_{\min}^{(1)}$ on $\mathbb{Z}_+^R$ by

$$c_{\min}^{(1)}(\boldsymbol{n}) = \min_{j \in \mathcal{J}} c_j^{(1)}(\boldsymbol{n}), \qquad \boldsymbol{n} \in \mathbb{Z}_+^R. \tag{1.14}$$

Now for each $\boldsymbol{n} \in \mathbb{Z}_+^R$, let $\mathcal{J}^{(1)}(\boldsymbol{n}) = \{j \in \mathcal{J} : c_j^{(1)}(\boldsymbol{n}) = c_{\min}^{(1)}(\boldsymbol{n})\}$ and let $\mathcal{R}^{(1)}(\boldsymbol{n}) = \{r \in \mathcal{R} : A_{jr} = 1 \text{ for some } j \in \mathcal{J}^{(1)}(\boldsymbol{n})\}$. For each $r \in \mathcal{R}^{(1)}(\boldsymbol{n})$ we let

$$b_r(\boldsymbol{n}) = w_r n_r c_{\min}^{(1)}(\boldsymbol{n}). \tag{1.15}$$

If $\mathcal{R}^{(1)}(\boldsymbol{n}) = \mathcal{R}$ then we have evaluated a bandwidth allocation for each call type

$r \in \mathcal{R}$. Otherwise, let $c_{j,k}$ be equal to the capacity available at resource j at the start of the k-th step in the algorithm. So in particular $c_{j,1} = c_j$ and

$$c_{j,2} = \begin{cases} 0 & \text{for} \quad j \in \mathcal{J}^{(1)}(\boldsymbol{n}) \\ c_{j,1} - c^{(1)}_{\min}(\boldsymbol{n}) \sum_{r \in \mathcal{R}^{(1)}(\boldsymbol{n})} A_{jr} w_r n_r & \text{for} \quad j \in \mathcal{J} \setminus \mathcal{J}^{(1)}(\boldsymbol{n}). \end{cases} \tag{1.16}$$

Now let

$$c^{(2)}_j(\boldsymbol{n}) = \frac{c_{j,2}}{\sum_{r \in \mathcal{R} \setminus \mathcal{R}^{(1)}(\boldsymbol{n})} A_{jr} w_r n_r}, \qquad j \in \mathcal{J} \setminus \mathcal{J}^{(1)}(\boldsymbol{n}). \tag{1.17}$$

and let

$$c^{(2)}_{\min}(\boldsymbol{n}) = \min_{j \in \mathcal{J} \setminus \mathcal{J}^{(1)}(\boldsymbol{n})} c^{(2)}_j(\boldsymbol{n}). \tag{1.18}$$

Let $\mathcal{J}^{(2)}(\boldsymbol{n}) = \{j \in \mathcal{J} \setminus \mathcal{J}^{(1)}(\boldsymbol{n})\colon c^{(2)}_j(\boldsymbol{n}) = c^{(2)}_{\min}(\boldsymbol{n})\}$ and let $\mathcal{R}^{(2)}(\boldsymbol{n}) = \{r \in \mathcal{R} \setminus \mathcal{R}^{(1)}(\boldsymbol{n})\colon A_{jr} = 1 \text{ for some } j \in \mathcal{J}^{(2)}(\boldsymbol{n})\}$. For each $r \in \mathcal{R}^{(2)}(\boldsymbol{n})$ we let

$$b_r(\boldsymbol{n}) = w_r n_r c^{(2)}_{\min}(\boldsymbol{n}). \tag{1.19}$$

If $\mathcal{R}^{(1)}(\boldsymbol{n}) \cup \mathcal{R}^{(2)}(\boldsymbol{n}) = \mathcal{R}$ then we now have evaluated a bandwidth allocation for each call type $r \in \mathcal{R}$. Otherwise, we continue the iterative algorithm in this fashion so as to allocated bandwidth to each call type $r \in \mathcal{R}$. We note that for each $\boldsymbol{n} \in \mathbb{Z}^R_+$ this iterative algorithm terminates in a finite number of steps since the sets $\mathcal{R}^{(1)}(\boldsymbol{n}), \mathcal{R}^{(2)}(\boldsymbol{n}), \ldots$ are non-empty.

Max-min fair control strategies are "fair" in the sense that for any $\boldsymbol{n} \in \mathbb{Z}^R_+$, we allocate the maximum minimum bandwidth to the call types which use resources where there is most competition first. These call types are then removed from the network and we iteratively repeat this process of allocating the maximum minimum bandwidth to call types which use resources where there is most competition.

Mo and Walrand [26] introduced the class of *weighted α-fair-sharing* control strategies or *α-bandwidth allocations*. These control strategies are defined as follows. For $\alpha > 0$, $\alpha \neq 1$, and positive weights w_r, $r \in \mathcal{R}$, a weighted α-fair-sharing control strategy, is given by taking, for each $\boldsymbol{n} \in \mathbb{Z}^R_+$, $\boldsymbol{b}$ to maximise the concave function

$$g^{\alpha}(\boldsymbol{b}) = \sum_{r\in\mathcal{R}} w_r n_r^{\alpha} \frac{(b_r(\boldsymbol{n}))^{1-\alpha}}{1-\alpha} \tag{1.20}$$

subject to the capacity constraints (1.4) and to the restrictions $b_r(\boldsymbol{n}) \geq 0$ (and $b_r(\boldsymbol{n}) = 0$ if $n_r = 0$) for every $r \in \mathcal{R}$. This class of control strategy can be further extended to each of the cases $\alpha = 0, 1, \infty$ by taking the limit of the α-fair-sharing control strategy as α tends to each of these values. For $\alpha = 1$ this is equivalent to maximising the objective function

$$g^{1}(\boldsymbol{b}) = \sum_{r\in\mathcal{R}} n_r w_r \log(b_r(\boldsymbol{n})) \tag{1.21}$$

subject to the capacity constraints (1.4) and to the restrictions $b_r(\boldsymbol{n}) \geq 0$ (and $b_r(\boldsymbol{n}) = 0$ if $n_r = 0$) for every $r \in \mathcal{R}$.

Mo and Walrand [26] have shown that all other known classes of fair-sharing control strategies are special cases of α-fair-sharing control strategies. When $\alpha = 0$, the control strategy is said to be a *maximum throughput* control strategy, when $\alpha = 1$, the control strategy is called a *proportionally fair* control strategy and when $\alpha = 2$, the control strategy is called a *potential delay* control strategy. Mo and Walrand [26] have shown that when $\alpha \to \infty$, the control strategy tends to the (weighted) max-min fair-sharing control strategy as defined and evaluated above.

Two papers by de Veciana *et al* [11] and Bonald and Massouliè [3] have considered a broad class of fair-sharing control strategies and have shown that under such control strategies, provided the various network resources individually have sufficient capacity (i.e. condition (1.7) holds), entrainment cannot arise and the network will remain stable. For proportionally fair and max-min fair-sharing control strategies de Veciana *et al* [11] have shown, using Lyapunov function techniques, that condition (1.7) is sufficient for the stability of these control strategies. Bonald and Massouliè [3] show that this result holds for general weighted α-fair-sharing control strategies by using fluid limits and appealing to a result of Dai [10].

Most recently the paper by Bonald *et al* [4] considered a model with a more general definition of a communication network model. Instead of using the incidence matrix A, they define a *capacity set* $\mathcal{C} \subset \mathbb{R}_+^R$. This set is compact, convex and coordinate convex (where a set $A \subset \mathbb{R}_+^R$ is *coordinate convex* if for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}_+^R$

such that $\boldsymbol{x} \geq \boldsymbol{y}$ component-wise, we have that $\boldsymbol{x} \in A$ implies that $\boldsymbol{y} \in A$). In place of condition (1.4) the paper assumes that $\boldsymbol{b} \in \mathcal{C}$ for all values of $\boldsymbol{n} \in \mathbb{Z}_+^R$ and in place of condition (1.7) the paper assumes that the input parameter is contained in the interior of $\mathcal{C}$. This paper shows that this model includes alternative routing models and can be applied to wireless networks. These are two models of huge practical importance. For this model the authors show that proportionally fair and max-min fair-sharing control strategies are stable. It is also shown that *balanced-fair* control strategies (see Chapter 6 for a definition of the *balance property*) are stable for this model.

One of the advantages of the above results for the stability of α-fair-sharing control strategies is that they hold for all network topologies. They are also completely robust in the sense that they are applicable without knowledge of the input parameter $\boldsymbol{\kappa}$. In applying control strategies to the Internet, engineers may only know the state of the system $\boldsymbol{n} \in \mathbb{Z}_+^R$ at any time and the capacity parameter $\boldsymbol{c}$, hence fair-sharing control strategies are useful in applying the simultaneous processor sharing model. Although some of the control strategies described in this thesis are only applicable with knowledge of the input parameter $\boldsymbol{\kappa}$, it is possible to couple most of these control strategies to ones which are robust in this sense.

The allocation of capacity by α-fair-sharing control strategies is very specific for each α and does not provide any insight into the problem of entrainment except for the fact that it cannot occur here. Solving the optimisation problem (1.20) does not provide insight into why a particular control strategy $\boldsymbol{b}$ may be stable or unstable. This is a limitation of α-fair-sharing control strategies. Indeed in this thesis we shall consider non-α-fair-sharing control strategies which are stable. Further, in this thesis we attempt to consider control strategies which provide an insight into the problem of entrainment.

Finally in this section, we give a brief outline to two recent papers which, for different communication models, apply similar techniques to the ones used in this thesis. MacPhee and Müller [24] consider a multiple queue model. For each job we have Poisson arrivals and there are various routing schemes for allocating job arrivals to a queue. There is a set of *servers* which can be configured in a variety of ways and *Jackson-type* feedback is allowed. The service times are exponentially distributed, and along with the feedback probabilities depend upon the configuration of servers. The finite collection of pairs of *server configurations* and routing schemes are called *management regimes.* The paper is concerned with the stability of a Markov process which records the vector of queue lengths together with a *policy scheme* (this is a recording of the history of management regimes used). Of interest are two types of policy scheme with a reasonable amount of homogeneity. The authors provide some interesting results regarding stability for these policy schemes using similar methods to (for example) Section 3.2, Chapter 4 and Section 6.6 of this thesis. Stolyar [32] considers a queueing model of multiple input flows each with its own queue which is served by a *generalized switch* i.e. the input flows are served by a switch, the states of which are random and follow a finite Markov chain. Each state of the switch chooses a *scheduling strategy.* The service rates between consecutive times (i.e. of time difference 1) depend on the state of the switch, the scheduling strategies and the queue lengths (all at time t). Of interest is a parameter called the *workload* of the system which for any time is the sum of each *weighted* queue length (according to a vector of weights, the components of which comprise of a specific weight for each queue). The author obtains some interesting results — using for example Lyapunov function techniques (see Chapters 3 and 4) — regarding the workload of the system. For example it is shown that under heavy traffic conditions this parameter converges to a *Brownian motion.*

1.4 Outline of the thesis

As stated previously, our aim in this thesis is to identify more general classes of stable control strategies, to provide insight into how the phenomenon of entrainment arises, and to show how control strategies may be modified if necessary so as to avoid it. The outline of the thesis is as follows.

Chapter 2 considers some simple general techniques useful in the analysis of stability, notably the use of coupling arguments and Lyapunov functions. In this chapter we also state and prove some basic results. Chapter 3 examines in detail the case when there are two call types which are constrained by one common resource and the capacity available for each call type is further constrained by a dedicated single unshared resource. Chapter 4 uses a workload-based approach to show that for some quite wide classes of network topologies, any control strategy is stable, provided *only* that for each $r \in \mathcal{R}$ $b_r(\boldsymbol{n})$ is modified so as to be small whenever n_r is small. Chapter 5 introduces new analytical techniques and uses these techniques to show stability for priority-based control strategies for arbitrary dimensions and network topologies. In Chapter 6 we study insensitivity of the stability and the stationary distribution of the network with respect to variation of the call size distribution.

Hansen *et al* [15] covers some of the topics of Chapters 2–5 and in preparation is a paper which will cover some of the topics of Chapter 6 (Hansen *et al* [16]). In particular, the work in Chapter 5 is joint with Jennie Hansen and Stan Zachary and is reported in Hansen *et al* [15]. Also, the work in Sections 6.5 and 6.6 is joint with Stan Zachary and will be reported in Hansen *et al* [16].

Chapter 2

Techniques and preliminary results

In this chapter we introduce techniques to establish sufficient conditions for the stability of a control strategy $\boldsymbol{b}$ for a given network $\mathcal{N}$. We adapt well-known Lyapunov function techniques for discrete time Markov chains and apply them to the simultaneous processor sharing model. Using these techniques, we provide a necessary and sufficient condition for the stability of any Pareto efficient control strategy $\boldsymbol{b}$ in the case when $J = 1$. For an arbitrary network $\mathcal{N}$ we establish a necessary and sufficient condition for the existence of a stable control strategy. Finally, we give some preliminary applications of Lyapunov function techniques to the simultaneous processor sharing model.

2.1 Lyapunov functions

We establish many of the stability results of this thesis via the use of Lyapunov functions. The simplest useful result here is Proposition 2.1.2 below, which gives a sufficient condition for stability and which is just the specialisation of Foster's criterion to the simultaneous processor sharing model. Proposition 2.1.4 is also useful in that it gives sufficient conditions for instability.

For any control strategy $\boldsymbol{b}$ for a network $\mathcal{N}$ and any real function f on $\mathbb{Z}_+^R$ define the real function $D_{\boldsymbol{b}} f$ on $\mathbb{Z}_+^R$ by

$$D_{\boldsymbol{b}} f(\boldsymbol{n}) = \sum_{r \in \mathcal{R}} \left[\nu_r \left(f(\boldsymbol{n} + \boldsymbol{e}_r) - f(\boldsymbol{n}) \right) + \mu_r b_r(\boldsymbol{n}) \left(f(\boldsymbol{n} - \boldsymbol{e}_r) - f(\boldsymbol{n}) \right) \right], \qquad (2.1)$$

where, for each $r \in \mathcal{R}$, $\boldsymbol{e}_r = (e_{rs},\, s \in \mathcal{R})$ denotes the unit vector given by $e_{rr} = 1$ and $e_{rs} = 0$ for $s \neq r$. Since for $r \in \mathcal{R}$ and $\boldsymbol{n} \in \mathbb{Z}_+^R$ such that $n_r = 0$, we have also $b_r(\boldsymbol{n}) = 0$, there is no problem arising in (2.1) from the lack of a formal definition of $f(\boldsymbol{n} - \boldsymbol{e}_r)$ in this case. Further, $D_{\boldsymbol{b}}$ may be thought of as the *generator* of the Markov process $\boldsymbol{n}(\cdot)$ under the control strategy $\boldsymbol{b}$.

Since Proposition 2.1.2 and Proposition 2.1.4 are specialisations of Propositions 5.3 (ii) and 5.4, Chapter 1 of Asmussen [1], which are concerned with discrete-time Markov chains, we use the following uniformisation argument to establish analogous results for continuous-time Markov processes. Suppose that we have a continuous-time irreducible Markov process $(x(t))_{t \geq 0}$ with countable state space E. For any states $i, j \in E$ such that $i \neq j$, let q_{ij} be the transition rate from state i to state j, and for all $i \in E$, let $q_{ii} = -\sum_{j \in E:\, j \neq i} q_{ij}$. Now suppose there exists $M < \infty$ such that

$$-q_{ii} \leq M \text{ for all } i \in E. \qquad (2.2)$$

Let $(X_n)_{n \in \mathbb{N}}$ be the discrete time Markov chain given by transition matrix $P = (p_{ij})_{i,j \in E}$ with transition probabilities given by

$$p_{ij} = \frac{q_{ij}}{M} + \delta_{ij} \text{ for all } i, j \in E \qquad (2.3)$$

where

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases} \qquad (2.4)$$

We note that we can construct the Markov process $(x(t))_{t \geq 0}$ by taking an independent Poisson process with rate M, and making transitions at the Poisson events in accordance with the Markov chain $(X_n)_{n \in \mathbb{N}}$. It follows from this coupling that the Markov process $(x(t))_{t \geq 0}$ is positive recurrent if and only if the Markov chain

$(X_n)_{n\in\mathbb{N}}$ is positive recurrent. Hence the stability of the continuous Markov process $(x(t))_{t\geq 0}$ is equivalent to the stability of the Markov chain $(X_n)_{n\in\mathbb{N}}$. This observation is the key to the proof of Proposition 2.1.2. This result is just the specialisation to the present problem of Proposition 5.3(ii), Chapter 1 of Asmussen [1] for a general Markov chain $(X_n)_{n\in\mathbb{N}}$, which we state for completeness.

Proposition 2.1.1. *(Asmussen [1].) Suppose that a Markov chain $(X_n)_{n\in\mathbb{N}}$ with state space E and transition matrix $P = (p_{ij})_{i,j\in E}$ is irreducible and let E_0 be a finite subset of E. Then $(X_n)_{n\in\mathbb{N}}$ is positive recurrent if for some function $f\colon E \to \mathbb{R}$ and some $\delta > 0$ we have*

$$\inf_{j\in E} f(j) > -\infty, \tag{2.5}$$

$$\sum_{k\in E} p_{jk} f(k) < \infty, \quad \textit{for all } j \in E_0, \quad \textit{and} \tag{2.6}$$

$$\sum_{k\in E} p_{jk} f(k) \leq f(j) - \delta \quad \textit{for all } j \notin E_0. \tag{2.7}$$

Remark We note that similar results to Proposition 2.1.1 appear in Brémaud [9] and Robert [28].

We now apply Proposition 2.1.1 to the simultaneous processor sharing model.

Proposition 2.1.2. *Suppose $\boldsymbol{b}$ is a feasible control strategy for a network $\mathcal{N}$ and suppose that there exists a function $f\colon \mathbb{Z}_+^R \to \mathbb{R}$, a finite set $F \subset \mathbb{Z}_+^R$ and some $\epsilon > 0$ such that*

$$\inf_{\boldsymbol{n}\in\mathbb{Z}_+^R} f(\boldsymbol{n}) > -\infty, \tag{2.8}$$

$$D_{\boldsymbol{b}} f(\boldsymbol{n}) \leq -\epsilon \quad \textit{for all} \quad \boldsymbol{n} \notin F. \tag{2.9}$$

Then $\boldsymbol{b}$ is stable.

Any function f which satisfies conditions (2.8) and (2.9) is often referred to as a *Lyapunov* function while the corresponding finite set F is often called the *refuge*.

Proof of Proposition 2.1.2. Let $\boldsymbol{b}$ be a feasible control strategy for a network $\mathcal{N}$. Let $\boldsymbol{n}(\cdot)$ be the Markov process associated with the given feasible control strategy $\boldsymbol{b}$ for the network $\mathcal{N}$. This Markov process has transition rates for $\boldsymbol{n} \in \mathbb{Z}_+^R$

$$q_{\boldsymbol{n}\boldsymbol{n}'} = \begin{cases} \nu_r & \text{if } \boldsymbol{n}' = \boldsymbol{n} + \boldsymbol{e}_r, \quad r \in \mathcal{R} \\ \mu_r b_r(\boldsymbol{n}) & \text{if } \boldsymbol{n}' = \boldsymbol{n} - \boldsymbol{e}_r, \quad r \in \mathcal{R} \\ -\sum_{r\in\mathcal{R}}(\nu_r + \mu_r b_r(\boldsymbol{n})) & \text{if } \boldsymbol{n}' = \boldsymbol{n} \\ 0 & \text{otherwise.} \end{cases} \tag{2.10}$$

Now, for some suitable $M < \infty$ such that (2.2) holds, let $(\boldsymbol{N}_k)_{k\in\mathbb{N}}$ be the Markov chain associated with the Markov process $\boldsymbol{n}(\cdot)$ constructed using the uniformisation technique described above. In particular $(\boldsymbol{N}_k)_{k\in\mathbb{N}}$ has state space $\mathbb{Z}_+^R$ and transition matrix $P = (p_{\boldsymbol{n}\boldsymbol{n}'})_{\boldsymbol{n},\boldsymbol{n}'\in\mathbb{Z}_+^R}$ with transition probabilities given by

$$p_{\boldsymbol{n}\boldsymbol{n}'} = \begin{cases} \frac{\nu_r}{M} & \text{if } \boldsymbol{n}' = \boldsymbol{n} + \boldsymbol{e}_r, \quad r \in \mathcal{R} \\ \frac{\mu_r b_r(\boldsymbol{n})}{M} & \text{if } \boldsymbol{n}' = \boldsymbol{n} - \boldsymbol{e}_r, \quad r \in \mathcal{R} \\ 1 - \frac{\sum_{r\in\mathcal{R}}(\nu_r + \mu_r b_r(\boldsymbol{n}))}{M} & \text{if } \boldsymbol{n}' = \boldsymbol{n} \\ 0 & \text{otherwise.} \end{cases} \tag{2.11}$$

Then for *any* function $f\colon \mathbb{Z}_+^R \to \mathbb{R}$ and for all $\boldsymbol{n} \in \mathbb{Z}_+^R$

$$D_{\boldsymbol{b}} f(\boldsymbol{n}) = \sum_{r\in\mathcal{R}} \left[\nu_r \left(f(\boldsymbol{n} + \boldsymbol{e}_r) - f(\boldsymbol{n})\right) + \mu_r b_r(\boldsymbol{n}) \left(f(\boldsymbol{n} - \boldsymbol{e}_r) - f(\boldsymbol{n})\right)\right]$$

$$= \sum_{\boldsymbol{n}'\in\mathbb{Z}_+^R} q_{\boldsymbol{n}\boldsymbol{n}'}(f(\boldsymbol{n}') - f(\boldsymbol{n})) \tag{2.12}$$

$$= M \sum_{\boldsymbol{n}'\in\mathbb{Z}_+^R} p_{\boldsymbol{n}\boldsymbol{n}'}(f(\boldsymbol{n}') - f(\boldsymbol{n})) \tag{2.13}$$

where (2.13) follows from (2.3).

Now suppose that for some function $f\colon \mathbb{Z}_+^R \to \mathbb{R}$, a finite set $F \subset \mathbb{Z}_+^R$ and $\epsilon > 0$, inequalities (2.9) and (2.10) hold. For all $\boldsymbol{n} \notin F$, we obtain from (2.9) and (2.13)

$$\sum_{\boldsymbol{n'} \in \mathbb{Z}_+^R} p_{\boldsymbol{n}\boldsymbol{n'}} f(\boldsymbol{n'}) \leq f(\boldsymbol{n}) - \frac{\epsilon}{M}. \tag{2.14}$$

Hence for the Markov chain $(\boldsymbol{N}_k)_{k\in\mathbb{N}}$ condition (2.7) of Proposition 2.1.1 is satisfied with $\delta = \epsilon/M$ and $E_0 = F$.

To show that condition (2.6) of Proposition 2.1.1 is satisfied we note for any $\boldsymbol{n} \in \mathbb{Z}_+^R$ the process $\boldsymbol{n}(\cdot)$ associated with $\boldsymbol{b}$ can only make a finite number of allowable transitions. It follows from (2.11)

$$\sum_{\boldsymbol{n'} \in \mathbb{Z}_+^R} p_{\boldsymbol{n}\boldsymbol{n'}} f(\boldsymbol{n'}) < \infty \quad \text{for all } \boldsymbol{n} \in \mathbb{Z}_+^R. \tag{2.15}$$

Hence condition (2.6) of Proposition 2.1.1 is satisfied. Finally, condition (2.5) of Proposition 2.1.1 is satisfied with $E = \mathbb{Z}_+^R$, and so $(\boldsymbol{N}_k)_{k\in\mathbb{N}}$ is positive recurrent. It follows that $\boldsymbol{n}(\cdot)$ is also positive recurrent and so the feasible control strategy $\boldsymbol{b}$ is stable. □

In light of Proposition 2.1.2 we consider the *workload function* for a network $\mathcal{N}$ as a candidate Lyapunov function. We define the workload function w on $\mathbb{Z}_+^R$ associated with a network $\mathcal{N}$ by

$$w(\boldsymbol{n}) = \sum_{r\in\mathcal{R}} \frac{n_r}{\mu_r} \quad \text{for all} \quad \boldsymbol{n} \in \mathbb{Z}_+^R. \tag{2.16}$$

We note that, for any control strategy $\boldsymbol{b}$ for the network $\mathcal{N}$, we have

$$D_{\boldsymbol{b}} w(\boldsymbol{n}) = \sum_{r\in\mathcal{R}} (\kappa_r - b_r(\boldsymbol{n})) \quad \text{for all } \boldsymbol{n} \in \mathbb{Z}_+^R \tag{2.17}$$

(recall that $b_r(\boldsymbol{n}) = 0$ whenever $n_r = 0$). We also note that for any control strategy $\boldsymbol{b}$ for a network $\mathcal{N}$, w satisfies condition (2.8) of Proposition 2.1.2. In the very simple network of Example 1.1.1 and for any Pareto efficient control strategy $\boldsymbol{b}$, the function $f = w$ satisfies the conditions (2.8) and (2.9) of Proposition 2.1.2, with $\epsilon = c_1 - (\kappa_1 + \kappa_2)$ and $F = \{\boldsymbol{n} \colon n_1 = 0\}$. However w fails to satisfy the hypotheses of Proposition 2.1.2, since in this case F is infinite and indeed we have observed that the Pareto efficient control strategy $\boldsymbol{b}$ considered in Example 1.1.1 is unstable when κ_2 is sufficiently close to c_2. However, as we shall see in Chapter 3, it is sufficient to modify the control strategy $\boldsymbol{b}$ when n_1 is close to 0 in order to achieve stability of the network. This stability may then be proved, using Proposition 2.1.2, by considering a modified version of the workload function w.

The following result complements Proposition 2.1.2. It describes conditions which are useful in determining the instability of a given control strategy $\boldsymbol{b}$ for a network $\mathcal{N}$. This result is just the specialisation to the present problem of Proposition 5.4, Chapter 1 of Asmussen [1] for a general Markov chain, which we state for completeness.

Proposition 2.1.3. *(Asmussen [1].) Suppose that a Markov chain $(X_n)_{n\in\mathbb{N}}$ with state space E is irreducible. Let E_0 be a finite subset of E and let $f\colon E \to \mathbb{R}$ be a function such that*

$$\sum_{k\in E} p_{jk} f(k) \geq f(j) \quad \text{for all} \quad j \notin E_0, \tag{2.18}$$

and that

$$f(i) > f(j) \quad \text{for some} \quad i \notin E_0 \quad \text{and all} \quad j \in E_0. \tag{2.19}$$

Then: (i) if f is bounded above and below the Markov chain $(X_n)_{n\in\mathbb{N}}$ is transient; (ii) if f is bounded below and

$$\sum_{k\in E} p_{jk} |f(k) - f(j)| \leq C \quad \text{for all} \quad j \in E, \tag{2.20}$$

and some $C < \infty$, the Markov chain $(X_n)_{n\in\mathbb{N}}$ is null recurrent or transient.

We apply Proposition 2.1.3 to the simultaneous processor sharing model to obtain the following result.

Proposition 2.1.4. *Let $\boldsymbol{b}$ be a feasible control strategy for a network $\mathcal{N}$ and suppose that there exists some finite set $F \subset \mathbb{Z}_+^R$ and a function $f : \mathbb{Z}_+^R \to \mathbb{R}$ such that*

$$D_{\boldsymbol{b}} f(\boldsymbol{n}) \geq 0 \text{ for all } \boldsymbol{n} \notin F, \tag{2.21}$$

and that

$$f(\boldsymbol{n'}) > f(\boldsymbol{n}) \text{ for some } \boldsymbol{n'} \notin F \text{ and all } \boldsymbol{n} \in F. \tag{2.22}$$

Then $\boldsymbol{b}$ is unstable if either f is bounded above and below or if, f is bounded below and for some $A < \infty$,

$$\sum_{r \in \mathcal{R}} \left(\nu_r \left| f(\boldsymbol{n} + \boldsymbol{e}_r) - f(\boldsymbol{n}) \right| + \mu_r b_r(\boldsymbol{n}) \left| f(\boldsymbol{n} - \boldsymbol{e}_r) - f(\boldsymbol{n}) \right| \right) \leq A, \quad \text{for all} \quad \boldsymbol{n} \in \mathbb{Z}_+^R, \tag{2.23}$$

where $\boldsymbol{e}_r = (e_{rs})_{s \in \mathcal{R}}$ is the R-dimensional unit with $e_{rr} = 1$ and $e_{rs} = 0$ for all $s \neq r$.

Proof. As in the proof of Proposition 2.1.2, let $\boldsymbol{n}(\cdot)$ be the Markov process associated with a given feasible control strategy $\boldsymbol{b}$ for a network $\mathcal{N}$ and with transition rates defined by (2.10). Let $(\boldsymbol{N}_k)_{k \in \mathbb{N}}$ be the Markov chain associated with $\boldsymbol{n}(\cdot)$ and constructed using the uniformisation technique as described above. In particular $(\boldsymbol{N}_k)_{k \in \mathbb{N}}$ has state space $\mathbb{Z}_+^R$ and transition matrix $P = (p_{\boldsymbol{n}\boldsymbol{n'}})_{\boldsymbol{n},\boldsymbol{n'} \in \mathbb{Z}_+^R}$ with transition probabilities $p_{\boldsymbol{n}\boldsymbol{n'}}$ defined by (2.11) (where $M < \infty$ is chosen such that (2.2) holds).

Then, as in the proof of Proposition 2.1.2, for any given function $f\colon \mathbb{Z}_+^R \to \mathcal{R}$ and for all $\boldsymbol{n} \in \mathbb{Z}_+^R$

$$D_{\boldsymbol{b}} f(\boldsymbol{n}) = M \sum_{\boldsymbol{n'} \in \mathbb{Z}_+^R} p_{\boldsymbol{n}\boldsymbol{n'}} (f(\boldsymbol{n'}) - f(\boldsymbol{n})). \tag{2.24}$$

Now suppose for a given finite set $F \subset \mathbb{Z}_+^R$ and function $f\colon \mathbb{Z}_+^R \to \mathbb{R}$ that conditions (2.21) and (2.22) hold. It follows from (2.21) and (2.24) that

$$\sum_{\boldsymbol{n'} \in \mathbb{Z}_+^R} p_{\boldsymbol{n}\boldsymbol{n'}} f(\boldsymbol{n'}) \geq f(\boldsymbol{n}) \text{ for all } \boldsymbol{n} \notin F \tag{2.25}$$

and so for the Markov chain $(\boldsymbol{N}_k)_{k \in \mathbb{N}}$ condition (2.18) of Proposition 2.1.3 is satisfied with $E_0 = F$. Clearly condition (2.19) of Proposition 2.1.3 also holds with $E_0 = F$.

Suppose that $f\colon \mathbb{Z}_+^R \to \mathbb{R}$ is bounded above and below. Then clearly by Proposition 2.1.3 the Markov chain $(\boldsymbol{N}_k)_{k \in \mathbb{N}}$ is not positive recurrent. Therefore $\boldsymbol{n}(\cdot)$ is also not positive recurrent, and the control strategy $\boldsymbol{b}$ is unstable. On the other hand, suppose instead $f\colon \mathbb{Z}_+^R \to \mathbb{R}$ is bounded below and condition (2.23) holds and observe that for all $\boldsymbol{n} \in \mathbb{Z}_+^R$

$$\begin{aligned}
&\sum_{r\in\mathcal{R}} (\nu_r \,|f(\boldsymbol{n}+\boldsymbol{e}_r) - f(\boldsymbol{n})| + \mu_r b_r(\boldsymbol{n})|f(\boldsymbol{n}-\boldsymbol{e}_r) - f(\boldsymbol{n})|) \\
&= \sum_{\boldsymbol{n'}\in\mathbb{Z}_+^R} q_{\boldsymbol{nn'}}|f(\boldsymbol{n'}) - f(\boldsymbol{n})| \\
&= M \sum_{\boldsymbol{n'}\in\mathbb{Z}_+^R} p_{\boldsymbol{nn'}}|f(\boldsymbol{n'}) - f(\boldsymbol{n})|. \qquad (2.26)
\end{aligned}$$

So it follows from conditions (2.23) and (2.26) that condition (2.20) of Proposition 2.1.3 holds for the Markov chain $(\boldsymbol{N}_k)_{k\in\mathbb{N}}$ with f as above, E_0 given by F and $C = A/M$. It follows from Proposition 2.1.3 that the Markov chain $(\boldsymbol{N}_k)_{k\in\mathbb{N}}$ is not positive recurrent. Therefore $\boldsymbol{n}(\cdot)$ is also not positive recurrent, and the control strategy $\boldsymbol{b}$ is unstable. □

2.2 Single resource networks

In this section we consider a network $\mathcal{N}$ with a single resource with capacity $c > 0$. In this case, condition (1.7) becomes

$$\sum_{r\in\mathcal{R}} \kappa_r < c. \qquad (2.27)$$

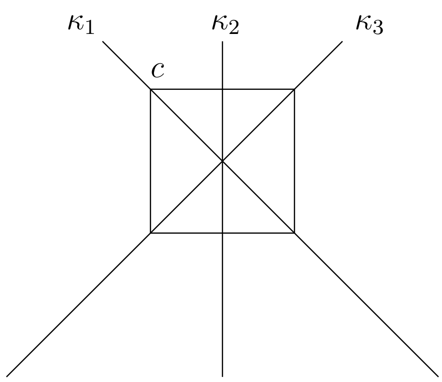

Figure 2.1 Single resource network with $R = 3$.

The following result is generally accepted and we give a short proof.

Theorem 2.2.1. *For any network $\mathcal{N}$ with $J = 1$, condition* (2.27) *is necessary for any feasible control strategy to be stable and sufficient for any Pareto efficient control strategy to be stable.*

Proof. Suppose $\boldsymbol{b}$ is a Pareto efficient control strategy for a network $\mathcal{N}$ with $J = 1$. Suppose (2.27) holds and let $\epsilon = c - \sum_{r\in\mathcal{R}} \kappa_r > 0$. We show that the workload function w defined by (2.16) satisfies the conditions of Proposition 2.1.2 for $\boldsymbol{b}$. We note that, for any $\boldsymbol{n} \neq \boldsymbol{0}$, $\sum_{r\in\mathcal{R}} b_r(\boldsymbol{n}) = c$, by Pareto efficiency. Thus for $\boldsymbol{n} \neq \boldsymbol{0}$,

$$D_{\boldsymbol{b}} w(\boldsymbol{n}) = \sum_{r\in\mathcal{R}} (\kappa_r - b_r(\boldsymbol{n}))$$

$$= \sum_{r\in\mathcal{R}} \kappa_r - c = -\epsilon. \tag{2.28}$$

So condition (2.9) is satisfied (with w as the Lyapunov function, $F = \{\boldsymbol{0}\}$ as the refuge and $\epsilon = c - \sum_{r\in\mathcal{R}} \kappa_r$). It follows from Proposition 2.1.2 that $\boldsymbol{b}$ is stable.

Now, for a network $\mathcal{N}$ with $J = 1$, suppose that (2.27) does not hold and suppose $\boldsymbol{b}$ a feasible control strategy for $\mathcal{N}$ such that $\sum_{r\in\mathcal{R}} b_r(\boldsymbol{n}) \leq c$ for all $\boldsymbol{n} \in \mathbb{Z}_+^R$. We apply Proposition 2.1.4 to show that $\boldsymbol{b}$ is unstable. In this case for all $\boldsymbol{n} \in \mathbb{Z}_+^R$,

$$D_{\boldsymbol{b}} w(\boldsymbol{n}) = \sum_{r\in\mathcal{R}} (\kappa_r - b_r(\boldsymbol{n}))$$

$$\geq \sum_{r\in\mathcal{R}} \kappa_r - c \geq 0. \tag{2.29}$$

It follows by (2.29) that condition (2.21) of Proposition 2.1.4 is satisfied with $f = w$ for all $\boldsymbol{n} \in \mathbb{Z}_+^R$. It is also trivial that w satisfies condition (2.22) of Proposition 2.1.4 since $w(\boldsymbol{n}) > w(\boldsymbol{0})$ for all $\boldsymbol{n} \neq \boldsymbol{0}$. Finally, it is clear that w satisfies condition (2.23) of Proposition 2.1.4, since for all $\boldsymbol{n} \in \mathbb{Z}_+^R$

$$\sum_{r\in\mathcal{R}} (\nu_r |w(\boldsymbol{n} + \boldsymbol{e}_r) - w(\boldsymbol{n})| + \mu_r b_r(\boldsymbol{n}) |w(\boldsymbol{n} - \boldsymbol{e}_r) - w(\boldsymbol{n})|) \leq A \tag{2.30}$$

for some $A < \infty$, where $\boldsymbol{e}_r = (e_{rs})_{s\in\mathcal{R}}$ is the R-dimensional unit with $e_{rr} = 1$ and $e_{rs} = 0$ for all $s \neq r$. Hence it follows from Proposition 2.1.4 (with w as the required function f and $F = \{\boldsymbol{0}\}$) that $\boldsymbol{b}$ is unstable. □

2.3 Existence of stable control strategies

In this section we develop necessary and sufficient conditions for the existence of a stable control strategy $\boldsymbol{b}$ for any network $\mathcal{N}$. For an arbitrary network $\mathcal{N}$ we also establish a sufficient condition for the existence of a stable Pareto efficient control strategy.

To give conditions for the existence of a stable control strategy we first consider a *complete partitioning* control strategy $\boldsymbol{b}$ defined as follows. Given a vector $\hat{\boldsymbol{b}} = (\hat{b}_r,\ r \in \mathcal{R})$ such that

$$\sum_{r\in\mathcal{R}} A_{jr}\hat{b}_r \leq c_j \quad \text{for all} \quad j \in \mathcal{J}, \tag{2.31}$$

we define the corresponding complete partitioning control strategy $\boldsymbol{b}$ by defining, for each $r \in \mathcal{R}$, $b_r(\boldsymbol{n}) = \hat{b}_r$ whenever $n_r > 0$ (and setting $b_r(\boldsymbol{n}) = 0$ otherwise).

Lemma 2.3.1. *Suppose $\mathcal{N}$ is a network such that* (1.7) *holds. Then there exists a stable feasible control strategy $\boldsymbol{b}$ on $\mathcal{N}$.*

Proof. It follows from (1.7) that there exists a vector $\hat{\boldsymbol{b}}$ which satisfies (2.31) and is such that for every $r \in \mathcal{R}$

$$\hat{b}_r > \kappa_r. \tag{2.32}$$

Let $\boldsymbol{b}$ be the corresponding complete partitioning control strategy for $\mathcal{N}$ that is based on $\hat{\boldsymbol{b}}$ and let $\boldsymbol{n}(\cdot) = (n_r(\cdot))_{r\in\mathcal{R}}$ denote the Markov process associated with the control strategy $\boldsymbol{b}$. We note that under this complete partitioning control strategy $\boldsymbol{b}$, the processes $\{n_r(\cdot), r \in \mathcal{R}\}$ are independent, and each process $n_r(\cdot)$ behaves as a (single-server) queue with arrival rate ν_r and departure rate $\mu_r\hat{b}_r$ (when $n_r(t) > 0$). By (2.32) it follows from standard queueing results that each process $n_r(\cdot)$ is positive recurrent (see for example Jones and Smith [17]). Hence the entire process $\boldsymbol{n}(\cdot)$ is positive recurrent and so $\boldsymbol{b}$ is stable. □

Next we establish a necessary condition for the existence of a feasible control strategy which is stable.

Lemma 2.3.2. *Let* $\boldsymbol{b}$ *be feasible control strategy for a network* $\mathcal{N}$*, then condition* (1.7) *is necessary for the stability of* $\boldsymbol{b}$*.*

Proof. Suppose that $\mathcal{N}$ is a network such that condition (1.7) does not hold and suppose $\boldsymbol{b}$ is a feasible control strategy for $\mathcal{N}$. Since (1.7) does not hold, there is some resource $j' \in \mathcal{J}$ such that

$$\sum_{r\in\mathcal{R}} A_{j'r}\kappa_r \geq c_{j'}. \tag{2.33}$$

Consider the *restricted workload function* associated with resource j' and defined by

$$w_{j'}(\boldsymbol{n}) = \sum_{r\in\mathcal{R}} A_{j'r}\frac{n_r}{\mu_r} \quad \text{for all} \quad \boldsymbol{n} \in \mathbb{Z}_+^R. \tag{2.34}$$

Then, by (1.4) and (2.33), for all $\boldsymbol{n} \in \mathbb{Z}_+^R$,

$$\begin{aligned} D_{\boldsymbol{b}}w_{j'}(\boldsymbol{n}) &= \sum_{r\in\mathcal{R}} A_{j'r}(\kappa_r - b_r(\boldsymbol{n})) \\ &\geq \sum_{r\in\mathcal{R}} A_{j'r}\kappa_r - c_{j'} \geq 0. \end{aligned} \tag{2.35}$$

Hence condition (2.21) of Proposition 2.1.4 is satisfied with $f = w_{j'}$ for all $\boldsymbol{n} \in \mathbb{Z}_+^R$.

It is also trivial that condition (2.22) of Proposition 2.1.4 is satisfied since, $w_{j'}(\boldsymbol{n}) > w_{j'}(\boldsymbol{0})$ for all $\boldsymbol{n} \neq \boldsymbol{0}$. Finally, it follows that condition (2.23) of Proposition 2.1.4 is satisfied, since for all $\boldsymbol{n} \in \mathbb{Z}_+^R$

$$\sum_{r\in\mathcal{R}} \left(\nu_r \left|w_{j'}(\boldsymbol{n}+\boldsymbol{e}_r) - w_{j'}(\boldsymbol{n})\right| + \mu_r b_r(\boldsymbol{n}) \left|w_{j'}(\boldsymbol{n}-\boldsymbol{e}_r) - w_{j'}(\boldsymbol{n})\right|\right) \leq A \tag{2.36}$$

for some $A < \infty$, where $\boldsymbol{e}_r = (e_{rs})_{s\in\mathcal{R}}$ is the R-dimensional unit with $e_{rr} = 1$ and $e_{rs} = 0$ for all $s \neq r$. Hence it follows from Proposition 2.1.4 (with $f = w_{j'}$ and $F = \{\boldsymbol{0}\}$) that $\boldsymbol{b}$ is unstable. □

Finally, in this section we show that, in some cases, we can couple two processes together such that the positive recurrence of one of these processes implies the positive recurrence of the other process (see for example Lindvall [22]). This simple coupling result is frequently useful in making elementary comparisons between processes driven by two different control strategies. We give an immediate application of Proposition 2.3.3 in Corollary 2.3.4

Proposition 2.3.3. *Suppose that* $\boldsymbol{b}$, $\boldsymbol{b}'$ *are two feasible control strategies for a network* $\mathcal{N}$ *such that for all* $r \in \mathcal{R}$,

$$b_r(\boldsymbol{n}) \geq b'_r(\boldsymbol{n'}) \quad \text{for all} \quad \boldsymbol{n}, \boldsymbol{n'} \in \mathbb{Z}_+^R \quad \text{such that} \quad n_r = n'_r,\, n_s \leq n'_s \quad \text{for } s \neq r. \tag{2.37}$$

Then $\boldsymbol{b}$ *is stable if* $\boldsymbol{b}'$ *is stable.*

Proof. Consider the two processes $\boldsymbol{n}(\cdot)$ and $\boldsymbol{n}'(\cdot)$ driven by $\boldsymbol{b}$ and $\boldsymbol{b}'$ respectively. It follows from (2.37) that we can construct the two processes $\boldsymbol{n}(\cdot)$ and $\boldsymbol{n}'(\cdot)$ on the same probability space such that whenever $\boldsymbol{n}(\cdot) = \boldsymbol{n}$ and $\boldsymbol{n}'(\cdot) = \boldsymbol{n}'$ where $n_r = n'_r$, $n_s \leq n'_s$ for $s \neq r$, then, if the process $\boldsymbol{n}'(t)$ has a transition to $\boldsymbol{n'} - \boldsymbol{e}_r$, the process $\boldsymbol{n}(t)$ has a transition to $\boldsymbol{n} - \boldsymbol{e}_r$. It follows from this coupling that if $n_r(0) \leq n'_r(0)$ for all $r \in \mathcal{R}$, then

$$n_r(t) \leq n'_r(t) \quad \text{for all} \quad r \in \mathcal{R} \quad \text{and} \quad t \geq 0. \tag{2.38}$$

Hence, if the process $\boldsymbol{n'}(\cdot)$ is positive recurrent, then the process $\boldsymbol{n}(\cdot)$ is also positive recurrent and the result follows. □

Corollary 2.3.4. *Suppose that* $\hat{\boldsymbol{b}}$ *is a vector which satisfies* (2.31) *and such that*

$$\hat{b}_r > \kappa_r \quad \text{for all} \quad r \in \mathcal{R}. \tag{2.39}$$

Suppose $\boldsymbol{b}$ *is a feasible control strategy for a network* $\mathcal{N}$ *such that* $b_r(\boldsymbol{n}) \geq \hat{b}_r$ *whenever* $n_r > 0$. *Then* $\boldsymbol{b}$ *is stable.*

Proof. The result follows from Proposition 2.3.3 by taking $\boldsymbol{b}'$ to be the stable complete partitioning control strategy based on $\hat{\boldsymbol{b}}$. □

We note that since, given the condition (1.7), there exists a vector $\hat{\boldsymbol{b}}$ and a Pareto efficient control strategy $\boldsymbol{b}$ which together satisfy the conditions of Corollary 2.3.4, it follows that the condition (1.7) is also necessary and sufficient for the existence of some stable Pareto efficient control strategy.

2.4 Further applications

In this section we give some further applications of Propositions 2.1.2 and 2.1.4. The following result shows that for some values of the parameters $\boldsymbol{\kappa}$ and $\boldsymbol{c}$, Pareto efficiency is sufficient for stability.

Lemma 2.4.1. *Suppose $\mathcal{N}$ is a network such that for the parameters $\boldsymbol{\kappa}$ and $\mathbf{c}$,*

$$\sum_{r\in\mathcal{R}} \kappa_r < \min_{j\in\mathcal{J}} c_j. \tag{2.40}$$

Then all Pareto efficient control strategies $\boldsymbol{b}$ for $\mathcal{N}$ are stable.

Proof. Suppose $\mathcal{N}$ is a network which satisfies (2.40) and suppose that $\boldsymbol{b}$ is a Pareto efficient control strategy for $\mathcal{N}$. We apply Proposition 2.1.2 to show that, under (2.40), the Pareto efficient control strategy $\boldsymbol{b}$ for the network $\mathcal{N}$ is stable. Let $\epsilon = \min_{j\in\mathcal{J}} c_j - \sum_{r\in\mathcal{R}} \kappa_r > 0$. Suppose $\boldsymbol{n} \neq \mathbf{0}$. Then, since $\boldsymbol{b}$ is Pareto efficient, there exists a saturated $j' \in \mathcal{J}$ such that

$$\sum_{r\in\mathcal{R}} A_{j'r} b_r(\boldsymbol{n}) = c_{j'}. \tag{2.41}$$

Hence, by (2.40) and (2.41), $\sum_{r\in\mathcal{R}} b_r(\boldsymbol{n}) \geq c_{j'} > \sum_{r\in\mathcal{R}} \kappa_r$. Now consider the workload function w defined by (2.16). Then, by (2.17), and (2.40)

$$\begin{aligned} D_{\boldsymbol{b}} w(\boldsymbol{n}) &= \sum_{r\in\mathcal{R}} (\kappa_r - b_r(\boldsymbol{n})) \\ &\leq \sum_{r\in\mathcal{R}} \kappa_r - c_{j'} \leq -\epsilon. \end{aligned} \tag{2.42}$$

Hence (2.9) is satisfied (with w as our required Lyapunov function, $F = \{\mathbf{0}\}$ as our refuge and $\epsilon = \min_{j\in\mathcal{J}} c_j - \sum_r \kappa_r$). It follows from Proposition 2.1.2 that $\boldsymbol{b}$ is stable. □

Now consider again the Example 1.1.1 where $b_1(\boldsymbol{n}) = c_1$ when $n_1 > 0$. Assume that $c_1 = 1$, $c_2 = 1/2$, $\nu_1 = 1/2$ and $\nu_2 < 1/2$. Again, for simplicity of exposition assume that $\mu_1 = \mu_2 = 1$. We note that these values of the parameters $\boldsymbol{\kappa}$ and $\boldsymbol{c}$ satisfy condition (1.10). For any $\lambda > 0$, define the function

$$f_\lambda(\boldsymbol{n}) = n_1 + \lambda n_2, \tag{2.43}$$

and observe $f_\lambda(\mathbf{0}) < f_\lambda(\boldsymbol{n})$, for all $\boldsymbol{n} \neq \mathbf{0}$. Hence (for any $\lambda > 0$) f_λ satisfies condition (2.22) of Proposition 2.1.4. Next, for $\boldsymbol{n} \neq \mathbf{0}$ we have

$$D_{\boldsymbol{b}} f_\lambda(\boldsymbol{n}) = \begin{cases} -\frac{1}{2} + \lambda\nu_2 & \text{if} \quad n_1 > 0, \\ \frac{1}{2} + \lambda(\nu_2 - \frac{1}{2}) & \text{if} \quad n_1 = 0, n_2 > 0 \end{cases} \tag{2.44}$$

and, in particular, $D_{\boldsymbol{b}} f_\lambda(\boldsymbol{n}) \geq 0$ for all $\boldsymbol{n} \neq \mathbf{0}$ if

$$\frac{1}{2\nu_2} \leq \lambda \leq \frac{1}{1 - 2\nu_2}. \tag{2.45}$$

It is easy to check that if $1/4 \leq \nu_2 < 1/2$, we can choose λ to satisfy condition (2.45) and so $D_{\boldsymbol{b}} f_\lambda(\boldsymbol{n}) \geq 0$ for $\boldsymbol{n} \neq \mathbf{0}$. Finally, it follows that condition (2.23) of Proposition 2.1.4 is satisfied here, since for all $\boldsymbol{n} \in \mathbb{Z}_+^2$ and $\lambda > 0$

$$\sum_{r \in \mathcal{R}} \left(\nu_r \left| f_\lambda(\boldsymbol{n} + \boldsymbol{e}_r) - f_\lambda(\boldsymbol{n}) \right| + \mu_r b_r(\boldsymbol{n}) \left| f_\lambda(\boldsymbol{n} - \boldsymbol{e}_r) - f_\lambda(\boldsymbol{n}) \right| \right) \leq A \tag{2.46}$$

for some $A < \infty$, where $\boldsymbol{e}_r = (e_{rs})_{s \in \mathcal{R}}$ is the 2-dimensional unit with $e_{rr} = 1$ and $e_{rs} = 0$ for all $s \neq r$.

Hence in the case $1/4 \leq \nu_2 < 1/2$, we can choose λ such that the conditions of Proposition 2.1.4 are satisfied (with $f = f_\lambda$ and $F = \{\mathbf{0}\}$). It follows that the Pareto efficient control strategy $\boldsymbol{b}$, defined by $b_1(\boldsymbol{n}) = 1$ when $n_1 > 0$, is unstable. (Similarly for $\nu_2 < 1/4$ we can choose λ to show, using Proposition 2.1.2, that $\boldsymbol{b}$ is stable in this case.)

We conclude this chapter by considering the α-fair-sharing control strategies introduced in Chapter 1. For the cases $\alpha = 1$ and $\alpha = \infty$, de Veciana *et al* [11] use Lyapunov function techniques to show that the condition (1.7) is sufficient for the stability of (weighted) α-fair-sharing control strategies. Bonald and Massoulié

[3] show that this result holds for general α by using fluid limits and appealing to a result of Dai [10] for multi-class queueing networks. (In fact it is not certain that Dai's result is directly applicable to the processor sharing network with simultaneous resource requirements.) The following result shows that for all $\alpha > 0$ it is possible to prove that α-fair-sharing control strategies are stable using Lyapunov function techniques.

Proposition 2.4.2. *(Bonald and Massoulié [3].) Suppose that* $\boldsymbol{b}$ *is an* α*-fair-sharing control strategy for a network* $\mathcal{N}$*. Then* $\boldsymbol{b}$ *is stable if and only if condition* (1.7) *is satisfied.*

We note that the function f_α defined by (2.47) below is the same as defined by Bonald and Massoulié [3] in their fluid limit approach and essentially their fluid limit technique is equivalent to the Lyapunov function technique described below. The following proof is an adaptation to Lyapunov functions of their fluid limit technique, following the same general line of argument except with greater attention to the necessary details required.

Proof of Proposition 2.4.2. Suppose $\alpha > 0$, and consider the function $f_\alpha \colon \mathbb{R}_+^R \to \mathbb{R}$ defined by

$$f_\alpha(\boldsymbol{n}) = \sum_{r \in \mathcal{R}} w_r \mu_r^{-1} \kappa_r^{-\alpha} \frac{n_r^{\alpha+1}}{\alpha+1} \quad \text{for} \quad \boldsymbol{n} \in \mathbb{R}_+^R. \tag{2.47}$$

where $\{w_r, r \in \mathcal{R}\}$ are positive weights. It follows that f_α satisfies condition (2.8) of Proposition 2.1.2 for all $\alpha > 0$.

Since f_α is twice differentiable in the interior of $\mathbb{R}_+^R$ it follows from the second (or extended) mean value theorem (see for example Quadling [27]) that, for each $\boldsymbol{n} \in \mathbb{Z}_+^R$ and for each $r \in \mathcal{R}$, there exists a $\theta_r(n_r)$ with $0 < \theta_r(n_r) < 1$ such that

$$f_\alpha(\boldsymbol{n} + \boldsymbol{e}_r) - f_\alpha(\boldsymbol{n}) = \frac{\partial}{\partial n_r} f_\alpha(\boldsymbol{n}) + \frac{1}{2} \frac{\partial^2}{\partial n_r^2} f_\alpha(\boldsymbol{n} + \theta_r(n_r)\boldsymbol{e}_r) \tag{2.48}$$

and further, for all $\boldsymbol{n} \in \mathbb{Z}_+^R$ and for each $r \in \mathcal{R}$ such that $n_r > 0$, there exists a $\hat{\theta}_r(n_r)$ with $0 < \hat{\theta}_r(n_r) < 1$ such that

$$f_\alpha(\boldsymbol{n} - \boldsymbol{e}_r) - f_\alpha(\boldsymbol{n}) = -\frac{\partial}{\partial n_r} f_\alpha(\boldsymbol{n}) - \frac{1}{2} \frac{\partial^2}{\partial n_r^2} f_\alpha(\boldsymbol{n} - \hat{\theta}_r(n_r)\boldsymbol{e}_r). \tag{2.49}$$

where $\boldsymbol{e}_r = (e_{rs})_{s\in\mathcal{R}}$ is the R-dimensional unit vector with $e_{rr} = 1$ and $e_{rs} = 0$ for all $s \neq r$. Hence from (2.1) we obtain, for any control strategy $\boldsymbol{b}$ and for all $\boldsymbol{n} \in \mathbb{Z}_+^R$,

$$D_{\boldsymbol{b}} f_\alpha(\boldsymbol{n}) = df_\alpha(\boldsymbol{n}) + \frac{\alpha}{2}\sum_{r\in\mathcal{R}} h_r(n_r) \tag{2.50}$$

where

$$df_\alpha(\boldsymbol{n}) = \sum_{r\in\mathcal{R}} w_r \kappa_r^{-\alpha} n_r^\alpha (\kappa_r - b_r(\boldsymbol{n})) \tag{2.51}$$

and, for each $r \in \mathcal{R}$,

$$h_r(n_r) = w_r\kappa_r^{-\alpha}(\kappa_r(n_r + \theta_r(n_r))^{\alpha-1} - b_r(\boldsymbol{n})(n_r - \hat{\theta}_r(n_r))^{\alpha-1}). \tag{2.52}$$

(Note for all $r \in \mathcal{R}$ that $b_r(\boldsymbol{n}) = 0$ whenever $n_r = 0$ so there is no problem with the lack of a formal definition of $(n_r - \hat{\theta}_r(n_r))$ in this case.)

Now suppose that $\boldsymbol{b} = \boldsymbol{b}(\boldsymbol{n})$ is a (weighted) α-fair-sharing control strategy for a network $\mathcal{N}$, with $\alpha > 0$ and $\alpha \neq 1$ and positive weights $\{w_r, r \in \mathcal{R}\}$. Fix $\boldsymbol{n} \in \mathbb{Z}_+^R$, then from the definition of α-fair-sharing control strategies $\boldsymbol{b}(\boldsymbol{n})$ is then obtained by maximising the function

$$g_{\boldsymbol{n}}^\alpha(\boldsymbol{x}) = \sum_{r\in\mathcal{R}} w_r n_r^\alpha \frac{x_r^{1-\alpha}}{1-\alpha} \tag{2.53}$$

over $\boldsymbol{x} = (x_r)_{r\in\mathcal{R}}$ subject to

$$\sum_{r\in\mathcal{R}} A_{jr} x_r \leq c_j \quad \text{for all} \quad j \in \mathcal{J}, \tag{2.54}$$

i.e. if $\boldsymbol{x}^*$ solves this optimisation problem, then we set $\boldsymbol{b}(\boldsymbol{n}) = \boldsymbol{x}^*$. Let $\boldsymbol{x} \in \mathbb{R}_+^R$ be such that $x_r = (1+\delta)\kappa_r$ for all $r \in \mathcal{R}$ where $\delta > 0$ is such that (2.54) holds. Then,

$$\sum_{r\in\mathcal{R}} \frac{\partial g(\boldsymbol{x})}{\partial x_r}(x_r - b_r(\boldsymbol{n})) \leq 0. \tag{2.55}$$

We note that in the case when $\alpha = 1$, $\boldsymbol{b}(\boldsymbol{n})$ is obtained by maximising the function

$$g^1_{\boldsymbol{n}}(\boldsymbol{x}) = \sum_{r\in\mathcal{R}} w_r n_r \log x_r \tag{2.56}$$

over $\boldsymbol{x} = (x_r)_{r\in\mathcal{R}}$ subject to (2.54). Then we see that (2.55) continues to hold for $\boldsymbol{x}$ defined as above. It follows by (2.55) for $\alpha > 0$

$$\frac{1}{(1+\delta)^\alpha}\sum_{r\in\mathcal{R}} w_r \kappa_r^{-\alpha} n_r^\alpha((1+\delta)\kappa_r - b_r(\boldsymbol{n})) \leq 0. \tag{2.57}$$

Hence, by (2.51) for all $\boldsymbol{n} \in \mathbb{Z}_+^R$

$$df_\alpha(\boldsymbol{n}) \leq -\delta \sum_{r\in\mathcal{R}} w_r \kappa_r^{-\alpha+1} n_r^\alpha. \tag{2.58}$$

It follows from (2.50), (2.52), (2.58) and since $b_r(\boldsymbol{n})$ is bounded that

$$D_{\boldsymbol{b}} f_\alpha(\boldsymbol{n}) \leq -\delta \sum_{r\in\mathcal{R}} w_r \kappa_r^{-\alpha+1} n_r^\alpha (1 + O((\sum_{r\in\mathcal{R}} n_r)^{-1})). \tag{2.59}$$

Hence, since $\alpha > 0$, we can choose some finite subset F of $\mathbb{Z}_+^R$ and $\epsilon > 0$ such that

$$D_{\boldsymbol{b}} f_\alpha(\boldsymbol{n}) \leq -\epsilon \quad \text{for all } \boldsymbol{n} \notin F. \tag{2.60}$$

Hence f_α is a Lyapunov function as required and therefore the α-fair sharing control strategies defined by (1.20) and (1.21) are stable under condition (1.7). □

Chapter 3

Two-dimensional case

In this chapter we consider stability for a network with two routes. In the most general case, two routes share one resource and, in addition, each route requires a single dedicated resource. The first stability result employs methods from Fayolle *et al* [12]. The main approach in this chapter is to characterise control strategies which Lyapunov functions detect as being stable. Having done this, we investigate the characteristics of a control strategy that contribute to it being stable. Finally, we study some specific examples of control strategies on this network.

3.1 General two-dimensional network

The most general two-dimensional network for the simultaneous processor sharing model, denoted $\mathcal{N}_2$, is described as follows. Let $\mathcal{R} = \{1, 2\}$ denote the routes, let $\mathcal{J} = \{0, 1, 2\}$ denote the set of (three) resources and suppose that the network topology is specified by

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.1}$$

For this network the condition (1.7) becomes

$$\kappa_1 < c_1,\ \kappa_2 < c_2,\ \kappa_1 + \kappa_2 < c_0. \tag{3.2}$$

We also assume without loss of generality, that $c_0 < \infty$ and that

$$c_1 \vee c_2 \leq c_0 \leq c_1 + c_2. \tag{3.3}$$

In particular if $c_1 > c_0$ then we can replace c_1 by $c_1^{'} = c_0$. Similarly if $c_0 > c_1 + c_2$ then we can replace c_0 by $c_0^{'} = c_1 + c_2$. (So we also obtain that $c_1 < \infty, c_2 < \infty$).

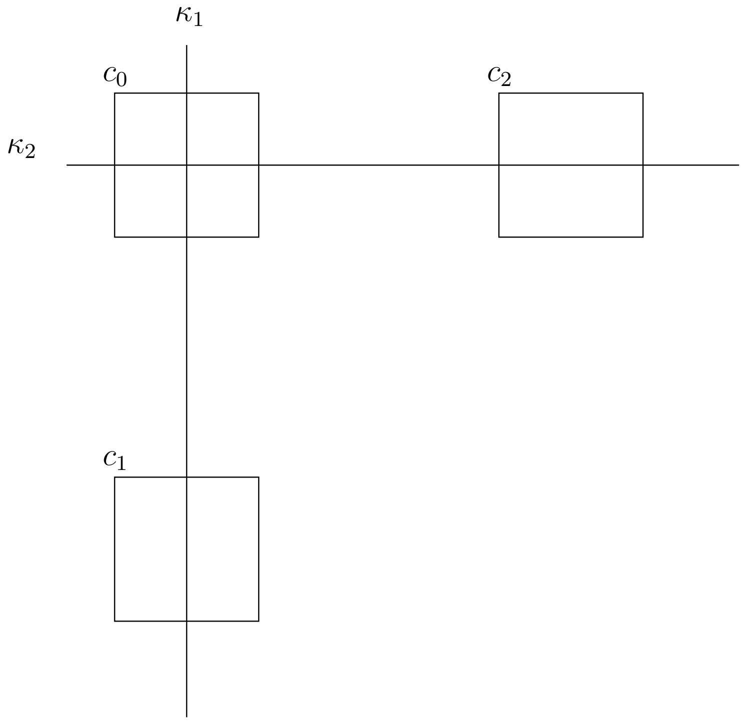

Figure 3.1 The network $\mathcal{N}_2$.

The feasibility condition (1.4) for a control strategy $\boldsymbol{b}$ given by $\boldsymbol{b}(\boldsymbol{n}) = (b_1(\boldsymbol{n}), b_2(\boldsymbol{n}))$ for this network becomes

$$b_1(\boldsymbol{n}) \leq c_1,\ b_2(\boldsymbol{n}) \leq c_2,\ b_1(\boldsymbol{n}) + b_2(\boldsymbol{n}) \leq c_0. \tag{3.4}$$

We note that any two-dimensional network for the simultaneous processor sharing model is a special case of the general network $\mathcal{N}_2$. For example, a network with network topology given by

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \tag{3.5}$$

is equivalent to the general network $\mathcal{N}_2$ with network topology (3.1) and $c_1 = c_0$, since, when $c_1 = c_0$, resource 1 does not put any additional constraint on the bandwidth allocated to calls of type 1.

Finally, we note that any Pareto efficient control strategy $\boldsymbol{b}$ for $\mathcal{N}_2$ satisfies

$$b_1(\boldsymbol{n}) + b_2(\boldsymbol{n}) = c_0. \tag{3.6}$$

when $n_1 \wedge n_2 > 0$ and

$$b_r(\boldsymbol{n}) = c_r, \tag{3.7}$$

whenever $n_r > 0$, $n_{r'} = 0$ for $r = 1, 2$ and $r' \neq r$.

3.2 Maximal spatial homogeneity

The first type of control strategies we consider for the network $\mathcal{N}_2$ are called *maximal spatial homogeneous* control strategies. In this case we have $\boldsymbol{b}(\boldsymbol{n}) = (b_1(\boldsymbol{n}), b_2(\boldsymbol{n}))$ defined as follows,

$$b_r(\boldsymbol{n}) = \begin{cases} b_r & \text{if} \quad n_r > 0 \quad \text{and} \quad n_{r'} > 0, \\ c_r & \text{if} \quad n_r > 0 \quad \text{and} \quad n_{r'} = 0, \\ 0 & \text{otherwise,} \end{cases} \tag{3.8}$$

for $r = 1, 2$ and $r' \neq r$ and where b_1 and b_2 are fixed positive constants. From (1.4) we require

$$b_1 \leq c_1,\, b_2 \leq c_2 \quad \text{and} \quad b_1 + b_2 \leq c_0. \tag{3.9}$$

If the control strategy $\boldsymbol{b}$ is Pareto efficient then we additionally require

$$b_1 + b_2 = c_0. \tag{3.10}$$

In order to state Theorem 3.2.1, we also define the constants

$$b_1^{\max} = \begin{cases} \frac{\kappa_1(c_0-c_2)}{\kappa_1+\kappa_2-c_2} & \text{if} \quad c_2 < \kappa_1 + \kappa_2, \\ \infty & \text{if} \quad c_2 \geq \kappa_1 + \kappa_2, \end{cases} \tag{3.11}$$

and,

$$b_2^{\max} = \begin{cases} \frac{\kappa_2(c_0-c_1)}{\kappa_1+\kappa_2-c_1} & \text{if} \quad c_1 < \kappa_1 + \kappa_2, \\ \infty & \text{if} \quad c_1 \geq \kappa_1 + \kappa_2, \end{cases} \tag{3.12}$$

and we note that, by (3.2), $b_r^{\max} > 0$ for $r = 1, 2$. The following result characterises values of b_1 and b_2 for which the corresponding maximal spatial homogeneous control strategy $\boldsymbol{b}$ is stable.

Theorem 3.2.1. *Suppose that $\boldsymbol{b}$ is a Pareto efficient maximal spatial homogeneous control strategy for $\mathcal{N}_2$. Then $\boldsymbol{b}$ is stable if and only if*

$$b_r < b_r^{\max} \tag{3.13}$$

for $r = 1, 2$.

Proof. Suppose $b_1, b_2 > 0$ are constants such that inequalities (3.9) and equality (3.10) are satisfied. Let $\boldsymbol{b}$ be a maximal spatial homogeneous control strategy for $\mathcal{N}_2$ corresponding to the constants b_1 and b_2 and let $\boldsymbol{n}(\cdot)$ be the process driven by $\boldsymbol{b}$. Let $\alpha_r = \nu_r - b_r\mu_r$ for $r = 1, 2$. Define also the constant

$$\beta_1 = \begin{cases} \alpha_1 & \text{if} \quad \alpha_2 \geq 0, \\ \nu_1 - c_1\mu_1\pi^{'}(0) - b_1\mu_1(1 - \pi^{'}(0)) & \text{if} \quad \alpha_2 < 0, \end{cases} \tag{3.14}$$

where $\boldsymbol{\pi}^{'} = (\pi^{'}(n))_{n\in\mathbb{Z}_+}$ is the stationary distribution of the Markov process on $\mathbb{Z}_+$ with transitions given by

$$n \to \begin{cases} n+1 & \text{at rate} \quad \nu_2, \\ n-1 & \text{at rate} \quad \mu_2 b_2 \mathbf{I}_{\{n\in\mathbb{Z}_+ :\, n\neq 0\}}. \end{cases} \tag{3.15}$$

where $\mathbf{I}_A$ denotes the the indicator function of a set A. From standard calculations for birth-death processes we obtain $\pi^{'}(0) = 1 - \kappa_2/b_2$ (see for example, Kelly [18]). Similarly define,

$$\beta_2 = \begin{cases} \alpha_2 & \text{if} \quad \alpha_1 \geq 0, \\ \nu_2 - c_2\mu_2\pi^{''}(0) - b_2\mu_2(1-\pi^{''}(0)) & \text{if} \quad \alpha_1 < 0. \end{cases} \tag{3.16}$$

where $\boldsymbol{\pi}^{''} = (\pi^{''}(n))_{n\in\mathbb{Z}_+}$ is the stationary distribution of the Markov process on $\mathbb{Z}_+$ given by

$$n \to \begin{cases} n+1 & \text{at rate} \quad \nu_1, \\ n-1 & \text{at rate} \quad \mu_1 b_1 \mathbf{I}_{\{n\in\mathbb{Z}_+:\, n\neq 0\}}. \end{cases} \tag{3.17}$$

Again, standard calculations for birth-death processes yield $\pi^{''}(0) = 1 - \kappa_1/b_1$.

The constant β_1 has an informal interpretation as the "average drift" rate of $n_1(\cdot)$ when $n_1(\cdot)$ is large and $n_2(\cdot)$ is in equilibrium. Similarly β_2 has an informal interpretation as the "average drift" rate of $n_2(\cdot)$ when $n_2(\cdot)$ is large and $n_1(\cdot)$ is in equilibrium. (See Fayolle *et al* [12].) A necessary and sufficient condition for stability can be given in terms of β_1 and β_2: the process $\boldsymbol{n}(\cdot)$ is positive recurrent if and only if $\beta_1 \vee \beta_2 < 0$. (This result follows most simply from Zachary [35], except for the critical case $\beta_1 \vee \beta_2 = 0$, for which see Fayolle *et al* [12].) So it is enough to show that

$$\beta_1 \vee \beta_2 < 0 \quad \text{if and only if} \quad b_r < b_r^{\max},\ r = 1, 2. \tag{3.18}$$

To establish (3.18) we make some simple observations:

(a) It follows from (3.2), that for $r = 1, 2$, $\kappa_r < b_r^{\max}$.

(b) If $b_2 \leq \kappa_2$, then $\alpha_2 \geq 0$ and $\beta_1 = \alpha_1 = \mu_1(\kappa_1 - b_1) < 0$, by (3.10).

(c) On the other hand if $b_2 > \kappa_2$, then $\alpha_2 < 0$ and

$$\begin{aligned} \beta_1 &= \nu_1 - c_1\mu_1\pi_0^{'} - b_1\mu_1(1-\pi_0^{'}) \\ &= \mu_1\left[\kappa_1 - c_1\left(1 - \frac{\kappa_2}{b_2}\right) - b_1\left(\frac{\kappa_2}{b_2}\right)\right] < 0 \end{aligned}$$

if and only if $b_2 < b_2^{\max}$.

It follows from the above observations that $\beta_1 < 0$ if and only if

$$b_2 < b_2^{\max}. \tag{3.19}$$

Similarly $\beta_2 < 0$ if and only if

$$b_1 < b_1^{\max}. \tag{3.20}$$

It follows from (3.19) and (3.20) that a Pareto efficient maximal spatial homogeneous control strategy $\boldsymbol{b}$ is stable if and only if, $b_r < b_r^{\max}$ for $r = 1, 2$. □

Remark We note the techniques that are used in the proof of Theorem 3.2.1 are only applicable for two-dimensional models. In higher dimensions these methods break down and newer methods are needed (see Chapter 5).

Using Theorem 3.2.1 we show that there exists at least one stable maximal spatial homogeneous control strategy for any values of the parameters $\boldsymbol{\kappa} = (\kappa_1, \kappa_2)$ and $\boldsymbol{c} = (c_0, c_1, c_2)$ which satisfy (3.2). This follows from the following claim: for any Pareto efficient maximal spatial homogeneous control strategy $\boldsymbol{b}$ and for $r = 1, 2$

$$b_r > \kappa_r \quad \text{implies that} \quad b_{r'} < b_{r'}^{\max} \quad r' \neq r. \tag{3.21}$$

We note from (3.2) and (3.9) that there exists some $\boldsymbol{b}$ with $b_r > \kappa_r$, $r = 1, 2$. So, in particular, it follows from (3.21) that there exists at least one stable maximal spatial homogeneous control strategy.

We show (3.21) holds for $r = 1$. Suppose first that $c_1 \geq \kappa_1 + \kappa_2$, then $b_2^{\max} = \infty$ and so for any Pareto efficient maximal spatial homogeneous control strategy $\boldsymbol{b}$, $b_2 < b_2^{\max}$. Now suppose instead that $c_1 < \kappa_1 + \kappa_2$. In this case, we have that $b_2^{\max} = (\kappa_2(c_0 - c_1))/(\kappa_1 + \kappa_2 - c_1)$, and we obtain from (3.2)

$$0 < \frac{(c_0 - (\kappa_1 + \kappa_2))(c_1 - \kappa_1)}{\kappa_1 + \kappa_2 - c_1} \leq \frac{c_0 c_1 - \kappa_1 c_0 - \kappa_2 c_1}{\kappa_1 + \kappa_2 - c_1} \tag{3.22}$$

Also since $\kappa_1 > 0$, we have

$$\kappa_1 + \frac{c_0 c_1 - \kappa_1 c_0 - \kappa_2 c_1}{\kappa_1 + \kappa_2 - c_1} > 0. \tag{3.23}$$

Now subtract $b_2^{\max}$ from both sides of this equation to obtain

$$\kappa_1 - c_0 > -b_2^{\max}. \tag{3.24}$$

Hence if $b_1 > \kappa_1$, then by (3.9) $b_2 < c_0 - \kappa_1 < b_2^{\max}$ as required. Similarly (3.21) holds for $r = 2$.

3.3 Stability via Lyapunov function techniques

Now we consider control strategies on $\mathcal{N}_2$ which are not maximally spatially homogeneous. The main tool of this section is to characterise stability for $\mathcal{N}_2$ by constructing suitable Lyapunov functions and showing that the conditions of Proposition 2.1.2 are satisfied. We use Lyapunov function techniques to show that under condition (3.2), any Pareto efficient control strategy is stable provided only that it is suitably modified for values of $\boldsymbol{n}$ close to the boundary of $\mathbb{Z}_+^2$. We begin by considering the *workload function* for $\mathcal{N}_2$ given by

$$w(\boldsymbol{n}) = \frac{n_1}{\mu_1} + \frac{n_2}{\mu_2} \quad \text{for} \quad \boldsymbol{n} \in \mathbb{Z}_+^2. \tag{3.25}$$

Note that for any control strategy $\boldsymbol{b}$,

$$D_{\boldsymbol{b}} w(\boldsymbol{n}) = \kappa_1 + \kappa_2 - (b_1(\boldsymbol{n}) + b_2(\boldsymbol{n})) \text{ for all } \boldsymbol{n} \in \mathbb{Z}_+^2. \tag{3.26}$$

Hence it is easy to check that, for *any* Pareto efficient control strategy $\boldsymbol{b}$, there exists $\epsilon > 0$ such that

$$D_{\boldsymbol{b}} w(\boldsymbol{n}) < -\epsilon \qquad \boldsymbol{n} \in \mathbb{Z}_+^2 \setminus B \tag{3.27}$$

where $B = \{\boldsymbol{n} \colon n_1 \wedge n_2 = 0\}$ is the (infinite) boundary set of $\mathbb{Z}_+^2$. Hence w is “almost” a Lyapunov function for any Pareto efficient control strategy $\boldsymbol{b}$ for $\mathcal{N}_2$. The difficulty is that w may fail to be a Lyapunov function because $\boldsymbol{b}$ is not stable or, in the case where $\boldsymbol{b}$ is stable, w simply fails to satisfy the conditions of Proposition 2.1.2. For example when $\boldsymbol{b}$ is any maximally spatial homogeneous control strategy w fails to satisfy the conditions of Proposition 2.1.2 when $\kappa_1 + \kappa_2 \geq c_1 \wedge c_2$ (this is the case even if $\boldsymbol{b}$ satisfies the conditions of Theorem 3.2.1 and is therefore stable). In the case when the control strategy $\boldsymbol{b}$ is unstable, we can consider how to modify $\boldsymbol{b}$ so as to obtain stability, whereas in the case when the control strategy $\boldsymbol{b}$ is stable, one can consider how to modify w so as to use Proposition 2.1.2 to prove stability. We shall see in Theorem 3.3.1 that it can be necessary to modify both $\boldsymbol{b}$ and w in order to be able to prove the stability of $\boldsymbol{b}$ via Proposition 2.1.2.

In order to prove Theorem 3.3.1 below, we now introduce the family of *modified workload functions* indexed by $a \in \mathbb{Z}_+$ and defined by

$$w^a(\boldsymbol{n}) = \frac{g^a(n_1)}{\mu_1} + \frac{g^a(n_2)}{\mu_2} \quad \text{for} \quad \boldsymbol{n} \in \mathbb{Z}_+^2, \tag{3.28}$$

where for $n \in \mathbb{Z}_+$,

$$g^a(n) = \begin{cases} n & \text{if } n \geq a, \\ \frac{a}{2}(1 + \frac{n^2}{a^2}) & \text{if} \quad n < a. \end{cases} \tag{3.29}$$

It follows from (3.28) and (3.29) that the modified workload function agrees with the workload function except when the process $\boldsymbol{n}(\cdot)$ is close to the boundary B of the state space $\mathbb{Z}_+^2$. We note that for $r = 1, 2$, the contribution of n_r in w tends to zero as $n_r \to 0$ and the contribution of n_r in w^a tends to $a/2$ as $n_r \to 0$.

In particular, when the process $\boldsymbol{n}(\cdot)$ is close to the subset $\{\boldsymbol{n} \in \mathbb{Z}_+^2 \colon n_2 = 0\}$ of $\mathbb{Z}_+^2$ the modified workload function w^a slowly diminishes the contribution of n_2. We show that for a wide class of Pareto efficient control strategies on the network $\mathcal{N}_2$, we can choose an $a > 0$ such that w^a is a Lyapunov function. These control strategies are characterised by the following result (which is similar in spirit to Theorem 3.3.1 (a)(i) of Fayolle *et al* [12]).

Theorem 3.3.1. *Given $\delta > 0$, there exists a constant $a_\delta \geq 0$ such that if $\boldsymbol{b}$ is a Pareto efficient control strategy for $\mathcal{N}_2$ which, for some $K < \infty$, satisfies*

$$\kappa_2 - b_2(\boldsymbol{n}) \leq -\delta \quad \textit{for} \quad \boldsymbol{n} \in \mathbb{Z}_+^2 \quad \textit{such that} \quad n_1 < a_\delta, n_2 \geq K \tag{3.30}$$

$$\kappa_1 - b_1(\boldsymbol{n}) \leq -\delta \quad \textit{for} \quad \boldsymbol{n} \in \mathbb{Z}_+^2 \quad \textit{such that} \quad n_2 < a_\delta, n_1 \geq K, \tag{3.31}$$

then, under condition (3.2), $\boldsymbol{b}$ *is stable.*

Proof. Given $\delta > 0$, fix $a_\delta > 0$ such that

$$a_\delta \geq \frac{1}{\delta}(\kappa_1 + \kappa_2 + c_0) \tag{3.32}$$

Assume without loss of generality that

$$\delta < \min(c_0 - (\kappa_1 + \kappa_2), c_1 - \kappa_1, c_2 - \kappa_2) \tag{3.33}$$

(otherwise we can just choose a smaller $\hat{\delta} > 0$ such that $\hat{\delta} < \min(c_0 - (\kappa_1 + \kappa_2), c_1 - \kappa_1, c_2 - \kappa_2)$ and set $a_\delta = a_{\hat{\delta}}$).

Now suppose $\boldsymbol{b}$ is a Pareto efficient control strategy for $\mathcal{N}_2$ such that conditions (3.30) and (3.31) are satisfied for the given δ and a_δ as chosen above. Note we can assume without loss of generality that in (3.30) and (3.31) we have $K \geq a_\delta$ To show that $\boldsymbol{b}$ is stable we apply Proposition 2.1.2. As a Lyapunov function for $\boldsymbol{b}$ we use the modified workload function w^{a_δ} on $\mathbb{Z}_+^2$ defined by (3.28) and (3.29). Then for all $\boldsymbol{n} \in \mathbb{Z}_+^2$

$$D_{\boldsymbol{b}} w^{a_\delta}(\boldsymbol{n}) = dw^{a_\delta}(\boldsymbol{n}) + \frac{1}{2a_\delta}[h_1^{a_\delta}(n_1) + h_2^{a_\delta}(n_2)], \tag{3.34}$$

where

$$dw^{a_\delta}(\boldsymbol{n}) = \min\left(\frac{n_1}{a_\delta}, 1\right)[\kappa_1 - b_1(\boldsymbol{n})] + \min\left(\frac{n_2}{a_\delta}, 1\right)[\kappa_2 - b_2(\boldsymbol{n})] \tag{3.35}$$

and

$$h_r^{a_\delta}(n_r) = \begin{cases} \kappa_r + b_r(\boldsymbol{n}) & \text{if } 0 \leq n_r < a_\delta, \\ b_r(\boldsymbol{n}) & \text{if } n_r = a_\delta, \\ 0 & \text{if } n_r > a_\delta, \end{cases} \tag{3.36}$$

for $r = 1, 2$. It is clear from the definition of a_δ that

$$\frac{1}{2a_\delta}[h_1^{a_\delta}(n_1) + h_2^{a_\delta}(n_2)] \leq \frac{\delta}{2} \quad \text{for all} \quad \boldsymbol{n} \in \mathbb{Z}_+^2. \tag{3.37}$$

Now suppose $\boldsymbol{n} \in \mathbb{Z}_+^2$ is such that $n_1 \geq K \geq a_\delta$. For any $\boldsymbol{n} \in \mathbb{Z}_+^2$ such that $n_1 \geq K$ it follows from (3.35) that

$$dw^{a_\delta}(\boldsymbol{n}) = \lambda^{a_\delta}(\boldsymbol{n})[\kappa_1 + \kappa_2 - (b_1(\boldsymbol{n}) + b_2(\boldsymbol{n}))] + (1 - \lambda^{a_\delta}(\boldsymbol{n}))[\kappa_1 - b_1(\boldsymbol{n})] \quad (3.38)$$

where

$$\lambda^{a_\delta}(\boldsymbol{n}) = \begin{cases} 1 & \text{if } n_2 \geq a_\delta \\ \frac{n_2}{a_\delta} & \text{if } n_2 < a_\delta \end{cases} \quad (3.39)$$

It then follows from (3.31), (3.38) and (3.39) that

$$dw^{a_\delta}(\boldsymbol{n}) \leq -\delta. \quad (3.40)$$

Likewise, when $\boldsymbol{n} \in \mathbb{Z}_+^2$ is such that $n_2 \geq K \geq a_\delta$ (3.40) continues to hold. So we obtain (3.40) for all $\boldsymbol{n} \notin F_K = \{\boldsymbol{n} \in \mathbb{Z}_+^2 : n_1 \vee n_2 < K\}$.

It now follows from (3.37) and (3.40) that

$$D_{\boldsymbol{b}} w^{a_\delta}(\boldsymbol{n}) < -\frac{\delta}{2} \quad \text{for all } \boldsymbol{n} \notin F_K. \quad (3.41)$$

Hence the conditions of Proposition 2.1.2 are satisfied (with w^{a_δ} as our required Lyapunov function, $\epsilon = \delta/2$ and F_K as our refuge) and it follows that $\boldsymbol{b}$ is stable.

□

Corollary 3.3.2. *Suppose $\boldsymbol{b}$ is a Pareto efficient control strategy for $\mathcal{N}_2$ such that*

$$\lim_{n_1 \to \infty} b_1(n_1, n_2) = c_1 \quad \textit{where} \quad n_2 \in \mathbb{Z}_+ \quad \textit{is fixed} \quad (3.42)$$

and

$$\lim_{n_2 \to \infty} b_2(n_1, n_2) = c_2 \quad \textit{where} \quad n_1 \in \mathbb{Z}_+ \quad \textit{is fixed.} \quad (3.43)$$

Then $\boldsymbol{b}$ is stable.

Proof. Let $\boldsymbol{b}$ be a Pareto efficient control strategy satisfying the conditions (3.42) and (3.43). Let $\delta > 0$ satisfy (3.33) and fix a_δ so that it satisfies (3.32).

It follows from (3.42) that for any $n_2 \in \mathbb{Z}_+$, there exists a $K(n_2) < \infty$ such that if $n_1 > K(n_2)$ then

$$b_1(\boldsymbol{n}) \geq c_1 - \delta. \tag{3.44}$$

Now let

$$K(a_\delta) := \max\{K(n_2)\colon 0 \leq n_2 < a_\delta\}. \tag{3.45}$$

Since a_δ is chosen to be finite it follows that $K(a_\delta) < \infty$. Then it follows from (3.2) and (3.44) that for any $\boldsymbol{n} \in \mathbb{Z}_+^2$ such that $n_1 \geq K(a_\delta)$ and $n_2 < a_\delta$

$$\kappa_1 - b_1(\boldsymbol{n}) \leq -\delta. \tag{3.46}$$

Similarly, it follows from (3.43) that for any $n_1 \in \mathbb{Z}_+$, there exists $K'(n_1) < \infty$ such that if $n_2 > K'(n_1)$ then

$$b_2(\boldsymbol{n}) \geq c_2 - \delta. \tag{3.47}$$

Now let

$$K'(a_\delta) := \max\{K'(n_1)\colon 0 \leq n_1 < a_\delta\}. \tag{3.48}$$

Since a_δ is chosen to be finite it follows that $K'(a_\delta) < \infty$. Then it follows from (3.2) and (3.47) that for any $\boldsymbol{n} \in \mathbb{Z}_+^2$ such that $n_2 \geq K'(a_\delta)$ and $n_1 < a_\delta$

$$\kappa_2 - b_2(\boldsymbol{n}) \leq -\delta \tag{3.49}$$

Finally let $K = K(a_\delta) \vee K'(a_\delta)$. Then it follows from (3.46) and (3.49) that the conditions of Theorem 3.3.1 are satisfied and hence $\boldsymbol{b}$ is stable. □

3.4 Partial spatial homogeneity

We say that a control strategy $\boldsymbol{b}$ given by $\boldsymbol{b}(\boldsymbol{n}) = (b_1(\boldsymbol{n}), b_2(\boldsymbol{n}))$ is a *partial spatial homogeneous* control strategy for the network $\mathcal{N}_2$ if, for some $a \in \mathbb{Z}_+$,

$$b_r(\boldsymbol{n}) = b_r(n_1 \wedge a, n_2 \wedge a) \quad \text{for} \quad r = 1, 2 \quad \text{and all} \quad \boldsymbol{n} \in \mathbb{Z}_+^2. \tag{3.50}$$

It should be noted that the maximal spatial homogeneous control strategies of Section 3.2 are a special case of these control strategies with $a = 1$. We consider a collection of partial spatial homogeneous control strategies $\boldsymbol{b}^a$ given by $\boldsymbol{b}^a(\boldsymbol{n}) = (b_1^a(\boldsymbol{n}), b_2^a(\boldsymbol{n}))$, indexed by $a \in \mathbb{Z}_+$ such that

$$b_1^a(a,a) + b_2^a(a,a) = c_0, \tag{3.51}$$

$$b_1^a(a,n_2) = c_1, \quad b_2^a(a,n_2) = (c_0 - c_1)\mathbf{I}_{\{\boldsymbol{n}\in\mathbb{Z}_+^2 \,:\, n_2\neq 0\}} \quad \text{for all} \quad n_2 < a, \tag{3.52}$$

$$b_1^a(n_1,a) = (c_0 - c_2)\mathbf{I}_{\{\boldsymbol{n}\in\mathbb{Z}_+^2 \,:\, n_1\neq 0\}}, \quad b_2^a(n_1,a) = c_2 \quad \text{for all} \quad n_1 < a, \tag{3.53}$$

$$b_1^a(\boldsymbol{n}) + b_2^a(\boldsymbol{n}) = c_0 \quad \text{for all} \quad \boldsymbol{n} \in \mathbb{Z}_+^2 \quad \text{such that} \quad n_1, n_2 < a, \tag{3.54}$$

(where $\mathbf{I}_A$ is the indicator function of a set A).

We show that for all a large enough , the corresponding control strategy $\boldsymbol{b}^a$ that satisfies (3.51) – (3.54) is stable.

Proposition 3.4.1. *Suppose* $\{\boldsymbol{b}^a : a \in \mathbb{Z}_+\}$ *is a collection of partial spatial homogeneous control strategies for* $\mathcal{N}_2$ *which satisfy conditions* (3.51) – (3.54)*. Then there exists* $\hat{a} \in \mathbb{Z}_+$ *such that for all* $a \geq \hat{a}$ *the corresponding control strategy* $\boldsymbol{b}^a$ *is stable.*

Proof. Given $\delta > 0$, set $\hat{a} = a_\delta$ to satisfy (3.32). Assume without loss of generality that δ satisfies (3.33). Now suppose that $\boldsymbol{b}^{\hat{a}}$ is a feasible Pareto efficient control strategy that satisfies (3.51) – (3.54) for $\hat{a}$. To show that $\boldsymbol{b}^{\hat{a}}$ is stable we apply Theorem 3.3.1. It follows from (3.53) that

$$\kappa_2 - b_2^{\hat{a}}(\boldsymbol{n}) = \kappa_2 - c_2 \leq -\delta \quad \text{for} \quad \boldsymbol{n} \in \mathbb{Z}_+^2 \quad \text{such that} \quad n_1 < \hat{a}, n_2 \geq \hat{a} \tag{3.55}$$

and similarly it follows from (3.52) that

$$\kappa_1 - b_1^{\hat{a}}(\boldsymbol{n}) = \kappa_1 - c_1 \leq -\delta \quad \text{for} \quad \boldsymbol{n} \in \mathbb{Z}_+^2 \quad \text{such that} \quad n_2 < \hat{a}, n_1 \geq \hat{a}. \tag{3.56}$$

Similarly, (3.55) and (3.56) both hold for all $a \in \mathbb{Z}_+$ such that $a \geq \hat{a}$.

Hence (3.30) and (3.31) are satisfied (with $K = a$) and so it follows from Theorem 3.3.1 that $\boldsymbol{b}^a$ is stable for all $a \geq \hat{a}$. □

3.5 Fair-sharing control strategies for $\mathcal{N}_2$

Finally in this chapter, we explicitly solve the optimisation problems that define α-fair-sharing control strategies for the network $\mathcal{N}_2$, in the cases $\alpha \neq 1$ and $\alpha = 1$, by using standard techniques (see for example, Whittle [33]). We show that α-fair-sharing control strategies for $\mathcal{N}_2$ satisfy the conditions of Corollary 3.3.2. This gives an alternative proof of the stability of α-fair-sharing control strategies for $\mathcal{N}_2$.

Proposition 3.5.1. *All α-fair-sharing control strategies for $\mathcal{N}_2$ satisfy the conditions of Corollary 3.3.2.*

Proof. First we solve the optimisation problems that define α-fair-sharing control strategies for $\mathcal{N}_2$. Suppose $\alpha \neq 1$. We can write the optimisation problem for α-fair-sharing control strategies on $\mathcal{N}_2$ as follows. For $\alpha > 0$ and positive weights w_1 and w_2, a control strategy $\boldsymbol{b}$ (given by $\boldsymbol{b}(\boldsymbol{n}) = (b_1(\boldsymbol{n}), b_2(\boldsymbol{n}))$) for $\mathcal{N}_2$ is α-fair-sharing if for each $\boldsymbol{n} \in \mathbb{Z}_+^2$, $\boldsymbol{b}(\boldsymbol{n})$ maximises the concave function

$$g_{\boldsymbol{n}}^{\alpha}(\boldsymbol{x}) = w_1 n_1^{\alpha} \frac{x_1^{1-\alpha}}{1-\alpha} + w_2 n_2^{\alpha} \frac{x_2^{1-\alpha}}{1-\alpha} \quad \text{for} \quad \boldsymbol{x} = (x_1, x_2) \in \mathbb{R}_+^2 \tag{3.57}$$

subject to the constraints

$$0 \leq x_1 \leq c_1, \quad 0 \leq x_2 \leq c_2, \quad x_1 + x_2 \leq c_0 \tag{3.58}$$

and to the restrictions that $x_r = 0$ if $n_r = 0$ (for $r = 1, 2$), i.e. if $\boldsymbol{x}^* = (x_1^*, x_2^*)$ solves this optimisation problem then we set $\boldsymbol{b}(\boldsymbol{n}) = \boldsymbol{x}^*$.

Suppose that $\boldsymbol{n} \in \mathbb{Z}_+^2$ is such that $n_1 = 0$ and $n_2 > 0$, then clearly $x_1^* = 0$ and $x_2^* = c_2$. Similarly if $\boldsymbol{n} \in \mathbb{Z}_+^2$ is such that $n_1 > 0$ and $n_2 = 0$, then clearly $x_1^* = c_1$ and $x_2^* = 0$. On the other hand suppose, $\boldsymbol{n} \in \mathbb{Z}_+^2$ is such that $n_1 \wedge n_2 > 0$. Then by Pareto efficiency $x_1 + x_2 = c_0$. It follows that

$$c_0 - c_2 \leq x_1 \leq c_1 \quad \text{and} \quad x_2 = c_0 - x_1. \tag{3.59}$$

Hence to calculate $\boldsymbol{x}^*$ we maximise the function

$$\hat{g}_{\boldsymbol{n}}^{\alpha}(x_1) = w_1 n_1^{\alpha} \frac{x_1^{1-\alpha}}{1-\alpha} + w_2 n_2^{\alpha} \frac{(c_0 - x_1)^{1-\alpha}}{1-\alpha} \tag{3.60}$$

subject to the constraint $c_0 - c_2 \leq x_1 \leq c_1$.

Observe by the standard theory of single variable calculus that on the interval $[0, \infty)$ the function $\hat{g}^{\alpha}_{\boldsymbol{n}}$ has a unique global maximum at

$$\hat{x}_1 = \frac{w_1^{1/\alpha} n_1 c_0}{w_1^{1/\alpha} n_1 + w_2^{1/\alpha} n_2}. \tag{3.61}$$

If $\hat{x}_1 \leq c_0 - c_2$, then $x_1^* = c_0 - c_2$. Likewise, if $\hat{x}_1 \geq c_1$, then $x_1^* = c_1$. Finally, if $c_0 - c_2 < \hat{x}_1 < c_1$, then $x_1^* = \hat{x}_1$. Combining the above results with (3.59) we obtain that $\boldsymbol{b}$ is an α-fair-sharing control strategy for $\mathcal{N}_2$ if

$$\boldsymbol{b}(\boldsymbol{n}) = \boldsymbol{x}^* = \begin{cases} (0,0) & \text{if} \quad n_1 = n_2 = 0 \\ (0, c_2) & \text{if} \quad n_1 = 0,\ n_2 > 0 \\ (c_1, 0) & \text{if} \quad n_1 > 0,\ n_2 = 0 \\ (c_0 - c_2, c_2) & \text{if} \quad \frac{n_1}{n_2} \leq \frac{w_1^{1/\alpha}}{w_2^{1/\alpha}} \frac{c_0 - c_2}{c_2} \\ \left(\frac{w_1^{1/\alpha} n_1 c_0}{w_1^{1/\alpha} n_1 + w_2^{1/\alpha} n_2}, \frac{w_2^{1/\alpha} n_2 c_0}{w_1^{1/\alpha} n_1 + w_2^{1/\alpha} n_2}\right) & \text{if} \quad \frac{w_1^{1/\alpha}}{w_2^{1/\alpha}} \frac{c_0 - c_2}{c_2} < \frac{n_1}{n_2} < \frac{w_1^{1/\alpha}}{w_2^{1/\alpha}} \frac{c_1}{c_0 - c_1} \\ (c_1, c_0 - c_1) & \text{if} \quad \frac{n_1}{n_2} \geq \frac{w_1^{1/\alpha}}{w_2^{1/\alpha}} \frac{c_1}{c_0 - c_1}. \end{cases} \tag{3.62}$$

If $\alpha = 1$ we can replace the quantities $x_r^{1-\alpha}/(1-\alpha)$ in (3.57) by $\log(x_r)$ (for $r = 1, 2$) and still obtain $\boldsymbol{x}^*$ as given by (3.62) as the solution to the new optimisation problem.

Now, we observe that the control strategies which satisfy (3.62) also satisfy conditions (3.42) and (3.43) of Corollary 3.3.2. $\square$

We note in the special case when $w_1 = w_2 = 1$ that α-fair-sharing control strategies for $\mathcal{N}_2$ are independent of the value of α.

Chapter 4

Workload-based techniques

In this chapter we use Lyapunov function techniques to establish the stability of Pareto efficient control strategies for various networks. We also compare these results with the so called fluid model. Finally we give an example of a network where these techniques can not be used to establish stability.

4.1 Workload function techniques

Suppose $\mathcal{N}$ is a general network. As in Chapter 3, we consider the problem of finding stable control strategies for $\mathcal{N}$. In particular, for a network $\mathcal{N}$ we consider Lyapunov function techniques based on the workload function w as defined by

$$w(\boldsymbol{n}) = \sum_{r \in \mathcal{R}} \frac{n_r}{\mu_r} \quad \text{for} \quad \boldsymbol{n} \in \mathbb{Z}_+^R. \tag{4.1}$$

Note that for any control strategy $\boldsymbol{b}$ for $\mathcal{N}$ we have

$$D_{\boldsymbol{b}} w(\boldsymbol{n}) = \sum_{r \in \mathcal{R}} (\kappa_r - b_r(\boldsymbol{n})) \text{ for all } \boldsymbol{n} \in \mathbb{Z}_+^R. \tag{4.2}$$

We recall that for any Pareto efficient control strategy $\boldsymbol{b}$ for the two-dimensional network $\mathcal{N}_2$, the workload function w satisfies the conditions (2.8) and (2.9) of Proposition 2.1.2, with $\epsilon = c_0 - (\kappa_1 + \kappa_2)$ and $F = \{\boldsymbol{n} \colon n_1 \wedge n_2 = 0\}$. Since F here is infinite, this is not sufficient for w to act as a Lyapunov function, and indeed we have observed that there are unstable Pareto efficient control strategies for this network.

A similar situation occurs in a wide variety of networks $\mathcal{N}$ with $\dim \mathcal{N} \geq 3$. Some simple function f, analogous to the workload function w satisfies the conditions of Proposition 2.1.2, either for some or for all Pareto efficient control strategies, with refuge F equal to some infinite subset of the boundary set $B = \{\boldsymbol{n}\colon n_{\min} = 0\}$ of $\mathbb{Z}_+^R$ where $n_{\min} = \min_{r \in \mathcal{R}} n_r$. However, for a Pareto efficient control strategy $\boldsymbol{b}$ for the network $\mathcal{N}$, the behaviour of the function $D_{\boldsymbol{b}} f$ may give insight into the question of the stability of $\boldsymbol{b}$. Specifically $D_{\boldsymbol{b}} f$ may give clues to the modification of $\boldsymbol{b}$ that is needed in order obtain stability and to the modification of f needed to obtain a corresponding Lyapunov function which can be used to establish stability of the modified $\boldsymbol{b}$. In particular, the stability of a Pareto efficient control strategy $\boldsymbol{b}$ can often be achieved by modifying the control strategy (and simultaneously the function f) close to the set B.

The intuitive idea here is that efficient use of the resources (as specified by the control strategy $\boldsymbol{b}$) ensures that the function $D_{\boldsymbol{b}} f$ is in general decreasing, but that such efficient use of the resources may not be possible for values of $\boldsymbol{n} \in B$. In such circumstances the control strategy $\boldsymbol{b}$ generally only requires small modifications in a neighbourhood of the boundary set B—thereby ensuring that the corresponding process $\boldsymbol{n}(\cdot)$ does not spend too much time on this boundary—in order to become stable.

We consider first those feasible control strategies $\boldsymbol{b}$ such that, for some $\delta_0 > 0$ and workload function w we have $D_{\boldsymbol{b}} w(\boldsymbol{n}) < -\delta_0$ for all $\boldsymbol{n} \in \mathbb{Z}_+^R$ such that $n_{\min} := \min_{r \in \mathcal{R}} n_r$ is sufficiently large. Proposition 4.1.2 below shows that, for such controls, only small modifications to the control strategy $\boldsymbol{b}$ are required to guarantee the stability of $\boldsymbol{b}$. As a first step, we require Lemma 4.1.1, which gives a simple condition for any feasible control strategy $\boldsymbol{b}$ for a network $\mathcal{N}$ to be stable. We note that Lemma 4.1.1 is a generalisation to higher dimension networks of the approach taken in Theorem 3.3.1.

To state Lemma 4.1.1 we define some notation as follows. For any $a \in \mathbb{Z}_+$, and for each $\boldsymbol{n} \in \mathbb{Z}_+^R$ define $\mathcal{R}_a(\boldsymbol{n}) = \{r\colon n_r \geq a\}$ and for any $\boldsymbol{n} \in \mathbb{Z}_+^R$, define $n_{\max} = \max_{r \in \mathcal{R}} n_r$.

Lemma 4.1.1. *Suppose that $\mathcal{N}$ is an arbitrary network. Then, given any $\delta > 0$,* *there exists a constant $a_\delta \geq 0$ (that depends only on the parameters of $\mathcal{N}$) such that if a feasible control strategy $\boldsymbol{b}$ for $\mathcal{N}$ which satisfies*

$$\sum_{r\in\mathcal{R}_{a'}(\boldsymbol{n})} (\kappa_r - b_r(\boldsymbol{n})) \leq -\delta \qquad \text{for all } 1 \leq a' \leq a_\delta, \text{ and } \boldsymbol{n} \in \mathbb{Z}_+^R \text{ such that } n_{\max} \geq a_\delta, \tag{4.3}$$

then $\boldsymbol{b}$ is stable.

Proof. Given $\delta > 0$, fix a_δ such that

$$a_\delta \geq \frac{1}{\delta} \sum_{r\in\mathcal{R}} (\kappa_r + c_{\max}) \tag{4.4}$$

where $c_{\max} = \max_{j\in\mathcal{J}}\{c_j\}$. Assume without loss of generality that

$$\delta < \min_{j\in\mathcal{J}} \left(c_j - \sum_{r\in\mathcal{R}} A_{jr}\kappa_r \right) := \hat{\delta}(\mathcal{N}) \tag{4.5}$$

(otherwise we can just choose a smaller $\bar{\delta} > 0$ such that $\bar{\delta} < \hat{\delta}(\mathcal{N})$ and set $a_\delta = a_{\bar{\delta}}$).

Now suppose that $\boldsymbol{b}$ is a feasible control strategy for a network $\mathcal{N}$ such that condition (4.3) is satisfied. To show that $\boldsymbol{b}$ is stable we apply Proposition 2.1.2.

As a Lyapunov function we use the *modified workload function* w^{a_δ} defined on $\mathbb{Z}_+^R$ by

$$w^{a_\delta}(\boldsymbol{n}) = \sum_{r\in\mathcal{R}} \frac{g^{a_\delta}(n_r)}{\mu_r}, \quad \text{for} \quad \boldsymbol{n} \in \mathbb{Z}_+^R. \tag{4.6}$$

where for $n \in \mathbb{Z}_+$,

$$g^{a_\delta}(n) = \begin{cases} n & \text{if } n \geq a_\delta, \\ \frac{a_\delta}{2}(1 + \frac{n^2}{a_\delta^2}) & \text{if } n < a_\delta. \end{cases} \tag{4.7}$$

We note that w^{a_δ} is a generalisation of the modified workload function introduced in Chapter 3. Then, as in Theorem 3.3.1,

$$D_{\boldsymbol{b}} w^{a_\delta}(\boldsymbol{n}) = dw^{a_\delta}(\boldsymbol{n}) + \frac{1}{2a_\delta} \sum_{r\in\mathcal{R}} h_r^{a_\delta}(n_r), \tag{4.8}$$

where

$$dw^{a_\delta}(\boldsymbol{n}) = \sum_{r\in\mathcal{R}} \min\left(\frac{n_r}{a_\delta}, 1\right) [\kappa_r - b_r(\boldsymbol{n})] \tag{4.9}$$

and for each $r \in \mathcal{R}$,

$$h_r^{a_\delta}(n_r) = \begin{cases} \kappa_r + b_r(\boldsymbol{n}) & \text{if } 0 \le n_r < a_\delta, \\ b_r(\boldsymbol{n}) & \text{if } n_r = a_\delta, \\ 0 & \text{if } n_r > a_\delta, \end{cases} \tag{4.10}$$

It is clear from the definition of a_δ and (4.10) that

$$\frac{1}{2a_\delta} \sum_{r\in\mathcal{R}} h_r^{a_\delta}(n_r) \le \frac{\delta}{2} \quad \text{for all} \quad \boldsymbol{n} \in \mathbb{Z}_+^R. \tag{4.11}$$

From (4.9) it is straightforward to check that

$$dw^{a_\delta}(\boldsymbol{n}) = \frac{1}{a_\delta} \sum_{a'=1}^{a_\delta} \sum_{r\in\mathcal{R}_{a'}(\boldsymbol{n})} (\kappa_r - b_r(\boldsymbol{n})) \quad \text{for all} \quad \boldsymbol{n} \in \mathbb{Z}_+^R. \tag{4.12}$$

It follows from (4.3) and (4.12) that

$$dw^{a_\delta}(\boldsymbol{n}) \le -\delta \quad \text{for} \quad \boldsymbol{n} \notin F_{a_\delta} \tag{4.13}$$

where $F_{a_\delta} = \{\boldsymbol{n} \colon n_{max} < a_\delta\}$. It now follows from (4.11) and (4.13) that

$$D_{\boldsymbol{b}} w^{a_\delta}(\boldsymbol{n}) < -\frac{\delta}{2} \quad \text{for} \quad \boldsymbol{n} \notin F_{a_\delta}. \tag{4.14}$$

It follows from (4.14) and Proposition 2.1.2 (with w^{a_δ} as our required Lyapunov function, $\epsilon = \delta/2$ and F_{a_δ} as our refuge) that $\boldsymbol{b}$ is stable. □

If $\delta < \hat{\delta}(\mathcal{N})$ then the lower bound over δ of the right hand side of (4.4) is affected by the slack in the inequalities (1.7). In particular, the lower bound for a_δ given by (4.4) is affected by $\hat{\delta}(\mathcal{N})$. The important observation here is that if, at each resource $j \in \mathcal{J}$, the total rate at which work arrives at j is close to the capacity of j then we need to have a relatively large value of a_δ to satisfy (4.4).

Proposition 4.1.2. *Suppose $\mathcal{N}$ is a network such that the condition* (1.7) *is satisfied. Let $\boldsymbol{b}$ be any feasible control strategy for $\mathcal{N}$ such that, for some $\delta_0 > 0$ and $a_0 \ge 1$ and workload function w,*

$$D_{\boldsymbol{b}} w(\boldsymbol{n}) \leq -\delta_0 \quad \textit{for all} \quad \boldsymbol{n} \in \mathbb{Z}_+^R \quad \textit{such that} \quad n_{\min} \geq a_0. \tag{4.15}$$

Then there exists $\hat{a} \in \mathbb{Z}_+$ and a feasible control strategy $\boldsymbol{b}'$ such that $\boldsymbol{b}'(\boldsymbol{n}) = \boldsymbol{b}(\boldsymbol{n})$ for all $\boldsymbol{n} \in \mathbb{Z}_+^R$ such that $n_{\min} \geq \hat{a}$ and such that $\boldsymbol{b}'$ is stable.

Proof. It follows from (1.7) that we can choose a vector $\hat{\boldsymbol{b}} = (\hat{b}_r, r \in \mathcal{R})$ such that

$$0 < \hat{\delta} := \min_{r \in \mathcal{R}} (\hat{b}_r - \kappa_r) \tag{4.16}$$

and

$$\sum_{r \in \mathcal{R}} A_{jr} \hat{b}_r \leq c_j \tag{4.17}$$

for all $j \in \mathcal{J}$. Let

$$\delta = \min(\delta_0, \hat{\delta}) \tag{4.18}$$

and let $a_\delta \geq a_0$ be as in Lemma 4.1.1. Set $\hat{a} = a_\delta$ and define the control strategy $\boldsymbol{b}'$ by

$$\boldsymbol{b}'(\boldsymbol{n}) = \boldsymbol{b}(\boldsymbol{n}), \qquad \text{if} \quad n_{\min} \geq \hat{a}, \tag{4.19}$$

$$b'_r(\boldsymbol{n}) = \hat{b}_r \mathbf{I}_{\{\boldsymbol{n} \in \mathbb{Z}_+^R :\, n_r > 0\}} \quad \text{for all} \quad r \in \mathcal{R}, \qquad \text{if} \quad n_{\min} < \hat{a}, \tag{4.20}$$

where $\mathbf{I}_A$ is the indicator function of the set A. Hence $\boldsymbol{b}'$ agrees with $\boldsymbol{b}$ when $n_{\min} \geq \hat{a}$ and otherwise $\boldsymbol{b}'$ acts as the complete partitioning control strategy based on the vector $\hat{\boldsymbol{b}}$.

To show that the control strategy $\boldsymbol{b}'$ is stable it is enough, by Lemma 4.1.1, to show that condition (4.3) holds for $\boldsymbol{b}'$ (with $a_\delta = \hat{a}$).

Suppose that $a' \leq \hat{a}$ and $n_{\max} \geq \hat{a}$. First, if we also have $n_{\min} \geq \hat{a}$, then $\mathcal{R}_{a'}(\boldsymbol{n}) = \mathcal{R}$ for all $a' \leq \hat{a}$. So by (4.2), (4.15) and (4.19)

$$\sum_{r \in \mathcal{R}_{a'}(\boldsymbol{n})} (\kappa_r - b'_r(\boldsymbol{n})) = \sum_{r \in \mathcal{R}} (\kappa_r - b'_r(\boldsymbol{n})) = \sum_{r \in \mathcal{R}} (\kappa_r - b_r(\boldsymbol{n})) \tag{4.21}$$

$$= D_{\boldsymbol{b}} w(\boldsymbol{n}) \leq -\delta_0 \leq -\delta \tag{4.22}$$

where the right-hand side of (4.21) is obtained since $\boldsymbol{b}'$ agrees with $\boldsymbol{b}$ in this case and the right-hand side of (4.22) follows from (4.15) and (4.18). On the other hand, if $n_{\min} < \hat{a}$, $n_{\max} \geq \hat{a}$, it follows from (4.16), (4.18) and (4.20) that

$$\sum_{r \in \mathcal{R}_{a'}(\boldsymbol{n})} (\kappa_r - b'_r(\boldsymbol{n})) = \sum_{r \in \mathcal{R}_{a'}(\boldsymbol{n})} (\kappa_r - b'_r) \leq -\hat{\delta} \leq -\delta \quad \text{for} \quad 1 \leq a' \leq a_\delta = \hat{a}. \tag{4.23}$$

So by (4.22) and (4.23), the control strategy $\boldsymbol{b}'$ satisfies the condition (4.3) of Lemma 4.1.1 and hence $\boldsymbol{b}'$ is stable. □

4.2 The tree network

There are many networks for which, under the condition (1.7), there exists $\delta_0 > 0$ such that, for $a_0 = 1$, condition (4.15) of Proposition 4.1.2 is satisfied for all Pareto efficient control strategies. One such class consists of networks, $\mathcal{N}_T$, with a *tree* topology. Such networks are defined as follows. Let $\mathcal{R} = \{1, \ldots, R\}$ denote the R routes, let $\mathcal{J} = \{0, \ldots, R\}$ denote the $R+1$ resources and suppose that the network topology is specified by,

$$A = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}. \tag{4.24}$$

In this network each route r requires capacity from a single dedicated resource $j = r$ of capacity c_r together with capacity from resource 0 which is shared by all routes and has capacity c_0.

As usual we assume condition (1.7) which for $\mathcal{N}_T$ becomes

$$\sum_{r \in \mathcal{R}} \kappa_r < c_0, \qquad \kappa_r < c_r, \qquad r = 1, \ldots, R. \tag{4.25}$$

We also assume, without loss of generality that $c_0 < \infty$ and that

$$\max_{r\in\mathcal{R}}\{c_r\} \le c_0 \le \sum_{r\in\mathcal{R}} c_r. \tag{4.26}$$

In particular, if for any $r \in \mathcal{R}$, we have $c_r > c_0$ then we can replace c_r by $c_r^{'} = c_0$. Similarly if $c_0 > \sum_{r\in\mathcal{R}} c_r$ then we can replace c_0 by $c_0^{'} = \sum_{r\in\mathcal{R}} c_r$. Therefore we also obtain $\max_{r\in\mathcal{R}}\{c_r\} < \infty$. We note that $\mathcal{N}_T$ is a natural generalisation of the network $\mathcal{N}_2$ considered in Chapter 3.

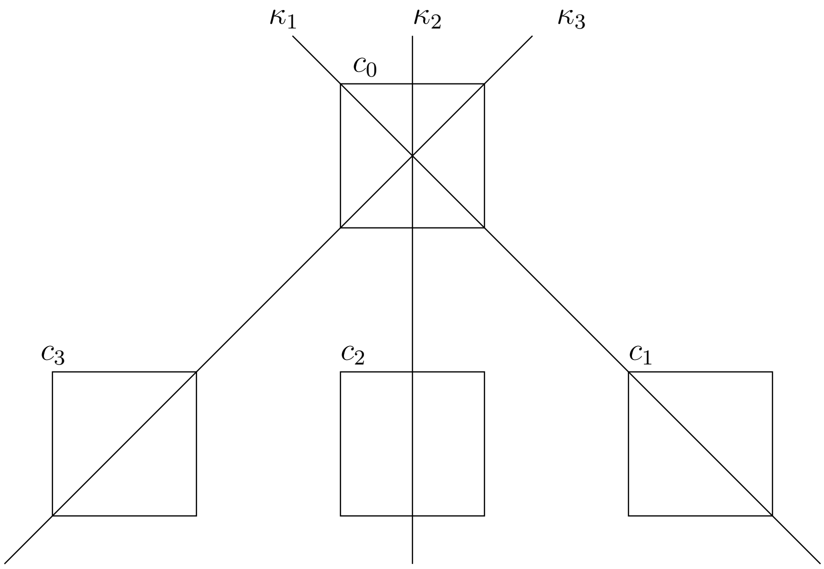

Figure 4.1 The tree network with $R = 3$.

The constraints (1.4) for feasible control strategies on $\mathcal{N}_T$ are given by

$$\sum_{r\in\mathcal{R}} b_r(\boldsymbol{n}) \le c_0 \tag{4.27}$$

$$b_r(\boldsymbol{n}) \le c_r, \qquad r \in \mathcal{R}. \tag{4.28}$$

It follows from (4.26) that for any Pareto efficient control strategy $\boldsymbol{b}$ for $\mathcal{N}_T$,

$$\sum_{r\in\mathcal{R}_1(\boldsymbol{n})} b_r(\boldsymbol{n}) = c_0 \wedge \left(\sum_{r\in\mathcal{R}_1(\boldsymbol{n})} c_r \right). \tag{4.29}$$

where $\tilde{\mathcal{R}}_0 = \{r \in \mathcal{R} \colon n_r > 0\}$. We note that if $n_{\min} > 0$ then $\tilde{\mathcal{R}}_0 = \mathcal{R}$ and so the right hand side of (4.29) becomes c_0. Hence, it follows from (4.2) and (4.29) that

$$D_{\boldsymbol{b}} w(\boldsymbol{n}) = \sum_{r \in \mathcal{R}} \kappa_r - c_0 \quad \text{for all } \boldsymbol{n} \text{ such that } n_{\min} > 0. \tag{4.30}$$

It now follows from (4.25), and from Proposition 4.1.2 with $\delta_0 = c_0 - \sum_{r \in \mathcal{R}} \kappa_r$ and $a_0 = 1$, that there exists some $\hat{a} \in \mathbb{Z}_+$ such that any Pareto efficient control strategy $\boldsymbol{b}$ only requires an appropriate modification on the set $\{\boldsymbol{n} \in \mathbb{Z}_+^R \colon n_{\min} < \hat{a}\}$ in order for $\boldsymbol{b}$ to achieve stability. In particular, for a tree network $\mathcal{N}_T$ a modified control strategy $\boldsymbol{b}'$ can be constructed by following the construction outlined in the proof of Proposition 4.1.2. We note that if $\boldsymbol{b}'$ is not Pareto efficient, we can modify $\boldsymbol{b}'$ on $\{\boldsymbol{n} \in \mathbb{Z}_+^R \colon n_{\min} < \hat{a}\}$ (as in Corollary 2.3.4) to obtain a Pareto efficient $\boldsymbol{b}''$ such that $\boldsymbol{b}''(\boldsymbol{n}) = \boldsymbol{b}(\boldsymbol{n})$ for all $\boldsymbol{n} \in \mathbb{Z}_+^R$ such that $n_{\min} \geq \hat{a}$. For the tree network $\mathcal{N}_T$, the application of Proposition 4.1.2 provides theoretical insight into why stability occurs (i.e. for a given Pareto efficient control strategy $\boldsymbol{b}$ it is only necessary to suitably modify $\boldsymbol{b}$ close to the set $\{\boldsymbol{n} \in \mathbb{Z}_+^R \colon n_{\min} = 0\}$). It should be noted that in practice, however the input parameter vector $\boldsymbol{\kappa}$ may not be known. For example, engineers working on the Internet may only know the capacity parameter vector $\boldsymbol{c}$ for some network $\mathcal{N}$ and the state $\boldsymbol{n}$ at any given time. Therefore it is desirable to choose a control strategy $\boldsymbol{b}$ which is stable for all values of the input parameter vector $\boldsymbol{\kappa}$. One such possibility is to consider a complete priority control strategy described below.

We say that $\boldsymbol{b}$ is a *complete priority control strategy* on a tree network $\mathcal{N}_T$ if there exists an $a > 0$ such that for all $a' \in [1, a]$

$$\sum_{r \in \mathcal{R}_{a'}(\boldsymbol{n})} b_r(\boldsymbol{n}) = c_0 \wedge \left(\sum_{r \in \mathcal{R}_{a'}(\boldsymbol{n})} c_r \right) \tag{4.31}$$

for all $\boldsymbol{n} \in \mathbb{Z}_+^R$ such that $n_{\min} < a$. Any control strategy for $\mathcal{N}_T$ satisfying (4.31) requires, for some $a > 0$, for all $\boldsymbol{n} \in \mathbb{Z}_+^R$ (such that $n_{\min} < a$) and for all $a' \in [1, a]$, that calls of type $r \in \mathcal{R}$, where $n_r \geq a'$ collectively have complete priority over calls of the remaining types. We note for complete priority control strategies that for $r \in \mathcal{R}$ and $a' \in [1, a]$ such that $n_r \geq a'$ we have that $b_r(\boldsymbol{n})$ is unaltered by $b_s(\boldsymbol{n})$ when $s \in \mathcal{R}$ and $n_s < a'$.

The following corollary of Lemma 4.1.1 shows that there exist values of $a \in$

$\mathbb{Z}_+$ such that the complete priority control strategy corresponding to a satisfies condition (4.3) of Lemma 4.1.1 and is therefore a stable control strategy.

Corollary 4.2.1. *There exists $\delta > 0$, such that if $\boldsymbol{b}$ is a complete priority control strategy for $\mathcal{N}_T$ with $a \geq a_\delta$ (where a_δ is defined by (4.4)), then $\boldsymbol{b}$ is stable.*

Remark Observe from the statement of Corollary 4.2.1 that for a given $\delta > 0$, the choice of a_δ which guarantees stability depends on the input parameter vector $\boldsymbol{\kappa}$, and in general increases as the values of this parameter increases. In this sense the corresponding complete priority control strategy $\boldsymbol{b}$ is not robust. However, except for values of the input parameter $\boldsymbol{\kappa}$ in which condition (1.7) is satisfied with approximate equality for at least one of the resources $j \in \mathcal{J}$, only a relatively small value of a_δ is required to ensure that the capacity is utilised with close to maximum efficiency, hence guaranteeing that entrainment is impossible and in turn guaranteeing stability of the control strategy $\boldsymbol{b}$. Further, in practice an engineer could deploy a complete priority control strategy and increase the value of a until the network "looks" stable.

Proof of Corollary 4.2.1. Let $\delta := \min_{j\in\mathcal{J}}(c_j - \sum_{r\in\mathcal{R}} A_{jr}\kappa_r) > 0$. For this value of δ define a_δ to satisfy (4.4).

Now suppose that $\boldsymbol{b}$ is a complete priority control strategy for $\mathcal{N}_T$ corresponding to $a \geq a_\delta$. For $a' \in [1, a]$, suppose first that $\boldsymbol{n} \in \mathbb{Z}_+^R$ is such that $\sum_{r\in\mathcal{R}_{a'}(\boldsymbol{n})} c_r < c_0$ then

$$\sum_{r\in\mathcal{R}_{a'}(\boldsymbol{n})} (\kappa_r - b_r(\boldsymbol{n})) = \sum_{r\in\mathcal{R}_{a'}(\boldsymbol{n})} \kappa_r - \sum_{r\in\mathcal{R}_{a'}(\boldsymbol{n})} c_r \leq -\delta. \tag{4.32}$$

On the other hand for $a' \in [1, a]$, suppose that $\boldsymbol{n} \in \mathbb{Z}_+^R$ is such that $c_0 \leq \sum_{r\in\mathcal{R}_{a'}(\boldsymbol{n})} c_r$, then

$$\sum_{r\in\mathcal{R}_{a'}(\boldsymbol{n})} (\kappa_r - b_r(\boldsymbol{n})) = \sum_{r\in\mathcal{R}_{a'}(\boldsymbol{n})} \kappa_r - c_0 \leq -\delta. \tag{4.33}$$

It follows from (4.32) and (4.33) that for all $\boldsymbol{n} \in \mathbb{Z}_+^R$ and $a' \in [1, a]$

$$\sum_{r\in\mathcal{R}_{a'}(\boldsymbol{n})} (\kappa_r - b_r(\boldsymbol{n})) \leq -\delta. \tag{4.34}$$

Therefore $\boldsymbol{b}$ satisfies condition (4.3) of Lemma 4.1.1 for $a \geq a_\delta$ and so the complete priority control strategy $\boldsymbol{b}$ corresponding to a is stable. $\square$

The following result for the tree network $\mathcal{N}_T$ is completely robust, in the sense that for any time $t \geq 0$ only the capacity parameter $\boldsymbol{c}$ and the state of the network $\boldsymbol{n} \in \mathbb{Z}_+^R$ need to be known in order to apply the control strategy provided that (4.25) holds. Corollary 4.2.2 says that stability is ensured for control strategies for $\mathcal{N}_T$ such that for each $\boldsymbol{n} \in \mathbb{Z}_+^R$ there is a hierarchy of prioritisation to certain call types, described as follows. For each $\boldsymbol{n} \in \mathbb{Z}_+^R$, a route $r \in \mathcal{R}$ with maximum value of n_r receives the maximum bandwidth available to it (in the case where there is more than one maximum value of n_r the maximum bandwidth available is received by one of these routes according to some prioritisation). We then "remove" this route from the network and the remaining bandwidth is then allocated iteratively in the same manner until all of the capacity is allocated. In order to state Corollary 4.2.2 we define for all $r \in \mathcal{R}$, $\boldsymbol{n} \in \mathbb{Z}_+^R$, the set $\mathcal{C}_r(\boldsymbol{n}) = \{r' \neq r \colon n_{r'} > n_r \text{ or } n_{r'} = n_r \text{ and } r' < r\}$.

Corollary 4.2.2. *Suppose that* $\boldsymbol{b}$ *is a control strategy for* $\mathcal{N}_T$ *such that for all* $r \in \mathcal{R}$

$$b_r(\boldsymbol{n}) = c_r \wedge \left(\left(c_0 - \sum_{r' \in \mathcal{C}_r(\boldsymbol{n})} c_{r'} \right) \vee 0 \right) \quad \textit{for} \quad \boldsymbol{n} \in \mathbb{Z}_+^R. \tag{4.35}$$

Then $\boldsymbol{b}$ *stable if and only if condition* (4.25) *holds.*

Proof of Corollary 4.2.2. Suppose that condition (4.25) does not hold. Then any feasible control strategy $\boldsymbol{b}$ for $\mathcal{N}_T$ is not stable by Lemma 2.3.2.

Now suppose that condition (4.25) holds. Let $\boldsymbol{b}$ be a control strategy that satisfies (4.35). It follows from (4.29) that $\boldsymbol{b}$ is a Pareto efficient control strategy. Let $\delta := \min_{j \in \mathcal{J}} (c_j - \sum_{r \in \mathcal{R}} A_{jr} \kappa_r) > 0$.

Next, suppose $a' \in \mathbb{Z}_+$ is fixed, but arbitrary. It follows from (4.35) that if $\sum_{r \in \mathcal{R}_{a'}(\boldsymbol{n})} c_r < c_0$ then

$$\sum_{r \in \mathcal{R}_{a'}(\boldsymbol{n})} (\kappa_r - b_r(\boldsymbol{n})) = \sum_{r \in \mathcal{R}_{a'}(\boldsymbol{n})} (\kappa_r - c_r) \leq -\delta. \tag{4.36}$$

On the other hand, if $\sum_{r\in\mathcal{R}_{a'}(\boldsymbol{n})} c_r \geq c_0$ then

$$\sum_{r\in\mathcal{R}_{a'}(\boldsymbol{n})} (\kappa_r - b_r(\boldsymbol{n})) = \sum_{r\in\mathcal{R}_{a'}(\boldsymbol{n})} (\kappa_r - c_0) \leq -\delta. \tag{4.37}$$

Since $a' \in \mathbb{Z}_+$ was arbitrary, it follows from (4.36) and (4.37) that for $\delta > 0$ as defined above, the control strategy $\boldsymbol{b}$ satisfies condition (4.3) of Lemma 4.1.1 and so $\boldsymbol{b}$ is stable. □

4.3 The backbone network

For other network topologies it is possible to obtain results that are analogous to Proposition 4.1.2 by following a similar approach. Again, we use Proposition 2.1.2 to determine stable control strategies for these networks.

We outline this approach for networks $\mathcal{N}_B$ with the *backbone* network topology. Such networks are are defined as follows. Let $\mathcal{R} = \{0, 1, \ldots, k\}$ denote the $k+1$ routes, let $\mathcal{J} = \{1, \ldots, k\}$ denote the k resources and suppose the network topology is specified by

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}. \tag{4.38}$$

In this network calls of each type $r = 1, \ldots, k$ require capacity from a single resource $j = r$ of capacity c_r, while calls of type 0 require capacity from each of the resources $1, \ldots, k$. For simplicity of exposition we assume that $\mu_r = 1$ for $r = 0, 1, \ldots k$ so that $\kappa_r = \nu_r$ for $r = 0, 1, \ldots k$.

Again we assume condition (1.7) which for $\mathcal{N}_B$ becomes

$$\kappa_0 + \kappa_r < c_r, \qquad r = 1, \ldots k. \tag{4.39}$$

When $\kappa_r = 0$ for $r = 2, \ldots k$ for the backbone network, this network again generalises the network of Example 1.1.1 with route 0 here playing the rôle of route 2 in Example 1.1.1.

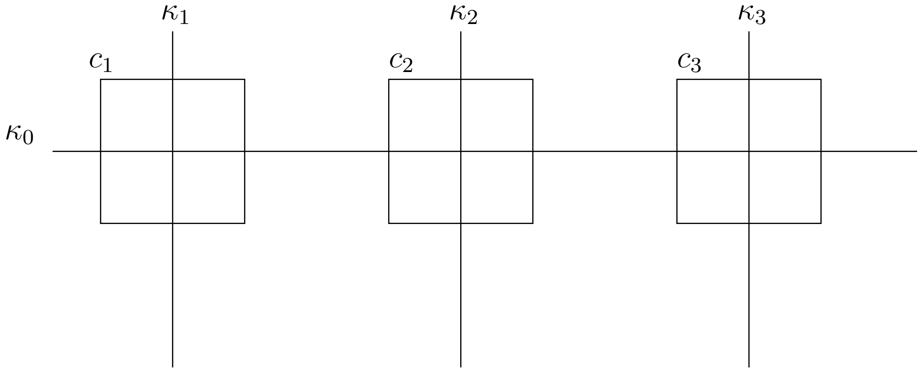

Figure 4.2 The backbone network with $k = 3$.

The state of the network is denoted by $\boldsymbol{n} = (n_0, n_1, \ldots, n_k)$. For any such $\boldsymbol{n} \in \mathbb{Z}_+^{k+1}$ we define $\hat{n}_{\min} = \min(n_1, \ldots, n_k)$ and $\hat{n}_{\max} = \max(n_1, \ldots, n_k)$.

Suppose that $\boldsymbol{b}$ is a feasible control strategy for $\mathcal{N}_B$. Then the feasibility constraints for this network are given by

$$b_0(\boldsymbol{n}) + b_r(\boldsymbol{n}) \le c_r, \qquad r = 1, \ldots k. \tag{4.40}$$

In addition for Pareto efficient control strategies $\boldsymbol{b}$, when $\hat{n}_{\max} > 0$,

$$b_0(\boldsymbol{n}) + b_r(\boldsymbol{n}) = c_r, \qquad \text{for all} \quad r = 1, \ldots, k \quad \text{such that} \quad n_r > 0. \tag{4.41}$$

On the other hand, when $\hat{n}_{\max} = 0$ and $n_0 > 0$, we have

$$b_0(\boldsymbol{n}) = \min_{r=1,\ldots,k} c_r. \tag{4.42}$$

Now consider the restricted workload function w_r corresponding to resource $r \in \mathcal{J}$ and defined by

$$w_r(\boldsymbol{n}) = n_0 + n_r \quad \text{for} \quad \boldsymbol{n} \in \mathbb{Z}_+^{k+1}. \tag{4.43}$$

Let

$$\delta' := \min_{1 \le r \le k} (c_r - \kappa_0 - \kappa_r) > 0. \tag{4.44}$$

Observe by (4.39) and (4.41) that for any Pareto efficient control strategy $\boldsymbol{b}$ for $\mathcal{N}_B$, for each $r = 1, \ldots, k$, and for any $\boldsymbol{n} \in \mathbb{Z}_+^{k+1}$ such that $n_r > 0$

$$D_{\boldsymbol{b}} w_r(\boldsymbol{n}) = \kappa_0 + \kappa_r - c_r \leq -\delta'. \tag{4.45}$$

Define the function $\hat{w}$ on $\mathbb{Z}_+^{k+1}$ by

$$\hat{w}(\boldsymbol{n}) = n_0 + \hat{n}_{\max} \quad \text{for} \quad \boldsymbol{n} \in \mathbb{Z}_+^R. \tag{4.46}$$

Observe from (4.45) that under (4.39) for any Pareto efficient control strategy $\boldsymbol{b}$ for $\mathcal{N}_B$

$$D_{\boldsymbol{b}} \hat{w}(\boldsymbol{n}) \leq -\delta' \tag{4.47}$$

except

(a) for $\boldsymbol{n} \in \mathbb{Z}_+^{k+1}$ such that $n_r = \hat{n}_{\max}$ for more than one $r \in \mathcal{R}$,

(b) for $\boldsymbol{n} \in \mathbb{Z}_+^R$ such that $\hat{n}_{\max} = 0$.

Since the case (a) defines an infinite set, $\hat{w}$ is almost a Lyapunov function.

In order to use Proposition 2.1.2 to obtain a stability result we modify the function $\hat{w}$ as follows.

(a) Smooth the function $\hat{n}_{\max}$ in the neighbourhood of those $\boldsymbol{n} \in \mathbb{Z}_+^{k+1}$ such that $n_r = \hat{n}_{\max}$ for several $r \in \mathcal{R}$, replacing it by a function ϕ, such that $D_{\boldsymbol{b}}\phi(\boldsymbol{n}) \leq -\delta'/2$ for all $\boldsymbol{n} \in \mathbb{Z}_+^{k+1}$ not belong to some finite set F.

(b) Define some modified candidate Lyapunov function $\hat{f}$ such that $\hat{f}(\boldsymbol{n}) = n_0 + g(\phi(\boldsymbol{n}))$ for some smooth function g such that for $r = 1, \ldots k$, the contribution of n_r in $\hat{f}$ tends to some constant as $n_r \to 0$.

In order for this function to be a Lyapunov function we need to restrict our consideration to control strategies which assign sufficient capacity to calls of type 0 when $\hat{n}_{\max}$ is small.

Theorem 4.3.1. *Suppose that $\mathcal{N}_B$ is a backbone network. Then, given any $\delta > 0$, there exists a constant $a_\delta \geq 0$ such that if a Pareto efficient control strategy $\boldsymbol{b}$ for $\mathcal{N}_B$ which satisfies*

$$\kappa_0 - b_0(\boldsymbol{n}) \leq -\delta \qquad \textit{whenever } \hat{n}_{\max} < a_\delta \textit{ and } n_0 > 0, \tag{4.48}$$

then $\boldsymbol{b}$ is stable.

Proof. Given a backbone network $\mathcal{N}_B$ and given $\delta > 0$, we show how a_δ can be chosen. Define $c^{\min} = \min_{r \in \mathcal{J}} c_r$. Assume without loss of generality that

$$\delta < c^{\min} - \kappa_0 \tag{4.49}$$

(otherwise we can just choose a smaller $\bar{\delta} > 0$ such that $\bar{\delta} < c^{\min} - \kappa_0$ and set $a_\delta = a_{\bar{\delta}}$). The choice of a_δ is based upon a careful analysis of the behaviour of a candidate Lyapunov function $\hat{f}_a$ where $a > 0$ is a parameter. As a first step, we outline the construction of $\hat{f}_a$.

We begin by noting that for any $a > 0$, by suitably smoothing the function $\hat{n}_{\max}$ in the region of those $\boldsymbol{n} \in \mathbb{Z}_+^{k+1}$ such that $n_r = \hat{n}_{\max}$ for more that one $r \in \mathcal{R}$, we may define a function ϕ_a on $\mathbb{R}_+^{k+1}$ with the following properties:

1. For all $\boldsymbol{n} \in \mathbb{Z}_+^{k+1} \subset \mathbb{R}_+^{k+1}$

$$\hat{n}_{\max} \leq \phi_a(\boldsymbol{n}) \leq \hat{n}_{\max} + a. \tag{4.50}$$

2. For all $\boldsymbol{n} \in \mathbb{R}_+^{k+1}$ and for all $r \in \mathcal{R}$

$$\frac{\partial \phi_a(\boldsymbol{n})}{\partial n_r} \geq 0. \tag{4.51}$$

3. For all $\boldsymbol{n} \in \mathbb{R}_+^{k+1}$

$$\sum_{r=1}^{k} \frac{\partial \phi_a(\boldsymbol{n})}{\partial n_r} = 1. \tag{4.52}$$

4. For $\boldsymbol{n} \in \mathbb{Z}_+^{k+1} \subset \mathbb{R}_+^{k+1}$ with $\hat{n}_{\max} \geq a$, for $r \in \mathcal{R}$ such that $n_r = 0$ we have

$$\frac{\partial \phi_a(\boldsymbol{n})}{\partial n_r} = 0. \tag{4.53}$$

5. For all $\boldsymbol{n} \in \mathbb{Z}_+^{k+1} \subset \mathbb{R}_+^{k+1}$ and for all $r \in \mathcal{R}$

$$\left| \frac{\partial \phi_a(\boldsymbol{n} + \boldsymbol{e}_r)}{\partial n_r} - \frac{\partial \phi_a(\boldsymbol{n})}{\partial n_r} \right| = O(a^{-1}) \tag{4.54}$$

uniformly on $\boldsymbol{n}$, where $\boldsymbol{e}_r = (e_{rs})_{s \in \mathcal{R}}$ is the $k+1$-dimensional unit vector with $e_{rr} = 1$ and $e_{rs} = 0$ for all $s \neq r$.

An example of such a function is given by the solution of the equation

$$\sum_{r=1}^{k} ((n_r - \phi_a(\boldsymbol{n}) + a)^+)^2 = a^2 \quad \text{for} \quad a \in \mathbb{Z}_+ \quad \text{and for} \quad \boldsymbol{n} \in \mathbb{R}_+^{k+1} \tag{4.55}$$

where $f(\boldsymbol{n})^+ := \max\{0, f(\boldsymbol{n})\}$ for any function $f : \mathbb{R}_+^{k+1} \to \mathbb{R}$.

For $k = 2$ and arbitrary $a > 0$, figure 4.3.2 illustrates the contours of ϕ_a, the "corners" of which are arcs of a circle of radius a.

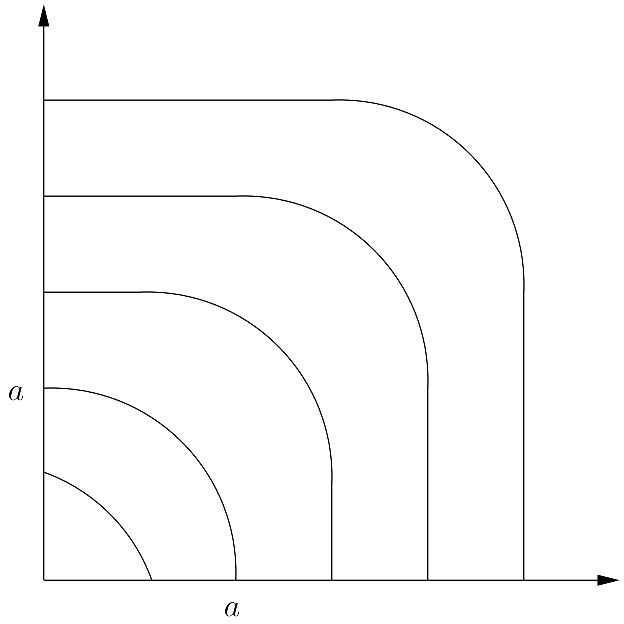

Figure 4.3 The contours of ϕ_a with $k = 2$.

We may deduce the properties 1, 2 ,4 and 5 from Figure 4.3.2. To deduce property 3 we use the following argument. For any $\gamma \in \mathbb{R}$ consider a $k+1$ dimensional vector of the form $\boldsymbol{\gamma} = (\gamma, \dots, \gamma) \in \mathbb{R}_+^{k+1}$. It follows from (4.55) that for any $\boldsymbol{n} \in \mathbb{R}_+^{k+1}$, for any $a \in \mathbb{Z}_+$ and for all $\gamma \in \mathbb{R}$

$$\phi_a(\boldsymbol{n} + \boldsymbol{\gamma}) = \phi_a(\boldsymbol{n}) + \gamma. \tag{4.56}$$

Now fix $\boldsymbol{n} \in \mathbb{R}_+^{k+1}$ and define the function $\psi_a^{\boldsymbol{n}} \colon \mathbb{R} \to \mathbb{R}$ by the composition

$$\gamma \to \boldsymbol{n} + \boldsymbol{\gamma} \to \phi_a(\boldsymbol{n} + \boldsymbol{\gamma}) = \phi_a(\boldsymbol{n}) + \gamma. \tag{4.57}$$

Then for any $\boldsymbol{n} \in \mathbb{R}_+^{k+1}$ and any $a \in \mathbb{Z}_+$

$$\frac{d}{d\gamma}(\psi_a^{\boldsymbol{n}}(\gamma)) = (1, 1, \ldots, 1) \begin{pmatrix} \frac{\partial \phi_a(\boldsymbol{n}+\boldsymbol{\gamma})}{\partial n_0} \\ \frac{\partial \phi_a(\boldsymbol{n}+\boldsymbol{\gamma})}{\partial n_1} \\ \vdots \\ \frac{\partial \phi_a(\boldsymbol{n}+\boldsymbol{\gamma})}{\partial n_k} \end{pmatrix} = \sum_{r=1}^{k} \frac{\partial \phi_a(\boldsymbol{n}+\boldsymbol{\gamma})}{\partial n_r} = 1, \tag{4.58}$$

where $(1, 1, \ldots 1)$ is the $k+1$-dimensional vector which has 1 for all its components. Now let $\gamma = 0$ to obtain (4.52).

Given a suitable function ϕ_a, we now define the corresponding Lyapunov function $\hat{f}_a$ on $\mathbb{R}_+^{k+1}$ by

$$\hat{f}_a(\boldsymbol{n}) = n_0 + g_a(\phi_a(\boldsymbol{n})) \quad \text{for} \quad \boldsymbol{n} \in \mathbb{R}_+^{k+1} \tag{4.59}$$

where g_a is a function on $\mathbb{R}$ defined by

$$g_a(n) = \begin{cases} n & \text{if} \quad n \geq 3a \\ \frac{a}{2}\left(1 + \frac{n^2}{a^2}\right) & \text{if} \quad 2a \leq n < 3a \\ \frac{a}{2} & \text{if} \quad n < 2a. \end{cases} \tag{4.60}$$

It follows from (4.54) that for any feasible control strategy $\boldsymbol{b}$ for $\mathcal{N}_B$

$$D_{\boldsymbol{b}}\hat{f}_a(\boldsymbol{n}) = d\hat{f}_a(\boldsymbol{n}) + O(a^{-1}) \tag{4.61}$$

uniformly on all $\boldsymbol{n} \in \mathbb{Z}_+^{k+1} \subset \mathbb{R}_+^{k+1}$, where

$$d\hat{f}_a(\boldsymbol{n}) = \kappa_0 - b_0(\boldsymbol{n}) + g_a'(\phi_a(\boldsymbol{n})) \sum_{r=1}^{k} \frac{\partial \phi_a(\boldsymbol{n})}{\partial n_r}(\kappa_r - b_r(\boldsymbol{n})) \tag{4.62}$$

$$= (1 - g_a'(\phi_a(\boldsymbol{n})))(\kappa_0 - b_0(\boldsymbol{n}))$$

$$+ g_a'(\phi_a(\boldsymbol{n})) \sum_{r=1}^{k} \frac{\partial \phi_a(\boldsymbol{n})}{\partial n_r}(\kappa_0 + \kappa_r - b_0(\boldsymbol{n}) - b_r(\boldsymbol{n})). \tag{4.63}$$

We note that (4.63) follows from (4.52). It follows from (4.61) that we may choose $\hat{a} \geq 0$ such that, for all $\boldsymbol{n} \in \mathbb{Z}_+^{k+1} \subset \mathbb{R}_+^{k+1}$ and feasible control strategies $\boldsymbol{b}$ for $\mathcal{N}_B$

$$D_{\boldsymbol{b}}\hat{f}_{\hat{a}}(\boldsymbol{n}) \leq d\hat{f}_{\hat{a}}(\boldsymbol{n}) + \frac{\delta}{2}. \tag{4.64}$$

Given $\hat{a}$ we set $a_\delta = 3\hat{a}$.

Consider the case when $\hat{n}_{\max} \geq a_\delta$, then by (4.50) and (4.60) $g'_{a_\delta}(\phi_{a_\delta}(\boldsymbol{n})) = 1$ and so

$$d\hat{f}_{a_\delta}(\boldsymbol{n}) = \sum_{r=1}^{k} \frac{\partial \phi_{a_\delta}(\boldsymbol{n})}{\partial n_r}(\kappa_0 + \kappa_r - b_0(\boldsymbol{n}) - b_r(\boldsymbol{n})) \leq -\delta, \tag{4.65}$$

by (4.51), (4.52) and (4.53). Now consider the case when $n_0 > 0$ and $\hat{n}_{\max} < a_\delta$. By the conditions of the theorem we have that $\kappa_0 - b_0(\boldsymbol{n}) \leq -\delta$. For $\hat{n}_{\max} < 2a_\delta/3$, we have by (4.50) and (4.60) that $g'_{a_\delta}(\phi_{a_\delta}(\boldsymbol{n})) = 0$ and so by (4.48)

$$d\hat{f}_{a_\delta}(\boldsymbol{n}) = \kappa_0 - b_0(\boldsymbol{n}) \leq -\delta. \tag{4.66}$$

Next, for $n_0 > 0$ and $2a_\delta/3 \leq \hat{n}_{\max} < a_\delta$, we have by (4.50), (4.52) and (4.60) that

$$d\hat{f}_{a_\delta}(\boldsymbol{n}) = \sum_{r=1}^{k} \frac{\partial \phi_{a_\delta}(\boldsymbol{n})}{\partial n_r} \left(\left(1 - \frac{\phi_{a_\delta}(\boldsymbol{n})}{a_\delta}\right)(\kappa_0 - b_0(\boldsymbol{n})) + \frac{\phi_{a_\delta}}{a_\delta}(\boldsymbol{n})\,(\kappa_0 + \kappa_r - b_0(\boldsymbol{n}) - b_r(\boldsymbol{n})) \right) \tag{4.67}$$

$$\leq -\delta \tag{4.68}$$

by the convexity of the right-hand side of (4.67). Hence by (4.64), (4.65), (4.66) and (4.68)

$$D\hat{f}_{a_\delta}(\boldsymbol{n}) \leq -\frac{\delta}{2} \quad \text{for} \quad \boldsymbol{n} \notin F_{a_\delta} \tag{4.69}$$

where $F_{a_\delta} = \{\boldsymbol{n} \in \mathbb{Z}_+^{k+1} : n_0 = 0, \hat{n}_{\max} < a_\delta\}$.

Hence (2.8) and (2.9) are satisfied and so it follows from Proposition 2.1.2 (with $\hat{f}_{a_\delta}$ as our required Lyapunov function, $\epsilon = \delta/2$ and F_{a_δ} as our refuge) that $\boldsymbol{b}$ is stable. □

Corollary 4.3.2. *Suppose that $\boldsymbol{b}$ is a Pareto control strategy for $\mathcal{N}_B$ such that for some $\delta > 0$ and the corresponding a_δ*

$$b_0(\boldsymbol{n}) = c^{\min} \quad \textit{whenever } \hat{n}_{\max} < a_\delta \quad \textit{and} \quad n_0 > 0, \tag{4.70}$$

then $\boldsymbol{b}$ is stable.

Proof. Let

$$\delta^{\max} = c^{\min} - \kappa_0 \tag{4.71}$$

then $\delta^{\max} > \delta > 0$ for any δ given by (4.49). Suppose that $\boldsymbol{b}$ is a Pareto efficient control strategy for $\mathcal{N}_B$ that satisfies (4.70). When $\hat{n}_{\max} < a_\delta$ and $n_0 \geq a_\delta$ we have

$$\begin{aligned} \kappa_0 - b_0(\boldsymbol{n}) =& \kappa_0 - c^{\min} \\ =& - \delta^{\max} < -\delta. \end{aligned} \tag{4.72}$$

Hence $\boldsymbol{b}$ satisfies (4.48) and so $\boldsymbol{b}$ is stable. □

4.4 The fluid model

In the simultaneous processor sharing model the call arrivals are Poisson distributed. This means that the call arrival process is non-smooth in the sense that "work" of each type $r \in \mathcal{R}$ does arrive steadily but instead arrives in random "chunks".

The stability and instability of the various control strategies may be further understood from the non-smooth nature of the call arrival process. Consider the analogous *fluid model* in which "work" of each type $r \in \mathcal{R}$ arrives steadily at rate κ_r and is processed at rate $b_r(\boldsymbol{n})$, where each n_r is now the volume of work of type $r \in \mathcal{R}$ in the network and where the control strategy $\boldsymbol{b}$ is again subject to constraints of the form (1.4). Consider again Example 1.1.1, then, under the condition (1.7) , it is easy to see that every Pareto efficient control strategy $\boldsymbol{b}$ is stable, in the sense here that the total volume of work in the system eventually becomes and remains zero. For our stochastic model, the possible modified control strategy discussed above (in the generalisation with the tree network topology of Example 1.1.1), in which $b_1(\boldsymbol{n})$ is kept small whenever n_1 is small and n_2 is large, may be seen as a smoothing operation forcing the behaviour of the stochastic model to follow more closely that of the fluid model.

In the analogous fluid model defined above, it is easily seen that, for networks with either the tree and backbone network topologies, condition (1.7) is sufficient for the stability of *any* Pareto efficient control strategy $\boldsymbol{b}$. In each case this follows, for example, by using the same Lyapunov function as for the stochastic model, except that in each case the functions g^{a_δ} and g_{a_δ} may both be replaced by the identity function. These examples illustrate a principle which seems likely to be true for more general network topologies, namely that when a control strategy is such that it is stable for the fluid model, then there is a closely approximating control strategy which is stable for the corresponding stochastic model.

4.5 The hypercube network

The results above show that for some network topologies condition (1.7) is sufficient for the stability of a Pareto efficient control strategy $\boldsymbol{b}$ as long as it is modified for values of $\boldsymbol{n} \in \mathbb{Z}_+^R$ close to the boundary of $\mathbb{Z}_+^R$. However this is not the case for all network topologies. Consider three-dimensional networks $\mathcal{N}_H$ with the *hypercube* topology which is defined as follows. Let $\mathcal{R} = \{1, 2, 3\}$ denote the set of three routes, let $\mathcal{J} = \{1, 2, 3\}$ denote the set of three resources and suppose that the network topology is specified by

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}. \tag{4.73}$$

where for simplicity of exposition $\nu_r = \nu$, $\mu_r = 1$ for $r = 1, 2, 3$ and $c_j = c$ for $j = 1, 2, 3$. On this network each route $r \in \mathcal{R}$ requires capacity from two resources (of capacity c). In addition each resource allocates bandwidth to a distinct pair of routes. Assume that the condition (1.7) is satisfied, i.e.

$$2\nu < c. \tag{4.74}$$

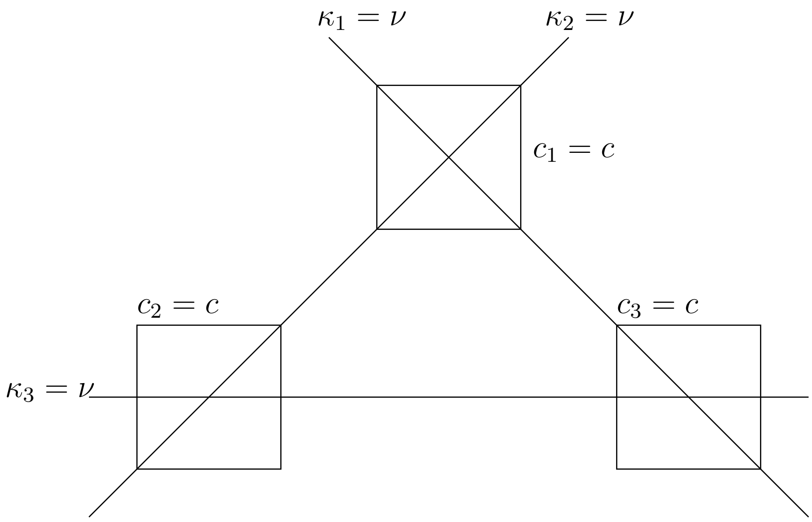

Figure 4.4 The hypercube network.

The feasibility constraints (1.4) for a control strategy $\boldsymbol{b}$ for $\mathcal{N}_H$ are given by

$$b_r(\boldsymbol{n}) + b_s(\boldsymbol{n}) \leq c \quad r = 1, 2, 3 \quad \text{for all} \quad s \neq r \quad \text{and} \quad \boldsymbol{n} \in \mathbb{Z}_+^3. \tag{4.75}$$

Consider any control strategy $\boldsymbol{b}'$ given by $\boldsymbol{b}'(\boldsymbol{n}) = (b'_1(\boldsymbol{n}), b'_2(\boldsymbol{n}), b'_3(\boldsymbol{n}))$ on $\mathcal{N}_H$ such that for all $\boldsymbol{n} \in \mathbb{Z}_+^3$ such that $\boldsymbol{n} \neq \boldsymbol{0}$ and for some fixed $r \in \mathcal{R}$ such that $n_r > 0$ we have

$$b'_r(\boldsymbol{n}) = c, \quad b'_s(\boldsymbol{n}) = 0 \quad \text{for all} \quad s \neq r \tag{4.76}$$

and when $n_r = 0$ we have $\sum_{s \neq r} b'_s(\boldsymbol{n}) = c$. We obtain that $\boldsymbol{b}'$ is a feasible and Pareto efficient control strategy for $\mathcal{N}_H$. However, Proposition 4.5.1 below shows that for some values of the parameter vectors $\boldsymbol{\kappa}$ and $\boldsymbol{c}$ the control strategy $\boldsymbol{b}'$ cannot be modified close to the boundary of $\mathbb{Z}_+^3$ to ensure stability. This can be seen as a consequence of the fact that for this control strategy there does not exist $\delta_0 > 0$ and $a_0 \geq 1$ which satisfy condition (4.15) of Proposition 4.1.2.

Proposition 4.5.1. *The Pareto efficient control strategy* $\boldsymbol{b}'$ *for* $\mathcal{N}_H$ *defined by* (4.76) *is unstable if*

$$2\nu < c \leq 3\nu. \tag{4.77}$$

Remark It follows conversely from Lemma 2.4.1 that if

$$3\nu < c \tag{4.78}$$

then any Pareto efficient control strategy $\boldsymbol{b}$ (including $\boldsymbol{b}'$) on $\mathcal{N}_H$ is stable.

Proof of Proposition 4.5.1. Suppose that $\boldsymbol{b}'$ is the Pareto efficient control strategy defined by (4.76). Consider the workload function w for the network

$$w(\boldsymbol{n}) = n_1 + n_2 + n_3 \quad \text{for} \quad \boldsymbol{n} \in \mathbb{Z}_+^3. \tag{4.79}$$

Then $w(\boldsymbol{n}) > w(\boldsymbol{0})$ and

$$D_{\boldsymbol{b}'} w(\boldsymbol{n}) = 3\nu - c \tag{4.80}$$

for $\boldsymbol{n} \neq \boldsymbol{0}$. Finally, for all $\boldsymbol{n} \in \mathbb{Z}_+^3$

$$\sum_{r \in \mathcal{R}} (\kappa_r |w(\boldsymbol{n} + \boldsymbol{e}_r) - w(\boldsymbol{n})| + b_r(\boldsymbol{n}) |w(\boldsymbol{n} - \boldsymbol{e}_r) - w(\boldsymbol{n})|) \leq A \tag{4.81}$$

for some $A < \infty$, where $\boldsymbol{e}_r = (e_{rs})_{s \in \mathcal{R}}$ is the three-dimensional unit with $e_{rr} = 1$ and $e_{rs} = 0$ for all $s \neq r$.

Hence if (4.77) holds conditions (2.21), (2.22) and (2.23) of Proposition 2.1.4 are satisfied (with $f = w$ and $F = \{\boldsymbol{0}\}$) and hence $\boldsymbol{b}'$ is unstable. □

In this network the instability of $\boldsymbol{b}'$ is not simply the result of a poor control strategy for $\boldsymbol{n} \in \mathbb{Z}_+^3$ close to the boundary of $\mathbb{Z}_+^3$, and this instability is equally present in the analogous fluid model.

Chapter 5

Stability of monotonic control strategies

Many control strategies likely to be of practical application possess a simple monotonicity property (for the definition of a monotonic control strategy $\boldsymbol{b}$ see (5.9) and (5.10) below). In this chapter we study stability for a wide class of such control strategies for general networks, giving sufficient conditions for stability, which for many classes of control strategy, are also close to being necessary.

5.1 Bounded control strategies

We say a control strategy $\boldsymbol{b}$ for a network $\mathcal{N}$ is *bounded* if, for all $r \in \mathcal{R}$, $b_r(\boldsymbol{n})$ is bounded in $\boldsymbol{n} \in \mathbb{Z}_+^R$. Note that the constraints (1.4) imply that all feasible control strategies are bounded. In this section we prove a key lemma for bounded control strategies. It asserts that for any route $r \in \mathcal{R}$, the long-term average input for $\mathcal{N}$ (i.e. $\kappa_r = \nu_r/\mu_r$) is at least as large as the long-term average bandwidth allocated to calls of type r.

Lemma 5.1.1. *Let $\boldsymbol{b}$ be a bounded control strategy for a network $\mathcal{N}$ and, as usual, let $\boldsymbol{n}(\cdot)$ denote the Markov process on $\mathbb{Z}_+^R$ defined by the parameters of $\mathcal{N}$ and the control strategy $\boldsymbol{b}$. Then, for any $r \in \mathcal{R}$,*

$$\limsup_{t\to\infty} \frac{1}{t}\int_0^t b_r(\boldsymbol{n}(u))\,du \le \kappa_r \quad \textit{almost surely.} \tag{5.1}$$

Proof. For any $r \in \mathcal{R}$, it can be shown that the process defined by

$$n_r^*(t) = n_r(t) + \int_0^t (b_r(\boldsymbol{n}(u)) - \kappa_r)du \quad t \geq 0 \tag{5.2}$$

is a zero-mean martingale (see for example Brémaud [8]). Since the process $\boldsymbol{n}(\cdot) = (n_r(\cdot), r \in \mathcal{R})$ is Markov with bounded transition rates, it is also straightforward to check that, for $r \in \mathcal{R}$, $\mathbf{E}(n_r^*(t))^2$ grows linearly in t. So, for some constant $M > 0$ and for $r \in \mathcal{R}$, we have $\mathbf{E}(n_r^*(t))^2 \leq Mt$ for all $t \geq 0$. Thus, for $r \in \mathcal{R}$ and for $1/2 < \alpha < 1$, the process $(n_r^*(t)/t^\alpha)_{t>0}$ is $\mathcal{L}^2$-bounded, and hence is a uniformly integrable supermartingale (see for example Williams [34]). So, for $1/2 < \alpha < 1$, there exists a random variable Z_r^α such that

$$\frac{n_r^*(t)}{t^\alpha} \to Z_r^\alpha \quad \text{almost surely as } t \to \infty. \tag{5.3}$$

It follows that for $r \in \mathcal{R}$

$$\frac{n_r(t)}{t} + \frac{1}{t}\int_0^t (b_r(\boldsymbol{n}(u)) - \kappa_r)du \to 0 \quad \text{almost surely as } t \to \infty. \tag{5.4}$$

Finally, since for $r \in \mathcal{R}$ we have $n_r(t) \geq 0$ for all $t \geq 0$, (5.1) follows from (5.3) and (5.4). □

In the case where the bounded control strategy $\boldsymbol{b}$ is stable, the Markov process $\boldsymbol{n}(\cdot)$ is positive recurrent, and so we have the stronger result that, for all $r \in \mathcal{R}$,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t b_r(\boldsymbol{n}(u))\, du = \mathbf{E}_\pi b_r = \kappa_r \quad \text{almost surely as } t \to \infty, \tag{5.5}$$

where

$$\mathbf{E}_\pi b_r = \sum_{\boldsymbol{n}\in\mathbb{Z}_+^R} \pi(\boldsymbol{n}) b_r(\boldsymbol{n}) \tag{5.6}$$

denotes the expectation of the function b_r with respect to the stationary distribution π of $\boldsymbol{n}(\cdot)$. The first equality in (5.5) follows from the ergodic theorem (see for example Norris [23]). The second equality of (5.5), readily deducible from the balance equations defining π, is simply the assertion that, under stationarity the expected arrival and departure rates for calls of type $r \in \mathcal{R}$ are equal.

5.2 Monotonic control strategies

We say that a bounded control strategy $\boldsymbol{b}$ for a network $\mathcal{N}$ is *monotonic* if, for all $r \in \mathcal{R}$ and for all $\boldsymbol{n} \in \mathbb{Z}_+^R$,

$$b_r(\boldsymbol{n}) \text{ is increasing in } n_r \text{ (with } n_s \text{ fixed for } s \in \mathcal{R} \text{ and } s \neq r\text{), and} \tag{5.7}$$

$$b_r(\boldsymbol{n}) \text{ is decreasing in } n_s \text{ (with } n_{s'} \text{ fixed for all } s' \in \mathcal{R} \text{ and } s' \neq s\text{),}$$

$$\text{for all } s \in \mathcal{R} \text{ and } s \neq r. \tag{5.8}$$

Note that, depending on the network topology, this property is natural in many applications. For instance, for networks $\mathcal{N}_T$ with the tree topology one can show that it is possessed by all the α-fair-sharing control strategies for $\mathcal{N}_T$. For more complex network structures, the results given below also apply for control strategies with similar properties to (5.7) and (5.8). These ideas are explored further in Section 5.4 below. Finally, we note that a related but somewhat different definition of monotonicity is used by Bonald and Proutière [6].

Given a monotonic control strategy $\boldsymbol{b}$ for a network $\mathcal{N}$, and $\mathcal{S} \subseteq \mathcal{R}$ (where $\mathcal{S}$ may be equal to the empty set $\emptyset$), define the function $\boldsymbol{b}^{\mathcal{S}} : \mathbb{Z}_+^S \to \mathbb{R}_+^R$ (where $S = |\mathcal{S}|$) by

$$b_r^{\mathcal{S}}(\boldsymbol{n}_{\mathcal{S}}) = \lim_{n_s \to \infty \,\forall\, s \notin \mathcal{S}} b_r(\boldsymbol{n}), \qquad r \in \mathcal{S}, \tag{5.9}$$

$$b_r^{\mathcal{S}}(\boldsymbol{n}_{\mathcal{S}}) = \lim_{n_r \to \infty} \lim_{n_s \to \infty \,\forall\, s \notin \mathcal{S} \cup \{r\}} b_r(\boldsymbol{n}), \qquad r \notin \mathcal{S}, \tag{5.10}$$

where $\boldsymbol{n}_{\mathcal{S}} = (n_s,\, s \in \mathcal{S})$ is the projection of $\boldsymbol{n} \in \mathbb{Z}_+^R$ onto the S-dimensional space $\mathbb{Z}_+^S$. Note that, by monotonicity of the control strategy $\boldsymbol{b}$, the function $\boldsymbol{b}^{\mathcal{S}}$ is well-defined. In particular, in (5.9), the order within $\mathcal{R} \backslash \mathcal{S}$ in which the limits are taken is irrelevant; however, in (5.10), the final limit must be taken as $n_r \to \infty$. We also note in the case when $\mathcal{S} = \mathcal{R}$ we have that $\boldsymbol{b}^{\mathcal{R}}$ equals $\boldsymbol{b}$. From the monotonicity of the control strategy $\boldsymbol{b}$ it follows that

$$b_s^{\mathcal{S}}(\boldsymbol{n}_{\mathcal{S}}) \leq b_s(\boldsymbol{n}) \quad \text{for all} \quad \boldsymbol{n} \in \mathbb{Z}_+^R \text{ and } s \in \mathcal{S}. \tag{5.11}$$

We use the function $\boldsymbol{b}^{\mathcal{S}}$ to define an auxiliary Markov process $\boldsymbol{n}^{\mathcal{S}}(\cdot) = (n_s^{\mathcal{S}}(\cdot),\ s \in \mathcal{S})$ with state space $\mathbb{Z}_+^S$ and transition rates given by

$$\boldsymbol{n}^{\mathcal{S}} \to \begin{cases} \boldsymbol{n}^{\mathcal{S}} + \boldsymbol{e}_s & \text{at rate } \nu_s, \\ \boldsymbol{n}^{\mathcal{S}} - \boldsymbol{e}_s & \text{at rate } \mu_s b_s^{\mathcal{S}}(\boldsymbol{n}_{\mathcal{S}}). \end{cases} \tag{5.12}$$

for $s \in \mathcal{S}$ (where $\boldsymbol{e}_s = (e_{ss'})_{s' \in \mathcal{S}}$ is the S-dimensional unit vector such that $e_{ss} = 1$ and $e_{ss'} = 0$ for $s' \neq s$). For a given $\mathcal{S} \subseteq \mathcal{R}$ we note that the parameters of the process $\boldsymbol{n}^{\mathcal{S}}(\cdot)$ are determined by a subset of the parameters of $\mathcal{N}$ and by $\boldsymbol{b}^{\mathcal{S}}$. The construction of the process $\boldsymbol{n}^{\mathcal{S}}(\cdot)$ in terms of this subset of the parameters of $\mathcal{N}$ along with inequality (5.11) enables us to couple the processes $\boldsymbol{n}^{\mathcal{S}}(\cdot)$ and $\boldsymbol{n}(\cdot)$ in a useful way. In particular, the proofs of Lemma 5.2.1 and Theorem 5.3.1 make use of the coupling of these processes.

We say that $\boldsymbol{b}^{\mathcal{S}}$ is *stable* if $\boldsymbol{n}^{\mathcal{S}}(\cdot)$ is a positive recurrent process. When $\boldsymbol{b}^{\mathcal{S}}$ is stable we let $\pi^{\mathcal{S}}$ denote the stationary distribution of $\boldsymbol{n}^{\mathcal{S}}(\cdot)$. For any function $f \colon \mathbb{Z}_+^S \to \mathbb{R}_+$, define

$$\mathbf{E}_{\pi^{\mathcal{S}}} f = \sum_{\boldsymbol{n}_{\mathcal{S}} \in \mathbb{Z}_+^S} \pi^{\mathcal{S}}(\boldsymbol{n}_{\mathcal{S}}) f(\boldsymbol{n}^{\mathcal{S}}) \tag{5.13}$$

to be the expected value of f under the distribution $\pi^{\mathcal{S}}$, provided the right hand side of (5.13) converges absolutely. In particular, expected values under the distribution $\pi^{\mathcal{S}}$ exist for bounded functions. In the case where $\mathcal{S}$ is the empty set $\emptyset$, we have that $\boldsymbol{b}^{\emptyset} = (b_r^{\emptyset},\ r \in \mathcal{R})$ is a vector of constants. We adopt the natural assumption that $\boldsymbol{b}^{\emptyset}$ is always stable (i.e. we assume that the upper bound of $\boldsymbol{b}$ is large enough). In this case the distribution $\pi^{\emptyset}$ is concentrated on a single point, and we have $\mathbf{E}_{\pi^{\emptyset}} b_r^{\emptyset} = b_r^{\emptyset}$ for all $r \in \mathcal{R}$.

For any time $t \geq 0$, let $\boldsymbol{n}_{\mathcal{S}}(t) = (n_s(t),\ s \in \mathcal{S})$ denote the projection of $\boldsymbol{n}(t) \in \mathbb{Z}_+^R$ onto the S-dimensional space $\mathbb{Z}_+^S$. With this notation we can now state and prove the following technical lemma which is required in the next section.

Lemma 5.2.1. *Suppose that $\boldsymbol{b}$ is a monotonic control strategy for a network $\mathcal{N}$ and suppose, for some non-empty $\mathcal{S} \subseteq \mathcal{R}$, that $\boldsymbol{b}^{\mathcal{S}}$ is stable. Then for any monotonic decreasing function $f \colon \mathbb{Z}_+^S \to \mathbb{R}_+$*

$$\liminf_{t\to\infty} \frac{1}{t}\int_0^t f(\boldsymbol{n}_{\mathcal{S}}(u))du \geq \mathbf{E}_{\pi^{\mathcal{S}}} f \quad \textit{almost surely.} \tag{5.14}$$

Proof. Since the control strategy $\boldsymbol{b}$ is monotonic it follows from (5.11) that we can couple the R-dimensional process $\boldsymbol{n}(\cdot)$ driven by $\boldsymbol{b}$ and the S-dimensional process $\boldsymbol{n}^{\mathcal{S}}(\cdot)$ driven by $\boldsymbol{b}^{\mathcal{S}}$ defined above, in such a way that

$$n_s(t) \leq n_s^{\mathcal{S}}(t) \quad \text{for all } t \geq 0 \quad \text{and for all } s \in \mathcal{S}. \tag{5.15}$$

Given such a coupling, we have for any monotonic decreasing function f,

$$\liminf_{t\to\infty} \frac{1}{t}\int_0^t f(\boldsymbol{n}_{\mathcal{S}}(u))du \geq \liminf_{t\to\infty} \frac{1}{t}\int_0^t f(\boldsymbol{n}^{\mathcal{S}}(u))du$$

$$= \mathbf{E}_{\pi^{\mathcal{S}}} f \tag{5.16}$$

where (5.16) follows from the ergodic theorem since $\boldsymbol{b}^{\mathcal{S}}$ is stable. □

5.3 Stability with prioritised call types

The first part of Theorem 5.3.1 below is similar in spirit to results of Borovkov [7] for *asymptotically spatially homogeneous* Markov processes. However in the case of the simultaneous processor sharing network with monotonic control strategies, it is the monotonicity of the control strategy itself that provides sufficient structure to obtain the results of Theorem 5.3.1. We now consider monotonic control strategies which give sufficient priority to routes belonging to some $\mathcal{S} \subset \mathcal{R}$ such that $\boldsymbol{b}^{\mathcal{S}}$ is stable. The first part of Theorem 5.3.1 provides a sufficient condition for stability of $\boldsymbol{b}^{\mathcal{S}\cup\{r\}}$ while the second part of the theorem provides a sufficient condition for the instability of $\boldsymbol{b}^{\mathcal{S}\cup\{r\}}$, for some $r \notin \mathcal{S}$.

Theorem 5.3.1. *Suppose that $\boldsymbol{b}$ is a monotonic feasible control strategy for a network $\mathcal{N}$ and suppose that $\mathcal{S} \subset \mathcal{R}$ is such that $\boldsymbol{b}^{\mathcal{S}}$ is stable.*

(i) If $r \notin \mathcal{S}$ is such that

$$\mathbf{E}_{\pi^{\mathcal{S}}} b_r^{\mathcal{S}} > \kappa_r, \tag{5.17}$$

then $\boldsymbol{b}^{\mathcal{S}\cup\{r\}}$ is stable.

(ii) If $r \notin \mathcal{S}$ is such that

$$\mathbf{E}_{\pi^{\mathcal{S}}} b_r^{\mathcal{S}} < \kappa_r, \tag{5.18}$$

then $\boldsymbol{b}^{\mathcal{S}\cup\{r\}}$ is unstable.

Remark Given the stability of $\boldsymbol{b}^{\mathcal{S}}$ for some $\mathcal{S} \subset \mathcal{R}$ (recall that, as already remarked, $\boldsymbol{b}^{\emptyset}$ is always stable), Theorem 5.3.1 gives criteria for determining the stability or otherwise of $\boldsymbol{b}^{\mathcal{S}\cup\{r\}}$ for any $r \notin \mathcal{S}$, except only in the case of equality in (5.17) or (5.18) (where the natural conjecture is that $\boldsymbol{b}^{\mathcal{S}\cup\{r\}}$ is unstable—see also the remarks at the end of Section 5.4). Recursive application of the theorem thus yields sufficient conditions both for the stability and the instability of monotonic control strategies. However, note that for example in the two-dimensional network $\mathcal{N}_2$ of Chapter 3 where $\mathcal{R} = \{1, 2\}$, $\boldsymbol{b}^{\{1\}}$ and $\boldsymbol{b}^{\{2\}}$ may both be unstable, while $\boldsymbol{b} = \boldsymbol{b}^{\{1,2\}}$ is stable, as is the case for α-fair-sharing control strategies for $\mathcal{N}_2$. In such circumstances Theorem 5.3.1 does not settle the question of the stability of $\boldsymbol{b}$. Rather its primary application is to control strategies in which there is a sufficient hierarchy of prioritisation as to permit the recursive application of the first part of the theorem to at least establish the stability of $\boldsymbol{b}^{\mathcal{S}}$ for $\mathcal{S} = \mathcal{R} \setminus \{r'\}$ for some $r' \in \mathcal{R}$. The theorem then also (in general) settles the question of the stability of $\boldsymbol{b}$ itself.

Proof of Theorem 5.3.1. Since, for given $\mathcal{S} \subset \mathcal{R}$ and $r \notin \mathcal{S}$, the stability of $\boldsymbol{b}^{\mathcal{S}\cup\{r\}}$ corresponds to the positive recurrence of the Markov process $\boldsymbol{n}^{\mathcal{S}\cup\{r\}}(\cdot)$ defined above, it is sufficient to prove the results (i) and (ii) in the case where $\mathcal{S} = \mathcal{R}\setminus\{r'\}$ for some $r' \in \mathcal{R}$ (i.e. $\mathcal{R} = \mathcal{S} \cup \{r'\}$) , and we henceforth assume this. The primary advantage of assuming this is that we avoid some unpleasant notational complexity. We identify each $\boldsymbol{n} \in \mathbb{Z}_+^R$ with the pair $(\boldsymbol{n}_{\mathcal{S}}, n_{r'})$ where $\boldsymbol{n}_{\mathcal{S}} = (n_s,\ s \in \mathcal{S})$. Also, recall that for each such $\boldsymbol{n}_{\mathcal{S}}$, we have

$$b_{r'}^{\mathcal{S}}(\boldsymbol{n}_{\mathcal{S}}) = \lim_{n_{r'}\to\infty} b_{r'}(\boldsymbol{n}_{\mathcal{S}}, n_{r'}). \tag{5.19}$$

(i) Suppose first that a given monotonic feasible control strategy $\boldsymbol{b}$ for a network $\mathcal{N}$ is such that for some $r' \in \mathcal{R}$ condition (5.17) holds. Let $\boldsymbol{n}(\cdot)$ be the corresponding Markov process on $\mathbb{Z}_+^R$ that is driven by $\boldsymbol{b}$. We require to show that

$\boldsymbol{b}$ is stable. The underlying idea of the proof of this part of the theorem is that the monotonicity of $\boldsymbol{b}$ and the stability of $\boldsymbol{b}^{\mathcal{S}}$ for $\mathcal{S} = \mathcal{R} \setminus \{r'\}$, ensure that the components $\{n_s(\cdot),\ s \in \mathcal{S}\}$ of the process $\boldsymbol{n}(\cdot)$ become and remain small, and the condition (5.17) then ensures that, except in some finite region A, the remaining component $n_{r'}(\cdot)$ of this process is decreasing at a rate which is bounded away from zero; thus the process $\boldsymbol{n}(\cdot)$ spends, in the long term, a nonzero proportion of time within A. To make this rigorous, we exploit a suitable coupling of the processes $\boldsymbol{n}(\cdot)$ and $\boldsymbol{n}^{\mathcal{S}}(\cdot)$ by using the monotonicity of $\boldsymbol{b}$.

To begin we make some simple observations. First, by condition (5.17) we can choose $\delta > 0$ such that

$$\sum_{\boldsymbol{n}_{\mathcal{S}} \in \mathbb{Z}_+^{R-1}} \pi^{\mathcal{S}}(\boldsymbol{n}_{\mathcal{S}}) b_{r'}^{\mathcal{S}}(\boldsymbol{n}_{\mathcal{S}}) = \mathbf{E}_{\pi^{\mathcal{S}}} b_{r'}^{\mathcal{S}} \geq \kappa_{r'} + 3\delta. \tag{5.20}$$

Now choose a compact set $A = \{\boldsymbol{n} \in \mathbb{Z}_+^R \colon n_r \leq \bar{n}_r,\ r \in \mathcal{R}\} \subset \mathbb{Z}_+^R$ — where $\{\overline{n}_r, r \in \mathcal{R}\}$ are fixed constants— such that

$$\sum_{\boldsymbol{n}_{\mathcal{S}} \in A^{\mathcal{S}}} \pi^{\mathcal{S}}(\boldsymbol{n}_{\mathcal{S}}) b_{r'}^{\mathcal{S}}(\boldsymbol{n}_{\mathcal{S}}) > \kappa_{r'} + 2\delta. \tag{5.21}$$

where $A^{\mathcal{S}}$ is the S-dimensional projection of A onto $\mathbb{Z}_+^{R-1}$. We also define the function $\bar{b}_{r'} \colon \mathbb{Z}_+^S \to \mathbb{R}_+$ by

$$\bar{b}_{r'}(\boldsymbol{n}_{\mathcal{S}}) = \begin{cases} 0 & \text{if } \boldsymbol{n}_{\mathcal{S}} \notin A^{\mathcal{S}} \\ b_{r'}(\boldsymbol{n}_{\mathcal{S}}, \bar{n}_{r'}) & \text{if } \boldsymbol{n}_{\mathcal{S}} \in A^{\mathcal{S}}, \end{cases} \tag{5.22}$$

where $\bar{n}_{r'}$ is a given constant, and we note it follows from the monotonicity of $\boldsymbol{b}$ that $\bar{b}_{r'}$ is decreasing in each of its arguments. We also have

$$b_{r'}(\boldsymbol{n}) \geq \bar{b}_{r'}(\boldsymbol{n}_{\mathcal{S}}) \qquad \text{for all } \boldsymbol{n} \notin A. \tag{5.23}$$

This follows trivially from (5.22), except for $\boldsymbol{n} \in \mathbb{Z}_+^R$ such that $n_s \leq \bar{n}_s$ for $s \in \mathcal{S}$ and $n_{r'} > \bar{n}_{r'}$. But in this case (5.23) follows from the monotonicity of $\boldsymbol{b}$. Finally, by choosing $\{\bar{n}_r, r \in \mathcal{R}\}$, all sufficiently large, we have that

$$\sum_{\boldsymbol{n}_{\mathcal{S}} \in \mathbb{Z}_+^{R-1}} \pi^{\mathcal{S}}(\boldsymbol{n}_{\mathcal{S}}) \bar{b}_{r'}(\boldsymbol{n}_{\mathcal{S}}) = \mathbf{E}_{\pi^{\mathcal{S}}} \bar{b}_{r'} > \kappa_{r'} + \delta, \tag{5.24}$$

by (5.17) and the monotone convergence theorem (see for example Freedman [13]) since, for any $\boldsymbol{n}_{\mathcal{S}}$,

$$\lim_{\bar{n}_{r'} \to \infty} b_{r'}(\boldsymbol{n}_{\mathcal{S}}, \bar{n}_{r'}) = b_{r'}^{\mathcal{S}}(\boldsymbol{n}_{\mathcal{S}}). \tag{5.25}$$

It now follows from Lemma 5.1.1 that, almost surely,

$$\begin{aligned}
\kappa_{r'} &\geq \limsup_{t \to \infty} \frac{1}{t} \int_0^t b_{r'}(\boldsymbol{n}(u))\, du \\
&\geq \limsup_{t \to \infty} \frac{1}{t} \int_0^t b_{r'}(\boldsymbol{n}(u))\, \mathbf{I}_{\{\boldsymbol{n}(u) \notin A\}}\, du \\
&\geq \limsup_{t \to \infty} \frac{1}{t} \int_0^t \bar{b}_{r'}(\boldsymbol{n}_{\mathcal{S}}(u))\, \mathbf{I}_{\{\boldsymbol{n}(u) \notin A\}}\, du \qquad (5.26) \\
&\geq \limsup_{t \to \infty} \frac{1}{t} \int_0^t \bar{b}_{r'}(\boldsymbol{n}_{\mathcal{S}}(u))\, du - b_{r'}^{\max} \liminf_{t \to \infty} \frac{1}{t} \int_0^t \mathbf{I}_{\{\boldsymbol{n}(u) \in A\}}\, du \qquad (5.27)
\end{aligned}$$

where $b_{r'}^{\max} = \max_{j \in \mathcal{J}} \{c_j \colon A_{jr'} = 1\}$ (and therefore is bounded), where $\mathbf{I}_E$ is the indicator function of a set E and where the inequality (5.26) follows from (5.23). Therefore by applying Lemma 5.2.1

$$\kappa_{r'} \geq \mathbf{E}_{\pi^{\mathcal{S}}} \bar{b}_{r'} - b_{r'}^{\max} \liminf_{t \to \infty} \frac{1}{t} \int_0^t I_{\{\boldsymbol{n}(u) \in A\}}\, du \tag{5.28}$$

Hence by (5.24)

$$\kappa_{r'} \geq \kappa_{r'} + \delta - b_{r'}^{\max} \liminf_{t \to \infty} \frac{1}{t} \int_0^t I_{\{\boldsymbol{n}(u) \in A\}}\, du. \tag{5.29}$$

It follows that

$$\liminf_{t \to \infty} \frac{1}{t} \int_0^t I_{\{\boldsymbol{n}(u) \in A\}}\, du \geq \frac{\delta}{b_{r'}^{\max}}. \tag{5.30}$$

Since A is finite it now follows from the ergodic theorem that the Markov process $\boldsymbol{n}(\cdot)$ is positive recurrent and so $\boldsymbol{b}$ is stable.

(ii) Now suppose instead that a given monotonic feasible control strategy $\boldsymbol{b}$ on a network $\mathcal{N}$ is such that for some $r' \in \mathcal{R}$ condition (5.18) holds. Let $\boldsymbol{n}(\cdot)$ be the corresponding Markov process on $\mathbb{Z}_+^R$ that is driven by $\boldsymbol{b}$. We show that $\boldsymbol{n}(\cdot)$ is transient and hence $\boldsymbol{b}$ is unstable.

The underlying idea of the proof of part (ii) is that whenever $n_{r'}(\cdot)$ is very large, the process $\boldsymbol{n}(\cdot)$ again behaves approximately as if it were controlled by $\boldsymbol{b}^{\mathcal{S}}$, and thus, from (5.18), we may expect that $\lim_{t\to\infty} n_{r'}(t) = \infty$ almost surely. To make this rigorous we again use the monotonicity of $\boldsymbol{b}$ to couple the process $\boldsymbol{n}(\cdot)$ to a process $\hat{\boldsymbol{n}}_{\bar{m}}(\cdot) = (\hat{n}_{r,\bar{m}}(\cdot), r \in \mathcal{R})$ with a control strategy which is sufficiently close to $\boldsymbol{b}^{\mathcal{S}}$. We then show that $\lim_{t\to\infty} \hat{n}_{r',\bar{m}}(t) = \infty$ almost surely, and use the coupling of $\boldsymbol{n}(\cdot)$ and $\hat{\boldsymbol{n}}_{\bar{m}}(\cdot)$ to ensure that also $\lim_{t\to\infty} n_{r'}(t) = \infty$ with strictly positive probability. This proves that $\boldsymbol{n}(\cdot)$ is transient as required.

Suppose $\bar{m} \in \mathbb{Z}_+$ it a fixed constant (its value to be chosen later), and define the Markov process $\hat{\boldsymbol{n}}_{\bar{m}}(\cdot) = (\hat{n}_{r,\bar{m}}(\cdot),\ r \in \mathcal{R})$, with state space $\mathbb{Z}_+^S \times \mathbb{Z}$, as follows: for each $s \in \mathcal{S}$, the component process $\hat{n}_{s,\bar{m}}(\cdot)$ has state space $\mathbb{Z}_+$ as usual, while $\hat{n}_{r',\bar{m}}(\cdot)$ has state space $\mathbb{Z}$; the transition rates for $\hat{\boldsymbol{n}}_{\bar{m}}(\cdot)$ are given by,

$$\hat{\boldsymbol{n}}_{\bar{m}} \to \begin{cases} \hat{\boldsymbol{n}}_{\bar{m}} + \boldsymbol{e}_r & \text{at rate} \quad \nu_r, \\ \hat{\boldsymbol{n}}_{\bar{m}} - \boldsymbol{e}_r & \text{at rate} \quad \mu_r \hat{b}_{r,\bar{m}}(\hat{\boldsymbol{n}}_{\bar{m}}) \end{cases} \tag{5.31}$$

for each $r \in \mathcal{R}$ (where $\boldsymbol{e}_r = (e_{rs'})_{s'\in\mathcal{R}}$ is the R-dimensional unit vector such that $e_{rr} = 1$ and $e_{rs'} = 0$ for $s' \neq r$), where $\hat{\boldsymbol{b}}_{\bar{m}} = (\hat{b}_{r,\bar{m}}(\boldsymbol{n}))_{r\in\mathcal{R}}$ is defined by

$$\hat{b}_{s,\bar{m}}(\boldsymbol{n}) := b_s(\boldsymbol{n}_{\mathcal{S}}, \bar{m}), \qquad s \in \mathcal{S},\ \boldsymbol{n} \in \mathbb{Z}_+^S \times \mathbb{Z} \tag{5.32}$$

$$\hat{b}_{r',\bar{m}}(\boldsymbol{n}) := b_{r'}^{\mathcal{S}}(\boldsymbol{n}_{\mathcal{S}}) \qquad \boldsymbol{n} \in \mathbb{Z}_+^S \times \mathbb{Z}. \tag{5.33}$$

From (5.32) and (5.33) we observe that the process $\hat{\boldsymbol{n}}_{\bar{m}}(\cdot)$ has uniformly bounded transition rates which are independent of the r'-th coordinate of the process $\boldsymbol{n}(\cdot)$. It follows from (5.32) and the monotonicity of $\boldsymbol{b}$ that

$$\hat{b}_{s,\bar{m}}(\boldsymbol{n}) \geq b_s^{\mathcal{S}}(\boldsymbol{n}_{\mathcal{S}}), \qquad s \in \mathcal{S},\ \boldsymbol{n} \in \mathbb{Z}_+^S \times \mathbb{Z}. \tag{5.34}$$

Now let $\hat{\boldsymbol{n}}_{\mathcal{S},\bar{m}}(\cdot) = (\hat{n}_{s,\bar{m}}(\cdot),\ s \in \mathcal{S})$ be the projection of the process $\hat{\boldsymbol{n}}_{\bar{m}}(\cdot)$ onto the state space $\mathbb{Z}_+^S$. As a first step in the proof we compare the projection process $\hat{\boldsymbol{n}}_{\mathcal{S},\bar{m}}(\cdot)$ to the process $\boldsymbol{n}^{\mathcal{S}}(\cdot)$ driven by $\boldsymbol{b}^{\mathcal{S}}$. We begin by constructing a coupling of the processes $\hat{\boldsymbol{n}}_{\mathcal{S},\bar{m}}(\cdot)$ and $\boldsymbol{n}^{\mathcal{S}}(\cdot)$. From (5.34) it is clear that we can construct both of these processes via the natural coupling such that

$$\hat{\boldsymbol{n}}_{\mathcal{S},\bar{m}}(0) = \boldsymbol{n}^{\mathcal{S}}(0) = \mathbf{0} \tag{5.35}$$

and such that

$$\hat{n}_{s,\bar{m}}(t) \leq n_s^{\mathcal{S}}(t), \quad \text{for all} \quad s \in \mathcal{S} \quad \text{and for all} \quad t \geq 0. \tag{5.36}$$

Since $\boldsymbol{n}^{\mathcal{S}}(\cdot)$ is assumed positive recurrent with stationary distribution $\pi^{\mathcal{S}}$, it follows from (5.36) that $\hat{\boldsymbol{n}}_{\mathcal{S},\bar{m}}(\cdot)$ is similarly positive recurrent with stationary distribution $\hat{\pi}_{\mathcal{S},\bar{m}}$, say.

Next we show that as $\bar{m}$ tends to infinity, $\hat{\pi}_{\mathcal{S},\bar{m}}$ converges in distribution to $\pi^{\mathcal{S}}$. We note it is enough to show that as $\bar{m}$ tends to infinity, $\hat{\pi}_{\mathcal{S},\bar{m}}(Z)$ converges to $\pi^{\mathcal{S}}(Z)$ for any set $Z \subseteq \mathbb{Z}_+^S$ of the form

$$Z = \{\boldsymbol{n} \colon n_s \geq z_s, s \in \mathcal{S}\}, \tag{5.37}$$

where $\{z_s, s \in \mathcal{S}\}$ are fixed constants. For each such $Z \subseteq \mathbb{Z}_+^S$ and each $\bar{m} \in \mathbb{Z}_+$, let $\hat{T}_{Z,\bar{m}}$ be the time spent by the process $\hat{\boldsymbol{n}}_{\mathcal{S},\bar{m}}(\cdot)$ in the set Z, prior to its first return to the state $\boldsymbol{n}_{\mathcal{S}} = \mathbf{0}$. Similarly, let $T_Z^{\mathcal{S}}$ be the time spent by the process $\boldsymbol{n}^{\mathcal{S}}(\cdot)$ in the set Z, prior to its first return to the state $\boldsymbol{n}_{\mathcal{S}} = \mathbf{0}$. In addition, for each $\bar{m} \in \mathbb{Z}$, let $\hat{T}_{\mathbb{Z}_+^S,\bar{m}}$ be the time that the process $\hat{\boldsymbol{n}}_{\mathcal{S},\bar{m}}(\cdot)$ takes to make its first return to the state $\boldsymbol{n}_{\mathcal{S}} = \mathbf{0}$ and let $T_{\mathbb{Z}_+^S}^{\mathcal{S}}$ be the time that the process $\boldsymbol{n}^{\mathcal{S}}(\cdot)$ takes to make its first return to the state $\boldsymbol{n}_{\mathcal{S}} = \mathbf{0}$. Thus from (5.36) it is clear that for each such $Z \subseteq \mathbb{Z}_+^S$ and $\bar{m} > 0$,

$$\hat{T}_{Z,\bar{m}} \leq T_Z^{\mathcal{S}}. \tag{5.38}$$

Furthermore from (5.9) and (5.32), for all $\boldsymbol{n} \in \mathbb{Z}_+^R$ and $s \in \mathcal{S}$ it follows that

$\hat{b}_{s,\bar{m}}(\boldsymbol{n}) = b_s(\boldsymbol{n}_{\mathcal{S}}, \bar{m})$ tends to $\boldsymbol{b}^{\mathcal{S}}(\boldsymbol{n}_{\mathcal{S}})$ as $\bar{m} \to \infty$. It follows that

$$p_{\bar{m}} := \mathbf{P}(\hat{T}_{Z,\bar{m}} \neq T_Z^{\mathcal{S}}) \to 0 \quad \text{as } \bar{m} \to \infty. \tag{5.39}$$

Now for each $\bar{m} \in \mathbb{Z}_+$, we also have

$$\begin{aligned} 0 \leq \mathbf{E}T_Z^{\mathcal{S}} - \mathbf{E}\hat{T}_{Z,\bar{m}} &= \mathbf{E}((T_Z^{\mathcal{S}} - \hat{T}_{Z,\bar{m}}) \cdot \mathbf{I}_{\{T_Z^{\mathcal{S}} \neq \hat{T}_{Z,\bar{m}}\}}) \\ &\leq \mathbf{E}(T_Z^{\mathcal{S}} \cdot \mathbf{I}_{\{T_Z^{\mathcal{S}} \neq \hat{T}_{Z,\bar{m}}\}}) \qquad (5.40) \\ &= \mathbf{E}(T_Z^{\mathcal{S}} \cdot \mathbf{I}_{\{T_Z^{\mathcal{S}} \neq \hat{T}_{Z,\bar{m}}\}} \cdot \mathbf{I}_{\{T_Z^{\mathcal{S}} \leq (p_{\bar{m}})^{-1/2}\}}) \\ &\quad + \mathbf{E}(T_Z^{\mathcal{S}} \cdot \mathbf{I}_{\{T_Z^{\mathcal{S}} \neq \hat{T}_{Z,\bar{m}}\}} \cdot \mathbf{I}_{\{T_Z^{\mathcal{S}} > (p_{\bar{m}})^{-1/2}\}}) \\ &\leq (p_{\bar{m}})^{1/2} + \mathbf{E}(T_Z^{\mathcal{S}} \cdot \mathbf{I}_{\{T_Z^{\mathcal{S}} > (p_{\bar{m}})^{-1/2}\}}). \qquad (5.41) \end{aligned}$$

where (5.40) follows from (5.38) and where $\mathbf{I}_E$ is the indicator function of a set E.

We note that it follows from (5.39) and dominated convergence that the right hand side of (5.41) tends to zero as $\bar{m}$ tends to infinity, and hence

$$\mathbf{E}\hat{T}_{Z,\bar{m}} \to \mathbf{E}T_Z^{\mathcal{S}} \quad \text{as } \bar{m} \to \infty, \tag{5.42}$$

for each set $Z \subseteq \mathbb{Z}_+^S$ of the form (5.37). In particular, by the same argument we have

$$\mathbf{E}\hat{T}_{\mathbb{Z}_+^S,\bar{m}} \to \mathbf{E}T_{\mathbb{Z}_+^S}^{\mathcal{S}} \quad \text{as } \bar{m} \to \infty. \tag{5.43}$$

Finally, since

$$\frac{\mathbf{E}\hat{T}_{Z,\bar{m}}}{\mathbf{E}\hat{T}_{\mathbb{Z}_+^S}} = \hat{\pi}_{\mathcal{S},\bar{m}}(Z) \quad \text{and} \quad \frac{\mathbf{E}T_Z^{\mathcal{S}}}{\mathbf{E}T_{\mathbb{Z}_+^s}^{\mathcal{S}}} = \pi^{\mathcal{S}}(Z) \tag{5.44}$$

for each set $Z \subseteq \mathbb{Z}_+^S$ of the form (5.37), it follows from (5.42), (5.43) and (5.44) that

$$\hat{\pi}_{\mathcal{S},\bar{m}}(Z) \to \pi^{\mathcal{S}}(Z) \quad \text{as } \bar{m} \to \infty. \tag{5.45}$$

Since (5.45) holds for all $Z \subset \mathbb{Z}_+^S$ of the form (5.37), it follows that $\hat{\pi}_{S,\bar{m}}$ converges in distribution to $\pi^{\mathcal{S}}$ as $\bar{m}$ tends to infinity. Since also $b_{r'}^{\mathcal{S}}$ is bounded, it now follows from (5.18) that we may choose the constant $\bar{m}' \in \mathbb{Z}$ sufficiently large (and fixed for the remainder of the proof) such that

$$\mathbf{E}_{\hat{\pi}_{\mathcal{S}}} b_{r'}^{\mathcal{S}} < \kappa_{r'}, \tag{5.46}$$

where

$$\hat{\pi}_{\mathcal{S}} := \hat{\pi}_{\mathcal{S},\bar{m}'}. \tag{5.47}$$

Finally, we couple the processes $\hat{\boldsymbol{n}}_{\bar{m}'}(\cdot)$ and $\boldsymbol{n}(\cdot)$ in such a way that the transience of $\hat{n}_{r',\bar{m}'}(\cdot)$, the r'-th component of $\hat{\boldsymbol{n}}_{\bar{m}'}(\cdot)$, implies the transience of $n_{r'}(\cdot)$ and hence the transience of $\boldsymbol{n}(\cdot)$. Define the random time

$$T_{\bar{m}'} = \min\{t > 0 \colon n_{r'}(t) < \bar{m}'\}. \tag{5.48}$$

We note that for all $\boldsymbol{n} \in \mathbb{Z}_+^R$ such that $n_{r'}(0) \geq \bar{m}'$, we have

$$\hat{b}_{r,\bar{m}'}(\boldsymbol{n}) \geq b_r(\boldsymbol{n}) \quad \text{for all} r \in \mathcal{R}. \tag{5.49}$$

Now suppose that the initial states of the processes $\hat{\boldsymbol{n}}_{\bar{m}'}(\cdot)$ and $\boldsymbol{n}(\cdot)$ are such that

$$\hat{\boldsymbol{n}}_{\bar{m}'}(0) = \boldsymbol{n}(0) \quad \text{and} \quad \hat{n}_{r',\bar{m}'}(0) = n_{r'}(0) > \bar{m}'. \tag{5.50}$$

Using the natural coupling, it follows from (5.48) and (5.49) that

$$\hat{n}_{r,\bar{m}'}(t) \leq n_r(t), \qquad \text{for} \quad 0 \leq t \leq T_{\bar{m}'} \quad \text{and for all} \quad r \in \mathcal{R}. \tag{5.51}$$

As noted above, the projection process $\hat{\boldsymbol{n}}_{\mathcal{S},\bar{m}'}(\cdot) = (\hat{n}_{s,\bar{m}'}(\cdot),\, s \in \mathcal{S})$ has stationary distribution $\hat{\pi}_{\mathcal{S}}$, while the process $\hat{n}_{r',\bar{m}'}(\cdot)$ maybe viewed as a Markov additive process modulated by the remaining components $\hat{\boldsymbol{n}}_{\mathcal{S},\bar{m}'}(\cdot)$ of $\hat{\boldsymbol{n}}_{\bar{m}'}(\cdot)$. From (5.46), the expectation of the increments of $\hat{n}_{r',\bar{m}'}(\cdot)$ between those times at which $\hat{\boldsymbol{n}}_{\mathcal{S},\bar{m}'}(\cdot)$ returns to any fixed state is strictly positive. It follows from (5.46) and the standard theory of Markov additive processes (see for example Asmussen [1]) that $\lim_{t\to\infty} \hat{n}_{r',\bar{m}'}(t) = \infty$ almost surely, and further that, under (5.50),

$$\mathbf{P}(\hat{n}_{r',\bar{m}'}(t) \geq \bar{m}' \text{ for all } t \geq 0,\ \lim_{t\to\infty} \hat{n}_{r',\bar{m}'}(t) = \infty) > 0, \tag{5.52}$$

and hence, from (5.51), that also

$$\mathbf{P}(n_{r'}(t) \geq \bar{m}' \text{ for all } t \geq 0,\ \lim_{t\to\infty} n_{r'}(t) = \infty) > 0. \tag{5.53}$$

Hence the process $\boldsymbol{n}(\cdot)$ is transient as required. □

5.4 Example with the hypercube network

Consider again the network with the hypercube topology $\mathcal{N}_H$ of Section 4.5 (with $\mu = 1$). As previously observed a necessary and sufficient condition for the existence of *some* stable control strategy is given by $2\nu < c$. Further, if $3\nu < c$, then Lemma 2.4.1 with the Lyapunov function f given by

$$f(\boldsymbol{n}) = \sum_{r=1}^{3} n_r \quad \text{for} \quad \boldsymbol{n} \in \mathbb{Z}_+^3 \tag{5.54}$$

shows that *any* Pareto efficient control $\boldsymbol{b}$ strategy is stable. Suppose now that $2\nu < c$ and that the Pareto efficient control strategy $\boldsymbol{b}$ is such that, for $r = 1, 2$, $b_r(\boldsymbol{n})$ is independent of n_3 and

$$b_1(\boldsymbol{n}) + b_2(\boldsymbol{n}) = c \quad \text{for all} \quad \boldsymbol{n} \in \mathbb{Z}_+^3 \quad \text{such that} \quad \max(n_1, n_2) > 0. \tag{5.55}$$

Hence

$$b_3(\boldsymbol{n}) = \begin{cases} \min\bigl(b_1(\boldsymbol{n}), b_2(\boldsymbol{n})\bigr) & \text{if } \max(n_1, n_2) > 0 \\ c & \text{if } \max(n_1, n_2) = 0. \end{cases} \tag{5.56}$$

Thus in particular calls of types 1 and 2 collectively have complete priority over calls of type 3. Although we do not, in this example, require any further monotonicity conditions on $\boldsymbol{b}$, it follows from the requirement of Pareto efficiency that the function $\boldsymbol{b}^{\{1,2\}} \colon \mathbb{Z}_+^2 \to \mathbb{R}_+^3$ is well-defined as before, being obtained from $\boldsymbol{b}$ by letting $n_3 \to \infty$. Since, for $r = 1, 2$, $b_r(\boldsymbol{n})$ is independent of calls of type 3, we see that calls of type 1 and 2 jointly have only one constraint. Thus in effect calls of types 1 and 2 behave as a network with one resource of capacity c and input rate 2ν. It follows from Theorem 2.2.1 and (5.55) that the condition

$$2\nu < c \tag{5.57}$$

is necessary and sufficient to ensure that $\boldsymbol{b}^{\{1,2\}}$ is stable. We use (a slight modification of) Theorem 5.3.1 to investigate the stability of $\boldsymbol{b}$. The stationary distribution $\pi^{\{1,2\}}$ on $\mathbb{Z}_+^2$ of the process $\boldsymbol{n}^{\{1,2\}}(\cdot)$ driven by $\boldsymbol{b}^{\{1,2\}}$ is, in this case, just that of the process $(n_1(\cdot), n_2(\cdot))$. Further, since, from (5.55), $(n_1 + n_2)(\cdot)$ is Markov, with a stationary distribution which is geometric and independent of any more detailed specification of $\boldsymbol{b}^{\{1,2\}}$, it follows that

$$\pi^{\{1,2\}}(0,0) = 1 - \frac{2\nu}{c}. \tag{5.58}$$

It follows from (1.4), (5.55) and the Pareto efficiency of $\boldsymbol{b}$ that

$$b_3^{\{1,2\}}(0,0) = c, \qquad b_3^{\{1,2\}}(n_1,0) = 0, \qquad b_3^{\{1,2\}}(0,n_2) = 0 \quad \text{for all } n_1, n_2 \geq 1. \tag{5.59}$$

We thus have that

$$\begin{aligned} \mathbf{E}_{\pi^{\{1,2\}}} b_3^{\{1,2\}} &= \sum_{(n_1,n_2)\in\mathbb{Z}_+^2} \pi^{\{1,2\}}(n_1,n_2) b_3^{\{1,2\}}(n_1,n_2) \\ &\geq c - 2\nu, \end{aligned} \tag{5.60}$$

with equality if and only if

$$b_3^{\{1,2\}}(n_1,n_2) = 0 \quad \text{for all } (n_1,n_2) \text{ such that } \min(n_1,n_2) \geq 1 \tag{5.61}$$

But from (5.56) condition (5.61) holds if and only if, for all (n_1, n_2) such that $\min(n_1, n_2) \geq 1$, we have

$$\min\bigl(b_1^{\{1,2\}}(n_1,n_2), b_2^{\{1,2\}}(n_1,n_2)\bigr) = 0 \tag{5.62}$$

i.e. in the case of the control strategy considered in Section 4.5 in which maximum resource is always allocated to calls of one type, and in which we have already observed that we have stability if and only if $3\nu < c$. Otherwise we have strict inequality in (5.60). Also since

$$b_3^{\{1,2\}}(n_1, n_2) \leq \frac{c}{2} \quad \text{for all} \quad (n_1, n_2) \quad \text{such that} \quad \min(n_1, n_2) \geq 1, \tag{5.63}$$

we have from (5.58) that

$$\mathbf{E}_{\pi^{\{1,2\}}} b_3^{\{1,2\}} < c - \nu. \tag{5.64}$$

Hence by (5.60) and (5.64) for each value of $\boldsymbol{b}^{\{1,2\}}$ there exists a constant k, with $0 \leq k < \nu$, such that

$$\mathbf{E}_{\pi^{\{1,2\}}} b_3^{\{1,2\}} = c - 2\nu + k. \tag{5.65}$$

Now note that, although $\boldsymbol{b}$ does not satisfy all the conditions for monotonicity given earlier, the assumption that $b_1(\boldsymbol{n})$ and $b_2(\boldsymbol{n})$ are independent of n_3 ensures that Theorem 5.3.1 continues to apply, indeed in a slightly improved form, to show that the condition

$$\mathbf{E}_{\pi^{\{1,2\}}} b_3^{\{1,2\}} > \nu \tag{5.66}$$

is necessary and sufficient for the stability of $\boldsymbol{b}$. For the sufficiency, note that the proof of part (i) of the theorem, with $\mathcal{S} = \{1, 2\}$ and $r' = 3$, goes through as before, except that the coupling between $\boldsymbol{n}(\cdot)$ and process $\boldsymbol{n}^{\mathcal{S}}(\cdot)$ is now obtained with equality i.e. (5.15) is replaced by: for $r = 1, 2$

$$n_r(t) = n_r^{\{1,2\}}(t) \quad \text{for all } t \geq 0, \tag{5.67}$$

and so we no longer require the condition that $\bar{b}_3$ is decreasing in each of its arguments in order to apply Lemma 5.2.1 to obtain (5.28). Similar obvious simplifications apply to the proof of part (ii), which here becomes a fairly standard argument. Further, consider the case when

$$\mathbf{E}_{\pi^{\{1,2\}}} b_3^{\{1,2\}} = \nu. \tag{5.68}$$

We show that the process $\boldsymbol{n}(\cdot)$ is not positive recurrent. Consider the process $\tilde{\boldsymbol{n}}(\cdot) = (\tilde{n}_r(\cdot))_{r=1,2,3}$ on the space $\mathbb{Z}_+^2 \times \mathbb{Z}$ which for $r = 1, 2, 3$ has transition rates

$$\tilde{\boldsymbol{n}} \to \begin{cases} \tilde{\boldsymbol{n}} + \boldsymbol{e}_r & \text{at rate} \quad \nu \\ \tilde{\boldsymbol{n}} - \boldsymbol{e}_r & \text{at rate} \quad b_r^{\{1,2\}}(n_1, n_2), \end{cases} \tag{5.69}$$

where $\boldsymbol{e}_r = (e_{rs})_{s=1,2,3}$ is the three-dimensional unit vector such that $e_{rr} = 1$ and $e_{rs} = 0$ for $s \neq r$. By the coupling of Theorem 5.3.1 notice that if $n_r(0) = \tilde{n}_r(0)$ for $r = 1, 2, 3$ then for all $t \geq 0$

$$n_r(t) = \tilde{n}_r(t) \quad \text{for } r = 1, 2 \quad \text{and} \quad \tilde{n}_3(t) \leq n_3(t). \tag{5.70}$$

Let

$$T = \inf_{t>0}\{t \colon \boldsymbol{n}(t) = \boldsymbol{0}\} \tag{5.71}$$

and let

$$\tilde{T} = \inf_{t>0}\{t \colon \tilde{n}_1(t) = \tilde{n}_2(t) = 0,\, \tilde{n}_3(t) \leq 0\}. \tag{5.72}$$

It follows from (5.70) that $T \geq \tilde{T}$ and in particular

$$\mathbf{E}T \geq \mathbf{E}\tilde{T}. \tag{5.73}$$

Now consider the one-dimensional process $\tilde{n}_3(\cdot)$ on $\mathbb{Z}$ which has transition rates defined by (5.69) for $r = 3$. Since the transition rates of $\tilde{n}_3(\cdot)$ do not depend on n_3 it is clear that $\tilde{n}_3(\cdot)$ is a homogeneous null-recurrent random walk on $\mathbb{Z}$, since it follows from condition (5.68) that the mean increment of $\tilde{n}_3(\cdot)$ is zero. In particular $\mathbf{E}\tilde{T} = \infty$ and from (5.73) we have that

$$\mathbf{E}T = \infty. \tag{5.74}$$

Hence the process $\boldsymbol{n}(t)$ is not positive recurrent.

Suppose now that c and $\boldsymbol{b}$ are held fixed and that ν is allowed to vary. For any value of ν, let $\boldsymbol{n}(\cdot)$ be the process corresponding to ν and for some $\nu' < \nu$

let $\boldsymbol{n}'(\cdot) = (n'_r(\cdot))_{r=1,2,3}$ be process corresponding to ν'. The obvious coupling argument shows that if $n_r(0) \leq n'_r(0)$ for $r = 1, 2, 3$ then

$$n'_r(t) \leq n_r(t) \quad \text{for all} \quad t \geq 0. \tag{5.75}$$

Therefore if $\boldsymbol{b}$ is stable for some ν then it is also stable for any $\nu' < \nu$. The above adaptation of Theorem 5.3.1, together with (5.65), shows that there is some critical parameter λ (depending on the detailed specification of $\boldsymbol{b}^{\{1,2\}}$ and hence $\pi^{\{1,2\}}$) such that $1/3 \leq \lambda \leq 1/2$ and $\boldsymbol{b}$ is stable if $\nu < \lambda c$ and similarly, unstable if $\nu > \lambda c$. For the control strategy described in Section 4.5 we already know that $\lambda = 1/3$; otherwise for the case $\nu = c/3$ we have $k = 0$ in (5.65) and hence stability; simple continuity arguments now give $\lambda > 1/3$ in this case.

Chapter 6

Insensitivity

In this chapter we consider the effect on the stability conditions for some control strategies when we relax the assumption that the call size distributions are exponential. We outline the property of insensitivity of the stationary distribution and show that this is equivalent to a seemingly weaker property which we refer to as scale insensitivity. Then we identify stability conditions for the general two dimensional network $\mathcal{N}_2$ which do not depend on the call size distribution.

6.1 The model

Throughout this chapter we consider a model for a process of call arrivals and departures, in which for each call type $r \in \mathcal{R}$ (where $\mathcal{R}$ is finite) calls arrive as a Poisson process with rate ν_r and have independent identically distributed call size (workload) distributions with mean μ_r^{-1}. (We no longer assume these distributions to be exponential.) For each $r \in \mathcal{R}$ let $\kappa_r := \nu_r/\mu_r$. As usual let $n_r(t)$ be the number of calls of type r in the system at time $t \geq 0$, and let $\boldsymbol{n}(t) = (n_r(t))_{r\in\mathcal{R}}$. For each $r \in \mathcal{R}$, the total workload of calls of type r is reduced at a rate $b_r(\boldsymbol{n}(t))$ where this *bandwidth* is divided equally amongst all calls of type r. Our choice of the function $\boldsymbol{b}$ defined by $\boldsymbol{b} = \boldsymbol{b}(\boldsymbol{n})$ is called the *control strategy* or the *bandwidth allocation* for the model.

6.2 Insensitivity of stationary distribution

This section is concerned with insensitivity of the stationary distribution of the process $\boldsymbol{n}(\cdot)$. We say that the stationary distribution of the process $\boldsymbol{n}(\cdot)$, driven by a control strategy $\boldsymbol{b}$, is *insensitive* if is invariant under variation of the call size distributions, subject to the mean of the latter being held constant. It has been shown by Bonald and Proutière [5] and by Serfozo [31] that, for any given control strategy $\boldsymbol{b}$, the stationary distribution of the process $\boldsymbol{n}(\cdot)$ driven by $\boldsymbol{b}$ is invariant under variation of the call size distribution as above, if and only if $\boldsymbol{b}$ satisfies the *balance property*, i.e. if and only if for all $\boldsymbol{n} \in \mathbb{Z}_+^R$ (where $R = |\mathcal{R}|$) such that for all $r, r' \in \mathcal{R}$ with $n_r > 0$ and $n_{r'} > 0$,

$$b_r(\boldsymbol{n})b_{r'}(\boldsymbol{n} - \boldsymbol{e}_r) = b_{r'}(\boldsymbol{n})b_r(\boldsymbol{n} - \boldsymbol{e}_{r'}). \tag{6.1}$$

where, for each $r \in \mathcal{R}$, $\boldsymbol{e}_r = (e_{rs})_{s\in\mathcal{R}}$ is the R-dimensional unit vector given by $e_{rr} = 1$ and $e_{rs} = 0$ for all $s \neq r$. The following proposition (to some extent implicit in the work of Bonald and Proutière [5]) shows that (6.1) is equivalent to the process $\boldsymbol{n}(\cdot)$ being reversible.

Proposition 6.2.1. *The balance property* (6.1) *is equivalent to the existence of a solution (*$\boldsymbol{\pi} = (\pi(\boldsymbol{n}))_{\boldsymbol{n}\in\mathbb{Z}_+^R}$*) for the detailed balance equations*

$$\pi(\boldsymbol{n})\mu_r b_r(\boldsymbol{n}) = \pi(\boldsymbol{n} - \boldsymbol{e}_r)\nu_r \qquad r \in \mathcal{R}. \tag{6.2}$$

Proof. First suppose that for a particular control strategy $\boldsymbol{b}$ the equations (6.2) have a solution $\boldsymbol{\pi}$. Then for any $\boldsymbol{n} \in \mathbb{Z}_+^R$ such that for all $r, r' \in \mathcal{R}$ with $n_r > 0$ and $n_{r'} > 0$

$$\pi(\boldsymbol{n})b_r(\boldsymbol{n})b_{r'}(\boldsymbol{n} - \boldsymbol{e}_r) = \pi(\boldsymbol{n} - \boldsymbol{e}_r - \boldsymbol{e}_{r'})\kappa_r\kappa_{r'}. \tag{6.3}$$

But similarly,

$$\pi(\boldsymbol{n})b_{r'}(\boldsymbol{n})b_r(\boldsymbol{n} - \boldsymbol{e}_{r'}) = \pi(\boldsymbol{n} - \boldsymbol{e}_r - \boldsymbol{e}_{r'})\kappa_r\kappa_{r'}. \tag{6.4}$$

Hence by combining (6.3) and (6.4) we obtain (6.1).

Conversely suppose that (6.1) holds. Given any value for $\pi(\mathbf{0})$ we choose $\pi(\boldsymbol{n})$ for a given $\boldsymbol{n} \in \mathbb{Z}_+^R$ as follows. Let $\boldsymbol{e}_{r_1} \dots \boldsymbol{e}_{r_k}$ be any sequence of R-dimensional unit vectors such that $\sum_{i=1}^k \boldsymbol{e}_{r_i} = \boldsymbol{n}$; define $\pi(\boldsymbol{n})$ by

$$\pi(\mathbf{0}) \prod_{r \in \mathcal{R}} \kappa_r^{n_r} = \pi(\boldsymbol{n}) b_{r_1}(\boldsymbol{n}) b_{r_2}(\boldsymbol{n} - \boldsymbol{e}_{r_1}) \dots b_{r_k}(\boldsymbol{n} - \boldsymbol{e}_{r_1} - \dots - \boldsymbol{e}_{r_{k-1}}). \qquad (6.5)$$

It is not difficult to see that it follows from (6.1) that $\pi(\boldsymbol{n})$ is well defined. It follows in particular that $\pi(\boldsymbol{n}) b_r(\boldsymbol{n}) = \pi(\boldsymbol{n} - \boldsymbol{e}_r)\kappa_r$. Hence $\boldsymbol{\pi}$ satisfies the detailed balance equations (6.2). □

In particular, we note for future reference that, in the case $R = 1$, the detailed balance equations (6.2) always have a solution. Therefore any stationary distribution which exists is always insensitive. In the case $b_1(n) = c$ for all $n \geq 1$ (for some constant c) the solution of these equations may be normalised to a probability distribution if and only if $\nu_1 < \mu_1 c$, this is given by

$$\pi(n) = \left(\frac{\nu_1}{\mu_1 c}\right)^n \left(1 - \frac{\nu_1}{\mu_1 c}\right), \quad n \geq 0. \qquad (6.6)$$

6.3 Scale insensitivity

We shall say that the stationary distribution (when it exists) of a process $\boldsymbol{n}(\cdot)$ driven by a control strategy $\boldsymbol{b}$ is *scale insensitive* if and only if it is invariant whenever, for any $r \in \mathcal{R}$, and any $\lambda > 0$, ν_r is replaced by $\lambda\nu_r$ and for each $\boldsymbol{n} \in \mathbb{Z}_+^R$, $b_r(\boldsymbol{n})$ is replaced by $\lambda b_r(\boldsymbol{n})$.

We now consider the following three properties for the process $\boldsymbol{n}(\cdot)$.

(a) Insensitivity of the stationary distribution of $\boldsymbol{n}(\cdot)$ under variation of the call size distributions subject to the mean being held constant.

(b) The existence of a solution to the detailed balance equations.

(c) Scale insensitivity of the stationary distribution of $\boldsymbol{n}(\cdot)$, as above.

We have already established, through Proposition 6.2.1, the equivalence of (a) and (b). Further Bonald and Proutière [5] show that (a) implies (c), the argument of which we sketch as follows. Consider the model of the present chapter in which the call size distributions are exponential (so that $\boldsymbol{n}(\cdot)$ is Markov). For $\lambda \in [0, 1]$ suppose, for some $r \in \mathcal{R}$, that the exponential (with mean μ_r^{-1}) distribution is replaced by a mixture distribution in which call sizes have

(i) exponential mean $\lambda^{-1}\mu_r^{-1}$ distribution with probability λ,

(ii) value zero with probability $1 - \lambda$.

(Note that under this new call size distribution call sizes of type $r \in \mathcal{R}$ continue to have mean μ_r^{-1}.) Then the process $\boldsymbol{n}(\cdot)$ remains a Markov process in which arrival and departure rates for calls of type $r \in \mathcal{R}$ are now $\lambda\nu_r$ and $\lambda\mu_r b_r(\boldsymbol{n})$. Hence distributional insensitivity implies scale insensitivity.

It follows from the above results that (b) implies (c), but this is also immediately obvious since the detailed balance equations for the process $\boldsymbol{n}(\cdot)$ remain unchanged under the above scaling by λ. We now show that (c) implies (b). Consider the full balance equations of $\boldsymbol{n}(\cdot)$

$$\sum_{r\in\mathcal{R}} \alpha_r \pi(\boldsymbol{n})(\nu_r + \mu_r b_r(\boldsymbol{n})) = \sum_{r\in\mathcal{R}} \alpha_r(\pi(\boldsymbol{n}+\boldsymbol{e}_r)\mu_r b_r(\boldsymbol{n}) + \pi(\boldsymbol{n}-\boldsymbol{e}_r)\nu_r). \qquad (6.7)$$

where $\alpha_r = 1$ for all $r \in \mathcal{R}$.

Now assume that the stationary distribution $\boldsymbol{\pi}$ of $\boldsymbol{n}(\cdot)$ is scale insensitive. Then we obtain the detailed balance equations by taking, for each $r \in \mathcal{R}$ in turn, $\alpha_r = 1$ and $\alpha_s = 0$ for $s \neq r$. Hence (c) implies (b). Hence the properties (a), (b) and (c) are equivalent. In particular insensitivity is equivalent to scale insensitivity.

6.4 An example with scale sensitivity

In this section we examine a two-dimensional process $\boldsymbol{n}^\lambda(\cdot) = (n_1^\lambda(\cdot), n_2^\lambda(\cdot))$ whose distribution is not scale insensitive as outlined in the last section. We study how the stationary distribution of $\boldsymbol{n}^\lambda(\cdot)$ varies with the scale parameter λ.

For $\lambda > 0$, consider the two-dimensional process $\boldsymbol{n}^\lambda(\cdot)$ in which calls of type 1 have Poisson arrivals at rate $\lambda\nu_1$ and departures at rate $\lambda\mu_1 b_1(\boldsymbol{n})$ and in which calls of type 2 have Poisson arrivals at rate ν_2 and exponential departures at rate $\mu_2 b_2(\boldsymbol{n})$, where the functions b_1 and b_2 on $\mathbb{Z}_+^2$ are given by

$$b_1(\boldsymbol{n}) = \begin{cases} b_1 & \text{if} \quad n_1 > 0 \quad \text{and} \quad n_2 > 0, \\ c_1 & \text{if} \quad n_1 > 0 \quad \text{and} \quad n_2 = 0, \\ 0 & \text{otherwise} \end{cases} \tag{6.8}$$

$$b_2(\boldsymbol{n}) = \begin{cases} b_2 & \text{if} \quad n_1 > 0 \quad \text{and} \quad n_2 > 0, \\ c_2 & \text{if} \quad n_1 = 0 \quad \text{and} \quad n_2 > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{6.9}$$

and where for some b_1, b_2 and $\boldsymbol{c} = (c_1, c_2)$ we take

$$b_r \le c_r \quad \text{for} \quad r = 1, 2 \tag{6.10}$$

and also

$$\kappa_r < c_r \quad \text{for} \quad r = 1, 2. \tag{6.11}$$

To ensure that, for each λ, the stationary distribution of $\boldsymbol{n}^\lambda(\cdot)$ exists we assume

$$\kappa_1 < b_1. \tag{6.12}$$

We note that for $\lambda \in (0, 1]$, as in the previous section, the process $\boldsymbol{n}^\lambda(\cdot)$ may be viewed as corresponding to an instance of the two-dimensional process $\boldsymbol{n}(\cdot) = (n_1(\cdot), n_2(\cdot))$ described in Chapter 3, in which the arrival rates are given by ν_1, ν_2, calls of type 1 have sizes which with probability λ are exponential with mean $(\lambda\mu_1)^{-1}$ and with probability $1 - \lambda$ have zero size, calls of type 2 have sizes which are exponential with mean μ_2^{-1}, and in which the bandwidth allocation is the function $\boldsymbol{b}$ given by $\boldsymbol{b}(\boldsymbol{n}) = (b_1(\boldsymbol{n}), b_2(\boldsymbol{n}))$ for $\boldsymbol{n} \in \mathbb{Z}_+^2$. With this interpretation, varying λ corresponds to one particular variation of the distribution of the size of calls of type 1, subject to the mean of this distribution being held constant.

For all $\lambda > 0$, it follows from the results of Theorem 3.2.1 for maximally spatially homogeneous control strategies, that the stationary distribution $\pi^\lambda(\cdot)$ of the process $\boldsymbol{n}^\lambda(\cdot)$ exists if and only if in addition to (6.12) we have that

$$b_2 > \kappa_2 \frac{(c_1 - b_1)}{(c_1 - \kappa_1)}, \tag{6.13}$$

by (6.10) and (6.11).

We note that this stability condition does not depend on the value of λ. For each $\lambda \in (0, 1]$ we let π^λ denote the stationary distribution of $\boldsymbol{n}^\lambda(\cdot)$ if it exists. For the remainder of this section we assume that condition (6.13) holds. We shall use informal theoretical arguments to suggest the limit, π^0, of the stationary distribution π^λ as $\lambda \to 0$. (This limit may be assumed to exist—see below—but may or may not itself be a distribution.) It is further intended to investigate numerically how π^λ varies for $\lambda \in (0, 1]$.

Consider first the case when $b_2 > \kappa_2$. When λ is very small, transitions in the first coordinate $n_1^\lambda(\cdot)$ of the process $\boldsymbol{n}^\lambda(\cdot)$ occur very infrequently, and hence the process $n_2^\lambda(\cdot)$ effectively has time to come to equilibrium between such transitions. Thus we expect that, as $\lambda \to 0$, the conditional distribution $\pi_{2|1}^\lambda(\cdot|n_1)$ of $n_2^\lambda(\cdot)$ given $n_1^\lambda(\cdot) = n_1$ should converge to the limiting distribution $\pi_{2|1}^0(\cdot|n_1)$ given by the solution of the balance equations

$$\pi_{2|1}^0(n_2|0)\nu_2 = \pi_{2|1}^0(n_2 + 1|0)\mu_2 c_2 \tag{6.14}$$

and

$$\pi_{2|1}^0(n_2|n_1)\nu_2 = \pi_{2|1}^0(n_2 + 1|n_1)\mu_2 b_2, \quad \text{for} \quad n_1 \geq 1, \tag{6.15}$$

where $\sum_{n_2 \geq 0} \pi^0(n_2|n_1) = 1$ for all $n_1 \in \mathbb{Z}_+$. Hence,

$$\pi_{2|1}^0(n_2|n_1) = \begin{cases} \left(1 - \frac{\kappa_2}{c_2}\right)\left(\frac{\kappa_2}{c_2}\right)^{n_2} & \text{if} \quad n_1 = 0 \\ \left(1 - \frac{\kappa_2}{b_2}\right)\left(\frac{\kappa_2}{b_2}\right)^{n_2} & \text{if} \quad n_1 \geq 1. \end{cases} \tag{6.16}$$

Consideration of the flux balance, under equilibrium, between consecutive values of n_1 implies that the marginal distribution $\pi_1^\lambda(\cdot)$ of $n_1^\lambda(\cdot)$ should converge to the distribution $\pi_1^0(\cdot)$ given by the solution of the balance equations

$$\pi_1^0(n_1)\lambda\nu_1 = \pi_1^0(n_1+1)\lambda\mu_1 \left[\pi_{2|1}^0(0|n_1+1)c_1 + \sum_{n_2 \geq 1} \pi_{2|1}^0(n_2|n_1+1)b_1 \right]. \quad (6.17)$$

Hence from (6.16) and (6.17)

$$\pi_1^0(n_1) = \left(1 - \frac{\kappa_1 b_2}{c_1 b_2 + \kappa_2(b_1 - c_1)}\right)\left(\frac{\kappa_1 b_2}{c_1 b_2 + \kappa_2(b_1 - c_1)}\right)^{n_1} \quad n_1 \in \mathbb{Z}_+. \quad (6.18)$$

Hence in this case we may expect the limiting distribution $\pi^0(\cdot)$ of $\boldsymbol{n}^\lambda(\cdot)$ to exist on $\mathbb{Z}_+^2$ and to be given by

$$\pi^0(\boldsymbol{n}) = \pi_1^0(n_1)\pi_{2|1}^0(n_2|n_1) \quad \boldsymbol{n} \in \mathbb{Z}_+^2. \quad (6.19)$$

We may also calculate the marginal distribution $\pi_2^0(\cdot)$ by

$$\pi_2^0(n_2) = \sum_{n_1 > 0} \pi^0(\boldsymbol{n}) = \sum_{n_1 > 0} \pi_1^0(n_1)\pi_{2|1}^0(n_2|n_1) \quad \boldsymbol{n} \in \mathbb{Z}_+^2. \quad (6.20)$$

Now consider the case when $b_2 \leq \kappa_2$. The following claims were made. For λ very small, transitions in the first coordinate $n_1^\lambda(\cdot)$ of the process again occur very infrequently, and we now expect $\pi_{2|1}^\lambda(n_2|n_1)$ to be close to 0 for all $n_1 > 0$ and finite n_2. By (6.12) and by considering the total probability flux in each direction between successive values of n_1, as $\lambda \to 0$, the marginal distribution $\pi_1^\lambda(\cdot)$ of $n_1^\lambda(\cdot)$ should converge to the proper distribution $\pi_1^0(\cdot)$ given by the solution of the balance equations

$$\pi_1^0(n_1)\nu_1 = \pi_1^0(n_1+1)\mu_1 b_1. \quad (6.21)$$

Hence

$$\pi_1^0(n_1) = \left(1 - \frac{\kappa_1}{b_1}\right)\left(\frac{\kappa_1}{b_1}\right)^{n_1}. \quad (6.22)$$

However, as remarked above, as $\lambda \to 0$, for $n_1 > 0$ the conditional distribution of n_2 given n_1 "drifts out to infinity"—i.e. the probability mass concentrated on any finite region of $\mathbb{Z}_+$ tends to zero. Therefore

$$\pi_{2|1}^\lambda(n_2|n_1) \to 0 \quad \text{as} \quad \lambda \to 0 \quad \text{for} \quad n_1 > 0. \quad (6.23)$$

It follows in this case that this limiting distribution is improper. However, the justifications for these claims are not clear. The validity of these claims are left as open questions.

We now investigate numerically how the joint distribution π^λ varies for general $\lambda \in (0, 1]$. It is convenient to define the function ρ on $(0, 1]$ by

$$\rho(\lambda) = \sum_{n_1 \geq 0} \sum_{n_2 \geq 0} |\pi^\lambda_{1,2}(n_1, n_2) - \pi^\lambda_1(n_1)\pi^\lambda_2(n_2)| \quad \text{for} \quad \lambda \in (0, 1]. \tag{6.24}$$

Note that, for all $\lambda \in (0, 1]$,

$$0 \leq \rho(\lambda) \leq 2. \tag{6.25}$$

The function ρ can be used as a measure of dependence between the first and second coordinates of the joint distribution π^λ. In particular, $\rho(\lambda) = 0$ if and only if these two coordinates are independent.

We now introduce an example where we investigate $\rho(\lambda)$ for $\lambda \in (0, 1]$. We take the following values for the parameters of the process $\boldsymbol{n}^\lambda(\cdot)$. Assume that, $c_1 = 1$, $c_2 = 1/2$, $\nu_1 = 1/2$ and $\nu_2 = 1/5$. For simplicity we assume that $\mu_1 = \mu_2 = 1$. We assume $b_1 = 1$ and $b_2 = 0$, so that calls of type 1 are given complete priority over calls of type 2. It follows that conditions (6.10), (6.11) and (6.12) hold. Further it follows that (6.13) is satisfied so that π^λ exists for all $\lambda \in (0, 1]$. Finally, note that here we are in the case in the case $b_2 \leq \kappa_2$.

For each $\lambda \in (0, 1]$ the stationary distribution π^λ of $\boldsymbol{n}^\lambda(\cdot)$ is computed numerically by solving the full balance equations. These were solved numerically for different values of λ using the iterative equation

$$\pi^{\lambda,(i+1)}(\boldsymbol{n}) = \frac{\sum_{\boldsymbol{n}' \neq \boldsymbol{n}} \pi^{\lambda,(i)}(\boldsymbol{n}') q_{\boldsymbol{n}',\boldsymbol{n}}}{\sum_{\boldsymbol{n}'} q_{\boldsymbol{n},\boldsymbol{n}'}} \tag{6.26}$$

where for each $\boldsymbol{n} \in \mathbb{Z}^2_+$ and for $r = 1, 2$

$$q_{\boldsymbol{n},\boldsymbol{n}'} = \begin{cases} \nu_r & \text{if} \quad \boldsymbol{n}' = \boldsymbol{n} + \boldsymbol{e}_r \\ b_r(\boldsymbol{n}) & \text{if} \quad \boldsymbol{n}' = \boldsymbol{n} - \boldsymbol{e}_r \\ 0 & \text{otherwise.} \end{cases} \tag{6.27}$$

where (for $r = 1, 2$) $\boldsymbol{e}_r = (e_{rs})_{s=1,2}$ is the two-dimensional unit vector with $e_{rr} = 1$ and $e_{rs} = 0$ for $s \neq r$.

Figure 6.1 below shows a plot of the estimated value of $\rho(\lambda)$ against λ for $\lambda \in (0, 1]$.

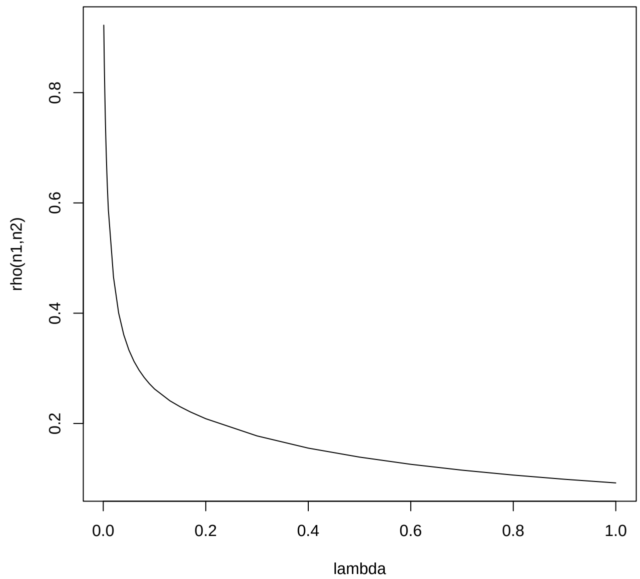

Figure 6.1 Plot of $\rho(\lambda)$ against λ.

From this plot we can make the following suggestions. For $\lambda = 1$ (when the distribution for type 1 call sizes is exponential) there is relatively little dependence between the two coordinates of π^λ. A small λ corresponds to the distribution for type 1 call sizes being far from exponential. Also when λ is small there is a high dependence between the two coordinates of π^λ. In this example , when the detailed balance equations do not hold, the stationary distribution of the process $\boldsymbol{n}(\cdot)$ is remarkably sensitive to variation in λ, and so also to variation in the call size distribution even when the mean of the latter is held fixed.

6.5 Insensitivity of stability

We now study the phenomenon of *stability insensitivity.* In the case of models in which call sizes are not necessarily exponentially distributed, we say that a given control strategy $\boldsymbol{b}$ is *stable* if and only if, starting from the state in which the network is empty, the expectation of the time until the first return to this state is finite. (In the exponential case this agrees with our earlier definition of stability as positive recurrence of the associated Markov process.) By *stability insensitivity* we mean that, for a given control strategy $\boldsymbol{b}$, the stability (or otherwise) of the process describing the state of the network is unchanged under variation of the call size distributions, again subject only to the means of the latter being held constant. We note that when the control strategy $\boldsymbol{b}$ is such that we have insensitivity in the sense discussed earlier, that is, that the stationary distribution of the process which records the number of calls of each type in the system is unchanged under variation of the call size distributions as above, then the existence of this stationary distribution means that we necessarily have stability insensitivity in the sense defined here. Thus the latter phenomenon may be regarded as a weaker property of a control strategy.

In order to investigate stability insensitivity, we refine the definition of the *state* of the network process at any time $t \geq 0$ by requiring it to record not only the number of calls of each type currently in progress, but also the used (or spent) call size (or workload) at time t associated with each of these calls; we denote this enhanced state of the network by $\boldsymbol{x}(t) = (x_r(t))_{r \in \mathcal{R}}$ where $x_r(t)$ records the number of calls of type r in the system at time t with their spent workloads. The process $\boldsymbol{x}(\cdot)$ is then again Markov, taking values in the enhanced state space $\mathcal{X}$. For any state $\boldsymbol{x} \in \mathcal{X}$, we denote by $\boldsymbol{n}^{\boldsymbol{x}}$ the corresponding reduced state recording only the number of calls of each type in progress. We also let $(\mathcal{F}_t)_{t \geq 0}$ denote any filtration with respect to which the process $\boldsymbol{x}(\cdot)$ is adapted.

For any finite set $A \subset \mathbb{Z}_+^R$ define the stopping time τ_A by

$$\tau_A = \inf_{t \geq 0} \{t \colon \boldsymbol{n}^{\boldsymbol{x}(t^-)} \notin A,\, \boldsymbol{n}^{\boldsymbol{x}(t)} \in A\} \tag{6.28}$$

(where $\boldsymbol{n}^{\boldsymbol{x}(t^-)} := \lim_{t' \uparrow t} \boldsymbol{n}^{\boldsymbol{x}(t)}$), i.e. the first time that $\boldsymbol{n}^{\boldsymbol{x}(\cdot)}$ hits or returns to the set A. It is convenient to write $\tau_{\mathbf{0}}$ for $\tau_{\{\mathbf{0}\}}$, the first time $\boldsymbol{n}^{\boldsymbol{x}(\cdot)}$ hits or returns to $\mathbf{0}$. We say a given control strategy $\boldsymbol{b}$ associated with a process $\boldsymbol{x}(\cdot)$ is *weakly stable* if there exists a finite set $A \subset \mathbb{Z}_+^R$ such that for all $\boldsymbol{x}$ with $\boldsymbol{n}^{\boldsymbol{x}} \notin A$

$$\mathbf{E}_{\boldsymbol{x}} \tau_A < \infty. \tag{6.29}$$

We say that a given control strategy $\boldsymbol{b}$ associated with a process $\boldsymbol{x}(\cdot)$ is *stable* if

$$\mathbf{E}_{\mathbf{0}}\tau_{\mathbf{0}} < \infty. \tag{6.30}$$

We note that in the Markov case (exponentially distributed workloads) stability and weak stability are equivalent, and coincide with positive recurrence of the process $\boldsymbol{n}^{\boldsymbol{x}(\cdot)}$.

To find stable control strategies for more general call size distributions we use Theorem 6.5.1, which is just the specialisation to the present problem in continuous time of Theorem 2.1.2 of Fayolle *et al* [12].

Theorem 6.5.1. *For a given control strategy* $\boldsymbol{b}$ *suppose there exists* $\epsilon > 0$, *a finite set* $A \subset \mathbb{Z}_+^R$, *an increasing sequence of stopping times* $0 = T_0 \leq T_1 \leq \dots$ *such that*

$$T_m \uparrow T_A \quad \textit{almost surely as} \quad m \to \infty, \tag{6.31}$$

where

$$T_A = \begin{cases} \min_{m \in \mathbb{Z}_+}(T_m : \boldsymbol{n}^{\boldsymbol{x}(T_m)} \in A), & \\ \infty, & \textit{if no such} \quad m \quad \textit{exists}, \end{cases} \tag{6.32}$$

and, finally, a function $f \colon \mathcal{X} \to \mathbb{R}_+$ *such that, for all initial states* $\boldsymbol{x}(0)$ *of the process* $\boldsymbol{x}(\cdot)$, *and for all* $m \geq 0$,

$$\mathbf{E}(f(\boldsymbol{x}(T_{m+1})) - f(\boldsymbol{x}(T_m))|\mathcal{F}_{T_m}) \leq -\epsilon\mathbf{E}(T_{m+1} - T_m|\mathcal{F}_{T_m}) \quad \textit{almost surely.} \tag{6.33}$$

Then $\boldsymbol{b}$ *is weakly stable.*

Proof. (Adapted from Fayolle *et al* [12], Theorem 2.1.2.) Let $\epsilon > 0$, $A \subset \mathbb{Z}_+^R$ and $f \colon \mathcal{X} \to \mathbb{R}_+$ satisfy the conditions of the present theorem. It follows from (6.33) that for $\boldsymbol{x}(0) \notin A$ and for all $m \geq 0$

$$\begin{aligned} \mathbf{E}(f(\boldsymbol{x}(T_{m+1})) - f(\boldsymbol{x}(T_m))) =& \mathbf{E}(\mathbf{E}(f(\boldsymbol{x}(T_{m+1})) - f(\boldsymbol{x}(T_m))|\mathcal{F}_{T_m})) \\ \leq& -\epsilon\mathbf{E}(T_{m+1} - T_m). \end{aligned} \tag{6.34}$$

Hence,

$$\mathbf{E}(f(\boldsymbol{x}(T_m))) \leq -\epsilon \mathbf{E}T_m + f(\boldsymbol{x}(0)) \tag{6.35}$$

since $\mathbf{E}T_m = \sum_{i=1}^{m}(\mathbf{E}T_{i+1} - \mathbf{E}T_i)$. It follows from (6.34) that for all $m \geq 1$

$$\mathbf{E}T_m \leq \frac{f(\boldsymbol{x}(0))}{\epsilon} < \infty \tag{6.36}$$

since $f(\boldsymbol{x}(T_m)) \geq 0$ almost surely. Then under the assumption (6.31) and using the monotone convergence theorem

$$\mathbf{E}T_A = \mathbf{E}(\lim_{m\to\infty} T_m) = \lim_{m\to\infty} \mathbf{E}T_m \leq \frac{f(\boldsymbol{x}(0))}{\epsilon}. \tag{6.37}$$

It follows from (6.28) and (6.32) that $\tau_A \leq T_A$ almost surely. Hence

$$\mathbf{E}\tau_A \leq \mathbf{E}T_A \leq \frac{f(\boldsymbol{x}(0))}{\epsilon} < \infty. \tag{6.38}$$

It follows that $\boldsymbol{b}$ is weakly stable. □

Corollary 6.5.2 below shows that, under the conditions of Theorem 6.5.1 and mild additional conditions, the control strategy $\boldsymbol{b}$ is in fact stable.

For any $r \in \mathcal{R}$, let F_r be the distribution function for the size of calls of type r; for any $y \geq 0$, define also $\overline{F}_r(y) = 1 - F_r(y)$. We shall say that this call size distribution is *light-tailed* if and only if the function e_r, defined by

$$e_r(y) = \frac{1}{\bar{F}(y)} \int_y^\infty \bar{F}(t)dt, \tag{6.39}$$

is bounded above in $y \geq 0$. The quantity $e_r(y)$ is the expected residual call size, conditional on the call size being at least y. The requirement that e_r is non-increasing is reasonably coincident with the the usual definition of light-tailed distributions.

Corollary 6.5.2. *Suppose that for some control strategy* $\boldsymbol{b}$*, the conditions of Theorem 6.5.1 are satisfied, and that additionally*

(a) all call size distributions are light-tailed,

(b) the function f, satisfying the conditions of Theorem 6.5.1, is bounded over all $\boldsymbol{x}$ such that $\boldsymbol{n}^{\boldsymbol{x}} = \boldsymbol{n}$, for each $\boldsymbol{n} \in \mathbb{Z}_+^R$

(c) there exists $b > 0$ such that

$$\sum_{r \in \mathcal{R}} b_r(\boldsymbol{n}) \geq b \quad \textit{for all} \quad \boldsymbol{n} \neq \mathbf{0}. \tag{6.40}$$

Then $\boldsymbol{b}$ is stable.

We note that condition (c) will certainly be satisfied for all reasonable control strategies.

Proof. Assume without loss of generality, that $\mathbf{0} \in A$ for the finite set $A \subset \mathbb{Z}_+^R$ defined in Theorem 6.5.1. Assume also without loss of generality that A is such that for $\boldsymbol{n} \in \mathbb{Z}_+^R$ such that $n_r > 0$ for all $r \in \mathcal{R}$ we have that

$$\boldsymbol{n} \in A \quad \text{implies} \quad \boldsymbol{n} - \boldsymbol{e}_r \in A \quad \text{for all} \quad r \in \mathcal{R} \tag{6.41}$$

where $\boldsymbol{e}_r = (e_{rs})_{s \in \mathcal{R}}$ is the R-dimensional unit vector such that $e_{rr} = 1$ and $e_{rs} = 0$ for all $s \neq r$.

We show first that there exists a constant $t_1 < \infty$ such that, for all $\boldsymbol{x}$ with $\boldsymbol{n}^{\boldsymbol{x}} \in A$,

$$\mathbf{E}_{\boldsymbol{x}} \tau_{A^C} < t_1, \tag{6.42}$$

where A^C is the compliment of the set A and where τ_{A^C} is defined by (6.28). To see this note that, given any $t_0 > 0$, there exists $p > 0$ such that, for all $\boldsymbol{x}$ with $\boldsymbol{n}^{\boldsymbol{x}} \in A$,

$$\mathbf{P}_{\boldsymbol{x}}(\boldsymbol{n}^{\boldsymbol{x}(t_0)} \notin A) \geq p. \tag{6.43}$$

Now observe the system at the successive times $t = t_0, 2t_0, \ldots$ and observe that the expected number of such observations needed to find $\boldsymbol{n}^{\boldsymbol{x}^{(t)}} \notin A$ is no greater than the mean of a geometric distribution with parameter p from which (6.42) follows easily.

Let

$$\partial A = \{\boldsymbol{n} \notin A \colon \boldsymbol{n} - \boldsymbol{e}_r \in A \text{ for some } r \in \mathcal{R}\}. \tag{6.44}$$

Since ∂A is finite it follows from condition (b) and (6.38) that there exists a constant $t_2 < \infty$ such that, for all $\boldsymbol{x}$ with $\boldsymbol{n}^{\boldsymbol{x}} \in \partial A$

$$\mathbf{E}_{\boldsymbol{x}} \tau_A < t_2. \tag{6.45}$$

It follows from the assumption (6.41) that there exists $q > 0$ such that, for all $\boldsymbol{x}$ with $\boldsymbol{n}^{\boldsymbol{x}} \in A$,

$$\mathbf{P}_{\boldsymbol{x}}(\boldsymbol{n}^{\boldsymbol{x}(t)} = 0 \text{ for some } t \leq \tau_{A^C}) \geq q. \tag{6.46}$$

This follows since, for a sufficiently large time t', the light-tailed assumption together with Markov's inequality (see for example Mitzenmacher and Upfal [25]) gives a non-zero lower bound (over all $\boldsymbol{x}$ as above) on the probability to clear all existing work from the system, and then there is a nonzero probability for the independent event that no further work comes in within time t'.

Finally it follows from (6.42), (6.45) and (6.46) that

$$\begin{aligned} \mathbf{E}_{\mathbf{0}}(\tau_{\mathbf{0}}) &\leq t_1 + (1-q)(t_2+t_1) + (1-q)^2(t_2+t_1) + \dots \\ &= t_1 + \frac{1-q}{q}(t_2+t_1) \\ &< \infty. \end{aligned} \tag{6.47}$$

Hence the control strategy $\boldsymbol{b}$ is stable as required. □

6.6 Stability conditions for $\mathcal{N}_2$

We continue to consider the model of this chapter with general call size distributions (with finite means). We consider in detail the case $\mathcal{R} = \{1, 2\}$. Our aim is to give, in Theorem 6.6.2, natural conditions in providing for the (weak) stability of a control strategy $\boldsymbol{b}$. Most of the work is done in Lemma 6.6.1 below.

For $r = 1, 2$ and for any real-valued function $\bar{b}_r$ on $\mathbb{Z}_+$ satisfying $\bar{b}_r(0) = 0$, consider the one-dimensional *Markov* process $\bar{n}_r(\cdot)$ with transition rates

$$\bar{n}_r \to \begin{cases} \bar{n}_r + 1 & \text{at rate} \quad \kappa_r \\ \bar{n}_r - 1 & \text{at rate} \quad \bar{b}_r(\bar{n}_r). \end{cases} \tag{6.48}$$

For $r = 1, 2$, define $\bar{\pi}_{r,\bar{b}_r}$ to be the stationary probability that this process takes the value 0 and be equal to zero when no such stationary distribution exists.

Lemma 6.6.1. *Suppose that, for some parameters $\delta, \beta_1, \beta_2 > 0$ and for all $\boldsymbol{n} \in \mathbb{Z}_+^R$ not belonging to some finite set A, a control strategy $\boldsymbol{b}$ is such that*

$$b_1(\boldsymbol{n}) + b_2(\boldsymbol{n}) \geq \kappa_1 + \kappa_2 + \delta \quad \textit{when} \quad n_1 \wedge n_2 > 0, \tag{6.49}$$

$$b_1(\boldsymbol{n}) \geq \beta_1 \quad \textit{when} \quad n_1 > 0 \quad \textit{and} \quad n_2 = 0, \tag{6.50}$$

and

$$b_2(\boldsymbol{n}) \geq \beta_2 \quad \textit{when} \quad n_1 = 0 \quad \textit{and} \quad n_2 > 0. \tag{6.51}$$

Suppose further that, for some functions $\bar{b}_1$, $\bar{b}_2$ on $\mathbb{Z}_+$ with $\bar{b}_1(0) = 0$, $\bar{b}_2(0) = 0$, and for some $a' > 0$

$$b_1(\boldsymbol{n}) \leq \bar{b}_1(n_1) \quad \textit{for all} \quad \boldsymbol{n} \in \mathbb{Z}_+^2 \quad \textit{such that} \quad n_2 \geq a', \tag{6.52}$$

and

$$b_2(\boldsymbol{n}) \leq \bar{b}_2(n_2) \quad \textit{for all} \quad \boldsymbol{n} \in \mathbb{Z}_+^2 \quad \textit{such that} \quad n_1 \geq a'. \tag{6.53}$$

Then a sufficient condition for the control strategy $\boldsymbol{b}$ associated with the process $\boldsymbol{x}(\cdot)$ to be weakly stable is given by

$$\bar{\pi}_{1,\bar{b}_1}(\beta_2 - (\kappa_1 + \kappa_2)) + (1 - \bar{\pi}_{1,\bar{b}_1})\delta > 0 \tag{6.54}$$

and

$$\bar{\pi}_{2,\bar{b}_2}(\beta_1 - (\kappa_1 + \kappa_2)) + (1 - \bar{\pi}_{2,\bar{b}_2})\delta > 0. \tag{6.55}$$

Proof. We use Theorem 6.5.1 to show that any control strategy $\boldsymbol{b}$, that satisfies the conditions of the present lemma, is weakly stable. For any $\boldsymbol{n} \in \mathbb{Z}_+^2$, let $n_{\min} = n_1 \wedge n_2$ and let $n_{\max} = n_1 \vee n_2$. Let $k \in \mathbb{Z}_+$ (to be chosen later) be such that $a' \le k < \infty$. Consider the finite set $A \subset \mathbb{Z}_+^2$ with $A = \{\boldsymbol{n} \in \mathbb{Z}_+^2 : n_{\max} \le k\}$. Given $t_0 > 0$ (again to be chosen later) consider the sequence of stopping times $0 = T_0 \le T_1 \le \ldots$ for the process $\boldsymbol{x}(\cdot)$ defined as follows (for $m \ge 1$): if $\boldsymbol{n}^{\boldsymbol{x}(T_{m-1})} \notin A$

$$T_m = \inf_t \{t > T_{m-1} \colon \textit{either} \quad n_{\min}^{\boldsymbol{x}(t^-)} > 0,\, n_{\min}^{\boldsymbol{x}(t)} = 0 \quad \textit{or} \quad t = T_{m-1} + t_0\} \tag{6.56}$$

i.e. T_m is the first time, after T_{m-1}, that the process $\boldsymbol{n}_{\min}^{\boldsymbol{x}(\cdot)}$ either returns to or hits the boundary set $\{\boldsymbol{n} \in \mathbb{Z}_+^2 : n_{\min} = 0\}$ or else endures for a further time t_0, whichever of these occurs first; if $\boldsymbol{n}^{\boldsymbol{x}(T_{m-1})} \in A$, we take $T_m = T_{m-1}$ (and hence also $T_{m'} = T_{m-1}$ for all $m' \ge m$). It follows that this sequence of stopping times satisfies (6.31) for A as defined above.

Consider the function f on $\mathcal{X}$ where $f(\boldsymbol{x})$ is defined to be the total expected residual size of calls of both type 1 and 2 when the state of the process is $\boldsymbol{x}$. Let $\boldsymbol{b}$ be a control strategy that satisfies the conditions of the present lemma. In using Theorem 6.5.1 it is clearly sufficient to verify condition (6.32) in the case $m = 1$ and for all initial states $\boldsymbol{x}(0)$ such that $\boldsymbol{n}^{\boldsymbol{x}(0)} \notin A$. Hence we must show that, for $\boldsymbol{x}(0)$ as above and for an appropriate choice of $\epsilon > 0$,

$$\mathbf{E}(f(\boldsymbol{x}(T_1)) - f(\boldsymbol{x}(0))) \le -\epsilon \mathbf{E} T_1. \tag{6.57}$$

In the case where $\boldsymbol{x}(0)$ is such that $n_{\min}^{\boldsymbol{x}(0)} > 0$, it follows from the definition of T_1 that (6.57) can clearly satisfied with $\epsilon = \delta$ (by (6.53) and (6.54)).

Hence, without loss of generality, it will be sufficient to show (6.57) for some $\epsilon > 0$ and for all initial states $\boldsymbol{x}(0)$ such that $n_1^{\boldsymbol{x}(0)} = 0$ and $n_2^{\boldsymbol{x}(0)} > k$. In this case it follows from the definition of T_1 and conditions (6.49), (6.51) and (6.54) that

$$\mathbf{E}(f(\boldsymbol{x}(T_1)) - f(\boldsymbol{x}(0))) \le -\delta \mathbf{E} T_1 + \nu_1^{-1}(\kappa_1 + \kappa_2 - \beta_2 + \delta). \tag{6.58}$$

Define

$$T' = \inf_t \{t > 0 \colon n_2^{\boldsymbol{x}(t)} = a'\}. \tag{6.59}$$

i.e. T' is the first time that the one-dimensional process $n_2^{\boldsymbol{x}(\cdot)}$ hits a' (recall that $n_2^{\boldsymbol{x}(0)} > k \geq a'$). Now consider a one-dimensional process $\bar{x}_1(\cdot)$, which behaves similarly to the first coordinate $x_1(\cdot)$ of $\boldsymbol{x}(\cdot)$ except only that the control strategy $b_1(\boldsymbol{n}^{\boldsymbol{x}(\cdot)})$ is replaced by $\bar{b}_1(n_1^{\bar{x}_1(\cdot)})$. Note that, by the insensitivity property for one-dimensional processes discussed earlier the stationary probability that $n_1^{\bar{x}_1(\cdot)}$ is equal to zero is given by $\bar{\pi}_{1,\bar{b}_1}$. We also start $\bar{x}_1(\cdot)$ such that $n_1^{\bar{x}_1(0)} = 0$. Define

$$\bar{T} = \inf_t \{t > 0 \colon n_1^{\bar{x}_1(t^-)} > 0,\, n_1^{\bar{x}_1(t)} = 0\} \tag{6.60}$$

i.e. $\bar{T}$ is the first time of the return of the process $n_1^{\bar{x}_1(\cdot)}$ to zero. By (6.52) we can couple the processes $\bar{x}_1(\cdot)$ and $\boldsymbol{x}(\cdot)$ in such a way that

$$n_1^{\bar{x}_1(t)} \leq n_1^{\boldsymbol{x}(t)} \quad \text{for} \quad t \in [0, T']. \tag{6.61}$$

It follows from the above coupling and condition (6.52) that

$$\mathbf{E}T_1 \geq \mathbf{E}(\bar{T} \wedge T' \wedge t_0). \tag{6.62}$$

In the case where $\bar{\pi}_{1,\bar{b}_1} = 0$, we have $\mathbf{E}\bar{T} = \infty$, and so, provided t_0 and k are chosen sufficiently large (recall that $n_2^{\boldsymbol{x}(0)} > k$), it follows from (6.58) and (6.62) that condition (6.57) may be satisfied with $\epsilon = \delta/2$.

In the case where $\bar{\pi}_{1,\bar{b}_1} > 0$, we have that $\mathbf{E}\bar{T} = (\nu_1 \bar{\pi}_{1,\bar{b}_1})^{-1}$, and so, for any $\lambda \in (0,1)$, it again follows from (6.62) that t_0 and k may be chosen sufficiently large such that $\mathbf{E}T_1 > \lambda(\nu_1 \bar{\pi}_{1,\bar{b}_1})^{-1}$, and hence, by (6.58), that

$$\begin{aligned}\mathbf{E}(f(\boldsymbol{x}(T_1)) - f(\boldsymbol{x}(0))) &\leq -\delta\lambda(\nu_1 \bar{\pi}_{1,\bar{b}_1})^{-1} + \nu_1^{-1}(\kappa_1 + \kappa_2 - \beta_2 + \delta) \\ &= -(\nu_1 \bar{\pi}_{1,\bar{b}_1})^{-1}(\bar{\pi}_{1,\bar{b}_1}(\beta_2 - (\kappa_1 + \kappa_2)) + \delta(\lambda - \bar{\pi}_{1,\bar{b}_1})).\end{aligned} \tag{6.63}$$

Thus it follows from (6.54), by choosing λ sufficiently close to 1, that (6.57) may once more be satisfied for some $\epsilon > 0$ such that $\epsilon < \lambda^{-1}(\bar{\pi}_{1,\bar{b}_1}(\beta_2 - (\kappa_1 + \kappa_2)) + \delta(\lambda - \bar{\pi}_{1,\bar{b}_1}))$ as required. □

Now consider again the general two-dimensional network $\mathcal{N}_2$ studied in Chapter 3. Let $\boldsymbol{\kappa} = (\kappa_1, \kappa_2)$ be the input parameter. Again let

$$(A_{jr})_{j\in\mathcal{J},r\in\mathcal{R}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{6.64}$$

define the network topology. We allow the possibility of non-exponential distributions for the size of calls. We provide, in Theorem 6.6.2, conditions for the weak stability of a control strategy $\boldsymbol{b}$. Further, in the case when (a) and (b) of Corollary 6.5.2 are satisfied, these conditions ensure the stability of $\boldsymbol{b}$. We define *capacity constraints* for $\boldsymbol{b}$ such that for all $\boldsymbol{n} \in \mathbb{Z}_+^2$

$$b_1(\boldsymbol{n}) + b_2(\boldsymbol{n}) \leq c_0 \quad \text{and} \quad b_r \leq c_r \quad \text{for} \quad r = 1, 2, \tag{6.65}$$

where $0 < c_j < \infty$ for $j = 0, 1, 2$ and $c_1 \vee c_2 \leq c_0 \leq c_1 + c_2$. We assume also the condition

$$\kappa_1 + \kappa_2 < c_0 \quad \text{and} \quad \kappa_r < c_r \quad \text{for} \quad r = 1, 2. \tag{6.66}$$

We show that any Pareto efficient control strategy $\boldsymbol{b}$, satisfying a natural condition is weakly stable.

Theorem 6.6.2. *Suppose $\mathcal{N}_2$ is a two-dimensional network. Then, given $\epsilon \in (0, \min(c_1 - \kappa_2, c_2 - \kappa_2)]$, there exists $a_{1,\epsilon}, a_{2,\epsilon}$ such that if $\boldsymbol{b}$ is a Pareto efficient control strategy for $\mathcal{N}_2$ which satisfies*

$$\liminf_{n_1\to\infty} b_1(\boldsymbol{n}) \geq \kappa_1 + \epsilon \quad \textit{for all} \quad n_2 < a_{2,\epsilon} \tag{6.67}$$

and

$$\liminf_{n_2\to\infty} b_2(\boldsymbol{n}) \geq \kappa_2 + \epsilon \quad \textit{for all} \quad n_1 < a_{1,\epsilon}, \tag{6.68}$$

then $\boldsymbol{b}$ is weakly stable.

Remark Note that, from Section 3.5, α-fair-sharing control strategies for $\mathcal{N}_2$ satisfy conditions (6.67) and (6.68) and hence are weakly stable.

Proof of Theorem 6.6.2. Given $\epsilon > 0$, let $\boldsymbol{b}$ be a Pareto efficient control strategy that satisfies conditions (6.67) and (6.68). We apply Lemma 6.6.1 to show that, for appropriately chosen, $a_{1,\epsilon}$ and $a_{2,\epsilon}$, independent of $\boldsymbol{b}$, the control strategy $\boldsymbol{b}$ is weakly stable. It follows from the Pareto efficiency of $\boldsymbol{b}$ that conditions (6.49), (6.50) and (6.51) are satisfied with $\delta = c_0 - (\kappa_1 + \kappa_2)$ and $\beta_r = c_r$ for $r = 1, 2$.

It follows from (6.67) that there exists $a_1^{'}$ such that

$$b_1(\boldsymbol{n}) \geq \kappa_1 + \epsilon \quad \text{for all} \quad \boldsymbol{n} \in \mathbb{Z}_+^2 \quad \text{such that} \quad n_2 < a_{2,\epsilon},\, n_1 \geq a_1^{'}. \tag{6.69}$$

Similarly, it follows from (6.68) that there exists $a_2^{'}$ such that

$$b_2(\boldsymbol{n}) \geq \kappa_2 + \epsilon \quad \text{for all} \quad \boldsymbol{n} \in \mathbb{Z}_+^2 \quad \text{such that} \quad n_1 < a_{1,\epsilon},\, n_2 \geq a_2^{'}. \tag{6.70}$$

We take the the functions $\bar{b}_1$ and $\bar{b}_2$ of Lemma 6.6.1 to be given by for $r = 1, 2$

$$\bar{b}_r(n_r) = \begin{cases} 0 & \text{whenever} \quad n_r = 0 \\ c_0 - (\kappa_{r'} + \epsilon) & \text{whenever} \quad 0 < n_r < a_{r,\epsilon} \\ c_r & \text{otherwise} \end{cases} \tag{6.71}$$

where $r' \neq r$. It then follows from (6.69) and (6.70) that for any Pareto efficient control strategy $\boldsymbol{b}$ on $\mathcal{N}_2$ satisfying (6.67) and (6.68) we have that (6.52) and (6.53) are satisfied with $a' = a_1^{'} \vee a_2^{'}$.

Consider first $r = 1$. Let $\bar{n}_1(\cdot)$ be the one-dimensional Markov process on $\mathbb{Z}_+$ with transition rates defined by (6.48) where $\bar{b}_1$ is defined by (6.71). Let $\bar{\pi}_1(\cdot)$ denote the stationary distribution of $\bar{n}_1(\cdot)$ if it exists. Note that for a finite $a_{1,\epsilon}$ the function f defined by $f(n) = n$ for all $n \in \mathbb{Z}_+$ acts as a Lyapunov function for $\bar{n}_1(\cdot)$ (with $F = \{n \in \mathbb{Z}_+ : n < a_{1,\epsilon}\}$ as the refuge). Hence for a finite $a_{1,\epsilon}$, the process $\bar{n}_1(\cdot)$ is positive recurrent and so $\bar{\pi}_{1,\bar{b}_1} = \bar{\pi}_1(0)$. By considering the detailed balance equations of $\bar{n}_1(\cdot)$ we obtain

$$\bar{\pi}_1(n) = \begin{cases} \bar{\pi}_{1,\bar{b}_1} \left(\frac{\kappa_1}{c_0 - (\kappa_2 + \epsilon)} \right)^n & \text{whenever} \quad n < a_{1,\epsilon} \\ \bar{\pi}_{1,\bar{b}_1} \left(\frac{\kappa_1}{c_0 - (\kappa_2 + \epsilon)} \right)^{a_{1,\epsilon}} \left(\frac{\kappa_1}{c_1} \right)^{n - a_{1,\epsilon}} & \text{whenever} \quad n \geq a_{1,\epsilon}. \end{cases} \tag{6.72}$$

Consider the case when $\epsilon \geq c_0 - (\kappa_1 + \kappa_2)$. It clear that $\bar{\pi}_{1,\bar{b}_1} \to 0$ as $a_{1,\epsilon} \to \infty$. Hence given $\epsilon_1' > 0$ there exists a sufficiently large $a_{1,\epsilon}$ such that

$$\bar{\pi}_{1,\bar{b}_1} < \epsilon_1'. \tag{6.73}$$

Hence

$$\begin{aligned}\bar{\pi}_{1,\bar{b}_1}(\beta_2 - (\kappa_1 + \kappa_2)) + (1 - \bar{\pi}_{1,\bar{b}_1})\delta &= \bar{\pi}_{1,\bar{b}_1}(c_2 - c_0) + c_0 - (\kappa_1 + \kappa_2) \\ &\geq \epsilon_1'(c_2 - c_0) + c_0 - (\kappa_1 + \kappa_2) > 0 \end{aligned} \tag{6.74}$$

for a sufficiently small ϵ_1' (i.e. for a sufficiently large $a_{1,\epsilon}$) and so condition (6.54) is satisfied.

On the other hand, consider the case when $\epsilon < c_0 - (\kappa_1 + \kappa_2)$. Then

$$\bar{\pi}_{1,\bar{b}_1} \sum_{n=0}^{a_{1,\epsilon}-1} \left(\frac{\kappa_1}{c_0 - (\kappa_2 + \epsilon)} \right)^n < 1 \tag{6.75}$$

and so

$$\bar{\pi}_{1,\bar{b}_1} < \frac{1 - \frac{\kappa_1}{c_0 - (\kappa_2 + \epsilon)}}{1 - \left(\frac{\kappa_1}{c_0 - (\kappa_2 + \epsilon)} \right)^{a_{1,\epsilon}}} \tag{6.76}$$

Hence $\bar{\pi}_{1,\bar{b}_1} \to 1 - \kappa_1/(c_0 - (\kappa_1 + \kappa_2))$ as $a_{1,\epsilon} \to \infty$. Hence given $\hat{\epsilon}_1 > 0$ there exists, a sufficiently large $a_{1,\epsilon}$ such that

$$\bar{\pi}_{1,\bar{b}_1} < \hat{\pi}_1 + \hat{\epsilon}_1 = 1 - \frac{\kappa_1}{c_0 - (\kappa_2 + \epsilon)} + \hat{\epsilon}_1 < 1 - \frac{\kappa_1}{c_0 - \kappa_2} + \hat{\epsilon}_1. \tag{6.77}$$

Hence

$$\begin{aligned}\bar{\pi}_{1,\bar{b}_1}(\beta_2 - (\kappa_1 + \kappa_2)) + (1 - \bar{\pi}_{1,\bar{b}_1})\delta &= \bar{\pi}_{1,\bar{b}_1}(c_2 - c_0) + c_0 - (\kappa_1 + \kappa_2) \\ &\geq \left(1 - \frac{\kappa_1}{c_0 - \kappa_2}\right)(c_2 - c_0) + c_0 - (\kappa_1 + \kappa_2) + \hat{\epsilon}_1(c_2 - c_0) \\ &= \frac{(c_2 - \kappa_2)(c_0 - (\kappa_1 + \kappa_2))}{c_0 - \kappa_2} + \hat{\epsilon}_1(c_2 - c_0) \\ &> 0 \end{aligned} \tag{6.78}$$

for a sufficiently small $\hat{\epsilon}_1$ (i.e. for a sufficiently large $a_{1,\epsilon}$) and so condition (6.54) is satisfied for all $\epsilon \in (0, \min(c_1 - \kappa_2, c_2 - \kappa_2)]$. Similarly (6.55) is satisfied for a sufficiently large $a_{2,\epsilon}$. Hence $\boldsymbol{b}$ is weakly stable. □

Bibliography

[1] Asmussen, S. (2003). *Applied Probability and Queues*, 2nd edition. Spinger-Verlag, New York.

[2] Bertsekas, D. and Gallager, R. (1987). *Data Networks*. Prentice Hall.

[3] Bonald, T. and Massoulié, L. (2001). Impact of Fairness on Internet Performance. In *Proceedings of ACM Sigmetrics 2001*. Cambridge, Massachusetts.

[4] Bonald, T., Massoulié, L., Proutière, A. and Virtamo, J. (2006). A queueing analysis of max-min fairness, proportional fairness and balanced fairness. *Queueing Systems*, **53**, 65–84.

[5] Bonald, T. and Proutière, A. (2003). Insensitive Bandwidth Sharing in Data Networks. *Queueing Systems*, **44**, 69–100.

[6] Bonald, T. and Proutière, A. (2004). On Stochastic Bounds for Monotonic Processor Sharing Networks. *Queueing Systems*, **47**, 81–106.

[7] Borovkov, A. A. (1998). *Ergodicity and Stability of Stochastic Processes.* Wiley, Chichester.

[8] Brémaud, P. (1981). *Point Processes and Queues.* Springer-Verlag.

[9] Brémaud, P. (1999). *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and queues.* Springer-Verlag New York, Inc.,

[10] Dai, J. (1995). On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *Annals of Applied Probability*, **5**, 49–77.

[11] de Veciana, G., Lee, T.-J. and Konstantopoulos, T. (2001). Stability and performance analysis of networks supporting elastic services. *IEEE/ACM Trans. on Networking*, **9**, 2–14.

[12] Fayolle, G., Malyshev, V. A. and Menshikov, M. V. (1995). *Topics in the Constructive Theory of Countable Markov Chains*. Cambridge University Press.

[13] Freedman, D. (1983). *Markov Chains*. Springer-Verlag.

[14] Gafni, E. M. and Bertsekas D. (1984). Dynamic control of session input rates in communication networks. *IEEE Trans. on Automatic Control*, **29**, 1009–1016.

[15] Hansen, J. C., Reynolds, C. J. and Zachary, S. (2007). Stability of processor sharing networks with simultaneous resource requirements. To appear in *Journal of Applied Probability*, **44**.

[16] Hansen, J. C., Reynolds, C. J. and Zachary, S. (2007). Insensitivity of stability in processor sharing networks with simultaneous resource requirements. (Provisional title.) (In preparation.)

[17] Jones, P. W. and Smith, P. (2001). *Stochastic Processes*. Arnold.

[18] Kelly, F. P. (1979). *Reversibility and Stochastic Networks*. Wiley.

[19] Kelly, F. P. (1991). Loss networks. *Annals of Applied Probability*, **1**, 319-378.

[20] Kelly, F. P., Mauloo, A. and Tan, D. (1998). Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, **49**, 237–252.

[21] Kelly, F. P. and Williams, R. J. (2004). Fluid model for bandwidth sharing. *Annals of Applied Probability*, **14**, 1055–1083.

[22] Lindvall, T. (1992). *Lectures on the coupling method*. Dover.

[23] Norris, J. R. (1997). *Markov Chains*. Cambridge University Press.

[24] MacPhee, I. M. and Müller, L. J. (2006). Stability criteria for controlled queueing systems. *Queueing Systems*, **52**, 215–229.

[25] Mitzenmacher, M. and Upfal, E. (2005). *Probability and Computing*. Cambridge University Press.

[26] Mo, J. and Walrand, J. (2000). Fair end-to-end window-based congestion control. *IEEE/ACM Trans. on Networking*, **8**, 556–567.

[27] Quadling, D. A. (1959). *Mathematical Analysis*. Oxford University Press.

[28] Robert, P. (2003). *Stochastic Networks and Queues*. Springer-Verlag Berlin Heidelberg.

[29] Roberts, J. W. and Massoulié, L. (2000). Bandwidth sharing and admission control for elastic traffic. *Telecommunications Systems*, **15**, 185–201.

[30] Ross, K. W. (1995). *Multiservice loss models for broadband telecommunication networks*. Springer, New York.

[31] Serfozo, R. (1999). *Introduction to Stochastic networks*. Springer.

[32] Stolyar, A. L. (2004). Maxweight scheduling in a generalized switch: state space collapse and workload minimization in heavy traffic. *Annals of Applied Probability*, **14**, 1–53.

[33] Whittle, P. (1971). *Optimization under Constraints: Theory and Applications of Nonlinear Programming*. John Wiley & Sons Ltd.

[34] Williams, D. (1991). *Probability with Martingales*. Cambridge University Press.

[35] Zachary, S. (1995). On Two-Dimensional Markov Chains in the Positive Quadrant with Partial Spatial Homogeneity. *Markov Processes and Related Fields*, **1**, 267–280.

[36] Zachary, S. and Ziedins, I. (1999). Loss networks and Markov random fields. *Journal of Applied Probability*, **36**, 403–414.