



HAL
open science

Infinite horizon optimal impulsive control with applications to Internet congestion control

Konstantin Avrachenkov, Oussama Habachi, Alexey Piunovskiy, Yi Zhang

► **To cite this version:**

Konstantin Avrachenkov, Oussama Habachi, Alexey Piunovskiy, Yi Zhang. Infinite horizon optimal impulsive control with applications to Internet congestion control. *International Journal of Control*, 2015, 88 (4), pp.703-716. 10.1080/00207179.2014.971436 . hal-01259259

HAL Id: hal-01259259

<https://inria.hal.science/hal-01259259>

Submitted on 20 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Infinite Horizon Optimal Impulsive Control with Applications to Internet Congestion Control

K. Avrachenkov, O. Habachi, A. Piunovskiy and Y. Zhang

Abstract

We investigate infinite horizon deterministic optimal control problems with both gradual and impulsive controls, where any finitely many impulses are allowed simultaneously. Both discounted and long run time average criteria are considered. We establish very general and at the same time natural conditions, under which the dynamic programming approach results in an optimal feedback policy. The established theoretical results are applied to the Internet congestion control, and by solving analytically and nontrivially the underlying optimal control problems, we obtain a simple threshold-based active queue management scheme, which takes into account the main parameters of the transmission control protocols, and improves the fairness among the connections in a given network.

AMS classification 49N25

1 Introduction

The impulsive optimal control theory has many real-life applications. For example many control problems in queueing theory, population dynamics, mathematical epidemiology, financial mathematics etc can be formulated as the impulsively controlled systems: see [10, 11, 14, 24, 25, 27] and the references therein. Roughly speaking, an impulse (or intervention) means the instant change of the state of the system. This results in discontinuous trajectories, leading to technical difficulties when solving optimal control problems. Nevertheless, it is possible to adjust the dynamic programming method for such models: see [3, 4, 5, 6, 22, 28]. The Pontryagin maximum principle (or the closely related Lagrangian approach) can also be used for solving impulsive optimal control problems: see [7, 11, 19, 26, 27].

In the current paper, we consider deterministic models similar to those investigated in [3, 11, 19, 22, 26, 27]. Note that in [11, 19, 22, 26] only the finite-horizon case was studied. The long-run average (per cycle) reward was considered in [27]. In the latter article, only a specific optimal control problem for a fish population was solved using the maximum principle. We underline that no gradual (or continuous, ordinary) control and no running cost/reward were considered in [11, 27]. In [3], the deterministic discounted model was studied. To guarantee no accumulation of impulses, usually a separated from zero cost is paid for any impulsive action; see [3, 4]. Note that our verification theorems remain valid when there is no impulse cost, which is the case in the considered applications. Therefore, the distinguishing features of the current work are like follows.

- We develop the dynamic programming approach to the infinite-horizon models both with the total discounted reward and the long-run average reward, which include the impulse-generated rewards along with the running reward. The impulse-generated reward/cost may be zero.
- Both the gradual and impulsive controls are considered.
- We allow any finite number of simultaneous impulses which was not allowed in [11, 22, 26, 27].
- We rigorously and nontrivially solve in closed-form two new problems of the Internet congestion control, which are of their own importance.

Let us elaborate a bit more on importance and interest of the application of our established theoretical results to the Internet congestion control. Recently, there has been a steady increase in the demand for QoS (Quality of Services) and fairness among the increasing number of IP (Internet Protocol) flows. Although the Transmission Control Protocol (TCP) gives efficient solutions to end-to-end error control and congestion control, the problem of fairness among flows is far from being solved: see, for example, [1, 21, 16] for the discussions of the unfairness among various TCP versions. The fairness can be improved by the Active Queue Management (AQM) through the participation of the links or routers in the congestion control. We measure the network fairness by the long-run average α -fairness and the discounted α -fairness, which can be specified to the total throughput, the proportional fairness and the max-min fairness maximization with the particular values of the tuning parameter α : see [20]. The network model together with its analysis in the present article is different from the existing literature on the network utility maximization, see e.g., [15, 12, 17], in at least the following three important aspects.

- We take into account the fine, saw-tooth like, dynamics of congestion control algorithms.
- We use per-flow control, which nowadays becomes feasible, see [23], and describe its form.
- By solving analytically the impulsive control problems, we propose a novel AQM scheme that takes into account not only the traffic transiting through bottleneck links but also end-to-end congestion control algorithms implemented at the edges of the network. More specifically, our scheme asserts that a congestion notification (packet drop or explicit congestion notification) should be sent out whenever the current sending rate is over a threshold, whose closed-form expression is obtained.

The rest of the article is organized as follows. In Section 2 we provide the verification theorems for the optimality of a control policy for the discounted and the long-run average problems, respectively. In Section 3 we apply the established theoretical results to solving rigorously two corresponding specific optimal impulsive control problems for the Internet congestion control. Section 4 concludes this article. The proofs of the statements presented in Section 3 are postponed to the appendix.

2 Dynamic programming for general optimal impulsive control problems

In this section, we establish the verification theorems for a general infinite horizon impulsive control problem under the long-run average criterion and the discounted criterion, which are then used to solve the concerned Internet congestion control problems in the next section.

2.1 Description of the controlled process

Let us consider the following dynamical system in $X \subseteq \mathbb{R}^n$ (with X being a nonempty measurable subset of \mathbb{R}^n , and some initial condition $x(0) = x_0 \in X$) governed by

$$dx = f(x, u)dt, \tag{1}$$

where $u \in U$ is the gradual control with U being an arbitrary nonempty Borel space. Suppose another nonempty Borel space V is given, and, at any time moment T , if he decides so, the decision maker can apply an impulsive control $v \in V$ leading to the following new state:

$$x(T) = j(x(T^-), v), \tag{2}$$

where j is a measurable mapping from $X \times V$ to X .

Below we let $c(x, u)$ be the reward rate if the controlled process is at the state x and the gradual control u is applied, and $C(x, v)$ be the reward earned from applying the impulsive control v , both being measurable real-valued functions.

Definition 1 A policy π is defined by a U -valued measurable mapping $u(t)$ and a sequence of impulses $\{T_i, v_i\}_{i=1}^{\infty}$ with $v_i \in V$ and $\dots \geq T_{i+1} \geq T_i \geq 0$, which satisfies $T_0 := 0$ and $\lim_{i \rightarrow \infty} T_i = \infty$. A policy π is called a feedback one if one can write $u(t) = u^f(x(t))$, $T_i^{\mathcal{L}} = \inf\{t > T_{i-1} : x(t) \in \mathcal{L}\}$, $v_i = v^{f, \mathcal{L}}(x(T_i^-))$, where u^f is a U -valued measurable mapping on X , and $\mathcal{L} \subset X$ is a specified (measurable) subset of X . A feedback policy is completely characterized and thus denoted by the triplet $(u^f, \mathcal{L}, v^{f, \mathcal{L}})$.

We underline that, since it is required in the above definition that $\lim_{i \rightarrow \infty} T_i = \infty$, under each policy there are no more than finitely many impulsive controls within each finite interval.

We are interested in the *admissible* policies π under which the following hold (with any initial state):

(a). $T_0 \leq T_1 < T_2 < \dots$. This requirement is not restrictive because, in case $n < \infty$ impulsive controls $v_{i+1}, v_{i+2}, \dots, v_{i+n}$ are applied simultaneously, i.e., $T_i < T_{i+1} = T_{i+2} = \dots = T_{i+n} < T_{i+n+1}$, we merge these impulsive controls as a single one \hat{v} by defining

$$j(x, \hat{v}) := j(j(\dots j(x, v_{i+1}), \dots, v_{i+n-1}), v_{i+n}) \quad (3)$$

and

$$C(x, \hat{v}) := C(x, v_{i+1}) + C(j(x, v_{i+1}), v_{i+2}) + \dots + C(j(\dots j(x, v_{i+1}), \dots, v_{i+n-1}), v_{i+n}). \quad (4)$$

Note that different orders of $v_{i+1}, v_{i+2}, \dots, v_{i+n}$ give rise to different \hat{v} , and since only finitely many impulses are admitted at any single time moment, the expressions on the right hand sides of (3) and (4) are always well defined.

(b). The controlled process $x(t)$ described by (1) and (2) is well defined: for any initial state $x(0) = x_0$, there is a unique piecewise differentiable function $x^\pi(t)$ with $x^\pi(0) = x_0$, satisfying (1) for all t , wherever the derivative exists; satisfying (2) for all $T = T_i$, $i = 1, 2, \dots$; and satisfying that $x^\pi(t)$ is continuous at each $t \neq T_i$.

The controlled process under such a policy π is denoted by $x^\pi(t)$.

Remark 1 We emphasize that not every arbitrary triplet $(u^f, \mathcal{L}, v^{f, \mathcal{L}})$ defines an admissible feedback policy: we must be sure that $\lim_{i \rightarrow \infty} T_i = \infty$. Otherwise, according to Definition 1, the objects u^f , $\{T_i^{\mathcal{L}}, v_i\}_{i=1}^{\infty}$ do not define a policy at all. The requirement $\lim_{i \rightarrow \infty} T_i = \infty$ appears in all cited works. Very often a positive penalty, bigger than $\varepsilon > 0$, for any one impulse is introduced: see Chapter 6, §1.1 in [4]; see also [3, 14]. In our notations, that means $C(x, v) < -\varepsilon$. As a result, any policy with a finite objective will be admissible. In the next section, we do not require the impulse reward C to be negative, and our verification theorems are of conditional nature: if one succeeds to find an admissible policy satisfying the corresponding equations and requirements, then that policy is optimal. Below we consider only admissible policies, and the word ‘admissible’ is omitted for brevity.

2.2 Verification theorems

Under policy π and initial state x_0 , the average reward is defined by

$$J(x_0, \pi) = \liminf_{T \rightarrow \infty} \frac{1}{T} \left\{ \int_0^T c(x^\pi(t), u(t)) dt + \sum_{i=1}^{N(T)} C(x^\pi(T_i^-), v_i) \right\}, \quad (5)$$

where and below $N(T) := \sup\{n > 0, T_n \leq T\}$, and $x(T_0^-) := x_0$; and the discounted reward (with the discount factor $\rho > 0$) is given by

$$J_\rho(x_0, \pi) = \liminf_{T \rightarrow \infty} J_\rho^T(x_0, \pi), \quad (6)$$

where

$$J_\rho^T(x_0, \pi) = \int_0^T e^{-\rho t} c(x^\pi(t), u(t)) dt + \sum_{i=1, 2, \dots, T_i \in [0, T]} e^{-\rho T_i} C(x^\pi(T_i^-), v_i).$$

We only consider the class of (admissible) policies π such that the right side of (5) (resp., (6)) is well defined under the average (resp., discounted) criterion, i.e., all the limits and integrals are finite. The optimal control problem under the average criterion reads

$$J(x_0, \pi) \rightarrow \max_{\pi}, \quad (7)$$

and the one under the discounted criterion reads

$$J_{\rho}(x_0, \pi) \rightarrow \max_{\pi}. \quad (8)$$

A policy π^* is called (average) optimal (resp., (discounted) optimal) if $J(x_0, \pi^*) = \sup_{\pi} J(x_0, \pi)$ (resp., $J_{\rho}(x_0, \pi^*) = \sup_{\pi} J_{\rho}(x_0, \pi)$) for each $x_0 \in X$. Below we consider both problems (5) and (8), and provide the corresponding verification theorems for an optimal feedback policy, see Theorems 1 and 2.

For the average problem (7), we consider the following condition.

Condition 1 *There are a continuous function $h(x)$ on X and a constant $g \in \mathbb{R}$ such that the following hold.*

(i) *The gradient $\frac{\partial h}{\partial x}$ exists everywhere apart from a subset $\mathcal{D} \subset X$, whereas under every policy π and for each initial state x_0 , $h(x^{\pi}(t))$ is absolutely continuous on $[T_i, T_{i+1})$, $i = 0, 1, \dots$; and $\{t \in [0, \infty) : x^{\pi}(t) \in \mathcal{D}\}$ is a null set with respect to the Lebesgue measure.*

(ii) *For all $x \in X \setminus \mathcal{D}$,*

$$\max \left\{ \sup_{u \in U} \left[c(x, u) - g + \left\langle \frac{\partial h}{\partial x}, f(x, u) \right\rangle \right], \sup_{v \in V} [C(x, v) + h(j(x, v)) - h(x)] \right\} = 0, \quad (9)$$

and for all $x \in \mathcal{D}$, $\sup_{v \in V} [C(x, v) + h(j(x, v)) - h(x)] \leq 0$.

(iii) *There are a measurable subset $\mathcal{L}^* \subset X$ and a feedback policy $\pi^* = (u^{f^*}, \mathcal{L}^*, v^{f, \mathcal{L}^*})$ such that for all $x \in X \setminus (\mathcal{D} \cup \mathcal{L}^*)$, $c(x, u^{f^*}(x)) - g + \left\langle \frac{\partial h}{\partial x}, f(x, u^{f^*}(x)) \right\rangle = 0$ and for all $x \in \mathcal{L}^*$, $C(x, v^{f, \mathcal{L}^*}(x)) + h(j(x, v^{f, \mathcal{L}^*}(x))) - h(x) = 0$, and $j(x, v^{f, \mathcal{L}^*}(x)) \notin \mathcal{L}^*$.*

(iv) *For any policy π and each initial state $x_0 \in X$, $\limsup_{T \rightarrow \infty} \frac{h(x^{\pi}(T))}{T} \geq 0$, whereas $\limsup_{T \rightarrow \infty} \frac{h(x^{\pi^*}(T))}{T} = 0$.*

Equation (9) is the Bellman equation for problem (7). The triplet (g, π^*, h) from Condition 1 is often called canonical, and the policy π^* is called a canonical policy. The next result asserts that any canonical policy is optimal for problem (7).

Theorem 1 *For the average problem (7), the feedback policy π^* in Condition 1 is optimal, and g in Condition 1 is the value function, i.e., $g = \sup_{\pi} J(x_0, \pi)$ for each $x_0 \in X$.*

Proof: For each arbitrarily fixed $T > 0$, initial state $x_0 \in X$ and policy π , it holds that

$$\begin{aligned} h(x^{\pi}(T)) &= h(x_0) + \int_0^T \left\{ \left\langle \frac{\partial h}{\partial x}(x^{\pi}(t)), f(x^{\pi}(t), u(t)) \right\rangle \right\} dt \\ &\quad + \sum_{i: T_i \in [0, T]} \{h(j(x^{\pi}(T_i^-), v_i)) - h(x^{\pi}(T_i^-))\}. \end{aligned} \quad (10)$$

Therefore,

$$\begin{aligned} &\int_0^T c(x^{\pi}(t), u(t)) dt + \sum_{i=1}^{N(T)} C(x^{\pi}(T_i^-), v_i) + h(x^{\pi}(T)) \\ &= h(x_0) + \int_0^T \left\{ c(x^{\pi}(t), u(t)) + \left\langle \frac{\partial h}{\partial x}(x^{\pi}(t)), f(x^{\pi}(t), u(t)) \right\rangle \right\} dt \\ &\quad + \sum_{i: T_i \in [0, T]} \{h(j(x^{\pi}(T_i^-), v_i)) - h(x^{\pi}(T_i^-)) + C(x^{\pi}(T_i^-), v_i)\} \leq h(x_0) + \int_0^T g dt, \end{aligned}$$

where the last inequality is because of (9) and the definition of g and h as in Condition 1. It follows that $\frac{1}{T} \left\{ \int_0^T c(x^\pi(t), u(t)) dt + \sum_{i=1}^{N(T)} C(x^\pi(T_i^-), v_i) \right\} + \frac{h(x^\pi(T))}{T} \leq \frac{h(x_0)}{T} + g$, and consequently, $J(x_0, \pi) + \limsup_{T \rightarrow \infty} \frac{h(x^\pi(T))}{T} \leq g$. Since $\limsup_{T \rightarrow \infty} \frac{h(x^\pi(T))}{T} \geq 0$ for each π , we obtain $J(x_0, \pi) \leq g$ for each policy π . For the feedback policy π^* from Condition 1, since $\limsup_{T \rightarrow \infty} \frac{h(x^{\pi^*}(T))}{T} = 0$, and we have $J(x_0, \pi^*) = g$. The statement is proved. \square The next remark is used in the proof of Theorem 3 below.

Remark 2 *It follows from the proof of the previous statement that if one can find a function h , a constant g and a feedback policy π^* such that Condition 1(i, ii, iii) is satisfied, and $\limsup_{T \rightarrow \infty} \frac{h(x^{\pi^*}(T))}{T} = 0$, then the policy π^* is optimal (with the value g) out of the class of policies π that satisfy $\limsup_{T \rightarrow \infty} \frac{h(x^\pi(T))}{T} \geq 0$.*

For the discounted problem (8), we formulate the following condition similar to Condition 1.

Condition 2 *There is a continuous function $W(x)$ on X such that the following hold.*

(i) *The gradient $\frac{\partial W}{\partial x}$ exists everywhere apart from a subset $\mathcal{D} \subset X \subset \mathbb{R}^n$; for any policy π and for any initial state x_0 , the function $W(x^\pi(t))$ is absolutely continuous on all intervals $[T_{i-1}, T_i]$, $i = 1, 2, \dots$; and the Lebesgue measure of the set $\{t \in [0, \infty) : x^\pi(t) \in \mathcal{D}\}$ equals zero.*

(ii) *The following Bellman equation*

$$\max \left\{ \sup_{u \in U} \left[c(x, u) - \rho W(x) + \left\langle \frac{\partial W}{\partial x}, f(x, u) \right\rangle \right], \sup_{v \in V} [C(x, v) + W(j(x, v)) - W(x)] \right\} = 0 \quad (11)$$

is satisfied for all $x \in X \setminus \mathcal{D}$ and $\sup_{v \in V} [C(x, v) + W(j(x, v)) - W(x)] \leq 0$ for all $x \in \mathcal{D}$.

(iii) *There are a measurable subset $\mathcal{L}^* \subset X$ and a feedback policy $\pi^* = (u^{f^*}, \mathcal{L}^*, v^{f, \mathcal{L}^*})$ such that $c(x, u^{f^*}(x)) - \rho W(x) + \left\langle \frac{\partial W}{\partial x}, f(x, u^{f^*}(x)) \right\rangle = 0$ for all $x \in X \setminus (\mathcal{D} \cup \mathcal{L}^*)$ and $C(x, v^{f, \mathcal{L}^*}(x)) + W(j(x, v^{f, \mathcal{L}^*}(x))) - W(x) = 0$ for all $x \in \mathcal{L}^*$; moreover, $j(x, v^{f, \mathcal{L}^*}(x)) \in X \setminus \mathcal{L}^*$.*

(iv) *For any initial state $x_0 \in X$, $\limsup_{T \rightarrow \infty} e^{-\rho T} W(x^\pi(T)) \geq 0$ for any policy π , whereas $\limsup_{T \rightarrow \infty} e^{-\rho T} W(x^{\pi^*}(T)) = 0$.*

Theorem 2 *For the discounted problem (8), the feedback policy π^* from Condition 2 is optimal, and $\sup_{\pi} J_{\rho}(x_0, \pi) = W(x_0) = J_{\rho}(x_0, \pi^*)$ for each $x_0 \in X$.*

Proof: The proof proceeds along the same line of reasoning as in that of Theorem 1; instead of (10), one should now make use of the representation

$$\begin{aligned} 0 &= W(x_0) + \int_0^T e^{-\rho t} \left\{ \left\langle \frac{\partial W(x^\pi(t))}{\partial x}, f(x^\pi(t), u(t)) \right\rangle - \rho W(x^\pi(t)) \right\} dt \\ &+ \sum_{i=1, 2, \dots, T_i \in [0, T]} e^{-\rho T_i} \{ W(g(x^\pi(T_i^-), v_i) - W(x^\pi(T_i^-))) \} - e^{-\rho T} W(x^\pi(T)). \end{aligned}$$

\square According to Theorem 2, if some function W satisfying Condition 2 is obtained, it must be unique.

Similarly to [4], our results are conditional: if one succeeds to obtain appropriate functions h or W satisfying Conditions 1 or 2, then the corresponding policy π^* is optimal for problem (7) or (8). We did not intend to investigate the existence of the solutions to the Bellman equations (9) and (11). That problem is rather delicate, and various sufficient conditions for similar problems can be found in e.g., [3, 4, 19]. In particular, to guarantee no accumulation of impulses, the authors usually require the (negative) impulse reward to be separated from zero. We emphasize that in Section 3, the verification theorems presented above are used to build the optimal policy for the problems with a zero impulse reward.

3 Applications to the Internet congestion control

In this section, we firstly informally describe the impulsive control problem for the Internet congestion control, which will then be later formalized in the framework of the previous section. In what follows,

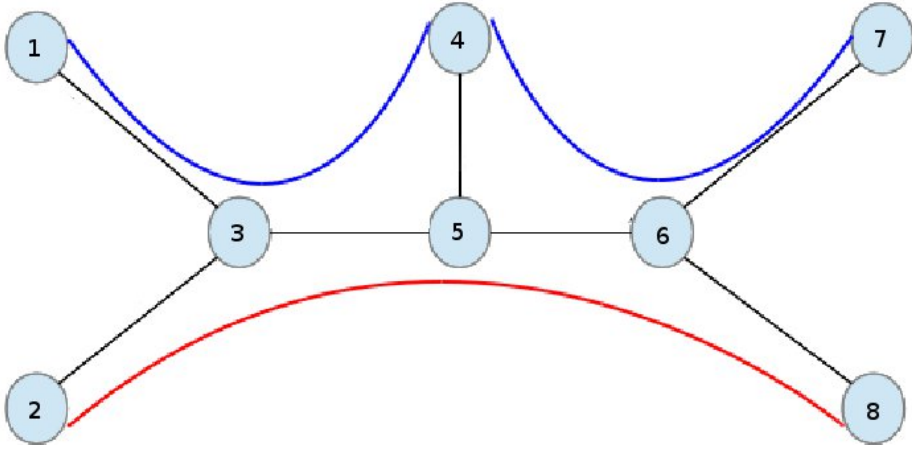


Figure 1: An example of a network with three connections and seven links.

when the context makes it clear, we do not explicitly indicate the underlying policy π , and often write x instead of x^π for brevity.

Let us consider n TCP connections operating in an Internet Protocol (IP) network of L links defined by a routing matrix A , whose element a_{lk} is equal to one if connection k goes through link l , or zero otherwise. Without loss of generality, we assume that each link is occupied by some connection, and each connection is routed through some link. Denote by $x_k(t)$ the sending rate of connection k at time t . We also denote by $P(k)$ the set of links constituting to the path of connection k . An example of such a network is sketched in Figure 1, where we specify a link with the two nodes it connects. For example, (1,3) denotes the link between nodes 1 and 3. Let us label all the links in the following way.

label	1	2	3	4	5	6	7
link	(1,3)	(3,5)	(4,5)	(5,6)	(6,7)	(2,3)	(6,8)

the connections 1, 2, 3 are routed over the paths $P(1) = \{1,2,3\}$, $P(2) = \{3,4,5\}$ and $P(3) = \{2,4,6,7\}$, respectively, so that the routing matrix is given by

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

In this section, the column vector notation $x(t) := (x_1(t), \dots, x_n(t))^T$ is in use, and the terms of connection and source are used interchangeably.

We point out that our network model is quite general at least in the following sense. The data sources are allowed to use different TCP versions, or if they use the same TCP, the TCP parameters (round-trip time, the increase-decrease factors) can be different. More precisely, we suppose that the sending rate of connection k evolves according to the following equation

$$\frac{d}{dt}x_k(t) = a_k x_k^{\gamma_k}(t), \quad a_k \in (0, \infty), \quad \gamma_k \in [0, 1], \quad (12)$$

in the absence of congestion notification, and the TCP reduces the sending rate abruptly if a congestion notification is sent to the source k , i.e., when a congestion notification is sent to the source k at time moment $T_{i,k}$ with $T_{0,k} := 0$ and $T_{i+1,k} \geq T_{i,k}$, its sending rate is reduced as follows

$$x_k(T_{i,k}) = b_k x_k(T_{i,k}^-) < x_k(T_{i,k}^-), \quad b_k \in (0, 1). \quad (13)$$

Here and below, a_k , b_k and γ_k are constants, which would cover at least two important versions of the TCP end-to-end congestion control; if $\gamma_k = 0$ we retrieve the AIMD congestion control mechanism (see [2]), and if $\gamma_k = 1$ we retrieve the Multiplicative Increase Multiplicative Decrease (MIMD) congestion control mechanism (see [13, 29]). Also note that (12) and (13) correspond to a hybrid model description that represents well the saw-tooth behavior of many TCP variants, see [9, 2, 29].

When $T_{i+2,k} > T_{i+1,k} = T_{i,k} > T_{i-1,k}$, multiple (indeed, two in this case) congestion notifications are being sent out simultaneously at $T_{i+1,k} = T_{i,k}$. We only allow finitely many congestion notifications to be sent out at any time moment. For now we write $T_i := (T_{i,1}, \dots, T_{i,n})$ for the i th time moments of the impulsive control for each of the n connections, and assume that the decision of reducing the sending rate of connection k is independent upon the other connections. Since there is no gradual control, we tentatively call the sequence of T_1, T_2, \dots a policy for the congestion control problem, which will be formalized below.

We will consider two performance measures of the system; namely the time average α -fairness function

$$\liminf_{T \rightarrow \infty} \frac{1}{1-\alpha} \sum_{k=1}^n \frac{1}{T} \int_0^T x_k^{1-\alpha}(t) dt,$$

and the discounted α -fairness function

$$\liminf_{T \rightarrow \infty} \frac{1}{1-\alpha} \sum_{k=1}^n \int_0^T e^{-\rho t} x_k^{1-\alpha}(t) dt,$$

to be maximized over the consecutive moments of sending congestion notifications $T_i, i = 1, 2, \dots$. In the meanwhile, due to the limited capacities of the links, the expression $\liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T Ax(t) dt$ (resp., $\liminf_{T \rightarrow \infty} \int_0^T e^{-\rho t} Ax(t) dt$) under the average (resp., discounted) criterion should not be too big. Therefore, after introducing the weight coefficients $\lambda_1, \dots, \lambda_L \geq 0$, we consider the following objective functions to be maximized:

$$Q = \sum_{k=1}^n \left\{ \liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T \frac{x_k^{1-\alpha}(t)}{1-\alpha} dt \right\} - \sum_{l=1}^L \lambda_l \sum_{k:l \in P(k)} \liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T x_k(t) dt \quad (14)$$

in the average case, and

$$Q_\rho = \sum_{k=1}^n \left\{ \liminf_{T \rightarrow \infty} \int_0^T e^{-\rho t} \frac{x_k^{1-\alpha}(t)}{1-\alpha} dt \right\} - \sum_{l=1}^L \lambda_l \sum_{k:l \in P(k)} \liminf_{T \rightarrow \infty} \int_0^T e^{-\rho t} x_k(t) dt \quad (15)$$

in the discounted case, where we recall that $P(k)$ indicates the set of links corresponding to connection k .

Below by using the verification theorems established earlier, we obtain the optimal policies for the problems

$$Q \rightarrow \max_{T_1, T_2, \dots} \quad (16)$$

and

$$Q_\rho \rightarrow \max_{T_1, T_2, \dots}, \quad (17)$$

respectively.

3.1 Solving the average optimal impulsive control problem

We consider in this subsection the average problem (16). Concentrated on policies satisfying

$$\liminf_{T \rightarrow \infty} \frac{1}{1-\alpha} \sum_{k=1}^n \frac{1}{T} \int_0^T x_k^{1-\alpha}(t) dt = \lim_{T \rightarrow \infty} \frac{1}{1-\alpha} \sum_{k=1}^n \frac{1}{T} \int_0^T x_k^{1-\alpha}(t) dt < \infty$$

and $\liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T x_k(t) dt < \infty$ for each $k = 1, \dots, n$, for problem (16) it is sufficient to consider the case of $n = 1$. Indeed, one can legitimately rewrite the function (14) as

$$Q = \sum_{k=1}^n \liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(\frac{x_k^{1-\alpha}(t)}{1-\alpha} - \lambda^k x_k(t) \right) dt,$$

where $\lambda^k := \sum_{l \in P(k)} \lambda_l$, which allows us to decouple different sources. Thus, we will focus on the case of $n = 1$, and solve the following optimal control problem

$$\tilde{J}(x_0) = \liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(\frac{x^{1-\alpha}(t)}{1-\alpha} - \lambda x(t) \right) dt \rightarrow \max_{T_1, T_2, \dots}, \quad (18)$$

where $x(t)$ is subject to (12), (13) and the impulsive controls T_1, T_2, \dots with the initial condition $x(0) = x_0$. Here and below the index $k = 1$ is omitted for convenience.

In the remaining part of this subsection, using the verification theorem (Theorem 1), we rigorously obtain the optimal policy and value to problem (18) in closed-forms.

Remark 3 *When more than one but finitely many congestion notifications are sent out simultaneously, in line with the treatment in the previous section, we will understand the resulting multiple reductions on the sending rate as a consequence of a single “big” impulsive control.*

Let us start with formulating the congestion control problem (18) in the framework given in the previous section, which also applies to the next subsection. Indeed, one can take the following system parameters; $X = (0, \infty)$, $j(x, v) = b^v x$, $C(x, v) = 0$ with $v \in V = \{1, 2, \dots\}$, $f(x, u) = ax^\gamma$, and $c(x, u) = \frac{x^{1-\alpha}}{1-\alpha} - \lambda x$ with $u \in U$, which is a singleton, i.e., there is no gradual control, so that in what follows, we omit $u \in U$ everywhere.

For the future reference and to improve the readability, we write down the Bellman equation (9) for problem (18) as follows;

$$\max \left\{ \left(\frac{x^{1-\alpha}}{1-\alpha} - \lambda x \right) - g + \frac{\partial h}{\partial x}(x) ax^\gamma, \sup_{m=1,2,\dots} \{h(b^m x) - h(x)\} \right\} = 0. \quad (19)$$

As will be seen in the next theorem, for each fixed x , the supremum inside the parenthesis of (19) is attained at a finite value of m .

Theorem 3 *Suppose $\lambda > 0$, $\gamma \in [0, 1]$, $\alpha > 0$, $\alpha \neq 1$, $2 - \alpha - \gamma \neq 0$, $a \in (0, \infty)$, and $b \in (0, 1)$. For the average congestion control problem (18), an optimal policy is given by $\pi^* = (\mathcal{L}^*, v^{\mathcal{L}^*})$ with $\mathcal{L}^* = [\bar{x}, \infty)$, and $v^{\mathcal{L}^*}(x) = k$ if $x \in [\frac{\bar{x}}{b^{k-1}}, \frac{\bar{x}}{b^k}) \subseteq \mathcal{L}^*$ for $k = 1, 2, \dots$, where*

$$\bar{x} = \left\{ \frac{(2-\gamma)(1-b^{2-\alpha-\gamma})}{(2-\alpha-\gamma)(1-b^{2-\gamma})\lambda} \right\}^{\frac{1}{\alpha}} > 0. \quad (20)$$

When $\gamma < 1$, the value function is given by

$$J(x_0, \pi^*) = g := \bar{x} \lambda \left(\frac{\alpha}{1-\alpha} \right) \frac{(1-\gamma)(1-b^{2-\gamma})}{(2-\gamma)(1-b^{1-\gamma})}; \quad (21)$$

and when $\gamma = 1$,

$$J(x_0, \pi^*) = g := \bar{x} \lambda \left(\frac{\alpha}{1-\alpha} \right) \frac{b-1}{\ln(b)}. \quad (22)$$

(Clearly, the constructed policy π^* is admissible.)

The proof of this theorem can be found in the appendix.

3.2 Solving the discounted optimal impulsive control problem

The discounted problem turns out more difficult to deal with, for which we consider that the sending rate increases additively, i.e., $\frac{dx_k(t)}{dt} = a_k > 0$, and decreases multiplicatively, i.e., $j(x_k, v) = b_k x_k$ with $b_k \in (0, 1)$ when a congestion notification is sent, see (12) and (13). Thus, the prevailing version of TCP New Reno in today's Internet is covered as a special case. Furthermore, we assume $\alpha \in (1, 2)$.

Similarly to the average case, we concentrate on policies under which $\liminf_{T \rightarrow \infty} \int_0^T e^{-\rho t} x_k(t) dt = \lim_{T \rightarrow \infty} \int_0^T e^{-\rho t} x_k(t) dt < \infty$, and upon rewriting the objective function in problem (17) as $Q_\rho = \sum_{k=1}^n \liminf_{T \rightarrow \infty} \int_0^T e^{-\rho t} \left(\frac{x_k^{1-\alpha}(t)}{1-\alpha} - \lambda^k x_k(t) \right) dt$, where $\lambda^k = \sum_{l \in P(k)} \lambda_l$, it becomes clear that there is no loss of generality to focus on the case of $n = 1$;

$$\tilde{J}_\rho(x_0) = \liminf_{T \rightarrow \infty} \int_0^T e^{-\rho t} \left(\frac{x^{1-\alpha}(t)}{1-\alpha} - \lambda x(t) \right) dt \rightarrow \max_{T_1, T_2, \dots}, \quad (23)$$

As in the average case, one can put this impulsive control problem in the framework of the previous section; see Remark 3 and the paragraph above it, so that Theorem 2 is applicable. Now the Bellman equation (11) has the form

$$\max \left\{ \frac{x^{1-\alpha}}{1-\alpha} - \lambda x - \rho W(x) + a \frac{dW}{dx}, \sup_{i \geq 1} [W(b^i x) - W(x)] \right\} = 0. \quad (24)$$

The linear differential equation

$$\frac{x^{1-\alpha}}{1-\alpha} - \lambda x - \rho \tilde{W}(x) + a \frac{d\tilde{W}}{dx} = 0 \quad (25)$$

can be integrated:

$$\tilde{W}(x) = e^{\frac{\rho}{a}(x-1)} \left(\frac{\lambda}{\rho} + \frac{\lambda a}{\rho^2} + \frac{1}{\rho(\alpha-1)} + \tilde{w}_1 \right) - \frac{x^{1-\alpha}}{\rho(\alpha-1)} - \frac{\lambda}{\rho} x - \frac{a\lambda}{\rho^2} - \frac{e^{\frac{\rho}{a}x}}{\rho} \int_1^x e^{-\frac{\rho}{a}u} u^{-\alpha} du. \quad (26)$$

Here $\tilde{w}_1 = \tilde{W}(1)$ is a fixed parameter.

Suppose for a moment that no impulses are allowed, so that $x(t) = x_0 + at$. We omit the π index because here is a single control policy. We have a family of functions $\tilde{W}(x)$ depending on the initial value \tilde{w}_1 , but only one of them coincides with the objective function

$$\liminf_{T \rightarrow \infty} \int_0^T \left\{ e^{-\rho t} \frac{(x(t))^{1-\alpha}}{1-\alpha} - \lambda x(t) \right\} dt = W^*(x_0).$$

In this situation, for the function \tilde{W} , all the parts of Condition 2 are obviously satisfied ($\mathcal{D} = \emptyset, T_1 = \infty, \mathcal{L}^* = \emptyset$) except for (iv).

Since $W^* < 0$, the case $\limsup_{T \rightarrow \infty} e^{-\rho T} W^*(x(T)) > 0$ is excluded and we need to find such an initial value w_1^* that

$$\lim_{T \rightarrow \infty} e^{-\rho T} \tilde{W}(x(T)) = 0, \quad \text{where } x(T) = x_0 + aT, \quad x_0 > 0. \quad (27)$$

Equation (27) is equivalent to the following:

$$\lim_{T \rightarrow \infty} e^{\frac{\rho}{a}x_0} \left\{ e^{-\frac{\rho}{a}x_0} \left(\frac{\lambda}{\rho} + \frac{\lambda a}{\rho^2} + \frac{1}{\rho(\alpha-1)} + \tilde{w}_1 \right) - \frac{1}{\rho} \int_1^{x_0+aT} e^{-\frac{\rho}{a}u} u^{-\alpha} du \right\} = 0.$$

Therefore,

$$w_1^* = \frac{e^{\frac{\rho}{a}}}{\rho} \left(\frac{\rho}{a} \right)^{\alpha-1} \Gamma \left(1 - \alpha, \frac{\rho}{a} \right) - \frac{1}{\rho(\alpha-1)} - \frac{\lambda(\rho+a)}{\rho^2}, \quad (28)$$

and $W^*(x_0)$ is given by (26) with $\tilde{w}_1 = w_1^*$. Here $\Gamma(y, z) = \int_z^\infty e^{-u} u^{y-1} du$ is the incomplete gamma function [8, 3.381-3].

For the discounted impulsive control problem (17), the solution is given in the following statement.

Theorem 4 *The following statements take place.*

(a) *Equation*

$$H(x) := \left(e^{\frac{\rho x(1-b)}{a}} - 1 \right) \frac{(1-b)\lambda a}{\rho} - (1-b)e^{\frac{\rho x}{a}} \int_{bx}^x e^{-\frac{\rho u}{a}} u^{-\alpha} du \quad (29)$$

$$- \left(e^{\frac{\rho x(1-b)}{a}} - b \right) \left[\frac{x^{1-\alpha}(1-b^{1-\alpha})}{\alpha-1} + \lambda x(1-b) \right] = 0$$

has a single positive solution \bar{x} .

(b) *Let*

$$w_1 = \frac{e^{\frac{\rho}{a}}}{\rho} \int_1^{\bar{x}} e^{-\frac{\rho u}{a}} u^{-\alpha} du - \frac{\lambda(\rho+a)}{\rho^2} - \frac{1}{\rho(\alpha-1)} \quad (30)$$

$$- \left[\frac{\bar{x}^{1-\alpha}(1-b^{1-\alpha})}{\rho(\alpha-1)} + \frac{(1-b)\lambda\bar{x}}{\rho} + \frac{e^{\frac{\rho b\bar{x}}{a}}}{\rho} \int_{b\bar{x}}^{\bar{x}} e^{-\frac{\rho u}{a}} u^{-\alpha} du \right] \Bigg/ \left(e^{\frac{\rho b\bar{x}-\rho}{a}} - e^{\frac{\rho\bar{x}-\rho}{a}} \right)$$

and, for $0 < x < \bar{x}$, put $W(x) = \tilde{W}(x)$, where \tilde{W} is given by formula (26) under $\tilde{w}_1 = w_1$. For the intervals $[\bar{x}, \frac{\bar{x}}{b}]$, $[\frac{\bar{x}}{b}, \frac{\bar{x}}{b^2}]$, \dots the function W is defined recursively: $W(x) := W(bx)$. Then items (i, ii, iii) of Condition 2 are satisfied by the function W and the feedback policy defined in part (c) of this statement.

(c) *The function $W(x_0) = \sup_{\pi} J_{\rho}(x_0, \pi) = J_{\rho}(x_0, \pi^*)$ is the Bellman function, where the (feedback) optimal policy π^* is given by*

$$\mathcal{L}^* = [\bar{x}, \infty), \quad v^{f, \mathcal{L}^*}(x) = i \quad \text{if } x \in \left[\frac{\bar{x}}{b^{i-1}}, \frac{\bar{x}}{b^i} \right).$$

See the appendix for the proof. The shape of the Bellman function W is plotted in Figure 2.

Remark 4 *Let us calculate the limit of \bar{x} when ρ approaches zero. One can easily show that, for any $x > 0$,*

$$\lim_{\rho \rightarrow 0} H(x) = 0 \quad \text{and} \quad \lim_{\rho \rightarrow 0} \frac{H(x)}{\rho} = \frac{x^2(1-b)}{a} \left[\frac{\lambda(b^2-1)}{2} + \frac{x^{-\alpha}(1-b^{2-\alpha})}{2-\alpha} \right].$$

Let

$$\bar{x}_0 = \left[\frac{2(1-b^{2-\alpha})}{\lambda(1-b^2)(2-\alpha)} \right]^{1/\alpha}, \quad (31)$$

i.e.

$$\lim_{\rho \rightarrow 0} \frac{H(x)}{\rho} \begin{cases} > 0, & \text{if } x < \bar{x}_0, \\ < 0, & \text{if } x > \bar{x}_0, \\ = 0, & \text{if } x = \bar{x}_0. \end{cases}$$

The function $\frac{H(x)}{\rho}$ is continuous with respect to ρ . Therefore, for any small enough $\varepsilon > 0$,

$$\exists \delta > 0 : \forall \rho \in (0, \delta) \quad \frac{H(\bar{x}_0 - \varepsilon)}{\rho} > 0 \quad \text{and} \quad \frac{H(\bar{x}_0 + \varepsilon)}{\rho} < 0$$

meaning that \bar{x}_{ρ} , the solution to (29) at $\rho \in (0, \delta)$, satisfies $\bar{x}_{\rho} \in (\bar{x}_0 - \varepsilon, \bar{x}_0 + \varepsilon)$. This means $\lim_{\rho \rightarrow 0+} \bar{x}_{\rho} = \bar{x}_0$. Note that (31) is the optimal threshold if we consider the long-run average reward with the same reward rate c .

Remark 5 *The two theorems established in this section define our proposed threshold-based AQM scheme, which asserts that if the sending rate is smaller than \bar{x} , then do not send any congestion notification, while if the sending rate is greater or equal to \bar{x} , then send (multiple, if needed) congestion notifications until the sending rate is reduced to some level below \bar{x} with \bar{x} given by (20) under the average criterion and by Theorem 4(a) under the discounted criterion.*

4 Conclusion

To sum up, in this paper, we studied optimal impulsive control problems on infinite horizon with both discounted and time average criteria, for which we established the dynamic programming approach. Our general theoretical results are then applied to construct a novel AQM scheme, which takes into account not only the traffic transiting through the bottleneck links but also the congestion control algorithms operating at the edges of the network. Our network model is deterministic, and the investigations are already nontrivial and technically involving. So a natural continuation of this work would be to consider a more realistic and detailed stochastic model for the underlying queueing network; see e.g., [18].

A

A.1 Proof of Theorem 3

Proof of Theorem 3. Suppose $\gamma < 1$. By Theorem 1, it suffices to show that Condition 1 is satisfied by the policy $\pi^* = (\mathcal{L}^*, v^{\mathcal{L}^*})$, the constant g given by (21) and the function

$$h(x) = \begin{cases} h_0(x), & \text{if } x \in (0, \bar{x}), \\ h_k(x) = h_0(b^k x), & \text{if } x \in [\bar{x}/b^{k-1}, \bar{x}/b^k], \end{cases} \quad (32)$$

where

$$h_0(x) = \frac{1}{a} \left[-\frac{x^{2-\alpha-\gamma}}{(1-\alpha)(2-\alpha-\gamma)} + \lambda \frac{x^{2-\gamma}}{2-\gamma} + g \frac{x^{1-\gamma}}{1-\gamma} \right],$$

and \bar{x} is given by (20). Standard analysis shows that the function h is bounded from below, and so Condition 1 is satisfied. Since part (i) of Condition 1 is trivially verified, we only verify its parts (ii,iii) as follows.

Consider firstly $x \in (0, \bar{x}) = X \setminus \mathcal{L}^*$. Then, we obtain from direct calculations that $\left(\frac{x^{1-\alpha}}{1-\alpha} - \lambda x\right) - g + \frac{\partial h(x)}{\partial x} a x^\gamma = \left(\frac{x^{1-\alpha}}{1-\alpha} - \lambda x\right) - g + \frac{\partial h_0(x)}{\partial x} a x^\gamma = 0$. Let us show that $\sup_{m=1,2,\dots} \{h(b^m x) - h(x)\} = \sup_{m=1,2,\dots} \{h_0(b^m x) - h_0(x)\} \leq 0$ for $x \in (0, \bar{x})$ as follows. Define $\Delta_1(x) := h_0(bx) - h_0(x)$ for each $x \in (0, \bar{x})$. Then one can show that $\Delta_1(x) < 0$ for each $b \in (0, 1)$. Indeed, direct calculations give

$$\Delta_1(x) = \frac{x^{1-\gamma}}{a} \left[-\frac{(b^{2-\alpha-\gamma} - 1)x^{1-\alpha}}{(1-\alpha)(2-\alpha-\gamma)} + \lambda \frac{(b^{2-\gamma} - 1)x}{2-\gamma} + g \frac{(b^{1-\gamma} - 1)}{1-\gamma} \right],$$

so that for the strict negativity of $\Delta_1(x)$, it is equivalent to showing it for the following expression

$$\tilde{\Delta}_1(x) := -\frac{(b^{2-\alpha-\gamma} - 1)x^{1-\alpha}}{(1-\alpha)(2-\alpha-\gamma)} + \lambda \frac{(b^{2-\gamma} - 1)x}{2-\gamma} + g \frac{(b^{1-\gamma} - 1)}{1-\gamma},$$

whose first order and second order derivatives (with respect to x) are given by

$$\tilde{\Delta}'_1(x) = -\frac{(b^{2-\alpha-\gamma} - 1)x^{-\alpha}}{2-\alpha-\gamma} + \lambda \frac{(b^{2-\gamma} - 1)}{2-\gamma}$$

and

$$\tilde{\Delta}''_1(x) = \frac{\alpha(b^{2-\alpha-\gamma} - 1)x^{-\alpha-1}}{2-\alpha-\gamma}.$$

Under the conditions of the parameters, $\tilde{\Delta}''_1(x) < 0$ for each $x \in (0, \bar{x})$, and thus the function $\tilde{\Delta}_1(x)$ is concave on $(0, \bar{x})$ achieving its unique maximum at the stationary point given by $x = \bar{x} = \left\{ \frac{(2-\gamma)(1-b^{2-\alpha-\gamma})}{(2-\alpha-\gamma)(1-b^{2-\gamma})\lambda} \right\}^{\frac{1}{\alpha}} > 0$. Note that $\tilde{\Delta}_1(\bar{x}) = 0$ and $\lim_{x \downarrow 0} \tilde{\Delta}_1(x) \leq 0$. It follows from the above observations and the standard analysis of derivatives that $\tilde{\Delta}_1(x) < 0$ and thus $\Delta_1(x) < 0$ for each $x \in (0, \bar{x})$.

Since $\frac{\partial \bar{x}}{\partial b} \leq 0$ for each $b \in (0, 1)$ as can be easily verified, one can replace b with b^m ($m = 2, 3, \dots$) in the above argument to obtain that $h_0(b^m x) - h_0(x) < 0$ for each $x \in (0, \bar{x})$, and thus

$$\sup_{m=1,2,\dots} \{h_0(b^m x) - h_0(x)\} \leq 0 \quad (33)$$

for $x \in (0, \bar{x})$, as desired. Hence, it follows that Condition 1(ii,iii) is satisfied on $(0, \bar{x})$.

Next, we show by induction that Condition 1(ii,iii) is satisfied on $[\frac{\bar{x}}{b^{k-1}}, \frac{\bar{x}}{b^k})$, $k = 1, 2, \dots$. Let us consider the case of $k = 1$, i.e., the interval $[\bar{x}, \frac{\bar{x}}{b})$. By the definition of the function $h(x)$, we have

$$\sup_{m=1,2,\dots} \{h(b^m x) - h(x)\} = 0 \quad (34)$$

for $x \in [\bar{x}, \bar{x}/b)$. Indeed, by the definition of $h(x)$, we have

$$h(bx) - h(x) = 0 \quad (35)$$

for $x \in [\bar{x}, \bar{x}/b)$, whereas for each $m = 2, 3, \dots$ and $x \in [\bar{x}, \frac{\bar{x}}{b})$, it holds that $h(b^m x) - h(x) = h_0(b^m x) - h_0(bx) \leq 0$, which follows from that $bx \in (0, \bar{x})$, $b^m x = b^{m-1}(bx) \in (0, \bar{x})$ for each $x \in [\bar{x}, \frac{\bar{x}}{b})$, and (33). Furthermore, one can show that

$$\Delta_2(x) := \frac{x^{1-\alpha}}{1-\alpha} - \lambda x - g + \frac{\partial h(x)}{\partial x} a x^\gamma \leq 0 \quad (36)$$

for each $x \in [\bar{x}, \bar{x}/b)$, which follows from the following observations. Since $h(x) = h_1(x) = h_0(bx)$, we see $\Delta_2(x) = -\frac{(b^{2-\alpha-\gamma}-1)x^{1-\alpha}}{1-\alpha} + \lambda(b^{2-\gamma}-1)x + g(b^{1-\gamma}-1)$ for each $x \in [\bar{x}, \frac{\bar{x}}{b})$, and in particular,

$$\Delta_2(\bar{x}) = 0, \quad (37)$$

as can be easily verified. The derivative of the function $\Delta_2(x)$ with respect to x is given by $\Delta_2'(x) = -(b^{2-\alpha-\gamma}-1)x^{-\alpha} + \lambda(b^{2-\gamma}-1)$. If $2-\gamma-\alpha < 0$, then $\Delta_2'(x) < 0$, which together with (37) shows $\Delta_2(x) \leq 0$ on $[\bar{x}, \frac{\bar{x}}{b})$. If $2-\gamma-\alpha > 0$, then $\Delta_2''(x) = \alpha(b^{2-\alpha-\gamma}-1)x^{-\alpha-1} < 0$ and thus, the function $\Delta_2(x)$ is concave with the maximum attained at the stationary point $x = \left(\frac{1-b^{2-\alpha-\gamma}}{(1-b^{2-\gamma})\lambda}\right)^{\frac{1}{\alpha}}$. Since $\left(\frac{1-b^{2-\alpha-\gamma}}{(1-b^{2-\gamma})\lambda}\right)^{\frac{1}{\alpha}} \leq \bar{x}$, (37) implies $\Delta_2(x) \leq 0$ on $[\bar{x}, \frac{\bar{x}}{b})$, as desired. By the way, for the later reference, the above observations actually show that

$$G(x) := -\frac{(b^{2-\alpha-\gamma}-1)x^{1-\alpha}}{1-\alpha} + \lambda(b^{2-\gamma}-1)x + g(b^{1-\gamma}-1) \leq 0 \quad (38)$$

for all $x \geq \bar{x}$. Thus, combining (34), (35), and (36) shows that Condition 1(ii,iii) is satisfied on $[\bar{x}, \frac{\bar{x}}{b})$.

Assume that for each $x \in [\frac{\bar{x}}{b^{k-1}}, \frac{\bar{x}}{b^k})$ and each $k = 1, 2, \dots, M$, relations (34) and (36) hold, together with

$$h(b^k x) - h(x) = 0 \quad (\text{the corresponding version of (35)}). \quad (39)$$

Now we consider the case of $k = M+1$, i.e., when $x \in [\frac{\bar{x}}{b^M}, \frac{\bar{x}}{b^{M+1}})$. For each $x \in [\frac{\bar{x}}{b^M}, \frac{\bar{x}}{b^{M+1}})$, when $m = 1, 2, \dots, M$, it holds that $b^m x \in [\frac{\bar{x}}{b^{M-m}}, \frac{\bar{x}}{b^{M+1-m}})$, and thus $h(b^m x) - h(x) = h_0(b^{M+1}x) - h_0(b^{M+1}(x)) = 0$; when $m = M+1, M+2, \dots$, $b^m x \in (0, \bar{x}) = \mathcal{L}^*$, and thus $h(b^m x) - h(x) = h_0(b^m x) - h_0(b^{M+1}x) = 0$ if $m = M+1$, and $h(b^m x) - h(x) = h_0(b^{m-(M+1)}(b^{M+1}x)) - h_0(b^{M+1}x) \leq 0$ if $m > M+1$, by (33). Thus, we see (34) holds for $x \in [\frac{\bar{x}}{b^M}, \frac{\bar{x}}{b^{M+1}})$. Note that in the above we have also incidentally verified the validity of (39) for the case of $k = M+1$.

Below we verify (36) for the case of $k = M+1$, which would complete the proof by induction. To this end, we first present some preliminary observations that hold for each $k = 1, 2, \dots$. For each $k = 1, 2, \dots$, since $h(x) = h_k(x) = h_0(b^k x)$ for each $x \in [\frac{\bar{x}}{b^{k-1}}, \frac{\bar{x}}{b^k})$, we have

$$\Delta_2(x) := -\frac{(b^{k(2-\alpha-\gamma)}-1)x^{1-\alpha}}{(1-\alpha)} + \lambda x(b^{k(2-\gamma)}-1) + g(b^{k(1-\gamma)}-1).$$

For the convenience of the future reference, let us introduce the notation

$$\begin{aligned}\tilde{\Delta}_k(x) &:= -\frac{(b^{k(2-\alpha-\gamma)} - 1)x^{1-\alpha}}{1-\alpha} + \lambda(b^{k(2-\gamma)} - 1)x + g(b^{k(1-\gamma)} - 1) \\ &= b^{k(1-\gamma)}\left(-\frac{b^{k(1-\alpha)}x^{1-\alpha}}{1-\alpha} + \lambda b^k x + g\right) - \left(-\frac{x^{1-\alpha}}{1-\alpha} + \lambda x + g\right)\end{aligned}$$

for each $x > 0$. Therefore, for $x \in [\frac{\bar{x}}{b^{k-2}}, \frac{\bar{x}}{b^{k-1}})$, we have

$$\Delta_2(x) = \tilde{\Delta}_{k-1}(x) = b^{(k-1)(1-\gamma)}\left(-\frac{b^{(k-1)(1-\alpha)}x^{1-\alpha}}{1-\alpha} + \lambda b^{k-1}x + g\right) - \left(-\frac{x^{1-\alpha}}{1-\alpha} + \lambda x + g\right).$$

Let us define

$$F(x) := -\frac{x^{1-\alpha}}{1-\alpha} + \lambda x + g$$

for each $x > 0$. We then have from the direct calculations that

$$\tilde{\Delta}_{k-1}\left(\frac{\bar{x}}{b^{k-2}}\right) = b^{(k-1)(1-\gamma)}F(b\bar{x}) - F\left(\frac{\bar{x}}{b^{k-2}}\right) \quad (40)$$

for each $k = 1, 2, \dots$. Focusing on $F\left(\frac{\bar{x}}{b^{k-2}}\right)$, we have

$$\begin{aligned}b^{1-\gamma}F\left(\frac{\bar{x}}{b^{k-2}}\right) &= -\frac{\bar{x}^{1-\alpha}}{1-\alpha} \frac{b^{1-\gamma}}{b^{(k-2)(1-\alpha)}} + \lambda \bar{x} \frac{b^{1-\gamma}}{b^{k-2}} + gb^{1-\gamma} \\ &= -\frac{\bar{x}^{1-\alpha}}{1-\alpha} \frac{b^{2-\alpha-\gamma}}{b^{(k-1)(1-\alpha)}} + \lambda \bar{x} \frac{b^{2-\gamma}}{b^{k-1}} + gb^{1-\gamma} \\ &= -\frac{\left(\frac{\bar{x}}{b^{(k-1)}}\right)^{1-\alpha}}{1-\alpha} b^{2-\alpha-\gamma} + \lambda \left(\frac{\bar{x}}{b^{k-1}}\right) b^{2-\gamma} + gb^{1-\gamma}.\end{aligned}$$

Recall that in the above, we have proved that $G(x) \leq 0$ for $x \geq \bar{x}$, see (38). Thus, we have $G\left(\frac{\bar{x}}{b^{k-1}}\right) \leq 0$, i.e., $-\frac{b^{2-\alpha-\gamma}\left(\frac{\bar{x}}{b^{k-1}}\right)^{1-\alpha}}{1-\alpha} + \lambda b^{2-\gamma}\left(\frac{\bar{x}}{b^{k-1}}\right) + gb^{1-\gamma} \leq -\frac{\left(\frac{\bar{x}}{b^{k-1}}\right)^{1-\alpha}}{1-\alpha} + \lambda\left(\frac{\bar{x}}{b^{k-1}}\right) + g$. Consequently,

$$b^{1-\gamma}F\left(\frac{\bar{x}}{b^{k-2}}\right) \leq -\frac{\left(\frac{\bar{x}}{b^{k-1}}\right)^{1-\alpha}}{1-\alpha} + \lambda\left(\frac{\bar{x}}{b^{k-1}}\right) + g = F\left(\frac{\bar{x}}{b^{k-1}}\right).$$

Now we verify (36) for the particular case of $k = M + 1$. By the inductive supposition, (36) holds for $x \in [\frac{\bar{x}}{b^{M-1}}, \frac{\bar{x}}{b^M})$, we thus have $\Delta_2\left(\frac{\bar{x}}{b^{M-1}}\right) \leq 0$, and

$$0 \geq \tilde{\Delta}_M\left(\frac{\bar{x}}{b^{M-1}}\right) = b^{M(1-\gamma)}F(b\bar{x}) - F\left(\frac{\bar{x}}{b^{M-1}}\right) \geq b^{(M)(1-\gamma)}F(b\bar{x}) - \frac{1}{b^{1-\gamma}}F\left(\frac{\bar{x}}{b^M}\right).$$

Therefore, we obtain that $b^{(M+1)(1-\gamma)}F(b\bar{x}) - F\left(\frac{\bar{x}}{b^M}\right) \leq 0$, and by (40),

$$\Delta_2\left(\frac{\bar{x}}{b^M}\right) \leq 0. \quad (41)$$

Furthermore, the derivative of the function $\tilde{\Delta}_{M+1}(x)$ with respect to x is given by $\tilde{\Delta}'_{M+1}(x) = -(b^{(M+1)(2-\alpha-\gamma)} - 1)x^{-\alpha} + \lambda(b^{(M+1)(2-\gamma)} - 1)$. If $2 - \gamma - \alpha < 0$, then $\tilde{\Delta}'_{M+1}(x) < 0$. Thus, by (41), we obtain that $\Delta_2(x) = \tilde{\Delta}_{M+1}(x) \leq 0$ for $x \in [\frac{\bar{x}}{b^M}, \frac{\bar{x}}{b^{M+1}})$. If $2 - \gamma - \alpha > 0$, then $\tilde{\Delta}''_{M+1}(x) = \alpha(b^{(M+1)(2-\alpha-\gamma)} - 1)x^{-\alpha-1} < 0$, and in turn, the function $\tilde{\Delta}_{M+1}(x)$ is concave with the maximum attained at the stationary point $x = \left(\frac{1-b^{(M+1)(2-\alpha-\gamma)}}{(1-b^{(M+1)(2-\gamma)})\lambda}\right)^{\frac{1}{\alpha}}$. Moreover, we have $\frac{\sum_{m=0}^M b^{m(2-\alpha-\gamma)}}{\sum_{m=0}^M b^{m(2-\gamma)}} \leq \frac{1}{b^{M\alpha}}$, which follows from the fact that for each $m = 0, 1, \dots, M$, $m(2 - \alpha - \gamma) + M\alpha \geq m(2 - \gamma)$, so that $b^{m(2-\alpha-\gamma)}b^{M\alpha} \leq b^{m(2-\gamma)}$. From this we see

$$\begin{aligned}&\frac{(1 - b^{2-\alpha-\gamma}) \sum_{m=0}^M b^{m(2-\alpha-\gamma)}}{(1 - b^{2-\gamma}) \sum_{m=0}^M b^{m(2-\gamma)}} \leq \frac{1}{b^{M\alpha}} \frac{1 - b^{2-\alpha-\gamma}}{1 - b^{2-\gamma}} \\ \Rightarrow &\frac{1 - b^{(M+1)(2-\alpha-\gamma)}}{(1 - b^{(M+1)(2-\gamma)})\lambda} \leq \frac{1}{b^{M\alpha}} \frac{(2 - \gamma)(1 - b^{2-\alpha-\gamma})}{(2 - \alpha - \gamma)(1 - b^{2-\gamma})\lambda} \\ \Leftrightarrow &\left(\frac{1 - b^{(M+1)(2-\alpha-\gamma)}}{(1 - b^{(M+1)(2-\gamma)})\lambda}\right)^{\left(\frac{1}{\alpha}\right)} \leq \frac{\bar{x}}{b^M}.\end{aligned}$$

Finally, it follows from the last line of the previous inequalities, the concavity of the function $\tilde{\Delta}_{M+1}$ and (41) that $\Delta_2(x) \leq 0$ for $x \in [\frac{\bar{x}}{b^M}, \frac{\bar{x}}{b^{M+1}})$, which verifies (36), and thus completes the proof.

For the case of $\gamma = 1$, we consider the function h in the form of (32) with \bar{x} being still given by (20), and h_0 being defined by

$$h_0(x) = \frac{1}{a} \left\{ -\frac{x^{1-\alpha}}{(1-\alpha)^2} + \lambda x + g \ln x \right\}.$$

Condition 1(i, ii, iii) can be verified similarly to the case of $\gamma < 1$. We now focus on the verification of Condition 1(iv). If $\alpha > 1$, then $g < 0$, and standard analysis shows that h_0 is bounded from below, so that Condition 1(iv) is verified. Consider the case of $\alpha < 1$. Then $g > 0$, and any policy π satisfying $\limsup_{t \rightarrow \infty} \frac{h(x^\pi(t))}{t} < 0$ cannot be optimal. Indeed, it follows from $\limsup_{t \rightarrow \infty} \frac{h(x^\pi(t))}{t} < 0$ that for each $\epsilon > 0$, there exists some $T > 0$ such that $h(x^\pi(t)) \leq -\epsilon t$ for all $t > T$. Therefore, $\lim_{t \rightarrow \infty} h(x^\pi(t)) = -\infty$. Since $-\frac{x^{1-\alpha}}{(1-\alpha)^2} + \lambda x$ is bounded on $[0, \bar{x}]$, necessarily $\lim_{t \rightarrow \infty} x^\pi(t) = 0$. But then $\liminf_{t \rightarrow \infty} \frac{1}{t} \int_0^t \left(\frac{(x^\pi(s))^{1-\alpha}}{1-\alpha} - \lambda x^\pi(s) \right) ds = 0 < g$ (c.f. (18)). Therefore, it suffices to consider policies π which verify Condition 1(iv), i.e., $\limsup_{t \rightarrow \infty} \frac{h(x^\pi(t))}{t} \geq 0$. Now the statement follows from Remark 2.

A.2 Proof of Theorem 4

Some comments and remarks are in position before we give the proof of this theorem. In fact, they help the reader follow the proof. For $b = 0.5$, $\rho = 1$, $\alpha = 1.3$, $\lambda = 2$, $a = 0.2$ the graph of the function W is presented in Figure 2. Here $\bar{x} = 0.7901$ and $w_1 = -4.9301$. The dashed line represents the graph of the function

$$z(x) = -\frac{1}{\rho} \left(\frac{x^{1-\alpha}}{\alpha-1} + \lambda x \right). \quad (42)$$

When $\tilde{W}(x) = z(x)$, we have $\frac{d\tilde{W}}{dx} = 0$; if $\tilde{W}(x) > z(x)$ (resp., $\tilde{W}(x) < z(x)$), the function \tilde{W} increases (resp., decreases). The dotted line represents the graph of the function

$$v(x) = \frac{a(x^{-\alpha} - \lambda)}{\rho^2} - \frac{1}{\rho} \left(\frac{x^{1-\alpha}}{\alpha-1} + \lambda x \right).$$

If $\tilde{W}(x) = v(x)$, then from (25) we have

$$\begin{aligned} a^2 \frac{d^2 \tilde{W}}{dx^2} &= a^2 \left[a\rho \frac{d\tilde{W}}{dx} + \lambda a - ax^{-\alpha} \right] \\ &= a^2 \left[\rho^2 \tilde{W}(x) + \rho\lambda x + \frac{\rho x^{1-\alpha}}{\alpha-1} + \lambda a - ax^{-\alpha} \right] = 0, \end{aligned}$$

that is, x is the point of inflection of the function \tilde{W} . This reasoning applies to any solution of equation (25).

In the graph, for $0 < x < \bar{x}$, the Bellman function $W(x) = \tilde{W}(x)$ has three parts, denoted below as I, II and III, where it increases, strictly decreases, and again increases. Correspondingly, the function $\tilde{W}(bx)$ also has three parts I, II and III, where it increases, strictly decreases and increases again, and $W(x) = \tilde{W}(bx)$ for $\bar{x} \leq x < \frac{\bar{x}}{b}$. The point \bar{x} is such that

$$\tilde{W}(\bar{x}) = \tilde{W}(b\bar{x}) \quad \text{and} \quad \left. \frac{d\tilde{W}(x)}{dx} \right|_{\bar{x}} = \left. \frac{d\tilde{W}(bx)}{dx} \right|_{\bar{x}}. \quad (43)$$

As is shown in the proof of Theorem 4 below, these two equations are satisfied if and only if \bar{x} solves equation (29).

Proof of Theorem 4. (a) Firstly, let us prove that no more than one positive number \bar{x} can satisfy equations (43). If \bar{x} satisfies (43), then the function \tilde{W} cannot have only one increasing branch above function v because the two increasing functions $\tilde{W}(x)$ and $\tilde{W}(bx)$ could not have common points.

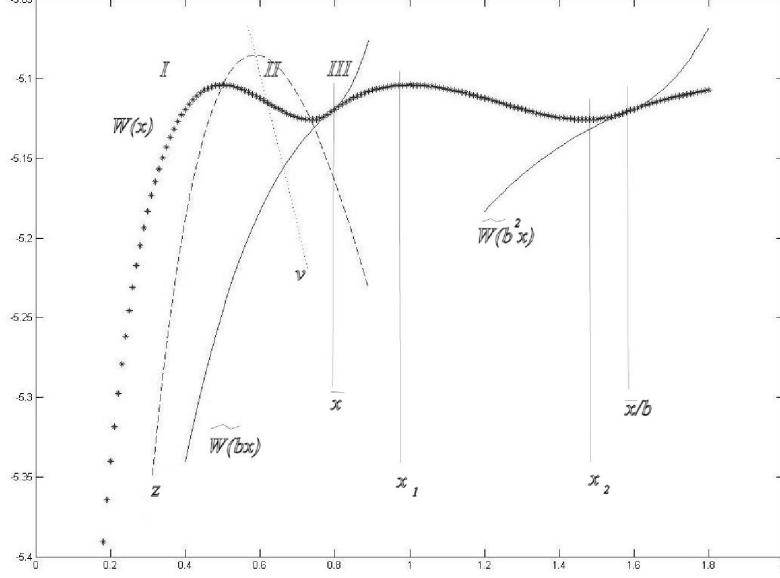


Figure 2: Graph of the Bellman function $W(x)$ (bright line with the star point markers).

The increasing part I of function $\tilde{W}(x)$ cannot intersect with $\tilde{W}(bx)$.

The strictly decreasing part II of the function $\tilde{W}(x)$ cannot intersect with the parts II and III of the function $\tilde{W}(bx)$. Possible common points with the part I of $\tilde{W}(bx)$ are of no interest because here $\frac{d\tilde{W}(x)}{dx} < 0$ and $\frac{d\tilde{W}(bx)}{dx} \geq 0$.

The increasing part III of function $\tilde{W}(x)$ can intersect with the parts I and II of function $\tilde{W}(bx)$, but again the latter case is of no interest because here $\frac{d\tilde{W}(x)}{dx} \geq 0$ and $\frac{d\tilde{W}(bx)}{dx} < 0$.

Thus, the only possibility to satisfy (43) is the case when the increasing part III of $\tilde{W}(x)$ touches the increasing part I of function $\tilde{W}(bx)$. The inflection line $v(x)$ is located between the increasing and decreasing branches of the function $z(x)$, so that the part III of $\tilde{W}(x)$ is convex and the part I of $\tilde{W}(bx)$ is concave, meaning that no more than one point \bar{x} can satisfy the equations (43).

Using formula (26), the equations (43) can be rewritten as follows:

$$\begin{aligned}
0 &= \tilde{W}(x) - \tilde{W}(bx) = \left(e^{\frac{\rho x}{a}} - e^{\frac{b\rho x}{a}} \right) \left[e^{-\frac{\rho}{a}} \left(\frac{\lambda}{\rho} + \frac{\lambda a}{\rho^2} + \frac{1}{\rho(\alpha-1)} + \tilde{w}_1 \right) \right. \\
&\quad \left. - \frac{1}{\rho} \int_1^x e^{-\frac{\rho u}{a}} u^{-\alpha} du \right] - \frac{x^{1-\alpha}(1-b^{1-\alpha})}{\rho(\alpha-1)} - (1-b) \frac{\lambda x}{\rho} + \frac{e^{\frac{b\rho x}{a}}}{\rho} \int_x^{bx} e^{-\frac{\rho u}{a}} u^{-\alpha} du; \quad (44) \\
0 &= \frac{d\tilde{W}(x)}{dx} - \frac{d\tilde{W}(bx)}{dx} = \left(\frac{\rho}{a} e^{\frac{\rho x}{a}} - \frac{b\rho}{a} e^{\frac{b\rho x}{a}} \right) \left[e^{-\frac{\rho}{a}} \left(\frac{\lambda}{\rho} + \frac{\lambda a}{\rho^2} + \frac{1}{\rho(\alpha-1)} + \tilde{w}_1 \right) \right. \\
&\quad \left. - \frac{1}{\rho} \int_1^x e^{-\frac{\rho u}{a}} u^{-\alpha} du \right] - (1-b) \frac{\lambda}{\rho} + \frac{be^{\frac{b\rho x}{a}}}{a} \int_x^{bx} e^{-\frac{\rho u}{a}} u^{-\alpha} du.
\end{aligned}$$

After we multiply these equations by factors $(1 - be^{(b-1)\frac{\rho x}{a}})$ and $\frac{a}{\rho} (1 - e^{(b-1)\frac{\rho x}{a}})$ correspondingly and subtract the equations, the variable \tilde{w}_1 is cancelled and we obtain equation

$$\begin{aligned}
0 &= \left(1 - be^{(b-1)\frac{\rho x}{a}} \right) \left[\frac{b^{1-\alpha} x^{1-\alpha}}{\rho(\alpha-1)} + \frac{e^{\frac{b\rho x}{a}}}{\rho} \int_x^{bx} e^{-\frac{\rho u}{a}} u^{-\alpha} du - \frac{x^{1-\alpha}}{\rho(\alpha-1)} - (1-b) \frac{\lambda x}{\rho} \right] \\
&\quad - \frac{a}{\rho} \left(1 - e^{\frac{(b-1)\rho x}{a}} \right) \left[\frac{be^{\frac{b\rho x}{a}}}{a} \int_x^{bx} e^{-\frac{\rho u}{a}} u^{-\alpha} du - (1-b) \frac{\lambda}{\rho} \right],
\end{aligned}$$

which is equivalent to $H(x) = 0$.

Equation (30) follows directly from the first one of equations (44): if we know the value of x (equal to \bar{x}), we can compute the value of $\tilde{w}_1 = w_1$.

To prove the solvability of the equation (29) we compute the following limits:

$$\begin{aligned}\lim_{x \rightarrow \infty} H(x) &\leq - \lim_{x \rightarrow \infty} e^{\frac{\rho x(1-b)}{a}} \cdot \lambda x(1-b) = -\infty; \\ \lim_{x \rightarrow 0} H(x) &= \lim_{x \rightarrow 0} (b-1) \left[\frac{x^{1-\alpha}(1-b^{1-\alpha})}{\alpha-1} + \int_{bx}^x e^{-\frac{\rho u}{a}} u^{-\alpha} du \right],\end{aligned}$$

and the positive expression in the square brackets does not exceed

$$\begin{aligned}\frac{x^{1-\alpha}(1-b^{1-\alpha})}{\alpha-1} + \int_{bx}^x \left[1 - \frac{\rho u}{a} + \frac{1}{2} \left(\frac{\rho u}{a} \right)^2 \right] u^{-\alpha} du &= \frac{x^{1-\alpha}(1-b^{1-\alpha})}{\alpha-1} \\ + \left[\frac{u^{1-\alpha}}{1-\alpha} - \frac{\rho u^{2-\alpha}}{a(2-\alpha)} + \frac{\rho^2 u^{3-\alpha}}{2a^2(3-\alpha)} \right]_{bx}^x &= \left[\frac{\rho^2 u^{3-\alpha}}{2a^2(3-\alpha)} - \frac{\rho u^{2-\alpha}}{a(2-\alpha)} \right]_{bx}^x \rightarrow 0 \text{ as } x \rightarrow 0,\end{aligned}$$

so that $\lim_{x \rightarrow 0} H(x) = 0$.

Finally,

$$\begin{aligned}\frac{dH}{dx} &= -\frac{\rho(1-b)}{a} e^{\frac{\rho x(1-b)}{a}} \left[\frac{x^{1-\alpha}(1-b^{1-\alpha})}{\alpha-1} + \lambda x(1-b) \right] \\ &+ \left(b - e^{\frac{\rho x(1-b)}{a}} \right) \left[\lambda(1-b) - x^{-\alpha}(1-b^{1-\alpha}) \right] + \lambda(1-b)^2 e^{\frac{\rho x(1-b)}{a}} \\ &- \frac{\rho}{a} (1-b) e^{\frac{\rho x}{a}} \int_{bx}^x e^{-\frac{\rho u}{a}} u^{-\alpha} du - (1-b) e^{\frac{\rho x}{a}} \left[e^{-\frac{\rho x}{a}} x^{-\alpha} - b e^{-\frac{\rho bx}{a}} (bx)^{-\alpha} \right] \\ &= -\frac{\rho(1-b)(1-b^{1-\alpha})}{a(\alpha-1)} x^{1-\alpha} - \left(b-1 - \frac{\rho x(1-b)}{a} \right) \left[(1-b^{1-\alpha})x^{-\alpha} - \lambda(1-b) \right] \\ &+ \lambda(1-b)^2 - \frac{\rho}{a} (1-b) \frac{x^{1-\alpha}(1-b^{1-\alpha})}{1-\alpha} \\ &- (1-b) \left(1 + \frac{\rho x}{a} \right) \left[x^{-\alpha} \left(1 - \frac{\rho x}{a} \right) - b \left(1 - \frac{\rho bx}{a} \right) (bx)^{-\alpha} \right] + \epsilon(x),\end{aligned}$$

where $\lim_{x \rightarrow 0} \epsilon(x) = 0$; so

$$\lim_{x \rightarrow 0} \frac{dH}{dx} = \lim_{x \rightarrow 0} \frac{\rho(1-b)(1-b^{2-\alpha})}{a} x^{1-\alpha} = +\infty$$

meaning that the continuous function $H(x)$ increases from the limit zero for small values of x and becomes negative for big values of x .

Therefore, equation (29) has a single positive solution \bar{x} , and part (a) of the statement has been proved.

(b) Item (i) of Condition 2 is obviously satisfied ($\mathcal{D} = \emptyset$).

For Item (ii), we consider the following three cases.

(α) Let $0 < x \leq \bar{x}$. The differential equation (25) holds for function W on the interval $0 < x < \bar{x}$. For these values of x ,

$$W(b^i x) < W(x), \forall i \geq 1. \quad (45)$$

Indeed, to prove this, note that the function $W(bx) = \tilde{W}(bx)$ is increasing (c.f. Figure 2), so that $W(bx) > W(b^2x) > \dots$. As explained in the proof of part (a), Part III of the function $W(x)$ is convex and the function $\tilde{W}(bx)$ touching smoothly $W(x)$ at the point \bar{x} , is concave, so that $W(bx) = \tilde{W}(bx) < W(x)$ here. The same inequality holds for smaller values of x where $W(x)$ decreases (part II) and $W(bx) = \tilde{W}(bx)$ increases. Part I of the function $W(x)$ is obviously bigger than $W(bx)$, too. Thus the Bellman equation (24) is satisfied on the interval $0 < x < \bar{x}$ and also on the interval $(0, \bar{x}]$.

(β) Consider $x \in (\bar{x}, \bar{x}/b]$ and denote x_1 and x_2 the points of the analytical maximum and minimum of the function $W(x) = \tilde{W}(bx)$. (See Fig.2.)

For $x \in (\bar{x}, x_1)$ the function $W(x)$ is concave; hence

$$a \frac{dW}{dx} \Big|_x < a \frac{dW}{dx} \Big|_{\bar{x}} = a \frac{d\tilde{W}}{dx} \Big|_{\bar{x}} = \rho \tilde{W}(\bar{x}) - \frac{\bar{x}^{1-\alpha}}{1-\alpha} + \lambda \bar{x} = \rho[\tilde{W}(\bar{x}) - z(\bar{x})].$$

(See formula (42).) Since $W(x)$ increases starting from $W(\bar{x}) = \tilde{W}(\bar{x})$ and $z(x)$ decreases, we have

$$a \frac{dW}{dx} \Big|_x < \rho[W(x) - z(x)] = \rho W(x) - \left(\frac{x^{1-\alpha}}{1-\alpha} - \lambda x \right),$$

and the Bellman equation (24) is satisfied because here $W(x) = \tilde{W}(bx) = W(bx)$ and $W(b^{i+1}x) < W(bx)$ for all $i \geq 1$ by (45).

For $x \in [x_1, x_2]$ we have $W(x) > z(x)$ and $a \frac{dW}{dx} \leq 0$: remember, $W(x) = \tilde{W}(bx)$ and the latter function is of type II for $x \in [x_1, x_2]$. Therefore, again

$$a \frac{dW}{dx} - \rho \left[W(x) - \frac{1}{\rho} \left(\frac{x^{1-\alpha}}{1-\alpha} - \lambda x \right) \right] < 0$$

and the Bellman equation (24) is satisfied.

For $x \in (x_2, \bar{x}/b]$, we have

$$a \frac{dW}{dx} = b \frac{d\tilde{W}}{dx} \Big|_{bx} < \frac{d\tilde{W}}{dx} \Big|_{bx}$$

because the function \tilde{W} increases here and $b \in (0, 1)$. Next,

$$\rho[W(x) - z(x)] = \rho[\tilde{W}(bx) - z(x)] > \rho[\tilde{W}(bx) - z(bx)]$$

because the function $z(x)$ decreases. Therefore,

$$a \frac{dW}{dx} - \rho[W(x) - z(x)] < \frac{d\tilde{W}}{dx} \Big|_{bx} - \rho[\tilde{W}(bx) - z(bx)] = 0$$

since $bx \leq \bar{x}$, and, for these values, equation (25) holds. We see that the Bellman equation (24) is satisfied.

(γ) Finally, let us consider the remaining interval $(\frac{\bar{x}}{b}, \infty)$. Suppose

$$a \frac{dW}{dx} - \rho[W(x) - z(x)] < 0 \text{ for } x \in \left(\frac{\bar{x}}{b^{i-1}}, \frac{\bar{x}}{b^i} \right],$$

for some natural $i \geq 1$. Then, for $x \in (\frac{\bar{x}}{b^i}, \frac{\bar{x}}{b^{i+1}}]$, we have

$$a \frac{dW}{dx} - \rho[W(x) - z(x)] = ba \frac{dW}{dx} \Big|_{bx} - \rho[W(bx) - z(x)].$$

If $\frac{dW}{dx} \Big|_{bx} < 0$ then the last expression is negative. Otherwise,

$$ba \frac{dW}{dx} \Big|_{bx} \leq a \frac{dW}{dx} \Big|_{bx} \text{ and } z(x) < z(bx),$$

so that

$$a \frac{dW}{dx} - \rho[W(x) - z(x)] < a \frac{dW}{dx} \Big|_{bx} - \rho[W(bx) - z(bx)] < 0$$

by the induction supposition.

Thus, the Bellman equation (24) is satisfied for all $x > 0$.

To finish the proof of part (b) of this statement, it remains to note that Item (iii) of Condition 2 is also obviously satisfied:

$$\mathcal{L}^* = [\bar{x}, \infty); \quad v^{f, \mathcal{L}^*}(x) = v_i \text{ if } x \in \left[\frac{\bar{x}}{b^{i-1}}, \frac{\bar{x}}{b^i} \right).$$

(c) Note that item (iv) of Condition 2 is not satisfied. Indeed, there is an admissible control policy such that, on any time interval $(T - 1, T]$, $x^\pi(\cdot)$ is so close to zero that $e^{-\rho T}W(x^\pi(T)) < -1$. (Remember that $\lim_{x \rightarrow 0} W(x) = -\infty$.)

Let us fix an arbitrary $x_0 > 0$ and modify the reward rate:

$$\hat{c}(x) = \begin{cases} c(x), & \text{if } x \geq \min\{x_0, b\bar{x}\} := \hat{x}; \\ c(\hat{x}), & \text{if } x < \hat{x}. \end{cases}$$

Note that $\hat{c} \geq c$. The function $\tilde{W}(x)$ given by (26) will change only for $x < \hat{x} \leq x_0$ and remains increasing in its part I, meaning that this modified function \hat{W} satisfies all items (i)–(iii) of Condition 2: the proof is identical to the one presented above. But now Condition 2 (iv) is also satisfied because the function \hat{W} is bounded. Therefore, according to Theorem 2, $\sup_\pi \hat{J}_\rho(x_0, \pi) = \hat{W}(x_0) = \hat{J}_\rho(x_0, \pi^*)$, where \hat{J}_ρ corresponds to the reward rate \hat{c} . But

$$\sup_\pi J_\rho(x_0, \pi) \leq \sup_\pi \hat{J}_\rho(x_0, \pi) = \hat{W}(x_0) = W(x_0),$$

and for the feedback policy π^* , which is independent of x_0 , we have

$$W(x_0) = \hat{W}(x_0) = \hat{J}_\rho(x_0, \pi^*) = J_\rho(x_0, \pi^*).$$

The last equality holds because, under the feedback policy π^* , starting from x_0 , the trajectory $x^{\pi^*}(t)$ satisfies $x^{\pi^*}(t) \geq \hat{x}$ for all $t \geq 0$, and in this region $\hat{c} = c$.

Acknowledgement

This work is partially funded by INRIA Alcatel-Lucent Joint Lab, ADR ‘‘Semantic Networking’’.

References

- [1] Eitan Altman, Konstantin Avrachenkov, and Balakrishna Prabhu. Fairness in MIMD congestion control algorithms. *Telecommunication Systems*, 30(4):387–415, 2005.
- [2] Konstantin Avrachenkov, Urtzi Ayesta, and Alexei Piunovskiy. Convergence of trajectories and optimal buffer sizing for AIMD congestion control. *Performance Evaluation*, 67:501–527, 2010.
- [3] M. Bardi and I. Capuzzo-Dolcetta. *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*. Birkhauser, Boston, 1997.
- [4] A. Bensoussan and J. Lions. *Impulse Control and Quasivariational Inequalities*. μ , Gauthier-Villars, Montrouge, 1984.
- [5] S. Christensen. On the solution of general impulse control problems using superharmonic functions. *Stochastic Processes and Their Applications*, 124:709–729, 2014.
- [6] M. Davis. *Markov Models and Optimization*. Chapman and Hall, London, 1993.
- [7] F. Dufour and B. Miller. Necessary conditions for optimal singular stochastic control problems. *Stochastics*, 79:469–504, 2007.
- [8] I. Gradshteyn and I. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, New York, 2007.
- [9] Joao P. Hespanha, Stephan Bohacek, Katia Obraczka, and Junsoo Lee. Hybrid modeling of TCP congestion control. *Hybrid Systems: Computation and Control, LNCS v. 2034*, pages 291–304, 2001.

- [10] A. Hordijk and F. van der Duyn Schouten. Average optimal policies in markov decision drift processes with applications to a queueing and a replacement model. *Advances in Applied Probability*, 15:274–303, 1983.
- [11] S. Hou and K. Wong. Optimal impulsive control problem with application to human immunodeficiency virus treatment. *Journal of Optimization Theory and Applications*, 151:385–401, 2011.
- [12] Frank P. Kelly, Aman K. Maulloo, and D.K.H. Tan. Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49:237–252, 1998.
- [13] Tom Kelly. Scalable TCP: Improving performance in high-speed wide area networks. *Computer Communication Review*, 33(2):83–91, 2003.
- [14] R. Korn. Some applications of impulse control in mathematical finance. *Mathematical Methods of Operations Research*, 50:493–518, 1999.
- [15] Srisankar Kunniyur and R. Srikant. End-to-end congestion control schemes: utility functions, random losses and ecn marks. *IEEE/ACM Transactions on Networking*, 11(5):689–702, October 2003.
- [16] Yee-Ting Li, D. Leith, and R.N. Shorten. Experimental evaluation of tcp protocols for high-speed networks. *IEEE/ACM Transactions on Networking*, 15(5):1109–1122, 2007.
- [17] Steven H. Low and David E. Lapsley. Optimization flow control i: basic algorithm and convergence. *IEEE/ACM Transactions on Networking*, 7:861–874, 1999.
- [18] A. Miller and B. Miller. Control of connected markov chains. application to congestion avoidance in the internet. In *Proceedings of IEEE CDC-ECC 2011*, pages 7242–7248, 2011.
- [19] B. Miller and E. Rubinovich. *Impulsive Control in Continuous and Discrete-Continuous Systems*. Springer, New York, 2003.
- [20] Jeonghoon Mo and Jean Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on Networking*, 8:556–567, 2000.
- [21] Niels Möller, Chadi Barakat, Konstantin Avrachenkov, and Eitan Altman. Inter-protocol fairness between TCP New Reno and TCP Westwood. In *3rd EuroNGI Conference on Next Generation Internet Networks*, pages 127–134, 2007.
- [22] M. Motta and F. Rampazzo. Dynamic programming for nonlinear systems driven by ordinary and impulsive controls. *SIAM Journal on Control and Optimization*, 34:199–225, 1996.
- [23] Ludovic Noirie, Emmanuel Dotaro, Giovanna Carofiglio, Arnaud Dupas, Pascal Pecci, Daniel Popa, and Georg Post. Semantic networking: Flow-based, traffic-aware, and self-managed networking. *Bell Labs Technical Journal*, 14(2):23–38, 2009.
- [24] J. Palczewski and L. Stettner. Impulsive control of portfolios. *Applied Mathematics and Optimization*, 56:67–103, 2007.
- [25] A. Piunovskiy. Multicriteria impulsive control of jump markov processes. *Mathematical Methods of Operations Research*, 60:125–144, 2004.
- [26] F. Taringoo and P. Caines. On the optimal control of impulsive hybrid systems on riemannian manifolds. *SIAM Journal on Control and Optimization*, 51:3127–3153, 2013.
- [27] Y. Xiao, D. Cheng, and H. Qin. Optimal impulsive control in periodic ecosystem. *Systems and Control Letters*, 55:558–565, 2006.

- [28] A. Yushkevich. Verification theorems for markov decision processes with controllable deterministic drift, gradual and impulse controls. *Theory of Probability and Its Applications*, 34:474–496, 1989.
- [29] Yi Zhang, Alexei Piunovskiy, Urtzi Ayesta, and Konstantin Avrachenkov. Convergence of trajectories and optimal buffer sizing for MIMD congestion control. *Computer Communications*, 33:149–159, 2010.