Applications of Borovkov's Renovation Theory to Non-Stationary Stochastic Recursive Sequences and their Control

Eitan ALTMAN
INRIA
2004 Route des Lucioles
BP93, 06902 Sophia-Antipolis Cedex
France

Arie HORDIJK
Dept. of Mathematics and Computer Science
Leiden University
P.O.Box 9512, 2300RA Leiden
The Netherlands

Submitted: October 1994, Revised: January 1996

Abstract

We investigate in this paper the stability of non-stationary stochastic processes, arising typically in applications of control. The setting is known as stochastic recursive sequences, which allows us to construct on one probability space stochastic processes that correspond to different initial states and even different control policies. It does not require any Markovian assumptions. A natural criterion for stability for such processes is that the influence of the initial state disappears after some finite time; in other words, starting from different initial states, the process will couple after some finite time to the same limiting (not necessarily stationary nor ergodic) stochastic process. We investigate this as well as other types of coupling, and present conditions for them to occur uniformly in some class of control policies. We then use the coupling results to establish new theoretical aspects in the theory of non Markovian control.
Keywords: coupling, renovating events

1 Introduction

Many different notions for stability of stochastic discrete event dynamic systems have been introduced in the last decades, and many different tools were developed for determining stability conditions. For systems described by homogeneous Markov chains, the concepts of stability studied are mostly those of ergodicity (for the countable state space) or Harris ergodicity (for the general state space) (see e.g. [38, 8] and references therein); geometric ergodicity [31, 38], which strengthens the notion of ergodicity and includes also rates of convergence; the μ-uniform ergodicity and μ-uniform recurrence (see [14, 15, 39], see also [38] for the generalization to the non-countable case) which provides for convergence in operator norm, and hence also convergence uniformly in the initial state;

coupling convergence to a stationary regime (see e.g. [35]); and the existence and the convergence of moments [27, 40]. For the non Markov setting, the notion of stability studied were the existence (and uniqueness) of a stationary regime and also coupling convergence to it [3, 5, 7, 8, 12, 10, 36].

All the above notions of stability are related to the stationarity of the process studied, or of the process to which it converges. For processes that are non-stationary, and do not converge to stationary ones, the concepts of stability studied in the past were mainly tightness and integrability uniformly in time (see e.g. [26, 14, 15, 39] for Markovian setting, and [9, 23] for some drift conditions for processes with marginal probabilities which have negative exponential tails).

The purpose of this paper is to study stability of a general class of non Markovian stochastic processes that include non-stationary and non-ergodic ones. The setting is known as stochastic recursive sequences, and was extensively studied in the stationary context by Brandt, Franken and Lisek [12] and by Borovkov and Foss [7, 8, 10, 20]. The study of stability of such sequences, in the stationary ergodic context, was done using either monotonicity assumptions (see [12, 36]), or renovating events, formulated by Borovkov [7]. The approach of renovating events already has its roots in the early work by Akhmarov and Leont'eva [1]; they were further developed in [2, 10, 20, 22]. The monotonicity and the renovation techniques have been applied successfully to queueing examples, e.g. the G/G/1 queue studied in [36], the G/G/s queue [7], and the generalized Jackson networks, see [4]. (In the latter reference, the existence of a stationary ergodic regime is established using the, so called, "saturation law" by Baccelli and Foss [3, 5]; coupling to it is then obtained using renovation techniques). Other applications of Borovkov's techniques of renovating events to the stability of queueing systems are [2, 20, 22].

We use a setting very similar to the stationary stochastic recursive sequences, which allows us to construct on one probability space stochastic processes that correspond to different initial states and even different control policies. We do not assume any exponential tails or moment conditions (as in [9, 23]). Our criterion of stability is related to coupling, and the tools that we use are those introduced by Borovkov [7] of renovating events. We are interested in the coupling to a "well-behaved" limit regime. A reasonable concept of stability for processes that do not converge to a stationary regime is that processes with the same dynamics but with different initial states, couple after some time which is almost surely finite. Thus, the system "forgets" the initial state after a time that is long enough.

For non-stationary processes occurring in controlled stochastic processes, we are further interested in stability conditions that are uniform in all or in some class of policies. The scrambling conditions introduced in [19], see also [18], the simultaneous Dœblin's conditions introduced by Hordijk [26], the Lyapunov type conditions [25], and the uniform μ-recurrence and uniform μ-ergodicity [14, 15, 39] all imply uniform tightness, and lead moreover to uniform integrability properties (with respect to unbounded cost functions), over all policies, including those for which the state process is non-stationary (see [15] Lemma 4.7). In this paper we investigate conditions for simultaneous coupling by different policies in finite time.

The coupling results serve as tools for studying theoretical problems in control. We are interested in the following question. Suppose we restrict to some subclass of policies, such as the stationary policies, or the periodic ones. Does there exist an optimal policy that minimizes some (expected) average cost? Using coupling arguments, we present sufficient conditions for a positive answer to that question.

2 The model and background

We consider a system in discrete time, $n \in \mathbb{Z}$, governed by the following dynamics:

$$X_{n+1} = f(X_n, \omega_n) \tag{1}$$

where typically, in control applications, $\omega_n = \{r_n, z_n, a_n\}$ and

- X_n is the state of the system at time n, taking values in some measurable space $(\mathbf{X}, \mathcal{X})$,
- $a = (..., a_{-1}, a_0, a_1, a_2, ...)$ is the policy, where a_n have values in some measurable space $(\mathbf{A}, \mathcal{B})$.
- r_n is the driving sequence, taking values in the measurable space $(\mathbf{R}, \mathcal{R})$.
- z_n is another sequence, related to the policy, taking values in the measurable space $(\mathbf{Z}, \mathcal{Z})$.

Let $\mathcal{F}^r_{n,l}$ and $\mathcal{F}^z_{n,l}$ be the σ-field generated by $r_n, ..., r_l$ and $z_n, ..., z_l$ respectively. Let $\mathcal{F}_{n,l}$ be the σ-field generated by $r_n, ..., r_l, z_n, ..., z_l$. Denote $\mathcal{F}^r_l = \mathcal{F}^r_{-\infty,l}$, and $\mathcal{F}^r = \mathcal{F}^r_{-\infty,\infty}$, with similar notation for $\mathcal{F}^z$ and $\mathcal{F}$. We assume that there is given a probability measure P on $\mathcal{F}^r \times \mathcal{F}^z$ for which the driving sequence r is stationary ergodic, and $\{z_n\}$ are i.i.d. and independent of r. Following [10], we call (1) a Controlled Stochastic Recursive Sequence (CSRS). Let $\omega = \{\omega_n\}$, $\Omega = (\mathbf{R} \times \mathbf{Z} \times \mathbf{A})^\infty$. Let $\pi_{r,z}$ be the projection of Ω to $(\mathbf{R} \times \mathbf{Z})^\infty$, let π_r be the projection from Ω to $\mathbf{R}^\infty$, and π_a the projection from Ω to $\mathbf{A}^\infty$. For all measurable sets $B \subset \Omega$ (with $\pi_{r,z} B \in \mathcal{F}$), we define $P(B) := P(\pi_{r,z} B)$, or $P(B) := P(\pi_r B)$ when the z_n sequence is not used.

We shall restrict (unless otherwise stated) to the set of Markov policies $\mathcal{M}$ where a are allowed to depend on the time and on the current state, i.e.

$$a_n = a_n(X_n). \tag{2}$$

We assume that a_n is measurable. We see that the state evolution is now defined simultaneously for all the policies, on the same probability space. We shall also consider the following subclasses of policies:

- $\mathcal{S}$: the stationary policies, where the decision a_n is a function of X_n (and not of the time): $a_n = a(X_n)$;

- $\mathcal{L}(K)$: the periodic policies, where for some K we have $a_{lK+j} = a_j$, for all $l \in \mathbb{Z}$ and for $j = 0, ..., K-1$;
- $\mathcal{M}_{ol}$: the Markov policies that do not depend on the state, i.e. a_n are constants (ol stands for "open loop"); We define $\mathcal{S}_{ol} = \mathcal{S} \cap \mathcal{M}_{ol}$ and $\mathcal{L}_{ol}(K) = \mathcal{L}(K) \cap \mathcal{M}_{ol}$.

Note that $\mathcal{S}_{ol} = \mathcal{L}_{ol}(1)$ and $\mathcal{S} = \mathcal{L}(1)$. The sequence z_n can further be used to define randomized versions of the above sets of policies, by adding to each one of these a dependence on z_n, see Example 2.2 below.

Remark 2.1 Stationary and periodic policies are naturally defined for all n. In many applications, however, policies (and sometimes also the dynamics) are defined on a fixed bounded horizon, e.g. for $n \in \{0, ..., N\}$ (see e.g. Derman [16] Section 2). In other applications, policies may be defined on a horizon that is bounded on one side; most often, for all $n \geq 0$ (see Derman [16] Sections 3, 6). Sometimes, policies are defined only for the negative n. An example for that is Markovian policies that are optimal for finite horizon problems, and are obtained by using successive approximations through backward induction (see Derman [16] Section 2). This approach yields a policy of the form $..., a_{-2}, a_{-1}$, such that for a problem with a fixed horizon of length N the optimal policy is $(a_{-N}, a_{-N+1}, ..., a_{-1})$. For any integer $k \leq N$, the k last components of the policy a are the optimal policy for a horizon of length k. In case that policies are defined only on a subset of $\mathbb{Z}$, we shall assume that there is some way of extending the definition to all $\mathbb{Z}$.

Example 2.1 Admission control to a single queue, with non-randomized policies
Consider admission control into a G/G/1 queueing system. Let σ_n denote the nth service time, and τ_n-the nth inter-arrival time. The sequence (σ_n, τ_n) is assumed to be stationary ergodic; the workload in the system at the instant of the n arriving customer is X_n. The decision $a_n = 1$ means accepting the nth arriving customer, and $a_n = 0$ means rejecting it. We first restrict to deterministic Markov policies. With $r_n := (\sigma_n, \tau_n)$ we have

$$X_{n+1} = \max(X_n + a_n\sigma_n - \tau_n, 0). \tag{3}$$

We shall assume that an initial state $X_0 = v$ is given.

■

Example 2.2 Admission control to a single queue, with randomized policies
Consider the model in Example 2.1; however, we further define z_n to be i.i.d. uniformly distributed on [0,1]. Let a_n take values in [0,1]. The decision a_n means accepting the nth arriving customer with probability a_n. The dynamics is then given by $X_{n+1} = \max(X_n + 1\{z_n \leq a_n\}\sigma_n - \tau_n, 0)$. Note that we could encorporate z_n in the r_n sequence.

■

Remark 2.2 In Example 2.1 and 2.2, the same CSRS could be used to model a situation where the state X_n is the number of customers in the queue at the nth arrival instant, σ_n is the number of customers in the nth batch arrival, τ_n are integer valued inter-arrival times, and service times have a unit duration.

We define the shift operators

$$\begin{aligned} T_r(\ldots, r_{-1}, \underline{r_0}, r_1, \ldots) &= (\ldots, r_{-2}, \underline{r_{-1}}, r_0, \ldots), \\ T_z(\ldots, z_{-1}, \underline{z_0}, z_1, \ldots) &= (\ldots, z_{-2}, \underline{z_{-1}}, z_0, \ldots), \\ T_a(\ldots, a_{-1}, \underline{a_0}, a_1, \ldots) &= (\ldots, a_{-2}, \underline{a_{-1}}, a_0, \ldots). \end{aligned}$$

Define also $T(r, z, a) = (T_r r, T_z z, T_a a)$. Let T^n be the nth iterate of T, and let T^{-1} be the inverse of T; for any function g from Ω to some space, we define

$$Ug(\omega) = g \circ T^{-1}(\omega) = g(T^{-1}\omega).$$

(Note that in general, $U^k X_n(\omega) \neq X_{n+k}(\omega)$). For any subset B of Ω, we denote

$$TB = \{T\omega : \omega \in B\} = \{\omega : T^{-1}\omega \in B\}.$$

When the subscripts r, z or a are added to the above operators, they are understood in the obvious way as related to T_r, T_z and T_a (instead of T).

We assume a fixed initial condition X_0. Following Borovkov [7], we define the following renovating events:

Definition 2.1 *The events $A_n \in \mathcal{F}_{n+L}$, $n \geq n_0 \in \mathbb{Z}$, (for some integer n_0) are called renovating on the intervals $[n, n+L]$ if for any $k > L$, and $\omega \in A_n$, the random variables $X_{n+k} = X_{n+k}(\omega)$ can be expressed in the form*

$$X_{n+k} = \phi_k(\omega_n, \ldots, \omega_{n+k-1}), \tag{4}$$

where $\phi_k : \Omega^k \to \mathbf{X}$ are some given functions (independent of n).

Note that any subset of a renovating event is also a renovating event. The renovating events play a role similar to regeneration events in the Markovian setting. In the controlled Markov chains, we know that stability (in the sense of tightness and integrability of the cost uniformly in time) are obtained uniformly over all policies by some conditions which are uniform over the subclass of pure stationary policies, relating to drift conditions to some "small" sets, see [14, 15, 39]. Under these conditions, the small sets become positive recurrent, and thus have positive probability to occur, infinitely often. In a similar way, for the non Markovian setting, we shall obtain stability under a given policy by the construction of an infinite sequence of renovating events that have positive probabilities (see [7] for the non-controlled case). We present several examples that illustrate the construction of renovating events (in our non-stationary setting).

Example 2.1 (continued):
We shall restrict first to $a \in \mathcal{M}_{ol}$. Assume $X_0 = v$ and let $\xi_n := a_n\sigma_n - \tau_n$,

$$\begin{aligned} A_0^a(v) = \{(\sigma,\tau,\tilde{a}) : \tilde{a} = a, \\ \xi_{-1} \le 0, \xi_{-1}+\xi_{-2} \le 0, ..., \xi_{-1}+...+\xi_{-\nu} \le 0, \xi_{-1}+...+\xi_{-\nu}+...+\xi_{-\nu-k} \le -v, \forall k \ge 1\}, \end{aligned} \tag{5}$$

for some integer $\nu \ge 0$, and let $A_n^a(v) = T^n A_0^a(v)$. Note that $A_0^a(v)$ depends also on ν, from the context it should be clear which ν is used. We recall that, since there is no z_n sequence, $P(A_0^a(v)) := P(\pi_r A_0^a(v))$. Then

$$\begin{aligned} \pi_r A_n^a(v) = \{(\sigma,\tau) :\ \xi_{n-1} \le 0, \xi_{n-1}+\xi_{n-2} \le 0, ..., \\ \xi_{n-1}+...+\xi_{n-\nu} \le 0, ..., \xi_{n-1}+...+\xi_{n-\nu}+...+\xi_{n-\nu-k} \le -v, \forall k \ge 1\}. \end{aligned}$$

Then for any choice of v in (5) and any policy a, $A_n^a(v)$ are renovating events for $n \ge \nu+1$. Indeed, on $A_n^a(v)$ we have $X_n = 0$ so that $\phi_0 = 0$ in (4) (L can be chosen to be zero). Borovkov and Foss considered in [10] the non-controlled SRS, which is the case when we use the policy $\overline{a}$ for which $a_n = 1$ for all n, and for which $A_n^{\overline{a}}(v)$ forms a stationary ergodic sequence of sets. They showed that if $E\sigma_0 < E\tau_0$, then ν can be chosen in (5) such that $P(A_0^{\overline{a}}(v)) > 0$. The proof of the last statement is not detailed in [10], and we present it below.

X_{n+1} satisfies,

$$X_{n+1} = \max\left(\xi_n, \xi_n+\xi_{n-1}, \xi_n+...+\xi_1, \xi_n+...+\xi_0+v, 0\right). \tag{6}$$

We show in the appendix that
Observation: Assume that $E\sigma_0 < E\tau_0$. If $P(A_0^{\overline{a}}(v)) = 0$ for all $\nu > 0$, then under $\overline{a}$, $\lim_{m\to\infty} P(X_n > 0, \forall n \ge m) = 1$.

Assume that $P(A_0^{\overline{a}}(v)) = 0$. Using Foss [21], we show that this contradicts the fact that $E\sigma_0 < E\tau_0$. Indeed, the above observation, together with (6), imply that there exists some integer N such that $P(\Pi_N) > 0$ where $\Pi_N \stackrel{\triangle}{=} \{\omega : X_n(\omega) > 0, \forall n \ge N\}$. On Π_N we have

$$X_{n+1} = X_N + \sum_{i=N}^{n} \xi_i \qquad \text{for all } n \ge N. \tag{7}$$

Since $E\sigma_0 < E\tau_0$ then we have by Birkhoff's ergodic Theorem, and from (7) that $\lim_{n\to\infty} \sum_{i=1}^{N} \xi_i = -\infty$, $P-a.s.$, so that on the set Π_N, $\lim_{n\to\infty} X_n = -\infty$. But this contradicts the fact that $X_n > 0, n \ge N$ on the set Π_N. We conclude that $P(A_0^{\overline{a}}) > 0$ for some choice of ν.

For an arbitrary policy $a \in \mathcal{M}_{ol}$, the sets $A_n^a(v)$ need not be stationary. However, we have for any a, $\pi_r A_n^{\overline{a}}(v) \subset \pi_r A_n^a(v)$, so we get

$$P(A_n^a(v)) \ge P(A_n^{\overline{a}}(v)) = P(A_0^{\overline{a}}(v)) > 0 \tag{8}$$

if $E\sigma_0 < E\tau_0$.

Finally, for $a \in \mathcal{M}/\mathcal{M}_{ol}$, we may take $A_0^a(v) := \pi_r A_0^{\overline{a}}(v) \times \{a\}$, with $A_n^a(v) := T^n A_0^a(v)$; $A_n^a(v)$ are still renovating events for $n \geq \nu + 1$, since on $A_n^a(v)$ we have $X_n = 0$ so that $\phi_0 = 0$ in (4). The discussion above then shows that if $E\sigma_0 < E\tau_0$, then $P(A_0^a(v)) > 0$. ∎

In the above example, the probability of the renovating events $A_n^a(v)$ are positive for all policies. Moreover, the sequence $A_n^{\overline{a}}(v)$ is stationary with respect to the probability P. There are cases where only under some policies one may construct renovating events A_n^a which have positive probability. We illustrate this in the following examples.

Example 2.3 Admission control policies with constraint
Consider the model in Example 2.1. Assume that $E\sigma_0 > E\tau_0$. It then follows that under the policy $\overline{a}$, which always accepts arriving customers, the queue grows to infinity (see [36]). The probability of the renovating event $A_0^{\overline{a}}$ (we omit in the notation the dependence on the starting state v) in (5) is zero. Define $\rho = E\sigma_0/E\tau_0$. Fix some positive integers K_1 and K_2, such that $\rho K_1 < K_2$, and denote $K = (K_1, K_2)$. We consider a subclass $\mathcal{A}^K$ of policies that have the following constraint: no more than K_1 customers are allowed to be admitted among any K_2 consecutive arriving customers. A policy in $\mathcal{A}^K$ is called feasible. Note that $\mathcal{A}^K$ includes non-stationary and non-periodic policies. We assume below that (σ_n, τ_n) are stationary ergodic with respect to both the unit shift as well as the K_2-step shift. We show in the Appendix that $P(\overline{\mathsf{A}}_0^a) > 0$ for any feasible a with period K_2, where the sets $\overline{\mathsf{A}}_0^a(v) \stackrel{\triangle}{=} \cup_{i=0}^{K-1} T^i A_0^a(v, \nu + i)$ are renovating ($A_0^a(v, \nu)$ are the sets defined in (5), note that here we write the dependence on ν explicitly). (On the set $\overline{\mathsf{A}}_0^a(v)$, $X_j = 0$ for some $j = 0, ..., K-1$.)

∎

$\overline{a}$ in Example 2.1 served to construct stationary minorant sets for the renovating events of other policies. These policies are, however, degenerate, in the sense that they do not depend on the state. Next we show that minorizing renovating sets can be constructed by state dependent policies.

Example 2.4 Renovating events for threshold policies
Consider again the setting of Example 2.1, with $E\sigma_0 > E\tau_0$. Consider the policy a^* such that a_n^* accepts the arriving customer if and only if $X_n \leq \overline{v}$, for some constant $\overline{v}$. Consider the class $\mathcal{A}^{\overline{v}} \subset \mathcal{M}$ of policies a for which a_n accepts the arriving customer if and only if $X_n \leq v^a(n)$, where $v^a(n) \leq \overline{v}$ is some sequence of real numbers, $n = ..., -1, 0, 1, 2, ...$ Assume that there exists some $\Delta > 0$ such that

$$P(\sigma_n + \Delta < \tau_n | \mathcal{F}_n^r) > 0, \qquad P - a.s.$$

for all n. Let $\overline{L}$ be some integer s.t. $\overline{L} > \overline{v}/\Delta$ and assume that the initial state is $v = 0$. Define

$$C_0 := \tag{9}$$
$$\{\sigma_{-1} + \Delta < \tau_{-1}, \sigma_{-1} + \sigma_{-2} + 2\Delta < \tau_{-1} + \tau_{-2}, ..., \sigma_{-1} + ... + \sigma_{-\overline{L}} + \overline{L}\Delta < \tau_{-1} + ... + \tau_{-\overline{L}}\}$$

and let $C_n = T_r^n C_0$. Then C_n is stationary, with $P(C_0) > 0$. For $a \in \mathcal{A}^{\overline{v}}$, define $A_0^a := C_0 \times \{a\}$. Then $A_n^a = T^n A_0^a$ are renovating events and have positive probabilities.

■

3 The stationary case

Before analyzing the general non-stationary setting, we begin by citing the results for the stationary one, for which the interpretation of the results (such as the process to which we have coupling) are simpler. We then discuss the interpretation of corresponding results for the the non-stationary case, before formalizing them explicitly in the following sections.

Borovkov established and proved the following result in [7] p. 260-262 for non-controlled SRS (i.e. for which $a = a_n$ is a constant for all $n \in \mathbb{Z}$) and for ergodic stationary driving sequence:

Theorem 3.1 *(Borovkov) Let A_n^a be a sequence of renovating events w.r.t. policy a, such that*

$$\lim_{n\to\infty} P\left(\bigcap_{l=l_0}^{\infty}\bigcup_{j=n_0}^{n} A_j^a T^{-l} A_{j+l}^a\right) = 1 \tag{10}$$

for any $n_0 > 0$ and $l_0 > 0$. Then there is a sequence X^k of random variables such that $X^k = U^k X^0$ satisfies (1), so that

$$\lim_{n\to\infty} P\left(U^{-n-k+1} X_{n+k+s} = X^{s+1} \ for\ all\ s \geq 0, k \geq 0\right) = 1. \tag{11}$$

In particular,

$$\lim_{n\to\infty} P\left(X_{n+s} = X^{n+s} \ for\ all\ s \geq 0\right) = 1 \tag{12}$$

and

$$\lim_{n\to\infty} P\left(U^{-n-k} X_{n+k} = X^0 \ for\ all\ k \geq 0\right) = 1. \tag{13}$$

Finally, if $A_n = T^n A_0$ then a sufficient and necessary condition for (10) is that $P(A_0) > 0$.

For a recent relaxation of the condition (10) see [11]. (In the above theorem and throughout the paper, we use Borovkov's notation $AB := A \cap B$ for sets). Equation (12) states that the forward coupling time

$$N^0(\omega) = \min_{n\geq 0}\{X_{n+l}(\omega) = X^{n+l}(\omega), \forall l \in \mathbb{N}_0\}$$

is finite P-a.s. ($\mathbb{N}_0$ is the set of nonnegative integers, including 0). Equation (13) implies that the backward coupling time

$$N_0(\omega) = \min_{k\geq 0}\{X_{l+k}(T^k\omega) = X^l(\omega), \forall l \in \mathbb{N}_0\} \tag{14}$$

is finite P-a.s. In order to further understand the above statements (and their future generalizations to the controlled case) we follow [10] and define the sequence of SRS $X_n^{[k]}$, $k = 0, 1, 2, ...$, by

$$X_n^{[k]}(\omega) = X_{n+k}(T^k\omega) = U^{-k}X_{n+k}(\omega), \; n \geq -k.$$

$X_n^{[k]}$ are thus determined by the initial value $X_{-k}^{[k]} = X_0$ at time $-k$, and evolve according to the dynamics

$$X_{n+1}^{[k]} = f(X_n^{[k]}, \omega_n) \tag{15}$$

for $n \geq -k$. With this interpretation, (14) can be rewritten as

$$N_0(\omega) = \min_{k\geq 0}\{X_l^{[k]} = X^l, \forall l \in \mathbb{N}_0\}. \tag{16}$$

Define the following coupling times for $k, s \in \mathbb{N}_0$:

$$\begin{aligned} N^k(\omega) \;&:=\; \min_{n\geq -k}\{X_{n+l}^{[k]}(\omega) = X^{n+l}(\omega), \forall l \in \mathbb{N}_0\} \qquad (17) \\ (\;&=\; \min_{n\geq -k}\{X_{n+k+l}(T^k\omega) = X^{n+l}(\omega), \forall l \in \mathbb{N}_0\}) \end{aligned}$$

$$N_s(\omega) \;:=\; \min_{k\geq 0}\{X_{s+l}^{[k]}(\omega) = X^{s+l}(\omega), \forall l \in \mathbb{N}_0\} \tag{18}$$

$$N^*(\omega) := \sup_{k\geq 0} N^k(\omega). \tag{19}$$

$$N_\#(\omega) := \sup_{s\geq 0} N_s(\omega). \tag{20}$$

Note that $N_s(\omega)$ are monotone non-increasing in s. Hence $N_\#(\omega) = N_0(\omega)$. In the controlled case, we shall add sometimes explicitly the policy and/or the initial state to the notation in the above coupling times. N^* was called in [10] the strong coupling time; it is shown in [10] that N^* is P-a.s. finite under suitable conditions, for the stationary ergodic case.

In the following sections we shall extend Theorem 3.1 to nonstationary policies, and to cases where P is not stationary ergodic; the statement of the theorem will be shown to hold under suitable conditions, even then. An obvious difference between the stationary ergodic framework and our framework is that in the stationary ergodic (noncontrolled) setting, the process X^n to which the coupling occurs, is a stationary process, and the process X_n couples to the stationary process in finite time, P-a.s.

In the non-stationary setting, the "limit" sequence X^n may not always have an obvious interpretation. Since it is constructed by a backward procedure, its interpretation is simpler for policies that are defined for negative times. For that case, fix some n. As k grows, the decision rule a_l (which are part of ω_l) used at any time $l \leq n$ in (15) are unchanged. In other words, a_l can be seen to be a function of the time till n, instead of the time from the initial operating time. Since the initial time decreases, this means that the time at which a_l is used relative to the initial time, grows with k. (15) thus seems as if it is driven by policies that are updated "backward" in time as k grows. This phenomena of going backwards in time is not new in control; we know in the framework of controlled Markov chains that a Markov optimal policy $a(k)$ for a problem with finite horizon k, is also growing "backward" in time as k grows; the optimal decision rule at any time is a function of the remaining time to go, independently of k, rather than of the time from the beginning (see Remark 2.1).

Assume now that a policy, or even ω_n, is defined only for non-negative n, and we extend it in some way to negative n. The process X^n may then depend in what way we extend ω to the negative n's. This of course motivates us to examine the question of whether the limit process is insensitive to the way we extend ω_n. This will be done in Theorem 5.3.

In the original Theorem by Borovkov [7] for the non-controlled case, the process X^n to which the coupling occurs, is a stationary process. The above theorem thus implies in particular that the process X_n couples to the stationary process in finite time, P-a.s. The following converse of Theorem 3.1 also holds, for the uncontrolled case. Let X^0 be such that (13) is satisfied. Then there are renovating events $A_n = T^n A_0$ with $P(A_0) > 0$. In the following counter example, we show that this converse does not hold for (12): (12) may hold, and yet there may exist no sequence of renovating events with positive probabilities, satisfying (10).

Counter-example 3.1 Non-existence of "good" renovating events
Let $\mathbf{X} = \{-1, 0, 1\}$, and let $X_0 = 0$. Consider $\mathbf{R} = \{-1, +1\}$ and assume that r_n are i.i.d. with $P(r_1 = 1) = P(r_n = -1) = 0.5$. For any $r' \in R$, let $f(0, r') = r'$, and $f(x, r') = -x, x \neq 0$. Then X_n couples at time 1 with an ergodic stationary process X^n whose samples are either $X^n = (-1)^n$ or $X^n = (-1)^{n+1}$ for all n (with equal probability).

Assume that A_n is a renovating event on the interval $[n, n+L]$ for some n and L, so that for all $k > L$, (4) holds. Since X_{n+k} does not depend on $r_l, l \geq 2$, it follows that ϕ_k is a function of k only. If $|\phi_k| \neq 1$ then clearly $P(A_n) = 0$. So, for $P(A_n)$ to be positive, we should have $|\phi_k| = 1$, and $\phi_{k+1} = -\phi_k$, $k > L$. Thus, assume (without loss of generality) that $\phi_k = 1$. Since $X_{n+k} = 1$ if and only if $r_1 = 1$, it follows that $A_n = \{r_1 = 1\}$.

Since this type of argument holds for any n, we have for all n either $A_n = \{r_1 = 1\}$ or $A_n = \{r_1 = -1\}$ (otherwise $P(A_n) = 0$). Hence, for any integers j and l, $A_j T^{-l} A_{j+l} = \{r_1 = r_l\}$. Since r_l are i.i.d., we conclude that the left hand side of (10) is equal to zero (due to the infinite intersection).

∎

4 Coupling results for the general non-stationary case

In this section we develop the general tools for establishing coupling to non-stationary processes. The result will hold for a general Ω space, and not just to the specific case in which $\omega_n = (r_n, z_n, a_n)$, which are suitable for the control applications. In the next section we shall go back to the precise construction of ω in order to present sufficient conditions for coupling and uniform coupling.

In order to establish the results we derive first some Lemmas. To that end, define for $n \in \mathbb{N}_0$ the sequence of sets

$$B_n := \{\omega | X_n(T^n\omega) = X_{n+k}(T^{n+k}\omega), \forall k \in \mathbb{N}_0\}.$$

(i.e. $B_n = \{\omega | X_0^{[n]} = X_0^{[n+k]}, \forall k \in \mathbb{N}_0\}$). Note that B_n may be empty, as well as any of the sets that we shall define below. Let

$$B := \bigcup_{n=0}^{\infty} B_n.$$

Clearly on B, the limit

$$X^0(\omega) := \lim_{n\to\infty} X_n(T^n\omega) = \lim_{n\to\infty} X_0^{[n]}(\omega) \tag{21}$$

exists. For $\omega \in T^s B$, denote

$$X^s(\omega) := X^0(T^{-s}\omega) \tag{22}$$

so that

$$X^s(\omega) = \lim_{n\to\infty} X_n(T^{n-s}\omega) = \lim_{n\to\infty} X_{n+s}(T^n\omega) = \lim_{n\to\infty} X_s^{[n]}(\omega).$$

Let

$$B^* := \bigcap_{s=0}^{\infty} T^s B.$$

Then

$$X^{s+1} = f(X^s(\omega), \omega_s), \qquad \omega \in B^*, \forall s \in \mathbb{N}_0. \tag{23}$$

To get relation (23), let $\omega \in B^*$, choose n sufficiently large so that

$$X^{s+1}(\omega) = X_{n+1}(T^{(n+1)-(s+1)}\omega)$$

and

$$X^s(\omega) = X_n(T^{n-s}\omega).$$

Then

$$X^{s+1}(\omega) = X_{n+1}(T^{n-s}\omega) = f(X_n(T^{n-s}\omega), (T^{n-s}\omega)_n) = f(X^s(\omega), \omega_s),$$

which establishes (23). Define

$$B_n^k := \{\omega \in T^n B^* | X_{n+k+l}(T^k\omega) = X^{n+l}(\omega),\ \forall l \in \mathbb{N}_0\}$$

(or, $B_n^k = \{\omega \in T^n B^* | X_{n+l}^{[k]}(\omega) = X^{n+l}(\omega),\ \forall l \in \mathbb{N}_0\}$). We will use the convention that the union of an empty collection of sets is the empty set.

Lemma 4.1 *Assume that*

$$A_n = T^n A_0,\ n \in \mathbb{Z}$$

are renovating on the intervals $[n, n+L]$ *for some* $L \in \mathbb{N}_0$. *Then*

$$\bigcup_{l=L+1}^{n} A_{-l} \subset B_n \tag{24}$$

$$\bigcup_{l=L+1}^{\infty} A_{-l} \subset B^* \tag{25}$$

$$\bigcup_{l=L+1-n}^{k} A_{-l} \subset B_n^k, \qquad \forall k, n \in \mathbb{N}_0 \tag{26}$$

Proof. Relation (24) is adapted from the proof of Theorem 1 Sec. 4.6 in Borovkov [7]: If $\omega \in A_n$ then by Definition 2.1,

$$X_{n+k}(\omega) = \phi_k(\omega_n, ..., \omega_{n+k-1}), \forall k > L.$$

Since $T^s\omega \in A_{n+s}$ for all $s \in \mathbb{N}_0$, we also have that

$$X_{n+s+k}(T^s\omega) = \phi_k((T^s\omega)_{n+s}, ..., (T^s\omega)_{n+s+k-1}) = \phi_k(\omega_n, ..., \omega_{n+k-1}),\ \forall k > L.$$

Consequently,

$$\omega \in A_n \ \Rightarrow\ X_{n+k}(\omega) = X_{n+k+s}(T^s\omega) \qquad \forall s \in \mathbb{N}_0, \forall k > L \tag{27}$$

(27) implies for $T^n\omega \in A_{n-l}$ (and hence for $\omega \in A_{-l}$) with $L < l \leq n$,

$$X_{n-l+k'}(T^n\omega) = X_{n-l+k'+s}(T^{n+s}\omega) \qquad \forall s \in \mathbb{N}_0, \forall k' > L.$$

With $k := k' - l$, we get

$$X_{n+k}(T^n\omega) = X_{n+k+s}(T^{n+s}\omega) \qquad \forall s, k \in \mathbb{N}_0.$$

Therefore,

$$A_{-l} \subset B_n \quad \text{ for all } L < l \leq n$$

from which (24) follows.

It now follows from (24) that

$$\bigcup_{l=L+1}^{\infty} A_{-l} \subset \bigcup_{n=0}^{\infty} B_n = B \tag{28}$$

and

$$\bigcup_{l=L+1}^{\infty} A_{-l} \subset \bigcup_{l=L+1-n}^{\infty} A_{-l} = \bigcup_{l=L+1}^{\infty} T^n A_{-l} \subset T^n B.$$

Hence,

$$\bigcup_{l=L+1}^{\infty} A_{-l} \subset \bigcap_{n=0}^{\infty} T^n B = B^*, \tag{29}$$

which is assertion (25).

From (25) and (23), it follows that

$$\left(\bigcup_{l=L+1}^{\infty} A_{-l} \right) \cap B_n^k = \left\{ \omega \in \bigcup_{l=L+1}^{\infty} A_{-l} | X_{n+k}(T^k\omega) = X^n(\omega) \right\}.$$

If $\omega \in T^n \left(B_{n+k} \cap \bigcup_{l=L+1}^{\infty} A_{-l} \right)$, then

$$X_{n+k}(T^k\omega) = X_{n+k}(T^{n+k}T^{-n}\omega) = X^0(T^{-n}\omega) = X^n(\omega)$$

and $\omega \in B_n^k$. Hence with (24),

$$\bigcup_{l=L+1-n}^{k} A_{-l} = T^n \left(\bigcup_{l=L+1}^{n+k} A_{-l} \right) \subset B_n^k$$

which is relation (26). ∎

The following Lemma relates the renovating events to the coupling times defined in (17) and (19) (for the general non-stationary case).

Lemma 4.2 *Let A_n be as in Lemma 4.1. The following inclusions hold:*

$$\bigcup_{l=L+1-n}^{k} A_{-l} \subset \{N^k \leq n\}, \tag{30}$$

$$\bigcup_{l=0}^{n-L-1} A_l \subset \{N^* \leq n\}. \tag{31}$$

Proof. Assume that $\omega \in \bigcup_{l=L+1-n}^{k} A_{-l}$.Then by Lemma 4.1 (26), $\omega \in B_n^k$, so that

$$X_{n+k+l}(T^k\omega) = X^{n+l}(\omega) \qquad \forall l \in \mathbb{N}_0.$$

Consequently, $\omega \in \{N^k \leq n\}$, which establishes assertion (30). Assertion (31) follows directly from (30) by taking the intersection over $k \geq 0$. ∎

Lemma 4.3 *Let A_n be as in Lemma 4.1. The following inclusion holds:*

$$\bigcup_{l=L+1}^{k} A_{-l} \subset \{N_{\#} \leq k\} \tag{32}$$

Proof. Assume that $\omega \in \bigcup_{l=L+1}^{k} A_{-l}$. Then $\omega \in B_0^k$, so that

$$X_{k+s+l}(T^k\omega) = X^{s+l}(\omega), \qquad \forall s, l \in \mathbb{N}_0.$$

Hence, $\omega \in \{N_\# \leq k\}$. This establishes the proof. ∎

Lemmas 4.2 and 4.3 imply

Corollary 4.1 *Fix a policy a, let $A_0^a = \{a\} \times A_0$ with A_0 a subset of $(\mathbf{R} \times \mathbf{Z})^\infty$ and assume that $A_n^a = T^n A_0^a$, $n \in \mathbb{Z}$ are renovating on the intervals $[n, n+L]$ for some $L \in \mathbb{N}_0$. Assume that*

$$\lim_{k \to \infty} P\left(\bigcup_{l=L+1}^{k} A_{-l}^{a}\right) = 1 \tag{33}$$

and

$$\lim_{n \to \infty} P\left(\bigcup_{l=0}^{n-L-1} A_l^a\right) = 1. \tag{34}$$

Then (i) N^* and $N_\#$ are finite P-a.s. (ii) If, moreover, (33) and (34) hold uniformly over some subclass $\overline{\mathcal{A}} \subset \mathcal{A}$, then coupling is obtained simultaneously in all policies in finite time, i.e. $\sup_{a \in \overline{\mathcal{A}}} N^*(a)$ and $\sup_{a \in \overline{\mathcal{A}}} N_\#(a)$ are finite P-a.s.

Proof. Obvious. ∎

In typical applications, as the ones in Examples 2.1 and 2.2, the state process corresponding to some non-stationary policy a is bounded by the state process corresponding to some stationary policies. (In Example 2.1, for any policy a, the stationary policy that always rejects customers produces a lower bound, and the one that always accepts customers produces an upper bound). Therefore, if we have also coupling results for the bounding stationary policies, this implies that the sequence X^n corresponding to a is tight. Often, the coupling times corresponding to some non-stationary policy can also be shown to be bounded by ones given by a stationary one.

5 Sufficient conditions for coupling

The conditions in Corollary 4.1 may seem difficult to check. However, motivated by (8) in Example 2.1, we have the following sufficient conditions for coupling.

Theorem 5.1 *Consider some class of policies $\overline{\mathcal{A}}$. Assume that for any policy $a \in \overline{\mathcal{A}}$, (i) $A_n^a = T^n A_0^a$, $n \in \mathbb{Z}$ are renovating on the intervals $[n, n+L]$ for some $L \in \mathbb{N}_0$, (ii) there exists some sequence of sets C_n such that $C_n \subset \pi_r A_n^a$ for all $a \in \overline{\mathcal{A}}$, $P(C_0) > 0$, and C_n are stationary ergodic under P.*
Then $\sup_{a \in \overline{\mathcal{A}}} N^*(a)$ and $\sup_{a \in \overline{\mathcal{A}}} N_\#(a)$ are finite P-a.s.

Proof. By the same argument as in Borovkov [7] p. 261, it follows from the ergodicity of the sequence C_n that

$$\lim_{k\to\infty} P\left(\bigcup_{l=L+1}^{k} C_{-l}\right) = 1 \qquad \text{and} \qquad \lim_{n\to\infty} P\left(\bigcup_{l=0}^{n-L-1} C_l\right) = 1.$$

This implies (33) and (34), so that the Theorem now follows from Corollary 4.1. ∎

Applying the above theorem to our examples, we see that the conclusion holds for our examples; in Example 2.1 (continued), we obtain simultaneous coupling in all policies within finite time, P-a.s. In Example 2.3 this coupling is obtained simultaneously over $\mathcal{A}^K$, and in Example 2.4 it is obtained uniformly over $\mathcal{A}^{\overline{v}}$. The sets C_n can be chosen as $T_r^n C_0$. This is also the case in Example 2.1 (continued) where we may choose $C_n = \pi_r A_n^{\overline{a}}(v)$.

Next we consider coupling uniformly in the initial state. Let $N^*(v)$ and $N_\#(v)$ denote the coupling times defined in (19) and (20) for initial state $X_0 = v$.

Theorem 5.2 *Fix some policy a.*
(i) Assume that for some set $V \subset \mathbf{X}$ of initial states, $A_n(v) = T^n A_0(v)$, $n \in \mathbb{Z}$ are renovating on the intervals $[n, n+L]$ for some $L \in \mathbb{N}_0$; there exists some sequence of sets C_n such that for all $v \in V$, $C_n \subset \pi_r A_n(v)$, $P(C_0) > 0$, and C_n are stationary ergodic under P. Then

$$\sup_{v\in V} N^*(v) \qquad \textit{and} \qquad \sup_{v\in V} N_\#(v) \quad \textit{are finite } P-a.s. \tag{35}$$

Moreover, the process X^n is the same for any initial state $X_0 = v \in V$.
(ii) Assume that there exist increasing sets of states $\{V_k\}$ with $\bigcup_{k=1}^\infty V_k = \mathbf{X}$, such that for any k, there exists a sequence of sets $C_n^k = T_r^n C_0^k$ which are stationary ergodic for each k, and satisfy $C_n^k \subset \pi_r A_n(v), \forall v \in V_k$, $P(C_0^k) > 0$. Then (35) holds for all V_k, and the process X^n does not depend on the initial value of the initial state $X_0 = v$.

Proof. The proof follows from arguments similar to the proof of Theorem 5.1, and are straightforward extensions of the proof for the non-controlled case, see [10] p. 22. ∎

Applying Theorem 5.2 to Example 2.1 we conclude that the limit process X^s to which coupling occurs does not depend on the initial state. The sets $C_n^k = T_r^n C_0^k$ can be taken as $C_0^k = \pi_r A_n^{\overline{a}}(v_k)$, where v_k is some increasing sequence of real numbers, converging to infinity.

Using the same arguments, we now consider the case where the policy is defined only for positive n.

Theorem 5.3 *Fix some initial state v. Fix some policy a, defined for positive n, and assume that we extend it to negative n's by concatenating it with some policy in some subclass of policies $\overline{\mathcal{A}}$.*

Assume that there exists some sequence of sets C_n such that for all $a' \subset \overline{\mathcal{A}}$, $A_n^{a',a} = T^n A_0^{a',a}$, $n \in \mathbb{Z}$ are renovating on the intervals $[n, n+L]$ for some $L \in \mathbb{N}_0$; $C_n \subset \pi_r A_n^{a',a}$, $P(C_0) > 0$, and C_n are stationary ergodic under P. Then

$$\sup_{a' \in \overline{\mathcal{A}}} N^*(a', a) \qquad \textit{and} \qquad \sup_{a' \in \overline{\mathcal{A}}} N_\#(a', a) \qquad \textit{are finite } P-a.s. \tag{36}$$

Note the following evident property, due to the dynamics (1). Fix some policy $a(0)$, defined for positive n, and some initial state v. Consider the policy $a = (a(1), a(0))$, for some arbitrary policy $a(1)$ defined for negative times (i.e. we consider the policy a that is obtained by concatenating $a(1)$ with policy $a(0)$). Then, for the same initial state and for any other policy $a(2)$, defined for negative times, the state process $X_n, n \geq 0$ corresponding to the concatenated policy $a' = (a(2), a(0))$ is the same as the one corresponding to a (pathwize). Therefore, if for a fixed initial state at time 0, there is (forward) coupling under some policy to some other process (e.g. to the one obtained by using the same policy but starting from a different initial state), then the time to coupling will not depend on the policy used before time 0. Thus, in most applications where one is interested in forward coupling results, policies need not be defined for negative times.

6 Periodic policies

We consider in this section a special case of “non-stationarity", where ω_n (and in particular, the policies) are periodic in some stochastic sense, with period K. (This will include, in particular, periodic policies). Denote $\mathbf{X}^{[K]} = \mathbf{X}^K$. Define the sequences of vectors $\underline{X}_n^{[K]} = (X_{nK}, ..., X_{(n+1)K-1})$, and $\underline{\omega}_n^{[K]} = (\omega_{nK}, ..., \omega_{(n+1)K-1})$. Let $\Omega^{[K]}$ be the set of sample paths of the form

$$(..., \underline{\omega}_{-n}^{[K]}, \underline{\omega}_{-n+1}^{[K]}, ..., \underline{\omega}_{-1}^{[K]}, \underline{\omega}_0^{[K]}, \underline{\omega}_1^{[K]}, ...).$$

Define the map $\Psi_\omega : \Omega \to \Omega^{[K]}$ by

$$\Psi_\omega(..., \omega_{-j}, \omega_{-j+1}, ..., \omega_{-1}, \omega_0, \omega_1, ...) = (..., \underline{\omega}_{-n}^{[K]}, \underline{\omega}_{-n+1}^{[K]}, ..., \underline{\omega}_{-1}^{[K]}, \underline{\omega}_0^{[K]}, \underline{\omega}_1^{[K]}, ...)$$

Define for a set $A \subset \Omega$, $\Psi_\omega A = \{\Psi_\omega(\omega) | \omega \in A\}$. Clearly, there exists some function $f^{[K]}$ such that (1) implies

$$\underline{X}_{n+1}^{[K]} = f^{[K]}(\underline{X}_n^{[K]}, \underline{\omega}_n^{[K]}), \qquad n \in \mathbb{Z}. \tag{37}$$

so that $\underline{X}^{[K]}$ is again a CSRS. Assume now that A_{nK}, $n \in \mathbb{N}_0$ are a sequence of renovating events on the intervals $[nK, (n+L)K]$. It is easily seen that the sets $\underline{A}_n := \Psi_\omega A_{nK}$ are renovating for the new CSRS (37). Assume that P is stationary ergodic under the K-step shift $T^{[K]} := T^K$, and that the new CSRS satisfies (10) of Theorem 3.1 (or that $P(A_0) > 0$, and $A_n = T^n A_0$). Then the results of Theorem 3.1 clearly hold for the new CSRS. Going back to the initial probability space, we then conclude in particular, that there are K sequences $\{X^{nK+j}\}_{n\in\mathbf{Z}}$, $j = 0, 1, ..., K-1$, which are all stationary ergodic under the shift T^K, so that

$$\lim_{n\to\infty} P(X_{(n+l)K+j} = X^{(n+l)K+j} \text{ for all } l \in \mathbb{N}_0 \text{ and } j = 0, 1, ..., K-1) = 1.$$

Note that assuming that P is stationary ergodic under T does not imply that it is stationary ergodic under T^K. Indeed, consider the sequence $\omega_n = \omega_0(-1)^n$, where $\omega_0 = 1$ w.p. 0.5 and $\omega_0 = -1$ w.p. 0.5. This sequence is stationary ergodic under the shift T, but not under T^2 (see discussion in [10] p. 22 Ex. 3).

7 Applications to control: optimality results

The framework that we developed in previous sections enables us to get new results on the relatively new area of control of non-Markovian stochastic processes. Non-Markovian control problems can be transformed into Markovian ones, by considering the whole history till time t as the state at time t. Optimal policies can then often be computed using dynamic programming (DP) techniques; however, these will typically depend on the whole history, since DP yields in general state dependent optimal policies, and in the above approach, the state correspond to the current history. This approach is thus not feasible from the computational point of view, and it is natural to restrict to simpler classes of policies. It is well known that even Markovian control problems may suffer from the above explosion of the state space, and may require using the whole histories as states, e.g. in case that the controller possesses only partial information (see e.g. Bertsekas [6]). The restriction to simpler classes of policies, especially the periodic ones $\mathcal{L}(K)$, and $\mathcal{L}_{ol}(K)$, are then a standard way of approaching such problems, see [28, 30, 32, 33].

We assume in this Section that $\mathbf{X}$ is a Borel space (a Borel subset of a complete separable metric space) and that $\mathbf{A}$ is a compact Borel space, with $\mathcal{X}$ resp. $\mathcal{B}$ the Borel σ-fields. We further assume that f (appearing in (1)) is continuous in its arguments. Introduce some bounded immediate cost : $c : \mathbf{X} \times \mathbf{A} \to \mathbb{R}$, and introduce the following costs, which are functions of the initial state v and the policy a:

$$C^T(v, a) = \frac{1}{T+1} \sum_{n=0}^{T} c(X_n, Y_n), \qquad J^T(v, a) = EC^T(v, a),$$

where $Y_n = a_n(X_n)$ is the action chosen at time n. Let

$$C(v, a) = \overline{\lim_{T\to\infty}}\, C^T(v, a), \qquad \text{and} \qquad J(v, a) = \overline{\lim_{T\to\infty}}\, J^T(v, a).$$

We assume that the immediate cost c is continuous in its arguments.

For any subclass of policies $\overline{\mathcal{A}}$ and a fixed initial state v, consider the following problems:

- $\mathbf{\Pi_1}(\overline{\mathcal{A}})$: find a policy $\overline{a} \in \overline{\mathcal{A}}$ that minimizes $J^T(v,a)$ over all $a \in \overline{\mathcal{A}}$; $\overline{a}$ is thus optimal if $J^T(v,\overline{a}) \leq J^T(v,a)\ \forall a \in \overline{\mathcal{A}}$.

- $\mathbf{\Pi_2}(\overline{\mathcal{A}})$: find a policy $\overline{a} \in \overline{\mathcal{A}}$ that minimizes $C(v,a)$ over all $a \in \overline{\mathcal{A}}$ in the stochastic sense; $\overline{a}$ is thus optimal if $C(v,\overline{a}) \leq_{st} C(v,a)\ \forall a \in \overline{\mathcal{A}}$.

- $\mathbf{\Pi_3}(\overline{\mathcal{A}})$: find a policy $\overline{a} \in \overline{\mathcal{A}}$ that minimizes $J(v,a)$ over all $a \in \overline{\mathcal{A}}$; $\overline{a}$ is thus optimal if $J(v,\overline{a}) \leq J(v,a)\ \forall a \in \overline{\mathcal{A}}$.

In all the above problems, the values of a_n for negative n do not influence the cost. In $\mathbf{\Pi_1}$, the values for $N > T$ do not influence the cost. Hence, we may restrict the definition of the policies to the interval where they are relevant to the cost (and extrapolate them in an arbitrary way outside that interval). We call this interval the effective time domain of the policy.

We begin by stating some continuity properties of the cost. We say that a sequence of policies $a^k \in \mathcal{M}_{ol}$ converges to $a \in \mathcal{M}_{ol}$ if for all n in the effective time domain of the policy,

$$\lim_{k\to\infty} a_n^k = a_n$$

We say that a sequence of policies $a^k \in \mathcal{M}$ converges to $a \in \mathcal{M}$ if for all n in the effective time domain of the policy, and for all $x \in \mathbf{X}$,

$$\lim_{k\to\infty} a_n^k(x) = a_n(x)$$

A function g over the policies is said to be continuous in a if for any sequence a^k of policies converging to a, $g(a^k)$ converges to $g(a)$.

Note that the sets of policies $\mathcal{M}_{ol}$, $\mathcal{S}_{ol}$, $\mathcal{L}_{ol}(K)$, $\mathcal{M}$, $\mathcal{S}$ and $\mathcal{L}(K)$ are compact.

Lemma 7.1 *(i) For any initial state v and any $T \in \mathbb{N}$, X_n, $c(X_n,a_n)$, $C^T(v,a)$ are continuous in a (for each r and z) and $J^T(v,a)$ are continuous in a, for all $a \in \mathcal{M}_{ol}$.*
(ii) Assume that $\mathbf{X}$ is countable. Then statement (i) holds for all $a \in \mathcal{M}$.

Proof. (i) is straightforward. (ii) Note that for the continuity of $c(X_n,Y_n), n \in \mathbb{N}_0$, we need that a_n, $n \in \mathbb{N}_0$ is continuously convergent, i.e. for $n \in \mathbb{N}_0$,

$$\lim_{k\to\infty} x_k = x \qquad \text{implies} \qquad \lim_{k\to\infty} a_n^k(x_k) = a_n(x).$$

Clearly, this is true if $\mathbf{X}$ is countable.

■

Corollary 7.1 *(i) For any $T \in \mathbb{N}$, there exists an optimal policy in $\mathcal{M}_{ol}$ to $\mathbf{\Pi_1}(\mathcal{M}_{ol})$.*
(ii) Assume that $\mathbf{X}$ is countable. Then for any $T \in \mathbb{N}$, there exists an optimal Markov policy to $\mathbf{\Pi_1}(\mathcal{M})$.

Proof. Follows from the continuity of $J^T(v,a)$ in a and the compactness of the set of policies. ∎

Next we consider the infinite horizon problems $\mathbf{\Pi_2}$ and $\mathbf{\Pi_3}$. For these problems we shall restrict to one of the following set of policies: $\overline{\mathcal{A}}(K) = \mathcal{L}_{ol}(K)$ or $\overline{\mathcal{A}}(K) = \mathcal{L}(K)$, $K \in \mathbb{N}$ (hence in particular to $\mathcal{S}$, $\mathcal{S}_{ol}$). If $\mathbf{X}$ is non-countable, we shall restrict to $\overline{\mathcal{A}}(K) = \mathcal{L}_{ol}(K)$.

Theorem 7.1 *Fix an integer K and an initial state v. Assume that for any policy $a \in \overline{\mathcal{A}}(K)$,*

- *Under P, r is stationary ergodic with respect to the shift T^K,*
- *$A_n^a := T^{nK} A_0$ are renovating on the interval $[nK, (n+L)K]$,*
- *There exists a sequence of sets C_n which are stationary ergodic under the shift T^K, and satisfy $C_n \subset \pi_r A_n^a, \forall a \in \overline{\mathcal{A}}(K)$; and $P(C_0) > 0$.*

Then
(i) $C(v,a) = J(v,a)$ P-a.s. and $J(v,a)$ is lower-semicontinuous in a for all $a \in \overline{\mathcal{A}}(K)$.
(ii) An optimal policy $a^* \in \overline{\mathcal{A}}(K)$ exists for $\mathbf{\Pi_2}(\overline{\mathcal{A}}(K))$ and $\mathbf{\Pi_3}(\overline{\mathcal{A}}(K))$.

Proof. From Section 6 and using Theorem 5.1, it follows that there exists some N^0 which is finite P-a.s., and a class of processes $X^n(a)$, $a \in \overline{\mathcal{A}}(K)$, which are stationary ergodic under the shift T^K, such that

$$X_n(v,a) = X^n(v,a) \qquad \forall n > N^0.$$

It follows from Lemma 7.1 that $X_n(v,a)$ is continuous in a for all n, P-a.s, and hence, for all $n > N^0$, $X^n(v,a)$ is continuous in a (P-a.s). This implies that $c(X^n, Y^n)$, as well as $(\sum_{j=n}^{n+K-1} c(X^j, Y^j))/K$ are continuous in a, for all $n > N^0$ (P-a.s.). (Here $Y^j := a_j(X^j)$). From the stationarity and ergodicity of X^n, it follows by the SLLN that for any a and v,

$$C(v,a) = \frac{E_v \sum_{j=0}^{K-1} c(X^j, Y^j)}{K} = J(v,a) \qquad P-a.s. \tag{38}$$

Let $a^k \in \overline{\mathcal{A}}(K)$ be a sequence of policies converging to some policy $a \in \overline{\mathcal{A}}(K)$. Define $Y^{j,k} := a_j^k(X^j)$. For all n,

$$\lim_{k\to\infty} J(v,a^k) = \lim_{k\to\infty} \frac{E_v \sum_{j=0}^{K-1} c(X^{nK+j}, Y^{nK+j,k})}{K}$$

$$= \lim_{k\to\infty}\left[\frac{E_v\sum_{j=0}^{K-1}c(X^{nK+j},Y^{nK+j,k})1\{n>N^0\}}{K}+\frac{E_v\sum_{j=0}^{K-1}c(X^{nK+j},Y^{nK+j,k})1\{n\le N^0\}}{K}\right]$$

$$\ge \frac{E_v\sum_{j=0}^{K-1}c(X^{nK+j},Y_{nK+j})1\{n>N^0\}}{K}+\frac{E_v\sum_{j=0}^{K-1}c(X^{nK+j},Y^{nK+j})1\{n\le N^0\}}{K}$$

where the inequality follows from Fatou's Lemma. Taking the limit as n tends to infinity, we get by the dominance convergence theorem (since c are bounded)

$$\lim_{k\to\infty} J(v,a^k)\ge J(v,a)$$

Combining this with (38) establishes (i). (ii) now follows from the compactness of $\overline{\mathcal{A}}(K)$. ∎

8 Example: optimality of Hajek sequences

In this section we present another class of policies that were introduced by Hajek in [24]. Consider Example 2.2, where τ_n and σ_n are i.i.d. sequences, exponentially distributed. We assume that we have no information about the state of the system. Let $c(x,a')=x+\gamma a'$, $a'\in\mathbf{A}$ be the immediate cost, where γ is some constant. Consider problem $\mathbf{\Pi_2}(\mathcal{A})$ or $\mathbf{\Pi_3}(\mathcal{A})$.

For the case that the immediate cost is of the form $c'(y,a')=y+\gamma' a'$, $a'\in\mathbf{A}$, where y is the number of customers in the queue, Hordijk and Loeve [29] have shown that in the class of open loop policies one may restrict to the class of policies $\overline{\mathcal{A}}_1$ for which the limit

$$q=\lim_{n\to\infty}\frac{1\{z_n\le a_n\}}{n}$$

exists, without loss of optimality. Consider now an arbitrary policy $a(q)$, achieving the above limit. Hajek further showed that $J(v,a^*(q))\le J(v,a)$, where $a^*(q)$ is defined by: $a_0=\lfloor\psi\rfloor$, and the kth action is given by

$$a_k=\lfloor kq+\psi\rfloor-\lfloor (k-1)q+\psi\rfloor; \tag{39}$$

$\psi\in[0,1]$ is an arbitrary number in $[0,1]$. Thus, in the setting of exponential service and interarrivals, one may restrict, without loss of optimality, to Hajek's policies (i.e., satisfying (39)), if the cost c' is considered. This optimality result also holds for the original cost c, since, due to the exponentiality of the service times, for any policy, the expected queue length is proportional to the expected workload in the queue. This also holds in the ergodic sense (the limit of the Cezaro sums, defining the cost function $C(v,a)$) due to the Strong Law of Large Numbers.

Motivated by this fact, we shall now restrict to Hajek's policies only. We relax the i.i.d. and exponential assumptions, allowing general stationary ergodic service and interarrival times. We show that among the Hajek's policy, there exists an optimal one (it will thus be an optimal policy among all open loop policies, in the special case that service times and inter-arrival times are exponential).

If q in (39) is irrational, then this implies by Weyl's ergodic Theorem [13] p. 159, that $a_k^*(q)$ are ergodic, under the shift T_a and hence, so is ω under T. We shall assume without loss of optimality that ψ is a uniformly distributed random variable over $[0,1]$. Unlike the setting of the previous section, the policy is thus random too. In order to handle this situation, we simply consider the original setting of Borovkov, (with no control), by allowing the driving sequence to include not only the sequences τ_n and σ_n, but also a_n. The process of the driving sequence $\omega = (\sigma_n, \tau_n, a_n)^\infty$ is stationary ergodic. If q is rational then $a^*(q)$ is periodic with some period K. Then, one may use again the above construction to include a in the driving sequence, which results in ω being stationary ergodic under the shift T^K. In both cases we may conclude by using Theorem 3.1, together with its extension for periodic policies (Section 6) that if

$$\rho(q) := \frac{qE\sigma_0}{E\tau} < 1$$

then X_n couples in finite time (P-a.s.) to a process that is stationary ergodic under a shift T^K, where K is finite. Hence, for such q,

$$C(v, a^*(q)) = J(v, a^*(q)) = \gamma q + \overline{X}(q), \qquad P - a.s. \tag{40}$$

where $\overline{X}(q)$ is the ergodic average workload when $a^*(q)$ is used.

Let $\overline{\mathcal{A}}_\rho$ be the set of Hajek policies $a^*(q)$ for all $\rho(q) \in [0, \rho]$. Under fairly general conditions (in particular, when the service and interarrival times are i.i.d. and not deterministic), one can show that

$$\lim_{\rho(q)\to 1} \overline{X}(q) = \infty. \tag{41}$$

It is therefore natural to restrict the optimization problem to $\overline{\mathcal{A}}_\rho$ with $\rho < 1$, for which (40) holds. It is easily seen that the set of policies $\overline{\mathcal{A}}_\rho$ is compact. Using the same type of arguments as in Theorem 7.1, we conclude that there exists a policy in $\overline{\mathcal{A}}_\rho$ which is optimal for both problems $\mathbf{\Pi_2}(\mathcal{A})$ and $\mathbf{\Pi_3}(\mathcal{A})$.

Remark 8.1 The policy $a^*(q)$ was shown to be stationary for irrational q. This stationarity is in the probabilistic sense, i.e. the distribution of such policies is invariant under the shift T. This is not to be confused with another notion of stationary policies. The class of policies $\mathcal{S}$ are often called stationary in the theory of control, see [17] p. 137; $a^*(q)$ is clearly not stationary in that sense.

9 Conclusions

We begin by interpreting the coupling results we obtained in Theorems 5.1, 5.2 and 5.3 in the context of control. The first important result is the forward coupling result, i.e. the fact that for any given

Markov control, we identified a limit process X^s such that the state process couples with a limiting process X^s, after some time $N^0 \leq N^*$, where N^* is P-a.s. finite (under suitable conditions). Under some additional conditions, the limiting process is independent of the initial state. Moreover, under further assumptions, the time till coupling is uniformly bounded for some classes of policies. All the uniformly boundedness of coupling times are obtained for the original probability space, on which trajectories are defined simultaneously for all initial states and policies. This type of result was the motivating reason for us to investigate coupling, as it seems to be a natural way to define stability for non-stationary stochastic processes in general, and in CSRS in particular.

A second interesting coupling consequence is on the finiteness of the backward coupling time N_0, (or, more generally, on the strong coupling time N^*). It has, as application in control, the coupling for backward policies, which are constructed as mentioned in the end of Remark 2.1. Hence, by using our probability space, we may couple the terminating state (or some terminating cost) that is obtained by simultaneously using the optimal policies a^N for a Markov decision process with horizon N (obtained by using the backward induction - successive approximation method, [16]), for all N sufficiently large.

As an application, we recall Example 2.1, and consider some terminating holding cost which is convex in the number of customers, and some additional reward for accepting customer. One can show in a Markovian setting in general, and in the particular case that the interarrival and service times are i.i.d. sequences, exponentially distributed, that the optimal policies are of threshold type (accept if and only if the number of customers are below some threshold v_n, that depends on the length n of the remaining horizon), and moreover, there is a constant v such that the optimal thresholds v_n are uniformly bounded: $v_n \leq v$. (Such type of results are known, see [34] who restricts to *increasing* convex holding cost and some additional reward per accepting customers; the result for general terminating convex holding cost, not necessarily increasing, can then be obtained by combining these arguments with those in [37], who studied general convex holding cost). From Example 2.4 we now conclude that the terminating cost (or terminating state), will couple, for all N sufficiently large, since the renovating events introduced there for the above policies have positive probabilities.

Finally, our coupling results enabled us to establish the existence of optimal policies within some subclasses of policies (Section 7).

Acknowledgement This research has been supported through visitor's grants from INRIA, and the Thomas Stieltjes Institute for Mathematics. The authors are greatful for the hospitality at INRIA and the Mathematical Institute of Leiden University during their visits.

References

[1] I. Akhmarov and N. P. Leont'eva, "Conditions for convergence to limit processes and the strong law of large numbers for queueing systems", *Teor. Veroyatnost i ee Primenen* **21**, (in Russian) pp. 559-570, MR 58 # 24597, 1976.

[2] S. Asmussen and S. G. Foss, "Renovation, regeneration, and coupling in multiple-server queues in continuous time", *Frontiers in Pure and Appl. Probab.* **1**, pp. 1-6, H. Niemi et al. (Eds), 1993.

[3] F. Baccelli and P. Brémaud, *Elements of Queueing theory*, Springer-Verlag, 1994.

[4] F. Baccelli, and S. Foss, "Stability of Jackson-type Queueing Networks", *QUESTA* **17**, pp. 5-72, 1994.

[5] F. Baccelli and S. Foss, "On the Saturation Rule for the Stability of Queues", *Journal of Applied Probability* **32**, pp. 494-507, 1995.

[6] Bertsekas, D. P., *Dynamic Programming. Deterministic and Stochastic Models*. Prentice Hall, Inc. Englewoods Cliff, 1987.

[7] A. A. Borovkov, *Asymptotic Methods in Queueing Theory*, John Wiley & Sons, 1984 (translated from Russian).

[8] A. A. Borovkov, *Ergodicity and stability of stochastic processes*, to appear (in Russian).

[9] A. A. Borovkov, "Remarks on boundedness of stochastic sequences and on estimates of their moments", preprint.

[10] A. A. Borovkov and S. G. Foss, "Stochastically recursive sequences and their generalizations", *Siberian Advances in Mathematics*, **2**, No. 1, pp. 16-81, 1992.

[11] A. A. Borovkov and S. G. Foss, "Two ergodicity criteria for stochastically recursive sequences", *ACTA Applicande Math.* **34**, pp. 125-134, 1994.

[12] A. Brandt, P. Franken and B. Lisek, *Stationary Stochastic Models*, Akademie-Verlag, Berlin, 1992.

[13] I. P. Cornfeld, S. V. Fomin and Ya. G. Sinai, *Ergodic Theory*, Springer-Verlag, New-York, 1982.

[14] R. Dekker and A. Hordijk, "Average, sensitive and Blackwell optimal policies in denumerable Markov decision chains with unbounded rewards", *Mathematics of Operations Research*, **13**, pp. 395-421, 1988.

[15] R. Dekker, A. Hordijk and F. M. Spieksma, "On the relation between recurrence and ergodicity properties in denumerable Markov decision chains", *Math. Operat. Res.*, **19**, pp. 539-559, 1994.

[16] C. Derman, *Finite State Markovian Decision Processes*, Academic Press, 1970.

[17] E. B. Dynkin and A. A. Yushkevich, *Controlled Markov Processes*, Springer-Verlag, 1979.

[18] A. Federgruen, A. Hordijk and H. C. Tijms, "A note on simultaneous recurrence conditions on a set of denumerable stochastic matrices", *J. Appl. Prob.* **15**, pp. 842-847, 1978.

[19] A. Federgruen and H. C. Tijms, "The optimality equation in average cost denumerable state semi-Markov decision problems, recurrency conditions and algorithms", *J. Appl. Prob.* **15**, pp. 356-373, 1978.

[20] S. G. Foss, "The method of renovating events and its applications in queueing theory", *Semi Markov Models, Theory and Applications*, Proc. 1-st Symp on Semi-Markov Processes, Brussel 1984 (Plenum 86).

[21] S. G. Foss, Private communication.

[22] S. G. Foss and V. V. Kalashnikov, "Regeneration and renovation in queues", *Queueing Systems* **8**, No. 3, pp. 211-223, 1991.

[23] B. Hajek, "Hitting-time and occupation-time bounds implied by drift analysis with applications", *Adv. Appl. Prob.*, **14**, pp. 502-525, 1982.

[24] B. Hajek, "Extremal splitting of point processes", *Mathematics of Operations Research*, **10**, pp. 543-556, 1985.

[25] A. Hordijk, "Regenerative Markov decision models", *Math. Prog. Study* **6**, pp. 49-72, 1976.

[26] A. Hordijk, *Dynamic Programming and Markov Potential Theory*, Second Edition, Mathematical Centre Tracts 51, Mathematisch Centrum, Amsterdam, 1977.

[27] A. Hordijk and P. J. Holewijn, "On the convergence of moments in stationary Markov chains", *Stoch. Proc. Appl.*, **3**, pp. 55-64, 1975.

[28] A. Hordijk and J. A. Loeve, "Undiscounted Markov decision chains with partial information; an algorithm for computing a locally optimal periodic policy", *ZOR - Mathematical Methods of Operations Research* **40**, pp. 163-181, 1994.

[29] A. Hordijk and J. A. Loeve, "Markov decision chains with partial information: optimality results", Technical report, University of Leiden.

[30] A. Hordijk, G. M. Koole and J. A. Loeve, "Analysis of a customer assignment model with no state information", *Probability in the Engineering and Informational Sciences*, **2**, pp. 419-429, 1994.

[31] D. G. Kendall, "Geometric ergodicity in the theory of queues", *Mathematical Methods in the Social Sciences*, eds. K. J. Arrow, S. Karlin and P. Suppes, Stanford University Press, pp. 176-195, 1960.

[32] V. G. Kulkarni and Y. Sering, “Optimal implementable policies: average cost and minimax criteria", *Technical report UNC/OR TR-90/9*, Department of Operations Research, University of North Carolina, Chapel Hill, 1992.

[33] V. G. Kulkarni and Y. Sering, “Optimal implementable policies: discounted cost case", *Computations with Markov chains*, pp. 283-307, Ed. W. Stewart, Kluwer Academic Publications, 1995.

[34] S. A. Lippman and S. Stidham, “Individual versus social optimization in exponential congestion systems", *Operations Research* **25**, pp. 233-247, 1977.

[35] T. Lindvall. *Lecture on the Coupling Method*, Wiley, 1992.

[36] R. Loynes, “The stability of a queue with non-independent inter-arrival and service times", *Proc. Cambr. Phil. Soc.* **58**, No. 3, pp. 497-520, 1962.

[37] D. J. Ma and A. M. Makowski, “Optimality results for a simple flow control problem", *Proceedings of the 26th IEEE Conference on Decision and Control*, Los Angeles, California, December 1987.

[38] S. Meyn and R. L. Tweedie, *Markov chains and stochastic stability* , Springer-Verlag, 1993.

[39] F. M. Spieksma, *Geometrically Ergodic Markov Chains and the Optimal Control of Queues*, Ph.D. thesis, 1990, University of Leiden.

[40] R. L. Tweedie, “The existence of moments for stationary Markov chains", *J. Applied Probability* **20**, pp. 191-196, 1983.

10 Appendix: on Ex. 2.1 and 2.3

Proof of Observation, in Ex. 2.1 (continued):

$X_j^{[n+1]}$ (see relation (15) for the notation) can be written as

$$X_j^{[n+1]} = \max\left(0, \sum_{i=k}^{j-1} \xi_i, k = -n, ..., j-1, \sum_{i=-n-1}^{j-1} \xi_i + v\right).$$

First consider the case $v = 0$. Then,

$$X_j^{[n+1]} = \max(X_j^{[n]}, \sum_{i=-n-1}^{j-1} \xi_i). \tag{42}$$

Hence $X_j^{[n]}$ is monotone increasing in n for all j.

Assume that $E\xi < 0$. Then due to the ergodicity of ξ, we have $\lim_{n\to\infty} \sum_{i=-n-1}^{-1} \xi_i = -\infty, P-a.s.$, and it follows from (42) that there exists some (random) N such that $X_0^{[n]} = X_0^{[N]}, \forall n \geq N, P-a.s.$ This implies (see (15)) that

$$X_j^{[n]} = X_j^{[N]} \qquad \forall n \geq N, \forall j \geq 0,\ P-a.s. \tag{43}$$

Assume now that $P(A_0^{\overline{a}}(0)) = 0$ for some $\nu > 0$. Then $P(\cup_{j\geq 0} A_j^{\overline{a}}(0)) = 0$ for that ν. On the complement of $\cup_{j\geq 0} A_j^{\overline{a}}(0)$ we have for all $j \geq 0$, $X_j^{[N_j]} > 0$ for some (random) $N_j \geq \nu$. Due to the monotonicity of $X_j^{[n]}$ in n, we have $X_j^{[n]} > 0$ for all $j \geq 0$ and $n > N_j$. (43) then implies that $X_j^{[n]} > 0$ for all $n \geq N$ and $j \geq 0$, so that $\lim_{n\to\infty} P(X_j^{[n]} > 0, \forall j \geq 0) = 1$. By stationarity, we finally obtain $\lim_{n\to\infty} P(X_{n+j} > 0, \forall j \geq 0) = 1$. This establishes the observation for $v = 0$.

In order to establish the Observation for arbitrary $v > 0$, we need the following Lemma. To indicate the dependence of the state trajectory on v, we add v as argument of the state.

Lemma 10.1 *Consider Example 2.1 (continued). Assume that $E\sigma_0 < E\tau_0$. Then under any policy, $X_0^{[n]}(v)$ couples with $X_0^{[n]}(0)$ within a time that is finite w.p.1. (yet may depend on the value of v).*

Proof. First note that $X_0^{[n]}(v) \geq X_0^{[n]}(0)$ since $v \geq 0$. Moreover, $X_0^{[n+1]}(v)$ can be written as

$$\begin{aligned}
X_0^{[n+1]}(v) &= \max\left(0, \sum_{i=-k}^{-1} \xi_i, k = 1, ..., n+1, \sum_{i=-n-1}^{-1} \xi_i + v\right) \\
&= \max\left(X_0^{[n+1]}(0), \sum_{i=-n-1}^{-1} \xi_i + v\right) \\
&\leq \max\left(X_0^{[n+1]}(0), \sum_{i=-n-1}^{-1} (\sigma_i - \tau_i) + v\right).
\end{aligned} \tag{44}$$

Since (σ_i, τ_i) are stationary ergodic and $E(\sigma_0 - \tau_0) < 0$, we have by Birkhoff's ergodic Theorem $\lim_{n\to\infty} \sum_{i=-n-1}^{-1} (\sigma_i - \tau_i) = -\infty$, P-a.s. This implies the Lemma. ∎

Next, we show that the Observation holds for any $v > 0$. Assume that $P(A_0^{\overline{a}}(v, \nu)) = 0$ for all ν. If $P(A_0^{\overline{a}}(0, \nu)) = 0$ for all ν, then we showed that $\lim_{m\to\infty} P(X_n(0) > 0, \forall n \geq m) = 1$, and the observation for v will follow since we have, samplewize, $X_n(v) \geq X_n(0)$. So, assume that $\delta \triangleq P(A_0^{\overline{a}}(0, \nu^*)) > 0$ for some ν^*. Since we assumed that $P(A_0^{\overline{a}}(v, \nu)) = 0$ for all ν, there exists a random increasing sequence n_i such that $P(X_0^{[n_i]}(v) > 0$ for all $i) = 1$ under $\overline{a}$. Since $X_0^{[n]}(v)$ and

$X_0^{[n]}(0)$ couple in a finite time P-a.s., this implies that

$$\overline{\lim_{n\to\infty}} X_0^{[n]}(0) > 0, \qquad P - a.s.$$

But this contradicts the following: under policy $\overline{a}$, $A_0^{\overline{a}}(0,\nu^*) \subset \{X_0^{[n]}(0) = 0 \text{ for all } n \geq \nu^*\}$, so that $P(X_0^{[n]}(0) = 0, \forall n \geq \nu^*) \geq \delta > 0$. Thus, we conclude that $P(A_0^{\overline{a}}(0,\nu)) = 0$ for all ν, which implies the Observation. ■

Remark 10.1 (i) For the case $v = 0$ it follows that we can replace the assumption on ν in the observation to "for some ν", instead of "for all ν".
(ii) For the case that $E\sigma_0 > E\tau_0$, it follows from Birkhoff's ergodic Theorem applied to (5) and (6), that $P(A_0^{\overline{a}}(v)) = 0$ for all $\nu > 0$, and that under $\overline{a}$, $\lim_{m\to\infty} P(X_n > 0, \forall n \geq m) = 1$.
(iii) The observation need not hold for the case $E\sigma_0 = E\tau_0$. Indeed, consider the sequence $\tau_n = \tau = 2$, $\sigma_n = 2 + Z(-1)^n$, where Z equals 1 w.p. 0.5 and -1 with probability 0.5. For $v = 1$ we have $P(A_0^{\overline{a}}(v)) = 0$ for all $\nu > 0$, since for any integer n, if $\sum_{i=-n}^{-1} \xi_i + v \leq 0$ then $\sum_{i=-n-1}^{-1} \xi_i + v = 1 > 0$. However, $P(X_n > 0, n \geq N) = 0.5$ for any N, since with probability 0.5 the queue empties every second time unit, and with probability half it never empties.

Lemma 10.2 *Consider Ex. 2.3.*
(i) The sets $\overline{\mathsf{A}}_n^a$ are renovating on the intervals $[n, n+K_2-1]$.
(ii) For any feasible a with period K_2, we have $P(\overline{\mathsf{A}}_0^a) > 0$.

Proof. We shall use the notation K instead of K_2.
(i) For $n > \nu$, we have on $A_n^a(v,\nu)$, $X_n = 0$, for an initial state $X_0 = v$. Hence on $\overline{\mathsf{A}}_n^a(v)$ we have $X_{n+K-j} = 0$ for some $j = 1, ..., K$, so that

$$X_{n+K} = \max\left(\xi_{n+K-1}, \xi_{n+K-1} + \xi_{n+K-2}, ..., \xi_{n+K-1} + ... + \xi_{n+K-j}\right). \tag{45}$$

But by (6),

$$X_{n+K} \geq \max\left(\xi_{n+K-1}, \xi_{n+K-1} + \xi_{n+K-2}, ..., \xi_{n+K-1} + ... + \xi_{n+K-m}\right)$$

for all $K \geq m > j$, which implies, in particular, that X_{n+K} satisfies the relation (45) with $j = K$. This easily implies that $\overline{\mathsf{A}}_n^a$ are indeed renovating for $n > \nu$ on the interval $[n, n+K-1]$.

(ii) Let a be a feasible policy with period $K = K_2$. Consider first $v = 0$ with ν fixed. On the complement of $\cup_j \overline{\mathsf{A}}_{jK}^a(0,\nu)$, we have for every $m = 0, K, 2K,$

$$X_m^{[n]} > 0, X_{m+1}^{[n]} > 0, ..., X_{m+K-1}^{[n]} > 0 \tag{46}$$

for some $n = N_m \geq \nu$, where N_m is random. As in the beginning of the appendix, $X_j^{[n]}$ are monotone increasing in n, which implies that (46) holds also for all $n > N_m$; moreover, there exists some random N such that (43) holds. If we assume that $P(\overline{\mathsf{A}}_0^a) = 0$ then it follows from stationarity that $P(\cup_j \overline{\mathsf{A}}_{jK}^a) = 0$. Hence, $\lim_{m\to\infty} P(X_n^{[m]} > 0, \forall n \geq 0) = 1$, from which we obtain, by stationarity, $\lim_{m\to\infty} P(X_n > 0, \forall n \geq m) = 1$. Hence there is some integer (nonrandom) N such that set Π_N defined above (7) has strictly positive probability, so that (7) holds with strictly positive probability. But this is in contradiction with the following. Since (σ_n, τ_n) are stationary ergodic with respect to the K-step shift, it follows from Birkhoff's ergodic theorem that $P - a.s.$,

$$\lim_{n\to\infty} \frac{\sum_{i=N}^{N-1+nK} \xi_i}{n} = \sum_{i=0}^{K-1} E\xi_i \leq K_1 E\sigma_0 - K_2 E\tau_0 < 0.$$

(The last inequality follows from the feasibility of a and the stationarity of σ_n and τ_n with respect to the one step shift.) Hence, on Π_N, X_n should converge to $-\infty$. This establishes (ii) for the case that $v = 0$ (by contradiction).

The case of $v > 0$ is obtained using a coupling argument as the one established in Lemma 10.1. To obtain the latter, note note that (44) still holds on $\overline{\mathsf{A}}_n^a(v)$. Substitute $n + 1 = mK + j$, $j = 0, ..., K - 1$, in (44), and take the limit as m goes to infinity. Using again Birkhoff's ergodic theorem, we conclude that $\lim_{n\to\infty} \sum_{i=-1}^{n} \xi_i = -\infty$, which implies that $X_0^{[n]}(v)$ couples to $X_0^{[n]}(0)$ in a time n that is finite a.s. ∎

Finally, we show the necessity of the assumption that σ_n be stationary ergodic with respect to the K_2 shift as well, in Ex. 2.3.

Example 2.3 (continued): Consider Ex. 2.3 with $K_2 = 2$, $K_1 = 1$, $\tau_n = 4$, and $\sigma_n = 5 + Z(-1)^n$, where Z equals 5 w.p. 0.5 and -5 with probability 0.5. Then $\rho = 1.2$, and the stability condition $K_1\rho < K_2$ holds. However, the policy that accepts every second customer is feasible, yet the queue length and the workload diverges to infinity w.p. 1/2. It turns out that the probability that relation (43) holds is less than 1, although the stability condition holds. In this example, $P(\overline{\mathsf{A}}_n^a(v)) > 0$ for some values of v, and is zero for large v. ∎