Asymptotic efficiency of restart and checkpointing

Antonio Sodre*
University of Texas at Austin

Abstract

Many tasks are subject to failure before completion. Two of the most common failure recovery strategies are `restart` and `checkpointing`. Under `restart`, once a failure occurs, it is restarted from the beginning. Under `checkpointing`, the task is resumed from the preceding checkpoint after the failure. We study asymptotic efficiency of `restart` for an infinite sequence of tasks, whose sizes form a stationary sequence. We define asymptotic efficiency as the limit of the ratio of the total time to completion in the absence of failures over the total time to completion when failures take place. Whether the asymptotic efficiency is positive or not depends on the comparison of the tail of the distributions of the task size and the random variables governing failures. Our framework allows for variations in the failure rates and dependencies between task sizes. We also study a similar notion of asymptotic efficiency for `checkpointing` when the task is infinite a.s. and the inter-checkpoint times are i.i.d.. Moreover, in `checkpointing`, when the failures are exponentially distributed, we prove the existence of an infinite sequence of universal checkpoints, which are always used whenever the system starts from any checkpoint that precedes them.

Key words: restart, checkpointing, failure recovery, dynamical systems, point process, point-shift.

MSC 2010 subject classification: Primary: 37A05, 60G55.

Introduction

In many situations, such as the execution of a computer program, the copy of a file from a remote location using a protocol such as FTP or HTTP, channel reservation in cognitive radio networks and others, tasks are subject

*asodre@math.utexas.edu

to failures. `Restart` and `checkpointing` are two of the most common ways to take into account failures in these context (see, among others, [15],[17], and [8]).

In `restart`, whenever a failure occurs, as the name suggests, the task is restarted. Accordingly, the *actual* time for completion is possibly larger than the *ideal* time. The latter is defined as the time for completion without failures. In `checkpointing`, the task is partitioned: when a failure occurs, it is resumed from the last element of the partition before the failure.

Here is a basic description of `restart`. Let D be the ideal task time. If no failure occurs, the actual time to complete the task is just D. If a failure occurs at $L_0 < D$, the task is restarted. Suppose there are $\nu > 0$ failures before the task is completed. Then the actual time is given by $T^R = \sum_{i=0}^{\nu} L_i + D$. Failures are modeled by a sequence of i.i.d. random variables $\{L_n\}_{n\geq 0}$, named *failure* times. The one-task `restart` model is studied in [2] and [4] for a random variable D with unbounded support (see Figure 1). Section 1 introduces the formalism for the one-task `restart` model.

In the one-task case, the actual time, T^R, may be heavy-tailed, even when the ideal time and the failure time have light tails. Moreover, the actual task time may have infinite expectations, even if both D and L_0 do not, depending on the comparison of the tail distributions of D and L_0 [2].

We extend the literature on `restart` by considering an infinite sequence of tasks, $\{D_n\}_{n\geq 0}$, introducing the concept of asymptotic efficiency. Let T_n^R be the actual time of task n. We define asymptotic efficiency as

$$e = \lim_{N\to\infty} \frac{\sum_{n=0}^{N-1} D_n}{\sum_{n=0}^{N-1} T_n^R}, \tag{0.1}$$

whenever the limit exists a.s.. The system is inefficient when $e = 0$.

In this sequential `restart` model, the ideal times is given by the distance between points of a simple stationary point process in $\mathbb{R}$. Such a point process can be seen as a random discrete sequence of distinct elements on $\mathbb{R}$, $\{X_n\}_{n\in\mathbb{Z}}$, such that $X_n < X_{n+1}$ for all n. The sequence of tasks sizes is given by $D_0 = X_1 - X_0$, $D_1 = X_2 - X_1$ and so on. We mark the point X_n with an i.i.d. sequence, $\{L_{n,i}\}_{i\geq 1}$, capturing the failure times of the n^{th}-task. We present the point process setting for modeling task sizes and failures in Section 2.

We prove that asymptotic efficiency exists when the point process is stationary, the failure sequence $\{L_{n,i}\}_{i\geq 1}$ is independent of D_n, and under some integrability conditions. We do not require the sequence $\{D_n\}_{n\geq 0}$ to be i.i.d.. In fact, our set-up allows for variations in the failure rates and dependencies between task sizes.

Moreover, we give conditions under which the asymptotic efficiency is positive or zero. Two special cases are considered: Markov renewal process, and the case in which there is a chance that tasks need to be repeated after completion (Section 5).

The `checkpointing` model can be described as follows [3]. Again, we have a random a task D of random length with infinite support, but finite a.s.. We partition $[0, D]$ into k intervals and label the endpoint of the l^{th} interval by X^l. We call $\{X^l\}_{l=1}^{k-1}$ the set of checkpoints. Once a checkpoint is reached and a failure occurs, the task is resumed from the latest checkpoint before the failure. More precisely, if a failure occurs before the first checkpoint, i.e., $L_1 < X^1$, the task is resumed from the beginning. If $L_1 > D$, i.e., there are no failures, the actual time to completion is simply D. Otherwise, if $X^1 < L_1 < D$, we check the partition in which L_1 falls. If $X^l \leq L^1 < X^{l+1}$, the task is resumed from X^l and the time spent so far is L_1. In that case, we start the clock again, representing it by the random variable L_2. If $L_2 < X^{l+1} - X^l$, the task does not leave the checkpoint X^l. Otherwise, we verify which checkpoint was reached or whether the task was completed. We repeat this procedure until the task is completed. Assume that there are $\tau > 0$ failures until completion. Then, the actual time is given by $T^C = \sum_{i=1}^{\tau} L_i + (D - X_\alpha)$, where $\alpha \in \{1, \dots, k\}$ is the last checkpoint visited (see Figure 2).

Regarding the sequential `checkpointing` considered here, we define and study a notion of asymptotic efficiency, in a similar way to (0.1). We consider a unique task, which is a.s. infinite, and the distances between checkpoints are given by the inter-arrivals of a point process. We give the precise definition of asymptotic efficiency for `checkpointing` in Section 2. We give a general condition for the existence of the asymptotic efficiency when the point process is a marked renewal process.

Moreover, in the renewal process model with exponentially distributed failure times, we show the existence of an infinite subsequence of *universal checkpoints*. If we start the system at any checkpoint preceding a universal checkpoint, the system will activate the latter a.s..

Section 1 reviews the *actual* time for one-task `restart` and `checkpointing`, and gives the conditions under which the *actual* time has finite moments. Section 2 presents a unified framework to study the asymptotic efficiency for both sequential `restart` and `checkpointing`. Section 3 presents our main results for sequential `restart`. Section 4 does the same for sequential `checkpointing`. Section 5 discusses some extensions. The appendix contains a technical proof.

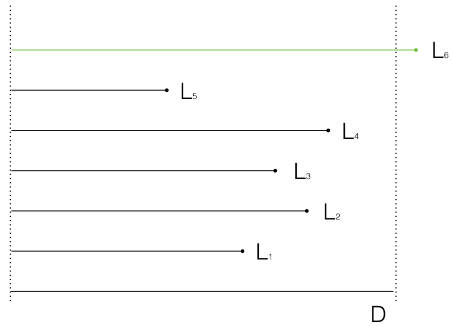

Figure 1: An instance of `restart`. The task size is D. Five attempts take place before the task is completed, i.e., $\nu = 5$. The time spent on each attempt is given by $L_1, \ldots, L_5$. In the sixth attempt the task is completed. The *actual* time spent on completing the task is then $T^R = \sum_{i=1}^{5} L_i + D$.

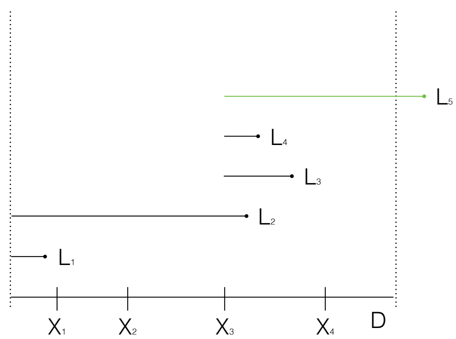

Figure 2: An instance of `checkpointing`. There is one failure before the first checkpoint, the second failure takes place after the third checkpoint and two more failures happen before the last checkpoint is surpassed. The *actual* time till completion is given by $T^C = \sum_{i=1}^{4} L_i + (D - X_3)$.

1 One task `restart` and `checkpointing`

In this section we recall known results on how to compute the actual time for `restart` and `checkpointing` with one task. The main result, which was obtained in [2], [4], and [5], is that the actual `restart` and `checkpointing` times have an infinite expectation if and only if the task size D has a tail heavier than L_0, the variable that captures failures. We build on this result in our study of the asymptotic efficiency in the next section.

Let D be a random variable in $\mathbb{R}^+$ which represents the ideal task time in both the `restart` model or the time up to the first checkpoint under the `checkpointing` model. Consider a sequence of i.i.d. random variables $\{L_n\}_{n\geq 0}$ defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ as D. Define

$$\tau = \inf\{k \geq 0 : L_k > D\}.$$

The actual time taken to complete the task D under `restart` is given by

$$T^R = \sum_{i=0}^{\tau-1} L_i + D,$$

and the actual time the system takes to pass the first checkpoint is

$$T^C = \sum_{i=0}^{\tau} L_i.$$

Assumption 1. D and L_0 are integrable and independent random variables with *right-unbounded support*, i.e, $\mathbb{P}[D > x], \mathbb{P}[L_0 > x] > 0$ for all $x \in [0, \infty)$.

Definition 1. Let V and W be random variables with right-unbounded support, defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say then that V has a $\mathbb{P}-$tail heavier than W if there exists z_0 such that $\mathbb{P}[V > z] \geq \mathbb{P}[W > z]$ for all $z \geq z_0$.

Theorem 2. Under Assumption 1,

$$\mathbb{E}[T^R] = \mathbb{E}[D] + \int_0^\infty \frac{\mathbb{E}[L_0 \mathbf{1}\{L_0 \leq z\}]}{\mathbb{P}[L_0 > z]} f_D(dz)$$

and

$$\mathbb{E}[T^C] = \mathbb{E}[L_0] \int_0^\infty \frac{1}{\mathbb{P}[L_0 > z]} f_D(dz),$$

where f_D is the distribution of D. Moreover, $\mathbb{E}[T^R], \mathbb{E}[T^C] = \infty$ if and only if D has a $\mathbb{P}-$tail heavier than L_0.

The proof is adapted from [3] and can be found in Appendix A.

As an application of Theorem 2, suppose $L_0 \sim \exp(\lambda_l)$ and $D \sim \exp(\lambda_d)$. Then $\mathbb{E}[T^R], \mathbb{E}[T^C] = \infty$ if and only if $\lambda_l \geq \lambda_d$.

A random variable Z is said to be heavy-tailed if for all $\gamma > 0$,

$$\lim_{t \to \infty} e^{\gamma t} \mathbb{P}[Z > t] = \infty,$$

and light-tailed if there exists $\gamma > 0$ such that the above limit is finite. By direct manipulations, one gets the following corollary of Theorem 2.

Corollary 3. Under Assumption 1:

1. If D is heavy-tailed and L_0 is light-tailed, $\mathbb{E}[T^C], \mathbb{E}[T^R] = \infty$;
2. If L_0 is heavy-tailed and D is light-tailed, $\mathbb{E}[T^C], \mathbb{E}[T^R] < \infty$.

2 Sequential `restart` and `checkpointing`

The goal of this section is to define the asymptotic efficiency under `restart` (resp. `checkpointing`) when there is a sequence of tasks (resp. a sequence of checkpoints) whose ideal times to completion (resp. distance between checkpoints) are given by the inter-arrival times of a stationary point process. We call the models introduced in this section *sequential* `restart` and `checkpointing`.

2.1 Point process and stationarity

First, we briefly review the necessary concepts in point process theory. For a more complete treatment on the subject see [10],[11],[9] among others. Consider a general probability space endowed with a measurable flow $(\Omega, \mathcal{F}, \mathbb{P}, \{\theta_t\}_{t \in \mathbb{R}})$. Let $\mathbf{N}(\mathbb{R} \ltimes (\mathbb{R}^+)^{\mathbb{N}})$ be the set of counting measures on $\mathbb{R}$ with marks in $(\mathbb{R}^+)^{\mathbb{N}}$. An element of $\mathbf{N}(\mathbb{R} \ltimes (\mathbb{R}^+)^{\mathbb{N}})$ is of the form $\psi = \sum_{n \in \mathbb{Z}} \delta_{(X_n, K_n)}(\cdot)$, in which $\delta_Z(\cdot)$ is the Dirac measure with mass at Z, $X_n \in \mathbb{R}$, $K_n \in (\mathbb{R}^+)^{\mathbb{N}}$, and the sequence $\{X_n\}_{n \in \mathbb{Z}}$ does not have accumulation points. We say K_n is the mark of X_n. For any $C \in \mathcal{B}(\mathbb{R}^+ \times (\mathbb{R}^+)^{\mathbb{N}})$: $\psi(C) = \sum_{n \in \mathbb{Z}} \delta_{(X_n, K_n)}(C)$. We write $X_n \in \psi$ whenever $\psi(\{X_n, K_n\}) \geq 1$.

We equip $\mathbf{N}(\mathbb{R} \ltimes (\mathbb{R}^+)^{\mathbb{N}})$ with the smallest $\sigma-$algebra $\mathcal{N}(\mathbf{N}(\mathbb{R} \ltimes (\mathbb{R}^+)^{\mathbb{N}}))$ that makes the family of mappings

$$\{\psi \mapsto \psi(C) : C \in \mathcal{B}(\mathbb{R}^+ \times (\mathbb{R}^+)^{\mathbb{N}}),\ C \text{ bounded}\}$$

measurable. A point process on $\mathbb{R}^+$ with marks in $(\mathbb{R}^+)^{\mathbb{N}}$ is a measurable mapping $\Phi : \Omega \to \mathbf{N}(\mathbb{R}^+ \ltimes (\mathbb{R}^+)^{\mathbb{N}})$.

The realization of a marked point process in $\mathbf{N}(\mathbb{R} \ltimes (\mathbb{R}^+)^{\mathbb{N}})$ corresponds to a sequence

$$\{X_n(\omega), K_n(\omega)\}_{n\in\mathbb{Z}} \subset \mathbb{R} \times (\mathbb{R}^+)^{\mathbb{N}}$$

such that $\{X_n(\omega)\}_{n\in\mathbb{Z}}$ has no accumulation points $\mathbb{P}-$a.s.. We often write X_n instead of $X_n(\omega)$. For all $C \in \mathcal{B}(\mathbb{R}\times\mathbb{R}^+)$, we let $\Phi(\omega, C) = \#\{(X_n, K_n)(\omega) \in C\}$. Moreover, we always label the points of Φ in $\mathbb{R}$ as follows:

$$\ldots \leq X_{-2} \leq X_{-1} \leq X_0 \leq 0 \leq X_1 \leq \ldots.$$

We assume Φ is θ_t-compatible, i.e., for all $t \in \mathbb{R}$,

1. $\mathbb{P} \circ (\theta_t)^{-1} = \mathbb{P}$,
2. For all $C \in \mathcal{B}(\mathbb{R})$ and $D \in \mathcal{B}((\mathbb{R}^+)^{\mathbb{N}})$:

$$\Phi(\theta_t\omega, C \times D) = \Phi(\omega, (C + t) \times D).$$

These, together with

$$\Phi(\omega, C \times (\mathbb{R}^+)^{\mathbb{N}}) < \infty \text{ for all } C \in \mathcal{B}(\mathbb{R}) \text{ bounded, } \mathbb{P} - \text{a.s.},$$

makes Φ a stationary marked point process.

Remark 4. It is most convenient for our purposes to take Ω to be $\mathbf{N}(\mathbb{R}^+ \ltimes (\mathbb{R}^+)^{\mathbb{N}})$ and $\mathcal{F}$ to be $\mathcal{N}(\mathbf{N}(\mathbb{R} \ltimes (\mathbb{R}^+)^{\mathbb{N}}))$.

In our setting, marked point processes are constructed in the following way. We start with a stationary point process in $\mathbb{R}$. Let $D_n = X_{n+1}-X_n$. We mark the point X_n with a sequence of i.i.d. random variables $K_n = \{L_{n,i}\}_{i\geq 1}$ that model failures as in Section 1.

We work with the point process under its Palm probability. Let $\lambda = \mathbb{E}[\Phi([0, 1] \times (\mathbb{R}^+)^{\mathbb{N}})]$ be the intensity of Φ. We assume $0 < \lambda < \infty$. The Palm probability of Φ is defined as, for all $A \in \mathcal{N}(\mathbf{N}(\mathbb{R} \ltimes (\mathbb{R}^+)^{\mathbb{N}}))$,

$$\mathbb{P}^0[A] = \frac{1}{\lambda|B|}\mathbb{E}\left[\sum_{n\in\mathbb{Z}} \mathbf{1}\{X_n \in B\}\mathbf{1}\{\Phi \circ \theta_{X_n} \in A\}\right], \tag{2.1}$$

for any $B \in \mathcal{B}(\mathbb{R})$ with positive Lebesgue measure $|B|$, where

$$\Phi \circ \theta_{X_n} = \{(X_m - X_n, K_m)\}_{m\in\mathbb{Z}}.$$

The probability measure $\mathbb{P}^0$ can be regarded as the distribution of the process given there is a point at the origin. In fact, $\mathbb{P}^0[0 \in \Phi] = 1$. For more on Palm probabilities, see [10], [11], [9], among others.

2.2 Point-shifts

To provide a unified definition of asymptotic efficiency for sequential `restart` and `checkpointing`, we resort to the theory of dynamics on point processes induced by point-shifts (for more on the subject, see [6] and [19]).

Define $\mathbf{N}^0(\mathbb{R}^+ \ltimes (\mathbb{R}^+)^{\mathbb{N}})$ as the subspace of $\mathbf{N}(\mathbb{R}^+ \ltimes (\mathbb{R}^+)^{\mathbb{N}})$ of all counting measures with mass at the origin. Let $\mathcal{N}(\mathbf{N}^0(\mathbb{R} \ltimes (\mathbb{R}^+))^{\mathbb{N}})$ be the corresponding trace $\sigma-$algebra.

Let $\theta : \mathbf{N}^0(\mathbb{R}^+ \ltimes (\mathbb{R}^+)^{\mathbb{N}}) \to \mathbf{N}^0(\mathbb{R}^+ \ltimes (\mathbb{R}^+)^{\mathbb{N}})$ be the discrete left-shift operator defined by

$$\theta\psi = \{(X_m - X_1, K_m)\}_{m\in\mathbb{Z}},$$

with $\theta^n\psi = \{(X_m - X_n, K_m)\}_{m\in\mathbb{Z}},\ n \in \mathbb{Z}$. Let $s : \mathbf{N}^0(\mathbb{R} \ltimes (\mathbb{R}^+)^{\mathbb{N}}) \to \mathbb{R}$ be a measurable function such that

$$s(\psi) = X_{\alpha_1}, \quad \text{where } \psi(\{X_{\alpha_1}, K_{\alpha_1}\}) \geq 1,$$

that is, s maps a counting measure to some element of its support. Such a map is called a *point-map*. A point-map s induces a *compatible point-shift*, S, that maps, in a translation invariant way, every point of a counting measure to another by

$$S(\psi, X_n) = s(\theta^n\psi) + X_n, \tag{2.2}$$

for all X_n in the support of ψ. Then, we define the translation by the point-shift s, $\theta_s : \mathbf{N}^0(\mathbb{R} \ltimes (\mathbb{R}^+)^{\mathbb{N}}) \to \mathbf{N}^0(\mathbb{R} \ltimes (\mathbb{R}^+)^{\mathbb{N}})$ as

$$\theta_s\psi := \{\psi - X_{\alpha_1}\} = \{(X_m - X_{\alpha_1}, K_m)\}_{m\in\mathbb{Z}}. \tag{2.3}$$

Inductively, assuming that $s^{n-1}(\psi)$ is defined and letting $\theta_s^{n-1}\psi = \{\psi - s^{n-1}(\psi)\}$, we let $s^n(\psi) = s(\theta_s^{n-1}\psi) + s^{n-1}(\psi)$ and $\theta_s^n(\psi) = \{\psi - s^n(\psi)\}$.

In words, s takes the counting measure and maps it to an element of its support, X_{α_1}. Then θ_s shifts the counting measure so that X_{α_1} is the origin. Applying s again to the shifted counting measure, we get some point on the support of ψ, say X_{α_2}, and θ_s^2 shifts ψ so that X_{α_2} is the origin, and so on.

2.3 Ideal times and actual times

As discussed in Section 1, in both the one task `restart` and `checkpointing` , we have the ideal time (when no failures take place) and the actual time (when accounting for failures). In our sequential models, we

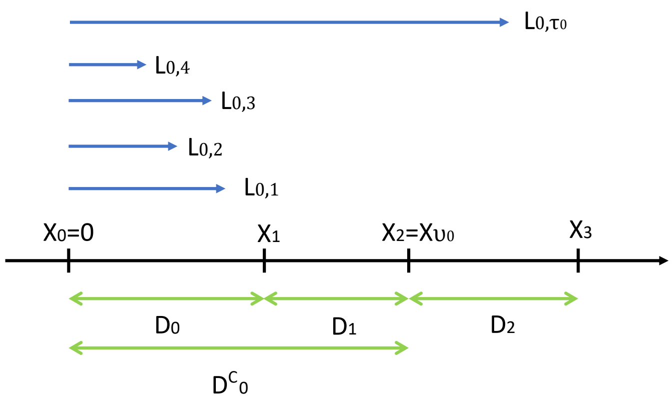

Figure 3: The progress of `checkpointing` at the $1^{st}-$iteration. There are four failures before the first checkpoint is surpassed, hence $\tau_0 = 5$. Then, there are no failures until the system is between checkpoints X_2 and X_3, so $\nu_1 = 2$. Here, the first ideal-time is $D_0^C = X_2 - X_0$ and the first actual time is $T_0^C = \sum_{i=1}^{5} L_{0,i}$.

have an ideal time and an actual time for each iteration. We define these using point-shifts.

First, let

$$\tau_n = \inf\{k \geq 1 : L_{n,k} > X_{n+1} - X_n\}. \tag{2.4}$$

The sequential `restart` point-map is given by $s_R(\Phi) = X_1$, so $S_R(X_n, \Phi) = X_{n+1}$ for all n. The translation by this point-map is simply the discrete left-shift operator, i.e., $\theta_{s_R}^n = \theta^n$. For the $n^{th}-$task, the ideal time is $D_n^R = D_n = X_{n+1} - X_n$ and the actual time is $T_n^R = \sum_{i=1}^{\tau_n - 1} L_{n,i} + D_n$.

The sequential `checkpointing` point-map is $s_C(\Phi) = X_{\nu_0}$, where

$$\nu_0 = \sup\{k \geq 1 : L_{0,\tau_0} \geq X_k\}. \tag{2.5}$$

Notice that $\tau_0 - 1$ is the number of failures before the first checkpoint is surpassed, and ν_0 the index of the next checkpoint secured once the system passes the first one. The $1^{st}-$ideal time is $D_0^C = X_{\nu_0}$ and the $1^{st}-$ actual time is $T_0^C = \sum_{i=1}^{\tau_0} L_{0,i}$. Figure 3 illustrates the first iteration in sequential `checkpointing`. Now, for each n, let

$$Z_n = L_{n,\tau_n} - D_n. \tag{2.6}$$

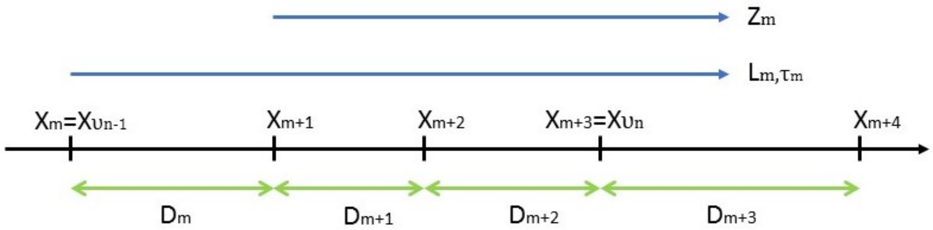

Figure 4: The progress of `checkpointing` at the $n^{th}-$iteration.

More generally, as illustrated in Figure 4, at the $n^{th}-$iteration, the $n^{th}-$ideal time is $D_n^C := X_{\nu_n} - X_{\nu_{n-1}}$, where

$$\nu_n = \sup\{k \geq \nu_{n-1} + 1 : Z_{\nu_{n-1}} > X_k - X_{\nu_{n-1}+1}\}, \tag{2.7}$$

with the $n^{th}-$actual time being $T_n^C = \sum_{i=1}^{\tau_{\nu_{n-1}}} L_{\nu_{n-1},i}$. We set $\nu_{-1} = 0$.

The table below summarizes our notation.

	`restart`	`checkpointing`
Point-map	$s_R(\Phi) = X_1$	$s_C(\Phi) = X_{\nu_0}$
$n^{th}-$ideal time	$D_n^R = X_{n+1} - X_n$	$D_n^C = X_{\nu_n} - X_{\nu_{n-1}}$
$n^{th}-$actual time	$T_n^R = \sum_{i=1}^{\tau_n - 1} L_{n,i} + D_n$	$T_n^C = \sum_{i=1}^{\tau_{\nu_{n-1}}} L_{\nu_{n-1},i}$
Point-map translation	$\theta_{s_R}^n = \theta^n$	$\theta_{s_C}^n = \theta^{\nu_n}$

2.4 Asymptotic efficiency

In this unified framework, we define asymptotic efficiency as the limit ratio of the sum of ideal times to the sum of actual times for both models.

Definition 5 (Asymptotic Efficiency)**.** The asymptotic efficiency is given by, for $i \in \{R, C\}$,

$$e(\omega) = \lim_{N \to \infty} \frac{\sum_{n=0}^{N-1} D_n^i(\omega)}{\sum_{n=0}^{N-1} T_n^i(\omega)} \ \mathbb{P}^0 - \text{a.s.}, \tag{2.8}$$

whenever the limit exists.

Notice that, when it exists, $0 \leq e \leq 1$ $\mathbb{P}^0-$a.s., as $T_n^i \geq D_n^i$ $n \geq 0$.

Let s be a point-map. Consider the sequence of probability measures on $(\mathbf{N}^0(\mathbb{R} \ltimes (\mathbb{R}^+)^{\mathbb{N}}), \mathcal{N}^0(\mathbb{R} \ltimes (\mathbb{R}^+)^{\mathbb{N}})$ defined by

$$\mathbb{P}^{s,n} = \mathbb{P}^0 \circ (\theta_s^n)^{-1}. \tag{2.9}$$

Then $\mathbb{P}^{s,n}$ can be interpreted as the distribution of the point process given that there is a point of the $n^{th}-$image of S at the origin. Suppose $\{\mathbb{P}^{s,n}\}_{n\geq 0}$ has a weak limit $\mathbb{P}^{s,\infty}$. As we shall see in detail, when the asymptotic efficiency exists, it is then the ratio of the expectations of D_0^i and T_0^i under $\mathbb{P}^{s_i,\infty}$, $i \in \{R, C\}$.

2.5 General Assumptions

In order to establish the existence of e in sequential `restart` we assume that the marked point process Φ is such that, under $\mathbb{P}^0$,

1. the i.i.d. sequence of failure marks, $\{L_{n,i}\}_{i\geq 1}$ is independent of D_n for all $n \geq 0$;
2. both D_0 and $L_{0,1}$ have right-unbounded support;
3. $\mathbb{E}^0[D_0]$, $\mathbb{E}^0[L_{0,1}] < \infty$.

This set of assumptions allows us to leverage the results of Section 1.

For `checkpointing` , besides items 1., 2., and 3. above, we assume that Φ is a marked renewal process, i.e., under $\mathbb{P}^0$, $\{D_n\}_{n\in\mathbb{Z}}$ is i.i.d.. Moreover, we assume that $\{L_{n,i}\}_{i\geq 1}$ is independent of D_m for all $m \geq n \geq 0$. In words, the failure marks of X_n are independent of the checkpoint intervals ahead.

3 Sequential `restart`: main result

Given that the translation by the sequential `restart` point-map is the discrete left-shift operator θ, the sequence $\{\mathbb{P}^{s_R,n}\}_{n\geq 0}$ (Equation (2.9)) is constant, with all its elements being equal to $\mathbb{P}^0$. This result holds as θ is bijective, so it preserves the Palm measure [20].

Moreover, from the fact that θ preserves $\mathbb{P}^0$, there exists a random variable L_0 such that

$$\mathbb{P}^0[L_0 > t] = \mathbb{P}^0[L_{n,i} > t], \quad \forall\, i, n \in \mathbb{N} \text{ and } t \in \mathbb{R}^+. \tag{3.1}$$

In the same vein, the sequence $\{D_n\}_{n\geq 0}$ is identically distributed (but not necessarily i.i.d.) under $\mathbb{P}^0$. Consequently, the `restart` actual time sequence, $\{T_n^R\}_{n\geq 0}$ is also identically distributed under $\mathbb{P}^0$.

Theorem 6. Let $\mathcal{I}$ be the invariant $\sigma-$algebra of $(\mathbb{P}^0, \theta)$. If D_0 does not have a $\mathbb{P}^0-$tail heavier than L_0, the asymptotic efficiency exists and it is given by the random variable

$$e = \frac{\mathbb{E}^0[D_0|\mathcal{I}]}{\mathbb{E}^0[T_0^R|\mathcal{I}]},\ \mathbb{P}^0 - \text{a.s.}. \tag{3.2}$$

If $(\mathbb{P}, \{\theta_t\}_{t\in\mathbb{R}})$ or, equivalently $(\mathbb{P}^0, \theta)$ is ergodic, we have $\mathbb{E}^0[D_0|\mathcal{I}] = \mathbb{E}^0[D_0]$ and $\mathbb{E}^0[T_0^R|\mathcal{I}] = \mathbb{E}^0[T_0^R]$, so that the asymptotic efficiency is constant. In this case, if D_0 does have a $\mathbb{P}^0-$tail heavier than L_0, $e = 0$ $\mathbb{P}^0 -$ a.s.

Proof. If D_0 does not have a $\mathbb{P}^0-$tail heavier than L_0, by Theorem 2, $\mathbb{E}^0[T_0^R] < \infty$, and, as $\mathbb{E}^0[D_0] < \infty$ by assumption, by Birkhoff's Pointwise Ergodic Theorem,

$$e = \lim_{N\to\infty} \frac{\sum_{n=0}^{N-1} D_n^i(\omega)}{\sum_{n=0}^{N-1} T_n^i(\omega)} = \lim_{N\to\infty} \frac{\sum_{n=0}^{N-1} D_0 \circ \theta^n}{\sum_{n=0}^{N-1} T_0^R \circ \theta^n} = \frac{\mathbb{E}^0[D_0|\mathcal{I}]}{\mathbb{E}^0[T_0^R|\mathcal{I}]},\ \mathbb{P}^0 - \text{a.s.}$$

When $(\mathbb{P}, \{\theta\}_{t\in\mathbb{R}})$ is ergodic, $\mathbb{E}^0[D_0|\mathcal{I}]$ (resp. $\mathbb{E}^0[T_0^R|\mathcal{I}]$) equals $\mathbb{E}^0[D_0]$ (resp. $\mathbb{E}^0[T_0^R]$) If D_0 does have a $\mathbb{P}^0-$tail heavier than L_0, $\mathbb{E}^0[T_n^R] = \infty$ for all $n \geq 0$. Suppose, by contradiction, that

$$\limsup_{N\to\infty} \frac{\sum_{n=0}^{N-1} D_n}{\sum_{n=0}^{N-1} T_n^R} > \epsilon\ \mathbb{P}^0 - \text{a.s.}$$

for some $\epsilon > 0$. It follows that

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} D_n > \epsilon \limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T_n^R.$$

Let $M > 0$ be a fixed integer. Then, as

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T_n^R > \limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \min\{T_n^R, M\},$$

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} D_n > \epsilon \limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \min\{T_n^R, M\}.$$

Now $\min\{T_n^R, M\}$ is integrable, so by by Birkhoff's Pointwise Ergodic Theorem, we have

$$\mathbb{E}^0[D_0] > \epsilon \mathbb{E}^0[\min\{T_0^R, M\}]. \tag{3.3}$$

Since $\mathbb{E}^0[D_0] < \infty$ and $\mathbb{E}^0[T_0^R] = \infty$, letting $M \to \infty$ on the RHS of (3.3) we have a contradiction. □

Remark 7. Notice that when $(\mathbb{P}, \{\theta_t\}_{t\in\mathbb{R}})$ is not ergodic, e can be zero with positive probability. Here is a simple example. Consider a stationary marked renewal process constructed in the following way. Let $D_0 \sim \exp(\lambda_d)$ under $\mathbb{P}^0$, and let c be a random variable taking values in $\{0, 1\}$ with, $\mathbb{P}^0[c = 0] = \mathbb{P}^0[c = 1] > 0$. Then, if $c = 0$, $L_0 \sim \exp(\lambda_1)$ under $\mathbb{P}^0$ and, otherwise $L_0 \sim \exp(\lambda_2)$. Assume $\lambda_1 \geq \lambda_d > \lambda_2$. Then $e = 0$ with probability $\mathbb{P}^0[c = 0]$.

Remark 8. Now since the set

$$A := \left\{ \lim_{N\to\infty} \frac{\frac{1}{N}\sum_{n=0}^{N-1} D_0 \circ \theta^n}{\frac{1}{N}\sum_{n=0}^{N-1} T_0^R \circ \theta^n} = \frac{\mathbb{E}^0[D_0|\mathcal{I}]}{\mathbb{E}^0[T_0^R|\mathcal{I}]} \right\}$$

is strictly $\theta-$invariant, i.e., $\theta A = A$, by property 1.6.1 in [7], $\mathbb{P}^0[A] = 1$ implies $\mathbb{P}[A] = 1$. Therefore, the above result also holds $\mathbb{P}-$a.s..

4 Sequential `checkpointing`: main results

In what follows:

- Φ satisfies the general assumptions for `checkpointing` in Section 2.5;
- $D_{\nu_n} = X_{\nu_n} - X_{\nu_n - 1}$, as illustrated in Figure 4, with ν_n defined in (2.7);
- $\mathbb{P}^{s_C,n} := \mathbb{P}^0 \circ (\theta^{\nu_n})^{-1}$, with $\{\mathbb{P}_+^{s_C,n}\}_{n\geq 0}$ being the restriction of $\mathbb{P}^{s_C,n}$ to $\mathbf{N}^0(\mathbb{R}^+ \ltimes (\mathbb{R}^+)^{\mathbb{N}})$.
- $\mathbb{P}_+^{s_C,\infty}$ denotes the weak limit of the sequence of distributions $\{\mathbb{P}_+^{s_C,n}\}_{n\geq 0}$, when it exists, and $\mathbb{E}_+^{s_C,\infty}$ is the expectation operator of $\mathbb{P}_+^{s_C,\infty}$.
- Assuming $\{D_{\nu_n}\}_{n\geq 0}$ converges weakly under the Palm distribution to a non-degenerate random variable D_∞ and letting $\hat{D}_\infty$ be an independent random variable distributed like the Palm distribution of D_∞, we set

$$\tau_\infty = \inf\{k \geq 1 : L_{0,k} > \hat{D}_\infty\} \tag{4.1}$$

and

$$\nu_\infty = \sup\{k \geq 1 : L_{0,\tau_\infty} \geq X_k\}. \tag{4.2}$$

Remark 9. Let $\mathbb{P}_+^0$ be the restriction of $\mathbb{P}^0$ to $\mathbf{N}^0(\mathbb{R}^+ \ltimes (\mathbb{R}^+)^{\mathbb{N}})$. Then, under the assumptions of Section 2.5, $\mathbb{P}_+^0$ is an independently marked renewal process and, therefore, satisfies the strong Markov property. In this section, all events consider under $\mathbb{P}^0$ belong to the trace $\sigma-$algebra $\mathcal{N}(\mathbf{N}^0(\mathbb{R}^+ \ltimes (\mathbb{R}^+)^{\mathbb{N}})$. Hence, we keep the notation $\mathbb{P}^0$ when there is no ambiguity.

Theorem 10. If $\{D_{\nu_n}\}_{n\geq 0}$ converges weakly under its Palm distribution to a non-degenerate random variable D_∞, then $\mathbb{P}_+^{s_C,\infty}$ exists.

Moreover, if $\mathbb{E}^0[D_\infty]$, $\mathbb{E}^0[\nu_\infty] < \infty$, and D_∞ does not have a $\mathbb{P}^0-$tail heavier than $L_{0,1}$, then the asymptotic efficiency exists and it is equal to

$$e = \frac{\mathbb{E}_+^{s_C,\infty}[D_0^C]}{\mathbb{E}_+^{s_C,\infty}[T_0^C]} \;\; \mathbb{P}^0 - \text{a.s.},$$

Otherwise, $e = 0$ $\mathbb{P}^0-$a.s..

For the sake of brevity, in what comes next, we denote the sequence of failure marks $\{L_{n,i}\}_{i\geq 1}$ by $\mathbf{L}_n$.

Lemma 11. Under $\mathbb{P}^0$, consider the filtration $\{\mathcal{F}_m\}_{m\geq 0}$ in which

$$\mathcal{F}_m = \sigma((D_0, \mathbf{L}_0), \ldots, (D_m, \mathbf{L}_m)) \;\forall\; m.$$

Then, for all $n \geq 0$, $\nu_n + 1$ is a stopping time with respect to $\{\mathcal{F}_m\}_{m\geq 0}$.

Proof. As $\mathbf{L}_n$ is independent of $\{D_m\}_{m\geq n}$, the stopping time property of $\nu_0 + 1$ follows from the fact that τ_0 is $\mathcal{F}_0-$measurable, $\nu_0 > 0$ a.s., and, for all $m \geq 1$,

$$\{\nu_0 + 1 = m\} = \{X_m \geq L_{0,\tau_0}\} \cap \{X_{m-1} < L_{0,\tau_0}\} \subset \mathcal{F}_m.$$

Now suppose ν_k+1 is a stopping time. Then, for $m \leq k$, $\{\nu_{k+1}+1 = m\} = \emptyset$, and, for $m > k$,

$$\{\nu_{k+1} + 1 = m\} = \cup_{i=1}^{m-1}(\{\nu_k + 1 = i\} \cap \{X_{m+1} < X_i + Z_{i-1} \leq X_m\}) \subset \mathcal{F}_m,$$

Z_{i-1} defined in (2.6). □

Let $\{\overline{D}_n\}_{n\in\mathbb{Z}}$ be a sequence of i.i.d random variables, independent of $\{D_n\}_{n\in\mathbb{Z}}$ such that $\overline{D}_n$ has the same distribution as D_0. Let the total lifetime of the renewal process $\{\overline{D}_n\}_{n\in\mathbb{Z}}$ be

$$\beta(t) = \overline{D}_n \text{ if } X_n < t \leq X_{n+1}. \tag{4.3}$$

Given $Z_0 = z$, we have $D_{\nu_0} = \beta(z)$. Therefore,

$$\mathbb{P}^0[D_{\nu_0} > x] = \int_0^\infty \mathbb{P}^0[\beta(t) > x] f_{Z_0}(dt),$$

where f_{Z_0} is the distribution of Z_0.

The interval D_{ν_0} tends to be larger than D_0, as failures are more likely to happen when checkpoints are more apart. In fact, D_{ν_0} stochastically dominates D_0, as $\mathbb{P}^0[\beta(t) > x] \geq \mathbb{P}^0[D_0 > x]$ for all $x, t \in \mathbb{R}^+$ [1]. This is an incarnation of the inspection paradox. Hence, in contrast with `restart`, θ_{s_C} does not preserve $\mathbb{P}^0$ and, consequently, $\{D_n^C\}_{n\geq 0}$ is not identically distributed under the Palm measure.

A sequence of inter-arrivals $\{\tilde{D}_n\}_{n\geq 0}$ is called a delayed renewal process if $\{\tilde{D}_n\}_{n\geq 0}$ is a sequence of independent and non-negative random variables and $\{\tilde{D}_n\}_{n\geq 1}$ is i.i.d.. In Proposition 12 below, we show that not only $\mathbb{P}_+^{s_C,n}$ is the distribution of an independently marked delayed renewal process, but also the distribution of the inter-arrivals after the first one is the same under $\mathbb{P}_+^{s_C,n}$ and $\mathbb{P}^0$. The result goes along with the interpretation of $\mathbb{P}^{s_C,n}$ as the distribution of the point process given there is a point of the $n^{th}-$iteration of the point-shift S_C at the origin. To illustrate our case, consider the point process shifted by θ^{ν_0}. As mentioned above, there is an inspection paradox effect in first interval D_{ν_0}. Nonetheless, as shown below, the inter-arrivals distributions $(X_{\nu_0+2} - X_{\nu_0+1})$, $(X_{\nu_0+3} - X_{\nu_0+2}), \ldots, (X_{\nu_0+j} - X_{\nu_0+j}), \ldots,$ are i.i.d. and have the same distribution under $\mathbb{P}^0$. This takes place in every iteration: the first inter-arrival interval after the shift is biased and the following ones maintain their distribution, which is that of a typical inter-arrival.

Proposition 12. For all $n \geq 0$, $\{\mathbb{P}_+^{s_C,n}\}_{n\geq 1}$ is the distribution of an independently marked delayed renewal process. Moreover,

$$\begin{aligned}&\mathbb{P}_+^{s_C,n}[D_1 \in A_1, \mathbf{L}_1 \in B_1, \ldots, D_m \in A_m, \mathbf{L}_m \in B_m, \ldots]\\&= \mathbb{P}^0[D_1 \in A_1, \mathbf{L}_1 \in B_1, \ldots, D_m \in A_m, \mathbf{L}_m \in B_m, \ldots]\end{aligned}$$

for all $\{A_i\}_{i\geq 1} \in \mathcal{B}(\mathbb{R}^+)$ and $\{B_i\}_{i\geq 1} \in \mathcal{B}((\mathbb{R})^{\mathbb{N}})$.

Proof. For $A_0, \ldots, A_j \in \mathcal{B}(\mathbb{R}^+)$ and $B_0, \ldots, B_j \in \mathcal{B}((\mathbb{R}^+)^{\mathbb{N}}))$,

$$\begin{aligned}&\mathbb{P}^{s_C,n}[D_j \in A_j, \mathbf{L}_j \in B_j, \ldots, D_0 \in A_0, \mathbf{L}_0 \in B_0]\\&= \mathbb{P}^0[D_{\nu_n+j} \in A_j, \mathbf{L}_{\nu_n+j} \in B_j, \ldots, D_{\nu_n} \in A_0, \mathbf{L}_{\nu_n} \in B_0]\\&= \mathbb{P}^0\left[D_{\nu_n+j} \in A_j, \mathbf{L}_{\nu_n+j} \in B_j | D_{\nu_n+j-1} \in A_{j-1}, \mathbf{L}_{\nu_n+j-1} \in B_{j-1},\right.\\&\quad \left.\ldots, D_{\nu_n} \in A_0, \mathbf{L}_{\nu_n} \in B_0\right]\\&\times \mathbb{P}^0[D_{\nu_n+j-1} \in A_{j-1}, \mathbf{L}_{\nu_n+j-1} \in B_{j-1} \ldots, D_{\nu_n} \in A_0, \mathbf{L}_{\nu_n} \in B_0].\end{aligned}$$

As $\nu_n + 1$ is a stopping time, by the strong Markov property of independently marked renewal processes, for every $j > 0$,

$$\mathbb{P}^0\left[D_{\nu_n+j} \in A_j, \mathbf{L}_{\nu_n+j} \in B_j | D_{\nu_n+j-1} \in A_{j-1}, \mathbf{L}_{\nu_n+j-1} \in B_{j-1}, \ldots, D_{\nu_n} \in A_0, \mathbf{L}_{\nu_n} \in B_{j-1}\right] = \mathbb{P}^0[D_j \in A_j]\mathbb{P}^0[\mathbf{L}_j \in B_j].$$

By keeping conditioning and applying the strong Markov property:

$$\begin{aligned} &\mathbb{P}^{s_C,n}[D_j \in A_j, \mathbf{L}_j \in B_j, \ldots, D_0 \in A_0, \mathbf{L}_0 \in B_0] \\ &= \mathbb{P}^0[D_{\nu_n} \in A_0, \mathbf{L}_{\nu_n} \in B_0] \prod_{i=1}^{j} \mathbb{P}^0[D_0 \in A_i]\mathbb{P}^0[\mathbf{L}_0 \in B_j]) \\ &= \mathbb{P}^0[D_{\nu_n} \in A_0]\mathbb{P}^0[\mathbf{L}_0 \in B_0] \prod_{i=1}^{j} \mathbb{P}^0[D_0 \in A_i]\mathbb{P}^0[\mathbf{L}_0 \in B_j], \end{aligned} \tag{4.4}$$

where the last equality follows from independent marking. □

Corollary 13. If $\{D_{\nu_n}\}_{n\geq 0}$ converges weakly under the Palm distribution to a non-degenerate random variable D_∞, then $\{\mathbb{P}_+^{s_C,n}\}_{n\geq 1}$ converges weakly to a distribution $\mathbb{P}_+^{s_C,\infty}$. Moreover, $\mathbb{P}_+^{s_C,\infty}$ is the distribution of an independently marked delayed renewal process in which the first inter-arrival interval is distributed as D_∞.

Proof. The results follow from Proposition 12 and taking the limit as $n \to \infty$ in (4.4). □

Lemma 14. For all $n \geq 1$,

$$\mathbb{P}^0[D_{\nu_n} > x] = \int_0^\infty \mathbb{P}^0[\beta(t) > x] f_{Z_{\nu_{n-1}}}(dt), \tag{4.5}$$

where $f_{Z_{\nu_{n-1}}}$ is the distribution of $Z_{\nu_{n-1}}$ under $\mathbb{P}^0$, with $\beta(t)$ defined in (4.3).

Proof. As $\mathbb{P}^0[D_{\nu_n} > x] = \mathbb{P}_+^{s_C,n-1}[D_{\nu_0} > x]$ and $\mathbb{P}_+^{s_C,n-1}$ is the distribution of a independently delayed renewal process such that $\{(D_i, \mathbf{L}_i)\}_{i\geq 1}$ has the same distribution under $\mathbb{P}_+^{s_C,n-1}$ and $\mathbb{P}^0$, we have

$$\mathbb{P}_+^{s_C,n-1}[D_{\nu_0} > x] = \int_0^\infty \mathbb{P}^0[\beta(t) > x] f_{Z_0}^{n-1}(dt),$$

where $f_{Z_0}^{n-1}$ is the distribution of Z_0 under $\mathbb{P}_+^{s_C,n-1}$. As $f_{Z_0}^{n-1} = f_{Z_{\nu_{n-1}}}$, the result follows. □

Remark 15. So far, we have defined $\{\mathbb{P}_+^{s_C,n}\}_{n\geq 0}$ and $\mathbb{P}_+^{s_C,\infty}$ (when it exists) on the space of counting measures. Once it is established that these distributions are concentrated on independently marked delayed renewal processes, we can, without loss of generality, define these measures on the space of discrete sequences in which each term belongs to $\mathbb{R}^+ \times (\mathbb{R}^+)^{\mathbb{N}}$, equipping it with the standard cylindrical Borel $\sigma-$algebra. We work on this space in the next proposition.

Proposition 16. If $\{D_{\nu_n}\}_{n\geq 0}$ converges weakly to a non-degenerate random variable, θ^{ν_0} preserves $\mathbb{P}_+^{s_C,\infty}$ and $(\mathbb{P}_+^{s_C,\infty}, \theta^{\nu_0})$ is mixing.

Proof. First we show θ^{ν_0} preserves $\mathbb{P}_+^{s_C,\infty}$. Consider the product cylinder set

$$C_{j_0,\ldots,j_l} := \{(D_n, \mathbf{L}_n)_{n\geq 0} : D_{j_0} \in A_{j_0}, \mathbf{L}_{j_0} \in B_{j_0},\ \ldots,\ D_{j_l} \in A_{j_l}, \mathbf{L}_{j_l} \in B_{j_l}\},$$

where $0 \geq j_0 > j_1, \ldots > j_l \in \mathbb{N}_+$.

$$\begin{aligned}
&\mathbb{P}_+^{s_C,n}[\theta^{\nu_0} C_{j_0,\ldots,j_l}] \\
&= \mathbb{P}^0\left[\theta^{\nu_0}\{(D_n, \mathbf{L}_n)_{n\geq 0} : D_{\nu_n+j_0} \in A_{j_0}, \mathbf{L}_{\nu_n+j_0} \in B_{j_0},\ \ldots,\right. \\
&\quad \left. D_{\nu_n+j_l} \in A_{j_l}, \mathbf{L}_{\nu_n+j_l} \in B_{j_l}\}\right] \\
&= \mathbb{P}^{s_C,1}\left[\{(D_n, \mathbf{L}_n)_{n\geq 0} : D_{\nu_n+j_0} \in A_{j_0}, \mathbf{L}_{\nu_n+j_0} \in B_{j_0},\ \ldots,\right. \\
&\quad \left. D_{\nu_n+j_l} \in A_{j_l}, \mathbf{L}_{\nu_n+j_l} \in B_{j_l}\}\right] \\
&= \mathbb{P}^0\left[\{(D_n, \mathbf{L}_n)_{n\geq 0} : D_{\nu_{n+1}+j_0} \in A_{j_0}, \mathbf{L}_{\nu_{n+1}+j_0} \in B_{j_0},\ \ldots,\right. \\
&\quad \left. D_{\nu_{n+1}+j_l} \in A_{j_l}, \mathbf{L}_{\nu_{n+1}+j_l} \in B_{j_l}\}\right].
\end{aligned}$$

Then, as $\{D_{\nu_n}\}_{n\geq 0}$ converges weakly to D_∞, by (4.4), taking the limit as $n \to \infty$ on both sides,

$$\mathbb{P}_+^{s_C,\infty}[\theta^{\nu_0} C_{j_0,\ldots,j_l}] = \mathbb{P}_+^{s_C,\infty}[C_{j_0,\ldots,j_l}]. \tag{4.6}$$

By standard extension arguments from product cylinder sets, we conclude that θ^{ν_0} preserves $\mathbb{P}_+^{s_C,\infty}$.

Next, we prove $(\mathbb{P}_+^{s_C,\infty}, \theta^{\nu_0})$ is mixing. First, we notice that D_{ν_n} is a function of $D_{\nu_{n-1}}$, $\{L_{0,\nu_{n-1}}\}_{i\geq 0}$, and $\{D_n\}_{n\geq \nu_{n-1}+1}$. By independent marking the i.i.d. sequence $\{L_{0,\nu_{n-1}}\}_{i\geq 0}$ is independent of ν_{n-1} and has the same distribution under $\mathbb{P}^0$ as $\{L_{0,i}\}_{i\geq 0}$. In the same way, by the strong Markov property the i.i.d. sequence $\{D_n\}_{n\geq \nu_{n-1}+1}$ is independent of ν_{n-1} and has the same distribution under $\mathbb{P}^0$ as $\{D_n\}_{i\geq 0}$. Therefore, $\{D_{\nu_n}\}_{n\geq 0}$ is a Markov Chain.

Let $C_{j_0,\dots,j_l}$ and $C_{j'_0},\dots,C_{j'_q}$ be two product cylinder sets. Following the same steps used to get (4.6), for all $m \geq 1$,

$$\begin{aligned}&\mathbb{P}_+^{s_C,\infty}[C_{j_0,\dots,j_l} \cap \theta_{s_C}^m C_{j'_0,\dots,j'_q}]\\ &= \lim_{n\to\infty} \mathbb{P}^0\left[D_{\nu_n+j_0} \in A_{j_0}, \mathbf{L}_{\nu_n+j_0} \in B_{j_0}, \dots, D_{\nu_n+j_l} \in A_{j_l}, \mathbf{L}_{\nu_n+j_l} \in B_{j_l} \cap\right.\\ &\left. D_{\nu_{n+m}+j'_0} \in A_{j'_0}, \mathbf{L}_{\nu_{n+m}+j'_0} \in B_{j'_0}, \dots, D_{\nu_{n+m}+j'_q} \in A_{j'_q}, \mathbf{L}_{\nu_{n+m}+j'_q} \in B_{j'_q}\right]\end{aligned} \tag{4.7}$$

Then, for all m such that $\nu_{n+m} + j'_0 > \nu_{n+1} + j_l$, as $\{D_{\nu_n}\}_{n\geq 0}$ is a Markov chain, (4.7) equals to

$$\begin{aligned}&\lim_{n\to\infty} \mathbb{P}^0\left[D_{\nu_n+j_0} \in A_{j_0}, \mathbf{L}_{\nu_n+j_0} \in B_{j_0}, \dots, D_{\nu_n+j_l} \in A_{j_l}, \mathbf{L}_{\nu_n+j_l} \in B_{j_l}\right]\\ &\times \mathbb{P}^0\left[D_{\nu_{n+m}+j'_0} \in A_{j'_0}, \mathbf{L}_{\nu_{n+m}+j'_0} \in B_{j'_0}, \dots, D_{\nu_{n+m}+j'_q} \in A_{j'_q}, \mathbf{L}_{\nu_{n+m}+j'_q} \in B_{j'_q}\right]\\ &= \mathbb{P}_+^{s_C,\infty}[C_{j_0,\dots,j_l}]\mathbb{P}_+^{s_C,\infty}[C_{j'_0,\dots,j'_q}].\end{aligned}$$

Again, invoking standard approximations arguments, we conclude that for all $C, C' \in \mathcal{B}((\mathbb{R}^+ \times (\mathbb{R}^+)^{\mathbb{N}})^{\mathbb{N}})$, $\lim_{m\to\infty} \mathbb{P}_+^{s_C,\infty}[C \cap \theta^{\nu_n} C'] = \mathbb{P}_+^{s_C,\infty}[C]\mathbb{P}_+^{s_C,\infty}[C']$, completing the proof. □

Lemma 17. Suppose $\{D_{\nu_n}\}_{n\geq 0}$ converges weakly to a non-degenerate random variable. Let A be a strictly $\theta^{\nu_0}-$invariant event in $\mathcal{N}(\mathbf{N}^0(\mathbb{R}^+ \ltimes (\mathbb{R}^+)^{\mathbb{N}})$. If $\mathbb{P}_+^{s_C,\infty}[A] = 1$, then $\mathbb{P}^0[A] = 1$.

Proof. By Corollary 13, as A is $\theta^{\nu_0}-$invariant, i.e., $\theta^{\nu_0} A = A$, and, hence, for all n, $\theta^{\nu_n} A = A$,

$$1 = \mathbb{P}_+^{s_C,\infty}[A] = \lim_{n\to\infty} \mathbb{P}^0[\theta^{\nu_n} A] = \lim_{n\to\infty} \mathbb{P}^0[A] = \mathbb{P}^0[A].$$

□

Proof of Theorem 10. As $\{D_{\nu_n}\}_{n\geq 0}$ converges in distribution to a non-degenerate random variable, by Corollary 13, $\mathbb{P}_+^{s_C,\infty}$ exists. Moreover, by Proposition 16, $(\mathbb{P}_+^{s_C,\infty}, \theta^{\nu_0})$ is mixing. Then, by Birkhoff's pointwise ergodic theorem, for any measurable function $h : \mathbf{N}^0(\mathbb{R}^+ \times (\mathbb{R}^+)^{\mathbb{N}}) \to \mathbb{R}^+$ such that $h \in \mathcal{L}^1(\mathbb{P}_+^{s_C,\infty})$,

$$\lim_{n\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} h \circ \theta^{\nu_n} = \mathbb{E}_+^{s_C,\infty}[h],\ \mathbb{P}_+^{s_C,\infty} - \text{a.s.}$$

Next, notice that

$$\mathbb{E}_+^{s_C,\infty}[D_C^0] = \mathbb{E}_+^{s_C,\infty}\left[\sum_{n=0}^{\nu_0-1} D_n\right]$$
$$= \mathbb{E}^0[D_\infty \mathbf{1}\{\nu_\infty = 1\}] + \mathbb{E}^0\left[\sum_{n=1}^{\nu_\infty-1} D_n \mathbf{1}\{\nu_\infty > 1\}\right].$$

Now, as $\mathbb{E}^0[D_\infty] < \infty$, $\mathbb{E}^0[D_\infty \mathbf{1}\{\nu_\infty = 1\}] < \infty$. Moreover, as $\{D_n\}_{n\geq 0}$ is i.i.d. under $\mathbb{P}^0$,

$$\mathbb{E}^0\left[\sum_{n=1}^{\nu_\infty-1} D_n \mathbf{1}\{\nu_\infty \geq 1\}\right] \leq \mathbb{E}^0\left[\sum_{n=0}^{\nu_\infty} D_n\right] \leq \mathbb{E}^0\left[\sum_{n=0}^{\nu_\infty+1} D_n\right].$$

Following the same reasoning as in Lemma 11, $\nu_\infty + 1$ is a stopping time with respect to the natural filtration of $\{(D_n, \mathbf{L}_n)\}_{n\geq 0}$. Hence, by the general version of Wald's equality for stopping times,

$$\mathbb{E}^0\left[\sum_{n=0}^{\nu_\infty+1} D_n\right] = \mathbb{E}^0[\nu_\infty + 1]\mathbb{E}^0[D_0],$$

which is finite as $\mathbb{E}^0[\nu_\infty] < \infty$ by assumption.

It follows that $\mathbb{E}_+^{s_C,\infty}[D_C^0] < \infty$. Hence, by Birkhoff's pointwise ergodic theorem,

$$\lim_{n\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} D_0^C \circ \theta^{\nu_n} = \mathbb{E}_+^{s_C,\infty}[D_C^0],\ \mathbb{P}_+^{s_C,\infty} - \text{a.s.} \tag{4.8}$$

As $\mathbb{E}_+^{s_C,\infty}[T_0^C] = \mathbb{E}^0[\sum_{i=1}^{\tau_\infty} L_{0,i}]$, if we assume that D_∞ does not have a $\mathbb{P}^0-$tail heavier than $L_{0,1}$, $T_0^C \in \mathcal{L}^1(\mathbb{P}_+^{s_C,\infty})$, so

$$\lim_{n\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T_0^C \circ \theta^{\nu_n} = \mathbb{E}_+^{s_C,\infty}[T_0^C],\ \mathbb{P}_+^{s_C,\infty} - \text{a.s.} \tag{4.9}$$

Therefore, by (4.8) and (4.9),

$$e := \lim_{N\to\infty} \frac{\sum_{n=0}^{N-1} D_n^C}{\sum_{n=0}^{N-1} T_n^C} = \lim_{N\to\infty} \frac{\sum_{n=0}^{N-1} D_0^C \circ \theta^{\nu_n}}{\sum_{n=0}^{N-1} T_0^C \circ \theta^{\nu_n}} = \frac{\mathbb{E}_+^{s_C,\infty}[D_0^C]}{\mathbb{E}_+^{s_C,\infty}[T_0^C]}. \tag{4.10}$$

Let

$$A := \left\{ \lim_{N\to\infty} \frac{\sum_{n=0}^{N-1} D_0^C \circ \theta^{\nu_n}}{\sum_{n=0}^{N-1} T_0^C \circ \theta^{\nu_n}} = \frac{\mathbb{E}_+^{s_C,\infty}[D_0^C]}{\mathbb{E}_+^{s_C,\infty}[T_0^C]} \right\}.$$

Since

$$\left(\lim_{N\to\infty} \frac{\sum_{n=0}^{N-1} D_0^C \circ \theta^{\nu_n}}{\sum_{n=0}^{N-1} T_0^C \circ \theta^{\nu_n}} \right) \circ \theta^{\nu_0} = \lim_{N\to\infty} \frac{\sum_{n=1}^{N-1} D_0^C \circ \theta^{\nu_n}}{\sum_{n=1}^{N-1} T_0^C \circ \theta^{\nu_n}}$$

and $\mathbb{E}_+^{s_C}$ is preserved by θ^{ν_0}, A is strictly $\theta^{\nu_0}-$invariant. Then, by Lemma 17, (4.10) holds $\mathbb{P}^0$-a.s., as desired.

If D_∞ does have a $\mathbb{P}^0-$tail heavier than $L_{0,1}$, $\mathbb{E}_+^{s_C,\infty}[T_0^C] = \infty$. Then, the same argument used in Theorem 6 applies to show that and $e = 0$. □

In the rest of this section, we focus on the case of exponential failure marks, in which we can say more about $\mathbb{P}_+^{s_C,\infty}$.

4.1 Exponential failures and universal checkpoints

Suppose that $\mathbb{P}^0[L_{0,0} \le x] = 1 - e^{-\lambda x}$, $\lambda > 0$, i.e., the failure marks are exponentially distributed. We show the conditions of Theorem 10 are then satisfied if $\mathbb{E}^0[\nu_1] < \infty$, so e exists. We also show there is a sequence of checkpoints that will be activated regardless of the initial checkpoint from which we start the system. These are called *universal checkpoints.*

Theorem 18. Suppose that Φ is a independently marked renewal process with exponentially distributed failure marks. Moreover, assume that $\mathbb{E}^0[D_{\nu_1}]$, $\mathbb{E}^0[\nu_1] < \infty$. Then e is well-defined.

Proof. In order to apply Theorem 10, we need to show that $\mathbb{P}_+^{s_C,\infty}$ exists, $\mathbb{E}^0[D_\infty]$, and $\mathbb{E}^0[\nu_\infty] < \infty$. Due to the memoryless property of the exponential distribution, the sequence of random variables $\{D_{\nu_n}\}_{n\ge1}$ is identically distributed, and thus converges in distribution. To see this, notice that Z_n defined in (2.6) is exponentially distributed with parameter λ for all n. Consequently, from (4.5),

$$\mathbb{P}^0[D_{\nu_n} > x] = \int_0^\infty \mathbb{P}^0[\beta(t) > x] \lambda e^{-\lambda t} dt, \ \forall\ n \ge 0.$$

By same reasoning, ν_∞ has the same distribution under Palm as ν_j for $j \ge 1$, so $\mathbb{E}^0[\nu_1] = \mathbb{E}^0[\nu_\infty]$. □

From (2.2) we have $s_C(\Phi, X_n) = X_{\kappa_n}$, where

$$\kappa_n = \sup\{k \geq n : L_{n,\tau_n} \geq X_k\}, \tag{4.11}$$

where τ_n as in (2.6). Then $S_c^2(\Phi, X_n) = S_C(\Phi, X_{\kappa_n})$, and so on. Let $\mathcal{H}_n = \{S_c^j(\Phi, X_n)\}_{j\geq 1}$. In words, $\mathcal{H}_n$ corresponds to the sequence of checkpoints if the system starts at X_n. Notice that $\mathcal{H}_0 = \{X_{\nu_n}\}_{n\geq 1}$.

Accordingly, we call $\mathcal{H}_n$ the set of *active checkpoints* of X_n or the *checkpoint trajectory* of X_n. We also say that if $X_m \in \mathcal{H}_n$, then X_m is *activated* by X_n. By convention, $X_n \notin \mathcal{H}_n$. We allow n to be negative. For example, the process may start from X_{-2}.

Definition 19 (Universal Checkpoints)**.** Suppose X_m is such that for all $k < m$ there exists $j \geq 1$ such that $S_c^j(\Phi, X_k) = X_m$ or, equivalently for all $k < m$, $X_m \in \mathcal{H}_k$. Then X_m is a universal checkpoint.

As the name suggests, universal checkpoints, if they exist, are activated if we start the system from any checkpoint that precedes them.

Theorem 20. Suppose that Φ is a independently marked renewal process with exponentially distributed failure marks. Moreover, assume $\mathbb{E}^0[\nu_1] < \infty$. Then there exists a sequence of universal checkpoints.

Let

$$N_n = \#\{m \in \mathbb{Z} : m < n \text{ and } S(\Phi, X_m) > X_n\},$$

and notice that if $N_n = 0$, then X_n is a universal checkpoint. Then, Theorem 20 follows directly from the proposition below.

Proposition 21. Under the assumptions of Theorem 20, for all $k \geq 0$, the process $\{N_n\}_{n\in\mathbb{Z}}$ admits a subsequence $\{N_{n_l}\}_{l\in\mathbb{Z}}$ such that $N_{n_l} = k$.

Proof. Let κ_n as in (4.11). Since Φ is an indepedently marked renewal process, the sequence $\{\kappa_n\}_{n\in\mathbb{Z}}$ is identically distributed. In particular κ_n has the same distribution as $\nu_1 = \kappa_0$

Define the events $A_n = \{S(\Phi, X_{-n}) > 0\}$. Then

$$\sum_{n=1}^{\infty} \mathbb{P}^0[A_n] = \sum_{n=1}^{\infty} \mathbb{P}^0[\kappa_{-n} > n] = \sum_{n=1}^{\infty} \mathbb{P}^0[\nu_1 > n] = \mathbb{E}^0[\nu_1] < \infty.$$

Therefore, by the Borel-Cantelli Lemma, $\mathbb{P}^0[A_n \text{ i.o.}] = 0$. Hence $N_0 < \infty$ a.s.. By the same reasoning, we conclude that $N_n < \infty$ a.s. for all n.

Thanks to the memoryless property of the exponential distribution, the value of N_n only depends on N_{n-1}, i.e., $\{N_n\}_{n\in\mathbb{Z}}$ is a Markov Chain. Now suppose $N_{n-1} = k$. That means that there are k checkpoint trajectories that do not use X_{n-1} as a checkpoint. Given $D_n = t$, a trajectory that does not use X_{n-1} as a checkpoint will not activate X_n with probability $e^{-\lambda t}$, and will activate X_n with probability $1 - e^{-\lambda t}$.

Considering the checkpoint trajectory of X_{n-1} we have that, for $1 \le j \le k$,

$$
\begin{aligned}
&\mathbb{P}^0[N_n = j | N_{n-1} = k] \\
&\quad = \mathbb{P}^0[N_n = j | N_{n-1} = k \text{ and } X_{n-1} \in \mathcal{H}_n]\mathbb{P}^0[X_{n-1} \in \mathcal{H}_n] \\
&\quad + \mathbb{P}^0[N_n = j | N_{n-1} = k \text{ and } X_{n-1} \notin \mathcal{H}_n]\mathbb{P}^0[X_{n-1} \notin \mathcal{H}_n] \\
&\quad = \left(\int_0^\infty \binom{k}{j} (e^{-\lambda t})^j (1 - e^{-\lambda t})^{k-j} f_D(dt)\right)\left(\int_0^\infty (1 - e^{-\lambda t}) f_D(dt)\right) \\
&\quad + \left(\int_0^\infty \binom{k}{j-1} (e^{-\lambda t})^{j-1} (1 - e^{-\lambda t})^{k-j+1} f_D(dt)\right)\left(\int_0^\infty e^{-\lambda t} f_D(dt)\right).
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
&\mathbb{P}^0[N_n = k+1 | N_{n-1} = k] = \left(\int_0^\infty (e^{-\lambda t})^k f_D(dt)\right)\left(\int_0^\infty e^{-\lambda t} f_D(dt)\right), \\
&\mathbb{P}^0[N_n = 0 | N_{n-1} = k] \\
&\quad = \left(\int_0^\infty (1 - e^{-\lambda t})^k f_D(dt)\right)\left(\int_0^\infty (1 - e^{-\lambda t}) f_D(dt)\right).
\end{aligned}
$$

Since $\{N_n\}_{n\in\mathbb{Z}}$ is a sequence of marks of the stationary ergodic marked point process Φ, if $\mathbb{P}^0[N_0 = k] > 0$, there exists a sequence $\{n_l\}_{l\in\mathbb{Z}}$ such that $N_{n_l} = k$. Since $N_n < \infty$, for all n, by the computations above, $\mathbb{P}^0[N_0 = k] > 0$ for all $k \ge 0$. □

5 Extensions

5.1 Markov renewal processes

Markov Renewal Processes give us instances in which the asymptotic efficiency exists, even though the process does not start is not at steady state when tasks start being executed. For simplicity, here we work out sequential `restart` case, although the same results can be achieved within the sequential `checkpointing` with exponential failure marks. First, we develop the results of a system that admits a steady state but the initial distribution is arbitrary. After that, we indicate how one can fit the Markov Renewal structure in the setting of Theorem 6.

Here we consider different states for the ideal task times and failures marks, which are are driven by a Markov Chain whose state space is countable. We briefly review the Markov Renewal process structure, adapting it to account for failure marks. For a more complete treatment of the subject see [11] and [1]. In short, the distribution of the task sizes and failure marks depends on the current state and the next state to be visited of a Markov Chain.

Consider a Markov Chain $\{Y_n\}_{n\geq 0}$ whose space state S is countable. Let $\mathbf{A} = (a_i)_{i\in S}$ be its initial distribution and $\mathbf{P} = (p_{ij})_{ij\in S}$ its transition matrix.

For $x, y \in \mathbb{R}^+$, let $G_{0,i,j}(x, y)$ and $G_{i,j}(x, y)$ be two joint distributions on $\mathbb{R}^2$. We then consider the trivariate sequence $\{(Y_n, D_n, L_n)\}_{n\in\mathbb{N}}$ defined in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that

1. $\mathbb{P}[Y_0 = k_0] = a_{k_0}$;
2. $\mathbb{P}[D_0 \leq x, L_0 \leq y, Y_1 = k_1 | Y_0 = k_0] = G_{0,k_0,k_1}(x, y) p_{k_0 k_1}$;
3. For $n \geq 1$:

$$\begin{aligned}&\mathbb{P}[D_n \leq x, L_n \leq y, Y_n = k_n | Y_0 = k_0, D_0 \leq x_0, L_0 \leq y_0, \ldots,\\ &\ Y_n = k_n, D_{n-1} \leq x_{n-1}, L_{n-1} \leq y_{n-1}]\\ &= \mathbb{P}[D_n \leq x, L_n \leq y, Y_n = k_n | Y_{n-1} = k_{n-1}] = G_{k_{n-1}k_n}(x, y) p_{k_{n-1}k_n}.\end{aligned}$$

We further assume that $G_{0,i,j}(x, y) = F^D_{0,i,j}(x) F^L_{0,i,j}(y)$ and $G_{i,j}(x, y) = F^D_{i,j}(x) F^L_{i,j}(y)$, namely, conditional on (Y_n, Y_{n-1}), L_n is independent of D_n. Notice that letting $y \to \infty$ we have the classical Markov Renewal Process [14].

Assumption 2. We impose conditions so that $X_n \to \infty$ as $n \to \infty$ $\mathbb{P}-$a.s.: (i) $\mathbf{P}$ is irreducible, (ii) there exists a non-trivial probability measure $(\pi_i)_{i\in S}$ such that $\pi P = P$ and $\sum_{i\in S} \pi_i \mu_i < \infty$, where

$$\mu_i = \sum_{j\in S} p_{ij} \int_0^\infty x F^D_{ij}(dx).$$

We write $\hat{D}_{ij}$ (respectively $\hat{L}_{ij}$) the random variable corresponding to the task size (failure mark) conditional on $Y_n = i$ and $Y_{n+1} = j$. It follows that $\hat{D}_{ij}$ (resp. $\hat{L}_{ij}$) has distribution $F^D_{i,j}(x)$ (resp. $F^L_{i,j}(x)$).

Assumption 3. For all $i, j \in S$, $\hat{D}_{ij}$ and $\hat{L}_{ij}$ are integrable random variables with right-unbounded support. Moreover, $\hat{D}_{ij}$ is independent of $\hat{L}_{ij}$.

Assumption 4. The embedded Markov Chain $\{Y_n\}_{n\in\mathbb{N}}$ is ergodic, with unique stationary distribution given by $\{\pi_i\}_{i\in S}$.

Then, we proceed to derive expressions for $\mathbb{E}[D_0]$ and $\mathbb{E}[T_0^R]$ when the chain is in steady state. First, for each pair of states (i, j) one can compute $\mathbb{E}[\hat{T}_{ij}^R]$, the expected value of the actual time $\hat{T}_{ij}^R$, when $Y_n = i$ and $Y_{n+1} = j$, as in Theorem 2, with $L_0 = \hat{L}_{ij}$ and $D = \hat{D}_{ij}$. By unconditioning,

$$\mathbb{E}[\hat{D}_i] = \sum_{j\in S} p_{ij}\mathbb{E}[\hat{D}_{ij}] \text{ and } \mathbb{E}[T_i^R] = \sum_{j\in S} p_{ij}\mathbb{E}[\hat{T}_{ij}^R],$$

where $\mathbb{E}[\hat{D}_i]$ (resp. $\mathbb{E}[T_i^R]$) is the expected size of the task (resp. actual time) when the chain is at state i.

Remark 22. A pair of states (i, j) is called *slow* if $\hat{D}_{ij}$ has a $\mathbb{P}-$tail heavier than $\hat{L}_{ij}$. It follows from Assumption 3 if that (i, j) is a slow pair of states then $\mathbb{E}[T_i^R] = \infty$.

Due to the Strong Law of Large Numbers for functionals of a Markov renewal process [14],

$$\lim_{n\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} D_n = \sum_{i\in S} \pi_i \mathbb{E}[\hat{D}_i],$$

and, if $\mathbb{E}^0[T_i^R] < \infty$

$$\lim_{n\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T_n^R = \sum_{i\in S} \pi_i \mathbb{E}[T_i^R].$$

Hence, if there are no slow states:

$$e = \frac{\sum_{i\in S} \pi_i \mathbb{E}[\hat{D}_i]}{\sum_{i\in S} \pi_i \mathbb{E}[T_i^R]}.$$

The next example shows that e can be equal to 0 even in the absence of slow states.

Example 1. Suppose $\hat{D}_{ij} \sim \exp\left(p_{ij}2^j\pi_i 2^i\right)$ and $\hat{L}_{ij} \sim \exp\left(p_{ij}2^j\pi_i(2^i-1)\right)$ for all $i \in \mathbb{N} = S$. We assume $0 < \pi_i < 1$ for all $i \geq 1$. Notice that each pair of states is not slow. However, $\sum_{i=1}^{\infty} \pi_i \mathbb{E}[\hat{D}_i] = 1$ and $\sum_{i\in S} \pi_i \mathbb{E}[T_i^R] = \sum_{i=1}^{\infty} 1 = \infty$, so $e = 0$. This holds as, if $\hat{D}_{ij} \sim \exp(\beta_{ij})$ and $\hat{L}_{ij} \sim \exp(\alpha_{ij})$ we have, by Theorem 2,

$$\mathbb{E}[T_{ij}^R] = \frac{1}{(\beta_{ij} - \alpha_{ij})}.$$

We discuss briefly how Markov Renewal Process can be fit into the framework of Theorem 6. Define a point process $\hat{\Phi}$ by $X_0 = 0$ and $D_n = X_{n+1} - X_n$ and a Semi-Markov Process $\{Y(t)\}_{t\in\mathbb{R}}$ by $Y(t) = Y_n$, when $X_n \leq t < X_{n+1}$.

Assume that $\sum_{i\in S} \pi_i \mathbb{E}[D_i] < \infty$ and $\hat{\Phi}([0,t)) < \infty$ for all t. Then let $\mathbb{P}^0 = \mathbb{P} \circ Y^{-1}$. Then there exists a probability space endowed with a flow, $(\hat{\mathbb{P}}, \Omega, \mathcal{F}, \{\theta\}_t)$, such that $\hat{\mathbb{P}}$ is θ_t-invariant, $\hat{\Phi}$ is a stationary point process whose Palm distribution is $\mathbb{P}^0$ ([7], Chapter 1).

Hence, by marking this process with the failure marks $\{L_{n,i}\}_{i\geq 1}$, Theorem 6 holds with $\mathbb{E}^0[D_0] = \sum_{i\in S} \pi_i \mathbb{E}[\hat{D}_i]$ and $\mathbb{E}^0[T_0^R] = \sum_{i\in S} \pi_i \mathbb{E}[T_i^R]$.

5.2 Repetition of tasks based on a Random Walk

We now proceed to study the asymptotic efficiency of `restart` when the tasks to be completed follows a transient simple random walk, i.e., there is a chance $p < \frac{1}{2}$ that, once a task is completed, progress is lost and the system returns to the previous task. For example, after completing task D_m, progress might be lost and the system resumes from task D_{m-1}, which is again subject to failures. We assume that this extra failure source is independent of the failure marks and ideal times and Φ is a renewal process. We show that the asymptotic efficiency exists and we provide a lower bound for it.

Remark 23. In order to simplify the proofs and, without loss of generality, we do not model any sort of *boundary effect*, i.e., the tasks $D_{-1} = X_0 - X_{-1}$, $D_{-2} = X_{-1} - X_{-2}$ and so on are well-defined.

Let $\{\xi_n\}_{n\geq 1}$ be an i.i.d. sequence of Bernoulli random variables such that $\mathbb{P}[\xi_n = -1] = p$ and $\zeta_n = \sum_{i=1}^{n} \xi_i$ (with $\xi_0 = 0$). At the $n^{th}-$iteration, the task being completed is $D_n^W = D_0 \circ \theta^{\zeta_{n-1}}$.

We assume that the failure marks are such that $L_{m,0}$ is independent of $L_{n,0}$ for all $n, m \in \mathbb{Z}$ and are re-drawn if a task is repeated. Under this assumption, instead of marking each point of the process with a sequence of i.i.d. random variables $\{L_{n,i}\}_{i\geq 1}$, we can simplify matters by considering a single sequence of i.i.d. random variables $\{L_i\}_{i\geq 1}$, such that L_i has the same distribution as L_0 under $\mathbb{P}^0$ to compute the actual time spent on task D_n^W. Then let $\tau_0^W := \inf\{i \geq 1 : L_i > D_0^W\}$ and $\tau_n^W := \inf\{i > \tau_{n-1}^W : L_i > D_n^W\}$. So, the actual time spent on task D_n^W is given by $T_n^W = \sum_{i=\tau_{n-1}^W}^{\tau_n^W+1} L_i + D_n^W$.

Let $\vartheta_n = \inf\{j \geq 1 : \zeta_j = n\}$, i.e., $\{\vartheta_n\}_{n\geq 0}$ is the sequence of ladder epochs of the random walk ζ_n with $\vartheta_0 = 0$ [13]. It follows that $D_n = X_{n+1} - X_n = X_{\zeta_{\vartheta_{n+1}}} - X_{\zeta_{\vartheta_n}}$.

Given those definitions, one can regard $T_n^R = \sum_{j=\vartheta_{n-1}}^{\vartheta_n - 1} T_j^W$ as the total actual time necessary to complete the task D_n. As before, the asymptotic efficiency is

$$e_p = \lim_{N\to\infty} \frac{\sum_{n=0}^{N-1} D_n}{\sum_{n=0}^{N-1} T_n^R}, \quad \mathbb{P}^0 - \text{a.s.},$$

whenever the limit exists.

Using the fact that Φ is an independently marked renewal process, $\mathbb{E}^0[T_j^W] = \mathbb{E}^0[T_l^W]$ for all $l, j \geq 0$, as the Palm expectation of T_j^W only depends on the distribution of L_0 and D_0 under $\mathbb{P}^0$.

Proposition 24. Suppose D_0 does not have a $\mathbb{P}^0-$tail heavier than L_0. Then

$$e \geq e_p = \frac{\gamma}{\rho} e,$$

where e is the asymptotic efficiency when $p = 0$, ρ is the probability that the random walk $\{\zeta_n\}_{n\geq 0}$ never returns to zero and γ is the probability that the walk never goes below zero. On the other hand, if D_0 has a $\mathbb{P}^0-$tail heavier than L_0, $e_p = 0$.

To prove Proposition 24 we rely on the next lemmas, following a similar reasoning as in [16].

We say that υ is a regeneration epoch if it is a ladder epoch and, moreover, $\zeta_j \geq \zeta_\upsilon$ for all $j > \upsilon$. That is, a regeneration epoch takes place when the walk reaches a certain level for the first time and never returns below it.

Lemma 25 below is proved in a more general context, namely, for the nearest-neighbor random walk on $\mathbb{Z}$ with site-dependent transition probabilities [12]. Lemma 26 is a classical result on transient random walks whose proof can be found in [18].

Lemma 25. If $p < \frac{1}{2}$ there exists a sequence $\{\upsilon_m\}_{m\geq 1}$ of regeneration epochs. The sequence $\{\upsilon_{n+1} - \upsilon_n\}_{n\geq 1}$ is i.i.d. as well as the sequence $\{\zeta_{\upsilon_{n+1}} - \zeta_{\upsilon_n}\}_{n\geq 1}$.

Lemma 26. Let R_n be the number of distinct sites visited by the walk $\{\zeta_k\}_{k\geq 0}$ after n steps. Then

$$\lim_{n\to\infty} \frac{\mathbb{E}^0[R_n]}{n} = \rho, \tag{5.1}$$

where ρ is the expected number of visits of the random walk to 0.

Lemma 27. $\mathbb{E}^0[\upsilon_1 - \upsilon_0] < \infty$

Proof. Fix any k and assume $\zeta_k = m$. By definition

$$\begin{aligned} &\{\zeta_k \text{ is a regeneration epoch}\} \\ &= \{\forall\ l < k,\ \zeta_l < m \text{ and } \zeta_k = m\} \cap \{\zeta_l > m,\ \forall\ l > k\} \\ &= \{\zeta_k \text{ is a ladder epoch}\} \cap \{\zeta_l > m\ \forall\ l > k\}. \end{aligned}$$

Then, by the Markov property, and noticing that $\gamma = \mathbb{P}^0[\{\zeta_l > m\ \forall\ l > k\}]$

$$\begin{aligned} &\mathbb{P}^0[\{\zeta_k \text{ is a regeneration epoch}\}] \\ &= \mathbb{P}^0[\{\zeta_k \text{ is a ladder epoch}\}]\mathbb{P}^0[\{\zeta_l > m,\ \forall\ l > k\}] \\ &= \gamma\mathbb{P}^0[\{\zeta_k \text{ is a ladder epoch}\}]. \end{aligned}$$

Now let $\hat{R}_n$ be the number of regeneration epochs on the interval $[1, n]$, i.e.,

$$\begin{aligned} \hat{R}_n &= \sum_{k=1}^{n} \mathbf{1}\{\zeta_k \text{ is a regeneration epoch}\} \Rightarrow \\ \mathbb{E}^0[\hat{R}_n] &= \gamma \sum_{k=1}^{n} \mathbb{P}^0[\{\zeta_k \text{ is a ladder epoch}\}] = \gamma\mathbb{E}^0[R_n]. \end{aligned} \tag{5.2}$$

.

By the Lemma 26 and (5.2),

$$\lim_{n\to\infty} \frac{\mathbb{E}^0[\hat{R}_n]}{n} = \frac{\gamma}{\rho}. \tag{5.3}$$

Next, notice that, by definition, $\mathbb{E}^0[R_n] \leq n$ and $\upsilon_n \geq n$, so, using (5.3),

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n}(\upsilon_i - \upsilon_{i-1})}{n} = \lim_{n\to\infty} \frac{\upsilon_n}{n} \leq \lim_{n\to\infty} \frac{n}{\mathbb{E}^0[\hat{R}_n]} \leq \frac{\rho}{\gamma}. \tag{5.4}$$

Now $\{\upsilon_n - \upsilon_{n-1}\}_{n>0}$ is an i.i.d. sequence. Hence, by the strong law of large numbers, we conclude from (5.4) that $\mathbb{E}^0[\upsilon_1 - \upsilon_0] < \infty$. ☐

Proof of Proposition 24. First, suppose D_0 has a $\mathbb{P}^0-$tail heavier than L_0. As the asymptotic efficiency when tasks may be repeated is less or equal to the asymptotic efficiency of when this is not the case, e_p, the former exists and it is zero.

Now suppose D_0 does not have a $\mathbb{P}^0-$tail heavier than L_0. Let

$$\mathbb{T}_n^R = \sum_{j=\upsilon_{n-1}+1}^{\upsilon_n} T_j^R,$$

for $n > 0$. Notice that $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n T_n^R = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \mathbb{T}_n^R$.

Now, by Lemma 27, since $\{\upsilon_n - \upsilon_{n-1}\}_{n\geq 0}$ is an i.i.d. sequence, Φ is a renewal process, the sequence $\{\mathbb{T}_n^R\}_{n\geq 1}$ is i.i.d..

The next step is to show that $\mathbb{E}^0[\mathbb{T}_0^R]$ is finite as long as D_0 does not have a $\mathbb{P}^0-$tail heavier than L_0, and can be bounded from above. First, notice that, since the discrete left-shift θ preserves $\mathbb{P}^0$ and the random walk is independent of Φ,

$$\mathbb{E}^0[\mathbb{T}_0^R] = \mathbb{E}^0[\mathbb{T}_0^R \circ \theta^{-\upsilon_0}] = \mathbb{E}^0\left[\sum_{i=1}^{\upsilon_1-\upsilon_0} T_i^R\right].$$

Now we use the following version of Wald's equality.

Lemma 28. If $\{Z_j\}_{j\in\mathbb{N}}$ is a sequence of positive random variables and η is a positive integer-valued random variables satisfying

1. $\mathbb{E}^0[Z] := \mathbb{E}^0[Z_j] < \infty$ for all $j \in \mathbb{N}$;
2. $\mathbb{E}^0[Z_j \mathbf{1}\{\eta \geq j\}] = \mathbb{E}^0[Z]\mathbb{P}^0[\eta \geq j]$ for all $j \in \mathbb{N}$,

then

$$\mathbb{E}^0\left[\sum_{i=1}^{\eta} Z_i\right] = \mathbb{E}^0[Z]\mathbb{E}^0[\eta].$$

Since Φ is a renewal process and the $\{L_i\}_{i\geq 1}$ is an i.i.d. sequence, $\mathbb{E}^0[T_j^R] = \mathbb{E}^0[T_i^R]$ for all $i, j \in \mathbb{N}$. Moreover, if $\mathbb{E}^0[T_0^R] < \infty$ Assumption 1 of Lemma 28 is satisfied. Assumption 2 of the same lemma is satisfied since the random walk $\{\zeta_k\}_{k\in\mathbb{N}}$ is independent of the point process and of the sequence $\{L_i\}_{i\geq 1}$.

As $\{\upsilon_n - \upsilon_{n-1}\}_{n>0}$ is an i.i.d. sequence (Lemma 25), and $\mathbb{E}^0[\upsilon_0 - \upsilon_1] < \infty$, by the strong law of large numbers,

$$\lim_{n\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \hat{\mathbb{T}}_n^R = \mathbb{E}^0 \left[\sum_{i=\upsilon_0+1}^{\upsilon_1} \hat{T}_i^R \right] = \mathbb{E}^0[T_0^R]\mathbb{E}[\upsilon_1 - \upsilon_0] < \infty, \quad \mathbb{P} - \text{a.s.}.$$

Therefore, we conclude that

$$e_p = \frac{\mathbb{E}^0[D_0]}{\mathbb{E}^0[T_0^R]\mathbb{E}^0[\upsilon_1 - \upsilon_0]}, \mathbb{P} - \text{a.s.}.$$

By (5.4), $\mathbb{E}^0[\upsilon_1 - \upsilon_0] \le \frac{\rho}{\gamma}$, completing the proof. □

A Proofs

Proof of Theorem 2. In this proof, we assume that D and L_0 admit densities. Let $T^R(z)$ be the *actual* `restart` time given $D = z$. Then

$$T^R(z) = z\mathbf{1}\{L_0 > z\} + (\hat{T}_R(z) + L_0)\mathbf{1}\{L_0 \le z\},$$

where $\hat{T}^R(z)$ is an independent copy of $T^R(z)$. Let $m_R(z)$ be the expectation of the actual `restart` time given $D = z$. Then

$$\begin{aligned} m_R(z) &= z\mathbb{P}[L_0 > z] + m_R(z)\mathbb{P}[L_0 \le z] + \mathbb{E}[L_0\mathbf{1}\{L_0 \le z\}] \Rightarrow \\ m_R(z) &= z + \frac{\mathbb{E}[L_0\mathbf{1}\{L_0 \le z\}]}{\mathbb{P}[L_0 > z]}. \end{aligned}$$

Therefore,

$$\mathbb{E}[T^R] = \mathbb{E}[m_R(z)] = \mathbb{E}[D] + \int_0^\infty \frac{\mathbb{E}[L_0\mathbf{1}\{L_0 \le z\}]f_D(dz)}{\mathbb{P}[L_0 > z]}.$$

In the same vein, let let $T^C(z)$ be the *actual* `checkpointing` time given $D = z$. Then $T^C(z) = L_0\mathbf{1}\{L_0 > z\} + (\hat{T}_R(z) + L_0)\mathbf{1}\{L_0 \le z\}$, where $\hat{T}_C(z)$ is an independent copy of $T^C(z)$. Let $m_C(z)$ be the expectation of the actual `checkpointing` time given $D = z$. Then

$$\begin{aligned} m_C(z) &= \mathbb{E}[U\mathbf{1}\{L_0 > z\} + m_C(z)\mathbb{P}[L_0 \le z] + \mathbb{E}[U\mathbf{1}\{L_0 \le z\}]) \\ &= \mathbb{E}[U] + m_C(z)\mathbb{P}[L_0 \le z] \Rightarrow \\ m_C(z) &= \frac{\mathbb{E}[L_0]}{\mathbb{P}[L_0 > z]}. \end{aligned}$$

Therefore,

$$\mathbb{E}[T^C] = \int_0^\infty \frac{\mathbb{E}[L_0] f_D(dz)}{\mathbb{P}[L_0 > z]}.$$

By construction, $T^C \geq T^R$, $\mathbb{P}-$a.s.. Suppose D has a $\mathbb{P}-$tail heavier than L_0. We show that $\mathbb{E}[T^R] = \infty$, which implies that $\mathbb{E}[T^C] = \infty$ as well. Suppose there exist $z_0 > 0$ so that $\mathbb{P}[L_0 > z] \leq \mathbb{P}[D > z]$ for all $z \geq z_0$. Let

$$Y(z_0) := \int_{z_0}^\infty \frac{f_D(dz)}{\mathbb{P}[L_0 > z]}.$$

Then, as D has a $\mathbb{P}-$tail heavier than L_0,

$$Y(z_0) \geq \int_{z_0}^\infty \frac{f_D(dz)}{\mathbb{P}[D > z]}.$$

Promoting the change of variable $w = \mathbb{P}[D \leq z]$ we have:

$$Y(z_0) = \int_{w_0}^1 \frac{1}{1-w} dw = \infty,$$

where $w_0 = \mathbb{P}[D \leq z_0] > 0$.

For the converse, it suffices to show that $\mathbb{E}[T^C] < \infty$. which is the case if D does not have a $\mathbb{P}-$tail heavier than L_0. Assume there is a sequence $\{z_n\}_{n\in\mathbb{N}}$ such that $z_n \to \infty$ and $\mathbb{P}[L_0 > z_n] > \mathbb{P}[D > z_n]$. Then, there is a sequence $\{\epsilon_n\}_{n\in\mathbb{N}}$ such that (i) $\epsilon_n > 0$ for all n, (ii) $\lim_{n\to\infty} \epsilon_n = 0$ and (iii) $\mathbb{P}[L_0 > z_n] = (\mathbb{P}[D > z_n])^{1+\epsilon_n}$. As before, using the change of variable $w = \mathbb{P}[D \leq z]$,

$$\begin{aligned}
\mathbb{E}[T^C] &\leq \mathbb{E}[U] \int_0^1 \frac{1}{(1-w)^{1+\epsilon_n}} dw \\
&= \mathbb{E}[L_0] \int_0^1 \frac{1}{(1-w)^{1+\epsilon_n}} dw \\
&\leq \mathbb{E}[L_0] \left(\int_0^1 \frac{1}{(1-w)^{2+2\epsilon_n}} dw \right)^{\frac{1}{2}}. \qquad \text{(A.1)}
\end{aligned}$$

The last inequality holding due to Holder's inequality. Letting $n \to \infty$ on the RHS of (A.1) and invoking Monotone Convergence, we conclude:

$$\mathbb{E}[T^C] \leq \mathbb{E}[L_0] \left(\int_0^1 \frac{1}{(1-w)^2} dw \right)^{\frac{1}{2}} < \infty.$$

□

Acknowledgments

This work was supported by a grant of the Simons Foundations (#197982 to the University of Texas at Austin). I would like to thank Abishek Sankararaman for his very valuable comments on this work.

References

[1] Søren Asmussen, *Applied probability and queues*, vol. 51, Springer, 2003.

[2] Søren Asmussen, Pierre Fiorini, Lester Lipsky, Tomasz Rolski, and Robert Sheahan, *Asymptotic behavior of total times for jobs that must start over if a failure occurs*, Mathematics of Operations Research **33** (2008), no. 4, 932–944.

[3] Søren Asmussen, Lester Lipsky, and Stephen Thompson, *Failure recovery in computing and data transmission: Restart and checkpointing.*

[4] Søren Asmussen, Anders Rønn-Nielsen, et al., *Failure recovery via restart: Wallclock models*, Thiele Research Reports, Department of Mathematical Sciences, University of Aarhus (2010).

[5] ———, *Failure recovery via restart: Wallclock models*, Thiele Research Reports, Department of Mathematical Sciences, University of Aarhus (2010).

[6] F. Baccelli and M.-O. Haji-Mirsadeghi, *Point-Map-Probabilities of a Point Process and Mecke's Invariant Measure Equation*, ArXiv e-prints (2013).

[7] Francois Baccelli and Pierre Brémaud, *Elements of queueing theory: Palm martingale calculus and stochastic recurrences*, vol. 26, Springer Science & Business Media, 2003.

[8] S Lirio Castellanos-Lopez, Felipe A Cruz-Pérez, and Genaro Hernandez-Valdez, *Channel reservation in cognitive radio networks with the restart retransmission strategy*, Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), 2012 7th International ICST Conference on, IEEE, 2012, pp. 65–70.

[9] Sung Nok Chiu, Dietrich Stoyan, Wilfrid S Kendall, and Joseph Mecke, *Stochastic geometry and its applications*, John Wiley & Sons, 2013.

[10] Daryl J Daley and David Vere-Jones, *An introduction to the theory of point processes*, vol. 1, Springer, 1988.

[11] ______, *An introduction to the theory of point processes*, vol. 2, Springer, 1988.

[12] Amir Dembo, Yuval Peres, and Ofer Zeitouni, *Tail estimates for one-dimensional random walk in random environment*, Communications in mathematical physics **181** (1996), no. 3, 667–683.

[13] Willliam Feller, *An introduction to probability theory and its applications*, vol. 2, John Wiley & Sons, 2008.

[14] Jacques Janssen and Raimondo Manca, *Applied semi-markov processes*, Springer Science & Business Media, 2006.

[15] Balachander Krishnamurthy, *Web protocols and practice: Http/1.1, networking protocols, caching, and traffic measurement*, Addison-Wesley Professional, 2001.

[16] Russell Lyons, Robin Pemantle, and Yuval Peres, *Biased random walks on galton–watson trees*, Probability theory and related fields **106** (1996), no. 2, 249–264.

[17] Jayakrishnan Nair, Martin Andreasson, Lachlan LH Andrew, Steven H Low, and John C Doyle, *File fragmentation over an unreliable channel*, INFOCOM, 2010 Proceedings IEEE, IEEE, 2010, pp. 1–9.

[18] Frank Spitzer, *Principles of random walk*, vol. 34, Springer Science & Business Media, 2013.

[19] H Thorisson, *Point-stationarity in d dimensions and palm theory. bernoulli 5 797–831*, Mathematical Reviews (MathSciNet) **10** (1999), 3318445.

[20] Hermann Thorisson, *Coupling, stationarity, and regeneration*, Springer New York, 2000.