



**HAL**  
open science

# Conditional failure occurrence rates for semi-Markov chains

Irene Votsi

► **To cite this version:**

Irene Votsi. Conditional failure occurrence rates for semi-Markov chains. *Journal of Applied Statistics*, 2019, 10.1080/02664763.2019.1610164 . hal-01761067v2

**HAL Id: hal-01761067**

**<https://hal.science/hal-01761067v2>**

Submitted on 23 Jan 2019

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

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

## Conditional failure occurrence rates for semi-Markov chains

Irene Votsi

Received: date / Accepted: date

**Abstract** In the present paper, we aim at providing plug-in-type empirical estimators that enable us to quantify the contribution of each operational or/and non-functioning state to the failures of a system described by a semi-Markov model. In the discrete time and finite state space semi-Markov framework, we study different conditional versions of an important reliability measure for random repairable systems, the failure occurrence rate based on counting processes. The aforementioned estimators are characterised by appealing asymptotic properties such as strong consistency and asymptotic normality. We further obtain detailed analytical expressions for the covariance matrices of the random vectors describing the conditional failure occurrence rates. As particular cases we present the failure occurrence rates for hidden (semi-) Markov models. We illustrate our results by means of an academic example based on simulated data. An application to real data is further presented that models sustainable vibration levels in a semi-Markov framework.

**Keywords** ROCOF · semi-Markov models · empirical estimators · asymptotic properties

**Mathematics Subject Classification (2000)** 60K15 · 60K20 · 62M05

### 1 Introduction

Semi-Markov models (SMMs) are state-of-the-art models that are widely used in many scientific fields such as reliability and DNA analysis ([3]), seismology ([22]) etc. One of the main distinguishing features of SMMs is that contrary to Markov models, they enable us to describe systems that evolve based not only on their last visited state (Markov property) but also on the time elapsed since this state. Due to this feature, popular “memory-full” distributions, such as the Weibull distribution, could be employed to describe sojourn (or interevent) times between successive events. We refer the interested reader to [8] or [12] for an introduction to homogeneous SMMs and to [18, 17] for non-homogeneous SMMs, respectively.

In the semi-Markov context, many reliability indicators have been introduced, including mean times to failure, hazard rates, availability functions etc. For recent advances in the topic concerning discrete time SMMs, see [1], [3], [5] and [6]. For continuous time SMMs, we address the interested reader to [10] and [11]. For advances in estimation methods of nonparametric semi-Markov models, see [15] and the references therein.

Here we focus on the rate of occurrence of failures (ROCOF), which is a fundamental reliability indicator for repairable, random systems subject to multiple failures. Yeh ([23]) was

---

I.Votsi  
Laboratoire Manceau de Mathématiques  
Le Mans Université  
72000 Le Mans, France  
Tel.: +33-0243833227  
E-mail: Eirini.Votsi@univ-lemans.fr

the first to investigate ROCOF for first-order Markov processes defined in a finite state space. Concerning semi-Markov processes in a finite or a general state space, ROCOF was studied by [13] and [9], respectively. The discrete time counterpart of ROCOF, was evaluated by [21] for SMMs and by [20] for hidden Markov renewal models.

ROCOF is rather a “global” reliability indicator in the sense that it does not distinguish neither on the starting functioning state nor on the ending inoperative state. It takes into account all the starting operational and all the ending defective states. Here we introduce the conditional counterparts of the failure occurrence rate, which allow quantifying the impact of any starting functioning state or/and any ending unworkable state. The study of these elementary rates could shed light on the functioning states that mostly influence the ROCOF and give insight into the dynamics of failures. Our objective is to provide indicators that enable us to identify opportunities for improving reliability performance, which is of special interest for real-life applications. The employment of reliability indicators with a more refined structure could guide industry action to improve operating practices and eventually provide decision support tools. Detailed analytical expressions for the covariance matrices of the random vectors describing the conditional failure occurrence rates are obtained. Consistent and asymptotically normal empirical estimators are presented and numerical examples based on real vibration data are provided.

The organization of the paper is as follows. In Section 2 the notation and preliminaries of SMMs are presented. Section 3 describes the definition, evaluation and statistical estimation of the conditional failure occurrence rates. Section 4 discusses numerical examples based on simulated and real data. Finally, in Section 5, we give some concluding remarks.

## 2 Notation and Preliminaries in Semi-Markov Models

We briefly recall the main definitions from the theory of discrete time SMMs (see, e.g., [3]). We consider a random system with finite state space  $E = \{1, 2, \dots, s\}$  described by an SMM. The stochastic evolution of the system is described by the following random sequences defined on a complete probability space (Fig.1):

1. The sequence  $\mathbf{J} = (J_n)_{n \in \mathbb{N}}$  with state space  $E$ , where  $J_n$  is the state visited by the system at the  $n$ -th jump time;
2. The  $\mathbb{N}$ -valued sequence  $\mathbf{S} = (S_n)_{n \in \mathbb{N}}$ , where  $S_n$  is the  $n$ -th jump time. We suppose that  $S_0 = 0$  and  $0 < S_1 < S_2 < \dots < S_n < S_{n+1} < \dots$  almost surely (a.s.);
3. The  $\mathbb{N}$ -valued sequence  $\mathbf{X} = (X_n)_{n \in \mathbb{N}}$  defined by  $X_0 = 0$  a.s. and  $X_n = S_n - S_{n-1}$  for all  $n \in \mathbb{N}$ . Thus for all  $n \in \mathbb{N}$ ,  $X_n$  is the sojourn time in state  $J_{n-1}$ , before the  $n$ -th jump.

We consider that the stochastic process  $(\mathbf{J}, \mathbf{S}) = (J_n, S_n)_{n \in \mathbb{N}}$  is a Markov renewal chain (MRC), that is for all  $n \in \mathbb{N}$ , for all  $i, j \in E$ , and for all  $k \in \mathbb{N}$  it satisfies a.s.

$$P(J_{n+1} = j, X_{n+1} = k | J_0, S_0, \dots, J_n, S_n) = P(J_{n+1} = j, X_{n+1} = k | J_n). \quad (1)$$

If equation (1) is independent of  $n$ , then  $(\mathbf{J}, \mathbf{S})$  is called homogeneous and its associated semi-Markov kernel (SMK)  $\mathbf{q} = (q_{ij}(k); i, j \in E, k \in \mathbb{N})$  is defined by

$$q_{ij}(k) = P(J_{n+1} = j, X_{n+1} = k | J_n = i),$$

where  $i, j \in E$  and  $k, n \in \mathbb{N}$ . It is worth mentioning that  $(\mathbf{J}, \mathbf{S})$  is also a Markov chain with state space  $E \times \mathbb{N}$  and therefore  $\mathbf{J}$  is a Markov chain with state space  $E$ . Thereafter, we refer to  $\mathbf{J}$  as the embedded Markov chain (EMC).

The semi-Markov chain (SMC)  $\mathbf{Z} = (Z_k)_{k \in \mathbb{N}}$  associated with the MRC  $(\mathbf{J}, \mathbf{S})$  describes the system's state at each time  $k \in \mathbb{N}$  and is defined by  $Z_k = J_{N(k)}$ , where  $N(k) = \max\{n \geq 0 : S_n \leq k\}$ . The transition kernel of the EMC is defined by  $\mathbf{P} = (p_{ij}; i, j \in E)$ , where  $p_{ij} = P(J_{n+1} = j | J_n = i)$ ,  $i, j \in E$ ,  $n \in \mathbb{N}$ , and its initial distribution by  $\boldsymbol{\alpha} = (\alpha(i); i \in E)$ , where  $\alpha(i) = P(J_0 = i)$ . At the initial time,  $t = 0$ , the state of the SMC coincides with the state of the EMC, i.e.  $Z_0 = J_0$ . We assume that there are neither instantaneous transitions ( $q_{ij}(0) = 0$ ,  $i, j \in E$ ) nor self-transitions for the EMC ( $p_{ii} = 0$ ,  $i \in E$ ).

The conditional distribution of the sojourn time in state  $i \in E$  given that the next visited state is  $j \in E$  is given by  $f_{ij}(k) = P(X_{n+1} = k | J_n = i, J_{n+1} = j)$ ,  $k \in \mathbb{N}$ . Moreover we define the survival function of the sojourn time distribution in state  $i \in E$  by  $\bar{H}_i(k) = P(X_{n+1} > k | J_n = i)$ ,  $k \in \mathbb{N}$ , and by  $m_i = \mathbb{E}(S_1 | Z_0 = i)$  the mean sojourn time of the SMC in state  $i \in E$ . The next assumptions have to be fulfilled in the following:

- A1** The SMC is irreducible and aperiodic [3];  
**A2** The mean sojourn times are finite, i.e.,  $m_i < \infty$ , for any  $i \in E$ .

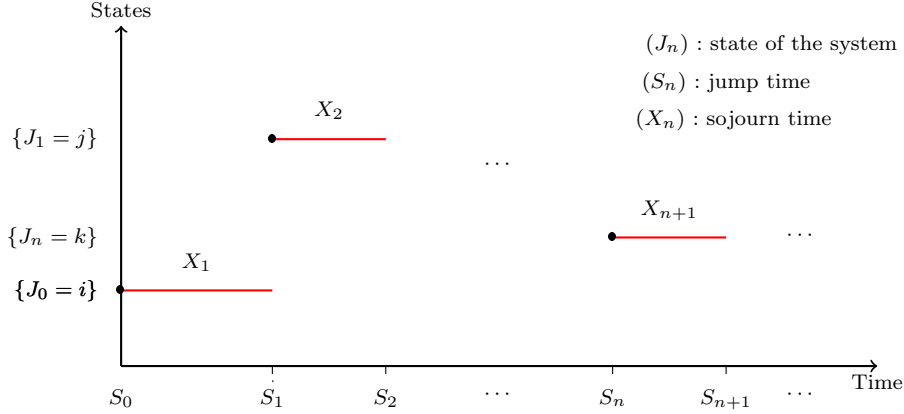


Fig. 1: A typical trajectory of a Markov renewal chain.

### 3 Conditional Failure Occurrence Rates

#### 3.1 General definitions

The failure occurrence rate is one of the most significant indicators in reliability theory, in particular in the study of repairable systems. In a semi-Markov framework, we first consider two non-empty, disjoint subsets of  $E$ ,  $U$  and  $D$  ( $U, D \subset E$ ;  $U, D \neq E$ ,  $U \cup D = E$ ), where  $U = \{1, \dots, r\}$  consists of the up (or functioning) states, and  $D = \{r + 1, \dots, s\}$  contains the down (or unworkable) states.

Second we denote by  $\mathbf{U} = (U_k)_{k \in \mathbb{N}}$  the sequence of the backward recurrence times of  $\mathbf{Z}$ , that is,  $U_k = k - S_{N(k)}$ . Following [4], the stochastic process  $(\mathbf{Z}, \mathbf{U}) = (Z_k, U_k)_{k \in \mathbb{N}}$  is a (time) homogeneous Markov chain governed by its initial distribution,  $\tilde{\boldsymbol{\alpha}} = (\tilde{\alpha}(i, 0); i \in E)$ , and its transition matrix  $\tilde{\mathbf{P}} = (\tilde{P}((i, t_1), (j, t_2))); (i, t_1), (j, t_2) \in E \times \mathbb{N}$ . We further denote  $T_M = \{0, 1, \dots, M\}$ , where  $M \in \mathbb{N}$ . In what follows, the matrix  $\Xi$ , of dimension  $(s(M + 1))^2 \times (s(M + 1))^2$ , is defined by

$$\Xi = \begin{pmatrix} \Xi(1, 0) & \mathbf{0} & \dots & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \Xi(i, t_1) & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \dots & \dots & \Xi(s, M) \end{pmatrix},$$

where the block  $\Xi(i, t_1)$  is the square matrix

$$\Xi(i, t_1) = \left( \delta_{(j, t_2)(r, t_3)} \tilde{P}((i, t_1), (j, t_2)) - \tilde{P}((i, t_1), (j, t_2)) \tilde{P}((i, t_1), (r, t_3)) \right),$$

and

$$\delta_{(j,t_2)(r,t_3)} = \begin{cases} 1, & \text{if } (j, t_2) = (r, t_3), \\ 0, & \text{elsewhere,} \end{cases}$$

for every  $(i, t_1), (j, t_2), (r, t_3) \in E \times T_M$ . Moreover, we define by  $\Delta$  the matrix:

$$\Delta = \begin{pmatrix} \frac{1}{\tilde{\pi}(1,0)}\Xi(1,0) & \underline{0} & \cdots & \cdots & \underline{0} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \frac{1}{\tilde{\pi}(i,t_1)}\Xi(i,t_1) & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \underline{0} & \cdots & \cdots & \cdots & \frac{1}{\tilde{\pi}(s,M)}\Xi(s,M) \end{pmatrix},$$

where  $\tilde{\pi} = (\tilde{\pi}(i, t_1); (i, t_1) \in E \times T_M)$ , is the stationary distribution of  $(\mathbf{Z}, \mathbf{U})$ . Consider now a trajectory  $\mathcal{H}(M)$  of an MRC with a fixed length  $M$ . For all  $i, j \in E$  and  $1 \leq k \leq M$ , we further define the counting processes:

1.  $N_i(M) = \sum_{n=1}^{N(M)-1} \mathbf{1}_{\{J_n=i\}}$  is the number of visits to state  $i$  of the EMC, before time  $M$ ;
2.  $N_{ij}(k, M) = \sum_{n=1}^{N(M)} \mathbf{1}_{\{J_{n-1}=i, J_n=j, X_n=k\}}$  is the number of transitions from  $i$  to  $j$ , up to time  $M$ , with sojourn time in state  $i$  equal to  $k$ .

Then, following [3], the empirical estimators of the semi-Markov kernel,  $q_{ij}(k)$ , and the survival function,  $H_i(k)$ , are defined by  $\hat{q}_{ij}(k, M) = \frac{N_{ij}(k, M)}{N_i(M)}$ , and  $\hat{H}_i(k, M) = 1 - \sum_{j \in E} \sum_{\ell=0}^k \hat{q}_{ij}(\ell, M)$ , respectively. Additionally, following [4], we obtain directly the plug-in type empirical estimators of the transition probabilities of the Markov chain  $(\mathbf{Z}, \mathbf{U})$ :

$$\hat{P}_M((i, t_1), (j, t_2)) = \begin{cases} \hat{q}_{ij}(t_1 + 1, M) / \hat{H}_i(t_1, M), & \text{if } t_2 = 0, \\ \hat{H}_i(t_1 + 1, M) / \hat{H}_i(t_1, M), & \text{if } i = j, t_2 - t_1 = 1, \\ 0, & \text{elsewhere.} \end{cases}$$

### 3.2 Full conditional ROCOF

In the sequel we are interested in studying the impact of both the starting up and the ending down states on the ROCOF, simultaneously. In order to do that, we define the full conditional ROCOF as the expected number of transitions of the SMC to the set  $D$  at time  $k$ , given that it starts in state  $i \in U$  and ends in state  $j \in D$ , i.e.,

$$r_{ij}^\#(k) = \mathbb{E}[N_{ij}^\#(k) - N_{ij}^\#(k-1)],$$

where

$$N_{ij}^\#(k) = \sum_{\ell=1}^k \mathbf{1}_{\{Z_{\ell-1}=i, Z_\ell=j\}}.$$

**Theorem 1** *The full conditional ROCOF at time  $k \in \mathbb{N}^*$  is given by*

$$r_{ij}^\#(k) = \sum_{m=0}^{k-1} [(\tilde{\alpha}\tilde{P}^{k-1})(i, m)] \tilde{P}((i, m), (j, 0)),$$

for any fixed arbitrary states  $i \in U, j \in D$ .

*Proof* For fixed  $i \in U, j \in D$  and  $k \in \mathbb{N}^*$ ,

$$\begin{aligned} r_{ij}^\#(k) &= \mathbb{E}[N_{ij}^\#(k) - N_{ij}^\#(k-1)] \\ &= \mathbb{E}\mathbf{1}_{\{Z_{k-1}=i, Z_k=j\}} \\ &= P(Z_{k-1} = i, Z_k = j). \end{aligned} \tag{2}$$

Moreover, following [21], we have that

$$P(Z_{k-1} = i, Z_k = j) = \sum_{m=0}^{k-1} [(\tilde{\alpha}\tilde{P}^{k-1})(i, m)] \tilde{P}((i, m), (j, 0)), \tag{3}$$

for  $i \neq j$ . Therefore, from equations (2) and (3), we get the desired result.

For given  $i \in U$ ,  $j \in D$ , we estimate the full conditional ROCOF by means of the plug-in type estimator

$$\widehat{r}_{ij}^{\#}(k, M) = \sum_{m=0}^{k-1} [(\widehat{\alpha} \widehat{P}_M^{k-1})(i, m)] \widehat{P}_M((i, m), (j, 0)),$$

where  $(\widehat{\alpha} \widehat{P}_M^{k-1})(i, m)$  is the  $(i, m)$ -th element of the vector  $\widehat{\alpha} \widehat{P}_M^{k-1}$ .

**Remark 1** The symbol  $\widehat{\alpha}(i, m)$  represents an estimator of the initial law of the stochastic process  $(Z_k, U_k)_{k \in \mathbb{N}}$ , that is an estimator of  $\tilde{\alpha}(i, m) = P(Z_0 = i, U_0 = m)$ . If we consider that  $(Z_k, U_k)_{k \in \mathbb{N}}$  is irreducible, then its initial law could be replaced by its stationary law that is given by the formula [4]  $\tilde{\pi}(i, m) = \frac{1 - \bar{H}_i(m)}{\mu_{ii}}$ , where  $\mu_{ii}$  is the mean recurrence time of state  $i$  of the SMC.

**Proposition 1** The empirical estimator of the stationary distribution of the process  $(Z_k, U_k)_{k \in \mathbb{N}}$  is strongly consistent in the sense that

$$\widehat{\pi}_M(i, m) \xrightarrow[M \rightarrow \infty]{a.s.} \pi(i, m).$$

*Proof* From the definition of  $\widehat{\pi}_M(i, m)$  we have that

$$\widehat{\pi}_M(i, m) = \frac{1 - \sum_{\ell=1}^m \sum_{j \in E} \widehat{q}_{ij}(\ell, M)}{\widehat{\mu}_{ii}(M)}.$$

First we notice that [3]

$$\widehat{q}_{ij}(\ell, M) \xrightarrow[M \rightarrow \infty]{a.s.} q_{ij}(\ell).$$

Second by the definition of the estimator of the mean recurrence time of the state  $i$  of the SMC we can obtain the empirical estimator

$$\widehat{\mu}_{ii}(M) = \frac{\sum_{j \in E} \widehat{\nu}_j(M) \widehat{m}_j(M)}{\widehat{\nu}_i(M)},$$

where

$$\widehat{\nu}_j(M) = \frac{N_j(M)}{N(M)}$$

and

$$\widehat{m}_j(M) = \sum_{k \geq 0} \widehat{H}_j(k, M).$$

Then since we deal with finite sums and [2]

$$\widehat{\nu}_j(M) \xrightarrow[M \rightarrow \infty]{a.s.} \nu_j \quad \text{and} \quad \widehat{m}_j(M) \xrightarrow[M \rightarrow \infty]{a.s.} m_j,$$

we obtain the strong consistency of the estimator of the mean recurrence time. The previous results lead directly to the strong consistency of the estimator of the stationary distribution of the process  $(Z_k, U_k)_{k \in \mathbb{N}}$ .

**Proposition 2** For any states  $i \in U$ ,  $j \in D$  and any fixed arbitrary positive integer  $k \in \mathbb{N}$ , the estimator of the full conditional ROCOF at time  $k$  is strongly consistent in the sense that

$$\widehat{r}_{ij}^{\#}(k, M) \xrightarrow[M \rightarrow \infty]{a.s.} r_{ij}^{\#}(k).$$

*Proof* Following [14] for the Markov chain  $(\mathbf{Z}, \mathbf{U})$  we have  $\widehat{P}_M \xrightarrow[M \rightarrow \infty]{a.s.} \widetilde{P}$  and since we deal with finite sums, we can write directly:

$$\begin{aligned} \widehat{r}_{ij}^\#(k, M) &= \sum_{m=0}^{k-1} [(\widehat{\alpha\widetilde{P}}_M^{k-1})(i, m)] \widehat{P}_M((i, m), (j, 0)) \\ &\xrightarrow[M \rightarrow \infty]{a.s.} \sum_{m=0}^{k-1} [(\widetilde{\alpha\widetilde{P}}^{k-1})(i, m)] \widetilde{P}((i, m), (j, 0)) \\ &= r_{ij}^\#(k). \end{aligned}$$

**Proposition 3** For any fixed  $i \in U$ ,  $j \in D$ ,  $k \geq 1$ , the random variable  $\sqrt{M}(\widehat{r}_{ij}^\#(k, M) - r_{ij}^\#(k))$  converges in distribution, as  $M$  tends to infinity, to a zero mean normal random variable with variance  $\Phi'_{ij} \Delta \Phi_{ij}^\top$ . Moreover  $\Phi_{ij} : [0, 1]^d \rightarrow \mathbb{R}^+$  ( $d = (s(M+1))^2$ ) is the function

$$\begin{aligned} &\Phi_{ij}(\widetilde{P}((i', m'), (j', t'))); (i', m'), (j', t') \in E \times T_M \\ &= \sum_{m=0}^{k-1} \left( (\widetilde{\alpha\widetilde{P}}^{k-1})(i, m) \right) \widetilde{P}((i, m), (j, 0)) \\ &= \sum_{m=0}^{k-1} \left( \sum_{s \in E} \widetilde{\alpha}(s, 0) \widetilde{P}^{k-1}((s, 0), (i, m)) \right) \widetilde{P}((i, m), (j, 0)) \end{aligned}$$

and

$$\Phi'_{ij} = \left( \frac{\partial \Phi_{ij}}{\partial \widetilde{P}((i', m'), (j', t'))}; (i', m'), (j', t') \in E \times T_M \right),$$

is the  $d$ -dimensional row vector of first derivatives of  $r_{ij}^\#(k)$  with respect to  $\widetilde{P}((i', m'), (j', t'))$ .

*Proof* Following the definition of the full conditional ROCOF, we have

$$\begin{aligned} &\sqrt{M}(\widehat{r}_{ij}^\#(k, M) - r_{ij}^\#(k)) \\ &= \sqrt{M} \sum_{m=0}^{k-1} \left( [(\widehat{\alpha\widetilde{P}}_M^{k-1})(i, m)] \widehat{P}_M((i, m), (j, 0)) \right. \\ &\quad \left. - [(\widetilde{\alpha\widetilde{P}}^{k-1})(i, m)] \widetilde{P}((i, m), (j, 0)) \right) \\ &= \sqrt{M} \left( \Phi_{ij}(\widehat{P}_M((i', m'), (j', t'))) - \Phi_{ij}(\widetilde{P}((i', m'), (j', t'))) \right), \end{aligned}$$

where  $\Phi_{ij}$  is a polynomial function and consequently differentiable. The vector

$$\sqrt{M} \left( \widehat{P}_M((i', m'), (j', t')) - \widetilde{P}((i', m'), (j', t')) \right)$$

converges, as  $M$  tends to infinity, to the normal distribution centered at the origin with covariance matrix  $\Delta$ . Then by means of the delta method (see, e.g., [16]), we obtain that

$$\sqrt{M} \left( \Phi_{ij}(\widehat{P}_M((i', m'), (j', t'))) - \Phi_{ij}(\widetilde{P}((i', m'), (j', t'))) \right) \xrightarrow[M \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \Phi'_{ij} \Delta \Phi_{ij}^\top).$$

At a next step we prove the asymptotic normality of the random vector  $\mathcal{R}(k) = (\mathcal{R}_{ij}(k))_{i,j \in E}$  where  $\mathcal{R}_{ij}(k) = \sqrt{M}(\widehat{r}_{ij}^\#(k, M) - r_{ij}^\#(k))$ , for any fixed  $k \geq 3$ , and obtain an explicit formula for the corresponding covariance matrix.

**Lemma 1** For  $n \geq 2$ , the random vector  $F^n = (f_{(i',m'),(j',t')}^n)_{(i',m'),(j',t') \in E \times T_M}$  where

$$f_{(i',m'),(j',t')}^n = \sqrt{M} (\widehat{P}^n((i',m'),(j',t')) - \widetilde{P}^n((i',m'),(j',t')))$$

has the same limit in distribution as the random vector

$$G^n = (g_{(i',m'),(j',t')}^n)_{(i',m'),(j',t') \in E \times T_M},$$

where

$$\begin{aligned} g_{(i',m'),(j',t')}^n &= \sum_{(j,t) \in E \times T_M} \left( \widetilde{P}^{n-1}((i',m'),(j,t)) f_{(j,t),(j',t')} + \widetilde{P}^{n-1}((j,t),(j',t')) f_{(i',m'),(j,t)} \right) \\ &+ \mathbf{1}_{\{n \geq 3\}} \sum_{k=2}^{n-1} \sum_{(i_1,t_1) \in E \times T_M} \sum_{(i_2,t_2) \in E \times T_M} \widetilde{P}^{n-k}((i',m'),(i_1,t_1)) \widetilde{P}^{k-1}((i_2,t_2),(j',t')) f_{(i_1,t_1),(i_2,t_2)}. \end{aligned}$$

Moreover,  $F^n$  converges, as  $M$  tends to infinity, to a centered normal random vector with covariance matrix  $\Sigma = \Sigma_f \Delta \Sigma_f^\top$ . The matrix

$$\Sigma_f = \Sigma_f((i',m'),(j',t'),(u',v'),(s',w'))_{(i',m'),(j',t'),(u',v'),(s',w') \in E \times T_M}$$

of dimension  $d \times d$  ( $d = s^2(M+1)^2$ ) is given by

$$\begin{aligned} \Sigma_f((i',m'),(j',t'),(u',v'),(s',w')) &= \delta_{(j',t'),(s',w')} \widetilde{P}^{n-1}((i',m'),(u',v')) \\ &+ \delta_{(i',m'),(u',v')} \widetilde{P}^{n-1}((s',w'),(j',t')) \\ &+ \sum_{k=2}^{n-1} \widetilde{P}^{n-k}((i',m'),(u',v')) \widetilde{P}^{k-1}((s',w'),(j',t')). \end{aligned}$$

*Proof* Since  $(\mathbf{Z}, \mathbf{U})$  is a Markov chain, we follow [14] and obtain the desired result.

**Theorem 2** For any fixed  $k \geq 3$ , the random vector  $\mathcal{R}(k) = (\mathcal{R}_{ij}(k))_{i,j \in E}$  where  $\mathcal{R}_{ij}(k) = \sqrt{M} (\widehat{r}_{ij}^\#(k, M) - r_{ij}^\#(k))$  converges in distribution, as  $M$  tends to infinity, to a centered normal random vector with covariance matrix  $\Sigma \Delta \Sigma^\top$ .

*Proof* First, we notice that

$$\begin{aligned} \mathcal{R}_{ij}(k) &= \sqrt{M} \sum_{m=0}^{k-1} \sum_{s \in E} \widetilde{\alpha}(s, 0) \left( \widehat{P}^{k-1}((s, 0), (i, m)) \widehat{P}((i, m), (j, 0)) \right. \\ &\quad \left. - \widetilde{P}^{k-1}((s, 0), (i, m)) \widetilde{P}((i, m), (j, 0)) \right) \\ &= \sum_{m=0}^{k-1} \sum_{s \in E} \widetilde{\alpha}(s, 0) \left( f_{(s,0),(i,m)}^{k-1} (\widehat{P}((i, m), (j, 0)) - \widetilde{P}((i, m), (j, 0))) \right. \\ &\quad \left. + \widetilde{P}^{k-1}((s, 0), (i, m)) f_{(i,m),(j,0)} + f_{(s,0),(i,m)}^{k-1} \widetilde{P}((i, m), (j, 0)) \right). \end{aligned}$$

Second, using Lemma 1 and following [14], we obtain that

$$f_{(s,0),(i,m)}^{k-1} \xrightarrow[M \rightarrow \infty]{\mathcal{L}} Z,$$

where  $Z$  is a centered normal random variable, and

$$\widehat{P}((i, m), (j, 0)) - \widetilde{P}((i, m), (j, 0)) \xrightarrow[M \rightarrow \infty]{\mathcal{P}} 0.$$



Then from Slutsky's theorem, we deduce that  $\mathcal{R}_{ij}(k)$  has the same limit in distribution as

$$\sum_{m=0}^{k-1} \sum_{s \in E} \tilde{\alpha}(s, 0) \left( \tilde{P}^{k-1}((s, 0), (i, m)) f_{(i, m), (j, 0)} + f_{(s, 0), (i, m)}^{k-1} \tilde{P}((i, m), (j, 0)) \right),$$

which in turn has the same limit in distribution as

$$\sum_{m=0}^{k-1} \sum_{s \in E} \tilde{\alpha}(s, 0) \left( \tilde{P}^{k-1}((s, 0), (i, m)) f_{(i, m), (j, 0)} + g_{(s, 0), (i, m)}^{k-1} \tilde{P}((i, m), (j, 0)) \right).$$

Since the last function is continuous and linear w.r.t. the vector  $f_{(i', m'), (j', t')}$ , we use the continuous mapping theorem and deduce that  $\mathcal{R}_{ij}$  converges in distribution to a centered normal random vector with covariance matrix  $\Sigma \Delta \Sigma^\top$ . The rectangular matrix

$$\Sigma = \Sigma((i, j), (u', v'), (s', w'))_{i, j \in E, (u', v'), (s', w') \in E \times T_M}$$

of dimension  $d$  ( $d = s^2 \times (s(M+1))^2$ ) is given by

$$\begin{aligned} \Sigma((i, j), (u', v'), (s', w')) &= \delta_{(u', i)} \mathbf{1}_{(v' \leq k-1)} \delta_{(s', w'), (j, 0)} \sum_{s \in E} \tilde{\alpha}(s, 0) \tilde{P}^{k-1}((s, 0), (u', v')) \\ &\quad + \delta_{(s', i)} \mathbf{1}_{(w' \leq k-1)} \tilde{P}((i, w'), (j, 0)) \sum_{s \in E} \tilde{\alpha}(s, 0) \tilde{P}^{k-2}((s, 0), (u', v')) \\ &\quad + \delta_{(u', v'), (s, 0)} \tilde{\alpha}(u', 0) \sum_{m=0}^{k-1} \tilde{P}((i, m), (j, 0)) \tilde{P}^{k-2}((s', w'), (i, m)) \\ &\quad + \sum_{m=0}^{k-1} \sum_{s \in E} \sum_{r=2}^{k-2} \tilde{\alpha}(s, 0) \tilde{P}((i, m), (j, 0)) \tilde{P}^{k-1-r}((s, 0), (u', v')) \tilde{P}^{r-1}((s', w'), (i, m)). \end{aligned}$$

### 3.3 Left conditional ROCOF

We first define the left conditional ROCOF as the expected number of transitions of the SMC to the set  $D$  at time  $k$ , given that it starts from the fixed state  $i \in U$ , i.e.,

$$\tilde{r}_i(k) = \mathbb{E}[\tilde{N}_i(k) - \tilde{N}_i(k-1)],$$

where

$$\tilde{N}_i(k) = \sum_{l=1}^k \mathbf{1}_{\{Z_{l-1}=i, Z_l \in D\}}.$$

**Theorem 3** *The left conditional ROCOF at time  $k \in \mathbb{N}^*$  is given by*

$$\tilde{r}_i(k) = \sum_{j \in D} \sum_{m=0}^{k-1} [(\tilde{\alpha} \tilde{P}^{k-1})(i, m)] \tilde{P}((i, m), (j, 0)),$$

for any fixed, arbitrary state  $i \in U$ .

*Proof* For fixed  $i \in U$  and  $k \in \mathbb{N}^*$ ,

$$\begin{aligned} \tilde{r}_i(k) &= \mathbb{E}[\tilde{N}_i(k) - \tilde{N}_i(k-1)] \\ &= \mathbb{E} \mathbf{1}_{\{Z_{k-1}=i, Z_k \in D\}} = P(Z_{k-1} = i, Z_k \in D) \\ &= \sum_{j \in D} P(Z_{k-1} = i, Z_k = j). \end{aligned} \tag{4}$$

Therefore, from equations (3) and (4), we get the desired result.

We further estimate the left conditional ROCOF by means of the plug-in type estimator:

$$\widehat{r}_i(k, M) = \sum_{j \in D} \sum_{m=0}^{k-1} [(\widehat{\alpha} \widehat{P}_M^{k-1})(i, m)] \widehat{P}_M((i, m), (j, 0)).$$

**Proposition 4** For any fixed, arbitrary state  $i \in U$  and any fixed, arbitrary positive integer  $k \in \mathbb{N}$ , the estimator of the left conditional ROCOF at time  $k$  is strongly consistent in the sense that

$$\widehat{r}_i(k, M) \xrightarrow[M \rightarrow \infty]{a.s.} \widetilde{r}_i(k).$$

*Proof* To prove the consistency, we work as in Proposition 2.

**Proposition 5** For any fixed  $i \in U$ ,  $k \geq 1$ ,  $\sqrt{M}(\widehat{r}_i(k, M) - \widetilde{r}_i(k))$  converges in distribution, as  $M$  tends to infinity, to a zero mean normal random variable with variance  $\Phi'_i \Delta \Phi_i'^\top$ . Moreover  $\Phi_i : [0, 1]^d \rightarrow \mathbb{R}^+$  ( $d = (s(M+1))^2$ ) is the function

$$\begin{aligned} & \Phi_i \left( \widetilde{P}((i', m'), (j', t')); (i', m'), (j', t') \in E \times T_M \right) \\ &= \sum_{j \in D} \sum_{m=0}^{k-1} [(\widetilde{\alpha} \widetilde{P}^{k-1})(i, m)] \widetilde{P}((i, m), (j, 0)) \\ &= \sum_{j \in D} \sum_{m=0}^{k-1} \left( \sum_{s \in E} \widetilde{\alpha}(s, 0) \widetilde{P}^{k-1}((s, 0), (i, m)) \right) \widetilde{P}((i, m), (j, 0)) \end{aligned}$$

and

$$\Phi'_i = \left( \frac{\partial \Phi_i}{\partial \widetilde{P}((i', m'), (j', t'))}; (i', m'), (j', t') \in E \times T_M \right),$$

is the  $d$ -dimensional row vector of first derivatives of  $\widetilde{r}_i(k)$  with respect to  $\widetilde{P}((i', m'), (j', t'))$ .

*Proof* First, we have that

$$\sqrt{M}(\widehat{r}_i(k, M) - \widetilde{r}_i(k)) = \sqrt{M} \left( \Phi_i(\widehat{P}_M((i', m'), (j', t'))) - \Phi_i(\widetilde{P}((i', m'), (j', t'))) \right).$$

The function  $\Phi_i$  is a polynomial function, therefore it is differentiable. Moreover, following [14] for the Markov chain  $(\mathbf{Z}, \mathbf{U})$ , the vector

$$\sqrt{M} \left( \widehat{P}_M((i', m'), (j', t')) - \widetilde{P}((i', m'), (j', t')) \right)$$

converges, as  $M$  tends to infinity, to the normal distribution centered at the origin with covariance matrix  $\Delta$ . Then by means of the delta method, we obtain the desired result.

**Theorem 4** For any fixed  $k \geq 3$ , the random vector  $\sqrt{M}(\widehat{r}_i(k, M) - r_i(k))_{i \in U}$  converges in distribution, as  $M$  tends to infinity, to a centered normal random vector with covariance matrix  $\Sigma_U \Delta \Sigma_U^\top$ , where the  $d$ -dimensional ( $d = r \times (s(M+1))^2$ ) matrix  $\Sigma_U = \Sigma_U(i, (u', v'), (s', w'))_{i \in U, (u', v'), (s', w') \in E \times T_M}$  is given by

$$\begin{aligned} \Sigma_U(i, (u', v'), (s', w')) &= \delta_{(u', i)} \mathbf{1}_{(v' \leq k-1)} \mathbf{1}_{(s' \in D)} \delta_{(w', 0)} \sum_{s \in E} \widetilde{\alpha}(s, 0) \widetilde{P}^{k-1}((s, 0), (u', v')) \\ &+ \delta_{(s', i)} \mathbf{1}_{(w' \leq k-1)} \sum_{j \in D} \widetilde{P}((i, w'), (j, 0)) \sum_{s \in E} \widetilde{\alpha}(s, 0) \widetilde{P}^{k-2}((s, 0), (u', v')) \\ &+ \delta_{(u', v'), (s, 0)} \widetilde{\alpha}(u', 0) \sum_{j \in D} \sum_{m=0}^{k-1} \widetilde{P}((i, m), (j, 0)) \widetilde{P}^{k-2}((s', w'), (i, m)) \\ &+ \sum_{j \in D} \sum_{m=0}^{k-1} \sum_{s \in E} \sum_{r=2}^{k-2} \widetilde{\alpha}(s, 0) \widetilde{P}((i, m), (j, 0)) \widetilde{P}^{k-1-r}((s, 0), (u', v')) \widetilde{P}^{r-1}((s', w'), (i, m)). \end{aligned}$$

*Proof* Working as in Theorem 2, we get the desired result.

### 3.4 Right conditional ROCOF

We go one step further and aim at studying the impact of the ending down state on the ROCOF. In order to do that, we define the right conditional ROCOF as the expected number of transitions of the SMC to the set  $D$  at time  $k$ , given that it ends in state  $j \in D$ , i.e.,

$$\bar{r}_j(k) = \mathbb{E}[\bar{N}_j(k) - \bar{N}_j(k-1)],$$

where

$$\bar{N}_j(k) = \sum_{l=1}^k \mathbf{1}_{\{Z_{l-1} \in U, Z_l = j\}}.$$

**Theorem 5** *The right conditional ROCOF at time  $k \in \mathbb{N}^*$  is given by*

$$\bar{r}_j(k) = \sum_{i \in U} \sum_{m=0}^{k-1} [(\tilde{\alpha} \tilde{P}^{k-1})(i, m)] \tilde{P}((i, m), (j, 0)),$$

for any fixed, arbitrary state  $j \in D$ .

*Proof* For fixed  $j \in D$  and  $k \in \mathbb{N}^*$ ,

$$\begin{aligned} \bar{r}_j(k) &= \mathbb{E}[\bar{N}_j(k) - \bar{N}_j(k-1)] \\ &= \mathbb{E}[\mathbf{1}_{\{Z_{k-1} \in U, Z_k = j\}}] = P(Z_{k-1} \in U, Z_k = j) \\ &= \sum_{i \in U} P(Z_{k-1} = i, Z_k = j). \end{aligned} \tag{5}$$

Equations (3) and (5) lead directly to the result.

For a fixed state  $j \in D$ , we estimate the right conditional ROCOF by means of the plug-in type estimator :

$$\hat{\bar{r}}_j(k, M) = \sum_{i \in U} \sum_{m=0}^{k-1} [(\hat{\alpha} \hat{P}_M^{k-1})(i, m)] \hat{P}_M((i, m), (j, 0)).$$

**Proposition 6** *For any fixed arbitrary state  $j \in D$  and any fixed arbitrary positive integer  $k \in \mathbb{N}$ , the estimator of the conditional ROCOF at time  $k$  is strongly consistent in the sense that*

$$\hat{\bar{r}}_j(k, M) \xrightarrow[M \rightarrow \infty]{a.s.} \bar{r}_j(k).$$

*Proof* Working as in Proposition 2, we obtain the desired result.

**Proposition 7** *For any fixed  $j \in D$ ,  $k \geq 1$ , the random variable  $\sqrt{M}(\hat{\bar{r}}_j(k, M) - \bar{r}_j(k))$  converges in distribution, as  $M$  tends to infinity, to a zero mean normal random variable with variance  $\Phi_j' \Delta \Phi_j^{\top} \Phi_j : [0, 1]^d \rightarrow \mathbb{R}^+$  ( $d = (s(M+1))^2$ ) is the function*

$$\begin{aligned} &\Phi_j \left( \tilde{P}((i', m'), (j', t')); (i', m'), (j', t') \in E \times T_M \right) \\ &= \sum_{i \in U} \sum_{m=0}^{k-1} [(\tilde{\alpha} \tilde{P}^{k-1})(i, m)] \tilde{P}((i, m), (j, 0)) \\ &= \sum_{i \in U} \sum_{m=0}^{k-1} \left( \sum_{s \in E} \tilde{\alpha}(s, 0) \tilde{P}^{k-1}((s, 0), (i, m)) \right) \tilde{P}((i, m), (j, 0)) \end{aligned}$$

and

$$\Phi_j' = \left( \frac{\partial \Phi_j}{\partial \tilde{P}((i', m'), (j', t'))}; (i', m'), (j', t') \in E \times T_M \right),$$

is the  $d$ -dimensional row vector of first derivatives of  $\bar{r}_j(k)$  with respect to  $\tilde{P}((i', m'), (j', t'))$ .

*Proof* The proof derives in the same way as in Proposition 3.

**Theorem 6** For any fixed  $k \geq 3$ , the random vector  $\sqrt{M}(\tilde{r}_j(k, M) - r_j(k))_{j \in D}$  converges in distribution, as  $M$  tends to infinity, to a centered normal random vector with covariance matrix  $\Sigma_D \Delta \Sigma_D^\top$ , where

$$\Sigma_D = \Sigma_D(j, (u', v'), (s', w'))_{j \in D, (u', v'), (s', w') \in E \times T_M}$$

is the  $d$ -dimensional matrix ( $d = (s - r) \times (s(M + 1))^2$ ) given by

$$\begin{aligned} \Sigma_D(j, (u', v'), (s', w')) &= \delta_{((s', w'), (j, 0))} \mathbf{1}_{(u' \in U)} \mathbf{1}_{(v' \leq k-1)} \sum_{s \in E} \tilde{\alpha}(s, 0) \tilde{P}^{k-1}((s, 0), (u', v')) \\ &\quad + \mathbf{1}_{(s' \in U)} \mathbf{1}_{(w' \leq k-1)} \tilde{P}((i, w'), (j, 0)) \sum_{s \in E} \tilde{\alpha}(s, 0) \tilde{P}^{k-2}((s, 0), (u', v')) \\ &\quad + \delta_{(u', v'), (s, 0)} \tilde{\alpha}(u', 0) \sum_{i \in U} \sum_{m=0}^{k-1} \tilde{P}((i, m), (j, 0)) \tilde{P}^{k-2}((s', w'), (i, m)) \\ &\quad + \sum_{i \in U} \sum_{m=0}^{k-1} \sum_{s \in E} \sum_{r=2}^{k-2} \tilde{\alpha}(s, 0) \tilde{P}((i, m), (j, 0)) \tilde{P}^{k-1-r}((s, 0), (u', v')) \tilde{P}^{r-1}((s', w'), (i, m)). \end{aligned}$$

*Proof* To prove this theorem one can follow the same lines as in Theorem 2.

### 3.5 Particular cases

In the sequel we present two particular cases of the full conditional ROCOF in the context of hidden semi-Markov models and hidden Markov models. The evaluation and estimation of the ROCOF have been studied for hidden Markov renewal models (HMRMs) in [19] and [20]. However HMRMs present a specific limitation: they consider that observations occur only at the jump times  $n \in \mathbb{N}$  and not at the calendar times  $k \in \mathbb{N}$ , which is the case for the “classical” hidden semi-Markov models. Here this assumption is relaxed by allowing observations to be recorded at the calendar times and an explicit formula is obtained for hidden semi-Markov models and hidden Markov models.

In a hidden (semi-) Markov context, we consider that the states visited by the (semi-) Markov chain are “hidden”, i.e. inaccessible to direct observation. On the other side, what is available to observation is another stochastic process,  $\mathbf{Y} = (Y_k)_{k \in \mathbb{N}}$  which is defined in a finite state space  $A$ . We assume that the observations are conditionally independent given the underlying states of the (semi-) Markov chain, that is

$$P(Y_k = a | Y_0, \dots, Y_{k-1}, Z_0, \dots, Z_k = i) = P(Y_k = a | Z_k = i),$$

for any  $a \in A$ ,  $i \in E$ ,  $k \in \mathbb{N}$ . We further define by  $\mathbf{R} = (R_{i;a}; i \in E, a \in A)$  the emission kernel, where  $R_{i;a} = P(Y_k = a | Z_k = i)$ .

Then the ROCOF is defined by

$$r(k) = \mathbb{E}[N(k) - N(k-1)],$$

where

$$N(k) = \sum_{l=1}^k \mathbf{1}_{\{Y_{l-1} \in U, Y_l \in D\}}.$$

#### 3.5.1 Hidden semi-Markov models

**Theorem 7** The ROCOF at time  $k \in \mathbb{N}^*$  is given by

$$r(k) = \sum_{a_1 \in U} \sum_{a_2 \in D} \sum_{i \in E} \sum_{j \in E} \sum_{u_1 \in T_{k-1}} \sum_{u_2 \in T_k} R_{j;a_2} P((i, u_1), (j, u_2)) R_{i;a_1} (aP)^{k-1}(i, u_1).$$

*Proof* For fixed  $k \in \mathbb{N}^*$

$$\begin{aligned}
r(k) &= \mathbb{E}[N(k) - N(k-1)] \\
&= P(Y_{k-1} \in U, Y_k \in D) \\
&= \sum_{a_1 \in U} \sum_{a_2 \in D} P(Y_{k-1} = a_1, Y_k = a_2) \\
&= \sum_{a_1 \in U} \sum_{a_2 \in D} \sum_{i \in E} \sum_{j \in E} P(Y_{k-1} = a_1, Y_k = a_2, Z_{k-1} = i, Z_k = j). \tag{6}
\end{aligned}$$

On the other hand

$$\begin{aligned}
&P(Y_{k-1} = a_1, Y_k = a_2, Z_{k-1} = i, Z_k = j) \\
&= \sum_{u_1 \in T_{k-1}} \sum_{u_2 \in T_k} P(Y_{k-1} = a_1, Y_k = a_2, Z_{k-1} = i, Z_k = j, U_{k-1} = u_1, U_k = u_2) \\
&= \sum_{u_1 \in T_{k-1}} \sum_{u_2 \in T_k} P(Y_k = a_2 | Y_{k-1} = a_1, Z_{k-1} = i, Z_k = j, U_{k-1} = u_1, U_k = u_2) \\
&\quad \times P(Y_{k-1} = a_1, Z_{k-1} = i, Z_k = j, U_{k-1} = u_1, U_k = u_2) \\
&= \sum_{u_1 \in T_{k-1}} \sum_{u_2 \in T_k} P(Y_k = a_2 | Z_k = j) P(Z_{k-1} = i, U_{k-1} = u_1, Y_{k-1} = a_1) \\
&\quad \times P(Z_k = j, U_k = u_2 | Z_{k-1} = i, U_{k-1} = u_1, Y_{k-1} = a_1) \\
&= \sum_{u_1 \in T_{k-1}} \sum_{u_2 \in T_k} P(Y_k = a_2 | Z_k = j) P((i, u_1), (j, u_2)) \\
&\quad \times P(Y_{k-1} = a_1 | Z_{k-1} = i) P(Z_{k-1} = i, U_{k-1} = u_1) \\
&= \sum_{u_1 \in T_{k-1}} \sum_{u_2 \in T_k} R_{j;a_2} P((i, u_1), (j, u_2)) R_{i;a_1} (aP)^{k-1}(i, u_1). \tag{7}
\end{aligned}$$

Therefore, from equations (6) and (7), we obtain the desired result.

### 3.5.2 Hidden Markov models

In the case of hidden Markov models,  $\mathbf{Z} = (Z_k)_{k \in \mathbb{N}}$  is a Markov chain and denote by  $\mathbf{P} = (p_{ij}; i, j \in E)$ , where  $p_{ij} = P(Z_{k+1} = j | Z_k = i)$ , the corresponding transition kernel.

**Theorem 8** *The ROCOF at time  $k \in \mathbb{N}^*$  is given by*

$$r(k) = \sum_{a_1 \in U} \sum_{a_2 \in D} \sum_{i \in E} \sum_{j \in E} \alpha(i_1) p_{i_1 i}^{(k-1)} p_{ij} R_{i;a_1} R_{j;a_2}.$$

*Proof* For fixed  $k \in \mathbb{N}^*$

$$\begin{aligned}
r(k) &= \mathbb{E}[N(k) - N(k-1)] \\
&= \mathbb{E} \mathbf{1}_{\{Y_{k-1} \in U, Z_k \in D\}} \\
&= P(Y_{k-1} \in U, Y_k \in D) \\
&= \sum_{a_1 \in U} \sum_{a_2 \in D} P(Y_{k-1} = a_1, Y_k = a_2). \tag{8}
\end{aligned}$$

On the other hand

$$\begin{aligned}
P(Y_{k-1} = a_1, Y_k = a_2) &= \sum_{i \in E} \sum_{j \in E} P(Y_{k-1} = a_1, Y_k = a_2, Z_{k-1} = i, Z_k = j) \\
&= \sum_{i \in E} \sum_{j \in E} P(Y_k = a_2 | Y_{k-1} = a_1, Z_{k-1} = i, Z_k = j) \\
&\quad \times P(Y_{k-1} = a_1, Z_{k-1} = i, Z_k = j) \\
&= \sum_{i \in E} \sum_{j \in E} \alpha(i_1) p_{i_1 i}^{(k-1)} p_{ij} R_{i; a_1} R_{j; a_2}. \tag{9}
\end{aligned}$$

Therefore, from equations (8) and (9), we get the desired result.

The estimation of the ROCOF for hidden (semi-) Markov models could be made by means of plug-in type estimators. These estimators could be expressed in terms of the maximum likelihood estimators of the emission, transition and semi-Markov kernels that have appealing asymptotic properties (see [3],[14]).

## 4 Numerical Examples

### 4.1 Academic example

To validate our results and see how we can employ the conditional ROCOF to quantify the effects of the starting functioning state, we adopt simulation methods. In particular, we use the algorithm presented by [3], which generates trajectories of an MRC in a fixed time interval  $[0, M]$  (Algorithm 1).

```

Initialization;
Set  $k = 0$ ,  $S_0 = 0$  and sample  $J_0$  from the initial distribution  $\alpha$ ;
Iteration;
while  $S_k \leq M$  do
  Sample the random variable  $J \sim \mathbf{P}(J_k, \cdot)$  and set  $J_{k+1} = J(\omega)$ ;
  Sample the random variable  $X \sim f_{J_k J_{k+1}}(\cdot)$ ;
  Set  $S_{k+1} = S_k + X$ ;
  Set  $k = k + 1$ ;
end

```

**Algorithm 1:** Monte-Carlo algorithm for the generation of a trajectory of an MRC,  $\mathcal{H}(M)$ .

We generate a trajectory of the MRC with length equal to  $M = 100000$ . The initial law  $\alpha$  and the transition kernel  $\mathbf{P}$  are given by

$$\alpha = (0.5 \ 0.5 \ 0) \quad \text{and} \quad \mathbf{P} = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.7 & 0 & 0.3 \\ 0.5 & 0.5 & 0 \end{pmatrix},$$

respectively. The sojourn times follow the discrete time Weibull distribution

$$f_{ij}(k) = \begin{cases} q^{(k-1)b} - q^{kb}, & \text{if } k \geq 1, \\ 0, & \text{if } k = 0, \end{cases}$$

with parameters  $(q, b) = (0.1, 0.9)$  for the transitions  $1 \rightarrow 2$  and  $2 \rightarrow 1$ ,  $(q, b) = (0.1, 2.0)$  for the transition  $2 \rightarrow 3$  and  $(q, b) = (0.6, 0.9)$  for the remaining transitions. We further denote by  $U = \{1, 2\}$  the subset of up states and by  $D = \{3\}$  the subset of down states. Our objective here is to estimate the ROCOF, quantify the impact of the starting functioning states on the

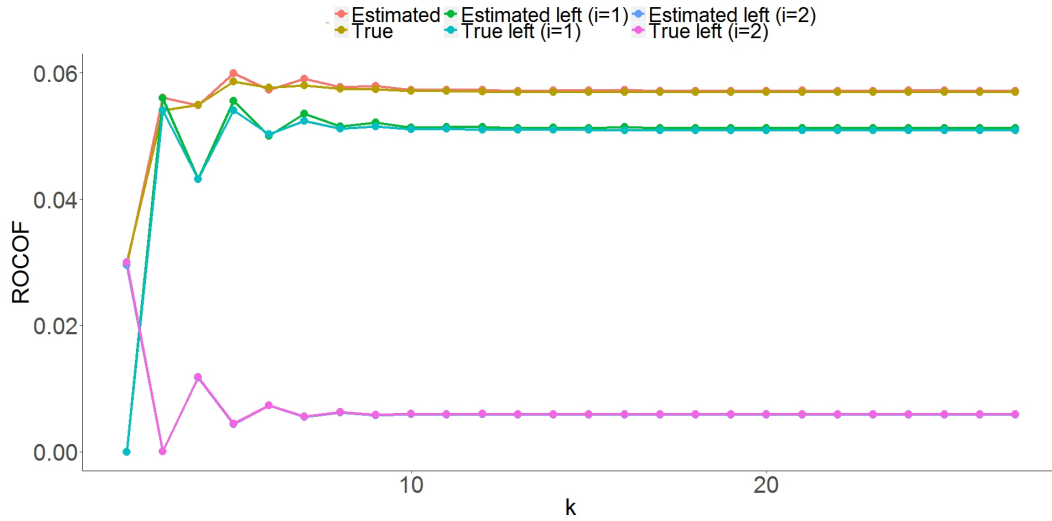


Fig. 2: (Conditional) ROCOF versus time based on simulated data.

estimator and show the strong consistency of the estimators. We first notice that the ROCOF tends to increase for  $k$  smaller than 8, which means that the system is deteriorating, and then it becomes constant due to the stationarity of the semi-Markov chain. The estimators tend to the corresponding true values as  $M$  increases. Moreover, we observe that the ROCOF is mostly affected by the up state  $i = 1$ , since the contribution of this state represents 83 % of its total value. The right conditional ROCOF in this example is equal to the ROCOF as there is only one down state. This example shows us concretely that the failure occurrence rate is mostly determined by the transition from state 1 to state 3. This information could be used to improve the performances of the system by adapting the maintenance policy.

#### 4.2 Real data example

The conditional ROCOF is further evaluated on real data representing ground accelerations measured in  $g$  ( $1g$  stands for an acceleration of  $9.81 \text{ m/s}^{-2}$ ) and retrieved from the online source <http://realvibrations.nipslab.org/>. It is well known that certain levels of vibration could degrade the good functioning of some small vibration sensitive equipments, such as microelectronics devices, and therefore could have a significant impact on their performance. In the present work, we suppose that an acceleration whose magnitude (absolute value) is superior to  $0.4g$  could lead to degradation. Fig. 3 represents the selected signal which is classified in four states (Table 1), where the subset of up states is  $U = \{2, 3\}$  and the subset of down states is  $D = \{1, 4\}$ .

State	Ground acceleration (in $g$ )
1	$[-4.0, -0.4)$
2	$[-0.4, +0.0)$
3	$[+0.0, +0.4)$
4	$[+0.4, +0.8)$

Table 1: States classification.

We first notice that the system is deteriorating for  $k$  smaller than 27 since the ROCOF tends to increase. Then its performance tends to stabilize since the ROCOF becomes constant.

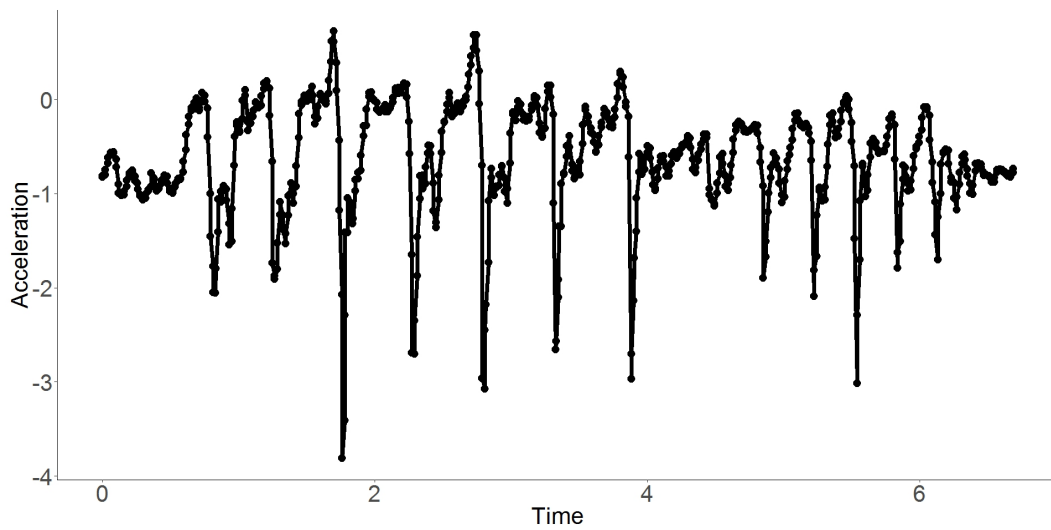
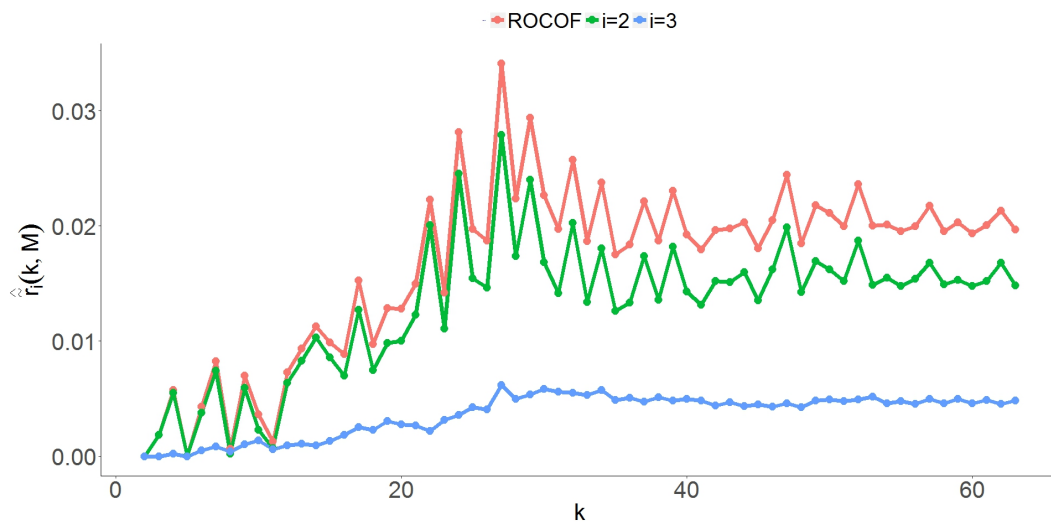
Fig. 3: Ground acceleration (in  $g$ ) versus time.

Fig. 4: (Conditional) ROCOF versus time based on real data.

Furthermore, we observe that the ROCOF is affected by both up states and mostly by the second up state. This information could be helpful in determining alert criteria associated with the deterioration of the system under study and passive vibration damping devices could be used to attenuate the vibration levels [7].

## 5 Concluding Remarks

In the present paper we study the sensitivity of the ROCOF in the current up and ending down states in a semi-Markov context. This is the first attempt to arise sensitivity issues of reliability indicators for SMMs. The results highlight that defining more precise reliability indicators could lead to more exact results for practitioners and researchers. Due to the flexibility and universality of the framework provided by SMMs, there exists a variety of extensions that could be adapted. Further research includes the extension of the results to higher order or non-homogeneous SMMs. ROCOF could further be investigated when independent and identical



copies of the process are observed, each over a fixed duration, instead of one single copy over a fixed time interval. Modelling the reliability of mechanical systems subjected to earthquake-like loads is a subject of major concern and SMMs can be used to achieve this goal. An application to non-linear oscillators under seismic forces is a topic of further work.

## Acknowledgements

The author is grateful to Mohamed Hamdaoui, Assistant Professor at the University of Lorraine, for fruitful discussions and kind assistance.

## References

1. Barbu, V., Karagrigoriou, A., Makrides, A. (2017), "Semi-Markov modelling for multi-state systems", *Methodol Comput Appl Probab*, 19(4), 1011–1028.
2. Barbu, V., Bulla, J., Maruotti, A. (2012), "Estimation of the stationary distribution of a semi-Markov chain", *Reliab. Stat. Stud.*, 5, 15–26.
3. Barbu, V., Limnios, N. (2008), Semi-Markov chains and hidden semi-Markov models toward applications: Their Use in Reliability and DNA Analysis, Lecture Notes in Statistics, Springer, New York.
4. Chryssaphinou, O., Karaliopoulou, M., Limnios, N. (2008), "On discrete time semi-Markov chains and applications in word occurrences", *Commun Stat Theory Methods*, 37, 1306–1322.
5. Georgiadis, S. (2017), "First hitting probabilities for semi-Markov chains and estimation", *Commun Stat Theory Method*, 46(5), 2435–2446.
6. Georgiadis, S., Limnios, N., Votsi, I. (2013), Reliability and probability of first occurred failure for discrete-time semi-Markov systems. In: Frenkel, I., Karagrigoriou, A., Lisnianski, A., Kleyner, A., eds. Applied Reliability Engineering and Risk Analysis: Probabilistic Models and Statistical Inference (chap. 12, 167179). New York: Wiley & Sons.
7. Xu, C., Wu, M-Z, Hamdaoui, M. (2016), "Mixed integer multi-objective optimization of composite structures with frequency-dependent interleaved viscoelastic damping layers", *Computers & Structures*, 172, 81–92.
8. Howard, R. (1971), Dynamic probabilistic systems, (Vol. II), New York, Wiley.
9. Limnios, N. (2012), "Reliability measures of semi-Markov systems with general state space", *Methodol Comput Appl Probab*, 14(4), 895–917.
10. Limnios, N., Oprisan, G. (2001), Convergence of probability measures, Wiley, New York, 2nd edition.
11. Limnios, N., Ouhbi, B. (2006), "Nonparametric estimation of some important indicators in reliability for semi-Markov processes", *Stat Methodol*, 3, 341–350.
12. Mode, C.J., Sleeman, C.K. (2000), Stochastic Processes in Epidemiology: HIV/AIDS, Other Infectious Diseases, and Computers, Singapore, World Scientific.
13. Ouhbi, B., Limnios, N. (2002), "The rate of occurrence of failures for semi-Markov processes and estimation", *Stat Probab Lett*, 59(3), 245–255.
14. Sadek, A., Limnios, N. (2002), "Asymptotic properties for maximum likelihood estimators for reliability and failure rates of Markov chains", *Commun Stat Theory Method*, 31(10), 1837–1861.
15. Trevezas, S., Limnios, N. (2011), "Exact MLE and asymptotic properties for nonparametric semi-Markov models", *J Nonparametr Stat*, 23(3), 719–739.
16. van der Vaart, A.W. (1998), Asymptotic Statistics, Cambridge University Press, New York.
17. Vassiliou, P.-C.G., Papadopoulou, A.A. (1992), "Asymptotic behavior of nonhomogeneous semi-Markov systems", *Linear Algebra Appl*, 210, 153–198.
18. Vassiliou, P.-C.G., Papadopoulou, A.A. (1992), "Non-homogeneous semi-Markov systems and maintainability of the state sizes", *J Appl Probab*, 29, 519–534.
19. Votsi, I., "Reliability indicators for hidden Markov renewal models", *Reliability Engineering: Theory and Applications*, edited by Vonta, I. and Ram, M., Boca Raton: CRC Press.
20. Votsi, I., Limnios, N. (2015), "Estimation of the intensity of the hitting time for semi-Markov chains and hidden Markov renewal chains", *J Nonparametr Stat*, 27 (2), 149–166.
21. Votsi, I., Limnios, N., Tsaklidis, G., Papadimitriou, E. (2014), "Hidden semi-Markov modeling for the estimation of earthquake occurrence rates", *Commun Stat Theory Methods*, 43, 1484–1502.
22. Votsi, I., Limnios, N., Tsaklidis, G., Papadimitriou, E. (2012), "Estimation of the expected number of earthquake occurrences based on semi-Markov models", *Methodol Comput Appl Probab*, 14(3), 685–703.
23. Yeh, L. (1997), "The rate of occurrence of failures", *J Appl Probab*, 34, 234–247.