

Article

Open Markov Type Population Models: From Discrete to Continuous Time

Manuel L. Esquível ^{1,*} , Nadezhda P. Krasii ²  and Gracinda R. Guerreiro ¹ ¹ Department of Mathematics, FCT NOVA, and CMA New University of Lisbon, Campus de Caparica, 2829-516 Caparica, Portugal; grg@fct.unl.pt² Department of Higher Mathematics, Don State Technical University, 344000 Rostov-on-Don, Russia; krasnad@yandex.ru

* Correspondence: mle@fct.unl.pt

Abstract: We address the problem of finding a natural continuous time Markov type process—in open populations—that best captures the information provided by an open Markov chain in discrete time which is usually the sole possible observation from data. Given the open discrete time Markov chain, we single out two main approaches: In the first one, we consider a calibration procedure of a continuous time Markov process using a transition matrix of a discrete time Markov chain and we show that, when the discrete time transition matrix is *embeddable* in a continuous time one, the calibration problem has optimal solutions. In the second approach, we consider semi-Markov processes—and open Markov schemes—and we propose a direct extension from the discrete time theory to the continuous time one by using a known structure representation result for semi-Markov processes that decomposes the process as a sum of terms given by the products of the random variables of a discrete time Markov chain by time functions built from an adequate increasing sequence of stopping times.



Citation: Esquível, M.L.; Krasii, N.P.; Guerreiro, G.R. Open Markov Type Population Models: From Discrete to Continuous Time. *Mathematics* **2021**, *9*, 1496. <https://doi.org/10.3390/math9131496>

Academic Editors: Panagiotis-Christos Vassiliou and Andreas C. Georgiou

Received: 31 May 2021
Accepted: 23 June 2021
Published: 25 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: Markov chains; open population Markov chain models; Semi-Markov processes

1. Introduction

After the first works introducing homogeneous open Markov population models in [1] followed by those in [2] and then in [3], further expanded by several authors and exposed in [4] and then in [5], the study of open populations in a finite state space in discrete time with a Markov chain structure became well established.

Following the pioneering work of Gani, introducing in [6] what now is known as *Cyclic Open Markov* population models, there were further extensions in [7], for non-homogeneous Markov chains and then, for cyclic non-homogeneous Markov systems or equivalently for non-homogeneous open Markov population processes, by the authors of [8,9]. Let us stress that continuous time non-homogeneous Markov systems have been studied lately in [10]. Furthermore, the recent work in [11] develops an approach to open Markov chains in discrete time—allowing a particle physics interpretation—for which there is a state space of the Markov chain—where distributions are studied by means of moment generating functions—there is an exit *reservoir*, which is tantamount to a cemetery state and, there is an incoming flow of particles, defined as a stochastic process in discrete time whose properties—e.g., stationarity—condition the distribution law of the particles in the state space.

Discrete time non-homogeneous semi-Markov systems or equivalently open semi-Markov population models were introduced and studied in [12,13]. The study of open populations in a finite state space in continuous time and governed by Markov laws, has already been carried in [14] and the references therein, and extensions to a general state space have been given in [15–17]. The continuous time framework has also been

addressed, for instance, in [18–20], for the case of semi-Markov processes and for non-homogeneous semi-Markov systems [21]. We may also refer a framework of open Markov chains with finite state space—see in [22] and references therein—that has already seen applications in Actuarial or Financial problems—as, for instance, in [23,24]—but also in population dynamics (see [25]). The weaker formalism open Markov schemes, in discrete time—developed in [26]—allows for influxes of new elements in the population to be given as general time series models.

Another example was motivated by the study of a continuous time non homogeneous Markov chain model for Long Term Care, based on an estimated Markov chain transition matrix with a finite state space, in [27], by means of a method for calibrating the intensities on the continuous time Markov chain using the discrete time transition matrix in the context of usual existence theorems for ordinary differential equations (ODE); this method will be considered, in Section 3.2, in the more general context of Caratheodory existence theorems for ODE.

The main contribution of the present work is to extend results on open Markov chains in discrete time to some continuous time process of Markov type using different methods of associating a continuous process to an observed process in discrete time. One of these methods—presented in Sections 3.2 and 3.3—is by calibration of the transition intensities. Another method considered for open Markov schemes—in Section 4.2 and also, briefly, for some particular cases, in Section 4.3—is to exploit a natural representation of the continuous time Markov type process, in Formula (2) of Section 2.

2. From Discrete Time to Continuous Time via a Structural Approach

We present the main ideas on a structural representation for continuous time process of Markov type that are crucial to our approach. The structure of continuous time processes—for instance, Markov, semi-Markov, and Markov type schemes processes—allows us to consider a fairly general representation formula—Formula (2)—decoupling the continuous time process as a discrete time process and a sequence of time functions depending on the sequence of the jump stopping times.

Consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a continuous time stochastic process $(Y_t)_{t \geq 0}$ defined on this probability space and $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ the natural filtration associated to this process, that is, such that $\mathcal{F}_t := \sigma(Y_s : s \leq t)$ is the algebra- σ generated by the variables of the process until time t . Consider also a sequence of random variables $(Z_n)_{n \geq 0}$ taking values in a finite state space $\Theta = \{\theta_1, \theta_2, \dots, \theta_r\}$, the sequence being adapted to the filtration \mathbb{F} and $0 \equiv \tau_0 < \tau_1 < \tau_2 < \dots < \tau_n < \dots$ an increasing sequence of \mathbb{F} -stopping times, denoted by \mathcal{T} , satisfying the following hypothesis:

Hypothesis 1. *Almost surely, $\lim_{n \rightarrow +\infty} \tau_n = +\infty$ and, for any $T \in \mathbb{R}_+$ and almost all $\omega \in \Omega$:*

$$\#\{k \geq 1 : \tau_k(\omega) \leq T\} < +\infty. \quad (1)$$

This hypothesis means that in every compact time interval $[0, T]$, for almost all $\omega \in \Omega$, there is only a finite number of stopping times realizations $\tau_k(\omega)$ in this interval.

Hypothesis 2. *The continuous time process $(Y_t)_{t \geq 0}$ admits a representation given, for $t \geq 0$, by*

$$Y_t = \sum_{n=0}^{+\infty} Z_n \mathbb{1}_{[\tau_n, \tau_{n+1}[}(t), \quad (2)$$

that is, a hypothesis on the structure of the continuous time process $(Y_t)_{t \geq 0}$.

It is well known—see in [28] (pp. 367–379) and in [29] (pp. 317–320)—that if $(Z_n)_{n \geq 0}$ is a Markov chain and the time intervals $(\tau_{n+1} - \tau_n)_{n \geq 0}$ are Exponentially distributed then $(Y_t)_{t \geq 0}$ can be taken to be a continuous time **homogeneous** Markov chain. If $(Z_n)_{n \geq 0}$ is a Markov chain and the time intervals $(\tau_{n+1} - \tau_n)_{n \geq 0}$ have a distribution that can depend

on the present state as well as on the one visited next then $(Y_t)_{t \geq 0}$ can be taken to be a **semi-Markov** process (see in [30] (pp. 261–262) and in [31] (pp. 295–299), for brief references). In the case of a semi-Markov processes, a nice result of Ronald Pyke (see in [32] (p. 1236)), reproduced ahead in Theorem A7, guarantees that when the state space is finite the process is *regular* implying that almost all paths of such a semi-Markov process are step-functions over $[0, +\infty[$ and so, the paths satisfy Formula (1). In another important case (see Theorems A5 and A6 ahead, or [30] (pp. 262–266) and [31] (pp. 195–244)), adequate hypothesis on the distribution of the stopping times and on the sequence $(Z_n)_{n \geq 0}$ implies that $(Y_t)_{t \geq 0}$ will be a **non homogeneous** Markov chain process in continuous time, whose trajectories are step functions also satisfying Formula (1). The representation in Formula (2), thus covers the cases of homogeneous and non homogeneous Markov processes in continuous time as well as semi-Markov processes, providing a desired connection between a continuous time process and a discrete one that is a component of the former. We observe that there is a practical justification for Hypothesis 1, namely, the *identifiability* of the process; as can be read in [33] (p. 3): “... Actually, in real systems the transition from one observable state into another takes some time.” Being so, the existence of accumulation points in a compact interval would preclude estimation procedures for instance of the distribution of the sequence $(\tau_{n+1} - \tau_n)_{n \geq 1}$.

3. From Discrete to Continuous Time Markov Chains: A Calibration Approach

In this section, we consider a calibration approach in order to determine a set of probability densities that best approaches a sequence of discrete time transition matrices with respect to a quadratic loss function. We then show that *embeddable* stochastic matrices, according to Definition 1, are solutions of the calibration problem. For the reader’s convenience, we recall in the first appendix the most important results on continuous time Markov chains with finite state space that are relevant for our study with emphasis on the crucial non-accumulation property of the jump times of a continuous time Markov chain (see Theorem A6 ahead). We will start by recalling the main information on embeddable chains. We then present one of the main contributions of this work, that is, a general result on the optimization problem of calibration and its relations with embeddable properties of discrete time Markov chains.

3.1. The Embedding of a Discrete Time Markov Chain in a Continuous One

The embedding of the discrete time Markov chain in a continuous one following the guidelines, for instance, in [34–40], can be considered as a method to connect a discrete time process with a continuous one. For notations on non-homogeneous continuous time Markov chains see Section 3.2.

Definition 1 (Embeddable stochastic matrix (see [38])). *A stochastic matrix R is said to be embeddable if there exists a time $t_R > 0$ and a family of stochastic matrices $P(s, t)$ continuously defined in the set of times $\{(s, t) \in \mathbb{R}^2 : 0 \leq s \leq t \leq t_R\}$ such that*

$$\begin{cases} P(s, t) = P(s, u)P(u, t) & 0 \leq s \leq u \leq t \leq t_R \\ P(s, s) = I & 0 \leq s \leq t_R. \\ P(0, t_R) = R. \end{cases} \quad (3)$$

We observe that by Theorem A2 ahead, the condition in Formulas (3) is tantamount to the definition of a continuous time Markov chain with transition probabilities given by $P(s, t)$.

Remark 1 (Intrinsic time for embeddable chains). *Goodman in [41]—aiming at a more general result for the Kolmogorov differential equations—showed that with the change of time given by*

$\varphi(u) := -\log \det P(0, u)$ —which amounts to a change in the matrix coefficients of $P(s, t)$ —we have that

$$t_R = -\log \det R. \quad (4)$$

This remarkable representation for the embedding time t_R will be useful for a result in Section 3.2 devoted to the calibration approach. It has also been used for estimation in [42] (p. 330).

See the work in [35] for a definition similar to Definition 1 and for a summary of many important results on this subject. The characterization of an embeddable stochastic matrix in a form useful for practical purposes was recently achieved in [43]. More useful results were obtained in [44]. The connections between this kind of embedding and the other approaches, for the association of a discrete time Markov chain and a continuous time process, deserve further study.

3.2. Continuous Time Markov Chains Calibration with a Discrete Time Markov Transition Matrix

The calibration of transition intensities of a non homogeneous Markov chain, with a discrete time Markov chain transition matrix estimated from data, was proposed in [27]. In this section, we establish a general formulation of the existence a unicity result that subsumes the approach and we establish a connection with the *embedding* approach of Section 3.1. Notation and needed essential results on non-homogeneous Markov processes in continuous time were recalled in Appendix A.

The procedure for calibration of intensities consists in finding the intensities of a non homogeneous continuous time Markov chain using a probability transition matrix of a discrete time Markov chain and a given loss function—having as arguments the transition probabilities of the continuous time Markov chain and some function of the transition matrix of the discrete time Markov chain—in such a way that the loss function is minimized.

Previously to the consideration of the theorem on the calibration of intensities we discuss some motivation for this approach. It may happen that a phenomena that could be dealt—due to its characteristics—with a continuous time Markov chain model can only be observed at regularly spaced time intervals. This is the case of the periodic assessments of the healthcare status of patients that can change at any time but are only object of a comprehensive evaluation on, say, a weekly basis. With the data originated by these observations we can only determine transition probabilities—for a defined period, say, a week—and, most importantly we cannot determine the time stamps for the patient status change. The question naturally poses itself: is it possible to associate—in some canonical way—to an estimated discrete time Markov chain transition matrix a process in continuous time that encompasses the discrete time process? First steps in this direction are provided by Theorem 1 that we now present and the following Theorems 2 and 3.

We formulate Theorem 1 in the context of Caratheodory's general existence theory of solutions of ordinary differential equations that we briefly recall. One reason for this choice is that according to [41] (p. 169) and we quote: “... This fact gives further evidence in support of the view that Caratheodory equations occupy a natural place in the theory of non-stationary Markov chains.” Another reason is the fact that Caratheodory existence theory is particularly suited for regime switching models and these models are the object of Theorem 3 ahead. Following the work in [45] (pp. 41–44), we consider the definition of an **extended solution** for a Cauchy problem of a differential equation,

$$Y'(t) = f(t, Y(t)), \quad Y(0) = \xi, \quad (5)$$

or formulated in an equivalent form,

$$Y(t) = \xi + \int_0^t f(s, Y(s)) ds, \quad (6)$$

for $f(t, y) : I \times \mathcal{D} \rightarrow \mathbb{R}^r$ a non-necessarily continuous function, with $I \subset [0, +\infty[$ and $\mathcal{D} \subset \mathbb{R}^r$, to be an **absolutely continuous** function $Y(t)$ (see [46], pp. 144–150) such that

$f(t, Y(t)) \in \mathcal{D}$ for $t \in I$ and Formula (5) is verified for all $t \in I$ possibly with the exception of a set of null Lebesgue measure. The well-known Caratheodory's existence theorem (see in [45], p. 43) ensures the existence of an extended solution with a given initial condition—given in a neighborhood of the initial time—under the conditions that $f(t, y)$ is measurable in the variable t , for fixed y , and continuous in the variable y , for fixed t , and moreover that there exists a Lebesgue integrable function $m(t)$, defined on a neighborhood of the initial time, let us say I , such that $|f(t, y)| \leq m(t)$ for $(t, y) \in I \times \mathcal{D}$. The question of unicity of the solution is dealt, usually, either directly using Theorem 18.4.13 in [47] (p. 337) or using Osgood's uniqueness theorem—as exposed, for instance, in [48] (p. 58) or in [49] (pp. 149–151)—to conclude that the extended solution—that with Caratheodory's theorem we know to exist—is unique in the sense that two solutions may only differ on a set of Lebesgue measure equal to zero. For our purposes we need an existence and unicity theorem for ordinary differential equations with solutions depending continuously on a parameter such as the general result of Theorem 4.2 in [45] (p. 53) with an omitted proof that follows for a lengthy previous exposition of related matters. For completeness we now establish a result that is suited to our purposes as it deals with the particular type of Kolmogorov equations for continuous time Markov chains.

Theorem 1 (Calibration of intensities with Caratheodory's type ODE existence theorem hypothesis). *Let, for $1 \leq n \leq N$, $\mathbf{R}^{\tau_n} = [r_{ij}^{(\tau_n)}]_{i,j=1,\dots,r}$ be the generic element of a sequence of numerical transition matrices taken at sequence of increasing dates $(\tau_n)_{1 \leq n \leq N}$. Consider a set of intensities $\mathbf{Q}(t, \lambda) = [q(u, i, j, \lambda)]_{i,j=1,\dots,r}$ —with $\lambda \in \Lambda \subset \mathbb{R}^d$ being a parameter and Λ being a compact set—satisfying the following conditions:*

1. *For every fixed λ the functions $q(u, i, j, \lambda)$ are measurable as functions of u .*
2. *For every fixed u the functions $q(u, i, j, \lambda)$ are continuous as functions of λ .*
3. *There exists a locally integrable function $M : [0, +\infty[\rightarrow [0, +\infty[$, such that for all $\lambda \in \Lambda$, $i \in \mathcal{I}$, $u \in [0, +\infty[$ and $0 \leq s \leq t$, the following conditions are verified:*

$$-q(u, i, i, \lambda) \leq M(u) \text{ and } \int_s^t M(u) du < +\infty. \quad (7)$$

Then, we have

1. *There exists $\mathbf{P}(s, t, \lambda) = [p(s, i, t, j, \lambda)]_{i,j=1,\dots,r}$ a probability transition matrix, with entries absolutely continuous in s and t , such that conditions in Definition A2, the Chapman–Kolmogorov equations in Theorem A1 and Theorem A3 are verified.*
2. *For each fixed s_0 , consider the loss function*

$$\mathcal{O}(s_0, \lambda) := \sum_{i,j=1,\dots,r} \sum_{n=1}^N \left(p(s_0, i, \tau_n, j, \lambda) - r_{ij}^{(\tau_n)} \right)^2. \quad (8)$$

Then, for the optimization problem $\inf_{\lambda \in \Lambda} \mathcal{O}(s_0, \lambda)$ there exists $\lambda_0 \in \Lambda$ such that

$$\mathcal{O}(s_0, \lambda_0) = \min_{\lambda \in \Lambda} \mathcal{O}(s_0, \lambda), \quad (9)$$

the unique minimum being attained at possibly several points $\lambda_0 \in \Lambda$.

Proof. We will prove, simultaneously, the existence of the probability transition matrix, the unicity in the extended solution sense and the continuous dependence of the parameter $\lambda \in \Lambda$ following the lines of the proof of the result denominated Hostinsky's representation (see in [29], pp. 348–349). As we suppose that Λ is compact, the continuity of $\mathbf{P}(s_0, t, \lambda)$, as a function of $\lambda \in \Lambda$ for every fixed t , will be enough to establish the second thesis.

We want to determine an extended solution of the Kolmogorov forward equation given in Formula (A11), that is an extended solution of

$$\begin{cases} \mathbf{P}'_t(s_0, t, \lambda) = \mathbf{P}(s_0, t, \lambda) \mathbf{Q}(t, \lambda) \\ \mathbf{P}(t, t) = \mathbf{I}, \end{cases} \quad (10)$$

an equation which, as seen in Formula (A12), can be read in integral form as,

$$\mathbf{P}(s_0, t, \lambda) = \mathbf{I} + \int_{[s_0, t]} \mathbf{P}(s_0, s, \lambda) \mathbf{Q}(s, \lambda) ds. \quad (11)$$

As previously said, we will now follow the general idea of successive approximations in the proof of the Picard–Lindelöf theorem for proving existence and unicity of solutions of ordinary differential equations for the forward Kolmogorov equation. By replacing $\mathbf{P}(s_0, s, \lambda)$ in the right-hand member of Equation (11) by this right-hand member we get,

$$\mathbf{P}(s_0, t, \lambda) = \mathbf{I} + \int_{[s_0, t]} \mathbf{Q}(s, \lambda) ds + \int_{[s_0, t]} \int_{[s_0, t_1]} \mathbf{P}(s_0, t_2, \lambda) \mathbf{Q}(t_1, \lambda) \mathbf{Q}(t_2, \lambda) dt_2 dt_1$$

and, by induction, we obtain

$$\begin{aligned} \mathbf{P}(s_0, t, \lambda) &= \mathbf{I} + \int_{[s_0, t]} \mathbf{Q}(s, \lambda) ds + \\ &+ \sum_{n=2}^k \int_{[s_0, t]} \int_{[t_1, t]} \cdots \int_{[t_{n-1}, t]} \mathbf{Q}(t_1, \lambda) \mathbf{Q}(t_2, \lambda) \cdots \mathbf{Q}(t_n, \lambda) dt_n \cdots dt_1 + \\ &+ \int_{[s_0, t]} \int_{[t_1, t]} \cdots \int_{[t_{k-1}, t]} \mathbf{P}(s_0, t_k, \lambda) \mathbf{Q}(t_1, \lambda) \mathbf{Q}(t_2, \lambda) \cdots \mathbf{Q}(t_k, \lambda) dt_k \cdots dt_1. \end{aligned}$$

Now, considering the function $M(t)$ in the third hypothesis stated above about the intensity matrix, we have that, by Lemma A1 (see also Lemma 8.4.1 in [29], p. 348), since $M(t)$ is integrable over any compact set, considering the (i, j) component of the $r \times r$ matrix, we have that

$$\begin{aligned} &\left| \left[\int_{[s_0, t]} \int_{[t_1, t]} \cdots \int_{[t_{k-1}, t]} \mathbf{P}(s_0, t_k, \lambda) \mathbf{Q}(t_1, \lambda) \mathbf{Q}(t_2, \lambda) \cdots \mathbf{Q}(t_k, \lambda) dt_k \cdots dt_1 \right]_{ij} \right| \leq \\ &\leq r^k \int_{[s_0, t]} \int_{[t_1, t]} \cdots \int_{[t_{k-1}, t]} M(t_1) M(t_2) \cdots M(t_k) dt_k \cdots dt_1 = \\ &= \frac{\left(r \int_{[s_0, t]} M(s) ds \right)^k}{k!}. \end{aligned}$$

Finally, as

$$\lim_{k \rightarrow +\infty} \frac{\left(r \int_{[s_0, t]} M(s) ds \right)^k}{k!} = 0,$$

we have that the series for which the sum represents $\mathbf{P}(x, t, \lambda)$, that is,

$$\mathbf{P}(s_0, t, \lambda) = \mathbf{I} + \sum_{n=1}^{+\infty} \left(\int_{[s_0, t]} \int_{[t_1, t]} \cdots \int_{[t_{n-1}, t]} \mathbf{Q}(t_1, \lambda) \mathbf{Q}(t_2, \lambda) \cdots \mathbf{Q}(t_n, \lambda) dt_n \cdots dt_1 \right),$$

is a series—of absolutely continuous functions of the variable t which are also continuous as functions of the parameter $\lambda \in \Lambda$ —converging normally and so the sum is an absolutely continuous function of the variable t and continuous function of the parameter λ . With a similar reasoning applied to the backward Kolmogorov equation we also have that $\mathbf{P}(s, t_0, \lambda)$ is absolutely continuous in the variable s and, obviously, continuous as a function

of the parameter $\lambda \in \Lambda$. We observe that it was stated in [41], pp. 166–167 (with a reference to a proof in [50] and proved also in [51]), that the separate absolute continuity of $P(s, t, \lambda)$ in the variables s and t ensures the uniqueness of the solution. \square

Remark 2 (An alternative path for the existence result). *We observe that, for every fixed value of the parameter λ , by a direct application of Caratheodory's existence theorem to the forward and backward Kolmogorov equations in Theorem A3, we obtain a probability transition matrix $P(s, t, \lambda) = [p(s, i, t, j, \lambda)]_{i,j=1,\dots,r}$, such that conditions in Definition A2 and the Chapman–Kolmogorov equations in Theorem A1 are verified, that in addition has entries absolutely continuous in s and t and such that Kolmogorov's equations are satisfied almost everywhere. With this approach the continuous dependence of the probability transition matrix on the parameter λ requires further proof.*

Remark 3 (On the parametrized intensities and transition probabilities). *In a first application to Long-Term Care of a simpler version of Theorem 1 presented in [27], we chose as intensities a parametrized family—of Gompertz–Makeham type (see, for instance, in [52], p. 62)—with a three dimensional parameter. We observe that, in its actual formulation, Theorem 1 contemplates the case of a set of intensities—and of associated transition probabilities—not necessarily with the same functional form with varying parameters but merely with a finite set of different functional forms indexed by the parameters.*

Remark 4 (Only one transition matrix observation). *In the case where we only have one estimated transition matrix R , we can consider the sequence of n step transition matrices given by the n fold product of the matrix R by itself. This situation will be addressed in Theorem 2 ahead, in the case of homogeneous Markov chains and in Theorem 3 for the non-homogeneous case.*

We also observe that in the case of a multidimensional parameter set Λ —say r_1 —and even in a reasonable state space of the discrete time Markov chain—say with r_2 states—the optimization problem of Formula (8) may require adequate algorithms to be solved as the number of variables is of the order of $r_1 \times r_2 \times (r_2 - 1)$. In [27] we opted for a modified grid search coupled with the numerical solutions of the Kolmogorov equations in order to recover the transition probabilities of the continuous time Markov chain.

Remark 5 (On the unicity of the solution of the calibration problem). *The unicity in law of the solution of the calibration problem deserves discussion. If there are several minimizers of the calibration problem, to each of these minimizers corresponds an intensity and to each intensity a , possible, different law for the stopping times of the continuous time Markov chain, as these laws are determined by the intensities (see Remark A2). The existence of criteria allowing to identify a distribution of inter-arrival times that stochastically dominates all other solutions is an open problem.*

We can establish a connection between the approach in Section 3.1 and Theorem 1 on calibration above, showing first—in Theorem 2—that, if a matrix is embeddable in a homogeneous continuous time Markov chain—with intensities depending continuously on a parameter—for a fixed value of the parameter, then this continuous time Markov chain solves the calibration problem in an optimum way. We recall that the continuous time Markov chain is homogeneous if, for all $0 \leq s, t$ the transition probabilities satisfy

$$P(s, s + t) = P(0, t),$$

and that the intensities matrix is constant as a function of time (see [41] (pp. 165–166) for definitions in this context).

Theorem 2 (Discrete chains embeddable in homogeneous continuous chains can be optimally calibrated). *Suppose that the matrix R is embeddable and let t_R and the transition probabilities $P(s, t, \lambda_1)$ satisfy Definition 1 in the case of a homogeneous continuous time Markov*

chain for some family of intensities $\mathbf{Q}(\lambda_1)$ where $\lambda_1 \in \Lambda$ is a given parameter. Then, with $\tau_n := nt_{\mathbf{R}}$ for $n \geq 1$ and $\mathbf{R}^{\tau_n} := \mathbf{R}^{(n)}$ —the n fold product of the matrix \mathbf{R} by itself—we have that the optimization problem, $\inf_{\lambda \in \Lambda} \mathcal{O}(\lambda)$ with respect to the loss function given by Formula (8) has an optimal solution $\mathbf{P}(s, t, \lambda_1)$ such that

$$\mathcal{O}(\lambda_1) = \min_{\lambda \in \Lambda} \mathcal{O}(\lambda) = 0.$$

Proof. It is enough to observe that by Formulas (3) in Definition 1 we have, as $\tau_2 - \tau_1 = \tau_1$,

$$\begin{aligned} P(0, \tau_2, \lambda_1) &= P(0, \tau_1, \lambda_1)P(\tau_1, \tau_2, \lambda_1) = P(0, \tau_1, \lambda_1)P(0, \tau_2 - \tau_1, \lambda_1) = \\ &= P(0, \tau_1, \lambda_1)P(0, \tau_1, \lambda_1) = \mathbf{R}^{(2)} = \mathbf{R}^{\tau_2}, \end{aligned}$$

and, by induction, that $P(0, \tau_n, \lambda_1) = \mathbf{R}^{\tau_n}$ and so in Formula (8) we have that $\mathcal{O}(\lambda_1) = 0$. \square

Remark 6 (On the skeletons of a homogeneous continuous time Markov chain). Another possible way to extend results from discrete time to continuous time is the approach of skeletons of Kingman and other authors (see [53,54], for instance). As we are more interested in non-homogeneous continuous time Markov chains we do not pursue this approach in the present work.

We now address the case of non homogeneous Markov chain. In Theorem 3, we show that if every element of a sequence, with no gaps, of matrix powers of a discrete time Markov chain is embeddable then there is a regime switching process of Markov type that solves optimally the calibration problem.

Theorem 3 (Discrete power-embeddable discrete chains can be optimally calibrated). Suppose that all the powers $\mathbf{R}^{(n)} = [r_{ij}^{(n)}]_{i,j=1,\dots,r}$, for $1 \leq n \leq N$, of a discrete time Markov chain transition matrix \mathbf{R} are embeddable and let $\mathbf{P}_n(s, t, \lambda_n)$ be the transition probabilities of the embedding continuous time Markov chain for $\mathbf{R}^{(n)}$ given in their intrinsic time—defined in Remark 1—in such a way that the respective embedding times verifies $t_{\mathbf{R}^{(n)}} = -n \log \det \mathbf{R}$ (according to Formula (4)). We suppose that the intensities $\mathbf{Q}_n(t, \lambda_n)$ for each of the transition probabilities $\mathbf{P}_n(s, t, \lambda_n)$ depend on parameters $\lambda_n \in \Lambda$, possibly different but all in a common parameter set Λ . With the convention $t_{\mathbf{R}^{(0)}} = 0$, and

$$\lambda(t) := \lambda_n, \quad t_{\mathbf{R}^{(n-1)}} \leq t \leq t_{\mathbf{R}^{(n)}},$$

let $\tilde{\mathbf{P}}(s, t, \lambda(t))$ be defined by

$$\tilde{\mathbf{P}}(s, t, \lambda(t)) := \mathbf{P}_n(s, t, \lambda_n), \quad 0 = t_{\mathbf{R}^{(0)}} \leq s \leq t_{\mathbf{R}^{(n)}}, \quad t_{\mathbf{R}^{(n-1)}} \leq t \leq t_{\mathbf{R}^{(n)}}, \quad s \leq t, \quad (12)$$

and thus satisfying $\tilde{\mathbf{P}}(0, t_{\mathbf{R}^{(n)}}, \lambda(t)) = \mathbf{P}_n(0, t_{\mathbf{R}^{(n)}}, \lambda_n) = \mathbf{R}^{(n)}$. Then, we have that the optimization problem, $\inf_{\lambda \in \Lambda} \mathcal{O}(\lambda)$ with respect to the loss function given by

$$\mathcal{O}(\lambda) := \sum_{i,j=1,\dots,r} \sum_{n=1}^N \left(\tilde{\mathbf{P}}(0, t_{\mathbf{R}^{(n)}}, \lambda(t))_{ij} - r_{ij}^{(n)} \right)^2, \quad (13)$$

has an optimal solution $\tilde{\mathbf{P}}(s, t, \lambda(t))$ such that

$$\mathcal{O}(\lambda(t)) = \min_{\lambda \in \Lambda} \mathcal{O}(\lambda) = 0.$$

Proof. We observe that the definition in Formula (12) is coherent—see Figure 1—and then it is a simple verification with the definitions proposed. \square

Remark 7 (An associated regime switching process). The function $\tilde{P}(s, t, \lambda(t))$ defined in Formula (12) was obtained by superimposing different transition probabilities for different Markov chains in continuous time. A natural question is to determine if there is—based on these different transitions probabilities—a regime switching Markov chain in continuous time that bears some connection with $\tilde{P}(s, t, \lambda(t))$. From a brief analysis of Figure 1 we can guess the natural definition of a regime switching Markov chain based on the probabilities $P_n(s, t, \lambda_n)$. Let

$$P(s, t, \lambda(t)) := P_n(s, t, \lambda_n), \quad t_{R^{(n-1)}} \leq s \leq t \leq t_{R^{(n)}}. \quad (14)$$

Formula (14) has the following interpretation. For each $1 \leq n \leq N$, consider continuous time Markov chain processes $(X_t^n)_{t \in [t_{R^{(n-1)}}, t_{R^{(n)}}]}$ with transition probabilities $P_n(s, t, \lambda_n)$ defined in the domains $\mathcal{R}_n := \{(s, t) \in \mathbb{R}^2 : t_{R^{(n-1)}} \leq s \leq t \leq t_{R^{(n)}}\}$ with the convention $t_{R^{(0)}} = 0$. The regime switching process $(Y_t)_{t \in [0, t_{R^{(N)}}]}$ is such that (compare with Formula (2)):

$$Y_t = X_t^n, \quad t \in [t_{R^{(n-1)}}, t_{R^{(n)}}],$$

that is, the process $(Y_t)_{t \in [0, t_{R^{(N)}}]}$ is obtained by gluing together $(X_t^n)_{t \in [t_{R^{(n-1)}}, t_{R^{(n)}}]}$, the paths of the processes which are bona fide continuous time Markov processes in each of their—non-random—time intervals $[t_{R^{(n-1)}}, t_{R^{(n)}}]$. It is clear that $P(s, t, \lambda(t))$ can be interpreted as a transition probability only when restricted to some domain \mathcal{R}_n and that, in general, it will not be a transition probability in the whole interval $[0, t_{R^{(N)}}]$.

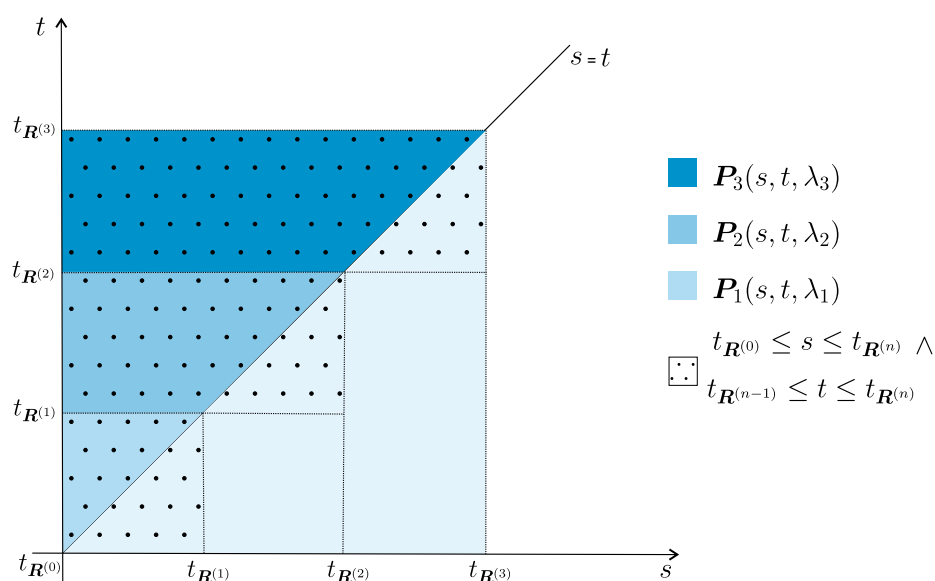


Figure 1. A representation of $\tilde{P}(s, t, \lambda(t))$ in Formula (12) for the first three initial times.

Remark 8. The regime switching process defined in Remark 7 deserves further study. We may, nevertheless, define transition probabilities $\tilde{P}(s, t, \lambda(t))$ for $t_{R^{(k-1)}} \leq s \leq t_{R^{(k)}} \leq t \leq t_{R^{(k+1)}}$ —with properties to be thoroughly investigated—by considering

$$\hat{P}(s, t, \lambda(t)) := P_k(s, t_{R^{(k)}}, \lambda_k) \cdot P_{k+1}(t_{R^{(k)}}, t, \lambda_{k+1}).$$

3.3. Conclusions on the Relations between Embeddable Matrices, Calibration, and Open Markov Chain Models

From Theorems 1–3, the following conclusions can be drawn. Given a discrete time Markov transition matrix,

- if the matrix is *embeddable*—according to Definition 1 of Section 3.1—there is an unique in law homogeneous Markov chain in continuous time that solves the calibration

problem optimally; the unicity is a consequence of Remark A2 that shows that the laws of the stopping times $(\tau_n)_{n \geq 0}$ in the representation of Formula (A13) only depend on the intensities and these are uniquely determined whenever the discrete time Markov chain is embeddable.

- if the matrix is *power-embeddable*—that is, if all the matrices of a finite sequence with no gaps of powers of the matrix are embeddable—then there is a unique regime switching continuous time non-homogeneous Markov chain—in the sense of Remark 7—that solves the calibration problem optimally. In this case, the unicity has a justification similar to the previously referred case, that is, the laws of the stopping times only depends on the intensities and these are determined by the fact that the matrix is power-embeddable.

As a consequence, for our purposes, it appears of fundamental importance to determine if a discrete time Markov chain transition matrix is embeddable and to determine—if possible, explicitly—the embedding continuous time Markov chain. Regarding this problem the results in [43,55] deserve further consideration.

Remark 9 (Applying Theorems 1–3). Suppose that discrete time Markov chain transition matrix, of a Markov chain process $(Z_n)_{n \geq 1}$ is embeddable in a continuous time Markov chain $(X_t)_{t \geq 0}$. We have, for this continuous time process and for a determined sequence of stopping times $(\tau_n)_{n \geq 1}$, the representation given in Formula (A13) of Theorem A5, that is,

$$X_t = \sum_{n=0}^{+\infty} X_{\tau_n} \mathbb{1}_{[\tau_n, \tau_{n+1}[}(t).$$

Now, as the Theorems referred to may consider that the process $(Z_n)_{n \geq 1}$ is suitably approximated by $(X_t)_{t \geq 0}$, we can also consider that the continuous time process defined by

$$\tilde{X}_t := \sum_{n=0}^{+\infty} Z_{\tau_n} \mathbb{1}_{[\tau_n, \tau_{n+1}[}(t), \quad (15)$$

is an approximation of $(Z_n)_{n \geq 1}$ in continuous time. For processes with a structural representation similar to the one of the process $(\tilde{X}_t)_{t \geq 0}$ we propose in Section 4.3 a method to extend from discrete to continuous time the open populations methodology.

4. More on Open Continuous Time Processes from Discrete Ones

In this section, we discuss an extension of the formalism of open Markov chains to the case of semi-Markov processes (sMp) and other continuous time processes, namely, the open Markov chain schemes introduced in [26]. For the reader's convenience we present in Appendix B a short summary on sMp and in the next Section 4.1 a review of the main results on the open Markov chain formalism for discrete time. Finally, we propose the second main contribution of this work, that is, an extension of the open Markov chain formalism in discrete time to continuous time in the case of sMp. We also briefly refer the case of open Markov schemes that, in some particular instances, can be dealt as the sMp case.

4.1. Open Markov Chain Modeling in Discrete Time: A Short Review

We now detail and comment the results that will be used in this paper on discrete time open Markov chains. The study of open Markov chain models we will present next relies on results and notations that were introduced in [56], further developed in [22] and that we reproduce next, for the readers convenience. We will suppose that, in general, the transition matrix of the Markov chain model may be written in the following form:

$$P = \begin{bmatrix} K & U_1 \\ 0 & V \end{bmatrix} \quad (16)$$

where \mathbf{K} is a $k \times k$ transition matrix between transient states, \mathbf{U}_1 a $k \times (r - k)$ matrix of transitions between the transient and the recurrent states, and \mathbf{V} a $(r - k) \times (r - k)$ matrix of transitions between the recurrent states. A straightforward computation then shows that

$$\mathbf{P}^{(n)} = \begin{bmatrix} \mathbf{K}^{(n)} & \mathbf{U}_n \\ \mathbf{0} & \mathbf{V}^{(n)} \end{bmatrix}, \quad n \in \mathbb{N}$$

with $\mathbf{U}_n = \mathbf{U}_{n-1}\mathbf{V} + \mathbf{K}^{(n-1)}\mathbf{U}_1 = \sum_{i=0}^{n-1} \mathbf{K}^{(i)} \mathbf{U}_1 \mathbf{V}^{(n-1-i)}$. We write the vector of the initial classification, for a time period i , as

$$\mathbf{c}_i^\top = [\mathbf{t}_i^\top | \mathbf{r}_i^\top], \quad i \in \mathbb{N} \quad (17)$$

with \mathbf{t}_i the vector of the initial allocation probabilities for the transient states and \mathbf{r}_i the vector of the initial allocation probabilities for the recurrent states. We suppose that at each epoch $i \geq 0$ there is an influx of new elements in the classes of the population—population that has its evolution governed by the Markov chain transition matrix—that is, a Poisson distributed with parameter λ_i . It is a consequence of the *randomized sampling* principle (see [57], pp. 216–217) that, if the incoming populations are distributed by the classes according with the multinomial distribution, then the sub-populations in the transient classes have independent Poisson distributions, with parameters given by the product of the Poisson parameter by the probability of the incoming new member being affected to the given class. With Formulas (16) and (17), we now notice that the vector of the Poisson parameters, for the population sizes in each state at an integer time N , may be written as

$$\lambda_N^{++\top} = \left[\sum_{i=1}^N \lambda_i \mathbf{t}_i^\top \mathbf{K}^{(N-i)} \mid \sum_{i=1}^N \lambda_i (\mathbf{t}_i^\top \mathbf{U}_{N-i} + \mathbf{r}_i^\top \mathbf{V}^{(N-i)}) \right]. \quad (18)$$

We observe that the first block corresponds to the transient states and the second block, the one in the right-hand side, corresponds to the recurrent states. From now on, as a first restricting hypothesis, we will also suppose that the transition matrix of the transient states, \mathbf{K} , is diagonalizable and so

$$\mathbf{K} = \sum_{j=1}^k \eta_j \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^\top,$$

with $(\eta_j)_{j \in \{1, \dots, k\}}$ the eigenvalues, $(\boldsymbol{\alpha}_j)_{j \in \{1, \dots, k\}}$ the left eigenvectors and $(\boldsymbol{\beta}_j)_{j \in \{1, \dots, k\}}$ the right eigenvectors of matrix \mathbf{K} . We observe that $j \in \{1, \dots, k\}$ corresponds to a transient state if and only if $|\eta_j| < 1$. We may write the powers of \mathbf{K} as

$$\mathbf{K}^{(n)} = \sum_{j=1}^k \eta_j^n \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^\top, \quad (19)$$

and so, as a consequence of (18), for the vector of the Poisson parameters corresponding only to the transient states, $\lambda_N^{+\top}$, we have

$$\lambda_N^{+\top} = \sum_{i=1}^N \lambda_i \mathbf{t}_i^\top \mathbf{K}^{(N-i)} = \sum_{j=1}^k \sum_{i=1}^N \lambda_i \eta_j^{N-i} \mathbf{t}_i^\top \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^\top. \quad (20)$$

The main result describing the asymptotic behaviour, established in [22], is the following.

Theorem 4 (Asymptotic behavior of Poisson parameters of an open Markov chain with Poisson distributed influxes). *Let a Markov chain driven system have a diagonalizable transition matrix between the transient states $\mathbf{K} = \sum_{j=1}^k \eta_j \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^\top$, written in its spectral decomposition form. Suppose the system to be fed by Poisson inputs with intensities $(\lambda_i)_{i \in \mathbb{N}}$ and such that the*

vector of initial classification of the inputs in the transient states converges to a fixed value, that is, $\lim_{i \rightarrow +\infty} \mathbf{t}_i^\top = \mathbf{t}_\infty^\top \neq \mathbf{0}$. Then, with $\lambda_n^{+\top}$ the vector of Poisson parameters of the transient sub-populations, at date $n \in \mathbb{N}$, we have the following:

1. If $\lim_{n \rightarrow +\infty} \lambda_n = \lambda \in \mathbb{R}_+$, then

$$\lambda_\infty^+ = \lim_{n \rightarrow +\infty} \lambda_n^{+\top} = \sum_{j=1}^k \frac{\lambda}{1 - \eta_j} \mathbf{t}_\infty^\top \alpha_j \beta_j^\top. \quad (21)$$

2. If $\lim_{n \rightarrow +\infty} \lambda_n = +\infty$ and there exists a constant $C > 0$ such that

$$\max_{1 \leq i \leq n} \left| \frac{\lambda_i - \lambda_{i+1}}{\lambda_n} \right| \leq C$$

then

$$\lim_{n \rightarrow +\infty} \frac{\lambda_n^{+\top}}{\lambda_n} = \sum_{j=1}^k \frac{1}{1 - \eta_j} \mathbf{t}_\infty^\top \alpha_j \beta_j^\top. \quad (22)$$

Remark 10. We observe that proportions in the Markov chain transient classes, on both statements of the Theorem 4, only depend on the eigenvalues $\eta_j, j = 1, \dots, k$. In fact, whenever using Formula (21) to compute proportions these proportions do not depend on the value of λ as we have that

$$\sum_{j=1}^k \frac{\lambda}{1 - \eta_j} \mathbf{t}_\infty^\top \alpha_j \beta_j^\top = \lambda \left[\mathbf{t}_\infty^\top \cdot \left(\sum_{j=1}^k \frac{1}{1 - \eta_j} \alpha_j \beta_j^\top \right) \right],$$

and the term in the right-hand side multiplying λ is a vector with the dimension equal to the number of transient classes k , which is equal to the dimension of the square matrix \mathbf{K} . As so, when computing proportions, by normalizing this vector with the sum of its components, $\lambda \neq 0$ disappears.

4.2. Open sMP from Discrete time Open Markov Chains

Let us suppose that the successive Poisson distributions of the influx of new members in the population are independent of the random time at which the influx of new members in the population occurs. For the notations used, see Appendix B. Consider a sMp given by the representation in Formula (A17), that is,

$$Y_t = \sum_{n=0}^{+\infty} Z_n \mathbb{1}_{[\tau_n, \tau_{n+1}[}(t),$$

in which $(Z_n)_{n \geq 0}$ is the embedded Markov chain and $(\tau_n)_{n \geq 0}$ are the jump times of the process. We now propose a method to extend the known method to study open Markov chains in discrete time to sMps.

- (1) In applications we usually consider that we have the influx of new members in the population being modeled by Poisson random variables that at each time t has a parameter $\lambda(t)$. Being so, Formula (20) may be rewritten as

$$\lambda_N^{+\top} = \sum_{i=1}^{i:t_i \leq N} \lambda(t_i) \mathbf{t}_i^\top \mathbf{K}^{(N-i)} = \sum_{j=1}^k \sum_{i=1}^{i:t_i \leq N} \lambda(t_i) \eta_j^{N-i} \mathbf{t}_i^\top \alpha_j \beta_j^\top, \quad (23)$$

where usually we can take $t_i = i$, as in a discrete time Markov chain, the actual time stamp is irrelevant as we only consider the sequence of epochs $i \geq 0$.

- (2) In a sMp the only difference we have with respect to a discrete time Markov chain is that the dates τ_i corresponding to each epoch i are random; altogether, the structure of the changes in the sub-populations in the transient states is governed by the transition matrix of the Markov chain. In a sMp, the only possible observable changes are those that occur at the random times where it jumps; as so, we will suppose that the

influxes of the new members of the population only occur at these random times. As a consequence, we should have that the vector parameter of the Poisson parameters, in the transient classes, is random since it depends on the random times in each we consider influxes and so, Formula (23) becomes

$$\lambda_N^{+T}(\omega) = \sum_{i=1}^{i:\tau_i(\omega) \leq N} \lambda(\tau_i(\omega)) \mathbf{t}_i^T \mathbf{K}^{(N-i)} = \sum_{j=1}^k \sum_{i=1}^{i:\tau_i(\omega) \leq N} \lambda(\tau_i(\omega)) \eta_j^{N-i} \mathbf{t}_i^T \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^T. \quad (24)$$

- (3) The parameters of interest will be the expected values of the random variables $\lambda_N^{+T}(\omega)$ —with the correspondent asymptotic behavior of these expected values when N grows indefinitely—and these expected values can be computed whenever the joint laws of $(\tau_0, \tau_1, \dots, \tau_i)$ are known, for $i \geq 0$. In fact, we observe that by Formula (24) we have

$$\begin{aligned} \mathbb{E}[\lambda_N^{+T} | \tau_1, \dots, \tau_i \dots] &= \mathbb{E}\left[\sum_{j=1}^k \sum_{i=1}^{i:\tau_i \leq N} \lambda(\tau_i) \eta_j^{N-i} \mathbf{t}_i^T \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^T | \tau_1, \dots, \tau_i \dots\right] = \\ &= \sum_{j=1}^k \sum_{i=1}^{i:\tau_i \leq N} \lambda(\tau_i) \eta_j^{N-i} \mathbf{t}_i^T \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^T. \end{aligned}$$

This formula has two consequences. The first one is that given an arbitrary strictly increasing sequence of dates $0 = t_0 < t_1 < \dots < t_i < \dots$ we have

$$\mathbb{E}[\lambda_N^{+T} | \tau_1 = t_1, \dots, \tau_i = t_i \dots] = \sum_{j=1}^k \sum_{i=1}^{i:t_i \leq N} \lambda(t_i) \eta_j^{N-i} \mathbf{t}_i^T \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^T,$$

thus justifying the assumption that given the strictly increasing of non accumulating stopping times dates $(\tau_1 = t_1, \dots, \tau_i = t_i \dots)$ we can proceed as with the usual open Markov chain model in discrete time. The second consequence deserving mention is that in order to compute the expected value of the vector parameters of the transient classes sub-populations, while preserving the Poisson distribution of the influx new members, we compute

$$\mathbb{E}[\lambda_N^{+T}] = \mathbb{E}\left[\mathbb{E}[\lambda_N^{+T} | \tau_1, \dots, \tau_i \dots]\right] = \mathbb{E}\left[\sum_{j=1}^k \sum_{i=1}^{i:\tau_i \leq N} \lambda(\tau_i) \eta_j^{N-i} \mathbf{t}_i^T \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^T\right],$$

using the joint laws of (τ_1, \dots, τ_i) for $i \geq 0$, laws we will suppose to be given.

Theorem 6, in the following, is one possible extension of the open Markov chain formalism to the sMp case taking as a starting point a discrete time Markov chain. To prove this result we will need Theorem 5—a generalization of Lebesgue dominated convergence theorem with varying measures—that we quote from Theorem 3.5 in [58] (p. 390).

Theorem 5 (Lebesgue dominated convergence theorem with varying measures). *Consider $(X, \mathcal{B}(X))$ a locally compact, separable topological space endowed with its Borel σ -algebra. Suppose that the sequence of probability measures $(\mu_n)_{n \geq 1}$ —each one of them defined in $(X, \mathcal{B}(X))$ —converges weakly to μ on $(X, \mathcal{B}(X))$ and that the sequence of measurable functions $(f_n)_{n \geq 1}$ converges continuously to f . Suppose additionally that, for some sequence of measurable functions $(g_n)_{n \geq 1}$ defined on X :*

1. *For all $\mathbf{t} \in X$ and $n \geq 1$, we have that $|f_n(\mathbf{t})| \leq g_n(\mathbf{t})$.*
2. *With the function g defined on X by*

$$g(\mathbf{t}) := \inf_{(t_n)_{n \geq 1}, \lim_{n \rightarrow +\infty} t_n = \mathbf{t}} \left\{ \liminf_{n \rightarrow +\infty} g_n(t_n) \right\}$$

we have that

$$\limsup_{n \rightarrow +\infty} \int g_n(t) d\mu_n(t) \leq \int g(t) d\mu(t) < +\infty.$$

Then, we have

$$\lim_{n \rightarrow +\infty} \int f_n(t) d\mu_n(t) = \int f(t) d\mu(t) < +\infty.$$

As said, we will suppose that we only observe the influx of the new members of the population into the sMp classes at the random times where it jumps—but, of course, accounting the state before the jump and the state after the jump—which is a hypothesis that makes sense under the perspective that we usually observe trajectories of the process. We then have the following extension of Theorem 4 to the case of sMp.

Theorem 6 (On the stability of open sMp transient states). *Let a sMp given by the representation in Formula (A17), that is,*

$$Y_t = \sum_{n=0}^{+\infty} Z_n \mathbb{1}_{[\tau_n, \tau_{n+1}[}(t),$$

in which $(Z_n)_{n \geq 0}$ is the embedded Markov chain and $(\tau_i)_{i \geq 0}$ are the jump times of the process. For the embedded Markov chain $(Z_n)_{n \geq 0}$, consider the notations of Section 4.2 and of Theorem 4 in this subsection. Suppose that the influx of new members in the population is modeled by Poisson random variables that at each time $t \in [0, +\infty[$ have a parameter $\lambda(t)$, with λ a **continuous** function. Suppose, furthermore, that the following hypothesis are verified.

1. The stopping times $(\tau_i)_{i \geq 0}$ are integrable, that is, $\mathbb{E}[\tau_i] < +\infty$ for all $i \geq 1$.
2. There exists $\lambda_\infty > 0$ such that, for every sequence of positive real numbers $(t_i)_{i \geq 1}$ such that $\lim_{i \rightarrow +\infty} t_i = +\infty$ we have

$$\lim_{i \rightarrow +\infty} \lambda(t_i) = \lambda_\infty \quad (25)$$

Then, we have that the asymptotic behavior of the expected value vector of parameters of Poisson distributed sub-populations in the transient classes of an open sMp, submitted to a Poisson influx of new members at the jump times of the sMp, is given by

$$\lim_{N \rightarrow +\infty} \mathbb{E}[\lambda_N^{+\top}] = \lim_{N \rightarrow +\infty} \mathbb{E} \left[\sum_{j=1}^k \sum_{i: \tau_i \leq N} \lambda(\tau_i) \eta_j^{N-i} \mathbf{t}_i^\top \alpha_j \beta_j^\top \right] = \sum_{j=1}^k \frac{\lambda_\infty}{1 - \eta_j} \mathbf{t}_\infty^\top \alpha_j \beta_j^\top. \quad (26)$$

Proof. For each $n \geq 1$, let $F_{(\tau_1, \dots, \tau_n)}$ be the joint distribution function of (τ_1, \dots, τ_n) . We want to compute the following limit of expectations:

$$\begin{aligned} \lim_{N \rightarrow +\infty} \mathbb{E}[\lambda_N^{+\top}] &= \lim_{N \rightarrow +\infty} \mathbb{E}[\lambda_N^{+\top}, \tau_1 < \dots < \tau_i \leq N] = \\ &= \lim_{N \rightarrow +\infty} \int_{0 < t_1 < \dots < t_i \leq N} \lambda_N^{+\top} dF_{(\tau_1, \dots, \tau_n)}(t_1, \dots, t_n) = \\ &= \lim_{N \rightarrow +\infty} \int_{0 < t_1 < \dots < t_i \leq N} \left(\sum_{j=1}^k \sum_{i: t_i \leq N} \lambda(t_i) \eta_j^{N-i} \mathbf{t}_i^\top \alpha_j \beta_j^\top \right) dF_{(\tau_1, \dots, \tau_n)}(t_1, \dots, t_n), \end{aligned} \quad (27)$$

and we observe that by Theorem 4 and by the first hypothesis, for every sequence of positive real numbers $(t_i)_{i \geq 1}$ such that $\lim_{i \rightarrow +\infty} t_i = +\infty$ and $t_1 < t_2 < \dots < t_i < \dots$, we have that

$$\lim_{N \rightarrow +\infty} \left(\sum_{j=1}^k \sum_{i: t_i \leq N} \lambda(t_i) \eta_j^{N-i} \mathbf{t}_i^\top \alpha_j \beta_j^\top \right) = \sum_{j=1}^k \frac{\lambda_\infty}{1 - \eta_j} \mathbf{t}_\infty^\top \alpha_j \beta_j^\top. \quad (28)$$

The limit in the last term of Formula (27) requires a result of Lebesgue convergence theorem type but with varying measures. For the purpose of applying Theorem 5, we introduce the adequate context and notations and then we will apply the referred theorem.

Consider the space $X = [0, +\infty[^{\mathbb{N}_0}$ defined to be the space of infinite sequences of numbers in $[0, +\infty[$, that is,

$$X = \{t = (t_1, \dots, t_i, \dots) : \forall i \geq 1, t_i \in [0, +\infty[\}.$$

Recall that with the metric d given by

$$\forall t = (t_1, \dots, t_i, \dots), t' = (t'_1, \dots, t'_i, \dots) \in X, d(t, t') := \sum_{i=1}^{+\infty} \frac{\min(1, |t_i - t'_i|)}{2^i},$$

X is a metric space, locally compact, separable and complete (see, for instance, in [59], pp. 9–10). We will consider $X = [0, +\infty[^{\mathbb{N}_0}$ endowed with the Borel σ -algebra $\mathcal{B}(X)$ generated by the family \mathcal{P}_f given by

$$\mathcal{P}_f = \left\{ A_{i_1} \times A_{i_2} \times \dots \times A_{i_p} : p \geq 1, A_{i_1} \in \mathcal{B}([0, +\infty[) \right\},$$

with $\mathcal{B}([0, +\infty[)$ the Borel σ -algebra of $[0, +\infty[$. We now take $(\tau_i)_{i \geq 0}$ the sequence of the jump times of the process represented in Formula (A17). First, we define the sequence of measures $(\mu_n)_{n \geq 1}$ where for each $n \geq 1$ we have that μ_n is defined on the measurable space $([0, +\infty[^n, \mathcal{B}([0, +\infty[^n))$ by considering, for $A_1 \times A_2 \times \dots \times A_n$ with $A_i \in \mathcal{B}([0, +\infty[)$, that

$$\mu_n(A_1 \times A_2 \times \dots \times A_n) = \mathbb{P}[\tau_1 \in A_1, \dots, \tau_n \in A_n] = \int_{t_1 \in A_1, \dots, t_n \in A_n} dF_{(\tau_1, \dots, \tau_n)}(t_1, \dots, t_n). \quad (29)$$

Being so, μ_n is the probability joint law of (τ_1, \dots, τ_n) and the last integral in the last term of Formula (27) is exactly an integration with respect to the measure μ_n . As a consequence of Formula (29), the sequence $(\mu_n)_{n \geq 1}$ verifies the compatibility conditions of Kolmogorov extension theorem (see [60], p. 46) and so there is a probability measure μ , defined on $(X, \mathcal{B}(X))$, having as finite dimensional distributions the measures of the sequence $(\mu_n)_{n \geq 1}$.

Now, for each $n \geq 1$, we can consider $\tilde{\mu}_n$ the extension of μ_n to the measurable space $(X, \mathcal{B}(X))$ in the following way:

$$\forall A \in \mathcal{B}(X) \quad \tilde{\mu}_n(A) = \int_{\{t=(t_1, \dots, t_i, \dots) \in A : t_1, \dots, t_n \in [0, +\infty[\}} dF_{(\tau_1, \dots, \tau_n)}(t_1, \dots, t_n). \quad (30)$$

In fact, with this definition the restriction of $\tilde{\mu}_n$ to $\mathcal{B}([0, +\infty[^n)$ is exactly μ_n . An important observation is the following. Consider $A := A_{i_1} \times A_{i_2} \times \dots \times A_{i_p} \in \mathcal{P}_f$. Then, for $m \geq i_p$ we have that

$$\begin{aligned} \tilde{\mu}_m(A) &= \int_{\{t=(t_1, \dots, t_i, \dots) \in A : t_1, \dots, t_m \in [0, +\infty[\}} dF_{(\tau_1, \dots, \tau_m)}(t_1, \dots, t_m) = \\ &= \int_{\{t=(t_1, \dots, t_i, \dots) \in A : t_1, \dots, t_{i_p} \in [0, +\infty[\}} dF_{(\tau_1, \dots, \tau_{i_p})}(t_1, \dots, t_{i_p}) = \\ &= \tilde{\mu}_{i_p}(A) = \mu_{i_p}(A) = \mu(A), \end{aligned} \quad (31)$$

thus showing that for every $A \in \mathcal{P}_f$ the sequence $(\tilde{\mu}_m(A))_{m \geq 1}$ converges to $\mu(A)$. Now, by Theorem 2.2 in [59] (p. 17), as \mathcal{P}_f is a π -system and every open set in the metric space (X, d) is a countable union of elements of \mathcal{P}_f , we have that the sequence $(\tilde{\mu}_m)_{m \geq 1}$ converges weakly to μ . In order to apply Theorem 5 to compute the limit, we may consider two approaches to deal with the fact that λ_N^{+T} is a vector of finite dimension k . Either we

proceed component wise or we consider norms. Let us follow the second path. Define, for integer N , and some constant M ,

$$f_N(\mathbf{t}) = f_N(t_1, \dots, t_i, \dots) := \sum_{i=1}^{t_i \leq N} \lambda(t_i) \eta_j^{N-i} \mathbf{t}_i^\top \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^\top,$$

and also,

$$g_N(\mathbf{t}) \equiv g := \left\| \sum_{j=1}^k \frac{\lambda_\infty}{1 - \eta_j} \mathbf{t}_\infty^\top \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^\top \right\| + M,$$

in such a way that $\|f_N(\mathbf{t})\| \leq g$; such choice of M is possible as a consequence of Formula (28). We can verify that the sequence $(f_N)_{N \geq 1}$ converges continuously to a function f by using Theorem 4.1.1 in [22] (p. 373). In fact, let us consider a sequence $(\mathbf{t}_N)_{n \geq 1}$ converging to some $(\mathbf{t}^\infty = (t_1^\infty, \dots, t_i^\infty, \dots))$ in the metric space (X, d) . With $(\mathbf{t}_N = (t_1^N, \dots, t_i^N, \dots))$ we surely have that $\lim_{N \rightarrow +\infty} t_i^N = t_i^\infty$ for all $i \geq 1$. As a consequence of the continuity of λ and of Theorem 4.1.1 in [22] (p. 373), we have that

$$\lim_{N \rightarrow +\infty} f_N(\mathbf{t}_N) = \lim_{N \rightarrow +\infty} \sum_{i=1}^{t_i^N \leq N} \lambda(t_i^N) \eta_j^{N-i} \mathbf{t}_i^\top \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^\top = \sum_{j=1}^k \frac{\lambda(\lim_{i \rightarrow +\infty} t_i^\infty)}{1 - \eta_j} \mathbf{t}_\infty^\top \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^\top =: f(\mathbf{t}^\infty).$$

It is clear now that the sequences $(f_N)_{N \geq 1}$, $(g_N)_{N \geq 1}$ and $(\tilde{\mu}_n)_{n \geq 1}$ satisfy together with μ the hypothesis of Theorem 5 and so the announced result in Formula (25) follows. \square

Remark 11 (Alternative proof for the weak convergence of the sequence $(\tilde{\mu}_n)_{n \geq 1}$). There is another proof the weak convergence of the sequence $(\tilde{\mu}_n)_{n \geq 1}$ to μ that we now present. We proceed by showing that the sequence $(\tilde{\mu}_n)_{n \geq 1}$ is relatively compact—as a consequence of Prohorov theorem (see [59], pp. 59–63)—because, as we will show next, this sequence is tight. Let an arbitrary $0 < \epsilon < 1$ be given and consider a sequence of positive numbers $(\xi_i)_{i \geq 1}$ such that, by Tchebychev inequality and using the fact that the stopping times τ_i have finite integrals,

$$\mathbb{P}[\tau_i > \xi_i] \leq \frac{\mathbb{E}[\tau_i]}{\xi_i},$$

in such a way that

$$\sum_{i=1}^{+\infty} \frac{\mathbb{E}[\tau_i]}{\xi_i} < \epsilon.$$

Now consider the Borel set $K_\epsilon = \prod_{i=1}^{+\infty} [0, \xi_i] \subset X$ which is compact by Tychonov theorem. We now have that

$$\begin{aligned} \tilde{\mu}_n(K_\epsilon) &= \int_{\{\mathbf{t}=(t_1, \dots, t_i, \dots) \in K_\epsilon : t_1, \dots, t_n \in [0, +\infty[\}} dF_{(\tau_1, \dots, \tau_n)}(t_1, \dots, t_n) = \\ &= \int_{\prod_{i=1}^n [0, \xi_i]} dF_{(\tau_1, \dots, \tau_n)}(t_1, \dots, t_n) = \\ &= \mathbb{P} \left[(\tau_1, \dots, \tau_n) \in \prod_{i=1}^n [0, \xi_i] \right] = \mathbb{P} \left[\bigcap_{i=1}^n \{\tau_i \leq \xi_i\} \right] = 1 - \mathbb{P} \left[\bigcup_{i=1}^n \{\tau_i > \xi_i\} \right] \geq \\ &\geq 1 - \sum_{i=1}^n \frac{\mathbb{E}[\tau_i]}{\xi_i} \geq 1 - \sum_{i=1}^{+\infty} \frac{\mathbb{E}[\tau_i]}{\xi_i} \geq 1 - \epsilon, \end{aligned}$$

thus showing that the sequence of probability measures $(\tilde{\mu}_n)_{n \geq 1}$ is tight in the measurable space $(X, \mathcal{B}(X))$. As said, by Prokhorov's theorem, this implies that the sequence $(\tilde{\mu}_n)_{n \geq 1}$ is relatively compact, that is, for every subsequence of $(\tilde{\mu}_n)_{n \geq 1}$, there exists a further subsequence and a probability measure such that this subsequence converges weakly to the said probability measure. Now, as, by construction, the probability measure μ has, as finite dimensional distributions the

probability measures $(\tilde{\mu}_n)_{n \geq 1}$ we can say that for $n \geq 1$, the finite dimensional distributions of $\tilde{\mu}_n$ converge weakly to the finite dimensional distributions of μ . As a consequence, following the observation in [59] (p. 58), the sequence $(\tilde{\mu}_n)_{n \geq 1}$ converges weakly to μ .

Remark 12 (Applying Theorem 6). If we manage to estimate a discrete time Markov chain transition matrix and if we manage to fit some function f —such that $\lim_{t \rightarrow +\infty} f(t) = \lambda_\infty$ —to the number of new incoming members in the population at a set of non accumulating non-evenly spaced dates (as done with a statistical procedure in [22] or, with a simple fitting in [25]) then, Theorem 6 allows us to get the asymptotic expected number of elements in the transient classes of a sMp having as embedded Markov chain the estimated one.

4.3. Open Continuous Time Processes from Open Markov Schemes

We may follow the approach of open Markov schemes in [26] and define a process in continuous time after getting a process in random discrete times describing, at least on average, the evolution of the elements in each transient class. Let us briefly recall the main idea. A population model is driven by a Markov chain defined by a sequence of initial distributions given, for $n \geq 1$, by $(\mathbf{q}^n)^\top = (q_1^n, q_2^n, \dots, q_r^n)$ and a transition matrix $\mathbf{P} = [p_{ij}]$, $1 \leq i, j \leq r$. After the first transition, the new values of the proportions in all states, after one transition, can be recovered from $\mathbf{P}^\top \mathbf{q} = (\mathbf{q}^\top \mathbf{P})^\top$ and, after n transitions, by $(\mathbf{P}^{(n)})^\top \mathbf{q} = (\mathbf{q}^\top \mathbf{P}^{(n)})^\top$. We want to account for the evolution of the **expected** number of elements in each class supposing that, at each **random** date τ_k , a random number X_{τ_k} of new elements enters the population. Just after the second cohort enters the population, a first transition occurs in the first cohort driven by the Markov chain law and so on and so forth. Table 1 summarizes this accounting process in which, at each step k , we distribute multinomially the new random arrivals X_{τ_k} according to the probability vector \mathbf{q}^k and the elements in each class are redistributed according to the Markov chain transition matrix \mathbf{P} .

Table 1. Accounting of n Markov cohorts each with an initial distribution.

Date	τ_1	τ_2	...	τ_{n-1}	τ_n
τ_1	$\mathbb{E}[X_{\tau_1}](\mathbf{q}^1)^\top$	$\mathbb{E}[X_{\tau_1}](\mathbf{q}^1)^\top \mathbf{P}$...	$\mathbb{E}[X_{\tau_1}](\mathbf{q}^1)^\top \mathbf{P}^{(n-2)}$	$\mathbb{E}[X_{\tau_1}](\mathbf{q}^1)^\top \mathbf{P}^{(n-1)}$
τ_2	—	$\mathbb{E}[X_{\tau_2}](\mathbf{q}^2)^\top$...	$\mathbb{E}[X_{\tau_2}](\mathbf{q}^2)^\top \mathbf{P}^{(n-3)}$	$\mathbb{E}[X_{\tau_2}](\mathbf{q}^2)^\top \mathbf{P}^{(n-2)}$
...
τ_n	—	—	—	—	$\mathbb{E}[X_{\tau_n}](\mathbf{q}^n)^\top$

At date τ_k , if we suppose that each new set of individuals in the population, a cohort, evolves independently from any one of the already existing sets of individuals but, accordingly, to the same Markov chain model, we may recover the total **expected** number of elements in each class at date τ_k by computing the sum:

$$\bar{\mathbf{K}}_n = \sum_{k=1}^n \mathbb{E}[X_{\tau_k}](\mathbf{q}^k)^\top \mathbf{P}^{(n-k)}. \quad (32)$$

Each vector component corresponds precisely to the **expected** number of elements in each class. In order to further study the properties of $(\bar{\mathbf{K}}_n)_{n \geq 1}$, given the properties of a stochastic process $\mathbb{X} = (X_{\tau_k})_{k \geq 1}$, we will randomize formula (32) by considering, instead, for $n \geq 1$:

$$\mathbf{K}_n = \sum_{k=1}^n X_{\tau_k}(\mathbf{q}^k)^\top \mathbf{P}^{(n-k)}, \quad (33)$$

and we observe that in any case $\mathbb{E}[\mathbf{K}] = \bar{\mathbf{K}}_n$. It is known that if the vector of classification probabilities is constant $\mathbf{c}^k = \mathbf{c}$ and if the \mathbb{X} is an ARMA, ARIMA, or SARIMA process, then the populations in each of the transient classes can be described by a sum of a deterministic trend, plus an ARMA process plus an evanescent process, that is a centered process $(Y_k)_{k \geq 1}$ such that $\lim_{k \rightarrow +\infty} \mathbb{E}[|Y_k|^2] = 0$ (see Theorems 3.1 and 3.2 in [26]).

The step process in continuous time naturally associated with the discrete time one would be then defined by for $t \geq 0$ by

$$\mathbf{K}_t := \sum_{n=0}^{+\infty} \mathbf{K}_n \mathbb{1}_{[\tau_n, \tau_{n+1}[}(t) = \sum_{n=0}^{+\infty} \left(\sum_{k=1}^n X_{\tau_k}(\mathbf{q}^k)^\top \mathbf{P}^{(n-k)} \right) \mathbb{1}_{[\tau_n, \tau_{n+1}[}(t).$$

In order to study this process we will have to take advantage of the properties of \mathbb{X} and of the family of stopping times $(\tau_k)_{k \geq 0}$. It should be noticed that if the process $\mathbb{X} = (X_t)_{t \geq 0}$ is Poisson distributed and the laws of the sequence $(\tau_k)_{k \geq 0}$ are known and it possible to determine the expected value of \mathbf{K}_t for $t \geq 0$ with a result similar to Theorem 6.

5. Conclusions

In this work, we studied several ways to associate, to an open Markov chain process in discrete time—which is often the sole accessible fruit of observation—a continuous time Markov or semi-Markov process that bears some natural relation with the discrete time process. Furthermore, we expect that association to allow the extension of the study of open populations from the discrete to the continuous time model. For that purpose, we consider three approaches: the first, for the continuous time Markov chains; the second, for the semi Markov case; and the third, for the open Markov schemes (see in [26]). For the semi-Markov case, under the hypothesis that we only observe the influx of new individuals in the population at the times of the random jumps, in the main result we determine the expected value of the vector of parameters of the conditional Poisson distributions in the transient classes when the influx of new members is Poisson distributed. The third approach, dealing with open Markov schemes is similar to the second one whenever we consider a similar context hypothesis, that is, distributed incoming new members of the population with known distributions and observation of this influx of new individuals at the times of the random jumps. In the case of the first approach, that is, for the case of Markov chain in continuous time, we propose a calibration procedure for which the embeddable Markov chains provide optimal solutions. In this case also, the study of open populations models relies on the main result proved for the semi-Markov case approach. Future work encompasses applications to real data and the determination of criteria to assess the quality of the association of the continuous model to the observed discrete time model.

Author Contributions: All authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: For the second author, this work was done under partial financial support of RFBR (Grant n. 19-01-00451). For the first and third author this work was partially supported through the project of the Centro de Matemática e Aplicações, UID/MAT/00297/2020 financed by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology). The APC was funded by the insurance company Fidelidade.

Acknowledgments: This work was published with financial support from the insurance company Fidelidade. The authors would like to thank Fidelidade for this generous support and also, for their interest in the development of models for insurance problems in Portugal. The authors express gratitude to Professor Panagiotis C.G. Vassiliou for his enlightening comments on a previous version of this work and to the comments, corrections and questions of the referees, in particular, to the one question that motivated the inclusion of Remark 5.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Some Essential Results on Continuous Time Markov Chains

In this exposition of the most relevant results pertinent to our purposes, we follow mainly the references [29–31]. As this exposition is a mere reminder of needed notions and results, the proofs are omitted unless the result is essential for our purposes.

Definition A1 (Continuous time Markov chain). Let \mathcal{I} be some *finite* set; for instance, $\Theta = \{\theta_1, \theta_2, \dots, \theta_r\}$ of Section 2. A stochastic process $(X_t)_{t \geq 0}$ is a **continuous time Markov chain** with state space \mathcal{I} if and only if the following **Markov property** is verified, namely, for all $i_0, i_1, \dots, i_n \in \mathcal{I}$ and $0 = t_0 < t_1 < \dots < t_n < \dots$ we have that

$$\begin{aligned} \mathbb{P}[X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}, \dots, X_{t_1} = i_1, X_{t_0} = i_0] = \\ = \mathbb{P}[X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}]. \end{aligned}$$

We observe that by force of the Markov property in Definition A1 the law of a continuous time Markov chain depends only on the following transition probabilities. Let I be the identity matrix with dimension $\#\mathcal{I}$ the Kronecker's delta be given by

$$\delta_i^j = \begin{cases} 0 & i \neq j \\ 1 & i = j. \end{cases}$$

Definition A2 (Transition probabilities). Let \mathcal{I} be the state space of $(X_t)_{t \geq 0}$ a continuous time Markov chain. The **transition probabilities** are defined by

$$\forall i, j \in \mathcal{I}, s < t, p(s, i, t, j) = \mathbb{P}[X_t = j | X_s = i] \text{ and } p(t, i, t, j) = \delta_i^j.$$

Let $\mathcal{L}(\mathbb{R}^{\#\mathcal{I}})$ be the space of square matrices with coefficients in \mathbb{R} . The **transition probability matrix function** $P: \mathbb{R}_+ \times \mathbb{R}_+ \mapsto \mathcal{L}(\mathbb{R}^{\#\mathcal{I}})$ is defined by

$$\forall i, j \in \mathcal{I}, s < t, P(s, t) = [p(s, i, t, j)]_{i, j \in \mathcal{I}} \text{ and } P(t, t) = I. \quad (\text{A1})$$

Transition probabilities of Markov processes in general satisfy a very important functional equation that results from the Markov property.

Theorem A1 (Chapman-Kolmogorov equations). Consider a NH-CT-MC as given in Definition A1. Let P its transition probability matrix function as given in Definition A2. We then have

$$\forall s, u, t, 0 \leq s < u < t, P(s, t) = P(s, u)P(u, t) \quad (\text{A2})$$

As an application of the celebrated existence theorem of Kolmogorov (in the form exposed in [61], pp. 8–10) we have that, under a set of natural hypothesis, there exists a NH-CT-MC such as the one in Definition A1.

Theorem A2 (On the existence of NH-CT-MC). Let p_0 be an initial probability over \mathcal{I} . Consider a matrix valued function $P: \mathbb{R}_+ \times \mathbb{R}_+ \mapsto \mathcal{L}(\mathbb{R}^{\#\mathcal{I}})$ denoted by $P(s, t) = [p(s, i, t, j)]_{i, j \in \mathcal{I}}$ and satisfying Formulas (A3) and (A4) below, that is,

1. For all $s < t$ and for all $i \in \mathcal{I}$

$$\sum_{j \in \mathcal{I}} p(s, i, t, j) = 1. \quad (\text{A3})$$

2. Formula (A2) in Theorem A1, namely,

$$\forall s, u, t, s < u < t, P(s, t) = P(s, u)P(u, t). \quad (\text{A4})$$

Define, for all $i_0, i_1, \dots, i_n \in \mathcal{I}$ and $0 = t_0 < t_1 < \dots < t_n < \dots$, the function

$$\begin{aligned} v_{i_0, t_1, \dots, t_n}(i_0, i_1, \dots, i_n) = \\ = p_0(i_0)p(t_0, i_0, t_1, i_1)p(t_1, i_1, t_2, i_2) \cdots p(t_{n-1}, i_{n-1}, t_n, i_n), \end{aligned} \quad (\text{A5})$$

and extend this definition to all possible $t_0, t_1, \dots, t_n, \dots$ by considering, with the adequate ordering permutation σ of $\{0, 1, 2, \dots, \#\mathcal{I}\}$ such that we have $t_{\sigma(0)} < t_{\sigma(1)} < \dots, < t_{\sigma(n)}$,

$$\nu_{t_{\sigma(0)}, t_{\sigma(1)}, \dots, t_{\sigma(n)}}(i_0, i_1, \dots, i_n) = \nu_{t_0, t_1, \dots, t_n}(i_{\sigma^{-1}(0)}, i_{\sigma^{-1}(1)}, \dots, i_{\sigma^{-1}(n)}). \quad (\text{A6})$$

Then, $(\nu_{t_0, t_1, \dots, t_n})_{t_0, t_1, \dots, t_n, n \geq 1}$ is a family of probability measures satisfying the compatibility conditions of Kolmogorov existence theorem and so, there exists \mathbb{P} a probability measure over the canonical probability space (Ω, \mathcal{A}) —with $\Omega = \mathcal{I}^{\mathbb{R}_+}$ and $\mathcal{A} = \mathcal{P}(\mathcal{I})^{\mathbb{R}_+}$ —such that if the stochastic process $(X_t)_{t \geq 0}$ is denoted by

$$\forall \omega = (i_t)_{t \geq 0} \in \Omega, X_t(\omega) = i_t,$$

then,

$$\forall i, j \in \mathcal{I}, s < t, p(s, i, t, j) = \mathbb{P}[X_t = j | X_s = i] \text{ and } p(t, i, t, j) = \delta_i^j, \quad (\text{A7})$$

that is, $(X_t)_{t \geq 0}$ has $P(s, t) = [p(s, i, t, j)]_{i, j \in \mathcal{I}}$ —together with $P(t, t) = I$ —as its transition probabilities.

A natural and useful way of defining transition probabilities is by means of the transition intensities that act like differential coefficients of transition probability functions.

Definition A3 (Transition intensities). Let $\mathcal{L}(\mathbb{R}^{\#\mathcal{I}})$ be the space of square matrices with coefficients in \mathbb{R} . A function $Q : \mathbb{R} \mapsto \mathcal{L}(\mathbb{R}^{\#\mathcal{I}})$ denoted by

$$Q(t) = [q(t, i, j)]_{i, j \in \mathcal{I}},$$

is a **transition intensity** iff for almost all $t \geq 0$ it verifies

- (i) $\forall i \in \mathcal{I}, t \geq 0, q(t, i, i) \leq 0;$
- (ii) $\forall i \in \mathcal{I}, t \geq 0, q(t, i, j) - q(t, i, i) \geq 0;$
- (iii) $\forall i \in \mathcal{I} \sum_{j \in \mathcal{I}} q(t, i, j) = 0.$

There is a way to write differential equations—the Kolmogorov backward and forward equations—useful for recovering the transition probability matrix from the intensities matrix and to study important properties of these transition probabilities.

Theorem A3 (Backward and Forward Kolmogorov equations). Suppose that $P(s, t)$ is continuous at s , that is,

$$\lim_{t \downarrow 0} P(0, t) = I \text{ and } \lim_{t \downarrow s} P(s, t) = \lim_{t \uparrow s} P(t, s) = I. \quad (\text{A8})$$

If there exists Q such that

$$\begin{aligned} Q(t) &= \lim_{k+h \rightarrow 0, k \equiv 0 \vee h \equiv 0} \frac{P(t-k, t+h) - I}{k+h} = \lim_{h \downarrow 0, h > 0} \frac{P(t, t+h) - I}{h} = \\ &= \lim_{k \downarrow 0, k > 0} \frac{P(t-k, t) - I}{k}, \end{aligned} \quad (\text{A9})$$

then we have the backward Kolmogorov (matrix) equation:

$$\frac{\partial}{\partial s} P(s, t) = -Q(s)P(s, t), P(s, s) = I, \quad (\text{A10})$$

and the forward Kolmogorov (matrix) equation:

$$\frac{\partial}{\partial t} P(s, t) = P(s, t)Q(s), P(t, t) = I. \quad (\text{A11})$$

Remark A1. The general theory of Markov processes shows that the condition that $P(s, t)$ is continuous in both s and t is sufficient to ensure the existence of the matrix intensities Q given in Formulas (A9) (see [31], p. 232). By means of a change of time Goodman (see [41]) proved that the existence of solutions of Kolmogorov equations is amenable to an application of Caratheodory's existence theorem for differential equations.

Given transition intensities satisfying an integrability condition there are transition probabilities uniquely associated with these transition intensities.

Theorem A4 (Transition probabilities from intensities). *Let Q be a transition intensity as in Definition A3 such that Theorem A3 holds. Then, we have that*

$$P(s, t) = I + \int_s^t Q(u)P(u, t)du \text{ and } P(s, t) = I + \int_s^t P(s, u)Q(u)du. \quad (\text{A12})$$

The existence of a NH-CT-MC can also be guaranteed by a constructive procedure that we now present and that is most useful for simulation.

Remark A2 (Constructive definition). *Given a transition intensity Q define*

$$p^*(t, i, j) = \begin{cases} \frac{1-\delta_i^j}{-q(t, i, i)}q(t, i, j) & q(t, i, i) \neq 0 \\ \delta_i^j & q(t, i, i) = 0. \end{cases}$$

1. Let $X_0 = i$, according to some initial distribution on \mathcal{I} ; the sequence $(\tau_n)_{n \geq 0}$ is defined by induction as follows; $\tau_0 \equiv 0$.
2. τ_1 time of first jump with Exponential distribution function:

$$F_{\tau_1}(t) = \mathbb{P}[\tau_1 \leq t] = 1 - \exp\left(-\int_0^t q(u, i, i)du\right),$$

and

$$\mathbb{P}[X_{s_1} = j | \tau_1 = s_1, X_0 = i] = p^*(s_1, i, j),$$

and so $X_t = i$ for $0 \equiv \tau_0 \leq t < \tau_1$. We note that this distribution of the stopping time is mandatory as a consequence of a general result on the distribution of sojourn times of a continuous time Markov chain (see Theorem 2.3.15 in [31], p. 221).

3. Given that $\tau_1 = s_1$ and $X_{s_1} = j$, τ_2 time of the second jump with Exponential distribution function

$$F_{\tau_2|\tau_1=s_1}(t) = \mathbb{P}[\tau_2 \leq t | \tau_1 = s_1] = 1 - \exp\left(-\int_0^t q(u + s_1, j, j)du\right)$$

and

$$\mathbb{P}[X_{s_2} = k | \tau_1 = s_1, X_0 = i, \tau_2 = s_2, X_{s_1} = j] = p^*(s_1 + s_2, j, k),$$

and so $X_t = j$ for $\tau_1 \leq t < \tau_2$.

The following result ensures that the preceding construction yields the desired result.

Theorem A5 (The continuous time Markov chain). *Let the intensities satisfy condition given by Formula (A12) in Theorem A4. Then, given the times $(\tau_0)_{n \geq 1}$, we have that with the sequence $(Y_n)_{n \geq 1}$ defined by $Y_n = X_{\tau_n}$, the process defined by:*

$$X_t = \sum_{n=0}^{+\infty} Y_n \mathbb{1}_{[\tau_n, \tau_{n+1}[}(t) = \sum_{n=0}^{+\infty} X_{\tau_n} \mathbb{1}_{[\tau_n, \tau_{n+1}[}(t) \quad (\text{A13})$$

is a continuous time Markov chain with transition probabilities \mathbf{P} given by Definition A2 and transition intensities \mathbf{Q} given by Definition A3 and Theorem A3.

Proof. This theorem is stated and proved, in the general case of Markov continuous time Markov processes in [31] (p. 229). \square

Lemma A1. Let $q : \mathbb{R}_+ \mapsto \mathbb{R}$ a measurable function integrable over every bounded interval of \mathbb{R}_+ . Then, we have that

$$\int_s^t \int_{s_1}^t \cdots \int_{s_{n-1}}^t q(s_1)q(s_2) \dots q(s_n) ds_n \dots ds_2 ds_1 = \frac{\left(\int_s^t q(u) du\right)^n}{n!},$$

for all $0 \leq s \leq t$, $n \geq 1$.

Proof. Let us observe that, for $n = 2$, we have that

$$\begin{aligned} \left(\int_s^t q(u) du\right)^2 &= \int_s^t \int_s^t q(v)q(u) dudv = \\ &= \int_s^t \int_s^t \mathbb{1}_{\{u \leq v\}} q(v)q(u) dudv + \int_s^t \int_s^t \mathbb{1}_{\{v \leq u\}} q(v)q(u) dudv. \end{aligned}$$

By induction we have for all $n \geq 1$, and for every permutation $\sigma \in \mathfrak{S}_n$

$$\begin{aligned} \left(\int_s^t q(u) du\right)^n &= \\ &= \sum_{\sigma \in \mathfrak{S}_n} \int_s^t \cdots \int_s^t \mathbb{1}_{\{u_{\sigma(1)} \leq u_{\sigma(2)} \leq \dots \leq u_{\sigma(n)}\}} q(u_1) \dots q(u_n) du_n \dots du_1 = \\ &= n! \int_s^t \cdots \int_s^t \mathbb{1}_{\{u_1 \leq u_2 \leq \dots \leq u_n\}} q(u_1) \dots q(u_n) du_n \dots du_1 = \\ &= \int_s^t \int_{u_1}^t \cdots \int_{u_{n-1}}^t q(u_1)q(u_2) \dots q(u_n) du_n \dots du_2 du_1, \end{aligned}$$

as all the integrals in the sum are equal by the symmetry of the integrand function, and then, by Fubini theorem. \square

Remark A3 (On a fundamental condition). The condition on q stated in Lemma A1 and reformulated in Formula (7) is the key to the proof of important results. In fact we have that this condition is sufficient to ensure that the associated Markov process has no discontinuities of the second type (see [31], p. 227) and, most important for the goals in this work, that the trajectories of the associated Markov process are step functions, that is, any trajectory has only a finite number of jumps in any compact subinterval of $[0, +\infty[$; we will detail this last part of the remark in Theorem A6.

Under the perspective of our main motivation the following result is crucial.

Theorem A6 (The non accumulation property of the jump times of a Markov chain). Let the intensities satisfy condition given by the statement of Lemma A1. Then, given the times $(\tau_n)_{n \geq 1}$, we have that:

$$\mathbb{P} \left[\sum_{n=1}^{+\infty} \tau_n = +\infty \right] = 1, \quad (\text{A14})$$

and so the trajectories of the process are step functions.

Proof. Property in Formula (A14) has non immediate proof. We present a proof based on a result in [62] (p. 160), stating that the condition given by:

$$\limsup_{h \downarrow 0} \sum_{t,i} p(t, i, t+h, j) = 0, \quad (\text{A15})$$

guarantees that the process has a stochastic equivalent that is a step process, meaning that for any trajectory ω the set of jumps of this trajectory has no limit points in the interval $[0, \zeta(\omega)[$, with $\zeta(\omega)$ being the end date of the trajectory. This result is based on a thorough analysis (see [62], pp. 149–159) of the conditions for a Markov process not to have discontinuities of the second type, meaning that the right-hand side and left-hand side limits exists for every date point and every trajectory. Now, with,

$$q(t) := \max_{1 \leq i \leq \#I} |q(t, i, i)|,$$

by virtue of the condition on q in Lemma A1—that is reformulated more precisely in Formula (7) of the statement in Theorem 1—we have that:

$$p(t, i, t+h, j) \leq \sum_{k=1}^{+\infty} \frac{\left(\#I \int_t^{t+h} q(u) du \right)^k}{k!}.$$

Therefore, for almost all $t \in [0, T]$,

$$\begin{aligned} \limsup_{h \downarrow 0} \sum_{t,i} p(t, i, t+h, j) &= (\#I - 1) \limsup_{h \downarrow 0} \sum_{t, k=0}^{+\infty} \frac{\left(\#I \cdot \int_t^{t+h} q(u) du \right)^k}{k!} = \\ &= (\#I - 1) \limsup_{h \downarrow 0} \sum_{t, k=0}^{+\infty} \frac{\left(h \cdot \#I \cdot \frac{1}{h} \int_t^{t+h} q(u) du \right)^k}{k!} = 0, \end{aligned}$$

as the series is uniformly convergent and for almost all $t \in [0, T]$,

$$\lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} q(u) du = q(t),$$

by Lebesgue's differentiation theorem. \square

Remark A4 (Negative properties). *The following negative properties suggest the alternative calibration approach that we propose in Section 3.2. Given $(X_{\tau_n})_{n \geq 0}$, the successive states occupied by the process, we observe that*

- the times $(\tau_n)_{n \geq 1}$ are **not** independent;
- the sequence $(Y_n)_{n \geq 1}$ defined by $Y_n = X_{\tau_n}$ is **not** a Markov chain.

Appendix B. Semi-Markov Processes: A Short Review

For the reader's convenience we present a short summary of the most important results semi-Markov processes (sMp), needed in this work, following [63] (pp. 189–200). The main foundational references for the theory of sMp are [32,64,65]. Important developments can be read in [33,66,67]. Among the many works with relevance for applications we refer, for instance, [68–73]. Let us consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The approach of Markov and semi-Markov processes via kernels is fruitful and so we are lead to the following definitions and results for what we will now follow, mainly, the works in [67] (pp. 7–15) and in [33]. Consider a general measurable state space $(\Theta, \mathcal{A}(\Theta))$. The σ -algebra $\mathcal{A}(\Theta)$ may be seen as the observable sets of the state space of the process Θ .

Definition A4 (Semi-Markov transition kernel). A map $Q : \Theta \times \mathcal{A}(\Theta) \times [0, +\infty[\rightarrow [0, 1]$ such that $(x, B, t) \mapsto Q(x, B, t)$ is a **semi-Markov transition kernel** if it satisfies the following properties.

- (i) $Q(x, \cdot, t)$ is measurable with respect to $\mathcal{A}(\Theta) \times \mathcal{B}([0, +\infty[)$ with $\mathcal{B}([0, +\infty[)$ the Borel σ -algebra of $[0, +\infty[$.
- (ii) For fixed $t > 0$, $Q(\cdot, \cdot, t) : \Theta \times \mathcal{A}(\Theta) \rightarrow [0, 1]$ is a **semistochastic kernel**, that is,
 - (ii.1) For fixed $\theta \in \Theta$ and $t > 0$, the map $Q(\theta, \cdot, t) : \mathcal{A}(\Theta) \rightarrow [0, 1]$ is a measure and we have $Q(\theta, \Theta, t) \leq 1$; if $Q(\theta, \Theta, t) = 1$ we have that $Q(\cdot, \cdot, t)$ is a **stochastic kernel**.
 - (ii.2) For a fixed $T \in \Theta$ we have that $Q(\cdot, T, t) : \Theta \rightarrow [0, 1]$ is measurable with respect to $\mathcal{A}(\Theta)$.
- (iii) For fixed $(\theta, T) \in \Theta \times \mathcal{A}(\Theta)$ we have that the function $Q(\theta, T, t) : [0, +\infty[\rightarrow [0, 1]$ is a nondecreasing function, continuous from the right and such that $Q(\theta, T, 0) = 0$.
- (iv) $P(\cdot, \cdot) : \Theta \times \mathcal{A}(\Theta) \rightarrow [0, 1]$ defined to be: $P(\cdot, \cdot) = Q(\cdot, \cdot, +\infty) = \lim_{t \rightarrow +\infty} Q(\cdot, \cdot, t)$ is a stochastic kernel.
- (v) For any $\theta \in \Theta$ we have that the function defined for $t \in [0, +\infty[$ by $F_\theta(t) := Q(\theta, \Theta, t)$ is a probability distribution function.

Now, consider Q a semi-Markov transition kernel, a continuous time stochastic process $(Y_t)_{t \geq 0}$ defined on this probability space and $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ the natural filtration associated to this process, i.e., $\mathcal{F}_t := \sigma(Y_s : s \leq t)$ is the algebra- σ generated by the variables of the process until time t . We now consider a sequence of random variables $(Z_n)_{n \geq 0}$ —taking values in a state space Θ , that for our purposes will, in general, be finite state space $\Theta = \{\theta_1, \theta_2, \dots, \theta_r\}$ and sometimes an infinite one $\Theta = \{\theta_1, \theta_2, \dots, \theta_r, \dots\}$ —the sequence being adapted to the filtration \mathbb{F} . We consider also $0 \equiv \tau_0 < \tau_1 < \tau_2 < \dots < \tau_n < \dots$ an increasing sequence of \mathbb{F} -stopping times, denoted by \mathcal{T} and $\Delta_n := \tau_n - \tau_{n-1}$ for $n \geq 1$.

Definition A5 (Markov renewal process). A two dimensional discrete time process $(Z_n, \Delta_n)_{n \geq 0}$ with state space $\Theta \times [0, +\infty[$ verifying,

$$\mathbb{P}[Z_{n+1} = \theta_j, \Delta_n \leq t | Z_0, \dots, Z_n, \Delta_1, \Delta_2, \dots, \Delta_n] = \mathbb{P}[Z_{n+1} = \theta_j, \Delta_n \leq t | Z_n],$$

for all $\theta_j \in \Theta$, $t \geq 0$ and almost surely that is, an homogeneous two dimensional Markov Chain, is a **Markov renewal process** if its transition probabilities are given by:

$$Q(\theta, T, t) = \mathbb{P}[Z_{n+1} \in T, \Delta_n \leq t | Z_n = \theta].$$

Remark A5 (Markov chains and Markov renewal processes). The transition probabilities of a Markov renewal process do not depend on the second component; as so, a Markov renewal process is a process of different type of a two dimensional Markov chain process. The first component of a Markov renewal process is a Markov chain, denoted the embedded Markov chain, with transition probabilities given by:

$$P(\theta, T) = Q(\theta, T, +\infty) = \lim_{t \rightarrow +\infty} Q(\theta, T, t) = \mathbb{P}[Z_{n+1} \in T | Z_n = \theta].$$

Definition A6 (Markov renewal times). The **Markov renewal times** of the Markov renewal process $(\tau_n)_{n \geq 0}$ are defined by

$$\tau_n = \sum_{k=1}^n \Delta_k,$$

and the probability distribution functions F_θ of the Markov renewal times depend on the states of the embedded Markov chain, as, by definition we have

$$F_\theta(t) := Q(\theta, \Theta, t) = \mathbb{P}[\Delta_n \leq t | Z_n = \theta].$$

Proposition A1. Consider a general measurable state space $(\Theta, \mathcal{A}(\Theta))$. Let Q be a semi-Markov transition kernel and P the associated stochastic kernel according to Definition A4. Then, there exists a function $F_\theta(\gamma, t)$ such that:

$$Q(\theta, T, t) = \int_T F_\theta(\gamma, t) P(\theta, d\gamma). \quad (A16)$$

Proof. As we have for $\theta \in \Theta$ and $T \in \mathcal{A}(\Theta)$ that $P(\theta, T) = Q(\theta, T, +\infty)$, we may conclude that $Q(\theta, T, +\infty) \leq P(\theta, T)$ and so, the measure $Q(\theta, \cdot, +\infty)$ is absolutely continuous with respect to the probability measure $P(\theta, \cdot)$ on $(\Theta, \mathcal{A}(\Theta))$ and so, by the Radon–Nicodym theorem, there exists a density $F_\theta(\gamma, t)$ verifying Formula (A16). \square

Remark A6 (Semi-Markov kernel for discrete space state). In the case of a discrete state space, say $\Theta = \{\theta_1, \theta_2, \dots, \theta_r, \dots\}$, we may consider $\mathcal{A}(\Theta) = \mathcal{P}(\Theta)$ the maximal σ -algebra of all the subsets of Θ and, with this condition, a semi-Markov kernel Q is defined by a matrix function $Q = [q(i, j, t)]_{i,j \geq 1, t \geq 0}$ such that

- (i) For $i, j \geq 1$ fixed the function $q(i, j, \cdot) : [0, +\infty[\rightarrow [0, 1]$ is nondecreasing.
- (ii) For $i \geq 1$ fixed the function $F_i(t) := \sum_{j \geq 1} q(i, j, t)$ is a probability distribution function.
- (iii) The matrix $P = [p(i, j)]_{i,j \geq 1, t \geq 0}$ with $p(i, j) := q(i, j, +\infty) = \lim_{t \rightarrow +\infty} q(i, j, t)$ is a stochastic matrix.

Definition A7 (Semi-Markov process). The process $(Y_t)_{t \geq 0}$ is a **semi-Markov process** if:

- (i) The process admits a representation given, for $t \geq 0$, by

$$Y_t = \sum_{n=0}^{+\infty} Z_n \mathbb{1}_{[\tau_n, \tau_{n+1}[}(t). \quad (A17)$$

- (ii) For $n \geq 0$ we have that $Z_n = Y_{\tau_n}$.
- (iii) The process $(Z_n, \tau_n)_{n \geq 0}$ is a **Markov renewal process** (Mrp), that is, it verifies

$$\begin{aligned} \mathbb{P}[Z_{n+1} = \theta_j, \tau_{n+1} - \tau_n \leq t | Z_0, \dots, Z_n, \tau_1, \tau_2, \dots, \tau_n] = \\ = \mathbb{P}[Z_{n+1} = \theta_j, \tau_{n+1} - \tau_n \leq t | Z_n], \end{aligned} \quad (A18)$$

for all $\theta_j \in \Theta$, $t \geq 0$ and almost surely—as it is a conditional expectation.

Proposition A2 (The sMp as a Markov chain). The process $(Z_n, \tau_n)_{n \geq 0}$ is a Markov chain with state space $\Theta \times [0, +\infty[$ and with semi-Markov transition kernel given by:

$$q(i, j, t) := \mathbb{P}[Z_{n+1} = \theta_j, \tau_{n+1} - \tau_n \leq t | Z_n = \theta_i]. \quad (A19)$$

Proposition A3 (The embedded Markov chain of the Mrp). The process $(Z_n)_{n \geq 0}$ is a Markov chain with state space Θ with transition probabilities given by:

$$p(i, j) := q(i, j, +\infty) = \mathbb{P}[Z_{n+1} = \theta_j | Z_n = \theta_i], \quad (A20)$$

and is denoted as the **embedded** Markov chain of the Mrp.

Proposition A4 (The conditional distribution function of the time between two successive jumps). Let $Q = [q(i, j, t)]_{i,j \in \{1,2,\dots,r\}, t \geq 0}$ be the semi-Markov kernel as in Proposition A20. Let the times between successive jumps be $\Delta_n := \tau_n - \tau_{n-1}$ have the conditional distribution function of the time between two successive jumps be given by

$$F_{ij}(t) := \mathbb{P}[\Delta_n \leq t | Z_n = \theta_i, Z_{n+1} = \theta_j]. \quad (A21)$$

Then, the semi-Markov kernel verifies,

$$q(i, j, t) := \mathbb{P}[Z_{n+1} = \theta_j, \Delta_n \leq t | Z_n = \theta_i] = p(i, j)F_{ij}(t), \quad (\text{A22})$$

with $p(i, j)$ as defined in Proposition A3.

Proof. It is a consequence of Proposition A1. \square

Remark A7 (Homogeneous Markov chains as semi Markov processes). Let $(X_t)_{t \geq 0}$ be a homogeneous Markov chain in continuous time with state space $\Theta = \{\theta_1, \theta_2, \dots, \theta_r, \dots\}$ and with—time independent—transition intensities given by $Q(t) = [q(i, j)]_{i, j \geq 1}$ (see Definition A3). Then, by the well known results on homogeneous Markov chains (see [29] pp. 317, 318) and by the representation given by Formula (A22), we have that

$$q(t, i, j) = \begin{cases} \frac{q(i, j)}{-q(i, i)} \left(1 - e^{q(i, i)t}\right) & i \neq j, \\ 0 & i = j \text{ or } q(i, i) = 0, \end{cases} \quad (\text{A23})$$

is the semi Markov kernel of a sMp. Being so, comparing Formula (A23) with Formulas (A21) and (A22), we can see that the main difference between a sMp and a continuous time Markov process is the fact that in the sMp case the conditional distribution function of the time between two successive jumps depend not only on the initial state of the jump but also on the final state, while in the homogeneous Markov chain case the dependence is only on the initial state of the jump.

Definition A8 (The sojourn time distribution in a state). The *sojourn time distribution* in the state $\theta_i \in \Theta = \{\theta_1, \theta_2, \dots, \theta_r, \dots\}$, is defined by:

$$H_i(t) := \sum_{j=1}^{+\infty} q(i, j, t) = \sum_{j=1}^{+\infty} p(i, j)F_{ij}(t). \quad (\text{A24})$$

Its mean value represent the mean sojourn time in state θ_i of the sMp $(Y_t)_{t \geq 0}$.

Definition A9 (Regular sMp). A sMp $(Y_t)_{t \geq 0}$ is **regular**, with $N(t)$ the number of jumps of the process in the time interval $]0, t]$ given by:

$$N(t) := \sup\{n \geq 0 : \tau_n \leq t\}, \quad (\text{A25})$$

defined for $t > 0$ verifies for all $\theta_i \in \Theta$,

$$\mathbb{P}_i[N(t) < +\infty] := \mathbb{P}[N(t) < +\infty | Z_0 = \theta_i] = 1. \quad (\text{A26})$$

Proposition A5 (Jumps times of a regular sMp do not have accumulation points). Let the sMp $(Y_t)_{t \geq 0}$ be regular. Then, almost surely, $\lim_{n \rightarrow +\infty} \tau_n = +\infty$ and, for any $T \in \mathbb{R}_+$ and almost all $\omega \in \Omega$:

$$\#\{k \geq 1 : \tau_k(\omega) \leq T\} < +\infty. \quad (\text{A27})$$

This means that in every compact time interval $[0, T]$, for almost all $\omega \in \Omega$ there is only a finite number of times $\tau_k(\omega)$ in this interval.

The following fundamental theorem ensures that for sMp with finite state space the sequence of stopping times do not accumulate in a compact interval.

Theorem A7 (A sufficient condition for regularity of a sMp). Let $\alpha > 0$ and $\beta > 0$ be constants such that or every state θ_i the sojourn time distribution in this state $H_i(t)$ defined in Definition A8 verifies:

$$H_i(\alpha) < 1 - \beta.$$

Then, the sMp is regular. In particular, any sMp with a finite state space is regular.

Proof. See in [74] (p. 88). \square

Remark A8 (On the estimation of sMp). The estimation of sMp is dealt, for instance, in [75,76].

References

- Vajda, S. The stratified semi-stationary population. *Biometrika* **1947**, *34*, 243–254. [[CrossRef](#)] [[PubMed](#)]
- Young, A.; Almond, G. Predicting Distributions of Staff. *Comput. J.* **1961**, *3*, 246–250. [[CrossRef](#)]
- Bartholomew, D.J. A multi-stage renewal process. *J. R. Statist. Soc. Ser. B* **1963**, *25*, 150–168. [[CrossRef](#)]
- Bartholomew, D.J. *Stochastic Models for Social Processes*, 2nd ed.; Wiley Series in Probability and Mathematical Statistics; John Wiley & Sons: London, UK; New York, NY, USA; Sydney, Australia, 1973.
- Bartholomew, D.J. *Stochastic Models for Social Processes*, 3rd ed.; Wiley Series in Probability and Mathematical Statistics; John Wiley & Sons, Ltd.: Chichester, UK, 1982.
- Gani, J. Formulae for Projecting Enrolments and Degrees Awarded in Universities. *J. R. Stat. Soc. Ser. A* **1963**, *126*, 400–409. [[CrossRef](#)]
- Bowerman, B.; David, H.T.; Isaacson, D. The convergence of Cesaro averages for certain nonstationary Markov chains. *Stoch. Process. Appl.* **1977**, *5*, 221–230. [[CrossRef](#)]
- Vassiliou, P.C.G. Cyclic behaviour and asymptotic stability of nonhomogeneous Markov systems. *J. Appl. Probab.* **1984**, *21*, 315–325. [[CrossRef](#)]
- Vassiliou, P.C.G. Asymptotic variability of nonhomogeneous Markov systems under cyclic behaviour. *Eur. J. Oper. Res.* **1986**, *27*, 215–228. [[CrossRef](#)]
- Dimitriou, V.A.; Georgiou, A.C. Introduction, analysis and asymptotic behavior of a multi-level manpower planning model in a continuous time setting under potential department contraction. *Commun. Statist. Theory Methods* **2021**, *50*, 1173–1199. [[CrossRef](#)]
- Salgado-García, R. Open Markov Chains: Cumulant Dynamics, Fluctuations and Correlations. *Entropy* **2021**, *23*, 256. [[CrossRef](#)]
- Vassiliou, P.C.G.; Papadopolou, A.A. Nonhomogeneous semi-Markov systems and maintainability of the state sizes. *J. Appl. Probab.* **1992**, *29*, 519–534. [[CrossRef](#)]
- Papadopolou, A.A.; Vassiliou, P.C.G. Asymptotic behavior of nonhomogeneous semi-Markov systems. *Linear Algebra Appl.* **1994**, *210*, 153–198. [[CrossRef](#)]
- Vassiliou, P.C.G. Asymptotic Behavior of Markov Systems. *J. Appl. Probab.* **1982**, *19*, 851–857. [[CrossRef](#)]
- Vassiliou, P.C.G. Markov Systems in a General State Space. *Commun. Stat. Theory Methods* **2014**, *43*, 1322–1339. [[CrossRef](#)]
- Vassiliou, P.-C.G. Rate of Convergence and Periodicity of the Expected Population Structure of Markov Systems that Live in a General State Space. *Mathematics* **2020**, *8*, 1021. [[CrossRef](#)]
- Vassiliou, P.-C.G. Non-Homogeneous Markov Set Systems. *Mathematics* **2021**, *9*, 471. [[CrossRef](#)]
- McClellan, S.I. A continuous-time population model with Poisson recruitment. *J. Appl. Probab.* **1976**, *13*, 348–354. [[CrossRef](#)]
- McClellan, S.I. Continuous-time stochastic models of a multigrade population. *J. Appl. Probab.* **1978**, *15*, 26–37. [[CrossRef](#)]
- McClellan, S.I. A Semi-Markov Model for a Multigrade Population with Poisson Recruitment. *J. Appl. Probab.* **1980**, *17*, 846–852. [[CrossRef](#)]
- Papadopolou, A.A.; Vassiliou, P.C.G. Continuous time nonhomogeneous semi-Markov systems. In *Semi-Markov Models and Applications (Compiègne, 1998)*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1999; pp. 241–251.
- Esquivel, M.L.; Fernandes, J.M.; Guerreiro, G.R. On the evolution and asymptotic analysis of open Markov populations: application to consumption credit. *Stoch. Models* **2014**, *30*, 365–389. [[CrossRef](#)]
- Guerreiro, G.R.; Mexia, J.A.T.; de Fátima Miguens, M. Statistical approach for open bonus malus. *Astin Bull.* **2014**, *44*, 63–83. [[CrossRef](#)]
- Afonso, L.B.; Cardoso, R.M.R.; Egídio dos Reis, A.D.; Guerreiro, G.R. Ruin Probabilities And Capital Requirement for Open Automobile Portfolios With a Bonus-Malus System Based on Claim Counts. *J. Risk Insur.* **2020**, *87*, 501–522. [[CrossRef](#)]
- Esquivel, M.L.; Patrício, P.; Guerreiro, G.R. From ODE to Open Markov Chains, via SDE: an application to models for infections in individuals and populations. *Comput. Math. Biophys.* **2020**, *8*, 180–197. [[CrossRef](#)]
- Esquivel, M.; Guerreiro, G.; Fernandes, J. Open Markov chain scheme models. *REVSTAT* **2017**, *15*, 277–297.
- Esquivel, M.L.; Guerreiro, G.R.; Oliveira, M.C.; Corte Real, P. Calibration of Transition Intensities for a Multistate Model: Application to Long-Term Care. *Risks* **2021**, *9*. [[CrossRef](#)]
- Resnick, S.I. *Adventures in Stochastic Processes*; Birkhäuser: Boston, MA, USA, 1992.
- Rolski, T.; Schmidli, H.; Schmidt, V.; Teugels, J. *Stochastic Processes for Insurance and Finance*; Wiley Series in Probability and Statistics; John Wiley & Sons Ltd.: Chichester, UK, 1999. [[CrossRef](#)]
- Iosifescu, M. *Finite Markov Processes and Their Applications*; Wiley Series in Probability and Mathematical Statistics; John Wiley & Sons, Ltd.: Chichester, UK; Editura Tehnică: Bucharest, Romania, 1980; p. 295.
- Iosifescu, M.; Tăutu, P. *Stochastic Processes and Applications in Biology and Medicine. I: Theory*; Biomathematics; Editura Academiei RSR: Bucharest, Romania; Springer: Berlin, Germany; New York, NY, USA, 1973; Volume 3, p. 331.

32. Pyke, R. Markov renewal processes: Definitions and preliminary properties. *Ann. Math. Statist.* **1961**, *32*, 1231–1242. [\[CrossRef\]](#)
33. Korolyuk, V.S.; Korolyuk, V.V. Stochastic Models of Systems. In *Mathematics and its Applications*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1999; Volume 469. [\[CrossRef\]](#)
34. Kingman, J.F.C. The imbedding problem for finite Markov chains. *Probab. Theory Relat. Fields* **1962**, *1*, 14–24. [\[CrossRef\]](#)
35. Johansen, S. The Imbedding Problem for Finite Markov Chains. In *Geometric Methods in System Theory*, 1st ed.; Mayne, D.Q.B.R.W., Ed.; D. Reidel Publishing Company: Dordrecht, The Netherlands; Boston, MA, USA, 1973; Volume 1, Chapter 13, pp. 227–237.
36. Johansen, S. A central limit theorem for finite semigroups and its application to the imbedding problem for finite state Markov chains. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **1973**, *26*, 171–190. [\[CrossRef\]](#)
37. Johansen, S. Some Results on the Imbedding Problem for Finite Markov Chains. *J. Lond. Math. Soc.* **1974**, *2*, 345–351. [\[CrossRef\]](#)
38. Fuglede, B. On the imbedding problem for stochastic and doubly stochastic matrices. *Probab. Theory Relat. Fields* **1988**, *80*, 241–260. [\[CrossRef\]](#)
39. Guerry, M.A. On the Embedding Problem for Discrete-Time Markov Chains. *J. Appl. Probab.* **2013**, *50*, 918–930. [\[CrossRef\]](#)
40. Jia, C. A solution to the reversible embedding problem for finite Markov chains. *Stat. Probab. Lett.* **2016**, *116*, 122–130. [\[CrossRef\]](#)
41. Goodman, G.S. An intrinsic time for non-stationary finite Markov chains. *Probab. Theory Relat. Fields* **1970**, *16*, 165–180. [\[CrossRef\]](#)
42. Singer, B. Estimation of Nonstationary Markov Chains from Panel Data. *Sociol. Methodol.* **1981**, *12*, 319–337. [\[CrossRef\]](#)
43. Lencastre, P.; Raischel, F.; Rogers, T.; Lind, P.G. From empirical data to time-inhomogeneous continuous Markov processes. *Phys. Rev. E* **2016**, *93*, 032135. [\[CrossRef\]](#)
44. Ekhsuehi, V.U. On the use of Cauchy integral formula for the embedding problem of discrete-time Markov chains. *Commun. Stat. Theory Methods* **2021**, 1–15. [\[CrossRef\]](#)
45. Coddington, E.A.; Levinson, N. *Theory of Ordinary Differential Equations*; McGraw-Hill Book Company, Inc.: New York, NY, USA; Toronto, ON, Canada; London, UK, 1955.
46. Rudin, W. *Real and Complex Analysis*, 3rd ed.; McGraw-Hill Book Co.: New York, NY, USA, 1987.
47. Kurzweil, J. Ordinary differential equations. In *Studies in Applied Mechanics*; Introduction to the theory of ordinary differential equations in the real domain, Translated from the Czech by Michal Basch; Elsevier Scientific Publishing Co.: Amsterdam, The Netherlands, 1986; Volume 13, p. 440.
48. Teschl, G. Ordinary differential equations and dynamical systems. In *Graduate Studies in Mathematics*; American Mathematical Society: Providence, RI, USA, 2012; Volume 140. [\[CrossRef\]](#)
49. Nevanlinna, F.; Nevanlinna, R. *Absolute Analysis*; Translated from the German by Phillip Emig, Die Grundlehren der mathematischen Wissenschaften, Band 102; Springer: New York, NY, USA; Heidelberg, Germany, 1973.
50. Severi, F.; Scorza Dragoni, G. *Lezioni di analisi. Vol. 3. Equazioni Differenziali Ordinarie e Loro Sistemi, Problemi al Contorno Relativi, Serie Trigonometriche, Applicazioni Geometriche*; Cesare Zuffi: Bologna, Italy, 1951.
51. Dobrušin, R.L. Generalization of Kolmogorov's equations for Markov processes with a finite number of possible states. *Matematicheskii Sbornik* **1953**, *33*, 567–596.
52. Pritchard, D.J. Modeling Disability in Long-Term Care Insurance. *N. Am. Actuar. J.* **2006**, *10*, 48–75. [\[CrossRef\]](#)
53. Kingman, J.F.C. Ergodic properties of continuous-time Markov processes and their discrete skeletons. *Proc. Lond. Math. Soc.* **1963**, *13*, 593–604. [\[CrossRef\]](#)
54. Conner, H. A note on limit theorems for Markov branching processes. *Proc. Am. Math. Soc.* **1967**, *18*, 76–86. [\[CrossRef\]](#)
55. Israel, R.B.; Rosenthal, J.S.; Wei, J.Z. Finding generators for Markov chains via empirical transition matrices, with applications to credit ratings. *Math. Financ.* **2001**, *11*, 245–265. [\[CrossRef\]](#)
56. Guerreiro, G.R.; Mexia, J.A.T. Stochastic vortices in periodically reclassified populations. *Discuss. Math. Probab. Stat.* **2008**, *28*, 209–227. [\[CrossRef\]](#)
57. Feller, W. *An Introduction to Probability Theory and Its Applications. Vol. I*, 3rd ed.; John Wiley & Sons, Inc.: New York, NY, USA; London, UK; Sydney, Australia, 1968.
58. Serfozo, R. Convergence of Lebesgue integrals with varying measures. *Sankhyā Ser. A* **1982**, *44*, 380–402.
59. Billingsley, P. *Convergence of Probability Measures*, 2nd ed.; Wiley Series in Probability and Statistics: Probability and Statistics; John Wiley & Sons, Inc.: New York, NY, USA, 1999. [\[CrossRef\]](#)
60. Durrett, R. *Probability—Theory and Examples*. In *Cambridge Series in Statistical and Probabilistic Mathematics*; Cambridge University Press: Cambridge, UK, 2019; Volume 49. [\[CrossRef\]](#)
61. Skorokhod, A.V. *Lectures on the Theory of Stochastic Processes*; VSP: Utrecht, The Netherlands; TBM Scientific Publishers: Kiev, Ukraine, 1996.
62. Dynkin, E.B. *Theory of Markov Processes*; Translated from the Russian by D. E. Brown and edited by T. Kőváry, Reprint of the 1961 English translation; Dover Publications, Inc.: Mineola, NY, USA, 2006.
63. Iosifescu, M.; Limnios, N.; Oprisan, G. *Introduction to Stochastic Models*; Applied Stochastic Methods Series; Translated from the 2007 French original by Vlad Barbu; ISTE: London, UK; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2010. [\[CrossRef\]](#)
64. Pyke, R. Markov renewal processes with finitely many states. *Ann. Math. Statist.* **1961**, *32*, 1243–1259. [\[CrossRef\]](#)
65. Feller, W. On semi-Markov processes. *Proc. Nat. Acad. Sci. USA* **1964**, *51*, 653–659. [\[CrossRef\]](#)
66. Kurtz, T.G. Comparison of semi-Markov and Markov processes. *Ann. Math. Statist.* **1971**, *42*, 991–1002. [\[CrossRef\]](#)

-
67. Korolyuk, V.; Swishchuk, A. Semi-Markov random evolutions. In *Mathematics and its Applications*; Translated from the 1992 Russian original by V. Zayats and revised by the authors; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1995; Volume 308. [[CrossRef](#)]
 68. Janssen, J.; de Dominicis, R. Finite non-homogeneous semi-Markov processes: Theoretical and computational aspects. *Insur. Math. Econ.* **1984**, *3*, 157–165. [[CrossRef](#)]
 69. Janssen, J.; Limnios, N. (Eds.) *Semi-Markov Models and Applications*; Selected papers from the 2nd International Symposium on Semi-Markov Models: Theory and Applications held in Compiègne, December 1998; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1999. [[CrossRef](#)]
 70. Janssen, J.; Manca, R. *Applied Semi-Markov Processes*; Springer: New York, NY, USA, 2006.
 71. Janssen, J.; Manca, R. *Semi-Markov Risk Models for Finance, Insurance and Reliability*; Springer: New York, NY, USA, 2007.
 72. Barbu, V.S.; Limnios, N. Semi-Markov chains and hidden semi-Markov models toward applications. In *Lecture Notes in Statistics*; Springer: New York, NY, USA, 2008; Volume 191.
 73. Grabski, F. *Semi-Markov Processes: Applications in System Reliability and Maintenance*; Elsevier: Amsterdam, The Netherlands, 2015.
 74. Ross, S.M. *Applied Probability Models with Optimization Applications*; Reprint of the 1970 original; Dover Publications, Inc.: New York, NY, USA, 1992.
 75. Moore, E.H.; Pyke, R. Estimation of the transition distributions of a Markov renewal process. *Ann. Inst. Stat. Math.* **1968**, *20*, 411. [[CrossRef](#)]
 76. Ouhbi, B.; Limnios, N. Nonparametric Estimation for Semi-Markov Processes Based on its Hazard Rate Functions. *Stat. Inference Stoch. Process.* **1999**, *2*, 151–173. [[CrossRef](#)]