COMPUTO

ISSN 2824-7795

Efficient simulation of individual-based population models

The R package IBMPopSim

Daphné Giorgi Laboratoire de Probabilités, Statistique et Modélisation, Sorbonne Université, CNRS
Vincent Lemaire [1] Laboratoire de Probabilités, Statistique et Modélisation, Sorbonne Université, CNRS
Sarah Kaakai [2] Laboratoire Manceau de Mathématiques, Le Mans Université, CNRS, Institut du Risque et de l'Assurance

Date published: 2024-03-17 Last modified: 2024-03-17

Abstract

The R Package `IBMPopSim` aims to simulate the random evolution of heterogeneous populations using stochastic Individual-Based Models (IBMs). The package enables users to simulate population evolution, in which individuals are characterized by their age and some characteristics, and the population is modified by different types of events, including births/arrivals, death/exit events, or changes of characteristics. The frequency at which an event can occur to an individual can depend on their age and characteristics, but also on the characteristics of other individuals (interactions). Such models have a wide range of applications in fields including actuarial science, biology, ecology or epidemiology. `IBMPopSim` overcomes the limitations of time-consuming IBMs simulations by implementing new efficient algorithms based on thinning methods, which are compiled using the `Rcpp` package while providing a user-friendly interface.

Keywords: Individual-based models, stochastic simulation, population dynamics, Poisson measures, thinning method, actuarial science, insurance portfolio simulation

Contents

[1] Corresponding author: vincent.lemaire@sorbonne-universite.fr
[2] Corresponding author: sarah.kaakai@univ-lemans.fr

1 Introduction

In various fields, advances in probability have contributed to the development of a new mathematical framework for so-called individual-based stochastic population dynamics, also called stochastic Individual-Based Models (IBMs).

Stochastic IBMs allow the modeling in continuous time of populations dynamics structured by age and/or characteristics. In the field of mathematical biology and ecology, a large community has used this formalism for the study of the evolution of structured populations (see e.g. (Ferrière and Tran 2009; Collet, Méléard, and Metz 2013; Bansaye and Méléard 2015; Costa et al. 2016; Billiard et al. 2016; Lavallée et al. 2019; Méléard, Rera, and Roget 2019; Calvez et al. 2020)), after the pioneer works (Fournier and Méléard 2004; Champagnat, Ferrière, and Méléard 2006; Tran 2008).

IBMs are also useful in demography and actuarial sciences, for the modeling of human populations dynamics (see e.g. (Bensusan 2010; Boumezoued 2016; El Karoui, Hadji, and Kaakai 2021)). They allow the modeling of heterogeneous and complex population dynamics, which can be used to compute demographic indicators or simulate the evolution of insurance portfolios in order to study the basis risk, compute cash flows for annuity products or pension schemes, or for a fine assessment of mortality models (Barrieu et al. 2012). There are other domains in which stochastic IBMs can be used, for example in epidemiology with stochastic compartmental models, neurosciences, cyber risk, or Agent-Based Models (ABMs) in economy and social sciences, which can be seen as IBMs. Many mathematical results have been obtained in the literature cited above, for quantifying the limit behaviors of IBMs in long time or in large population. In particular, pathwise representations of IBMs have been introduced in (Fournier and Méléard 2004) (and extended to age-structured populations in (Tran 2008)), as measure-valued pure jumps Markov processes, solutions of SDEs driven by Poisson measures. These pathwise representations are based on the *thinning* and projection of Poisson random measures defined on extended spaces. However, the simulation of large and interacting populations is often referred as computationally expensive.

The aim of the `R` package `IBMPopSim` is to meet the needs of the various communities for efficient tools in order to simulate the evolution of stochastic IBMs. `IBMPopSim` provides a general framework for the simulation of a wide class of IBMs, where individuals are characterized by their age and/or a set of characteristics. Different types of events can be included in the modeling by users, depending on their needs: births, deaths, entry or exit in/to the population and changes of characteristics (swap events). Furthermore, the various events that can happen to individuals in the population can occur at a non-stationary frequency, depending on the individuals' characteristics and time, and also including potential interactions between individuals.

We introduce a unified mathematical and simulation framework for this class of IBMs, generalizing the pathwise representation of IBMs by thinning of Poisson measures, as well as the associated population simulation algorithm, based on an acceptance/rejection procedure. In particular, we provide general sufficient conditions on the event intensities under which the simulation of a particular model is possible.

We opted to implement the algorithms of the `IBMPopSim` package using the `Rcpp` package, a tool facilitating the seamless integration of high-performance `C++` code into easily callable `R` functions (Eddelbuettel and Francois 2011). With just a few lines of `C++` code, `IBMPopSim` offers user-friendly R functions for defining IBMs. Once events and their associated intensities are specified, an automated procedure creates the model. This involves integrating the user's source code into the primary `C++` code using a template mechanism. Subsequently, `Rcpp` is invoked to compile the model and integrate it into the `R` session. Following this process, the model becomes callable with varying parameters, enabling the generation of diverse population evolution scenarios. Combined with the design of the simulation algorithms, the package structure yields very competitive simulation runtimes for IBMs, while staying user-friendly for `R` users. Several outputs function are also implemented in `IBMPopSim`. For instance the package allows the construction and visualization of age pyramids, as well as the construction of death and exposures table from the censored individual data, compatible with `R` packages concerned with mortality modelling, such as (Hyndman et al. 2023) or (Villegas, Millossovich, and Kaishev Hyndman 2018). Several examples are provided in the form of `R` vignettes on the website, and in recent works of (El Karoui, Hadji, and Kaakai 2021) and (Roget et al. 2022).

Designed for applications in social sciences, the `R` package `MicSim` (Zinn 2014) can be used for continuous time microsimulation. In continuous-time microsimulation, individual life-courses are usually specified by sequences of state transitions (events) and the time spans between these transitions. The state space is usually discrete and finite, which is no necessarily the case in `IBMPopSim`, where individuals can have continuous characteristics. But most importantly, microsimulation does not allow for interactions between individuals. Indeed, microsimulation produces separately the life courses of all individuals in the populations, based on the computation of the distribution functions of the waiting times in the distinct states of the state space, for each individual (Zinn 2014). This can be slow in comparison to the simulation by thinning of event times occurring in the population, which is based on selecting event times among some competing proposed event times. Finally, `MicSim` simplifies the Mic-Core microsimulation tool implemented in Java (Zinn et al. 2009). However, the implementation in `R` of simulation algorithms yields longer simulation run times than when using `Rcpp`. To the best of our knowledge, there are no other `R` packages currently available addressing the issue of IBMs efficient simulation.

In Section 2, we introduce the mathematical framework that characterizes the class of Stochastic Individual-Based Models (IBMs) that can be implemented in the `IBMPopSim` package. In particular, a general pathwise representation of IBMs is presented. The population dynamics is obtained as the solution of an SDE driven by Poisson measures, for which we obtain existence and uniqueness results in Theorem 1. Additionally, a succinct overview of the package is provided. In Section 3 the two main algorithms for simulating the population evolution of an IBM across the interval $[0, T]$ are detailed. In **?@sec-package** we present the main functions of the `IBMPopSim` package, which allow for the definition of events and their intensities, the creation of a model, and the simulation of scenarios. Two examples are detailed in **?@sec-insurance-portfolio** and sec-example-interaction, featuring applications involving an heterogeneous insurance portfolio characterized by entry and exit events, and an age and size-structured population with intricate interactions.

2 Stochastic Individual-Based Models (IBMs) in IBMPopSim

Stochastic Individual-Based Models (IBMs) represent a broad class of random population dynamics models, allowing the description of populations evolution on a microscopic scale. Informally, an IBM can be summarized by the description of the individuals constituting the population, the various types of events that can occur to these individuals, along with their respective frequencies. In `IBMPopSim`, individuals can be characterized by their age and/or a collection of discrete or continuous characteristics. Moreover, the package enables users to simulate efficiently populations in which one or more of the following event types may occur:

- **Birth event**: addition of an individual of age 0 to the population.
- **Death event**: removal of an individual from the population.
- **Entry event**: arrival of an individual in the population.
- **Exit (emigration) event**: exit from the population (other than death).
- **Swap event**: an individual changes characteristics.

Each event type is linked to an associated event kernel, describing how the population is modified following the occurrence of the event. For some event types, the event kernel requires explicit specification. This is the case for entry events when a new individual joins the population. Then,the model should specify how the age and characteristics of this new individual are chosen. For instance, the characteristics of a new individual in the population can be chosen uniformly in the space of all characteristics, or can depend on the distribution of his parents or those of the other individuals composing the population.

The last component of an IBM are the event intensities. Informally, an event intensity is a function $\lambda_t^e(I, Z)$ describing the frequency at which an event e can occur to an individual I in a population Z at a time t. Given a history of the population $(\mathcal{F}_t)$, the probability of event e occurring to individual I during a small interval of time $(t, t+dt]$ is proportional to $\lambda^e(I, t)$:

$$\mathbb{P}(\text{event } e \text{ occurring to } I \text{ during } (t, t+dt] | \mathcal{F}_t) \simeq \lambda_t^e(I, Z)dt.$$

The intensity function λ^e can include dependency on the individual's I age and characteristics, the time t, or the population composition Z in the presence of interactions.

2.1 Brief package overview

Prior to providing a detailed description of an Individual-Based Model (IBM), we present a simple model of birth and death in an age-structured *human* population. We assume no interactions between individuals, and individuals are characterized by their gender, in addition to their age. In this simple model, all individuals, regardless of gender, can give birth when their age falls between 15 and 40 years, with a constant birth rate of 0.05. The death intensity is assumed to follow a Gompertz-type intensity depending on age. The birth and death intensities are then given by

$$\lambda^b(t, I) = 0.05 \times \mathbf{1}_{[15,40]}(a(I, t)), \quad \lambda^d(t, I) = \alpha \exp(\beta a(I, t)),$$

with $a(I, t)$ the age of individual I at time t. Birth events are also characterized with a kernel determining the gender of the newborn, who is male with probability p_{male}.

2.1.1 Model creation

To implement this model in IBMPopSim, it is necessary to individually define each event type. In this example, the `mk_event_individual` function is used. The creation of an event involves a few lines of cpp instructions defining the intensity and, if applicable, the kernel of the event. For a more in depth description of the event creation step and its parameters, we refer to **?@sec-package_events**.

The events of this simple model are for example defined through the following calls.

```
birth_event <- mk_event_individual(
  type = "birth",
  intensity_code = "result = birth_rate(I.age(t));",
  kernel_code = "newI.male = CUnif(0,1) < p_male;")

death_event <- mk_event_individual(
  type = "death",
  intensity_code = "result = alpha * exp(beta * I.age(t));")
```

In the cpp codes, the names `birth_rate`, `p_male`, `alpha` and `beta` refer to the model parameters defined in the following list.

```
params <- list(
  "alpha" = 0.008, "beta" = 0.02,
  "p_male" = 0.51,
  "birth_rate" = stepfun(c(15, 40), c(0, 0.05, 0)))
```

In a second step, the model is created by calling the function `mk_model`. A cpp source code is automatically created through a template mechanism based on the events and parameters, subsequently compiled using the `sourceCpp` function from the `Rcpp` package.

```
birth_death_model <- mk_model(
  characteristics = c("male" = "bool"),
  events = list(death_event, birth_event),
  parameters = params)
```

2.1.2 Simulation

Once the model is created and compiled, the `popsim` function is called to simulate the evolution of a population according to this model. To achieve this, an initial population must be defined. In this example, we extract a population from a dataset specified in the package (a sample of 100 000 individuals based on the population of England and Wales in 2014). It is also necessary to set bounds for the events intensities. In this example, they are obtained by assuming that the maximum age for an individual is 115 years.

```
a_max <- 115
events_bounds = c(
  "death" = params$alpha * exp(params$beta * a_max),
  "birth" = max(params$birth_rate))
```

The function `popsim` can now be called to simulate the population starting from the initial population `population(EW_pop_14$sample)` up to time $T = 30$.

```
sim_out <- popsim(
  birth_death_model,
  population(EW_pop_14$sample),
  events_bounds,
  parameters = params, age_max = a_max,
  time = 30)
```

The data frame `sim_out$population` contains the information (birth, death, gender) on individuals who lived in the population over the period $[0, 30]$. Functions of the package allows to provide aggregated information on the population.

In the remainder of this section, we define rigorously the class of IBMs that can be simulated in `IBMPopSim`, along with the assumptions that are required in order for the population to be simulatable. The representation of age-structured IBMs based on measure-valued processes, as introduced in (Tran 2008), is generalized to a wider class of abstract population dynamics. The modeling differs slightly here, since individuals are *kept in the population* after their death (or exit), by including the death/exit date as an individual trait.

2.2 Population

2.2.1 Notations

In the remainder of the paper, the filtered probability space is denoted by $(\Omega, \{\mathcal{F}_t\}, \mathbb{P})$, under the usual assumptions. All processes are assumed to be càdlàg and adapted to the filtration $\{\mathcal{F}_t\}$ (for instance the history of the population) on a time interval $[0, T]$. For a càdlàg process X, we denote $X_{t^-} := \lim_{\substack{s \to t \\ s<t}} X_s$.

2.2.2 Individuals

An individual is represented by a triplet $I = (\tau^b, \tau^d, x) \in \mathcal{I} = \mathbb{R} \times \bar{\mathbb{R}} \times \mathcal{X}$ with:

- $\tau^b \in \mathbb{R}$ the date of birth,
- $\tau^d \in \bar{\mathbb{R}}$ the death date, with $\tau^d = \infty$ if the individual is still alive,
- a collection $x \in \mathcal{X}$ of characteristics where $\mathcal{X}$ is the space of characteristics.

Note that in IBMs, individuals are usually characterized by their age $a(t) = t - \tau^b$ instead of their date of birth τ^b. However, using the latter is actually easier for the simulation, as it remains constant over time.

2.2.3 Population process

The population at a given time t is a random set

$$Z_t = \{I_k \in \mathcal{I};\ k = 1, \dots, N_t\},$$

composed of all individuals (alive or dead) who have lived in the population before time t. As a random set, Z_t can be represented by a random counting measure on $\mathcal{I}$, that is an integer-valued measure $Z : \Omega \times \mathcal{I} \to \bar{\mathbb{N}}$ where for $A \in \mathcal{I}$, $Z(A)$ is the (random) number of individuals I in the subset A. With this representation:

$$Z_t(\mathrm{d}\tau^b, \mathrm{d}\tau^d, \mathrm{d}x) = \sum_{k=1}^{N_t} \delta_{I_k}(\tau^b, \tau^d, x),$$
$$\text{with} \int_{\mathcal{I}} f(\tau^b, \tau^d, x) Z_t(\mathrm{d}\tau^b, \mathrm{d}\tau^d, \mathrm{d}x) = \sum_{k=1}^{N_t} f(I_k).$$

The number of individuals present in the population *before time* t is obtained by taking $f \equiv 1$:

$$N_t = \int_{\mathcal{I}} Z_t(\mathrm{d}\tau^b, \mathrm{d}\tau^d, \mathrm{d}x) = \sum_{k=1}^{N_t} \mathbf{1}_{\mathcal{I}}(I_k).$$

Note that $(N_t)_{t\geq 0}$ is an increasing process since dead/exited individuals are kept in the population Z. The number of alive individuals in the population at time t is:

$$N_t^a = \int_{\mathcal{I}} 1_{\{\tau^d > t\}} Z_t(\mathrm{d}\tau^b, \mathrm{d}\tau^d, \mathrm{d}x) = \sum_{k=1}^{N_t} 1_{\{\tau_k^d > t\}}. \tag{1}$$

Another example is the number of alive individuals of age over a is

$$N_t([a,+\infty)) := \int_{\mathcal{I}} \mathbf{1}_{[a,+\infty)}(t-\tau^b) 1_{]t,\infty]}(\tau^d) Z_t(\mathrm{d}\tau^b, \mathrm{d}\tau^d, \mathrm{d}x) = \sum_{k=1}^{N_t} \mathbf{1}_{\{t-\tau_k^b \geq a\}} 1_{\{\tau_k^d \geq t\}}.$$

2.3 Events

The population composition changes at random dates following different types of events. `IBMPopSim` allows the simulation of IBMs with the following events types:

- A **birth** event at time t is the addition of a new individual $I' = (t, \infty, X)$ of age 0 to the population. Their date of birth is $\tau^b = t$, and characteristics is X, a random variable of distribution defined by the birth kernel $k^b(t, I, \mathrm{d}x)$ on $\mathcal{X}$, depending on t and its parent I. The population size becomes $N_t = N_{t^-} + 1$, and the population composition after the event is

$$Z_t = Z_{t^-} + \delta_{(t,\infty,X)}.$$

- An **entry** event at time t is also the addition of an individual I' in the population. However, this individual is not of age 0. The date of birth and characteristics of the new individual $I' = (\tau^b, \infty, X)$ are random variables of probability distribution defined by the entry kernel $k^{en}(t, \mathrm{d}s, \mathrm{d}x)$ on $\mathbb{R}\times\mathcal{X}$. The population size becomes $N_t = N_{t^-}+1$, and the population composition after the event is:

$$Z_t = Z_{t^-} + \delta_{(\tau^b,\infty,X)}.$$

- A **death** or **exit** event of an individual $I = (\tau^b, \infty, x) \in Z_{t^-}$ at time t is the modification of its death date τ^d from $+\infty$ to t. This event results in the simultaneous addition of the individual (τ^b, t, x) and removal of the individual I from the population. The population size is not modified, and the population composition after the event is

$$Z_t = Z_{t^-} + \delta_{(\tau^b,t,x)} - \delta_I.$$

- A **swap** event (change of characteristics) results in the simultaneous addition and removal of an individual. If an individual $I = (\tau^b, \infty, x) \in Z_{t^-}$ changes of characteristics at time t, then it is removed from the population and replaced by $I' = (\tau^b, \infty, X)$. The new characteristics X is a random variable of distribution $k^s(t, I, \mathrm{d}x)$ on $\mathcal{X}$, depending on time, the individual's age and previous characteristics x. In this case, the population size is not modified and the population becomes:

$$Z_t = Z_{t^-} + \delta_{(\tau^b,\infty,X)} - \delta_{(\tau^b,\infty,x)}.$$

To summarize, the space of event types is $E = \{b, en, d, s\}$, and the jump $\Delta Z_t = Z_t - Z_{t^-}$ (change in the population composition) generated by an event of type $e \in \{b, en, d, s\}$ is denoted by $\phi^e(t, I)$. We thus have the following rules summarized in the table Table 1.

Table 1: Action in the population for a given event name

Event	Type	$\phi^e(t, I)$	New individual
Birth	b	$\delta_{(t,\infty,X)}$	$\tau^b = t,\ X \sim k^b(t, I, \mathrm{d}x)$
Entry	en	$\delta_{(\tau^b,\infty,X)}$	$(\tau^b, X) \sim k^{en}(t, \mathrm{d}s, \mathrm{d}x)$
Death/Exit	d	$\delta_{(\tau^b,t,x)} - \delta_{(\tau^b,\infty,x)}$	$\tau^d = t$
Swap	s	$\delta_{(\tau^b,\infty,X)} - \delta_{(\tau^b,\infty,x)}$	$X \sim k^s(t, I, \mathrm{d}x)$

Remark 2.1 (Composition of the population).

- At time T, the population Z_T contains all individuals who lived in the population before T, including dead/exited individuals. If there are no swap events, or entries,the population state Z_t for any time $t \leq T$ can be obtained from Z_T. Indeed, if $Z_T = \sum_{k=1}^{N_T} \delta_{I_k}$, then the population at time $t \leq T$ is simply composed of the individuals born before t:

$$Z_t = \sum_{k=1}^{N_T} \mathbf{1}_{\{\tau_k^b \leq t\}} \delta_{I_k}.$$

- In the presence of entries (open population), a characteristic x can track the individuals' entry dates. Then, the previous equation can be easily modified in order to obtain the population Z_t at time $t \leq T$ from Z_T.

2.4 Events intensity

Once the different event types have been defined in the population model, the frequency at which each event occur in the population e have to be specified. Informally, the intensity $\Lambda_t^e(Z_t)$ at which an event e can occur is defined by

$$\mathbb{P}\big(\text{event } e \text{ occurs in the population } Z_t \in (t, t + \mathrm{d}t] | \mathcal{F}_t\big) \simeq \Lambda_t^e(Z_t)\mathrm{d}t.$$

For a more formal definition of stochastic intensities, we refer to (Brémaud 1981) or (Kaakai and El Karoui 2023). The form of the intensity function $(\Lambda_t^e(Z_t))$ determines the population simulation algorithm in `IBMPopSim`:

- When the event intensity does not depend on the population state,

$$\big(\Lambda_t^e(Z_t)\big)_{t\in[0,T]} = \big(\mu^e(t)\big)_{t\in[0,T]}, \tag{2}$$

  with μ^e a deterministic function, the events of type e occur at the jump times of an inhomogeneous Poisson process of intensity function $(\mu^e(t))_{t\in[0,T]}$. When such an event occurs, the individual to whom the event happens to is drawn uniformly among alive individuals in the population. In a given model, the set of events $e \in E$ with Poisson intensities will be denoted by $\mathcal{P}$.
- Otherwise, we assume that the global intensity $\Lambda_t^e(Z_t)$ at which the events of type e occur in the population can be written as the sum of individual intensities $\lambda_t^e(I, Z_t)$:

$$\Lambda_t^e(Z_t) = \sum_{k=1}^{N_t} \lambda_t^e(I_k, Z_t),$$
$$\text{with } \mathbb{P}\big(\text{event } e \text{ occurs to an individual } I \in (t, t + \mathrm{d}t] | \mathcal{F}_t\big) \simeq \lambda_t^e(I, Z_t)\mathrm{d}t.$$

Obviously, nothing can happen to dead or exited individuals, i.e. individuals $I = (\tau^b, \tau^d, x)$ with $\tau^d \leq t$. Thus, individual event intensities are assumed to be null for dead/exited individuals:

$$\lambda_t^e(I, Z_t) = 0, \text{ if } \tau^d \leq t, \text{ so that } \Lambda_t^e(Z_t) = \sum_{k=1}^{N_t^a} \lambda_t^e(I_k, Z_t),$$

with N_t^a the number of alive individuals at time t.

The event's individual intensity $\lambda_t^e(I, Z_t)$ can depend on time (for instance when there is a mortality reduction over time), on the individual's age $t - \tau^b$ and characteristics, but also on the population composition Z_t. The dependence of λ^e on the population Z models interactions between individuals in the populations. Hence, two types of individual intensity functions can be implemented in `IBMPopSim`:

1. *No interactions:* The intensity function λ^e does not depend on the population composition. The intensity at which the event of type e occur to an individual I only depends on its date of birth and characteristics:
$$\lambda_t^e(I, Z_t) = \lambda^e(t, I), \tag{3}$$
where $\lambda^e : \mathbb{R}_+ \times \mathcal{I} \to \mathbb{R}^+$ is a deterministic function. In a given model, we denote by $\mathcal{E}$ the set of event types with individual intensity Equation 3.

2. *"Quadratic" interactions:* The intensity at which an event of type e occurs to an individual I depends on I and on the population composition, through an interaction function W^e. The quantity $W^e(t, I, J)$ describes the intensity of interactions between two alive individuals I and J at time t, for instance in the presence of competition or cooperation. In this case, we have
$$\lambda_t^e(I, Z_t) = \sum_{j=1}^{N_t} W^e(t, I, I_j) = \int_{\mathcal{I}} W^e(t, I, (\tau^b, \tau^d, x)) Z_t(\mathrm{d}\tau^b, \mathrm{d}\tau^d, \mathrm{d}x), \tag{4}$$
where $W^e(t, I, (\tau^b, \tau^d, x)) = 0$ if the individual $J = (\tau^b, \tau^d, x)$ is dead, i.e. $\tau^d \leq t$. In a given model, we denote by $\mathcal{E}_W$ the set of event types with individual intensity Equation 4.

To summarize, an individual intensity in IBMPopSim can be written as:

$$\lambda_t^e(I, Z_t) = \lambda^e(t, I)\mathbf{1}_{\{e \in \mathcal{E}\}} + \left(\sum_{j=1}^{N_t} W^e(t, I, I_j)\right)\mathbf{1}_{\{e \in \mathcal{E}_W\}}. \tag{5}$$

Example 2.1.

1. An example of death intensity without interaction for an individual $I = (\tau^b, \tau^d, x)$ alive at time t, $t < \tau^d$, is:
$$\lambda^d(t, I) = \alpha_x \exp(\beta_x a(I, t)), \text{ where } a(I, t) = t - \tau^b$$
is the age of the individual I at time t. In this classical case, the death rate of an individual I is an exponential (Gompertz) function of the individual's age, with coefficients depending on the individual's characteristics x.

2. In the presence of competition between individuals, the death intensity of an individual I also depend on other individuals J in the population. For example, if $I = (\tau^b, \tau^d, x)$, with x its size, then we can have:
$$W^d(t, I, J) = (x_J - x)^+ \mathbf{1}_{\{\tau_J^d > t\}}, \quad \forall\, J = (\tau_J^b, \tau_J^d, x_J). \tag{6}$$
This can be interpreted as follows: if the individual I meets randomly an individual J alive at time t, and of bigger size $x_J > x$, then he can die at the intensity $x_J - x$. If J is smaller than

I, then he cannot kill I. The bigger is the size x of I, the lower is his death intensity $\lambda_t^d(I, Z_t)$ defined by

$$\lambda_t^d(I, Z_t) = \sum_{\substack{J \in Z_t, \\ x_J > x}} (x_J - x) \mathbf{1}_{\{\tau_J^d > t\}}.$$

3. `IBMPopSim` can simulate IBMs that include intensities expressed as a sum of Poisson intensities and individual intensities of the form $\Lambda^e(Z_t) = \mu_t^e + \sum_{k=1}^{N_t} \lambda^e(I_k, Z_t)$. Other examples are provided in **?@sec-insurance-portfolio** and **?@sec-example-interaction**. Finally, the global intensity at which an event can occur in the population is defined by:

$$\Lambda_t(Z_t) = \sum_{e \in \mathscr{P}} \mu^e(t) + \sum_{e \in \mathscr{E}} \Big(\sum_{k=1}^{N_t} \lambda^e(t, I_k) \Big) + \sum_{e \in \mathscr{E}_W} \Big(\sum_{k=1}^{N_t} \sum_{j=1}^{N_t} W^e(t, I_k, I_j) \Big). \tag{7}$$

An important point is that for events $e \in \mathscr{E}$ without interactions, the global event intensity $\Lambda_t^e(Z_t) = \sum_{k=1}^{N_t} \lambda^e(t, I_k)$ is *of order* N_t^a defined in Equation 1 (number of alive individuals at time t). On the other hand, for events $e \in \mathscr{E}_W$ with interactions, $\Lambda_t^e(Z_t) = \sum_{k=1}^{N_t} \sum_{j=1}^{N_t} W^e(t, I_k, I_j)$ is of order $(N_t^a)^2$. Informally, this means that when the population size increases, events with interaction are more costly to simulate. Furthermore, the numerous computations of the interaction kernel W^e can also be quite costly. The randomized Algorithm [algo::rzndomized], detailed in Section 2.3, allows us to overcome these limitations.

2.4.1 Events intensity bounds

The simulation algorithms implemented in `IBMPopSim` are based on an acceptance/rejection procedure, which requires to specify bounds for the various events intensities $\Lambda_t^e(Z_t)$. These bounds are defined differently depending on the expression of the intensity.

Assumption 2.1. *For all events $e \in \mathscr{P}$ with Poisson intensity Equation 2, the intensity is assumed to be bounded on $[0, T]$:*

$$\forall t \in [0, T], \quad \Lambda_t^e(Z_t) = \mu^e(t) \leq \bar{\mu}^e.$$

When $e \in \mathscr{E} \cup \mathscr{E}_W$, $\Lambda_t^e(Z_t) = \sum_{k=1}^{N_t} \lambda_t^e(I_k, Z_t)$, assuming that $\Lambda_t^e(Z_t)$ is uniformly bounded is too restrictive since the event intensity depends on the population size. In this case, the assumption is made on the individual intensity λ^e or on the interaction function W^e, depending on the situation.

Assumption 2.2. *For all event types $e \in \mathscr{E}$, the associated individual event intensity λ^e with no interactions,* i.e. *λ^e verifies Equation 3, is assumed to be uniformly bounded:*

$$\lambda^e(t, I) \leq \bar{\lambda}^e, \quad \forall\, t \in [0, T],\ I \in \mathscr{I}.$$

In particular,

$$\forall t \in [0, T], \quad \Lambda_t^e(Z_t) = \sum_{k=1}^{N_t} \lambda^e(t, I) \leq \bar{\lambda}^e N_t. \tag{8}$$

Assumption 2.3. *For all event types $e \in \mathscr{E}_W$, the associated interaction function W^e is assumed to be uniformly bounded:*

$$W^e(t, I, J) \leq \bar{W}^e, \quad \forall\, t \in [0, T],\ I, J \in \mathscr{I}.$$

In particular, $\forall t \in [0, T]$,

$$\lambda_t^e(I, Z_t) = \sum_{j=1}^{N_t} W^e(t, I, I_j) \leq \bar{W}^e N_t, \quad \textit{and} \quad \Lambda_t^e(Z_t) \leq \bar{W}^e (N_t)^2.$$

Assumption 2.1, Assumption 2.2 and Assumption 2.3 yield that events in the population occur with the global event intensity $\Lambda_t(Z_t)$, given in Equation 7, which is dominated by a polynomial function in the population size:

$$\Lambda_t(Z_t) \leq \bar{\Lambda}(N_t), \quad \text{with } \bar{\Lambda}(n) = \sum_{e \in \mathscr{P}} \bar{\mu}^e + \sum_{e \in \mathscr{E}} \bar{\lambda}^e n + \sum_{e \in \mathscr{E}_W} \bar{W}^e n^2. \tag{9}$$

This bound is linear in the population size if there are no interactions, and quadratic if there at least is an event including interactions. This assumption is the key to the algorithms implemented in `IBMPopSim`. Before presenting the simulation algorithm, we close this section with a rigorous definition of an IBM, based on the pathwise representation of its dynamics a Stochastic Differential Equation (SDE) driven by Poisson random measures.

2.5 Pathwise representation

Since the seminal paper of (Fournier and Méléard 2004), it has been shown in many examples that a stochastic IBM dynamics can be defined rigorously as the unique solution of an SDE driven by Poisson measures, under reasonable non explosion conditions. In the following, we introduce a unified framework for the pathwise representation of the class of stochastic IBMs introduced above. Some recalls on Poisson random measures are presented in the Appendix Section 3.4, and for more details on these representations on particular examples, we refer to the abundant literature on the subject.

In the following we consider an individual-based stochastic population $(Z_t)_{t \in [0,T]}$, keeping the notations introduced in Section 2.3 and Section 2.4 for the events and their intensities. In particular, the set of events types that define the population evolution is denoted by $\mathscr{P} \cup \mathscr{E} \cup \mathscr{E}_W \subset E$, with $\mathscr{P}$the set of events types with Poisson intensity verifying Assumption 2.1, $\mathscr{E}$ the set of events types with individual intensity and no interaction, verifying Assumption 2.2 and finally $\mathscr{E}_W$ the set of event types with interactions, verifying Assumption 2.3.

2.5.1 Non explosion criterion

First, one has to ensure that the number of events occurring in the population will not explode in finite time, leading to an infinite simulation time. Assumption 2.2 and Assumption 2.3 are not sufficient to guarantee the non explosion of the event number, due to the potential explosion of the population size in the presence of interactions. An example is the case when only birth events occur, with an intensity $\Lambda_t^b(Z_t) = C_b(N_t^a)^2$ ($W^b(t, I, J) = C_b$). Then, the number of alive individuals $(N_t^a)_{t \geq 0}$ is a well-known pure birth process of intensity function $g(n) = C_b n^2$ (intensity of moving from state n to $n+1$). This process explodes in finite time, since g does not verify the necessary and sufficient non explosion criterion for pure birth Markov processes: $\sum_{n=1}^{\infty} \frac{1}{g(n)} = \infty$ (see e.g. Theorem 2.2 in (Bansaye and Méléard 2015)). There is thus an explosion in finite time of birth events.

This example shows that the important point for non explosion is to control the population size. We give below a general sufficient condition on birth and entry event intensities, in order for the population size to stay finite in finite time. This ensures that the number of events does not explode in finite time. Informally, the idea is to control the intensities by a pure birth intensity function verifying the non-explosion criterion.

Assumption 2.4. *Let $e = b$ or en, a birth or entry event type. If the intensity at which the events of type e occur in the population are not Poissonian, i.e. $e \in \mathscr{E} \cup \mathscr{E}_W$, then there exists a function $f^e : \mathbb{N} \to (0, +\infty)$, such that*

$$\sum_{n=1}^{\infty} \frac{1}{n f^e(n)} = \infty,$$

and for all individual $I \in \mathcal{J}$ and population measure $Z = \sum_{k=1}^{n} \delta_{I_k}$ of size n,

$$\lambda_t^e(I, Z) \leq f^e(n), \ \forall\, 0 \leq t \leq T.$$

If $e \in \mathcal{E}$, $\lambda_t^e(I, Z) = \lambda^e(t, I) \leq \bar{\lambda}^e$ by the domination Assumption 2.3, then Assumption 2.4 is always verified with $f^e(n) = \bar{\lambda}^e$.

Assumption 2.4 yields that the global intensity $\Lambda_t^e(\cdot)$ of event e is bounded by a function g^e only depending on the population size:

$$\Lambda_t^e(Z) \leq g^e(n) := n f^e(n), \quad \text{with} \sum_{n=1}^{\infty} \frac{1}{g^e(n)} = \infty.$$

If $e \in \mathcal{P}$ has a Poisson intensity, then $\Lambda_t^e(Z) = \mu_t^e$ always verifies the previous equation with $g^e(n) = \bar{\mu}^e$.

Before introducing the IBM SDE, let us give an idea of the equation construction. Between two successive events, the population composition Z_t stays constant, since the population process $(Z_t)_{t\geq 0}$ is a pure jump process. Furthermore, since each event type is characterized by an intensity function, the jumps occurring in the population can be represented by restriction and projection of a Poisson measure defined on a larger state space. More precisely, we introduce a random Poisson measure Q on $\mathbb{R}^+ \times \mathcal{J} \times \mathbb{R}^+$, with $\mathcal{J} = \mathbb{N} \times (\mathcal{E} \cup \mathcal{E}_W)$. Q is composed of random quadruplets (τ, k, e, θ), where τ represents a potential event time for an individual I_k and event type e. The last variable θ is used to accept/reject this proposed event, depending on the event intensity. Hence, the Poisson measure is restricted to a certain random set and then projected on the space of interest $\mathbb{R}^+ \times \mathcal{J}$. If the event is accepted, then a jump $\phi^e(\tau, I_k)$ occurs.

The proof of Theorem 2.1 is detailed in the Appendix **?@prf-thm-eq-Z**. Note that Equation 10 is an SDE describing the evolution of the IBM, the intensity of the events in the right hand side of the equation depending on the population process Z itself. The main idea of the proof of Theorem 2.1 is to use the non explosion property of Lemma 2.1, and to write the r.h.s of Equation 10 as a sum of simple equations between two successive events, solved by induction.

The proof of Lemma 2.1 , detailed in Appendix **?@prf-lem-non-explosion** is more technical and rely on pathwise comparison result, generalizing those obtained in (Kaakai and El Karoui 2023). An alternative pathwise representation of the population process, inspired by the randomized Algorithm [algo::rzndomized], is given as well in Theorem [ThEqZrandomized].

Theorem 2.1 (Pathwise representation). *Let $T \in \mathbb{R}^+$ and $\mathcal{J} = \mathbb{N} \times (\mathcal{E} \cup \mathcal{E}_W)$. Let Q be a random Poisson measure on $\mathbb{R}^+ \times \mathcal{J} \times \mathbb{R}^+$, of intensity $\mathrm{d}t\delta_{\mathcal{J}}(\mathrm{d}k, \mathrm{d}e)\mathbf{1}_{[0,\bar{\lambda}^e]}(\theta)\mathrm{d}\theta$, with $\delta_{\mathcal{J}}$ the counting measure on $\mathcal{J}$. Finally, let $Q^{\mathcal{P}}$ be a random Poisson measure on $\mathbb{R}^+ \times \mathcal{P} \times \mathbb{R}^+$, of intensity $\mathrm{d}t\delta_P(\mathrm{d}e)\mathbf{1}_{[0,\bar{\mu}^e]}(\theta)\mathrm{d}\theta$, and $Z_0 = \sum_{k=1}^{N_0} \delta_{I_k}$ an initial population. Then, under Assumption 2.4 , there exists a unique measure-valued population process Z, strong solution on the following SDE driven by the Poisson measure Q:*

$$\begin{aligned} Z_t = Z_0 &+ \int_0^t \int_{\mathcal{J}\times\mathbb{R}^+} \phi^e(s, I_k)\mathbf{1}_{\{k\leq N_{s^-}\}}\mathbf{1}_{\{\theta\leq\lambda_s^e(I_k,Z_{s^-})\}}Q(\mathrm{d}s, \mathrm{d}k, \mathrm{d}e, \mathrm{d}\theta) \\ &+ \int_0^t \int_{\mathcal{P}\times\mathbb{R}^+} \phi^e(s, I_{s^-})\mathbf{1}_{\{\theta\leq\mu^e(s)\}}Q^{\mathcal{P}}(\mathrm{d}s, \mathrm{d}e, \mathrm{d}\theta), \qquad \forall 0 \leq t \leq T, \end{aligned}$$

and where I_{s^-} is an individual, chosen uniformly among alive individuals in the population Z_{s^-}.

Lemma 2.1. *Let Z be a solution of Equation 10 on $\mathbb{R}^+$, with $(T_n)_{n\geq 0}$ its jump times, $T_0 = 0$. If Assumption 2.4 is satisfied, then*

$$\lim_{n\to\infty} T_n = \infty, \quad \mathbb{P}\text{-a.s.}$$

3 Population simulation

We now present the main algorithm for simulating the evolution of an IBM over $[0, T]$.The algorithm implemented in `IBMPopSim` allows the exact simulation of Equation 10, based on an acceptance/reject algorithm for simulating random times called *thinning*. The exact simulation of event times with this acceptance/reject procedure is closely related to the simulations of inhomogeneous Poisson processes by the so-called thinning algorithm, often attributed to (Lewis and Shedler 1979). The simulation methods for inhomogeneous Poisson processes can be adapted to IBMs, and we introduce in this section a general algorithm extending those by (Fournier and Méléard 2004) (see also (Ferrière and Tran 2009), (Bensusan 2010)).

The algorithm is based on exponential "candidate" event times, chosen with a (constant) intensity which must be greater than the global event intensity $\Lambda_t(Z_t)$ (Equation 3). Starting from time t, once a candidate event time $t+\bar{T}_\ell$ has been proposed, a candidate event type e (birth, death,...) is chosen with a probability p^e depending on the event intensity bounds $\bar{\mu}^e$, $\bar{\lambda}^e$ and $\bar{W}^e$, as defined in Assumption 2.2 and Assumption 2.3. An individual I is then drawn from the population. Finally, it remains to accept or reject the candidate event with a probability $q^e(t, I, Z_t)$ depending on the true event intensity. If the candidate event time is accepted, then the event e occurs at time $t + \bar{T}_\ell$ to the individual I. The main idea of the algorithm implemented can be summarized as follows:

1. Draw a candidate time $t + \bar{T}_\ell$ and candidate event type e.
2. Draw a uniform variable $\theta \sim \mathcal{U}([0, 1])$ and individual I.
3. **If** $\theta \leq q^e(t, I, Z_t)$ **then** event e occur to individual I, **else** Do nothing and start again from $t + \bar{T}_\ell$.

Before introducing the main algorithms in more details, we recall briefly the thinning procedure for simulating inhomogeneous Poisson processes, as well as the links with pathwise representations. Some recalls on Poisson random measures are presented in Section 3.4. For a more general presentation of thinning of a Poisson random measure, see (Devroye 1986; Çinlar 2011; Kallenberg 2017).

3.1 Thinning of Poisson measure

Let us start with the simulation and pathwise representation of an inhomogeneous Poisson process on $[0, T]$ with intensity $(\Lambda(t))_{t\in[0,T]}$. The thinning procedure is based on the fundamental assumption that $\Lambda(t) \leq \bar{\Lambda}$ is bounded on $[0, T]$. In this case, the inhomogeneous Poisson can be obtained from an homogeneous Poisson process of intensity $\bar{\Lambda}$, which can be simulated easily.

First, the Poisson process can be extended to a Marked Poisson measure $\bar{Q} := \sum_{\ell\geq 1} \delta_{(\bar{T}_\ell, \bar{\Theta}_\ell)}$ on $(\mathbb{R}^+)^2$, defined as follow:

- The jump times of $(\bar{T}_\ell)_{\ell\geq 1}$ of $\bar{Q}$ are the jump times of a Poisson process of intensity $\bar{\Lambda}$.
- The marks $(\bar{\Theta}_\ell)_{\ell\geq 1}$ are *i.i.d.* random variables, uniformly distributed on $[0, \bar{\Lambda}]$.

By Proposition 3.6 , $\bar{Q}$ is a Poisson random measure with mean measure

$$\bar{\mu}(\mathrm{d}t, \mathrm{d}\theta) := \bar{\Lambda}\mathrm{d}t\frac{\mathbf{1}_{[0,\bar{\Lambda}]}(\theta)}{\bar{\Lambda}}\mathrm{d}\theta = \mathrm{d}t\mathbf{1}_{[0,\bar{\Lambda}]}(\theta)\mathrm{d}\theta.$$

In particular, the average number of atoms $(\bar{T}_\ell, \bar{\Theta}_\ell)$ in $[0, t] \times [0, h]$ is

$$\mathbb{E}[Q([0, t] \times [0, h])] = \mathbb{E}[\sum_\ell \mathbf{1}_{[0,t]\times[0,h]}(\bar{T}_\ell, \bar{\Theta}_\ell)] = \int_{(\mathbb{R}^+)^2} \bar{\mu}(\mathrm{d}t, \mathrm{d}\theta) = t(\bar{\Lambda} \wedge h).$$

The thinning is based on the restriction property for Poisson measure: for a measurable set $\Delta \subset \mathbb{R}^+ \times \mathbb{R}^+$, the restriction $Q^\Delta := \mathbf{1}_\Delta \bar{Q}$ of $\bar{Q}$ to Δ (by taking only atoms in Δ) is also a Poisson random measure of mean measure $\mu^\Delta(\mathrm{d}t, \mathrm{d}\theta) = \mathbf{1}_\Delta(t, \theta)\bar{\mu}(\mathrm{d}t, \mathrm{d}\theta)$.

In order to obtain an inhomogeneous Poisson measure of intensity $(\Lambda(t))$, the "good" choice of Δ is the hypograph of Λ: $\Delta = \{(t,\theta) \in [0,T] \times [0,\bar{\Lambda}];\ \theta \leq \Lambda(t)\}$ (see Figure 1). Then,

$$Q^{\Delta} = \sum_{\ell \geq 1} \mathbf{1}_{\{\bar{\Theta}_\ell \leq \Lambda(\bar{T}_\ell)\}} \delta_{(\bar{T}_\ell, \bar{\Theta}_\ell)},$$

and since $\Lambda(t) \leq \bar{\Lambda}$, on $[0,T]$:

$$\mu^{\Delta}(\mathrm{d}t, \mathrm{d}\theta) = \mathbf{1}_{\{\theta \leq \Lambda(t)\}} \mathrm{d}t \mathbf{1}_{[0,\bar{\Lambda}]}(\theta) \mathrm{d}\theta = \mathbf{1}_{\{\theta \leq \Lambda(t)\}} \mathrm{d}t \mathrm{d}\theta.$$

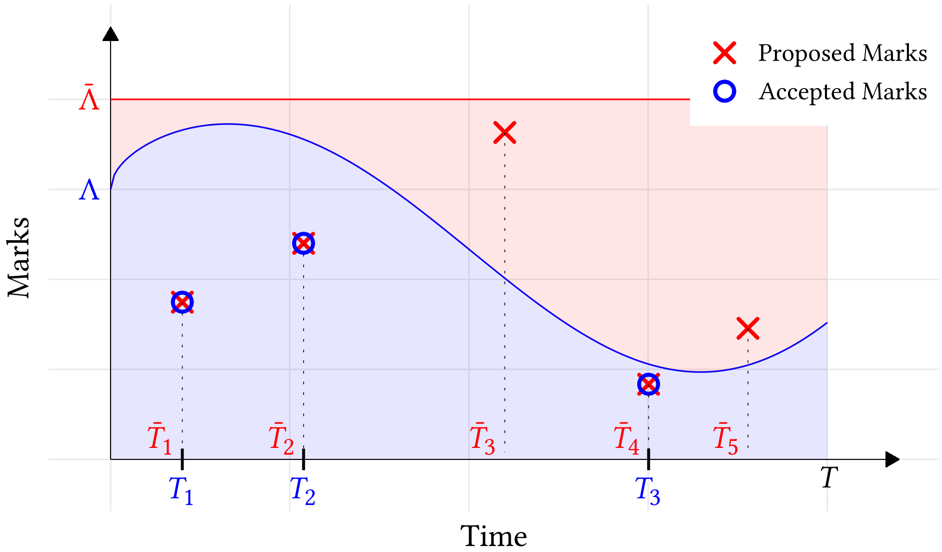

Figure 1: test

Finally, the inhomogeneous Poisson process is obtained by the projection Proposition 3.5, which states that the jump times of Q^{Δ} are the jump times of an inhomogeneous Poisson process of intensity $(\Lambda(t))$.

Proposition 3.1. *The counting process N^{Λ}, projection of Q^{Δ} on the time component and defined by,*

$$N_t^{\Lambda} := Q^{\Delta}([0,t] \times \mathbb{R}^+) = \int_0^t \int_{\mathbb{R}^+} \mathbf{1}_{\{\theta \leq \Lambda(s)\}} \bar{Q}(\mathrm{d}s, \mathrm{d}\theta) = \sum_{\ell \geq 1} 1_{\{\bar{T}_\ell \leq t\}} 1_{\{\bar{\Theta}_\ell \leq \Lambda(\bar{T}_\ell)\}}, \quad \forall t \in [0,T], \tag{10}$$

is an inhomogeneous Poisson process on $[0,T]$ of intensity function $(\Lambda(t))_{t \in [0,T]}$. The thinning Equation 10 is a pathwise representation of N^{Λ} by restriction and projection *of the Poisson measure Q on $[0,T]$.*

The previous proposition yields a straightforward thinning algorithm to simulate the jump times $(T_k)_{k \geq 1}$ of an inhomogeneous Poisson process of intensity $\Lambda(t)$, by selecting jump times $\bar{T}_\ell$ such that $\bar{\Theta}_\ell \leq \Lambda(\bar{T}_\ell)$.

3.1.1 Multivariate Poisson process

This can be extended to the simulation of multivariate inhomogeneous Poisson processes, which is an important example before tackling the simulation of an IBM.

Let $(N^j)_{j \in \mathcal{J}}$ be a (inhomogeneous) multivariate Poisson process indexed by a finite set $\mathcal{J}$, such that $\forall j \in \mathcal{J}$, the intensity $(\lambda_j(t))_{t \in [0,T]}$ of N_j is bounded on $[0,T]$:

$$\sup_{t \in [0,T]} \lambda_j(t) \leq \bar{\lambda}_j, \text{ and let } \bar{\Lambda} = \sum_{j \in \mathcal{J}} \bar{\lambda}_j.$$

Recall that such multivariate counting process can be rewritten as a Poisson random measure $N = \sum_{k\geq 1} \delta_{(T_k, J_k)}$ on $\mathbb{R}^+ \times \mathcal{J}$ (see e.g. Sec. 2 of Chapter 6 in (Çinlar 2011)), where T_k is the kth jump time of $\sum_{j\in\mathcal{J}} N^j$ and J_k corresponds to the component of the the vector which jumps. In particular, $N_t^j = N([0,t] \times \{j\})$.

Once again the simulation of such process can be obtained from the simulation of a (homogeneous) multivariate Poisson process of intensity vector $(\bar{\lambda}_j)_{j\in\mathcal{J}}$, extended into a Poisson measures by adding marks on $\mathbb{R}^+$. Thus, we introduce the Marked Poisson measure $\bar{Q} = \sum \delta_{(\bar{T}_\ell, \bar{J}_\ell, \bar{\Theta}_\ell)}$ on $\mathbb{R}^+ \times \mathcal{J} \times \mathbb{R}^+$, such that:

- The jump times $(\bar{T}_\ell)$ of $\bar{Q}$ are the jump times of a Poisson measure of intensity $\bar{\Lambda}$.
- The variables $(\bar{J}_\ell)$ are *i.i.d.* random variables on $\mathcal{J}$, with $p_j = \mathbb{P}(\bar{J}_1 = j) = \bar{\lambda}_j/\bar{\Lambda}$ and representing the component of the vector which jumps.
- The marks $(\bar{\Theta}_\ell)$ are independent variables with $\bar{\Theta}_\ell$ a uniform random variable on $[0, \bar{\lambda}_{\bar{J}_\ell}]$, $\forall \ell \geq 1$.

By Proposition 3.6 and Proposition 3.5, each measure $\bar{Q}_j(\mathrm{d}t, \mathrm{d}\theta) = \bar{Q}(\mathrm{d}t, \{j\}, \mathrm{d}\theta) = \sum_{\ell\geq 1} 1_{\{\bar{J}_\ell = j\}} \delta_{(\bar{T}_\ell, \bar{\Theta}_\ell)}$ is a marked Poisson measure of intensity

$$\bar{\mu}_j(\mathrm{d}t, \mathrm{d}\theta) = \bar{\Lambda} p_j \mathrm{d}t \frac{1_{\{\theta \leq \bar{\lambda}_j\}}(\theta)}{\bar{\lambda}_j} \mathrm{d}\theta = \mathrm{d}t 1_{\{\theta \leq \bar{\lambda}_j\}}(\theta) \mathrm{d}\theta.$$

As a direct application of Proposition 3.1 , the inhomogeneous multivariate Poisson process is obtained by restriction of each measures $\bar{Q}_j$ to $\Delta_j = \{(t,\theta) \in [0,T] \times [0, \bar{\lambda}_j];\ \theta \leq \lambda_j(t)\}$ and projection.

Proposition 3.2. *The multivariate counting process $(N^j)_{j\in\mathcal{J}}$, defined for all $j \in \mathcal{J}$ and $t \in [0,T]$ by thinning and projection of $\bar{Q}$:*

$$N_t^j := \int_0^t \int_{\mathbb{R}^+} \mathbf{1}_{\{\theta \leq \lambda_j(s)\}} \bar{Q}(\mathrm{d}s, \{j\}, \mathrm{d}\theta) = \sum_{\ell\geq 1} 1_{\{\bar{T}_\ell \leq t\}} 1_{\{\bar{J}_\ell = j\}} 1_{\{\bar{\Theta}_\ell \leq \lambda_j(\bar{T}_\ell)\}},$$

is an inhomogeneous Poisson process of intensity vector $(\lambda_j(t))_{j\in\mathcal{J}}$ on $[0,T]$.

Proposition 3.2 yields the following simulation Algorithm 1 for multivariate Poisson processes.

Algorithm 1 Thinning algorithm for multivariate inhomogeneous Poisson processes.

Input: Functions $\lambda_j : [0,T] \to [0, \bar{\lambda}]$ and $\bar{\lambda}_j$, $\bar{\Lambda} = \sum_{j\in\mathcal{J}} \bar{\lambda}_j$
Output: Points (T_k, J_k) of Poisson measure N on $[0,T] \times \mathcal{J}$
Initialization $T_0 \longleftarrow 0$, $\bar{T}_0 \longleftarrow 0$
while $T_k < T$ **do**
    repeat
        increment iterative variable $\ell \longleftarrow \ell + 1$
        compute next proposed time $\bar{T}_\ell \longleftarrow \bar{T}_{\ell-1} + S_\ell$ with $S_\ell \sim \mathcal{E}(\bar{\Lambda})$
        draw $\bar{J}_\ell \sim \mathcal{U}\{\bar{\lambda}_j/\bar{\Lambda}, j \in \mathcal{J}\}$ i.e. $\mathbb{P}(\bar{J}_\ell = j) = \bar{\lambda}_j/\bar{\Lambda}$
        draw $\bar{\Theta}_\ell \sim \mathcal{U}([0, \bar{\lambda}_{\bar{J}_\ell}])$
    until accepted event $\bar{\Theta}_\ell \leq \lambda_{\bar{J}_\ell}(\bar{T}_\ell)$
    record $(T_k, J_k) \longleftarrow (\bar{T}_\ell, \bar{J}_\ell)$ as accepted point
end while

Remark 3.1. The acceptance/rejection Algorithm 1 can be efficient when the functions λ_j are of different order, and thus bounded by different $\bar{\lambda}_j$. However, it is important to note that the simulation of the discrete random variables $(\bar{J}_\ell)$ can be costly (compared to a uniform law) when $\mathcal{J}$ is large, for instance when an individual is drawn from a large population. In this case, an alternative is to choose the same bound $\bar{\lambda}_j = \bar{\lambda}$ for all $j \in \mathcal{J}$. Then the marks $(\bar{J}_\ell, \bar{\Theta}_\ell)$ are *i.i.d.* uniform variables on $\mathcal{J} \times [0, \bar{\lambda}]$, faster to simulate.

3.2 Simulation algorithm

Let us now come back to the simulation of the IBM introduced in **?@sec-ibm**. For ease of notations, we assume that there are no event with Poisson intensity ($\mathcal{P} = \emptyset$), so that all events that occur are of type $e \in \mathcal{E} \cup \mathcal{E}_W$, with individual intensity $\lambda_t^e(I, Z_t)$ depending on the population composition Z_t ($e \in \mathcal{E}_W$) or not ($e \in \mathcal{E}$), as defined in **?@eq-indivual-intensity** and verifying either Assumption 2.2 or Assumption 2.3. The global intensity **?@eq-global-event-intensity** at time $t \in [0, T]$ is thus

$$\Lambda_t(Z_t) = \sum_{e\in\mathcal{E}} \Big(\sum_{k=1}^{N_t} \lambda^e(t, I_k) \Big) + \sum_{e\in\mathcal{E}_W} \Big(\sum_{k=1}^{N_t} \sum_{j=1}^{N_t} W^e(t, I_k, I_j) \Big) \leq \bar{\Lambda}(N_t), \tag{11}$$

with $\bar{\Lambda}(n) = \big(\sum_{e\in\mathcal{E}} \bar{\lambda}^e \big) n + \big(\sum_{e\in\mathcal{E}_W} \bar{W}^e \big) n^2$.

One of the main difficulty is that the intensity of events is not deterministic as in the case of inhomogeneous Poisson processes, but a function $\Lambda_t(Z_t)$ of the population state, bounded by a function which also depends on the population size. However, the Algorithm 1 can be adapted to simulate the IBM. The construction is done by induction, by conditioning on the state of the population Z_{T_k} at the kth event time T_k ($T_0 = 0$).

We first present the construction of the first event at time T_1.

3.2.1 First event simulation

Before the first event time, on $\{t < T_1\}$, the population composition is constant : $Z_t = Z_0 = \{I_1, \dots, I_{N_0}\}$. For each type of event e and individual I_k, $k \in \{1, \dots N_0\}$, we denote by $N^{k,e}$ the counting process of intensity $\lambda_t^e(I_k, Z_t)$, counting the occurrences of the events of type e happening to the individual I_k. Then, the first event T_1 is the first jump time of the multivariate counting vector $(N^{(k,e)})_{(k,e)\in\mathcal{J}_0}$, with $\mathcal{J}_0 = \{1, \dots, N_0\} \times \big(\mathcal{E} \cup \mathcal{E}_W\big)$.

Since the population composition is constant before the first event time, each counting process N^j coincides on $[0, T_1[$ with an inhomogeneous Poisson process, of intensity $\lambda_t^e(I_k, Z_0)$. Thus (conditionally to Z_0), T_1 is also the first jump time of an inhomogeneous multivariate Poisson process $N^0 = (N^{0,j})_{j\in\mathcal{J}_0}$ of intensity function $(\lambda_j)_{j\in\mathcal{J}_0}$, defined for all $j = (k, e) \in \mathcal{J}_0$ by:

$$\lambda_j(t) = \lambda_t^e(I_k, Z_0) \leq \bar{\lambda}_0^e \quad \text{with} \quad \bar{\lambda}_0^e = \bar{\lambda}^e \mathbf{1}_{e\in\mathcal{E}} + \bar{W}^e N_0 \mathbf{1}_{e\in\mathcal{E}_W},$$

by Assumption 2.2 and Assumption 2.3. In particular, the jump times of N^0 occur at the intensity

$$\Lambda(t) = \sum_{j\in\mathcal{J}_0} \lambda_j(t) = \sum_{e\in\mathcal{E}\cup\mathcal{E}_W} \sum_{k=1}^{N_0} \lambda_t^e(I_k, Z_0) \leq \bar{\Lambda}(N_0) = N_0 \sum_{e\in\mathcal{E}\cup\mathcal{E}_W} \bar{\lambda}_0^e.$$

By Proposition 3.2, N^0 can be obtained by thinning of the marked Poisson measure $\bar{Q}^0 = \sum_{\ell\geq 1} \delta_{(\bar{T}_\ell, (\bar{K}_\ell, \bar{E}_\ell), \bar{\Theta}_\ell)}$ on $\mathbb{R}^+ \times \mathcal{J}_0 \times \mathbb{R}^+$, with:

- $(\bar{T}_\ell)_{\ell\in\mathbb{N}^*}$ the jump times of a Poisson process of rate $\bar{\Lambda}(N_0)$.
- $(\bar{K}_\ell, \bar{E}_\ell)_{\ell\in\mathbb{N}^*}$ discrete *i.i.d.* random variables on $\mathcal{J}_0 = \{1, \dots, N_0\} \times \big(\mathcal{E} \cup \mathcal{E}_W\big)$, with K_ℓ representing the index of the chosen individual and E_ℓ the event type for the proposed event, such that:

  $$\mathbb{P}(\bar{K}_1 = k, \bar{E}_1 = e) = \frac{\bar{\lambda}_0^e}{\bar{\Lambda}(N_0)} = \frac{1}{N_0} \frac{\bar{\lambda}_0^e N_0}{\bar{\Lambda}(N_0)},$$

  i.e. $(\bar{K}_1, \bar{E}_1)$ are distributed as independent random variables where $\bar{K}_1 \sim \mathcal{U}(\{1, \dots, N_0\})$ and $\bar{E}_1$ such that

  $$p_e := \mathbb{P}(\bar{E}_1 = e) = \frac{\bar{\lambda}_0^e N_0}{\bar{\Lambda}(N_0)}.$$

- $(\bar{\Theta}_\ell)_{\ell\in\mathbb{N}^*}$ are independent uniform random variables, with $\bar{\Theta}_\ell \sim \mathcal{U}([0, \bar{\lambda}^{\bar{E}_\ell}])$.

Since the first event is the first jump of N^0, by Proposition 3.2 and Algorithm 1 , the first event time T_1 is the first jump time $\bar{T}_\ell$ of $\bar{Q}^0$ such that $\bar{\Theta}_\ell \leq \lambda^{\bar{E}_\ell}_{\bar{T}_\ell}(I_{\bar{K}_\ell}, Z_0)$.

At $T_1 = \bar{T}_\ell$, the event $\bar{E}_\ell$ occurs to the individual $I_{\bar{K}_\ell} = (\tau^b, \infty, x)$. For instance, if $\bar{E}_\ell = d$, a death/exit event occurs, so that $Z_{T_1} = Z_0 + \delta_{(\tau^b, T_1, x)} - \delta_{I_{\bar{K}_\ell}}$ and $N_{T_1} = N_0$. If $\bar{E}_\ell = b$ or en, a birth or entry event occurs, so that $N_{T_1} = N_0 + 1$, and a new individual I_{N_0+1} is added to the population, chosen as described in Table 1. Finally, if $\bar{E}_\ell = s$, a swap event occurs, the population size stays constant and $I_{\bar{K}_\ell}$ is replaced by an individual $I'_{\bar{K}_\ell}$, chosen as described in Table 1.

The steps for simulating the first event in the population can be iterated in order to simulate the population. At the kth step, the same procedure is repeated to simulate the kth event, starting from a population $Z_{T_{k-1}}$ of size $N_{T_{k-1}}$.

Algorithm 2 IBM simulation algorithm (without events of Poissonian intensity)

Input: Initial population Z_0, horizon $T > 0$, and events described by:
- Intensity functions and bounds $(\lambda^e, \bar{\lambda}^e)$ for $e \in \mathcal{E}$ and $(W^e, \bar{W}^e)$ for $e \in \mathcal{E}_W$
- Event action functions $\phi^e(t, I)$ for $e \in \mathcal{E} \cup \mathcal{E}_W$ (see @tbl-event-action)
Output: Population Z_T
Initialization $T_0 \longleftarrow 0$, $\bar{T}_0 \longleftarrow 0$
while $T_k < T$ **do**
  repeat
    increment iterative variable $\ell \longleftarrow \ell + 1$
    compute next proposed time $\bar{T}_\ell \longleftarrow \bar{T}_{\ell-1} + \mathcal{E}\big(\bar{\Lambda}(N_{T_k})\big)$
    draw a proposed event $\bar{E}_\ell \sim \mathcal{U}\{p_e\}$ with $p_e = \frac{\bar{\lambda}^e \mathbf{1}_{e\in\mathcal{E}} + \bar{W}^e N_{T_k} \mathbf{1}_{e\in\mathcal{E}_W}}{\sum_{e\in\mathcal{E}} \bar{\lambda}^e + \sum_{e\in\mathcal{E}_W} \bar{W}^e N_{T_k}}$
    draw an individual index $\bar{K}_\ell \sim \mathcal{U}(\{1, \dots, N_{T_k}\})$
    draw $\bar{\Theta}_\ell \sim \mathcal{U}([0, \bar{\lambda}^{\bar{E}_\ell}])$ if $\bar{E}_\ell \in \mathcal{E}$ or $\bar{\Theta}_\ell \sim \mathcal{U}([0, \bar{W}^{\bar{E}_\ell} N_{T_k}])$ if $\bar{E}_\ell \in \mathcal{E}_W$
  until accepted event $\bar{\Theta}_\ell \leq \lambda^{\bar{E}_\ell}_{\bar{T}_\ell}(I_{\bar{K}_\ell}, Z_{T_k})$
  increment iterative variable $k \longleftarrow k + 1$
  record $(T_k, E_k, I_k) \longleftarrow (\bar{T}_\ell, \bar{E}_\ell, I_{\bar{K}_\ell})$ as accepted time, event and individual
  update the population $Z_{T_k} = Z_{T_{k-1}} + \phi^{E_k}(T_k, I_k)$
end while

Theorem 3.1. *Algorithm 2 are exact simulations of solution of the SDE Equation 10.*

The proof of Theorem 3.1 is detailed in the Appendix **?@prf-no-interaction**.

Remark 3.2. The population Z_{T_k} includes dead/exited individuals before the event time T_k. Thus, $N_{T_k} > N^a_{T_k}$ is greater than the number of alive individuals at time T_k. When a dead individual $I_{\bar{K}_l}$ is drawn from the population during the rejection/acceptance phase of the algorithm, the proposed event $(\bar{T}_\ell, \bar{E}_\ell, I_{\bar{K}_\ell})$ is automatically rejected since the event intensity is $\lambda^{\bar{E}_\ell}_{\bar{T}_\ell}(I_{\bar{K}_\ell}, Z_{T_k}) = 0$ (nothing can happen to a dead individual). This can slow down the algorithm, especially when the proportion of dead/exited individuals in the population increases. However, the computational cost of keeping dead/exited individuals in the population is much lower than the cost of removing an individual from the population at each death/exit event, which is linear in the population size.

Actually, dead/exited individuals are regularly removed from the population in the `IBMPopSim` algorithm, in order to optimize the trade-off between having to many dead individuals and removing dead individuals from the population too often. The frequency at which dead individuals are "removed from the population" can be chosen by the user, as an optional argument of the main function `popsim` (see details in Section 3).

Remark 3.3. In practice, the bounds $\bar{\lambda}^e$ and $\bar{W}^e$ should be chosen as sharp as possible. It is easy to see that conditionally to $\{\bar{E}_\ell = e, \bar{T}_\ell = t, \bar{K}_\ell = l\}$ the probability of accepting the event is, depending if there are interactions,

$$\mathbb{P}\big(\bar{\Theta}_\ell \leq \lambda_t^e(I_l, Z_{T_k})|\mathcal{F}_{T_k}\big) = \frac{\lambda^e(t, I_l)}{\bar{\lambda}^e}\mathbf{1}_{e\in\mathcal{E}} + \frac{\sum_{j=1}^{N_{T_k}} W^e(t, I_l, I_j)}{\bar{W}^e N_{T_k}}\mathbf{1}_{e\in\mathcal{E}_W}.$$

The sharper the bounds $\bar{\lambda}^e$ and $\bar{W}^e$ are, the higher is the acceptance rate. For even sharper bounds, an alternative is to define bounds $\bar{\lambda}^e(I_l)$ and $\bar{W}^e(I_l)$ depending on the individuals' characteristics. However, the algorithm is modified and the individual I_l is not chosen uniformly in the population anymore. Due to the population size, this is way more costly than choosing uniform bounds, as explained in Remark 3.1.

3.3 Simulation algorithm with randomization

Let $e \in E_W$ be an event with interactions. In order to evaluate the individual intensity $\lambda_t^e(I, Z_t) = \sum_{j=1}^{N_t} W^e(t, I, I_j)$ one must compute $W^e(t, I_l, I_j)$ for all individuals in the population. This step can be computationally costly, especially for large populations. One way to avoid this summation is to use randomization (see also (Fournier and Méléard 2004) in a model without age). The randomization consists in replacing the summation by an evaluation of the interaction function W^e using an individual J drawn uniformly from the population.

More precisely, if $J \sim \mathcal{U}(\{1, \dots, N_{T_k}\})$ is independent of $\bar{\Theta}_\ell$, we have

$$\mathbb{P}\Big(\bar{\Theta}_\ell \leq \sum_{j=1}^{N_{T_k}} W^e(t, I_l, I_j)|\mathcal{F}_{T_k}\Big) = \mathbb{P}\big(\bar{\Theta}_\ell \leq N_{T_k} W^e(t, I_l, I_J)|\mathcal{F}_{T_k}\big). \tag{12}$$

Equivalently, we can write this probability as $\mathbb{P}\big(\tilde{\Theta}_\ell \leq W^e(t, I_l, I_J)\big)$ where $\tilde{\Theta}_\ell = \frac{\bar{\Theta}_\ell}{N_{T_k}} \sim \mathcal{U}([0, \bar{W}^e])$ is independent of $J \sim \mathcal{U}(\{1, \dots, N_{T_k}\})$.

The efficiency of the randomization procedure increases with the population homogeneity. If the function W^e varies little according to the individuals in the population, the randomization approach is very efficient in practice, especially when the population is large.

We now present the main algorithm implemented in the `popsim` function of the `IBMPopSim` package in the case where events arrive with individual intensities, but also with interactions (using randomization) and Poisson intensities. In this general case, $\bar{\Lambda}(n)$ is defined by Equation 9.

Proposition 3.3.

Appendix

3.4 Recall on Poisson random measures

We recall below some useful properties of Poisson random measures, mainly following Chapter 6 of (Çinlar 2011). We also refer to (Kallenberg 2017) for a more comprehensive presentation of random counting measures.

Definition 3.1 (Poisson Random Measures)**.** Let μ be a σ-finite diffuse measure on a Borel subspace $(E, \mathcal{E})$ of $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. A random counting measure $Q = \sum_{k\geq 1} \delta_{X_k}$ is a Poisson (counting) random measure of *mean measure* μ if

1. $\forall A \in \mathscr{E}$, $Q(A)$ is a Poisson random variable with $\mathbb{E}[Q(A)] = \mu(A)$.
2. For all disjoints subsets $A_1, \dots, A_n \in \mathscr{E}$, $Q(A_1), \dots, Q(A_n)$ are independent Poisson random variables.

Let us briefly recall here some simple but useful operations on Poisson measures. In the following, Q is a Poisson measure of mean measure μ, unless stated otherwise.

Proposition 3.4 (Restricted Poisson measure). *If $B \in \mathscr{E}$, then, the restriction of Q to B defined by*

$$Q^B = \mathbf{1}_B Q = \sum_{k\geq 1} \mathbf{1}_B(X_k)\delta_{X_k}$$

is also a Poisson random measure, of mean measure $\mu^B = \mu(\cdot \cap B)$.

Proposition 3.5 (Projection of Poisson measure). *If $E = F_1 \times F_2$ is a product space, then the projection*

$$Q_1(\mathrm{d}x) = \int_{F_2} Q(\mathrm{d}x, \mathrm{d}y)$$

is a Poisson random measure of mean measure $\mu_1(\mathrm{d}x) = \int_{F_2} \mu(\mathrm{d}x, \mathrm{d}y)$.

3.4.1 Link with Poisson processes

Let $Q = \sum_{k\geq 1} \delta_{T_k}$ a Poisson random measure on $E = \mathbb{R}^+$ with mean measure $\mu(\mathrm{d}t) = \Lambda(t)\mathrm{d}t$ absolutely continuous with respect to the Lebesgue measure, $\mu(A) = \int_A \Lambda(t)\mathrm{d}t$. The counting process $(N_t)_{t\geq 0}$ defined by

$$N_t = Q([0,t]) = \sum_{k\geq 1} \mathbf{1}_{\{T_k \leq t\}}, \quad \forall\, t \geq 0, \tag{13}$$

is an inhomogeneous Poisson process with intensity function (or rate) $t \mapsto \Lambda(t)$. In particular, when $\Lambda(t) \equiv c$ is a constant, N is a homogeneous Poisson process with rate c. Assuming that the atoms are ordered $T_1 < T_2 < \dots$, we recall that the sequence $(T_{k+1} - T_k)_{k\geq 1}$ is a sequence of *i.i.d.* exponential variables of parameter c.

3.4.2 Marked Poisson measures on $E = \mathbb{R}^+ \times F$

We are interested in the particular case when E is the product space $\mathbb{R}^+ \times F$, with $(F, \mathscr{F})$ a Borel subspace of $\mathbb{R}^d$. Then, a random counting measure is defined by a random set $S = \{(T_k, \Theta_k), k \geq 1\}$. The random variables $T_k \geq 0$ can be considered as time variables, and constitute the jump times of the random measure, while the variables $\Theta_k \in F$ represent space variables.

We recall in this special case the Theorem VI.3.2 in (Çinlar 2011).

Proposition 3.6 (Marked Poisson measure). *Let m be a σ–finite diffuse measure on $\mathbb{R}^+$, and K a transition probability kernel from $(\mathbb{R}^+, \mathscr{B}(\mathbb{R}^+))$ into $(F, \mathscr{F})$. Assume that the collection $(T_k)_{k\geq 1}$ forms a Poisson process $(N_t) = (\sum_{k\geq 1} 1_{\{T_k \leq t\}})$ with mean $m(\mathrm{d}t) = \Lambda(t)\mathrm{d}t$, and that given $(T_k)_{k\geq 1}$, the variables Θ_k are conditionally independent and have the respective distributions $K(T_k, \cdot)$.*

1. *Then, $\{(T_k, \Theta_k);\ k \geq 1\}$ forms a Poisson random measure $Q = \sum_{k\geq 1} \delta_{(T_k, \Theta_k)}$ on $(\mathbb{R}^+ \times F, \mathscr{B}(\mathbb{R}^+) \otimes \mathscr{F})$, called a* Marked point process *, with mean μ defined by*

$$\mu(\mathrm{d}t, \mathrm{d}y) = \Lambda(t)\mathrm{d}tK(t, \mathrm{d}y).$$

2. *Reciprocally let Q be a Poisson random measure of mean measure $\mu(\mathrm{d}t, \mathrm{d}y)$, admitting the following disintegration with respect to the first coordinate: $\mu(\mathrm{d}t, \mathrm{d}y) = \tilde{\Lambda}(t)\mathrm{d}t\nu(t, \mathrm{d}y)$, with $\nu(t, F) < \infty$. Let $K(t, \mathrm{d}y) = \dfrac{\nu(t, \mathrm{d}y)}{\nu(t, F)}$ and $\Lambda(t) = \nu(t, F)\tilde{\Lambda}(t)$. Then, $Q = \sum_{k\geq 1} \delta_{(T_k, \Theta_k)}$ is a marked Poisson*

measure with $(T_k, \Theta_k)_{k\in\mathbb{N}^*}$ defined as above. In particular, the projection $N = (N_t)_{t\geq 0}$ of the Poisson measure on the first coordinate,

$$N_t = Q([0,t]\times F) = \sum_{k\geq 1} \mathbf{1}_{[0,t]\times F}(T_k, \Theta_k) = \sum_{k\geq 1} \mathbf{1}_{\{T_k\leq t\}}, \quad \forall\, t\geq 0,$$

is an inhomogeneous Poisson process of rate $\Lambda(t) = \nu(t,F)\tilde{\Lambda}(t)$.

When the transition probability kernel K does not depend on the time: $K(t,A) = \nu(A)$ for some probability measure ν, then the marks $(\Theta_k)_{k\geq 1}$ form an *i.i.d.* sequence with distribution ν, independent of $(T_k)_{k\geq 1}$.

The preceding proposition thus yields a straight forward iterative simulation procedure for a Marked Poisson process on $[0,T]\times F$ with mean measure $\mu(\mathrm{d}t,\mathrm{d}y) = c\mathrm{d}tK(t,\mathrm{d}y)$ and $c>0$. The procedure is described in Algorithm **??** .

3.5 Pathwise representation of IBMs

3.5.1 Notation reminder

The population's evolution is described by the measure valued process $(Z_t)_{t\geq 0}$. Several types of events e can occur to individuals denoted by I. In an event of type e occur to the individual I at time t, then the population state Z_{t^-} is modified by $\phi^e(t,I)$. If $e\in\mathscr{E}\cup\mathscr{E}_W$, then events of type e occur with an intensity $\sum_{k=1}^{N_t}\lambda_t^e(I,Z_t)$, with $\lambda_t^e(I,Z_t)$ defined by **?@eq-individual-intensity**. If $e\in\mathscr{P}$, then events of type e occur in the population at a Poisson intensity of (μ_t^e).

3.5.2 Proof of Theorem 2.1

For ease of notation, we prove the case when $\mathscr{P}=\emptyset$ (there are no events with Poisson intensity).

Step 1 The existence of a solution to Equation 10 is obtained by induction. Let Z^1 be the unique solution the thinning equation:

$$Z_t^1 = Z_0 + \int_0^t\int_{\mathcal{J}\times\mathbb{R}^+} \phi^e(s,I_k)\mathbf{1}_{\{k\leq N_0\}}\mathbf{1}_{\{\theta\leq\lambda_s^e(I_k,Z_0)\}}Q(\mathrm{d}s,\mathrm{d}k,\mathrm{d}e,\mathrm{d}\theta), \quad \forall 0\leq t\leq T.$$

Let T_1 be the first jump time of Z^1. Since $Z_{s^-}^1 = Z_0$ and $N_{s^-} = N_0$ on $[0,T_1]$, Z^1 is solution of Equation 10 on $[0,T_1]$.

Let us now assume that Equation 10 admits a solution Z^n on $[0,T_n]$, with T_n the n–th event time in the population. Let Z^{n+1} be the unique solution of the thinning equation:

$$Z_t^{n+1} = Z_{t\wedge T_n}^n + \int_{t\wedge T_n}^t\int_{\mathcal{J}\times\mathbb{R}^+} \phi^e(s,I_k)\mathbf{1}_{\{\theta\leq\lambda_s^e(I_k,Z_{T_n}^n)\}}\mathbf{1}_{\{k\leq N_{T_n}^n\}}Q(\mathrm{d}s,\mathrm{d}k,\mathrm{d}e,\mathrm{d}\theta).$$

First, observe that Z^{n+1} coincides with Z^n on $[0,T_n]$. Let T_{n+1} be the $(n+1)$–th jump of Z^{n+1}. Furthermore, $Z_{s^-}^{n+1} = Z_{T_n}^n$ and $N_{s^-}^{n+1} = N_{T_n}^n$ on $[T_n, T_{n+1}]$ (nothing happens between two successive event times), Z^{n+1} verifies for all $t\leq T_{n+1}$:

$$Z_t^{n+1} = Z_{t\wedge T_n}^n + \int_{t\wedge T_n}^t\int_{\mathcal{J}\times\mathbb{R}^+} \phi^e(s,I_k)\mathbf{1}_{\{\theta\leq\lambda_s^e(I_k,Z_{s^-}^{n+1})\}}\mathbf{1}_{\{k\leq N_{s^-}^{n+1}\}}Q(\mathrm{d}s,\mathrm{d}k,\mathrm{d}e,\mathrm{d}\theta).$$

Since, Z^n is a solution of Equation 10 on $[0,T_n]$ coinciding with Z^{n+1}, this achieves to prove that Z^{n+1} is solution of Equation 10 on $[0,T_{n+1}]$. Finally, let $Z=\lim_{n\to\infty}Z^n$. For all $n\geq 1$, T_n is the n–th event time of Z, and Z is solution of Equation 10 on all time intervals $[0,T_n\wedge T]$ by construction.

By Lemma 2.1 $T_n \underset{n\to\infty}{\longrightarrow} \infty$. Thus, by letting $n \to \infty$ we can conclude that Z is a solution of Equation 10 on $[0, T]$.

Step 2 Let $\tilde{Z}$ be a solution of Equation 10. Using the same arguments than in Step 1, it is straight forward to show that $\tilde{Z}$ coincides with Z^n on $[0, T_n]$, for all $n \geq 1$. Thus, $\tilde{Z} = Z$, with achieves to prove uniqueness.

References

Bansaye, Vincent, and Sylvie Méléard. 2015. *Stochastic Models for Structured Populations*. Springer International Publishing.

Barrieu, Pauline, Harry Bensusan, Nicole El Karoui, Caroline Hillairet, Stéphane Loisel, Claudia Ravanelli, and Yahia Salhi. 2012. "Understanding, Modelling and Managing Longevity Risk: Key Issues and Main Challenges." *Scandinavian Actuarial Journal* 2012 (3): 203–31.

Bensusan, Harry. 2010. "Interest rate and longevity risk: dynamic model and applications to derivative products and life insurance." Theses, Ecole Polytechnique X.

Billiard, Sylvain, Pierre Collet, Régis Ferrière, Sylvie Méléard, and Viet Chi Tran. 2016. "The Effect of Competition and Horizontal Trait Inheritance on Invasion, Fixation, and Polymorphism." *Journal of Theoretical Biology* 411: 48–58.

Boumezoued, Alexandre. 2016. "Micro-macro analysis of heterogenous age-structured populations dynamics.Application to self-exciting processes and demography." Theses, Université Pierre et Marie Curie.

Brémaud, Pierre. 1981. *Point Processes and Queues: Martingale Dynamics*. Vol. 66. Springer.

Calvez, Vincent, Susely Figueroa Iglesias, Hélène Hivert, Sylvie Méléard, Anna Melnykova, and Samuel Nordmann. 2020. "Horizontal Gene Transfer: Numerical Comparison Between Stochastic and Deterministic Approaches." *ESAIM: Proceedings and Surveys* 67: 135–60.

Champagnat, Nicolas, Régis Ferrière, and Sylvie Méléard. 2006. "Unifying Evolutionary Dynamics: From Individual Stochastic Processes to Macroscopic Models." *Theoretical Population Biology* 69 (3): 297–321.

Çinlar, Erhan. 2011. *Probability and Stochastics*. Springer New York.

Collet, Pierre, Sylvie Méléard, and Johan AJ Metz. 2013. "A Rigorous Model Study of the Adaptive Dynamics of Mendelian Diploids." *Journal of Mathematical Biology* 67: 569–607.

Costa, Manon, Céline Hauzy, Nicolas Loeuille, and Sylvie Méléard. 2016. "Stochastic Eco-Evolutionary Model of a Prey-Predator Community." *Journal of Mathematical Biology* 72: 573–622.

Devroye, Luc. 1986. *Nonuniform Random Variate Generation*. Springer-Verlag, New York.

Eddelbuettel, Dirk, and Romain Francois. 2011. "Rcpp: Seamless r and c++ Integration." *Journal of Statistical Software* 40 (8): 1–18. https://doi.org/10.18637/jss.v040.i08.

El Karoui, Nicole, Kaouther Hadji, and Sarah Kaakai. 2021. "Simulating Long-Term Impacts of Mortality Shocks: Learning from the Cholera Pandemic." *arXiv Preprint arXiv:2111.08338*.

Ferrière, Régis, and Viet Chi Tran. 2009. "Stochastic and Deterministic Models for Age-Structured Populations with Genetically Variable Traits." In, 27:289–310. ESAIM Proc. EDP Sci., Les Ulis.

Fournier, Nicolas, and Sylvie Méléard. 2004. "A Microscopic Probabilistic Description of a Locally Regulated Population and Macroscopic Approximations." *Ann. Appl. Probab.* 14 (4): 1880–1919.

Hyndman, Rob, Heather Booth Booth, Leonie Tickle Tickle, John Maindonald, Simon Wood Wood, and R Core Team. 2023. *demography: Forecasting Mortality, Fertility, Migration and Population Data*. https://cran.r-project.org/package=demography.

Kaakai, Sarah, and Nicole El Karoui. 2023. "Birth Death Swap Population in Random Environment and Aggregation with Two Timescales." *Stochastic Processes and Their Applications* 162: 218–48. https://doi.org/https://doi.org/10.1016/j.spa.2023.04.017.

Kallenberg, Olav. 2017. *Random Measures, Theory and Applications*. Vol. 77. Probability Theory and

Stochastic Modelling. Springer, Cham.

Lavallée, François, Charline Smadi, Isabelle Alvarez, Björn Reineking, François-Marie Martin, Fanny Dommanget, and Sophie Martin. 2019. "A Stochastic Individual-Based Model for the Growth of a Stand of Japanese Knotweed Including Mowing as a Management Technique." *Ecological Modelling* 413: 108828.

Lewis, Peter, and Gerald Shedler. 1979. "Simulation of Nonhomogeneous Poisson Processes by Thinning." *Naval Research Logistics Quarterly* 26 (3): 403–13.

Méléard, Sylvie, Michael Rera, and Tristan Roget. 2019. "A Birth–Death Model of Ageing: From Individual-Based Dynamics to Evolutive Differential Inclusions." *Journal of Mathematical Biology* 79: 901–39.

Roget, Tristan, Pierre Jolivet, Sylvie Méléard, and Michael Rera. 2022. "Positive Selection of Senescence Through Increased Evolvability: Ageing Is Not a by-Product of Evolution." *bioRxiv*, 2022–03.

Tran, Viet Chi. 2008. "Large Population Limit and Time Behaviour of a Stochastic Particle Model Describing an Age-Structured Population." *ESAIM: Probability and Statistics* 12: 345–86. https://doi.org/10.1051/ps:2007052.

Villegas, Andres, Pietro Millossovich, and Vladimir Kaishev Hyndman. 2018. *StMoMo: Stochastic Mortality Modelling*. https://cran.r-project.org/package=StMoMo.

Zinn, Sabine. 2014. "The MicSim package of R: an entry-level toolkit for continuous-time microsimulation." *International Journal of Microsimulation* 7 (3): 3–32.

Zinn, Sabine, Jutta Gampe, Jan Himmelspach, and Adelinde M Uhrmacher. 2009. "MIC-CORE: A Tool for Microsimulation." In *Proceedings of the 2009 Winter Simulation Conference (WSC)*, 992–1002. IEEE.

Session information

```
sessionInfo()
```