7

Piecewise-deterministic Markov Processes for Spatio-temporal Population Dynamics

7.1. Introduction

7.1.1. *Models for Population Dynamics*

Population dynamics is the study of the structure, the pattern and the biological and environmental drivers of populations. Studies of population dynamics are carried out at various scales, from the microscopic scale to the global scale, and are particularly relevant in ecology and epidemiology.

Numerous and diverse modeling approaches have been proposed to mathematically represent population dynamics. These modeling approaches are based on diverse mathematical tools adapted to (i) different resolutions at which the population dynamics are considered (e.g. individuals, groups, presence in quadrats, and numbers of individuals in districts), and (ii) different levels of perceptions (e.g. the population itself, its averaged characteristics, or more generally aggregated functions of the population patterns). For instance, ODEs were used to describe the average growth of populations [TUR 03, chap.3], branching processes were used to model the growth and adaptation of populations [MÉL 11], PDEs and integrodifferential equations were used to represent the spatio-temporal intensity of populations with local and non-local dispersal capacities [ROQ 10, ALF 13], SDEs were used to model trajectories of individuals [GLO 15], temporal point processes were used to build birth-death models [CHA 06], spatio-temporal point processes were used to model the temporal evolution of the spatial pattern of individuals forming a population [SOU 11], stochastic Markovian areal processes were used to model large-scale dynamics [SOU 09b], regressions (eventually including

---

Chapter written by Candy ABBOUD, Rachid SENOUSSI and Samuel SOUBEYRAND.

auto-regressive components) were used to take into account the effect of environmental variables on population characteristics [BOR 17].

Suppose that we are interested in fitting a spatio-temporal population dynamic model to data. There is, like in many other application fields, a trade-off between model realism and estimation complexity. For example, fitting a population dynamic model essentially constructed from a partial differential equation containing a few parameters [SOU 14b] is generally easier than fitting a (more flexible and realistic) hierarchical stochastic spatio-temporal Markovian model including a few parameters but numerous latent variables [SOU 09b]. In this example, two extreme cases are considered: (i) a model with a deterministic behavior and a few degrees of freedom, which may yield poor goodness-of-fit, and (ii) a model with a stochastic behavior and lots of degrees of freedom, which may induce identifiability issues. Intermediate models are required to achieve rapid, realistic and consistent inferences. Spatio-temporal PDMPs can play this role.

7.1.2. *Spatio-temporal PDMP for Population Dynamics*

Spatio-temporal PDMPs can be occasionally encountered in the theoretical and quantitative population dynamic literature, but these models are generally not called PDMPs. Here, we give three examples of spatio-temporal PDMPs built at three different levels: the population, the metapopulation (which is a set of populations) and the individual. These processes are illustrated in Figure 8.1.

The coalescing colony model [SHI 95], which was developed to represent stratified diffusion in biological invasions, is a PDMP. Stratified diffusion typically consists of two components: neighborhood diffusion and long-distance dispersal. The former component is modeled in the coalescing colony model by a deterministic expansion of colonies (this is the *flow*); The latter component is modeled by the random Markovian generation of new colonies away from the existing colonies (this is the *jump* process). The coalescing colony model was developed to investigate the impact of stratified dispersal on the rate of expansion of populations with several propagation modes.

The metapopulation epidemic model proposed in [SOU 09a] is another example of spatio-temporal PDMP representing a population dynamic. Here, the population of interest is a pathogen population whose hosts are spread in a set of disconnected areas, called host patches. In this model, host patches can be either healthy or infected by the pathogen; When a host patch is infected, the local pathogen population grows in a deterministic way (this is the *flow*); Infected patches can infect distant healthy patches in a stochastic manner (this is the *jump* process; the pathogen *jumps* from infected patches to healthy patches). The metapopulation epidemic model was fitted to presence/absence data of the pathogen in host patches at the end of successive epidemic seasons.

PDMPs can also provide concise mathematical descriptions of trajectories of individuals. Examples of such models are given in [CAI 17, chap. 1] under the term velocity-jump models. These models were used to carry out a statistical analysis of the expansion of the cane toad using data obtained by monitoring successive daily locations of a sample of toads. In the simplest model, each individual randomly alternates between encamped and running modes, whose durations are independently and exponentially distributed (this is the *jump* process). When an individual *jumps* towards a new running mode, the direction is randomly drawn in a specified distribution. When an individual is running, its movement is deterministic and linear given the random direction of the movement (this is the *flow*).

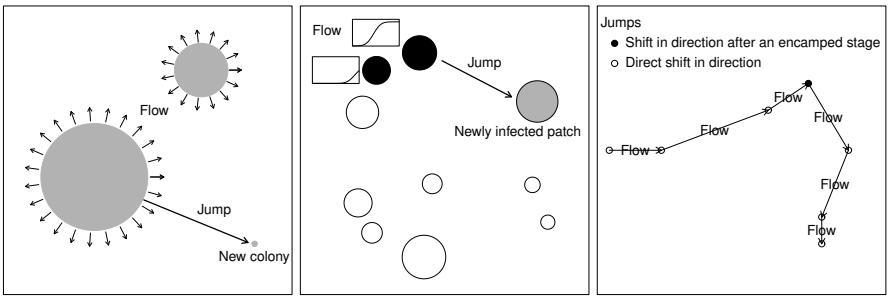

Figure 7.1: Illustrations of the flows and jumps for the coalescing colony model (left), the metapopulation epidemic model (center) and the simple velocity-jump model.

7.1.3. *Chapter Contents*

In the following, we describe three contexts where PDMPs arise for describing population dynamics at the population level, the metapopulation level and the individual level, respectively. Section 8.2 shows how the coalescing colony model was built and how it can be formulated as a PDMP. It also introduces a spatio-temporal PDMP based on a reaction-diffusion equation that could be used to model the dynamic of an invading pathogen (e.g. *Xylella fastidiosa* in Corsica) that might have been introduced at multiple points in space and time. Section 8.3 presents the metapopulation epidemic model mentioned above and gives details about how it was fitted to data. Section 8.4 describes a theoretical framework for building trajectory models with jumps, including PDMPs.

7.2. Stratified Dispersal Models

In this section, we briefly review some mathematical models describing spatio-temporal dynamics of populations. We are especially interested in the

dispersal modes incorporated in these models. Thus, we will consider some reaction-diffusion models including only short-distance dispersal processes, and coalescing colony models including both short-distance and long-distance dispersal processes. We will show how the latter model can be formulated as a PDMP. Finally we will present an original reaction-diffusion-based PDMP describing invasion dynamics with multiple introductions.

7.2.1. *Reaction-diffusion Equations for Modeling Short-distance Dispersal*

There are typically three stages arising successively during a biological invasion process: (1) establishment where a few individuals arrive and succeed to settle, (2) linear expansion when the invasion occurs by neighborhood diffusion as in this section or biphasic expansion when the invasion is driven by stratified diffusion (see Section 8.2.2), and (3) concentration of the invasive species in the area of invasion until saturation [COL 04, RIC 00]. When one aims to model dispersal phenomena such as biological invasions, reaction-diffusion equations are frequently used and have been exploited in many domains, especially in medicine, ecology and epidemiology [GAT 96, ROQ 13, MUR 96]. Reaction-diffusion equations are partial differential equations of parabolic type [EVA 98]. Here, we describe some reaction-diffusion equations, in which dispersal is considered as a random diffusion process.

Random diffusion at the population level can be derived from random walks at the individual level. Random walks are often used to describe invasions by species that move via short-distance dispersal. Basic random walk models describe the path of an individual moving in a spatial domain via a succession of random steps. Typically, in a unidimensional space, as illustrated in Figure 8.2, the individual located at x can move to the left and reach $x-d$ with probability $\mathbb{P}_L$, move to the right and reach $x+d$ with probability $\mathbb{P}_R$ or stay at the same place with probability $\mathbb{P}_S = 1 - \mathbb{P}_L - \mathbb{P}_R$. Such a microscopic and individual-based description of movements can be used to obtain diffusion equations at the population level [ROQ 13, SHI 97, SKE 51]. In particular, the 1D random walk without directional bias and with constant and non-persistant increments leads to the following form of diffusion equation: $\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}$, where u is the density of population.

In 1937, Fisher analyzed the rate of advance of advantageous genes with a PDE [FIS 37], which has been generalized into:

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}u + \underbrace{u(r-bu)}_{f(u)}, \quad t \geq 0 \qquad [7.1]$$

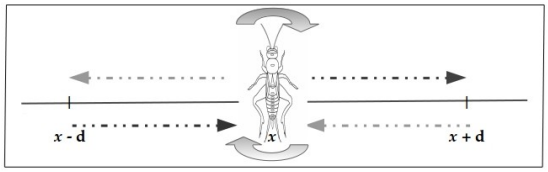

Figure 7.2: Unidimensional random walk model.

where $u = u(t, x)$ is the frequency of the advantageous gene at time t and spatial location x in a unidimensional space; $D > 0$ is the coefficient measuring the rate of dispersal; r stands for the intrinsic growth rate of the species; and b corresponds to the coefficient measuring the effect of intra-specific competition; $f(u)$ is the population growth term.

In the line with Fisher's work, Skellam [SKE 51] proposed two-dimensional PDEs for describing population dynamics. The so-called Skellam model, in particular, allowed him to theoretically study population spread with Malthusian growth. This model incorporates two terms, namely the population dispersal term and the population growth term, and assumes that there is no intra-specific competition:

$$\frac{\partial u}{\partial t} = D\Delta u + ur, \quad t \geq 0 \qquad [7.2]$$

Figure 8.3 presents the solution of Equation [8.2] in a two-dimensional space, for specific values of parameters, initial conditions and boundary conditions.

Positive wavefront type-solutions exist for Equation [8.2]. One simplified form of a traveling wave (in a unidimensional space) is a function of the form:

$$u(t, x) = U(x - ct)$$

where $c \in \mathbb{R}$ is the speed of the front $U \in \mathcal{C}^2(\mathbb{R})$. Note that a traveling wavefront can be defined not only when $t > 0$ but also for any $t \in \mathbb{R}$.

Skellam showed that the rate of spread at the front of the population range asymptotically approaches $c_0 = 2\sqrt{rD}$ when a small population is initially introduced at the origin. Furthermore, Luther [LUT 06] and Kolmogorov et al. [KOL 37] were the first to prove the existence of wavefront type-solutions for a diffusion equation with a logistic growth term $f(u) = ru(1 - u)$ (Fisher-KPP). Kolmogorov et al. showed that some initial distributions converge asymptotically to a traveling wave propagating to the right with a well defined, constant speed $c = 2\sqrt{rD}$. When the growth term includes an Allee effect as follows: $f(u) = ru(1 - u)(u - \theta)$, where $\theta \in]0, 1[$ is the Allee effect parameter, then there

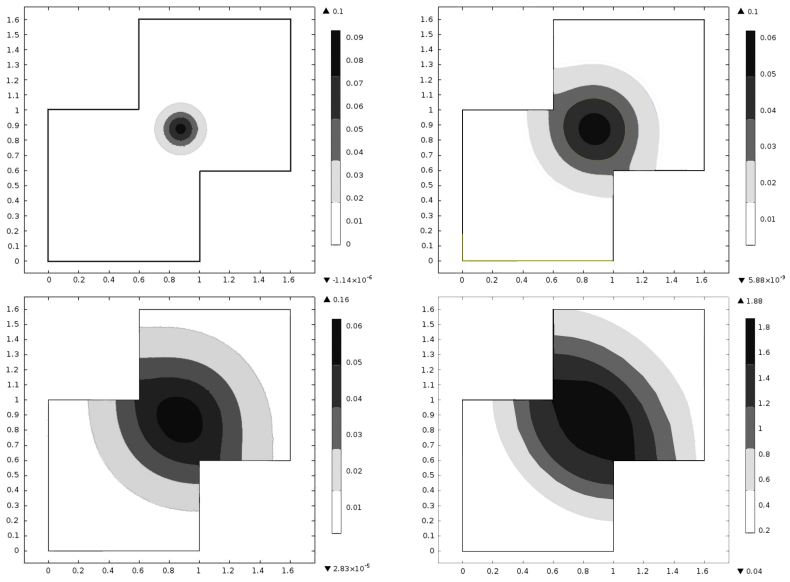

Figure 7.3: Numerical solution $u(t, \mathbf{x})$ of Skellam model [8.2] in a bi-dimensional space (where $\mathbf{x} = (x, y)$) with Neumann boundary conditions, at time 0 (top left), 3 (top right), 6 (bottom left) and 12 (bottom right). The dispersal coefficient and the intrinsic growth rate were fixed at $(D, r) = (5 \times 10^{-3}, 0.5)$. The initial condition was $u(0, \mathbf{x}) = 0.1 \exp(-(10\|\mathbf{x} - \tilde{\mathbf{x}}_0\|)^2)$, where $\tilde{\mathbf{x}}_0 = (\tilde{x}_0, \tilde{y}_0) = (0.8, 0.8)$.

exists a unique positive wavefront-type solution with $\lim_{x \longrightarrow -\infty} U = 1$, $\lim_{x \longrightarrow +\infty} U = 0$. In addition, the speed of the front is [HAD 75, ROT 81, LEW 93]:

$$c = \sqrt{2rD}(\frac{1}{2} - \theta) \qquad [7.3]$$

7.2.2. *Stratified Diffusion*

The models introduced above are generally not adapted to describe the dynamics of populations that expand their range not only by neighborhood dispersal but also by long-distance dispersal, which can corresponds to rare but significant events. The term *stratified diffusion* was used to describe this twofold dispersal process [HEN 89].

Shigesada et al. [SHI 95] proposed stratified diffusion models (derived from Skellam's equation for neighborhood dispersal) and studied their properties. In theses models, the population of interest is in a homogeneous environment and expands its range continuously in time for neighborhood dispersal and at discrete random times

for long-distance dispersal (i.e. colonization events). Two frameworks were considered: (i) the nuclei of colonization created by long-distance migrants are located far enough to assume that their ranges do not overlap, mutually and with the mother colony, for a long time; (ii) the nuclei of colonization created by long-distance migrants merge with the mother colony as soon as they *touch* the mother colony (because of their own expansion and the expansion of the mother colony), but the merging of two nuclei of colonization is neglected. Framework (ii) led to the *coalescing colony model* [SHI 95] that we revisit in the next section by incorporating an Allee effect.

7.2.3. *Coalescing Colony Model with Allee effect*

Model description and properties

Suppose that a few individuals invade a given location of the 2D Euclidean space at $t = 0$, succeed to settle, and form a so-called *mother colony* with a disk shape whose radius increases at a constant rate c by neighborhood diffusion (the establishment phase is neglected). By setting $c = \sqrt{2rD}(\frac{1}{2} - \theta)$, the expansion of the mother colony is an approximation of the population expansion governed by the following PDE incorporating an Allee effect (see Equation [8.3]):

$$\frac{\partial u}{\partial t} = D\Delta u + ru(1-u)(u-\theta),$$

given adequate initial conditions.

The expansion of the mother colony is augmented by long-distance dispersal events generating child colonies. More precisely, the mother colony releases long-distance dispersers that settle at a distance $L > 0$ of the border of the mother colony and produce child colonies. The rate of generation of child colonies, say $\tilde{\lambda}$, is assumed to depend on the current radius z of the mother colony. Typically, $\tilde{\lambda}$ is a non-decreasing function of z. Shigesada et al. considered three cases:

- $\tilde{\lambda}(z) = \lambda_0$, i.e. the mother colony produces long-distance migrants at a constant rate;
- $\tilde{\lambda}(z) = \lambda_1 z$, i.e. the mother colony produces long-distance migrants at a time-varying rate proportional to its perimeter;
- $\tilde{\lambda}(z) = \lambda_2 z^2$, i.e. the mother colony produces long-distance migrants at a time-varying rate proportional to its area.

Additionally, every child colony expands its range circularly at the constant rate c, like the mother colony, but do not release long-distance migrants. When the mother colony and a child colony collide, the area covered by the child colony is instantaneously assigned to the mother colony, which remains a disk with same center but with a

larger radius. Collisions between child colonies are neglected. An illustration of this process is provided in Figure 8.4.

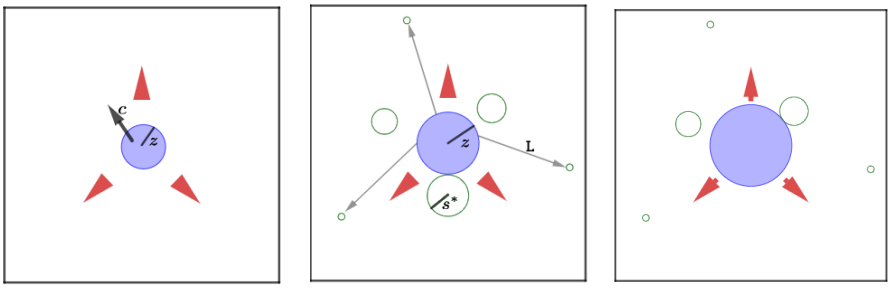

Figure 7.4: Illustration for the coalescing colony model. First, from $t = 0$, the range of the mother colony (disks) expands by short-distance dispersal with a constant rate c (left). Then, the mother colony generates long-distance dispersers to the distance L from its border at the rate $\tilde{\lambda}(z(t))$. The child colonies (circles) expands their range at the rate c until they collides with the mother colony after a period of duration $\frac{L}{2c}$. Finally (right), at the time of coalescence, the range of the blue including the green colony is immediately reshaped into a circular pattern while the total area of both colonies remains the same.

The coalescing colony model is characterized by the following properties [SHI 95]. The expectation of the number of child colonies having radius s at time t, say $n(s,t)$, satisfies the following von Foerster equation and initial / boundary conditions:

$$\begin{cases} \frac{\partial n}{\partial t}(s,t) + c\frac{\partial n}{\partial s}(s,t) = 0 \quad \text{for } s \in (0, s^*(t)) \\ n(s,0) = 0 \\ cn(0,t) = \tilde{\lambda}(z(t)), \end{cases} \qquad [7.4]$$

where $z(t)$ is the radius of the mother colony at time t and $s^*(t)$ is the radius of the first child colony coalescing with the mother colony immediately before the collision. Equation [8.4] has an explicit solution:

$$n(s,t) = \frac{1}{c}\tilde{\lambda}\Big(z\big(t - \frac{s}{c}\big)\Big)\mathbb{1}_{\{ct \geq s > 0\}}(s,t).$$

The area $\pi z(t)^2$ of the mother colony satisfies, before and after collision with a child colony:

$$\frac{d}{dt}\pi z^2 = \begin{cases} 2\pi z c & \text{for } t \in (0, t_1) \\ 2\pi z c + \pi {s^*}^2 n(s^*, t)(c - \frac{ds^*}{dt}) & \text{for } t \geq t_1, \end{cases}$$

where $t_1 = \frac{L}{2c}$ is the time when the first mother-child collision occurs.

Finally, $z(t)$ and $s^*(t)$ are linked by the following equation when $t \geq t_1$:

$$L = z(t) - z\Big(t - \frac{s^*(t)}{c}\Big) + s^*(t),$$

where $t - \frac{s^*(t)}{c}$ is the time when the collided child colony was at a distance L of the mother colony (for further details see Shigesada et al. [SHI 95]).

PDMP Formulation of the Coalescing Colony Model with Allee Effect:

The coalescing colony model can be seen as a precursory example of PDMPs modeling spatio-temporal population dynamics. In this case, the PDMP is the Boolean process formed by the union of the mother and child colonies:

$$X_t = \mathcal{B}(O, z(t)) \cup \Big(\bigcup_{i=1}^{m(t)} A_i(t) \Big)$$

$$A_i(t) = \begin{cases} \mathcal{B}(O_i, s_i(t)) & \text{if } d(O, O_i) > z(t) + s_i(t) \\ \emptyset & \text{otherwise,} \end{cases}$$

where $\mathcal{B}(O, z(t))$ is the ball with center O and radius $z(t)$ covered by the mother colony, $m(t)$ is the number of child colonies generated until time t, and $\mathcal{B}(O_i, s_i(t))$ is the ball with center O_i and radius $s_i(t)$ covered by child colony i until its collision with the mother colony, that is to say while $z(t) + s_i(t) < d(O, O_i)$, and $d(\cdot, \cdot)$ is the inter-point Euclidean distance. Between collision times (thereafter called *jump times*), the radii of the mother and child colonies grow at the constant speed c given by Equation [8.3]. We remind, in addition, that the coalescence of two child colonies and the generation of grandchild colonies by child colonies (i.e. secondary colonizations) are neglected.

Let T_j be the j-th jump time corresponding to the time of generation of child colony j. Let τ_j be the time of collision between the mother colony and child colony j. Over $[T_j, T_{j+1})$, $m(t) = j$, eventual collisions following the expansion of colonies occur in a deterministic way and describing the dynamic of X_t is equivalent to describing the dynamics of the radii $z(t)$ and $s_i(t)$, $i = 1, \ldots, j$, because the centers O and O_i are fixed. For $t \in [0, T_1)$,

$$z(t) = ct$$

and for $t \in [T_j, T_{j+1})$, $j \geq 1$, the radii of the mother and child colonies satisfy:

$$\begin{aligned} z(t) =& z(T_j) + c(t - T_j) \\ &+ \sum_{i=1}^{j} \left[\left(s_i(\tau_i^-)^2 + z(\tau_i^-)^2 \right)^{1/2} - z(\tau_i^-) \right] \mathbb{1}(t \geq \tau_i > T_j) \\ s_i(t) =& \{ s_i(T_j) + c(t - T_j) \} \mathbb{1}(t < \tau_i), \quad \forall i = 1, \cdots, j, \end{aligned}$$

where

$$\begin{aligned} s_i(\tau_i^-) =& s_i(T_j) + c(\tau_i - T_j) \\ z(\tau_i^-) =& z(\max\{\tau_{i-1}, T_j\}) + c(\tau_i - \max\{\tau_{i-1}, T_j\}) \end{aligned}$$

with the conventions $\tau_0 = 0$ and $s_i(t) = 0$ when child colony i has merged with the mother colony. We now give the expression of τ_i for i such that $T_j < \tau_i < T_{j+1}$. Let $t_0 = \max\{\tau_{i-1}, T_j\}$ be the time of the event (i.e. a collision or the generation of a child colony) preceeding τ_i. If a collision occured at t_0 and if the resulting instantaneous growth of the mother colony led the mother colony to touch or overlap colony i, then $\tau_i = t_0$ (i.e. multiple instantaneous collisions occur). Otherwise, τ_i satisfies the following equation:

$$d(O, O_i) = L + z(T_i) = z(t_0) + c(\tau_i - t_0) + s_i(t_0) + c(\tau_i - t_0),$$

whose solution is:

$$\tau_i = t_0 + \frac{d(O, O_i) - z(t_0) - s_i(t_0)}{2c}. \qquad [7.5]$$

In the case of instantaneous collisions, the fraction in Equation [8.5] is non-positive (since the sum of radii $z(t_0) + s_i(t_0)$ is larger than or equal to $d(O, O_i)$). Thus, whatever the event at t_0,

$$\begin{cases} \tau_i = t_0 + \max\left\{0, \dfrac{d(O, O_i) - z(t_0) - s_i(t_0)}{2c}\right\} \\ t_0 = \max\{\tau_{i-1}, T_j\}. \end{cases}$$

Therefore, τ_i can be recursively defined as a function of radii and center locations at time T_j, which are functions of X_{T_j}

To demonstrate that X_t can be viewed as a PDMP, we will now give the expression of the *flow function* Φ, the *jump rate* λ and the *jump kernel* Q. Let

$$\mathbf{x} = \mathcal{B}(O_{\mathbf{x}}, z_{\mathbf{x}}) \cup \left(\bigcup_{k=1}^{K_{\mathbf{x}}} \mathcal{B}(O_{\mathbf{x}k}, s_{\mathbf{x}k}) \right)$$

be in the set $\mathcal{X}$ of unions of disjoint balls included in $\mathbb{R}^2$ and suppose that k is ordered such as the sequence of $d(O_{\mathbf{x}}, O_{\mathbf{x}k})$ increases with k. Note that knowing $\mathbf{x}$ is equivalent to knowing $\{O_{\mathbf{x}}, z_{\mathbf{x}}, O_{\mathbf{x}k}, s_{\mathbf{x}k}; k = 1, \ldots, K_{\mathbf{x}}\}$. Define Φ over $\mathcal{X} \times \mathbb{R}_+$ as follows:

$$\Phi(\mathbf{x}, t) = \mathcal{B}(O_{\mathbf{x}}, \phi_1(\mathbf{x}, t)) \cup \left(\bigcup_{k=1}^{K_{\mathbf{x}}} \mathcal{B}(O_{\mathbf{x}k}, \phi_2(\mathbf{x}, t, k)) \right),$$

with the convention $\mathcal{B}(O_{\mathbf{x}k}, 0) = \emptyset$ and

$$\begin{aligned}
\phi_1(\mathbf{x}, t) &= z_{\mathbf{x}} + ct \\
&\quad + \sum_{k=1}^{K_{\mathbf{x}}} \Big[\Big((s_{\mathbf{x}k} + ct)^2 + (\phi_1(\mathbf{x}, \tau_{\mathbf{x},k-1}) + c(t - \tau_{\mathbf{x},k-1}))^2 \Big)^{1/2} \\
&\qquad - (\phi_1(\mathbf{x}, \tau_{\mathbf{x},k-1}) + c(t - \tau_{\mathbf{x},k-1})) \Big] \mathbb{1}(t \geq \tau_{\mathbf{x}k}) \\
\phi_2(\mathbf{x}, t, k) &= (s_{\mathbf{x}k} + ct)\mathbb{1}(t < \tau_{\mathbf{x}k}), \quad \forall k = 1, \cdots, K_{\mathbf{x}} \\
\tau_{\mathbf{x}0} &= 0 \\
\tau_{\mathbf{x}k} &= \tau_{\mathbf{x},k-1} + \max \left\{ 0, \frac{d(O_{\mathbf{x}}, O_{\mathbf{x}k}) - \phi_1(\mathbf{x}, \tau_{\mathbf{x},k-1}) - \phi_2(\mathbf{x}, \tau_{\mathbf{x},k-1}, k)}{2c} \right\} \\
&\quad \forall k = 1, \ldots, K_{\mathbf{x}}.
\end{aligned}$$

Thus, X_t is a PDMP with flow function Φ:

$$X_t = \begin{cases} \Phi(X_{T_j}, t) & \text{if } t \in [T_j, T_{j+1}) \\ U_{j+1} & \text{if } t = T_{j+1}, \end{cases}$$

where the inter-jump duration $S_{j+1} = T_{j+1} - T_j$ (with $j \geq 1$ and the convention $T_0 = 0$) has a survival function satisfying:

$$P(S_{j+1} \geq t) = \exp\left(- \int_0^t \lambda(\Phi(X_{T_j}, v)) dv \right);$$

the rate function $\lambda : \mathcal{X} \to \mathbb{R}_+$ satisfies:

$$\lambda(\mathbf{x}) = \tilde{\lambda}(z_{\mathbf{x}}),$$

with $\tilde{\lambda}(z_{\mathbf{x}}) = \lambda_1 z_{\mathbf{x}}$ for example as proposed in Section 8.2.3; and U_{j+1} is drawn from the jump kernel $Q(\Phi(X_{T_j}, S_{j+1}), \cdot)$ such that:

$$U_{j+1} = \Phi(X_{T_j}, S_{j+1}) \cup \mathcal{B}(O_{\text{new}}, 0)$$

with O_{new} uniformly drawn on the circle centered around O and radius $z(T_{j+1}) + L$.

7.2.4. *A PDMP Based on Reaction-Diffusion for Modeling Invasions with Multiple Introductions*

Section 8.2.1 presented the use of reaction-diffusion equations for modeling population dynamics with short-distance dispersal and Section 8.2.2 presented the combination of a jumping process and an approximation of a reaction-diffusion equation to obtain a model with both short and long-distance dispersal. The latter model was shown to be a PDMP. Here, we introduce an other spatio-temporal PDMP based on reaction-diffusion for modeling dynamics with short-distance dispersal only but with multiple introductions of the species of interest. In this model, the flow represented by a reaction-diffusion equation with an Allee effect will be stochastically disrupted at random times to mimic introductions having a limited extent in space. This model will be used in a future study to describe the dynamics of the plant-pathogenic bacterium *Xylella fastidiosa* (Xf) in Corsica. Figure 8.5 shows the pattern of plants which have been detected as infected by Xf in Corsica between August 2015 and May 2017. This map displays several clusters of infected plants with different sizes, which may have been induced by several introductions of the pathogen in different areas of Corsica and at different times.

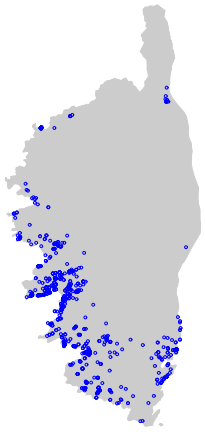

Figure 7.5: Pattern of plants which have been detected as infected by *Xylella fastidiosa* in Corsica between August 2015 and May 2017.

In what follows, we introduce a candidate model for describing the invasion of Corsica by Xf and lay some track to estimate the unknown parameters and latent variables of the model. Assume that $u(t, \mathbf{x})$, which will be used to model the

probability that a plant located at $\mathbf{x} \in \Omega \subset \mathbb{R}^2$ is infected at time t, satisfies between two introductions of the invading species:

$$\begin{cases} \frac{\partial u}{\partial t} = D\Delta u + bu(u-\theta)(1-u) & \text{in } \Omega \\ \nabla u.n = 0 & \text{on } \partial\Omega, \end{cases} \quad [7.6]$$

where D is the dispersal rate, b the intrinsic growth rate of Xf, and $\theta \in]0; \frac{1}{2}[$ the reaction threshold in Ω which induces an Allee effect (Ω, in the Xf application, will be the area covered by the Corsican territory).

The progression of u will be interrupted at each introduction time and *re-initialized*. At the first introduction time, i.e. $t = \tau_0$, u is initialized as follows:

$$u_0(\mathbf{x}) = u(\tau_0, \mathbf{x}) = f(\mathbf{x}, \mathbf{x}_0) \quad \text{in } \Omega,$$

where $f : \Omega \mapsto [0,1]$ is a continuous function, which is typically decreasing with the distance from $\mathbf{x}_0$ to $\mathbf{x}$ (like a kernel function). Thus, the invading species is first introduced around $\mathbf{x}_0$ at τ_0.

The subsequent introductions (i.e. the jumps) are assumed to be governed by a spatio-temporal homogeneous Poisson point process Ψ with constant intensity λ over $\Omega \times (\tau_0, \tau_{\text{end}})$. Let $\{\psi_0^1, \cdots, \psi_0^N\}$ be a realization of Ψ where $\psi_0^i = (\mathbf{x}_0^i, T_i)$, and set $(\mathbf{x}_0^0, T_0) = (\mathbf{x}_0, \tau_0)$ and $T_{N+1} = \tau_{\text{end}}$. We define the spatio-temporal PDMP $\{X_t\}_{\tau_0 \leq t < \tau_{\text{end}}}$ by:

$$X_t(\mathbf{x}) = \begin{cases} f(\mathbf{x}, \mathbf{x}_0^0) & \text{if } t = T_0 = \tau_0 \\ u(t, \mathbf{x}) & \text{if } t \in (T_i, T_{i+1}),\ i = 0, \ldots, N \\ u(t, \mathbf{x}) + f(\mathbf{x}, \mathbf{x}_0^i) & \text{if } t = T_i,\ i = 1, \ldots, N \end{cases}$$

where u is governed by Equation [8.6] over $(T_i, T_{i+1}]$ with initial state at T_i being X_{T_i}. Then, $\min\{1, \max\{0, X_t(\mathbf{x}))\}\}$ is viewed as the probability that a plant located at $\mathbf{x} \in \Omega$ is infected at time t. The $\min - \max$ operator is used because X_t may sporadically go out of $[0,1]$.

In the application of interest, namely the invading dynamic of Xf in Corsica, the estimation of model parameters (D, b, θ, λ and eventual parameters arising in f) and latent variables (jump times T_i and introduction locations $\mathbf{x}_0^i$) will be carried out in a mechanistic-statistical framework, which can cope with various types of data [ROQ 11, SOU 09a, SOU 09b, WIK 03a, WIK 03b]. Consider, for instance, that data collection consists of independently sampling plants in $\Omega \times (\tau_0, \tau_{\text{end}})$ and diagnosing their health statuses. Let $Z(\mathbf{s}_j, t_j) \in \{0, 1\}$ be the observed health status of plant j sampled at location $\mathbf{s}_j$ and time t_j, $j = 1, \ldots, n$, where 0 stands for the *observed healthy status* and 1 for the *observed infected status*. Let ϵ_{FN} be the probability of

diagnosing a plant as healthy whereas it is infected (false-negative rate) and ϵ_{FP} be the probability of diagnosing a plant as infected whereas it is healthy (false-postive rate). Then, $Z(\mathbf{s}_j, t_j)$ can be assumed to be Bernoulli distributed as follows:

$$Z(\mathbf{s}_j, t_j) \mid \{X_t\} \overset{\text{indep.}}{\sim} \text{Bernoulli}\Big(\epsilon_{\text{FN}}(1-\epsilon_{\text{FP}})\min\{1, \max\{0, X_{t_j}(\mathbf{s}_j))\}\}\Big),$$

and the estimation of model parameters and latent variables can be made, in a frequentist or Bayesian framework, with the resulting likelihood and an appropriate algorithm (an example of Bayesian algorithm will be given in the next section for a different model).

7.3. Metapopulation Epidemic Model

7.3.1. *Spatially Realistic Levins Model*

In ecology, the class of Stochastic Patch Occupancy Models (SPOM) has been developed to characterize and infer the dynamics of metapopulations. A metapopulation is a set of spatially separated populations of the same species which interact via between-population migrations of individuals. Among this class of models, the spatially realistic Levins model (SRLM) is a major reference [OVA 04].

Consider a set of n circular habitat patches with areas $a_i > 0$ and centers $x_i \in \mathbb{R}^2$, $i \in \mathcal{I} = \{1, \ldots, n\}$. Let $d_{i,j}$ denote the Euclidean distance between x_i and x_j. The binary variable $Y_i(t) \in \{0, 1\}$ gives the occupation status of patch i at time $t \in \mathbb{R}$: $Y_i(t) = 1$ if patch i is occupied by the species of interest at t, $Y_i(t) = 0$ otherwise. The random vector $\mathbf{Y}(t) = \{Y_1(t), \ldots, Y_n(t)\}$ follows a binary-state continuous-time Markov process with inhomogeneous transition rates. Local extinctions independently occur with a constant rate e_i, which is typically proportional to the patch area a_i:

$$\mathbb{P}(Y_i(t+dt) = 0 \mid Y_i(t) = 1) = e_i dt.$$

Colonizations of unoccupied patches occur with a time-varying rate depending on the occupation status of the other patches and their distance with respect to the focal patch:

$$\mathbb{P}(Y_i(t+dt) = 1 \mid Y_i(t) = 0) = \sum_{\substack{j=1 \\ j \neq i}}^{n} c_{ij} Y_j(t) dt,$$

where c_{ij} is typically a function of the distance d_{ij} and other patch characteristics such as the areas a_i and a_j. In general, the larger d_{ij}, the lower c_{ij} (source patches send more migrants to close patches than to further patches), and the larger a_i and a_j, the larger c_{ij} (large patches send more migrants and have a higher propensity to receive migrants).

7.3.2. *A Colonization Piecewise-Deterministic Markov Process*

Here, we are interested in a pathogen metapopulation. Thus, in what follows, we adopt the vocabulary of epidemiology. In particular, thereafter, a patch is a set of hosts for the pathogen of interest, an occupied patch is a patch that is infected by a pathogen population, and an unoccupied patch is said to be healthy.

This section presents the metapopulation model proposed in [SOU 09a], which differs from the Levins model mainly because (i) extinctions and colonizations occur on distinct periods, (ii) the binary occupation status $Y_i(t)$ is augmented by a time-varying quantitative variable providing the size of the pathogen population within patch i, and (iii) observation variables are explicitly introduced in the model. To simplify the presentation of the model, we focus on the metapopulation dynamic during one year, which is assumed to consist of two successive periods: the *dormancy* period and the *growing season* period. Without loss of generality, we assume that dormancy occurs during the time interval $[-1, 0)$ while the growing season occurs during the interval $[0, 1)$. The initial time $t = -1$ is just after the end of the previous growing season, while time $t = 1$ corresponds to the beginning of the next season.

In the following, *infection times* T_i ($i \in \mathcal{I}$) denote the times of initiation of local epidemics in the year under consideration; let $\mathbf{T} = \{T_i : i \in \mathcal{I}\}$. As a local epidemic can only occur during the growing season, $T_i \geq 0$. We assume that the pathogen survived in patch i during the dormancy if and only if $T_i = 0$. In the case of local epidemics not due to survival of the pathogen in patch i the infection time is a *colonization time*. By convention, we set $T_i \geq 1$ if patch i is still healthy at time $t = 1$.

Observation variables

Â The metapopulation dynamic is observed at the patch level at times $t = -1$ and $t = 1$, i.e. the end of successive years. Given that sampling is not complete (there are some patches whose health statuses are not observed) and that infections are not always detected, we introduce the observation variables $Y^{\text{obs}}_{i,t}$, $i \in \mathcal{I} = \{1, \ldots, n\}$ and $t \in \{0, 1\}$:

$$Y^{\text{obs}}_{i,t} = \begin{cases} 0 & \text{if the meadow is observed as healthy} \\ 1 & \text{if the meadow is observed infected} \\ \texttt{NA} & \text{if the meadow is not sampled.} \end{cases}$$

There are no false-positives (i.e. healthy patches observed as infected). In addition, vectors of explanatory variables are observed at the patch level, namely the patch coordinates x_i, the area a_i covered by the patch and $\{B_i, C_i, D_i\}$ that will arise in the model as regressors.

In the model, the response variables are the observations $\mathbf{Y}_1^{\text{obs}} = \{Y_{i,1}^{\text{obs}} : i \in \mathcal{I}\}$ at time $t = 1$, and we work conditionally on past observations $\mathbf{Y}_{-1}^{\text{obs}} = \{Y_{i,-1}^{\text{obs}} : i \in \mathcal{I}\}$ and covariates $\{x_i, a_i, B_i, C_i, D_i : i \in \mathcal{I}\}$. The observed final health statuses $Y_{i,1}^{\text{obs}}$ are assumed to be independently drawn from $\{0, 1, \texttt{NA}\}$ with unequal probabilities, given actual final health statuses:

$$Y_{i,1}^{\text{obs}} \mid Y_i(1) \sim \alpha_1 \text{Dirac}(0) + \alpha_2 \text{Dirac}(1) + (1 - \alpha_1 - \alpha_2)\text{Dirac}(\texttt{NA}),$$

where α_1 and α_2 account for misclassification and incompleteness in the observation process at $t = 1$ and satisfy:

$$\begin{aligned}
\alpha_1 &= r_1 \frac{p_1}{p_1 + q_1(1 - p_1)} \\
\alpha_2 &= r_1 \left(1 - \frac{p_1}{p_1 + q_1(1 - p_1)}\right) \\
p_1 &= \mathbb{P}(Y_{i1}^{\text{obs}} = 1 \mid Y_{i1}^{\text{obs}} \neq \texttt{NA}) \\
q_1 &= \mathbb{P}(Y_{i,1} = 1 \mid Y_{i1}^{\text{obs}} = 0) \\
r_1 &= \mathbb{P}(Y_{i1}^{\text{obs}} \neq \texttt{NA}).
\end{aligned}$$

Probabilities p_1, q_1 and r_1 are *observation parameters* whose values are assessed before fitting the model to data and plugged in the model.

Extinctions

Extinctions of the pathogen in infected patches can only occur during the dormancy period $[-1, 0)$. Times of extinction are not explicitly introduced into the model. We simply assume that extinctions between times -1 and 0 are, conditionally on observations $Y_{i,-1}^{\text{obs}}$, the result of independent Bernoulli draws for the infection statuses $Y_i(0)$ of patches:

$$\begin{aligned}
Y_i(0) \mid Y_{i,-1}^{\text{obs}} &\sim \text{Bernoulli}(b_i s(Y_{i,-1}^{\text{obs}})) \\
b_i &= \text{logit}^{-1}(B_i^T \beta) \\
s(Y_{i,-1}^{\text{obs}}) &= \begin{cases} 1 & \text{if } Y_{i,-1}^{\text{obs}} = 1 \\ q_{-1} & \text{if } Y_{i,-1}^{\text{obs}} = 0 \\ p_{-1} + q_{-1}(1 - p_{-1}) & \text{if } Y_{i,-1}^{\text{obs}} = \texttt{NA}, \end{cases}
\end{aligned} \qquad [7.7]$$

where b_i gives the conditional probability of pathogen survival given that patch i was infected in the beginning of dormancy, and s deals with misclassification and incompleteness of the observation process at time $t = -1$. b_i is a function of observed covariates B_i and a vector of parameters β (B_i^T is the transpose of B_i), $p_{-1} = \mathbb{P}(Y_{i,-1}^{\text{obs}} = 1 \mid Y_{i,-1}^{\text{obs}} \neq \texttt{NA})$ and $q_{-1} = \mathbb{P}(Y_{i,-1} = 1 \mid Y_{i,-1}^{\text{obs}} = 0)$. Probabilities p_{-1} and q_{-1} are *observation parameters* whose values are assessed before fitting the model to data and plugged in the model. By convention, $Y_i(0) = 1$ if and only if $T_i = 0$.

Colonizations

Healthy patches are immune during the dormancy and susceptible within the growing season. Infected patches are infectious only during the growing season. The degrees of susceptibility and infectiousness depend on explanatory variables and time as described below. In addition, already infected patches cannot be over-infected during the growing season.

The spread of the pathogen during the growing season is modeled as a spatio-temporal piecewise-Poisson point process [ILL 08] . In this process, point (t, x) specifies a time and a location at which the numbers of dispersing incoming pathogen are large enough to potentially initiate a local epidemic in a healthy patch with a standard degree of susceptibility. Thus, each point stands for a potential colonization event.

The point process is governed by an intensity function $\tilde{\lambda}$ quantifying the risk of infection at each space-time location, this risk being generated by the already infected patches. Therefore, $\tilde{\lambda}$ varies in time and space with the number, the spatial locations and the infectiousness of these patches. The expression of $\tilde{\lambda}$ at time t and location x is given by:

$$\tilde{\lambda}(t, x) = \sum_{j \in \mathcal{I}_t} c_j g_j(t - T_j) h(x, x_j), \qquad [7.8]$$

where $\mathcal{I}_t = \{j \in \mathcal{I} : T_j < t\}$ is the set of patches infected before time t; c_j encodes characteristics of patch j such as its physiological state and features of the surrounding habitat, which are expected to partly determine the infectiousness of j; g_j is a deterministic standardized disease progress function, which gives the shape of the pathogen growth within patch j; h is a dispersal function, which models pathogen dispersal as a function of the source location x_j and the location of the receiving patch x. The product $c_j g_j(t - T_j)$ specifies the degree of infectiousness of patch j at time t. In the beginning of the growing season, just after time zero, $\tilde{\lambda}$ is generated only by those patches in which the pathogen survived during the dormancy.

The standardized disease progress function is specified with a thresholded quadratic form:

$$g_j(t) = \min\{t^2, \omega a_j\} \mathbb{1}(t \geq 0), \qquad [7.9]$$

where ω is a positive parameter. The threshold ωa_j takes into account possible saturation effects, which are assumed to be proportional to the patch area a_j.

The dispersal function h is specified as an anisotropic exponential dispersal function parameterized by $\eta = (\eta_1, \ldots, \eta_5)$ [SOU 07]:

$$h(x, x') = \frac{h_1\{\phi(x - x')\}}{h_2\{\phi(x - x')\}^2} \exp\left(-\frac{||x - x'||}{h_2\{\phi(x - x')\}}\right),$$

where $\phi(x-x')$ is the angle made by the vector $x-x'$, $||x-x'||$ is the distance between x and x', $h_1(\phi)$ gives the probability that a spore is dispersed in direction ϕ, and $h_2(\phi)$ gives the expected distance travelled by a spore dispersed in direction ϕ. The angular function h_1 is assumed to be a von Mises density function [FIS 95] parameterized by a mean direction parameter $\eta_1 \in \mathbb{R}$ and a dispersion parameter $\eta_2 > 0$:

$$h_1(\phi) = \{2\pi I_0(\eta_2)\}^{-1} \exp\{\eta_2 \cos(\phi - \eta_1)\},$$

with $I_0(u) = (2\pi)^{-1} \int_0^{2\pi} \exp\{u \cos(\phi)\} d\phi$. The angular function h_2 is assumed to be proportional to a von Mises density function parameterized by a mean direction parameter $\eta_3 \in \mathbb{R}$, a dispersion parameter $\eta_4 > 0$

$$h_2(\phi) = \eta_5\{2\pi I_0(\eta_4)\}^{-1} \exp\{\eta_4 \cos(\phi - \eta_3)\},$$

where $\eta_5 > 0$ is the constant of proportionality.

A healthy patch i is colonized during the growing season if a point of the piecewise Poisson point process is deposited in i and it succeeds in initiating a local epidemic. The intensity of points deposited in i at time t is given by the product $a_i \tilde{\lambda}(t, x_i)$; a_i is considered as the effective capture area of patch i and $x \mapsto \tilde{\lambda}(t, x)$ is assumed to be approximately constant over patch i. Any deposited point is assumed to initiate a local epidemic with probability d_i, which reflects the degree of susceptibility of i and encodes individual characteristics such as local climatic conditions.

Quantities c_j and d_i always appear in the model as the product $c_j d_i$. They are jointly modeled as a function of explanatory variables: $c_j d_i = \exp(C_j^T \gamma + D_i^T \delta)$, where C_j and D_i are vectors of covariates, and γ and δ are vectors of parameters.

PDMP formulation of the colonization dynamic

Let $\mathbf{X}_t \in \mathcal{X}$, $t \in [0, 1]$, be the $[2 \times n]$ matrix satisfying:

$$\mathbf{X}_t = \begin{pmatrix} X_{11}(t) & \cdots & X_{1n}(t) \\ X_{21}(t) & \cdots & X_{2n}(t) \end{pmatrix} = \begin{pmatrix} c_1 g_1(t - T_1) & \cdots & c_n g_n(t - T_n) \\ Y_1(t) & \cdots & Y_n(t) \end{pmatrix},$$

where each column provides, for a given patch, the size of the pathogen population at time t and the health status of the patch at time t (remind that $Y_i(t) = \mathbb{1}(t \geq T_i)$).

We introduce the function $\Phi = (\Phi_1, \ldots, \Phi_n) : \mathcal{X} \times \mathbb{R}_+ \to \mathcal{X}$ whose j-th component satisfies:

$$\Phi_j(\mathbf{x}, t) = \begin{cases} \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } \mathbf{x}_{2j} = 0 \\ \begin{pmatrix} c_j \min\{(t + \sqrt{\mathbf{x}_{1j}/c_j})^2, \omega a_j\} \\ 1 \end{pmatrix} & \text{if } \mathbf{x}_{2j} = 1. \end{cases} \quad [7.10]$$

Let T_i and $T_{i'}$ be two successive colonization times (i.e. $0 < T_i < T_{i'}$ and no colonization occurred in the time interval $(T_i, T_{i'})$), called *jump times* in the PDMP framework. The inter-jump duration $S_{i'} = T_{i'} - T_i$ has a survival function detailed in Equation [8.16] that takes an exponential form depending on the multivariate *jump rate* $\lambda : \mathcal{X} \mapsto \mathbb{R}_+^n$:

$$\lambda(\mathbf{X}_t) = \begin{pmatrix} d_1 a_1 \tilde{\lambda}(t, x_1)(1 - Y_1(t)) \\ \vdots \\ d_n a_n \tilde{\lambda}(t, x_n)(1 - Y_n(t)) \end{pmatrix},$$

where $\tilde{\lambda}$ was defined in Equation [8.8] and can be expressed as a function of $\mathbf{X}_t$, and the variables $Y_1(t), \ldots, Y_n(t)$ are the components of the 2nd row of $\mathbf{X}_t$.

Using Equations [8.9] and [8.10], $\mathbf{X}_t$ is a PDMP with *flow function* Φ:

$$\mathbf{X}_t = \begin{cases} \Phi(\mathbf{X}_{T_i}, t) & \text{if } t \in [T_i, T_{i'}) \\ \mathbf{U}_{i'} & \text{if } t = T_{i'}, \end{cases}$$

where $\mathbf{U}_{i'}$ is drawn from the *jump kernel* $Q_{i'}(\Phi(\mathbf{X}_{T_i}, S_{i'}), \cdot)$. In the simplest case (the one which is considered thereafter), the jump kernel is a Dirac distribution, which changes only the health status $X_{2i'}(t) = Y_{i'}(t)$ of i' from healthy to infected:

$$\mathbf{U}_{i'} = \Phi(\mathbf{X}_{T_i}, S_{i'}) + \begin{pmatrix} \mathbf{0}_n \\ \mathbf{1}_n(i') \end{pmatrix},$$

where $\mathbf{0}_n$ is the raw vector with n zeros and $\mathbf{1}_n(i')$ is the raw vector whose i'-th element is equal to 1 and the $n - 1$ other elements are equal to 0. This form could be generalized by drawing a random value for the size of the pathogen population $X_{1i'}(t)$ in patch i' when this patch is colonized:

$$\mathbf{U}_{i'} = \Phi(\mathbf{X}_{T_i}, S_{i'}) + \begin{pmatrix} \min\{U_{i'}, \omega a_{i'} c_{i'}\} \mathbf{1}_n(i') \\ \mathbf{1}_n(i') \end{pmatrix},$$

where the real variable $U_{i'}$ should be randomly drawn in $\mathbb{R}_+$. As mentioned above, we use thereafter the simplest case:

$$\mathbf{X}_t = \begin{cases} \Phi(\mathbf{X}_{T_i}, t) & \text{if } t \in [T_i, T_{i'}) \\ \Phi(\mathbf{X}_{T_i}, S_{i'}) + \begin{pmatrix} \mathbf{0}_n \\ \mathbf{1}_n(i') \end{pmatrix} & \text{if } t = T_{i'}. \end{cases}$$

7.3.3. *Bayesian Inference Approach*

We aim to infer infection times $\mathbf{T}$ and parameters $\boldsymbol{\Theta} = (\omega, \eta, \beta, \gamma, \delta)$ given observed health statuses $\mathbf{Y}_{i,-1}^{\text{obs}}$ and $\mathbf{Y}_{i,1}^{\text{obs}}$, covariates $\mathbf{Z} = \{x_i, a_i, B_i, C_i, D_i : i \in \mathcal{I}\}$ and observation parameters $\kappa_{-1} = (p_{-1}, q_{-1})$ and $\kappa_1 = (p_1, q_1)$ (we will see below, in *Remark 1*, that the observation parameter r_1 can be removed from the model in the inference stage). The inference is made by using the probability distribution $P(\mathbf{Y}_1^{\text{obs}} \mid \mathbf{Y}_{-1}^{\text{obs}}, \mathbf{Z})$, which can be written as follows:

$$P(\mathbf{Y}_1^{\text{obs}} \mid \mathbf{Y}_{-1}^{\text{obs}}, \mathbf{Z}) = \int_{\mathbf{T}} P_{\kappa_1}(\mathbf{Y}_1^{\text{obs}} \mid \mathbf{T}) dP_{\boldsymbol{\Theta},\kappa_{-1}}(\mathbf{T} \mid \mathbf{Y}_{-1}^{\text{obs}}, \mathbf{Z}). \quad [7.11]$$

Equation [8.11] highlights the hierarchical structure of the model. In the first stage, the term $P_{\boldsymbol{\Theta},\kappa_{-1}}(\mathbf{T} \mid \mathbf{Y}_{-1}^{\text{obs}}, \mathbf{Z})$ gives the distribution of infection times given the observed initial statuses and covariates. This term incorporates the survival process during dormancy and the colonization PDMP parameterized by $\boldsymbol{\Theta}$, and the observation process at time $t = -1$ parameterized by κ_{-1}. In the second stage, the term $P_{\kappa_1}(\mathbf{Y}_1^{\text{obs}} \mid \mathbf{T})$ gives the distribution of the observed final statuses given infection times. This term corresponds to the observation process at time $t = 1$ parameterized by κ_1. Note that when $\mathbf{T}$ is known, $\mathbf{Y}_{-1}^{\text{obs}}$ and $\mathbf{Z}$ bring no information on $\mathbf{Y}_1^{\text{obs}}$, i.e. $P_{\kappa_1}(\mathbf{Y}_1^{\text{obs}} \mid \mathbf{T}, \mathbf{Y}_{-1}^{\text{obs}}, \mathbf{Z}) = P_{\kappa_1}(\mathbf{Y}_1^{\text{obs}} \mid \mathbf{T})$.

Equation [8.11] can be used to infer the unknowns $\mathbf{T}$ and $\boldsymbol{\Theta}$. However, the integral at the right-hand-side cannot be calculated analytically. To overcome this difficulty, the infection times $\mathbf{T}$ can be considered as latent variables, whose distribution is specified by $P_{\boldsymbol{\Theta},\kappa_{-1}}(\mathbf{T} \mid \mathbf{Y}_{-1}^{\text{obs}}, \mathbf{Z})$, and inference can be carried out with a Markov chain Monte Carlo (MCMC) method in the Bayesian context [ROB 99] or a Monte Carlo expectation maximization method in the frequentist context [WEI 90].

In this study, we chose the Bayesian approach and we applied MCMC using a Metropolis-Hastings algorithm to draw a sample from the posterior distribution of the parameters and the infection times. The posterior distribution, up to a normalizing constant, can be written as

$$P_{\kappa_{-1},\kappa_1}(\boldsymbol{\Theta}, \mathbf{T} \mid \mathbf{Y}_{-1}^{\text{obs}}, \mathbf{Y}_1^{\text{obs}}, \mathbf{Z}) \propto P_{\kappa_1}(\mathbf{Y}_1^{\text{obs}} | \mathbf{T}) P_{\boldsymbol{\Theta},\kappa_{-1}}(\mathbf{T} \mid \mathbf{Y}_{-1}^{\text{obs}}, \mathbf{Z}) \pi(\boldsymbol{\Theta}), [7.12]$$

where π is the prior distribution of $\boldsymbol{\Theta}$ and the symbol '$\propto$' means 'proportional to'. The following paragraphs provide the expressions of the terms appearing in Equation [8.12].

Expression of $P(\mathbf{T} \mid \mathbf{Y}_{-1}^{obs}, \mathbf{Z})$

Here, we give the expression of the conditional probability of any space-time configuration $\mathbf{T}$, describing what patches are infected at what times, given the observed initial health statuses $\mathbf{Y}_{-1}^{\text{obs}}$ and covariates $\mathbf{Z}$. Thereafter, for the sake of convenience, we omit the conditioning covariates and the conditioning parameters in the notation.

We make the three following assumptions in addition to those made above. First, the infection potential is constant within each patch. Second, the degree of susceptibility of a healthy patch at time zero is independent of the initial health status at time $t = -1$. Third, points of the Poisson point process located in susceptible patches independently succeed in initiating local epidemics. The success of a point in initiating a local epidemic is patch dependent. It is measured by the success probability d_i which reflects the degree of susceptibility of i and encodes individual characteristics such as local climatic conditions.

Let $t_1, \ldots, t_n$ be times in $[0, 1]$ and $\mathcal{I}_A = \{i \in \mathcal{I} : t_i = 0\}$, $\mathcal{I}_B = \{i \in \mathcal{I} : 0 < t_i < 1\}$ and $\mathcal{I}_C = \{i \in \mathcal{I} : t_i = 1\}$. $\mathcal{I}_A$, $\mathcal{I}_B$ and $\mathcal{I}_C$ are associated, respectively, with the sets of patches where the pathogen survived during the dormancy, which were colonized during the season and which remained healthy. We show below that:

$$\begin{aligned} P(\{T_i = 0 : i \in \mathcal{I}_A\}, \{T_i = t_i : i \in \mathcal{I}_B\}, \{T_i \geq 1 : i \in \mathcal{I}_C\} \mid \mathbf{Y}_{-1}^{\text{obs}}) \\ = \prod_{i \in \mathcal{I}_A} b_i s(Y_{i,-1}^{\text{obs}}) \prod_{i \in \mathcal{I}_B} \{1 - b_i s(Y_{i,-1}^{\text{obs}})\} e^{-d_i a_i \tilde{\Lambda}(t_i, x_i)} d_i a_i \tilde{\lambda}(t_i, x_i) \\ \times \prod_{i \in \mathcal{I}_C} \{1 - b_i s(Y_{i,-1}^{\text{obs}})\} e^{-d_i a_i \tilde{\Lambda}(1, x_i)}, \end{aligned} \quad [7.13]$$

where $\tilde{\Lambda}(t, x) = \int_0^t \tilde{\lambda}(s, x) ds$ is the time-cumulated infection risk affecting location x. Quantities d_i and c_j are only contained in $d_i a_i \tilde{\lambda}(t_i, x_i)$ and $d_i a_i \tilde{\Lambda}(t_i, x_i)$ as the product form $d_i c_j$. This product was directly modeled (instead of separately modeling d_i and c_j) to avoid identifiability difficulties in parameter estimation.

In Equation [8.13], the term $b_i s(Y_{i,-1}^{\text{obs}})$ is the probability of pathogen survival in i during the dormancy. In the second product of [8.13], the term $1 - b_i s(Y_{i,-1}^{\text{obs}})$ is the probability of pathogen extinction in i during the dormancy. The term $e^{-d_i a_i \tilde{\Lambda}(t_i, x_i)} d_i a_i \tilde{\lambda}(t_i, x_i)$ is the probability that i remained susceptible during $[0, t_i)$ and was infected at t_i. The product $d_i a_i \tilde{\lambda}(t, x_i)$ of the degree of susceptibility d_i, the capture area a_i, and the infection risk $\tilde{\lambda}(t, x_i)$ measures the instantaneous risk of

infection of patch i at time t. Finally, in the third product of [8.13], $1-b_i s(Y_{i,-1}^{\text{obs}})$ is the probability of pathogen extinction in i during the dormancy and $e^{-d_i a_i \tilde{\Lambda}(1,x_i)}$ is the probability that i remained healthy during the epidemic period $[0,1]$.

Proof of Equation [8.13]

Let $\tau_0, \ldots, \tau_{n+1}$ be $n+2$ ordered times in $[0,1]$ satisfying

$$0 = \tau_0 = \cdots = \tau_q < \cdots < \tau_r = \cdots = \tau_{n+1} = 1,$$

and $\mathcal{I}^* = \{i_1, \ldots, i_n\}$ be a permutation of $\mathcal{I} = \{1, \ldots, n\}$. We want to determine the conditional probability that, given the observed initial statuses $\mathbf{Y}_{-1}^{\text{obs}}$ and the covariates $\mathbf{Z}$,

– patch i_k ($k \leq q$) is infected at time $\tau_k = 0$ (survival of the pathogen during dormancy),

– patch i_k ($q < k < r$) is the k-th patch to be infected and its infection time is $\tau_k \in (0,1)$ (colonization),

– patch i_k, $k \geq r$, is still susceptible at time $\tau_k = 1$.

In other words, we want to determine

$$p(\mathcal{I}^*, \boldsymbol{\tau}; \mathbf{Y}_{-1}^{\text{obs}}) = P(\{T_{i_k} = \tau_k : k < r\}, \{T_{i_k} > \tau_k : k \geq r\} \mid \mathbf{Y}_{-1}^{\text{obs}})$$

where $\boldsymbol{\tau} = \{\tau_1, \ldots, \tau_n\}$. Note that times $\tau_{q+1}, \cdots, \tau_{r-1}$ corresponding to colonization events are mutually different and different from one under the Poisson assumption.

Let $\mathcal{A} = \{T_{i_k} = \tau_k : k \leq q\}$, $\mathcal{B} = \{T_{i_k} = \tau_k : q < k < r\}$, $\mathcal{C} = \{T_{i_k} > \tau_k : k \geq r\}$ and $\mathcal{D} = \{T_{i_k} > 0 : k > q\}$. As $\{T_{i_k} = \tau_k : k < r\} = \mathcal{A} \cap \mathcal{B}$ and the event $\mathcal{D}$ is included in $\mathcal{B} \cap \mathcal{C}$,

$$\begin{aligned} p(\mathcal{I}^*, \boldsymbol{\tau}; \mathbf{Y}_{-1}^{\text{obs}}) &= P(\mathcal{A}, \mathcal{B}, \mathcal{C} \mid \mathbf{Y}_{-1}^{\text{obs}}) \\ &= P(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D} \mid \mathbf{Y}_{-1}^{\text{obs}}) \\ &= P(\mathcal{C} \mid \mathcal{A}, \mathcal{B}, \mathcal{D}, \mathbf{Y}_{-1}^{\text{obs}}) P(\mathcal{B} \mid \mathcal{A}, \mathcal{D}, \mathbf{Y}_{-1}^{\text{obs}}) P(\mathcal{A} \mid \mathcal{D}, \mathbf{Y}_{-1}^{\text{obs}}) P(\mathcal{D} \mid \mathbf{Y}_{-1}^{\text{obs}}). \end{aligned}$$

The two last terms at the right-hand-side of the previous equation correspond to survivals and extinctions during the dormancy and can be written as

$$P(\mathcal{A} \mid \mathcal{D}, \mathbf{Y}_{-1}^{\text{obs}}) = P(\mathcal{A} \mid \mathbf{Y}_{-1}^{\text{obs}}) = \prod_{k \leq q} P(T_{i_k} = 0 \mid Y_{i_k,-1}^{\text{obs}}) = \prod_{k \leq q} b_{i_k} s(Y_{i_k,-1}^{\text{obs}})$$

[7.14]

$$P(\mathcal{D} \mid \mathbf{Y}_{-1}^{\text{obs}}) = \prod_{k > q} P(T_{i_k} > 0 \mid Y_{i_k,-1}^{\text{obs}}) = \prod_{k > q} \{1 - b_{i_k} s(Y_{i_k,-1}^{\text{obs}})\}. \qquad [7.15]$$

where function s, satisfying $s(y) = (q_{-1})^{\mathbb{1}(y=0)}\{p_{-1} + q_{-1}(1-p_{-1})\}^{\mathbb{1}(y=\texttt{NA})}$, $y \in \{0, 1, \texttt{NA}\}$, comes from [8.7].

The term $P(\mathcal{B} \mid \mathcal{A}, \mathcal{D}, \mathbf{Y}_{-1})$ is the conditional probability density function of the colonization times. So, it corresponds to the pathogen spread during the season modeled using a piecewise spatio-temporal Poisson point process with intensity $\tilde{\lambda}$ (see eq. [8.8]). Assuming that the degree of susceptibility of a patch not infected at time zero is not affected by the initial health status, $P(\mathcal{B} \mid \mathcal{A}, \mathcal{D}, \mathbf{Y}^{\text{obs}}_{-1})$ can be decomposed into

$$\begin{aligned} P(\mathcal{B} \mid \mathcal{A}, \mathcal{D}, \mathbf{Y}^{\text{obs}}_{-1}) &= \prod_{q<k<r} P(T_{i_k} = \tau_k \mid \{T_{i_j} = \tau_j : j < k\}) \\ &= \prod_{q<k<r} P(T_{i_k} = \tau_k, \{T_{i_j} > \tau_k : j > k\} \mid \{T_{i_j} = \tau_j : j < k\}) \\ &= \prod_{q<k<r} \left(- \frac{\partial P(T_{i_k} > t, \{T_{i_j} > \tau_k : j > k\} \mid \{T_{i_j} = \tau_j : j < k\})}{\partial t} \bigg|_{t=\tau_k} \right). \end{aligned}$$

$P(T_{i_k} > t, \{T_{i_j} > \tau_k : j > k\} \mid \{T_{i_j} = \tau_j : j < k\})$ is the probability that the k-th patch to be infected is not infected during the time interval $[\tau_{k-1}, t]$, and that the other remaining susceptible patches are not infected during the time interval $[\tau_{k-1}, \tau_k]$. Hence,

$$\begin{aligned} P(T_{i_k} > t, &\{T_{i_j} > \tau_k : j > k\} \mid \{T_{i_j} = \tau_j : j < k\}) \\ &= P(N_{i_k}(\tau_{k-1}, t) = 0, \{N_{i_j}(\tau_{k-1}, \tau_k) = 0 : j > k\} \mid \{T_{i_j} = \tau_j : j < k\}), \end{aligned}$$

where $N_i(t_1, t_2)$ is the number of points —of the Poisson point process— which (i) are located in the spatial surface A_i covered by patch i, (ii) are located in the time interval $[t_1, t_2]$, and (iii) are effectively efficient for initiating a local epidemic. Condition (iii) depends on the degree of susceptibility of the patch in question. We assume that the filter due to (iii) is an independent thinning operator [DIG 83, STO 95] with the probability d_i of thinning which depends on local characteristics. From the Poisson and thinning assumptions, $N_i(t_1, t_2)$ is Poisson distributed with mean value $d_i \int_{A_i} \int_{t_1}^{t_2} \tilde{\lambda}(t, x) dt dx$. Assuming that the infection risk is constant on the spatial surface A_i (with area a_i and centroid x_i) yields

$$N_i(t_1, t_2) \mid \{T_j : j \in \mathcal{I}_{t_2}\} \sim \text{Poisson}(d_i a_i \tilde{\Lambda}(t_1, t_2, x_i))$$

$$\tilde{\Lambda}(t_1, t_2, x_i) = \int_{t_1}^{t_2} \tilde{\lambda}(s, x_i) ds.$$

The distribution of $N_i(t_1, t_2)$ is conditional on infection times in the past of t_2 because $\tilde{\lambda}$ is a function of these times on $[t_1, t_2]$ (see eq. [8.8]). Moreover, $N_{i_j}(\tau_{k-1}, t)$ $(j \geq k)$ are independent for $t \in [\tau_{k-1}, \tau_k]$. This yields

$$\begin{aligned} &P(T_{i_k} > t, \{T_{i_j} > \tau_k : j > k\} \mid \{T_{i_j} = \tau_j : j < k\}) \\ &\quad = \exp\{-d_{i_k} a_{i_k} \tilde{\Lambda}(\tau_{k-1}, t, x_{i_k})\} \prod_{j>k} \exp\{-d_{i_j} a_{i_j} \tilde{\Lambda}(\tau_{k-1}, \tau_k, x_{i_j})\}. \end{aligned} \quad [7.16]$$

It follows

$$P(\mathcal{B} \mid \mathcal{A}, \mathcal{D}, \mathbf{Y}_{-1}^{\text{obs}}) = \prod_{q<k<r} \left(d_{i_k} a_{i_k} \tilde{\lambda}(\tau_k, x_{i_k}) \prod_{j \geq k} \exp\{-d_{i_j} a_{i_j} \tilde{\Lambda}(\tau_{k-1}, \tau_k, x_{i_j})\} \right). \quad [7.17]$$

The term $P(\mathcal{C} \mid \mathcal{A}, \mathcal{B}, \mathcal{D}, \mathbf{Y}_{-1}^{\text{obs}})$ corresponds to the patches which remain susceptible at the end of the season. Its expression was also derived using the Poisson point process. Indeed, $P(\mathcal{C} \mid \mathcal{A}, \mathcal{B}, \mathcal{D}, \mathbf{Y}_{-1}^{\text{obs}})$ is the probability that patches i_k $(k \geq r)$ remain susceptible during the time interval $[\tau_{r-1}, 1]$, i.e. after the infection of the $(r-1)$-th patch to be infected. Thus,

$$P(\mathcal{C} \mid \mathcal{A}, \mathcal{B}, \mathcal{D}, \mathbf{Y}_{-1}^{\text{obs}}) = \prod_{k \geq r} P(N_{i_k}(\tau_{r-1}, 1) = 0 \mid \{T_{i_j} = \tau_j : j < r\}) \quad [7.18]$$

$$= \prod_{k \geq r} \exp\{-d_{i_k} a_{i_k} \tilde{\Lambda}(\tau_{r-1}, 1, x_{i_k})\} \quad [7.19]$$

From [8.14], [8.17] and [8.18], it follows

$$\begin{aligned} p(\mathcal{I}^*, \boldsymbol{\tau}; \mathbf{Y}_{-1}) = &\prod_{k \leq q} b_{i_k} s(Y_{i_k,-1}) \\ &\times \prod_{q<k<r} \{1 - b_{i_k} s(Y_{i_k,-1})\} d_{i_k} a_{i_k} \tilde{\lambda}(\tau_k, x_{i_k}) \exp\{-d_{i_k} a_{i_k} \tilde{\Lambda}(0, \tau_k, x_{i_k})\} \\ &\times \prod_{k \geq r} \{1 - b_{i_k} s(Y_{i_k,-1})\} \exp\{-d_{i_k} a_{i_k} \tilde{\Lambda}(0, 1, x_{i_k})\} \end{aligned}$$

Expression of $P(\mathbf{Y}_1^{obs} \mid \mathbf{T})$

It is assumed that infected patches remain infected until the end of the season, i.e. if $T_i < 1$, then $Y_i(1) = 1$. Moreover, we add one assumption to those made on the

observation process when survivals during the dormancy were modeled: the success in detecting an infection does not depend on the infection time.

Using material provided in the paragraph entitled *Observation variables* in Section 8.3.2, the distribution $P(\mathbf{Y}_1^{\text{obs}} \mid \mathbf{T})$ satisfies:

$$\begin{aligned} P(\mathbf{Y}_1^{\text{obs}} \mid \mathbf{T}) &= \prod_{i \in \mathcal{I}} P(Y_{i1}^{\text{obs}} \mid T_i) \\ &= \prod_{i:Y_{i1}^{\text{obs}}=1} \frac{p_1 \mathbb{1}(T_i < 1)}{p_1 + q_1(1-p_1)} \prod_{i:Y_{i1}^{\text{obs}}=0} \left(1 - \frac{p_1 \mathbb{1}(T_i < 1)}{p_1 + q_1(1-p_1)}\right) \\ &\quad \times (r_1)^{\sum_i \mathbb{1}(Y_{i1}^{\text{obs}} \neq \texttt{NA})} (1-r_1)^{\sum_i \mathbb{1}(Y_{i1}^{\text{obs}} = \texttt{NA})}, \end{aligned} \quad [7.20]$$

Remark 1. Assessing r_1 prior to the estimation procedure is not required since the term $(r_1)^{\sum_i \mathbb{1}(Y_{i1}^{\text{obs}} \neq \texttt{NA})}(1-r_1)^{\sum_i \mathbb{1}(Y_{i1}^{\text{obs}} = \texttt{NA})}$ in [8.20] brings no information on the dynamics and can be removed from the posterior distribution in the MCMC.

Remark 2. In [8.20], the fraction $p_1 \mathbb{1}(T_i < 1)/\{p_1 + q_1(1-p_1)\}$ is the probability that $Y_{i1}^{\text{obs}} = 1$ given the infection time T_i and given that the patch is sampled at time $t = 1$. It equals zero if $T_i \geq 1$ since a healthy patch is never observed as infected. It is less than one if $T_i < 1$ since the pathogen presence in an infected patch can be undetected.

7.3.4. *Markov Chain Monte Carlo (MCMC) Algorithm*

This section shows how to sequentially update the parameters and the infection times in the MCMC algorithm, by exploiting the decomposition properties of the posterior distribution (block updating).

The posterior distribution can be decomposed as follows. We split $\boldsymbol{\Theta}$ into two subsets: $\boldsymbol{\Theta} = (\theta_1, \theta_2)$, where θ_1 is the parameter vector used to specify the survival probabilities b_i ($i \in \mathcal{I}$), and θ_2 is the parameter vector used in the infection risk $\tilde{\lambda}$. Actually, θ_2 parameterize c_i, d_i, g and h which appear in $\tilde{\lambda}$. The posterior distribution $P_{\kappa_{-1},\kappa_1}(\boldsymbol{\Theta}, \mathbf{T} \mid \mathbf{Y}_{-1}^{\text{obs}}, \mathbf{Y}_1^{\text{obs}})$ can be decomposed into, up to a multiplicative constant,

$$P_{\kappa_{-1},\kappa_1}(\boldsymbol{\Theta}, \mathbf{T} \mid \mathbf{Y}_{-1}^{\text{obs}}, \mathbf{Y}_1^{\text{obs}}) \propto P_{\kappa_1}(\mathbf{Y}_1^{\text{obs}} | \mathbf{T}) Q_{\kappa_{-1}}(\mathbf{T}, \mathbf{Y}_{-1}^{\text{obs}}, \theta_1) Q(\mathbf{T}, \theta_2) \pi_1(\theta_1) \pi_2(\theta_2) \quad [7.21]$$

where π_1 and π_2 are the prior distributions for θ_1 and θ_2, $P_{\kappa_1}(\mathbf{Y}_1^{\text{obs}} \mid \mathbf{T})$ is given by [8.20], and

$$Q_{\kappa_{-1}}(\mathbf{T} \mid \mathbf{Y}_{-1}^{\text{obs}}, \theta_1) = \prod_{i:T_i=0} b_i s(Y_{i,-1}^{\text{obs}}) \prod_{i:T_i>0} \{1 - b_i s(Y_{i,-1}^{\text{obs}})\} \quad [7.22]$$

$$Q(\mathbf{T} \mid \theta_2) = \prod_{i:0<T_i<1} d_i a_i \tilde{\lambda}(T_i, x_i) e^{-d_i a_i \tilde{\Lambda}(T_i, x_i)} \prod_{i:T_i \geq 1} e^{-d_i a_i \tilde{\Lambda}(1, x_i)}, \quad [7.23]$$

are obtained from [8.13].

Let $\mathbf{T}^c$, θ_1^c and θ_2^c denote current values for the infection times and the parameters in the algorithm. Let $\mathbf{T}^*$, θ_1^* and θ_2^* be candidate values respectively drawn from the proposal distributions $q(\cdot \mid \mathbf{T}^c)$, $q(\cdot \mid \theta_1^c)$ and $q(\cdot \mid \theta_2^c)$. First, $\mathbf{T}^*$ replaces $\mathbf{T}^c$ with probability

$$\min\left\{1, \frac{P_{\kappa_1}(\mathbf{Y}_1^{\text{obs}}|\mathbf{T}^*) Q_{\kappa_{-1}}(\mathbf{T}^*, \mathbf{Y}_{-1}^{\text{obs}}, \theta_1^c) Q(\mathbf{T}^*, \theta_2^c) q(\mathbf{T}^c \mid \mathbf{T}^*)}{P_{\kappa_1}(\mathbf{Y}_1^{\text{obs}}|\mathbf{T}^c) Q_{\kappa_{-1}}(\mathbf{T}^c, \mathbf{Y}_{-1}^{\text{obs}}, \theta_1^c) Q(\mathbf{T}^c, \theta_2^c) q(\mathbf{T}^* \mid \mathbf{T}^c)}\right\}.$$

No significant simplification is possible in the calculation of this acceptance probability (only the priors disappear). Second, θ_1^* replaces θ_1^c with probability

$$\min\left\{1, \frac{Q_{\kappa_{-1}}(\mathbf{T}^c, \mathbf{Y}_{-1}^{\text{obs}}, \theta_1^*) \pi_1(\theta_1^*) q(\theta_1^c \mid \theta_1^*)}{Q_{\kappa_{-1}}(\mathbf{T}^c, \mathbf{Y}_{-1}^{\text{obs}}, \theta_1^c) \pi_1(\theta_1^c) q(\theta_1^* \mid \theta_1^c)}\right\}.$$

Here, only the new value of [8.22] and $\pi_1(\theta_1^*)$ must be computed. Third, θ_2^* replaces θ_2^c with probability

$$\min\left\{1, \frac{Q(\mathbf{T}^c, \theta_2^*) \pi_2(\theta_2^*) q(\theta_2^c \mid \theta_2^*)}{Q(\mathbf{T}^c, \theta_2^c) \pi_2(\theta_2^c) q(\theta_2^* \mid \theta_2^c)}\right\}.$$

Here, only the new value of [8.23] and $\pi_2(\theta_2^*)$ must be computed.

If the number of infection times is large, then the proposed infection times will certainly be always rejected. To overcome this issue, one can sequentially update subsets of infectious times. For any subset $\mathcal{J}$ of $\mathcal{I}$, we can draw candidate values $\mathbf{T}_{\mathcal{J}}^* = \{T_i^* : i \in \mathcal{J}\}$ from a proposal distribution $q(\cdot \mid \mathbf{T}_{\mathcal{J}}^c)$, where $\mathbf{T}_{\mathcal{J}}^c = \{T_i^c : i \in \mathcal{J}\}$, and accept it with probability

$$\min\left\{1, \frac{P_{\kappa_1}(\mathbf{Y}_1^{\text{obs}}|\mathbf{T}^*) Q_{\kappa_{-1}}(\mathbf{T}^*, \mathbf{Y}_{-1}^{\text{obs}}, \theta_1^c) Q(\mathbf{T}^*, \theta_2^c) q(\mathbf{T}_{\mathcal{J}}^c \mid \mathbf{T}_{\mathcal{J}}^*)}{P_{\kappa_1}(\mathbf{Y}_1^{\text{obs}}|\mathbf{T}^c) Q_{\kappa_{-1}}(\mathbf{T}^c, \mathbf{Y}_{-1}^{\text{obs}}, \theta_1^c) Q(\mathbf{T}^c, \theta_2^c) q(\mathbf{T}_{\mathcal{J}}^* \mid \mathbf{T}_{\mathcal{J}}^c)}\right\},$$

where component i of $\mathbf{T}^*$ is T_i^* if $i \in \mathcal{J}$, and T_i^c else. Note that a similar procedure can be applied for θ_1 and θ_2 if their dimensions are extensive.

7.3.5. *Example of Results*

The inference approach presented above was applied to infer the metapopulation dynamic of the powdery mildew *Podosphaera plantaginis*, which is a fungal pathogen of the host plant *Plantago lanceolata*, in the Åland Islands. Host plants are spread in more than 4000 meadows (i.e. patches) in this archipelago. Figure 8.6 shows patches observed as infected in 2003 and 2004. Details about data, prior distributions, MCMC tuning and results can be found in [SOU 09a]. Here, we only illustrate the type of output that can be obtained, namely the posterior distributions of the infection times in 2004 of six different patches; see Figure 8.7. Each of the six distributions shows a typical pattern, from the patch that was certainly infected in the beginning of the growing season (patch 1) to the patch that certainly remained healthy until the end of the season (patch 6).

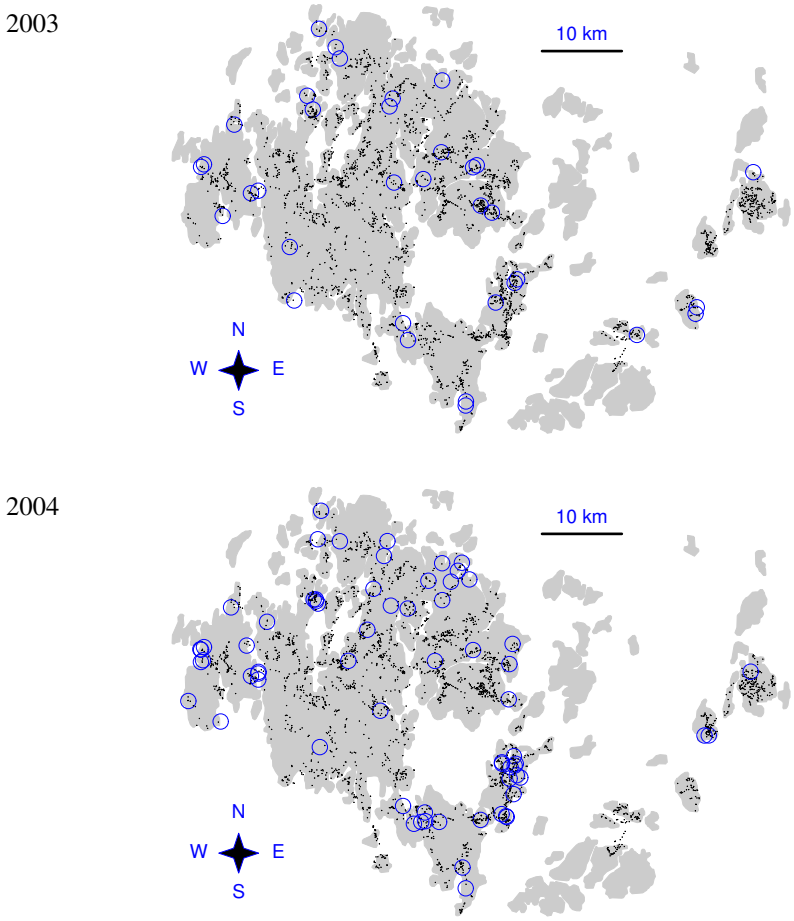

Figure 7.6: Map of the Åland Islands and patches of *Plantago lanceolata* that are healthy (dots) and infected (circles) in 2003 (top panel) and 2004 (bottom panel).

7.4. Stochastic Approaches for Modeling Spatial Trajectories

The study of animal movements informs on both individual behaviors of focal species and population-level dynamics. In particular, the characterization of territories used by individuals can be assessed via an estimation of the expected movements of animals, using discretely located data obtained at some given observation times. Many other application domains (e.g. physics of particles and transportation science) actually share the same questions regarding statistical inference of movements and trajectory reconstitution conditional on observations.

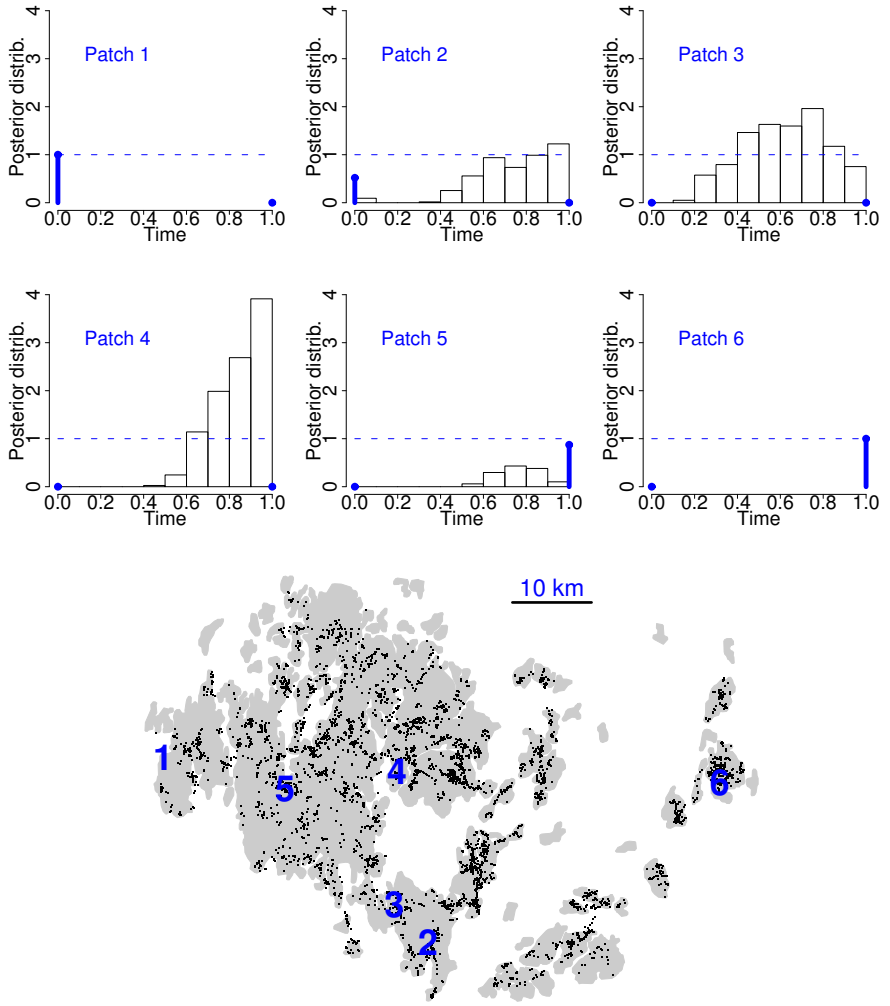

Figure 7.7: Zero-one-inflated posterior distributions of the infection times in year 2004 of six different patches (top panels). Locations of patches in the Åland Islands are indicated in the top panel. In each top panel, the dots at times zero and one give the posterior probabilities that the infection time is zero and one, respectively.

Various theoretical models for describing movements are available. Initially, continuous-time movements were often assumed to be simple Brownian motions [HOR 07], but then more and more general stochastic differential equations have

been proposed [IAC 08]. Other approaches consider movements in a discrete-time context, mainly using multivariate Markov chains.In this section and in connection with the topic of the book, we will only consider time-continuous processes. From the numeric and inferential point of view, several `R` packages are available for performing statistical analyses of trajectories (e.g. `Move`, `BBMM` and `MovementAnalysis`).

In what follows, we present the simple case of interpolating punctual observations along a d-dimensional Brownian motion giving rise to the so called Brownian bridge. Then, we show how one can use the stochastic machinery, namely the martingale machinery of predictable compensation for jumps, for building models of trajectories with jumps that can be viewed as PDMPs. We illustrate our approach by exhibiting the diversity of behaviors that elementary examples may exhibit.

Notation

We will assume that the continuous index set for processes is time. Naturally, depending on the topic, one can replace the time index by any other real variable that have a pertinent meaning with respect to the underlying dynamics. Scalar elements (either constants, functions or processes) will be denoted by capital letters (e.g., X), vector elements by bold letters ($\mathbf{x}$), and matrix elements by capital bold letters ($\mathbf{X}$). Moreover, note that random functions include deterministic ones, and that the term *process* is used with a generic meaning, whereas the term *sequence* denotes only discrete-time random processes.

7.4.1. *Conditioning a Brownian Motion by Punctual Observations*

Due to the lack of relevant knowledge or because of their characteristics, movements of animals or particles in spatial domains are often modeled as realizations of Brownian processes, which are viewed as reference models for trajectories. We recall that a standard d-dimensional Brownian motion $\mathbf{w}(t)$ in $\mathbb{R}^d$ simply consists of d independent copies of one-dimensional standard Brownian motions $W_i(t)$ with $W_i(0) = 0, i = 1, \cdots, d$. $\mathbf{w}(t)$ being Gaussian, it is entirely characterized by its first order moments: $E(\mathbf{w}(t)) = 0$ and $E(\mathbf{w}(t)\mathbf{w}^T(s)) = (t \wedge s)\mathbf{I}_d$ where $\mathbf{I}_d$ stands for the d-unit matrix.

Observations of a random processes $\mathbf{x}(t)$ representing a trajectory, even when they are dense in time, always yield a sequential data set $\mathbf{y}_n = \mathbf{x}(T_n^{\mathbf{y}})$ for observation time $T_n^{\mathbf{y}}$, $n = 1, 2, \ldots$. Assuming that these observation times are independent of the process, one can infer some statistical characteristics of $\mathbf{x}(t)$ and then take into account observations to simulate (i.e. reconstruct or interpolate) the non-observed part of the trajectory. In the case of the Brownian motion, the conditioning with respect to observations gives the so-called Brownian bridge.

Brownian bridge on $\mathbb{R}^d$

The Brownian bridge $X(t),\ t \in [0,1]$, in $\mathbb{R}$ is defined (in distribution) as a Brownian motion $W(t), t \in [0,1]$, conditional on the knowledge that at $t = 1$, $W(t) = 0$. A path-wise definition exists: $X(t) = W(t) - tW(1),\ t \in [0,1]$.

This definition can be straightly extended to any interval $[T_1, T_2]$. Using the specific properties of conditional expectation for Gaussian distribution, one can easily prove that conditionally on $\{W(T_1), W(T_2)\}$, the Brownian bridge $X(t), t \in [T_1, T_2]$, is a Gaussian process, with $E(X(t)) = \frac{W(T_1)(T_2-t)+W(T_2)(t-T_1)}{T_2-T_1}$ and $E(X(t)X(s)) = \frac{(T_2-t)(s-T_1)}{T_2-T_1} = C(t,s)$, independent of $W(T_1),\ W(T_2)$, for $T_1 \le s \le t \le T_2$.

In particular, $X(t)$ follows a Gaussian distribution with mean $\mu(t) = a_1 + \frac{t-T_1}{T_2-T_1}(a_2 - a_1)$ and variance $\sigma^2(t) = C(t,t)$ where $a_1 = W(T_1)$ and $a_2 = W(T_2)$.

A d-dimensional Brownian bridge $\mathbf{x}(t) = (X_1, \ldots, X_d)(t),\ t \in [T_1, T_2]$, with $\mathbf{x}(T_j) = \mathbf{a}_j = (a_{1,j}, \ldots, a_{d,j}) \in \mathbb{R}^d,\ j = 1, 2$, is defined as a vector of d independent Brownian bridges $X_i(t)$ with $X_i(T_j) = a_{i,j},\ j = 1, 2$. More explicitly, $\mathbf{x}(t)$ has a Gaussian density $\varphi(x|\mu(t), \Sigma(t))$ with mean $\mu(t) = \mathbf{a}_1 + \frac{t-T_1}{T_2-T_1}(\mathbf{a}_2 - \mathbf{a}_1)$ and covariance matrix $\Sigma(t) = \frac{(T_2-t)(s-T_1)}{T_2-T_1}\mathbf{I}_d$.

Brownian bridge with noisy extremal points

Due to measurement errors, the points $\mathbf{a}_j,\ j = 1, 2$, are generally random. If we assume these points to be independent with densities $f_{\mathbf{a}_j},\ j = 1, 2$, the distribution of $\mathbf{x}(t)$ can be written:

$$P\left(\mathbf{x}(t) \in \mathbb{D}\right) = \int_{\mathbb{R}^d \times \mathbb{R}^d} \left(\int_{\mathbb{D}} \varphi(x|\mu(t), \Sigma(t))dx \right) f_{\mathbf{a}_1}(u) f_{\mathbf{a}_2}(v) du dv,\ \mathbb{D} \subset \mathbb{R}^d.$$

In the case of Gaussian densities $f_{\mathbf{a}_j}(u) = \prod_{i=1}^{d} \varphi(u_i|a_{i,j}, \sigma_j^2),\ j = 1, 2$, the process $\mathbf{x}(t)$ remains Gaussian with mean $\mu(t, \mathbf{a}_1, \mathbf{a}_2) = \mathbf{a}_1 + \frac{t-T_1}{T_2-T_1}(\mathbf{a}_2 - \mathbf{a}_1)$ whereas its covariance matrix satisfies $\Sigma^*(t) = \sigma^{*2}(t)\mathbf{I}_d$ with

$$\sigma^*(t) = \frac{(T_2 - t)(t - T_1)}{T_2 - T_1} + \sigma_1^2 \left(\frac{T_2 - t}{T_2 - T_1} \right)^2 + \sigma_2^2 \left(\frac{t - T_1}{T_2 - T_1} \right)^2.$$

Mean occupation time

An important index in ecological studies is the mean occupation time of space domain $\mathbb{D}$ during a time interval $[t_1, t_2]$, which is defined as the non-negative random

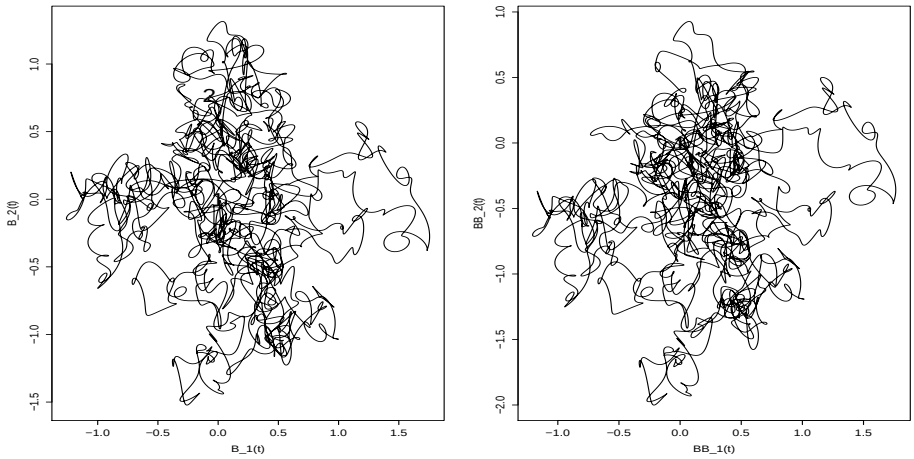

Figure 7.8: Paths in the plane of a standard Brownian motion starting at point (0,0) marked by "1" and arriving at an unconditioning point marked by "2" (left) and a standard Brownian bridge starting and arriving at point (0,0) marked by "1" (right).

variable $\tau_{\mathbb{D}} = \frac{1}{t_2-t_1}\int_{t_1}^{t_2} \mathbf{1}_{\{\mathbf{x}(t)\in\mathbb{D}\}}dt$. Its expectation $\nu(\mathbb{D}) = E(\tau_{\mathbb{D}})$ induces an absolutely continuous measure with density:

$$h(x) = \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} \varphi(x \mid \mu(t, \mathbf{a}_1, \mathbf{a}_2), \Sigma^*(t))dt.$$

Related statistical issues

For ecological and territory planing purposes, one can be interested in the estimation of the density $h(x)$ after collecting a set of observations $(T_j, \mathbf{x}(T_j) = \mathbf{a}_j),\ j = 1, \ldots, n+1$. Assume that these data are drawn from a d-dimensional Brownian motion with diffusion coefficient σ^2 and variances of measurement errors σ_j^2 depending on locations $\mathbf{a}_j$, and for $t \in [T_j, T_{j+1}]$, and let

$$\mu_j(t) = \mu(t, \mathbf{a}_j, \mathbf{a}_{j+1})$$

$$\sigma_j^*(t) = \sigma^2\frac{(T_{j+1}-t)(t-T_j)}{T_{j+1}-T_j} + \sigma_j^2\left(\frac{T_{j+1}-t}{T_{j+1}-T_j}\right)^2 + \sigma_{j+1}^2\left(\frac{t-T_j}{T_{j+1}-T_j}\right)^2.$$

Let the process $\mathbf{x}(t)$, $t \in [T_1, T_{n+1}]$, be formed by the set of independent Brownian bridges connecting $\mathbf{a}_j$ to $\mathbf{a}_{i+1}$ within time interval $[T_j, T_{j+1}]$, $j = 1, \ldots, n$. Then, the total mean occupation time of space has density

$$h(x) = \frac{1}{T_{n+1} - T_1} \sum_{j=1}^{n} \int_{T_j}^{T_{j+1}} \varphi(x \mid \mu_j(t), \Sigma_j^*(t)) dt. \qquad [7.24]$$

The variances of measurement errors σ_j^2 are generally specified and one only has to estimate the diffusion coefficient σ^2 to compute the occupation time density.

The following trick was used to build a simple conditional likelihood for data. Assume that n is even, then one can prove that observations $\mathbf{x}(T_{2k})$, $k = 1, \ldots, n/2$, conditional on the values of observations $\mathbf{x}(T_{2k-1})$, $k = 1, \ldots, n/2$, are independent Gaussian random vectors with mean vectors $\mu_{2k-1}(T_{2k})$ and covariance matrices $\Sigma^*_{2k-1}(T_{2k})$. Hence, we can get an estimate $\hat{\sigma}^2$ by maximizing the following likelihood:

$$\prod_{k=1}^{n/2} \varphi(\mathbf{a}_{2k} \mid \mu_{2k-1}(T_{2k}), \Sigma^*_{2k-1}(T_{2k})).$$

The estimation of density h can be performed with standard numerical methods approximating the integral form in Equation [8.24]. This approach was compared to kernel methods considering observed locations as i.i.d. random vectors drawn from h, and was proven to be much more efficient since it accounts for measurement errors and temporal dependencies between observed locations. Moreover, the first approach yields more realistic domains for level sets of h.

Extension to further movement dynamics

Beyond the Brownian bridge, there exists today a wide range of literature about more general (and more realistic) diffusion bridges in $\mathbb{R}^1$ and $\mathbb{R}^d$ related to some specific stochastic differential equations of the form:

$$d\mathbf{x}(t) = f(t, \mathbf{x})dt + \sigma(t, \mathbf{x})d\mathbf{w},$$

driven by a d-dimensionnal Brownian motion $\mathbf{w}(t)$ and a vectorial drift function f.

There are many results about the characterization (in distribution as well as in a path-wise sense) of these diffusions when they are considered conditionally on their values $\mathbf{x}(T_j) = \mathbf{a}_j$ at times T_j, $j = 1, 2$. These results are however more complicated to obtain since they are grounded on sophisticated tools such as the Girsanov theorem.

7.4.2. *Movements with Jumps, Including Mathematical Preliminaries*

Thereafter, we assume that there exists a complete probability space $(\Omega, \mathbb{F}, \mathbb{P})$ with a filtration (or history) $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ such that processes are $\mathbb{F}$-adapted, stopping times refer to $\mathbb{F}$ and martingales to $(\mathbb{F}, \mathbb{P})$. We shall neither develop the classical theory of the predictable σ-algebra nor insist on other definitions such as predictable processes. One has to know that a process with everywhere càglàd paths (i.e. left continuous with right limits) are predictable. We also adopt standard notations for a process $X(t)$ such as $X(t-) = \lim_{s\uparrow t} X(s)$ and $\Delta X(t) = X(t) - X(t-)$. For càdlàg processes X, the continuous part is defined as $X^c(t) = X(t) - \sum_{s\leq t} \Delta X(s)$.

Point processes and predictable projections

In what follows, an 1D temporal point process $N(t)$ is given by a strictly increasing sequence of stopping times $(T_i)_{i\geq 0}$ with the convention that $T_0 = 0$. The associated counting process is defined as $N(t) = \Sigma_{i>0} \mathbf{1}_{\{T_i\leq t\}}$. Under this definition N is adapted. Moreover, it is assumed to be a *simple* point process in the sense that all jumps are only 1-valued.

As an adapted increasing process, N is a submartingale (i.e. $E(N(t)|\mathcal{F}_s) \geq N(s)$; for all $t \geq s$) and by the Doob-Meyer Theorem [PRO 05], there exists an increasing predictable process $\tilde{N}(t)$ such that $M(t) = N(t) - \tilde{N}(t)$ is a martingale. $\tilde{N}(t)$ is called the predictable compensator of $N(t)$. Theoretically, it is defined as a conditional expectation with respect to the predictable σ-field. In most interesting cases, $\tilde{N}(t)$ is almost surely absolutely continuous with respect to the Lebesgue measure with a random density function $\lambda(t)$, called the intensity function, that is $\tilde{N}(t) = \int_0^t \lambda(s)ds$.

When the filtration reduces to natural history of the process $N(t)$, the intensity can be deduced as follows (see [DAL 88] for details): If regular versions $G_{i+1}(dt|\mathcal{F}_{T_i})$ of the conditional distribution functions of interval lengths $D_{i+1} = T_{i+1} - T_i$ exist, then $\tilde{N}(t) = \Sigma_{i>0}\, \Lambda_i(t)$, with

$$\Lambda_{i+1}(t) = \begin{cases} 0 & \text{if } t \leq T_i \\ \int_0^{(t-T_i)\wedge D_{i+1}} \frac{G_{i+1}(ds|\mathcal{F}_{T_i})}{1-G_{i+1}(s-|\mathcal{F}_{T_i})} & \text{if } t > T_i. \end{cases} \qquad [7.25]$$

If $G_i(dt|\mathcal{F}_{T_i}) \ll dt, i \geq 1$, then one can paste the different pieces into a single formula $\tilde{N}(t) = \int_0^t \lambda(s)ds$. As we shall see below, this expression is largely used in survival analysis, where $\Lambda_i(t)$ stands for the cumulative hazard function of the random variable D_i and its derivative is the hazard function.

For filtrations richer than the natural history, the calculation of the compensator is generally out of reach, but it conserves the same interpretation, namely the best

cumulative predictor of the jumps of N. Authors generally assume some specific forms for the intensity process grounded on relevant hypotheses for the application domain of interest, because in many interesting cases, the form of the compensator uniquely determines the probability distribution underlying the point process $N(t)$. For example a deterministic continuous compensator refers to a Poisson processes (see an example of sample paths for N and $\tilde{N}$, and the associated compensating process M in Figure 8.9).

A useful and *universal* property, under the natural history, is that a simple point process N with continuous and a.s. unbounded compensator $\tilde{N}$ undergoing the random time change $\tilde{N}^{-1}(t)$, yields a standard homogeneous Poisson process $N^*(t) = N(\tilde{N}^{-1}(t))$. A partial converse is that a standard Poisson process $N^*(t) = \sum_{i>0} \mathbf{1}_{\{T_i^* \leq t\}}$ and a positive function $\lambda(t)$ jointly give rise to a Poisson process $N(t) = \sum_{i>0} \mathbf{1}_{\{T_i \leq t\}}$ of intensity $\lambda(t)$ with $T_i = \int_0^{T_i^*} \lambda(s)ds = \Lambda(T_i^*)$.

Generalization to multivariate and marked point processes

A d-dimensional point process $\mathbf{N}(t) = (N_1, \ldots, N_d)(t)$ is defined similarly as above by a probability space with d sequences of stopping times (T_i^j), $j = 1, \ldots, d$, $i \geq 0$, with corresponding vectorial compensator $\tilde{\mathbf{N}}(t) = (\tilde{N}_1, \ldots, \tilde{N}_d)(t)$ and martingales $M^j(t) = (N^j - \tilde{N}^j)(t)$.

However, in the context of movements with random jump sizes, we need a wider generalization, namely the marked point processes and their dual predictable projections [JAC 75]. We avoid details of the theory by simply restricting our presentation of marked point processes within the following framework. A sequence of random vectors $(T_i, \varepsilon_i)_{i \geq 0}$ taking values in $\mathbb{R}_+ \times \mathbb{R}^d$ with $T_i < T_{i+1}$ defines a random measure $N(dt, dx) = \sum_i \delta_{(T_i, \varepsilon_i)}$. A stochastic machinery similar than above can be developed to enable us to assert that there exists a predictable random measure (on an extended probability space) $\tilde{N}(dt, dx)$ such that for every predictable process $Y(s, x)$, the process $M_Y(t) = \int_0^t \int_{\mathbb{R}^d} Y(s, x)(N - \tilde{N})(dt, dx)$ is a martingale.

In the case of a filtration corresponding to the natural history, a formulation similar to [8.25] gives the predictable projection, by replacing the previous conditional probability function $G_{i+1}(dt|\mathcal{F}_{T_i})$ by the distributions $G_{i+1}(dt \times dx|\mathcal{F}_{T_i})$ of the variable $(T_{i+1} - T_i, \varepsilon_{i+1})$ conditionally on $\mathcal{F}_{T_i}$. More precisely, we have $\tilde{N}(dt, dx) = \Sigma_{i>0} \Lambda_i(dt, dx)$, with

$$\Lambda_{i+1}(dt, dx) = \frac{G_{i+1}(dt - T_i, dx|\mathcal{F}_{T_i})}{G_{i+1}([t - T_i, \infty] \times \mathbb{R}^d|\mathcal{F}_{T_i})} \mathbf{1}_{\{T_i < t \leq T_{i+1}\}}.$$

Example 1

In the case of an 1D point process, let us assume that $G_{i+1}(ds|\mathcal{F}_{T_i})$ is the Weibull distribution $W(\alpha, \beta)$, with hazard function $h(t) = \beta \alpha^\beta t^{\beta - 1}$ and cumulative hazard

function $H(t) = (\alpha t)^\beta$. Parameters α and β are the scale and shape characteristics. According to Equation [8.25], the compensator is written $\tilde{N}(t) = \alpha^\beta(\sum_{i=1}^{n}(T_i - T_{i-1})^\beta + (t - T_n)^\beta)$ for $T_n < t \leq T_{n+1}$.

This compensator is stochastic since its expression depends on the stopping times T_i. Actually, N is a renewal process and it is not a Poisson process, unless $\beta = 1$ since for that case $\tilde{N}(t) = \alpha^\beta t$ is deterministic.

Example 2

We now consider the analogous Poisson process with intensity $\lambda(t) = \beta\alpha^\beta t^{\beta-1}$. Thus, for any interval $I = [\tau_1, \tau_2]$, the number of point events $N(I)$ in I is Poisson distributed with parameter $\Lambda(I) = \int_{\tau_1}^{\tau_2} \lambda(s)ds = \alpha^\beta(\tau_1^\beta - \tau_2^\beta)$. In addition, given the number of point events $N(I) = k$, its realization $\{Y_1, \ldots Y_k\}$ within I, are i.i.d random variables with probability density $g(t) = \beta\frac{t^{\beta-1}}{\tau_2^\beta - \tau_1^\beta}\mathbf{1}_{\{\tau_1 \leq t \leq \tau_2\}}$.

For simulation purpose, note that Y_j has the same distribution as $(U(\tau_2^\beta - \tau_1^\beta) + \tau_1^\beta)$, where U is uniformly distributed over $[0, 1]$. Note also that the time transformation $\Lambda^{-1}(t)$ makes $N^*(t) = N(\Lambda^{-1}(t)$ to be a standard Poisson process. Observing that $D_{i+1}^* = T_{i+1}^* - T_i^*$ is exponentially distributed with rate 1, one can prove that the inter-event length time $D_{i+1} = T_{i+1} - T_i$, conditionally on T_i (or T_i^*), has the following survival function:

$$S_{i+1}(t) = \exp\{-\alpha^\beta((T_i + t)^\beta - T_i^\beta)\}\ ,\ t \geq 0.$$

This formula states that T_{i+1} conditionally on the event $T_{i+1} > T_i$ behaves as a Weibull distributed random variable $Y \sim W(\alpha, \beta)$, conditioned by the event $Y > T_i$; this is the memory loss property of a Poisson process. Figure 8.9 illustrates sample paths for $N, \tilde{N}$ and M for parameter values $\alpha = 1$ and $\beta = 1.2$.

Stochastic integrals for purely discontinuous martingales

In the context of point processes, stochastic integration reduces to path-wise integrals (in the sense of Stieltjes-Lebesgue integrals for bounded variation integrands), but nevertheless requires care. For sake of completeness, let us first recall that a semimartingale $X(t)$ is defined by the identity $X(t) = M(t) + A(t)$, where $M(t)$ is a local martingale and $A(t)$ is a locally bounded variation process. For any semimartingale X, one can define its quadratic variation process $[X, X](t) = X^2(t) - 2\int_0^t X(s-)dX(s)$, which is also a locally bounded variation process and satisfies $\Delta[X, X](t) = (\Delta X(t))^2$. The continuous part of $[X, X]$ is defined by $[X, X]^c(t) = [X, X](t) - \sum_{0 \leq s \leq t}(\Delta X(s))^2$. The quadratic co-variation process of two semimartingales is defined by duality as $[X, Y](t) = ([X + Y, X + Y] - [X, X] - [Y, Y])/2$, and similarly satisfies $\Delta[X, Y](t) = \Delta X(t)\Delta Y(t)$.

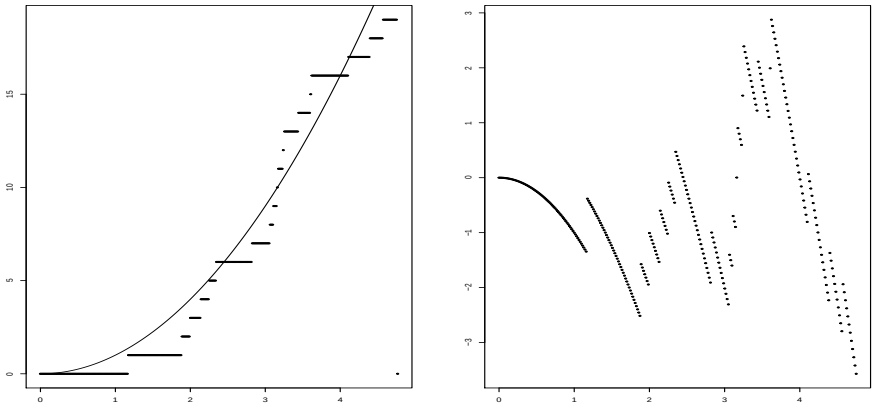

Figure 7.9: Left: Counting Poisson process $N(t)$ of Example 2 (broken line) and its compensator $\tilde{N}(t) = t^\beta$ (continuous line) with $\alpha = 1$ and $\beta = 1.2$. Right: The corresponding compensating martingale $M(t) = (N - \tilde{N})(t)$.

For a *simple* counting process N, we have $[N, N](t) = N(t)$. More generally $[X, X](t) = \sum_{s \leq 0} \Delta X^2(s)$ holds for any adapted process X with locally bounded variation, so that $[X, X](t) \equiv 0$ if in addition X is continuous. In fact, the machinery of stochastic calculus intervenes only when the martingale component M has a non-purely discontinuous part (i.e. $[M^c, M^c] \neq 0$).

Point processes, compensators and martingales

We recall that if $M(t) = N(t) - \tilde{N}(t)$ denotes the martingale compensating the jumps of a simple point process N, then any adapted, integrable predictable (in particular left continuous) $f(t)$, gives rise to a martingale $M_f(t) = \int_0^t f(s)dM(s)$. These processes are also purely discontinuous martingales and their quadratic co-variation processes satisfy the following formula $[M_f, M_g](t) = \int_0^t f(s)g(s)dN(s)$. As a by-product, we see that $[M_f, M_g](t)$ is compensated by $\int_0^t f(s)g(s)d\tilde{N}(s)$, so that for $t \geq s$, we have:

$$E(M_f(t)M_g(t) \mid \mathcal{F}_s) = E\left(\int_0^t f(u)g(u)d\tilde{N}(u) \mid \mathcal{F}_s\right).$$

This formula is particularly appealing for deterministic functions f and g and/or handy expressions of the compensator $\tilde{N}(dt)$ to explicitly calculate the covariance functions. Figure 8.10 shows two examples of 2D-trajectories whose coordinates are correlated martingales defined by stochastic integrals as above.

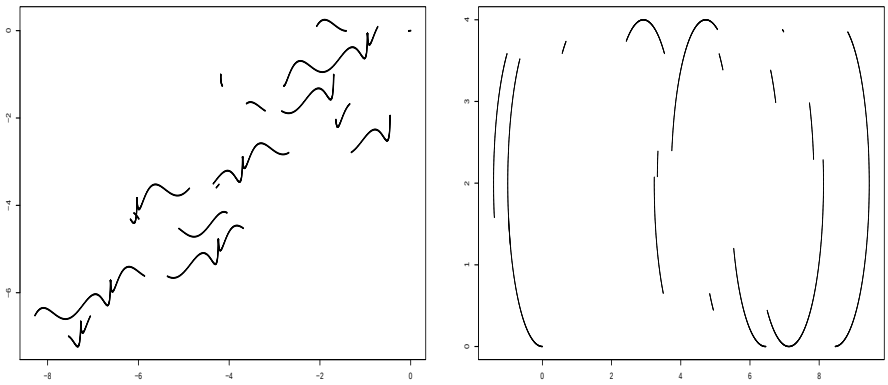

Figure 7.10: Realization of 2D martingales $X_i(t) = \int_0^t f_i(s)d(N - \tilde{N})(s)$, $i \in \{1, 2\}$, with $f_1(t) = 1 + \cos(t)$ and $f_2(t) = 0.5 - 2\sin(3t)$ (left panel), and $f_1(t) = \cos(0.3t)$ and $f_2(t) = 2\sin(0.3t)$ (right panel).

Example 3: Stochastic differential equations with impulsions

We now illustrate an other type of dynamical systems based on stochastic differential equations driven by compound point processes via a particular but nevertheless generic example for many dynamics.

We consider an autonomous system undergoing random shocks at random times. For $\mathbf{x} = (x_1, x_2)$, we consider the quadratic function $C(\mathbf{x}) = x_1^2 + \beta x_2^2$ on the plane $\mathbb{R}^2$. The level curves of C are either ellipses (when $\beta > 0$) or hyperbolas (when $\beta < 0$). This is obvious for $\beta > 0$. For $\beta < 0$, let $\beta = -\rho^2$, then the equation $C(\mathbf{x}) = c$ can be written $(x_1 - \rho x_2)(x_1 + \rho x_2) = c$, which reduces to $u_1 u_2 = c$ after a linear transformation.

Besides, using classical tricks for ordinary differential equations (ODE), one can prove that functions $\mathbf{x}(t)$ satisfying $C(\mathbf{x}(t)) = c$ are governed by the following homogeneous linear ODE:

$$\mathbf{x}'(t) = A\mathbf{x}(t), \text{ with } \mathbf{x}(0) = \mathbf{x}^* \qquad [7.26]$$

whose solution is $\mathbf{x}(t) = e^{At}\mathbf{x}^*$. More explicitly,

– if $\beta = \rho^2 > 0$, we have $A = \begin{pmatrix} 0 & -\rho^2 \\ 1 & 0 \end{pmatrix}$ and the solution of ODE [8.26] is:

$$\begin{aligned} x_1(t) &= x_1^* \cos(\rho t) - x_2^* \rho \sin(\rho t) \\ x_2(t) &= x_1^* \sin(\rho t)/\rho + x_2^* \cos(\rho t) \end{aligned}$$

– whereas for $\beta = -\rho^2 < 0$, we get $A = \begin{pmatrix} 0 & -\rho \\ -1/\rho & 0 \end{pmatrix}$, yielding the following solution of ODE [8.26]:

$$\begin{aligned} x_1(t) &= \quad x_1^*(e^t + e^{-t})/2 \quad + x_2^*\rho(-e^t + e^{-t})/2 \\ x_2(t) &= x_1^*(-e^t + e^{-t})/(2\rho) + \; x_2^*\rho(e^t + e^{-t})/2 \end{aligned}$$

Next, let us consider the marked point process $\sum_{i>0} \delta_{(T_i,\varepsilon_i)}$ in $\mathbb{R}_+ \times \mathbb{R}^2$, and the bi-dimensional stochastic differential equation (SDE):

$$d\mathbf{z}(t) = A(\mathbf{z}(t-))dt + dM(t) \qquad [7.27]$$

where $M(t) = \sum_{i>0} \varepsilon_i \mathbf{1}_{\{T_i \le t\}}$. The sequence $(\varepsilon_i)_{i\ge1}$ is formed by i.i.d. elements and is independent of $(T_i)_{i\ge1}$.

The solution of this SDE consists in a particle trajectory formed by a sequence of disjoint curve arcs, each being a solution of Eq. [8.26]: at random times T_i, the particle jumps by a size ε_i from its present orbit at a new location, initiates a new orbit, and so on.

Figure 8.11 illustrates sample paths for both ellipsoidal and hyperbolic orbits, depending on the sign of β, with standard Gaussian variables ε_i.

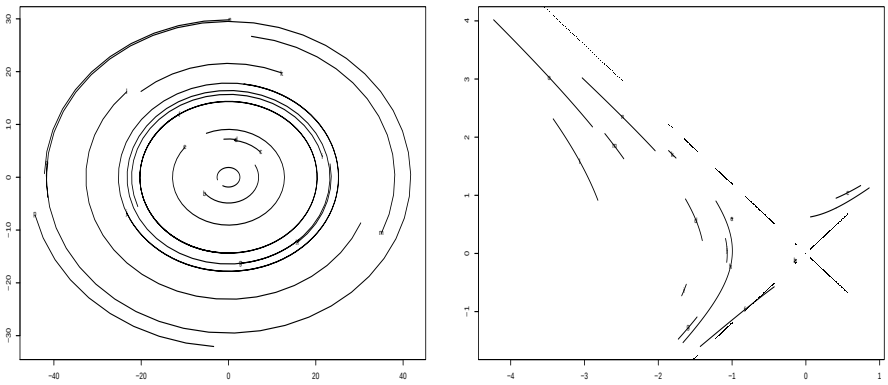

Figure 7.11: Realizations of path obtained with an SDE with jumps (Example 3). Left: Ellipsoidal orbits ($\beta = 2$). Right: Hyperbolic orbits ($\beta = -0.7$).

When are these stochastic differential equations PDMP?

As seen earlier in this chapter and the introductory chapter, piecewise deterministic Markov processes (PDMP) introduced by M.H.A Davis [DAV 84] enrich the usual classes of Markov Processes (diffusions, jump processes,...) by allowing a part of determinism in paths while inheriting the appealing Markovian properties [COS 08]. PMDP are time homogeneous $\mathbb{R}^d$-valued processes $\mathbf{x}(t)$ with $càdlàg$ sample paths.

PDMPs can be sequentially *constructed* via an increasing sequence of stopping times $(T_n)_{n\geq 0}$ with $T_0 = 0$ and $\mathbf{x}(0) = \mathbf{x}_0$. For $\mathbf{x}(T_n) = \mathbf{x}_n$ and $t \in [T_n, T_{n+1}[$, the process $\mathbf{x}(t)$ obeys a deterministic rule, e.g. an ODE $d\mathbf{x}(t) = V(\mathbf{x}(t))dt$, governed by a regular vector field V. Then, conditionally to the past $\mathcal{F}_{T_n}$, the lifetime $D_{n+1} = T_{n+1} - T_n$ has hazard function $\lambda(s) = h(\mathbf{x}(T_n + s))$, where h is a non-negative bounded measurable function on $\mathbb{R}^d$. At time T_{n+1}, the process $\mathbf{x}(t)$ jumps to a state $\mathbf{x}(T_{n+1}) = \mathbf{x}(T_{n+1}-) + \varepsilon_{n+1}$, in accordance with a probability transition $Q(d\varepsilon|\mathbf{x}(T_{n+1}-))$. The triplet (V, h, Q) characterizes entirely the probability distribution of the PMDP. Note that when h satisfies $< \nabla h(\mathbf{x}), F(\mathbf{x}) >\equiv 0$, for all $\mathbf{x}$, i.e. h is a first integral for dynamical system, then $\lambda(s)$ is constant on the deterministic parts of paths and therefore the D_n are exponentially distributed.

Piecewise deterministic processes presented in this paper are based on a little more general marked point processes $N(dt, dx) = \sum_{i\geq 0} \delta_{(T_i,\varepsilon_i)}$ and so are neither Markovian nor time homogeneous in general and, therefore, are not PDMP in general. For the class of processes developed here to be PDMP, it is sufficient that the conditional cumulative intensities are separable measures in dt and dx and have the following form:

$$\Lambda_i(dt - T_i, dx) = h(\mathbf{x}(T_i + t))dt \times Q(dx|\mathbf{x}(T_{n+1}-)).$$

7.4.3. *The Doléans Dade Exponential Semimartingales*

The following theorem is borrowed from Protter [PRO 05] and is a consequence of the change of variables theorem as regards to Ito calculus for semimartingales.

THEOREM 7.1.– If X is a semimartingale with $X(0) = 0$, then there exists a unique semimartingale Z satisfying the equation $dZ(t) = Z(t-)dX(t)$, with $Z(0) = 1$ which is given by:

$$Z(t) = \exp^{(X(t)-\frac{1}{2}[X,X]^c(t))} \prod_{s\leq t} \left((1 + \Delta X(s)) \exp^{-\Delta X(s)}\right). \qquad [7.28]$$

The solution $Z(t)$, usually denoted $\mathcal{E}_X(t)$, is called the stochastic (or Doléans Dade) exponential of X. This theorem encompasses many useful results and

applications. The formula reduces a lot for locally bounded variation processes X, since in this case $[X, X]^c(t) \equiv 0$ implies that

$$Z(t) = \exp^{X^c(t)} \prod_{s \leq t} (1 + \Delta X(s))) .$$

Under a mild integrability condition, if $X(t)$ is a martingale, then $Z(t)$ is also a martingale. A multivariate version of the theorem exists [JAC 82] and corresponds to the analog of deterministic linear differential equations $dZ(t) = Z(t-)dX(t)$ with a matrix process A and a vector semimartingale X.

In what follows, we present several applications of exponential martingales.

Example 4: Deterministic semimartingales

Theorem 8.1 includes extensions of the case of deterministic homogeneous linear differential equations. For instance, for any $d \times d$ matrix A, there exists a unique solution $Z(t) = \exp^{At} z_0$ to equation $dZ(t) = AZ(t)dt$, with $Z(0) = z_o \in \mathbb{R}^d$, taking here the deterministic matrix semimartingale $X(t) = At$.

Example 5: Cumulative hazard function

The probability distribution function $F(t) = P(T \leq t)$ and the survival function $S(t) = 1 - F(t)$ of a non-negative random variable T, with $dS(t) = -dF(t)$, are monotonic functions and have bounded variations. The cumulative hazard function $\Lambda(t) = \int_0^t \frac{dF(s)}{1-F(s-)}$ satisfies the equation $d\Lambda(t)(1 - F(s-)) = d(F(s))$. Conversely, given a positive increasing function Λ with $\Lambda(0) = 0$, there exists a unique function S with $S(0) = 0$, which satisfies $dS(t) = -S(t-)d\Lambda(t)$. Equation [8.28] implies that S satisfies:

$$S(t) = 1 - F(t) = \exp^{-\Lambda^c(t)} \prod_{s \leq t} (1 - \Delta\Lambda(s)) .$$

Note that the absolutely-continuous case $d\Lambda(t) = \lambda(s)ds$ yields $S(t) = e^{-\int_0^t \lambda(s)ds}$.

Example 6 : Survival analysis

Survival analysis in statistics is based on the simple case of a simple point process with at most one event at time T. Let $S(t)$ and $\Lambda(t)$ be respectively the survival and cumulative hazard functions of T; then according to formula [8.25], the associated compensating martingale is written $M(t) = \mathbf{1}_{\{T \leq t\}} - \int_0^t \mathbf{1}_{\{s<T\}} \lambda(s)ds = \mathbf{1}_{\{T \leq t\}} - \Lambda(t \wedge T)$.

Since M is a pure jump martingale, with $[M, M]^c(t) \equiv 0$ and $M^c(t) = -\Lambda(t \wedge T)$, its exponential is also a pure jump martingale and satisfies: $Z(t) = \exp^{-\Lambda(t \wedge T)} \left(1 + \mathbf{1}_{[T,\infty[}(t)\right) = S(t \wedge T) \left(1 + \mathbf{1}_{[T,\infty[}(t)\right)$.

For statistical purposes, we have however to deal with a little more sophisticated exponential martingale. Assume for example that T has hazard functions $\lambda_0(t)$ under probability $\mathbb{P}_0$ and $\lambda_\theta(t)$ under probability $\mathbb{P}_\theta$, such that $\lambda_\theta(t) = \mu_\theta(t)\lambda_0(s)$. Now, if we consider the $\mathbb{P}_0$ martingale $X_\theta(t) = \int_0^t (\mu_\theta(s) - 1) dM_0(s)$, we find that its stochastic exponential $Z_\theta(t) = \exp^{\int_0^t \log(\mu_\theta(s)) dN(s) - \int_0^t (\mu_\theta(s)-1)\lambda_\theta(s) ds}$ is also a $\mathbb{P}_0$ martingale that exactly corresponds to the likelihood ratio $L_\theta(t) = E(\frac{d\mathbb{P}_\theta}{d\mathbb{P}_0}|\mathcal{F}_t) = f_\theta(t)^{\mathbf{1}_{\{T \leq t\}}}(1 - F_\theta(t))^{\mathbf{1}_{\{T>t\}}}$.

This construction is in fact a major key for dealing with more general statistical contexts (see Section 8.4.4).

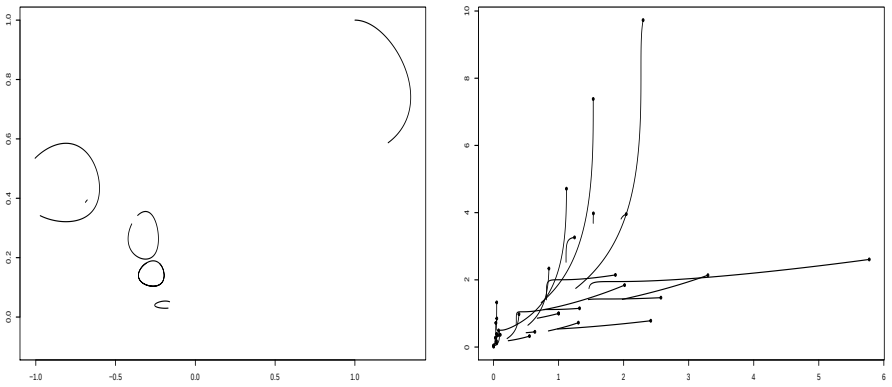

Figure 7.12: Sample paths of 2D stochastic exponentials $Z(t) = \mathcal{E}(M(t))$ (see Equation [8.28]) driven by compensating martingales $M_j(t); j = 1, 2$, based on a standard Poisson process $N(t)$. Left: $M_1(t) = \int_0^t \cos(s)(dN(s) - ds)$ and $M_2(t) = -\int_0^t \sin(s)(dN(s) - ds)$. Right: $M_1(t) = \int_0^t (1 + \cos(s))(dN(s) - ds)$ and $M_2(t) = \int_0^t (1 - \sin(s))(dN(s) - ds)$.

7.4.4. *Statistical Issues*

General case

We present here an important application of the exponential semimartingale theorem allowing a statistical approaches for marked point processes and related models such as PDMPs. It is a sort of Girsanov theorem characterizing the ratio of probability measures. Given two equivalent probability measures $\mathbb{P}$ and $\mathbb{Q}$ on some complete filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t \leq 0}$, we already knows that the Radon-Nycodym

derivative $Z(\omega) = \frac{d\mathbb{P}}{d\mathbb{Q}}$ is a positive $\mathbb{Q}$-integrable random variable implying therefore that the process $Z(t) = E(Z|\mathcal{F}_t)$ is a positive uniformly integrable martingale that equals $\frac{d\mathbb{P}_{|\mathcal{F}_t}}{d\mathbb{Q}_{|\mathcal{F}_t}}$, such that $Z(t)$ corresponds to a ratio of likelihoods.

In a statistical framework, considering a parametric set of probabilities $(\mathbb{P}_\theta, \theta \in \Theta)$ equivalent to $\mathbb{Q}$, such that $d\tilde{N}_\theta(t) = \mu_\theta(t)d\tilde{N}(t)$, where $\tilde{N}_\theta$ and $\tilde{N}$ are the respective compensators of N, one may expect to find , under some mild conditions, a particular $\mathbb{Q}$-martingale $W_\theta(t) = \int_0^t \rho_\theta(s)d(N - \tilde{N})(t)$ such that the likelihood ratio $Z_\theta(t)$ corresponds to the positive stochastic exponential of $W_\theta(t)$. Indeed, one can promptly and heuristically prove that it is true (and only true) for the process $\rho_\theta(t) = \lambda_\theta(t) - 1$.

In the context of multivariate/marked point processes, Jacod [JAC 75] gives a plain formula for the Radon-Nycodym derivative $\frac{d\mathbb{P}_{|\mathcal{F}_t}}{d\mathbb{Q}_{|\mathcal{F}_t}}$ under the natural filtration. This formula corresponds to the solution of the Doléan Dade equation for the martingale $W_\theta(t)$.

An Example

The statistical approach proposed above is applied here to the process presented in Example 3 of Section 8.4.2, which is piecewise driven by an ODE and randomly jumps at times T_i with jump amplitudes ε_i; in other words, this process statisfies the stochastic differential equation: $dX(t) = V(X(t-))dt + \sum_i \varepsilon_i \delta_{T_i}$.

Among the many potential measurements of the movement (eg length, kinetic energy,...), let us take the travel length L as a movement characteristic of a particle on orbits. For a particle starting from $\mathbf{x}_0$ at time $t = 0$, this is defined by:

$$L(t, X_0) = \int_0^t |V((X(s))|ds.$$

On one hand, let us assume that the random measure $N(dt, dx) = \sum_i \delta_{(T_i,\varepsilon_i)}$ has under probability $\mathbb{Q}$ the conditional intensities:

$$\Lambda_i(dt, dx \mid \mathcal{F}_{T_i}) = dt\mathbf{1}_{\{T_i \le t < T_{i+1}\}} \times \varphi(x|\mathbf{0}, \mathbf{I}_2)dx,$$

where $\varphi(x|\mathbf{m}, \Sigma)$ stands for the Gaussian density with mean $\mathbf{m}$ and covariance Σ in $\mathbb{R}^2$. In that case we obtain $\tilde{N}(dt, dx) = dt\varphi(x|\mathbf{0}, \mathbf{I}_2)dx$, meaning that $N(dt, dx)$ is a Poisson measure under $\mathbb{Q}$.

On the other hand, let us assume that under $\mathbb{P}_\theta$, the conditional intensities depends on paths as follows:

$$\Lambda_{i,\theta}(dt, dx \mid \mathcal{F}_{T_i}) = \alpha h_{\beta_1}(t - T_i, X(T_i))\mathbf{1}_{\{T_i \le t < T_{i+1}\}}dt \times \varphi(x|\mathbf{m}_\theta(t), \mathbf{I}_2)dx,$$

where for $T_i \leq t < T_{i+1}$, $\theta = (\alpha, \gamma, \beta_1, \beta_2)$ and $\gamma = (\gamma_1, \gamma_2)$, we define:

$$h_{\beta_1}(s, X_0)) = \frac{d}{ds} L^{\beta_1}(s, X_0)$$
$$\mathbf{m}_\theta(t) = \gamma\, L^{\beta_2}(t - T_i, X(T_i)).$$

The function $h_\beta(s, X_0)$, should be interpreted as the hazard function of the Weibull distribution $W(1, \beta)$ related to the positive travel length variable L on the orbit starting from X_0.

The previous equations ultimately tells that $\tilde{N}_\theta(dt, dx) = \lambda_\theta(t, x)\tilde{N}_\theta(dt, dx)$, with

$$\lambda_\theta(t, x) = \alpha \sum_{i \geq 0} h_{\beta_1}(t - T_i, X(T_i))$$
$$\exp^{-\frac{1}{2}[<\gamma,\gamma> L^{2\beta_2}(t-T_i, X(T_i)) - 2<x,\gamma> L^{\beta_2}(t-T_i, X(T_i))]}\, \mathbf{1}_{\{T_i \leq t < T_{i+1}\}}.$$

Next, for the sake of simplicity, let us suppose that the process is observed in the random time interval $[0, T_n]$, such that the likelihood ratio corresponds to the stochastic exponential of the $\mathbb{Q}$ martingale $W_\theta(t) = \int_0^t \int_{\mathbb{R}^2} (\lambda_\theta(s, x) - 1)(N - \tilde{N})(ds, dx)$. According to formula [8.28], the log-likelihood is equal to:

$$\begin{aligned}
\log(Z_\theta(T_n)) &= -\int_0^{T_n} \int_{\mathbb{R}^2} (\lambda_\theta(s, x) - 1)\tilde{N}(ds, dx) \\
&\quad + \int_0^{T_n} \int_{\mathbb{R}^2} \log(\lambda_\theta(s, x))N(ds, dx). \\
&= -\alpha \left(\sum_{i=0}^{n-1} L^{\beta_1}(T_{i+1} - T_i, X(T_i)) \right) - T_n + n \log(\alpha) \\
&\quad + \sum_{i=0}^{n-1} \log(h_{\beta_1}(T_{i+1} - T_i, X(T_{i+1}))) \\
&\quad - \frac{1}{2} < \gamma, \gamma > \sum_{i=0}^{n-1} L^{2\beta_2}(T_{i+1} - T_i, X(T_{i+1})) \\
&\quad + \sum_{i=0}^{n-1} < \gamma, \Delta X(T_{i+1}) > L^{\beta_2}(T_{i+1} - T_i, X(T_{i+1})).
\end{aligned}$$

One can therefore easily derive the set of equations for the maximum likelihood estimate (MLE) $\hat{\theta}$ and apply a classical optimization procedure. As an illustration, we

deal here with the simple case where the parameters β_1 and β_2 are known, which allows us to get explicit formulas for the MLE of α and $\gamma = (\gamma_1, \gamma_2)$:

$$\hat{\alpha} = \frac{n}{\sum_{i=0}^{n-1} L^{\beta_1}(T_{i+1} - T_i, X(T_i))}$$
$$\hat{\gamma} = \frac{\sum_{i=0}^{n-1} \Delta X(T_{i+1}) L^{\beta_2}(T_{i+1} - T_i, X(T_{i+1}))}{\sum_{i=0}^{n-1} L^{2\beta_2}(T_{i+1} - T_i, X(T_{i+1}))}.$$

As a perspective, one can expect to use asymptotic techniques for discrete time indexed martingales in order to derive the asymptotic behaviors (in almost sure and in distribution senses) of these estimators and, therefore, perform sensible null hypothesis testing such as $\gamma = 0$ and $\alpha = \alpha_0$.

7.5. Conclusion

This chapter gave an introduction to spatio-temporal PDMPs used to model population dynamics. Spatio-temporal PDMPs offer the possibility to build flexible models and achieve relatively realistic and consistent inferences. Thus, we presented three different modeling frameworks corresponding to three resolutions, namely the population, the metapopulation and the individual. We have seen that, depending on the dynamics of interest, the jumps in the PDMP can correspond to long-distance dispersal events, new introductions, or significant shifts in individual behaviors.

In the examples of models presented above, the spatio-temporal dependencies are contained in the flow function, whereas jumps are independent and identically distributed. However, for populations whose individuals can be transported in groups [SOU 11, SOU 14a], jumps should be correlated in space and time. For instance, in the metapopulation model of Section 8.3, a source patch could release a group of spores transported by wind towards a set of nearby patches. Such a process could lead to the simultaneous infection of several aggregated patches. Hence, developing PDMPs with dependent random jumps would be interesting for better taking into account specificities of some population dynamics. Moreover, it would be also challenging from the perspective of model construction, simulation and inference.

Acknowledgements

This work was partially funded by the HORIZON 2020 XF-ACTORS Project SFS-09-2016 and the INRA-DGAL Xylella Project n° 21000679.

7.6. Bibliography

[ALF 13] ALFARO M., COVILLE J., RAOUL G., "Travelling waves in a nonlocal reaction-diffusion equation as a model for a population structured by a space variable and a phenotypic trait", *Communications in Partial Differential Equations*, vol. 38, p. 2126–2154, 2013.

[BOR 17] BORDIER C., DECHATRE H., SUCHAIL S., PERUZZI M., SOUBEYRAND S., PIOZ M., PÉLISSIER M., CRAUSER D., LE CONTE Y., ALAUX C., "Colony adaptive response to simulated heat waves and consequences at the individual level in honeybees (*Apis mellifera*)", *Scientific Reports*, vol. 7, 2017.

[CAI 17] CAILLERIE N., Stochastic and deterministic kinetic equations in the context of mathematics applied to biology, Theses, Université de Lyon, July 2017.

[CHA 06] CHAMPAGNAT N., FERRIÈRE R., MÉLÉARD S., "Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models", *Theoretical population biology*, vol. 69, p. 297–321, 2006.

[COL 04] COLAUTTI R. I., RICCIARDI A., GRIGOROVICH I. A., MACISAAC H. J., "Is invasion success explained by the enemy release hypothesis?", *Ecology Letters*, vol. 7, p. 721–733, 2004.

[COS 08] COSTA O. L. V., DUFOUR F., "Stability and ergodicity of piecewise deterministic Markov processes", *SIAM Journal on Control and Optimization*, vol. 47, p. 1053–1077, 2008.

[DAL 88] DALEY D. J., VERE-JONES D., *An introduction to the theory of point processes*, Springer, 1988.

[DAV 84] DAVIS M. H. A., "Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models", *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 46, p. 353–388, 1984.

[DIG 83] DIGGLE P. J., *Statistical Analysis of Spatial Point Patterns*, Academic Press, London, 1983.

[EVA 98] EVANS L. C., *Partial differential equations*, Graduate studies in mathematics, American Mathematical Society, Providence, R.I, 1998.

[FIS 37] FISHER R. A., "The wave of advance of advantageous genes", *Annals of Eugenics*, vol. 7, p. 355–369, 1937.

[FIS 95] FISHER N. I., *Statistical Analysis of Circular Data*, Cambridge University Press, 1995.

[GAT 96] GATENBY R. A., GAWLINSKI E. T., "A reaction-diffusion model of cancer invasion", *Cancer research*, vol. 56, p. 5745–5753, 1996.

[GLO 15] GLOAGUEN P., ETIENNE M.-P., LE CORFF S., Stochastic differential equation based on a multimodal potential to model movement data in ecology, working paper or preprint, September 2015.

[HAD 75] HADELER K., ROTHE F., "Travelling fronts in nonlinear diffusion equations", *Journal of Mathematical Biology*, vol. 2, p. 251–263, 1975.

[HEN 89] HENGEVELD R., *Dynamics of biological invasions*, Springer Science & Business Media, 1989.

[HOR 07] HORNE J. S., GARTON E. O., KRONE S. M., LEWIS J. S., "Analyzing animal movements using Brownian bridges", *Ecology*, vol. 88, p. 2354–2363, 2007.

[IAC 08] IACUS S. M., *Simulation and inference for stochastic differential equations with R examples*, Springer Science & Business Media, New York, 2008.

[ILL 08] ILLIAN J., PENTTINEN A., STOYAN H., STOYAN D., *Statistical Analysis and Modelling of Spatial Point Patterns*, Wiley, 2008.

[JAC 75] JACOD J., "Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales", *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 31, p. 235–253, 1975.

[JAC 82] JACOD J., "Équations différentielles stochastiques linéaires: la méthode de variation des constantes", *Séminaire de Probabilité (Strasbourg)*, vol. 16, p. 442–446, 1982.

[KOL 37] KOLMOGOROV I., PETROVSKY A., PISCOUNO N., "Etude de l'équation de la diffusion avec croissance de la quantité de la matière et son application à un problème biologique", *Moscow University Bull. Math*, vol. 1, p. 1–25, 1937.

[LEW 93] LEWIS M., KAREIVA P., "Allee dynamics and the spread of invading organisms", *Theoretical Population Biology*, vol. 43, p. 141–158, 1993.

[LUT 06] LUTHER R., "Räumliche Ausbreitung chemischer Reaktionen", *Zeitschrift für Elektrochemie*, vol. 12, p. 596–600, 1906.

[MÉL 11] MÉLÉARD S., "Random modeling of adaptive dynamics and evolutionary branching", *The Mathematics of Darwin's Legacy*, p. 175–192, 2011.

[MUR 96] MURRAY J., KULESA P., "On a dynamic reaction–diffusion mechanism: The spatial patterning of teeth primordia in the alligator", *Journal of the Chemical Society, Faraday Transactions*, vol. 92, p. 2927–2932, 1996.

[OVA 04] OVASKAINEN O., HANSKI I., "Metapopulation dynamics in highly fragmented landscapes", HANSKI I., GAGGIOTTI O., Eds., *Ecology, Genetics, and Evolution of Metapopulations*, Chapter 4, p. 73-103, Elsevier, 2004.

[PRO 05] PROPTER P., *Stochastic Integration and Differential Equations: a New Approach*, Springer, Berlin, 2 edition, 2005.

[RIC 00] RICHARDSON D. M., PYŠEK P., REJMANEK M., BARBOUR M. G., PANETTA F. D., WEST C. J., "Naturalization and invasion of alien plants: concepts and definitions", *Diversity and distributions*, vol. 6, p. 93–107, 2000.

[ROB 99] ROBERT C. P., CASELLA G., *Monte Carlo Statistical Methods*, Springer, New York, 1999.

[ROQ 10] ROQUES L., HAMEL F., FAYARD J., FADY B., KLEIN E., "Recolonisation by diffusion can generate increasing rates of spread", *Theoretical population biology*, vol. 77, p. 205–212, 2010.

[ROQ 11] ROQUES L., SOUBEYRAND S., ROUSSELET J., "A statistical-reaction–diffusion approach for analyzing expansion processes", *Journal of Theoretical Biology*, vol. 274, p. 43–51, 2011.

[ROQ 13] ROQUES L., *Modèles de réaction-diffusion pour l'écologie spatiale: [avec exercices dirigés]*, Éd. Quae, Versailles, 2013, OCLC: 951442261.

[ROT 81] ROTHE F., "Convergence to pushed fronts", *The Rocky Mountain Journal of Mathematics*, p. 617–633, 1981.

[SHI 95] SHIGESADA N., KAWASAKI K., TAKEDA Y., "Modeling stratified diffusion in biological invasions", *The American Naturalist*, vol. 146, p. 229–251, 1995.

[SHI 97] SHIGESADA N., KAWASAKI K., *Biological invasions: theory and practice*, Oxford University Press, UK, 1997.

[SKE 51] SKELLAM J. G., "Random dispersal in theoretical populations", *Biometrika*, vol. 38, p. 196–218, 1951.

[SOU 07] SOUBEYRAND S., ENJALBERT J., SANCHEZ A., SACHE I., "Anisotropy, in density and in distance, of the dispersal of yellow rust of wheat: Experiments in large field plots and estimation", *Phytopathology*, vol. 97, p. 1315-1324, 2007.

[SOU 09a] SOUBEYRAND S., LAINE A., HANSKI I., PENTTINEN A., "Spatio-temporal structure of host-pathogen interactions in a metapopulation", *The American Naturalist*, vol. 174, p. 308-320, 2009.

[SOU 09b] SOUBEYRAND S., NEUVONEN S., PENTTINEN A., "Mechanical-statistical modelling in ecology: from outbreak detections to pest dynamics", *Bulletin of Mathematical Biology*, vol. 71, p. 318-338, 2009.

[SOU 11] SOUBEYRAND S., ROQUES L., COVILLE J., FAYARD J., "Patchy patterns due to group dispersal", *Journal of Theoretical Biology*, vol. 271, p. 87-99, 2011.

[SOU 14a] SOUBEYRAND S., MRKVIČKA T., PENTTINEN A., "A Nonstationary Cylinder–Based Model Describing Group Dispersal in a Fragmented Habitat", *Stochastic Models*, vol. 30, p. 48–67, 2014.

[SOU 14b] SOUBEYRAND S., ROQUES L., "Parameter estimation for reaction-diffusion models of biological invasions", *Population ecology*, vol. 56, p. 427–434, 2014.

[STO 95] STOYAN D., KENDALL W. S., MECKE J., *Stochastic Geometry and its Applications, 2nd Ed.*, Wiley, Chichester, 1995.

[TUR 03] TURCHIN P., *Complex population dynamics: a theoretical/empirical synthesis*, vol. 35, Princeton University Press, 2003.

[WEI 90] WEI G. C. G., TANNER M. A., "Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms", *Journal of the American Statistical Association*, vol. 85, p. 699-704, 1990.

[WIK 03a] WIKLE C. K., "Hierarchical Bayesian models for predicting the spread of ecological processes", *Ecology*, vol. 84, p. 1382–1394, 2003.

[WIK 03b] WIKLE C. K., "Hierarchical models in environmental science", *International Statistical Review*, vol. 71, p. 181–199, 2003.

Index