Hawkes processes with random environment

Felix Cheysson
Work in progress with Paul Bastide, Gabriel Lang, Sylvain Le Corff, Marie Perrot-Dockes, also Richard Aoun, Léo Micollet, and Viet Chi Tran.

Université Gustave Eiffel, CNRS, UMR 8050, **LAMA**.

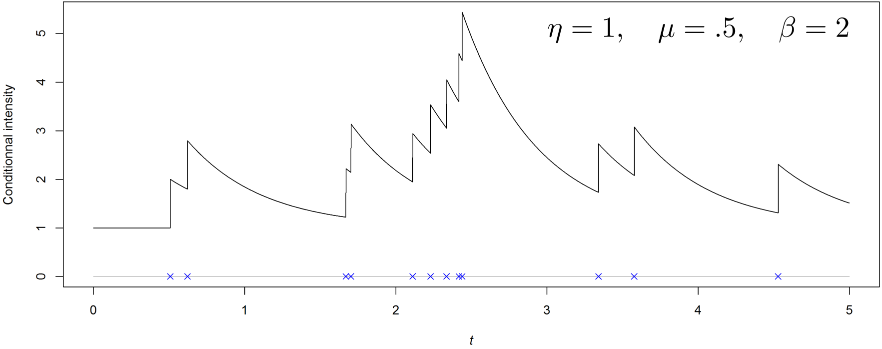
Statistiques au sommet de Rochebrune
March 23rd 2024

Hawkes process

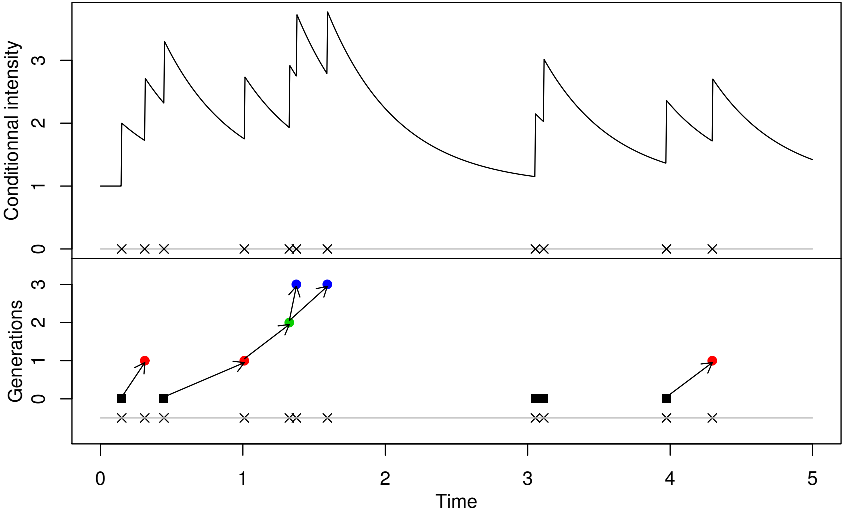
Linear Hawkes process on the real half-line

With exponentially decaying intensity:

$$\lambda(t) = \eta + \mu \int_{-\infty}^{t} \beta e^{-\beta(t-s)} N(\mathrm{d}s).$$



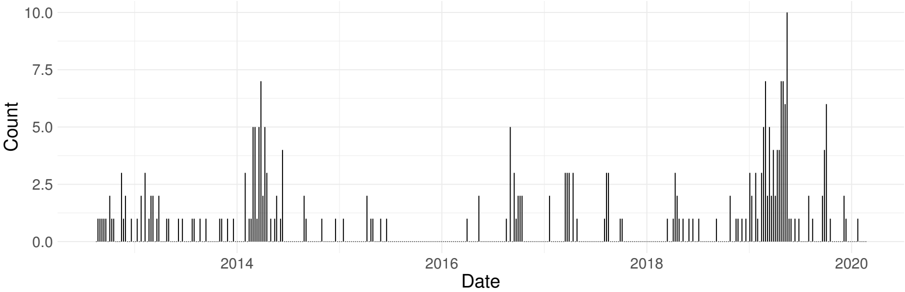
Hawkes process as a branching process



In the literature

- Self-exciting and clustering properties.
- Many disciplines of application:
  - seismology (Adamopoulos, 1976), neurophysiology, finance (Bacry, Mastromatteo, and Muzy, 2015), genomics (Reynaud-Bouret and Schbath, 2010), epidemiology, etc.
  - general review (Reinhart, 2018).
- Interesting properties:
  - Poisson cluster process: each cluster is a continuous-time Galton-Watson tree (Hawkes and Oakes, 1974).
  - Martingale properties of $N(t) - \int_0^t \lambda(s)\mathrm{d}s$ and $\left(N(t) - \int_0^t \lambda(s)\mathrm{d}s\right)^2 - \int_0^t \lambda(s)\mathrm{d}s$.
  - Erlang kernel $\rightarrow$ piecewise deterministic Markov process (Duarte, Löcherbach, and Ost, 2019).

Case-study: transmission of Measles in Tokyo[1]



Gaussian reproduction kernel: $h(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\nu)^2}{2\sigma^2}\right)$

- $\widehat{\nu} = 9.8$ days, $\widehat{\sigma} = 5.9$ days

Epidemiology (Centers for Disease Control and Prevention, 2015)

Incubation period: 10-12 days after exposure.

Transmission period: 4 days before to 4 days after rash onset.

[1] https://www.niid.go.jp/niid/en/survaillance-data-table-english.html

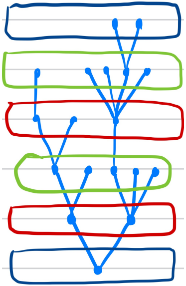
Our motivation

- In epidemiology, $\mu \sim \mathcal{R}_0$ (basic reproduction number).
- **Problem**: μ fixed, must be < 1 for sub-criticality of the process.
- Extension possible in the framework of local stationarity (Dahlhaus, 1996; Roueff, Sachs, and Sansonnet, 2016).
  - Strict upper limit $\sup_{t \in \mathbb{R}} \{\mu_t\} < 1$.

- Galton-Watson with random environment (Smith and Wilkinson, 1969; Athreya and Karlin, 1971).
  - Reproduction law different between each generation.
  - Extinction of the process almost sure if

$$\mathbb{E}\left[\log \mu(\omega)\right] \leq 0,$$

with $\mu(\omega)$ the mean reproduction at a given generation.



Outline

Hawkes process with random environment (HPRE)

- Usually, random environment with Hawkes process is meant as

$$\lambda(t) = \eta + \mu \int_0^t h(t-s) N(\mathrm{d}s) + \int_0^t \sigma(s) \mathrm{d}B_s.$$

  - CSBPs with random environments obtained as scaling limits of BPREs (Bansaye and Simatos, 2015; Bansaye, Caballero, and Méléard, 2019).
  - No direct dependency on further generations.
  - Strong assumptions of exponential kernel $h \equiv \exp$.
- Existing works (Dassios and Zhao, 2011; Lee, Lim, and Ong, 2016) as

$$\lambda(t) = \eta + \int_0^t X_s h(t-s) N(\mathrm{d}s).$$

  - No systematic characterisation of the process (existence, stationarity) when the reproduction number increases above 1.
  - Assumption of exponential kernel $h \equiv \exp$.

Existence of HPREs

Our framework for HPREs:

$$\lambda(t) = \eta + \int_0^t \phi(X_s) h(t-s) N(\mathrm{d}s),$$

with $\eta > 0$, $\int h = 1$, and ϕ a non-negative function.

Proposition

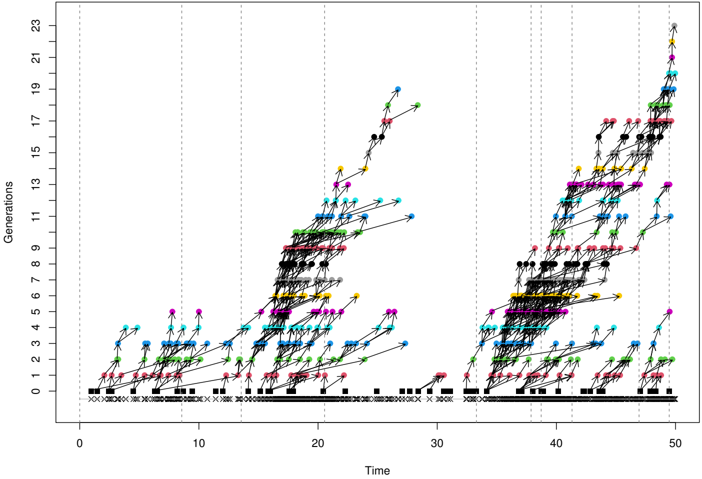
Let Q denote the unitary bivariate Poisson point process. Assume that $(X_t)_{t\in\mathbb{R}}$ is locally bounded almost surely. Then there exists a pathwise unique HPRE, with bounded intensity λ satisfying, for all $t \geq 0$,

$$\begin{cases} \lambda(t) = \eta + \int_0^t \phi(Y_s) h(t-s) N(\mathrm{d}s), \\ N\big([0,t]\big) = \int_0^t \int_0^\infty \mathbb{1}_{\{\theta \leq \lambda_s\}} Q(\mathrm{d}s, \mathrm{d}\theta). \end{cases}$$

Idea of the proof: Picard's iteration proof for ODEs.

A simulation of a HPRE

Random environment (X_t): alternating renewal process.



Discretisation scheme for HPRE

- Assume X_t satisfies the EDS

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t.$$

- Euler scheme X_t^n on $[0, T]$ with discretisation step T/n.
- HPRE N^n with intensity

$$\lambda_t^n = \lambda_0 + \int_{-\infty}^{t} \phi(X_s^n) h(t-s) N_t^n(\mathrm{d}t).$$

Proposition

Suppose that $b(\cdot)$ and $\sigma(\cdot)$ are globally Lipschitz. Then there exists an increasing function $K(\cdot)$ such that, for any $T > 0$ and $n \geq 1$,

$$\mathbb{P}\left(\int_0^T \left|N(\mathrm{d}t) - N^n(\mathrm{d}t)\right| \neq 0\right) \leq \frac{K(T)}{\sqrt{n}}$$

The EHPOU model

Exponential Hawkes Process driven by Ornstein-Uhlenbeck.

$$\begin{cases} \mathrm{d}X_t = -\alpha(X_t - \mu)\mathrm{d}t + \sigma \mathrm{d}W_t, \\ \lambda_t = \eta + \int_0^t \phi(X_s)\beta e^{-\beta(t-s)} N(\mathrm{d}s), \end{cases}$$

with initial conditions $X_0 \sim \mathcal{N}\big(\mu, \sigma^2/(2\alpha)\big)$ and $N(\mathbb{R}^-) = 0$.

Stability of EHPOU

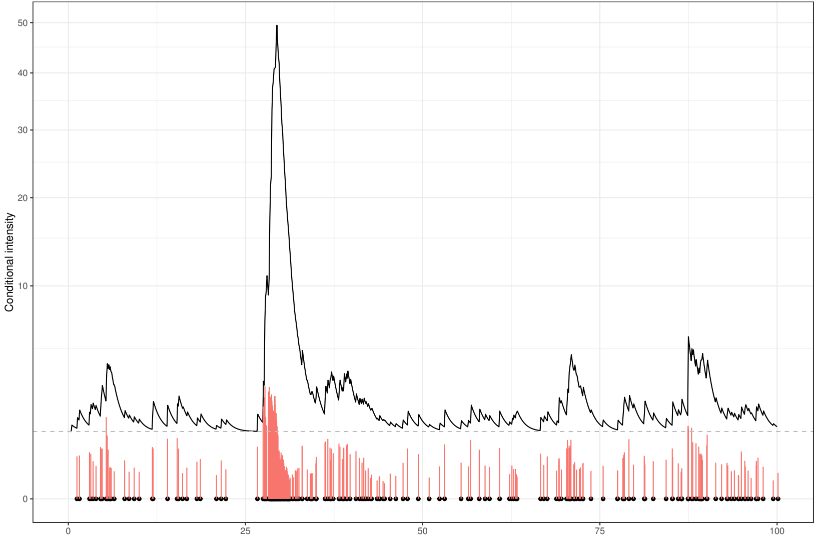
Let $l_t = \mathbb{E}\big[\lambda_t \mid (X_v)_{v \geq 0}\big]$. Then,

$$l_t = \eta \left[1 + \int_0^t \phi(X_s)\beta e^{-\beta(t-s)} \exp\left(\int_s^t \beta\phi(X_u)\mathrm{d}u \right) \mathrm{d}s \right].$$

¿The EHPOU is sub-critical if $\mathbb{E}[\phi(X_0)] < 1$?

Idea of the proof: Gronwall's lemma.

A simulation of an EHPOU

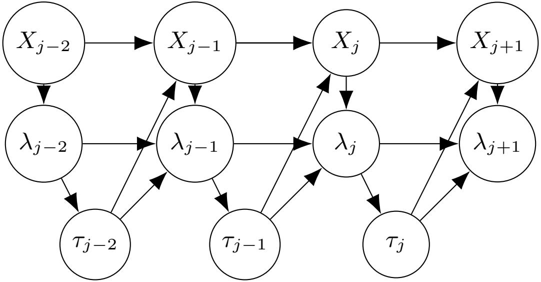


Embedded chain

Some properties

- The joint 3-uplet $(X_t, \lambda_t, N_t)_{t\geq 0}$ is jointly Markovian;
- Conditionally on $(X_t)_{t\geq 0}$, $(\lambda_t, N_t)_{t\geq 0}$ is a PDMP.

Let T_i denote the arrival times of the process, and $\tau_i = T_i - T_{i-1}$. Then, with abuse of notation, the chain $(X_i, \lambda_i, \tau_i) = (X_{T_i}, \lambda_{T_i}, \tau_i)$ is a discrete time Markov chain with transition satisfying the following graph:



Inference using EM

Estimation of the parameters $\theta = (\mu, \alpha, \sigma^2, \eta, \beta)$ through usual **Expectation-Maximisation** algorithm:

$$\log p_\theta\big((N_t)_{t\geq 0}\big) = \mathbb{E}_{p_\theta((X_i,\lambda_i)_{i=1}^n \mid (\tau_i)_{i=1}^n)}\big[\log p_\theta\big((X_i, \lambda_i, \tau_i)_{i=1}^n\big)\big],$$

where

$$\log p_\theta\big((X_i, \lambda_i, \tau_i)_{i=1}^n\big) = \sum_{j=1}^n \log p_\theta(X_j, \lambda_j, \tau_j \mid X_{j-1}, \lambda_{j-1}, \tau_{j-1}).$$

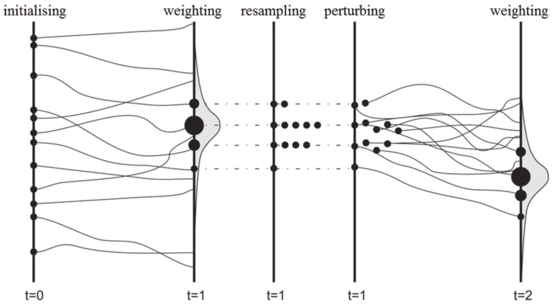
- Given θ_n obtained at n^{th} iteration:
- *E-step*: Compute $\mathbb{E}_{p_{\theta_n}((X_i,\lambda_i) \mid (\tau_i))}[\log p_\theta((X_i, \lambda_i, \tau_i))]$.
  - Or rather, its Monte-Carlo approx. $\sum_{k=1}^N \omega^{(k)} \log p_\theta((X_i^{(k)}, \lambda^{(k)}, \tau^{(k)}))$.
- *M-step*: Find $\theta_{n+1} = \arg\max_\theta \mathbb{E}_{p_{\theta_n}((X_i,\lambda_i) \mid (\tau_i))}[\log p_\theta((X_i, \lambda_i, \tau_i))]$.

Guarantees that $\log p_{\theta_{n+1}}\big((N_t)_{t\geq 0}\big) \geq \log p_{\theta_n}\big((N_t)_{t\geq 0}\big)$.

Particle filtering

Problem: $p_{\theta_n}((X_i, \lambda_i)_{i=1}^n \mid (\tau_i)_{i=1}^n)$ is unknown.

Estimation through **Sequential Monte Carlo** (Douc, Moulines, and Stoffer, 2014).



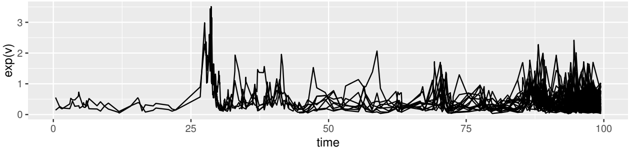
(Alvares, 2017)

For all j, this constructs samples $\{(X_j^{(k)}, \lambda_j^{(k)})\}_{k=1}^N$ under the filtering law $p_\theta(X_j, \lambda_j \mid (X_i, \lambda_i)_{i=1}^j)$.

Approximating the smoothing distribution

We want to approximate $p_\theta(X_j, \lambda_j \mid (X_i, \lambda_i)_{i=1}^n)$.

- Usually through some backward pass on the generated particles, e.g. *Poor man's smoother*.



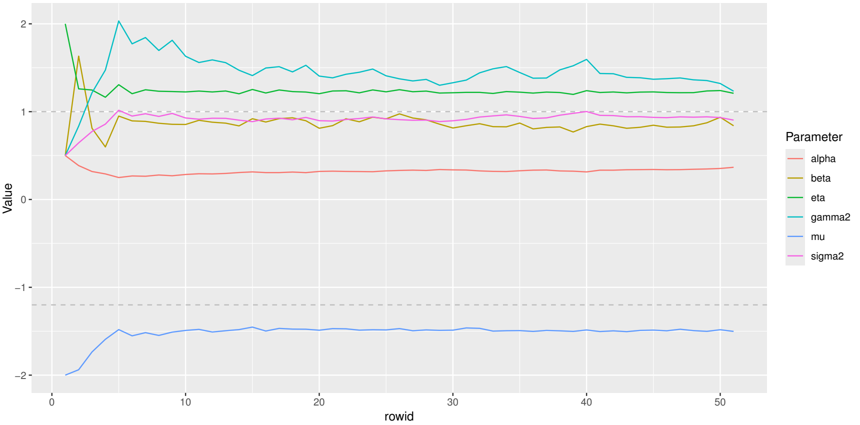
- FFBSm/i algorithms (Douc, Moulines, and Stoffer, 2014) fail here due to λ_i being conditionally deterministic given X_i, λ_{i-1} and τ_{i-1}.
- **Lagged SMC**: let $l > 0$, then

$$\mathbb{E}\big[h(Z_{j-1}, Z_j) \;\big|\; \tau_{1:n}\big] \simeq \mathbb{E}\big[h(Z_{j-1}, Z_j) \;\big|\; \tau_{1:j+l}\big].$$

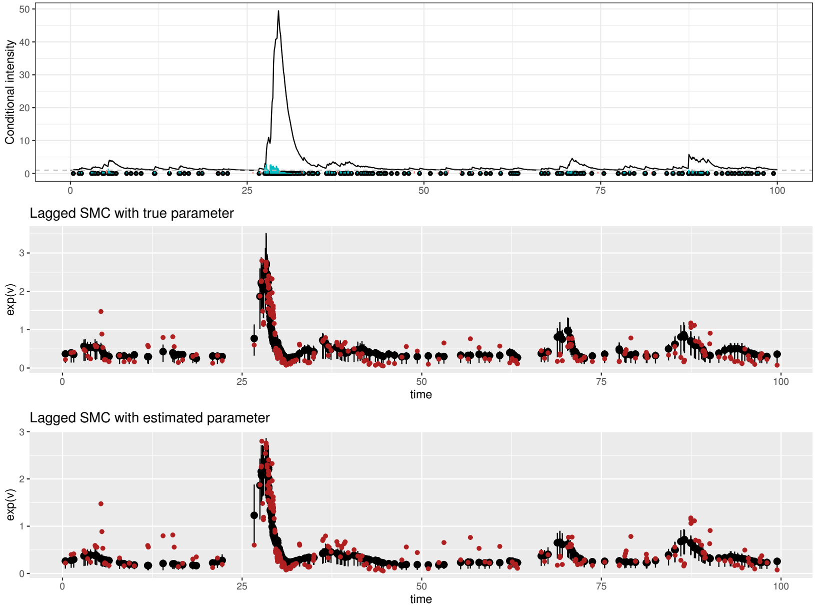
On Touille[2] tous ces objets

[2]Cette charmante chatonne de 8 ans s'appelle Touille.

Convergence of the EM-estimated parameters



Estimation of the OU trajectory



Conclusion

- Some first properties for the HPRE (existence, stability, ¿subcriticality?).
- "*Working*" statistical inference.
- Perspectives:
  - Study the tail distribution of the EHPOU.
  - Fine-tune the EM algorithm.

Thank you for your attention!

For Further Reading I

- Adamopoulos, L. (1976). "Cluster models for earthquakes: Regional comparisons". In: *J. Int. Assoc. Math. Geol.* 8.4, pp. 463–475. ISSN: 0020-5958. DOI: 10.1007/BF01028982.
- Alvares, Danilo (July 2017). "Sequential Monte Carlo methods in Bayesian joint models for longitudinal and time-to-event data". PhD thesis.
- Athreya, Krishna B. and Samuel Karlin (1971). "On Branching Processes with Random Environments: I: Extinction Probabilities". In: *The Annals of Mathematical Statistics* 42.5, pp. 1499–1520. ISSN: 0003-4851. DOI: 10.1214/aoms/1177693150.
- Bacry, Emmanuel, Iacopo Mastromatteo, and Jean-François Muzy (2015). "Hawkes Processes in Finance". In: *Mark. Microstruct. Liq.* 1.1, p. 1550005. ISSN: 2382-6266. DOI: 10.1142/S2382626615500057. arXiv: 1502.04592.

For Further Reading II

- Bansaye, Vincent, Maria Emilia Caballero, and Sylvie Méléard (2019). "Scaling limits of population and evolution processes in random environment". In: *Electronic Journal of Probability* 24. ISSN: 10836489. DOI: 10.1214/19-EJP262.
- Bansaye, Vincent and Florian Simatos (2015). "On the scaling limits of Galton–Watson processes in varying environments". In: *Electronic Journal of Probability* 20. ISSN: 10836489. DOI: 10.1214/EJP.v20-3812. arXiv: 1112.2547.
- Centers for Disease Control and Prevention (2015). *Epidemiology and Prevention of Vaccine-Preventable Diseases*. Ed. by Jennifer Hamborsky, Andrew Kroger, and Charles (Skip) Wolfe. 13th ed. Washington D.C.: Public Health Foundation.
- Dahlhaus, R. (1996). "On the Kullback-Leibler information divergence of locally stationary processes". In: *Stoch. Process. their Appl.* 62.1, pp. 139–168. ISSN: 03044149. DOI: 10.1016/0304-4149(95)00090-9.

For Further Reading III

- Dassios, Angelos and Hongbiao Zhao (2011). "A dynamic contagion process". In: *Adv. Appl. Probab.* 43.3, pp. 814–846. ISSN: 0001-8678. DOI: 10.1239/aap/1316792671.
- Douc, Randal, Eric Moulines, and David Stoffer (2014). *Nonlinear Time Series: Theory, Methods, and Applications with R Examples*. CRC Press.
- Duarte, Aline, Eva Löcherbach, and Guilherme Ost (2019). "Stability, convergence to equilibrium and simulation of non-linear Hawkes processes with memory kernels given by the sum of Erlang kernels". In: *ESAIM Probab. Stat.* 23, pp. 770–796. ISSN: 1262-3318. DOI: 10.1051/ps/2019005. arXiv: 1610.03300.
- Hawkes, Alan G. and David Oakes (1974). "A cluster process representation of a self-exciting process". In: *J. Appl. Probab.* 11.03, pp. 493–503. ISSN: 0021-9002. DOI: 10.2307/3212693.

For Further Reading IV

- Lee, Young, Kar Wai Lim, and Cheng Soon Ong (2016). "Hawkes processes with stochastic excitations". In: *33rd International Conference on Machine Learning, ICML 2016* 1, pp. 132–145. arXiv: 1609.06831.
- Reinhart, Alex (2018). "A Review of Self-Exciting Spatio-Temporal Point Processes and Their Applications". In: *Stat. Sci.* 33.3, pp. 299–318. ISSN: 0883-4237. DOI: 10.1214/17-STS629. arXiv: 1708.02647v2.
- Reynaud-Bouret, Patricia and Sophie Schbath (2010). "Adaptive estimation for hawkes processes; Application to genome analysis". In: *Ann. Stat.* 38.5, pp. 2781–2822. ISSN: 00905364. DOI: 10.1214/10-AOS806. arXiv: 0903.2919.
- Roueff, François, Rainer von Sachs, and Laure Sansonnet (2016). "Locally stationary Hawkes processes". In: *Stoch. Process. their Appl.* 126.6, pp. 1710–1743. ISSN: 03044149. DOI: 10.1016/j.spa.2015.12.003.

For Further Reading V

Smith, Walter L. and William E. Wilkinson (1969). "On Branching Processes in Random Environments". In: *The Annals of Mathematical Statistics* 40.3, pp. 814–827. ISSN: 0003-4851. DOI: 10.1214/aoms/1177697589.