A Fractional Hawkes process

J. Chen[*]

School of Mathematics, Cardiff University, UK

A. G. Hawkes[†]

School of Management, Swansea University, UK

E. Scalas[‡]

Department of Mathematics, University of Sussex, UK

Abstract

We modify ETAS models by replacing the Pareto-like kernel proposed by Ogata with a Mittag-Leffler type kernel. Provided that the kernel decays as a power law with exponent $\beta+1 \in (1,2]$, this replacement has the advantage that the Laplace transform of the Mittag-Leffler function is known explicitly, leading to simpler calculation of relevant quantities.

[*] chenj60@cardiff.ac.uk

[†] a.g.hawkes@swansea.ac.uk

[‡] e.scalas@sussex.ac.uk

I. INTRODUCTION

In 1971, Hawkes ([7, 8]) introduced a class of self-exciting processes to model contagious processes. In their simpler version, these are point processes with the following conditional intensity

$$\lambda(t|\mathcal{H}_t) = \lim_{h \to 0} \frac{\mathbb{E}(N(t, t+h)|\mathcal{H}_t)}{h} = \lambda + \alpha \int_{-\infty}^{t} f(t-u) \, dN(u),$$

where $\lambda > 0$, $N(t)$ is a Hawkes self-exciting counting process, $\mathcal{H}_t$ represents the history of the process, α is a branching ratio that must be smaller than 1 for stability, and $f(t)$ is a suitable kernel ($f(t)$ must be a probability density function for a positive random variable).

In 1988, Ogata ([11]) proposed the use of a power-law kernel for self-exciting processes of Hawkes type in order to reproduce the empirical Omori law for earthquakes. Ogata's models are also known as *Epidemic Type Aftershock Sequence* models or ETAS models. Within this framework, it is natural to replace Ogata's power-law kernel with a Mittag-Leffler kernel and this will be the main contribution of this chapter. We will first introduce the Mittag-Leffler distribution for positive random variables, then we will define a "fractional" version of Hawkes processes. Spectral properties and intensity expectation will be discussed using the fact that the Laplace transform of the one-parameter Mittag-Leffler function of argument $-t^\beta$ with $\beta \in (0, 1]$ is known analytically. Finally, we will present a simple algorithm based on the thinning method by Ogata [10] that simulates the conditional intensity process.

II. MITTAG-LEFFLER DISTRIBUTED POSITIVE RANDOM VARIABLES

Consider the one-parameter Mittag-Leffler function

$$E_\beta(z) := \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(n\beta + 1)}, \tag{II.1}$$

with $\beta \in (0, 1]$. If computed on $z = -t^\beta$ for $t \geq 0$, the Mittag-Leffler function $E_\beta(-t^\beta)$ has the meaning of survival function for a positive random variable T with infinite mean. This function interpolates between a stretched exponential for small times and a power-law with index β for large times. Its sign-changed first derivative

$$f_\beta(t) := -\frac{dE_\beta(-t^\beta)}{dt} = t^{\beta-1} E_{\beta,\beta}(-t^\beta) \tag{II.2}$$

2

is the probability density function of T, where $E_{\alpha,\beta}(z)$ is the two-parameter Mittag-Leffler function defined as

$$E_{\gamma,\delta}(z) := \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(n\gamma + \delta)}. \tag{II.3}$$

Notice that the one-parameter Mittag-Leffler function coincides with the two parameter one with $\gamma = \beta$ and $\delta = 1$. For a suitable function $f(t)$ defined for positive t, let us define its Laplace transform as

$$\tilde{f}(s) = \mathcal{L}(f(t),s) = \int_0^{\infty} f(t)e^{-st}\,dt.$$

The functions $E_\beta(-t^\beta)$ and $f_\beta(t)$ have explicit Laplace transforms. The survival function has the following Laplace transform

$$\mathcal{L}(E_\beta(-t^\beta);s) = \frac{s^{\beta-1}}{1+s^\beta}, \tag{II.4}$$

and the probability density function has the following Laplace transform

$$\mathcal{L}(f_\beta(t);s) = \frac{1}{1+s^\beta}. \tag{II.5}$$

Moreover, they have an explicit representation as an infinite (actually continuous) sum of exponential functions [5];

$$E_\beta(-t^\beta) = \int_0^{\infty} e^{-\theta t} K_\beta(\theta)\,d\theta, \tag{II.6}$$

with

$$K_\beta(\theta) = \frac{1}{\pi} \frac{\theta^{\beta-1} \sin(\beta\pi)}{\theta^{2\beta} + 2\theta^\beta \cos(\beta\pi) + 1}, \tag{II.7}$$

leading to

$$f_\beta(t) = \int_0^{\infty} \theta e^{-\theta t} K_\beta(\theta)\,d\theta. \tag{II.8}$$

These functions play an important role in fractional calculus. For instance $E_\beta(-t^\beta)$ is the solution of the following anomalous relaxation problem

$$\frac{d^\beta g(t)}{dt^\beta} = -g(t), \tag{II.9}$$

where d^β/dt^β is the Caputo derivative defined as

$$\frac{d^\beta g(t)}{dt^\beta} = \frac{1}{\Gamma(1-\beta)} \frac{d}{dt} \int_0^t \frac{g(\tau)}{(t-\tau)^\beta}\,dt - \frac{t^{-\beta}}{\Gamma(1-\beta)}g(0^+), \tag{II.10}$$

with initial condition $g(0^+) = 1$.

3

III. THE FRACTIONAL HAWKES PROCESSES

It becomes natural to use $f_\beta(t)$ as kernel for a version of Hawkes processes that we can call *fractional* Hawkes processes. The conditional intensity of the process is given by

$$\lambda(t|\mathcal{H}_t) = \lim_{h \to 0} \frac{\mathbb{E}(N(t, t+h)|\mathcal{H}_t)}{h} = \lambda + \alpha \int_{-\infty}^{t} f_\beta(t-u) \, dN(u), \qquad \text{(III.1)}$$

where $\lambda > 0$ and $N(t)$ is a Hawkes self-exciting point process, leading to

$$\lambda(t|\mathcal{H}_t) = \lambda + \alpha \sum_{t_i < t} f_\beta(t - t_i). \qquad \text{(III.2)}$$

with the branching ratio $\alpha < 1$ for stability. Hainaut [6] gives a different definition of fractional Hawkes process. Let us use his notation in the following remark. In his paper, he considers the time-changed intensity process λ_{S_t} where, in his case, the conditional intensity λ_t of the self-exciting process is the solution of a mean-reverting stochastic differential equation

$$d\lambda_t = \kappa(\theta - \lambda_t) \, dt + \eta \, dP_t,$$

where κ, θ and η are suitable parameters and the driving process P_t is given by

$$P_t := \sum_{k=1}^{N_t} \xi_k,$$

where N_t is a counting process and ξ_is are independent and identically distributed marks with finite positive mean and finite variance. The time-change S_t is the inverse of a β-stable subordinator. Our definition (III.1) is much simpler and it is directly connected with ETAS processes given that the kernel $f_\beta(t)$ has power-law tail with index $\beta + 1$. [9]. In particular, given the explicit Laplace transform of $f_\beta(t)$ and its representation in terms of infinite sum of exponentials, we can derive some explicit formulas.

A. Spectral properties

From equation (11) in [7], we get the following equation for the covariance density $\mu(\tau)$ for $\tau > 0$.

$$\mu(\tau) = \alpha \left[\Lambda f_\beta(\tau) + \int_0^\tau f_\beta(\tau - v) \mu(v) \, dv + \int_0^\infty f_\beta(\tau + v) \mu(v) \, dv \right], \qquad \text{(III.3)}$$

where Λ represents the asymptotic stationary value of the conditional intensity as derived in equation (3.9) below. If we now take the Laplace transform of (III.3), we get

$$\tilde{\mu}(s) = \alpha \left[\Lambda \tilde{f}_\beta(s) + \tilde{f}_\beta(s)\tilde{\mu}(s) + \mathscr{L}\left(\int_0^\infty f_\beta(\tau + v)\mu(v)\,dv; s \right) \right]. \tag{III.4}$$

Now, using (II.8) and setting

$$h(\theta) = \theta K_\beta(\theta) = \frac{1}{\pi} \frac{\theta^\beta \sin(\beta\pi)}{\theta^{2\beta} + 2\theta^\beta \cos(\beta\pi) + 1},$$

we can write

$$f_\beta(\tau) = \int_0^\infty h(\theta) e^{-\theta\tau}\,d\theta,$$

so that the last summand in (III.4) becomes

$$\mathscr{L}\left(\int_0^\infty f_\beta(\tau + v)\mu(v)\,dv; s \right) = \int_0^\infty e^{-s\tau} \left[\int_0^\infty \left[\int_0^\infty h(\theta) e^{-\theta(\tau+v)}\,d\theta \right] \mu(v)\,dv \right] d\tau$$

$$= \int_0^\infty h(\theta) \frac{1}{\theta + s} \tilde{\mu}(\theta)\,d\theta. \tag{III.5}$$

In principle, a numerical approximation of the integral in equation (III.5) plugged into (III.4) and coupled with equation (II.5) can lead to an explicit approximate expression for the Laplace transform $\tilde{\mu}(s)$. This will be the subject of a further paper. An alternative approach is given in [8] where the Bartlett spectrum is defined for real ω as

$$f(\omega) = \frac{1}{2\pi} \int e^{-i\omega\tau} \mu^{(c)}(\tau)\,d\tau, \tag{III.6}$$

where, because $\mathbb{E}[(dN(t))^2] = \mathbb{E}[dN(t)]$ if events cannot occur multiply, the complete covariance density contains a delta function

$$\mu^{(c)}(\tau) = \Lambda\delta(t) + \mu(t).$$

Then it is shown in [8] p. 441 that

$$f(\omega) = \frac{\Lambda}{2\pi(1 - G(\omega))(1 - G(-\omega))}, \tag{III.7}$$

where, in our case, we would have

$$G(\omega) = \int_0^\infty e^{-i\omega\tau} \alpha f_\beta(\tau)\,d\tau = \frac{\alpha}{1 + (i\omega)^\beta}.$$

5

The proof of this result depended on the assumption that the exciting kernel decays exponentially asymptotically. However, this is not true for the Mittag-Leffler distribution, which decays as a power law. Bacry and Muzy [2] prove a more general result using Laplace transforms in the complex plane, more easily digested from section 2.3.1 in [1]. Then the Laplace transform of the covariance density is given by

$$\tilde{\mu}^{(c)}(s) = \frac{\Lambda}{(1 - \tilde{\Phi}(s))(1 - \tilde{\Phi}(-s))}, \tag{III.8}$$

where

$$\tilde{\Phi}(s) = \int_0^\infty e^{-s\tau} \Phi(\tau) d\tau$$

is the Laplace transform of the exciting kernel. In our case for $\tau > 0$, we have

$$\Phi(\tau) = \alpha f_\beta(\tau)$$

and

$$\tilde{\Phi}(s) = \frac{\alpha}{1 + s^\beta}.$$

Equations (III.7) and (III.8) look much the same, apart from a change of notation and a factor 2π. The difference, however, is that in (III.7) we are dealing with real ω while, in (III.8), s is a general complex variable and we can choose its domain to obtain well-behaved functions.

B. Intensity expectation

Let us consider the expectation $\Lambda(t) = \mathbb{E}[\lambda(t|\mathcal{H}_t)]$ for both a stationary and non-stationary fractional Hawkes process. In the stationary case (process from $t = -\infty$), from equation (III.1), we get $\Lambda = \lambda + n\Lambda$, leading to

$$\Lambda = \frac{\lambda}{1 - \alpha}. \tag{III.9}$$

On the contrary in the non-stationary case (process from $t = 0$), we can modify equation (III.1) as follows

$$\lambda(t|\mathcal{H}_t) = \lambda + \alpha \int_0^t f_\beta(t - u) dN(u), \tag{III.10}$$

so that the time-dependent expectation obeys the equation

$$\Lambda(t) = \lambda + \alpha \int_0^t f_\beta(t - u) \Lambda(u) du. \tag{III.11}$$

Taking Laplace transforms, we get

$$\tilde{\Lambda}(s) = \frac{\lambda}{s} + \alpha \tilde{f}_\beta(s)\tilde{\Lambda}(s),$$

so that

$$\tilde{\Lambda}(s) = \frac{\lambda}{s}\frac{1}{1-\alpha\tilde{f}_\beta(s)}.$$

Using equation (II.5) yields

$$\tilde{\Lambda}(s) = \frac{\lambda}{s}\frac{1+s^\beta}{(1-\alpha)+s^\beta}. \tag{III.12}$$

Equation (III.12) can be inverted numerically (or analytically for $\beta = 1/2$) to give $\Lambda(t)$ as well as the expected number of events from 0 to t as

$$\mathbb{E}[N(t)] = \int_0^t \Lambda(\tau)d\tau.$$

Also, based on a continuous version of Hardy-Littlewood Tauberian theorem [3] we get

$$\lim_{t\to\infty}\Lambda(t) = \frac{\lambda}{1-\alpha}$$

as given by equation (III.9). This result is exemplified in Fig. 1 for $\beta = 1/2$, $\lambda = 1$ and $\alpha = 1/2$. In that case, we get, for $t > 0$

$$\Lambda(t) = 2 - e^{t/4}\text{erfc}(\sqrt{t}/2).$$

From Fig. 1, one can see that $\Lambda(t)$ goes up very fast at first and, then, slowly converges to its asymptotic value 2. This is presumably due to the long tail of the Mittag-Leffler kernel.

IV.   SIMULATION

In order to simulate the intensity process introduced in equation (III.1), we use the thinning algorithm introduced by Ogata [10] (see also Zhuang and Touati 2015 report [12]). The function `ml.m` described in [4] is needed to compute the Mittag-Leffler functions described above and can be retrieved from the Matlab file exchange. The algorithm is as follows

1. Set the initial time $t = 0$, a counter $i = 0$ and a final time T.

2. Compute $M = \lambda + \alpha\sum_{t_i < t+\varepsilon} f_\beta(t+\varepsilon-t_i)$, for some small value of ε.
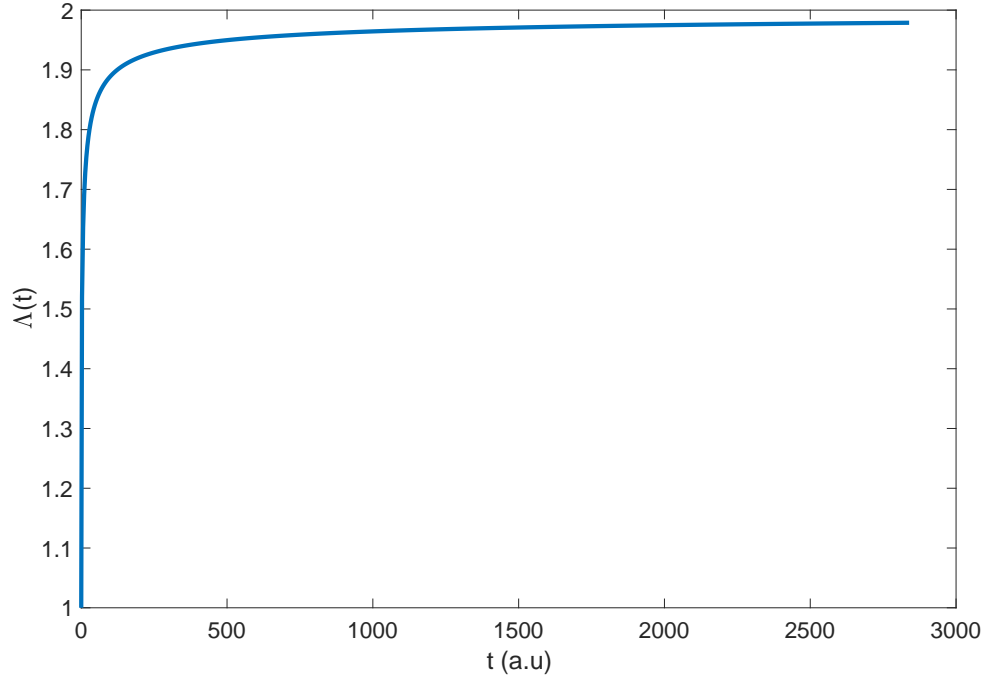
FIG. 1. $\Lambda(t)$ as a function of t for $\beta = 1/2$, $\lambda = 1$ and $\alpha = 1/2$.

3. Generate a positive exponentially distributed random variable E with the meaning of a waiting time, with rate $1/M$.

4. Set $\tau = t + E$.

5. Generate a uniform random variate U between 0 and 1.

6. If $U < [\lambda + \alpha \sum_{t_i < \tau} f_\beta(\tau - t_i)]/M$, set $t_{i+1} = \tau$ and update time to $t = \tau$, else, just set $t = \tau$.

7. Return to step 2 until t exceeds T.

8. Return the set of event times (or epochs) t_i

An implementation of this algorithm in Matlab is presented below.

```
T=5;
t=0;
```

```
n=0;
alpha=0.9; % For stability
mu=1;
epsilon=1e-10;
beta=0.7;


SimPoints=[];


while t<T
t
M=mu+sum(alpha*ml(-(t+epsilon-SimPoints).^beta,beta,beta,1).  ...
*(t+epsilon-SimPoints).^(beta-1));
E=exprnd(1/M,1,1);
t=t+E;
U=rand;
if (U<(((mu+sum(alpha*ml(-(t-SimPoints).^beta,beta,beta,1).  ...
*(t-SimPoints).^(beta-1))))/M))
n=n+1;
SimPoints = [SimPoints, t];
end
end
index=find(SimPoints<10);
SimPoints=SimPoints(index);
```

Two examples of intensity process simulated up to $t = 5$ are presented in Fig. 2 and Fig. 3 for $\beta = 0.9$ and $\beta = 0.7$, respectively.

V. OUTLOOK

In this paper, we defined a "fractional" Hawkes process and we studied its spectral properties and the expectation of its intensity. Thanks to explicit expressions for Laplace transforms, we
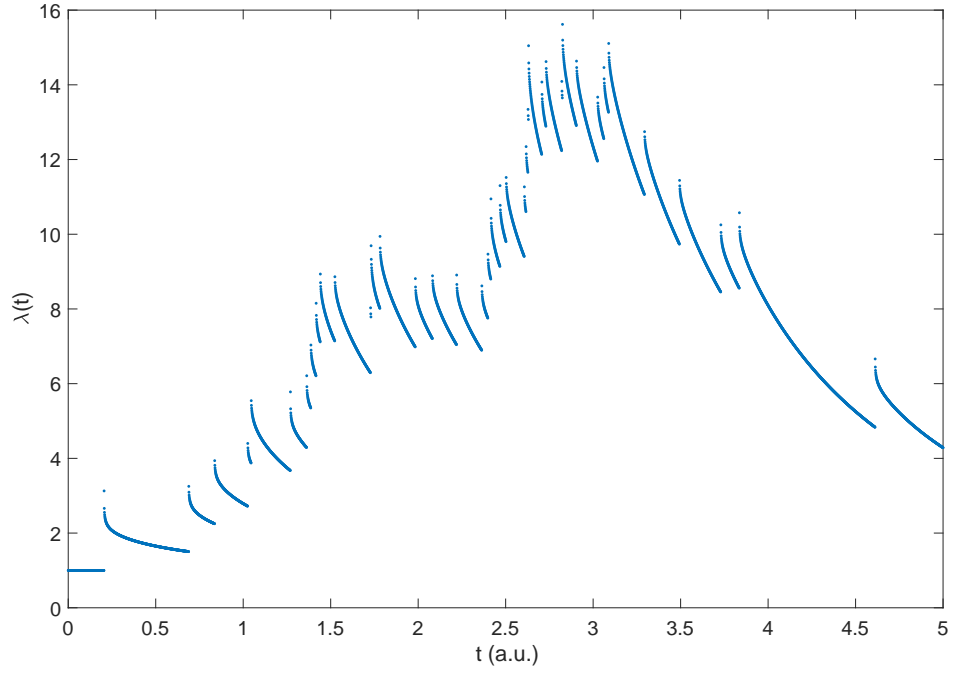
FIG. 2. $\lambda(t|\mathcal{H}_t)$ as a function of time t for $\beta = 0.9$, $\lambda = 1$ and $\alpha = 0.9$.
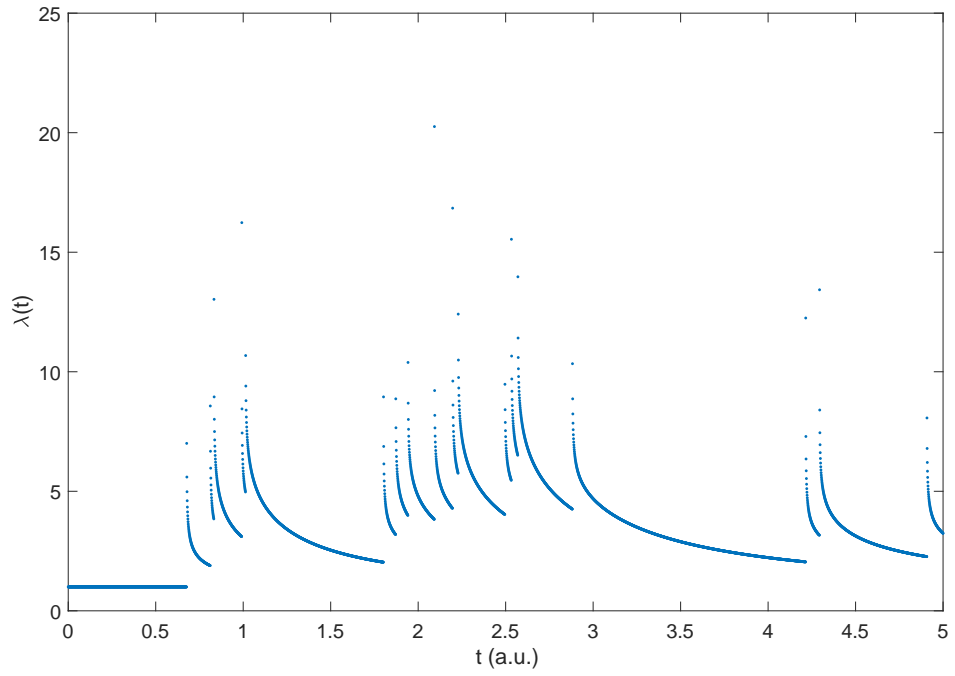

FIG. 3. $\lambda(t|\mathcal{H}_t)$ as a function of t for $\beta = 0.7$, $\lambda = 1$ and $\alpha = 0.9$.

10

could derive some analytical expressions that are only asymptotically available for power-law kernels of Pareto type as originally suggested by Ogata. We also presented an explicit simulation of the intensity process based on the so-called thinning method.

Further work is needed to better characterize our process. In particular, we did not deal with parameter estimation, the multivariate version of the process, and we did not use the model to fit earthquake data (or any other data for what matters including financial data). We do hope that all this and more can become the subject of an extensive future paper on this process.

---

[1] E. Bacry, I. Mastromatteo, and J.F. Muzy, Hawkes processes in finance, Market Microstructure and Liquidity, **1**, 2015. DOI: `http://dx.doi.org/10.1142/S2382626615500057`, pp59.

[2] E. Bacry, and J.F. Muzy, First- and second order statistics characterization of Hawkes processes and non-parametric estimation, IEEE Transactions On Information Theory, **62**(4), 2184–2201, 2016.

[3] W. Feller, An introduction to probability theory and its applications. Vol. II. Second edition, John Wiley & Sons, 1971.

[4] R. Garrappa, Numerical evaluation of two and three parameter Mittag-Leffler functions, SIAM Journal of Numerical Analysis, **53**(3), 1350–1369, 2015.

[5] R. Gorenflo and F. Mainardi, Fractional Calculus: Integral and Differential Equations of Fractional Order in Fractals and Fractional Calculus in Continuum Mechanics, A. Carpinteri and F. Mainardi (eds.) Springer 223–276, 1997.

[6] D. Hainaut, Fractional Hawkes processes, UCLouvain preprint, 2019.

[7] A.G. Hawkes, Spectra of some self-exciting and mutually exciting point processes, Biometrika **58**, 83–90, 1971.

[8] A.G. Hawkes, Spectra of some mutually exciting point processes, J. Royal Statistical Society B **33**(3), 438–443, 1971.

[9] F. Mainardi, M. Raberto, R. Gorenflo and E. Scalas, Fractional calculus and continuous-time finance II: the waiting-time distribution, Physica A, **287**, 468–481, 2000.

[10] Y. Ogata, On Lewis' simulation method for point processes, IEEE Transactions on Information Theory, **27**, 23–31, 1981.

[11] Y. Ogata, Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes,

83(401), 9–27, 1988.

[12] J. Zhuang and S. Touati, Stochastic simulation of earthquake catalogs, Community Online Resource for Statistical Seismicity Analysis, 2015. Available at `http://www.corssa.org`.