On the Truncation Error of a Superposed Gamma Process

Julyan Arbel and Igor Prünster

Abstract Completely random measures (CRMs) form a key ingredient of a wealth of stochastic models, in particular in Bayesian Nonparametrics for defining prior distributions. CRMs can be represented as infinite series of weighted random point masses. A constructive representation due to Ferguson and Klass provides the jumps of the series in decreasing order. This feature is of primary interest when it comes to sampling since it minimizes the truncation error for a fixed truncation level of the series. In this paper we focus on a general class of CRMs, namely the *superposed gamma process*, which suitably transformed has already been successfully implemented in Bayesian Nonparametrics, and quantify the quality of the approximation in two ways. First, we derive a bound in probability for the truncation error. Second, following [1], we study a moment-matching criterion which consists in evaluating a measure of discrepancy between actual moments of the CRM and moments based on the simulation output. To this end, we show that the moments of this class of processes can be obtained analytically.

Keywords Bayesian Nonparametrics · Completely random measures · Ferguson and Klass algorithm · Moment-matching · Normalized random measures · Posterior sampling · Superposed gamma process

1 Introduction

Completely random measures (CRMs), also known as independent increment processes, have blossomed in the last decades in modern stochastic modeling and inference as a basic building block of countless popular models. A prominent usage

J. Arbel (✉)
Inria Grenoble - Rhône-Alpes, Grenoble, France
e-mail: julyan.arbel@inria.fr

J. Arbel
Université Grenoble Alpes, Grenoble, France

I. Prünster
Department of Decision Sciences, BIDSA and IGIER, Bocconi University, Milan, Italy
e-mail: igor@unibocconi.it

of CRMs is within Bayesian nonparametric statistics (see [15, 18]). For instance, the popular Dirichlet process [9] can be obtained as normalization or exponentiation of suitable CRMs (see [10]). Survival analysis, [2, 17], random sparse networks, [6], biology [20, 28], are only a few of the various modern applications tackled with CRMs.

Implementation of CRM-based models usually requires to simulate the CRMs trajectories. As infinite dimensional objects, they need to be truncated, leading to an approximation error. The representation due to Ferguson and Klass [11] (see also [30]) is arguably one of the most useful ones in that it displays the weights in decreasing order. This implies that the approximation error is minimized over the whole sample space for a given truncation level. This appealing feature was exploited in many works, including [3, 4, 7, 8, 13, 22–26] to cite just a few in Bayesian Nonparametrics. The quality of the approximation, hardly addressed by those previous works, is the focus of this paper.

Many classical methods in statistics and econometrics use moments, such as, for instance, the method of simulated moments [21] and the general method of moments [12]. In this paper, we follow another research line and show how moments of the CRMs can be used in order to assess the quality of the approximation due to the truncation [1]. It is based on the observation that moments of CRMs are simple to compute, hence one can quantify the quality of the approximation by evaluating a measure of discrepancy between the actual moments of the CRM at issue and the moments computed based on the sampled realizations of the CRM. The truncation level is then selected so that the measure of discrepancy does not exceed a given threshold, say 5%. In Arbel and Prünster [1] the methodology is illustrated on two classes of CRMs, namely the *generalized gamma process* and the *stable-beta process*. In the present paper we focus on another broad class called the *superposed gamma process* (see [19, 27]). More specifically, after a brief presentation of CRMs and of the Ferguson and Klass algorithm in Sect. 2, we derive a bound in probability on the truncation error in Sect. 3.1 and then show the applicability of the moment-matching criterion by deriving analytically the moments of the superposed gamma process in Sect. 3.2.

2 Sampling Completely Random Measures

2.1 Completely Random Measures

A CRM $\tilde{\mu}$ on $\mathbb{X}$ is a random measure which spreads out mass independently in the space. More precisely, the random variables $\tilde{\mu}(A_1), \ldots, \tilde{\mu}(A_n)$ are mutually independent for any disjoint sets $A_1, \ldots, A_n$.

Kingman [16] showed that the only way to spread out mass in a *completely random* fashion (without deterministic components) is by randomly scattering point masses in the space. In other words, CRMs select (almost surely) discrete measures and

hence can be represented as

$$\tilde{\mu} = \sum_{i \geq 1} J_i \delta_{Z_i} \tag{1}$$

where the jumps J_i and locations Z_i are random. Both jumps and locations are controlled by the so-called Lévy intensity which characterizes the CRM. It is a measure on $\mathbb{R}^+ \times \mathbb{X}$ which can be written as $\nu(\mathrm{d}v, \mathrm{d}x) = \rho(\mathrm{d}v)\alpha(\mathrm{d}x)$ for so-called homogeneous CRM, which are considered here and correspond to the case of jumps independent of the locations. The function ρ controls the intensity of the jumps. The measure α, if the CRM is (almost surely) finite, which is assumed throughout, splits up in $\alpha = a P_0$ where $a > 0$ is called the total mass parameter and the probability distribution P_0 tunes the locations.

Ever-popular CRMs include the *generalized gamma process* introduced by Brix [5] and the *stable-beta process*, or *three-parameter beta process*, defined by Teh and Gorur [29] as an extension of the beta process [14]. Here we consider another large class of completely random measures called *superposed gamma process*, introduced by Regazzini et al. [27]. It is identified by the jump intensity

$$\rho(\mathrm{d}v) = \frac{1 - \mathrm{e}^{-\eta v}}{1 - \mathrm{e}^{-v}} \frac{\mathrm{e}^{-v}}{v} \, \mathrm{d}v, \quad \eta > 0. \tag{2}$$

As noted by Lijoi et al. [19], one usually restricts attention to the case of positive integer η. Under this assumption, the superposed gamma processtakes the form of a genuine *superposition* of independent gamma processes with increasing integer-valued scale parameter, with jump intensity $\rho(\mathrm{d}v) = \frac{1}{v} \left(\mathrm{e}^{-v} + \mathrm{e}^{-2v} + \ldots + \mathrm{e}^{-\eta v} \right) \mathrm{d}v/v$. The specification of integer values for η has also the advantage to lead to analytical computation of the moments. Note that the special case $\eta = 1$ reduces to the gamma process, which gives rise to the Dirichlet process by normalization. Alternatively, the normalization of the superposed gamma processfor unspecified η provides the so-called generalized Dirichlet process [19].

2.2 Ferguson and Klass Algorithm

Ferguson and Klass [11] devise a constructive representation of a CRM which produces the jumps in decreasing order. This corresponds to the (almost surely unique) ordering of the sum elements in (1) where $J_1 > J_2 > \cdots$. Indeed, the jumps are obtained as $\xi_i = N(J_i)$, where $N(v) = \nu([v, \infty), \mathbb{X})$ is a decreasing function, and $\xi_1, \xi_2, \ldots$ are jump times of a standard Poisson process (PP) of unit rate: $\xi_1, \xi_2 - \xi_1, \ldots \overset{\text{i.i.d.}}{\sim} \mathrm{Exp}(1)$. Figure 1 illustrates the function $N(\cdot)$ which takes the following form in the superposed gamma process case
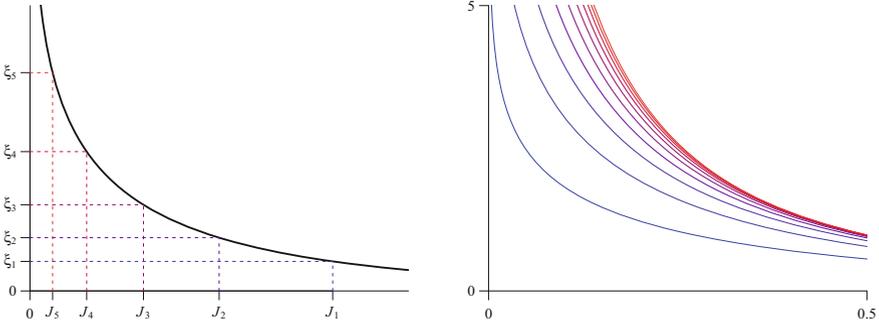
Fig. 1 *Left* illustration of Ferguson and Klass representation through the inversion of the jumps times $\xi_1, \ldots, \xi_5$ for a homogeneous Poisson process on $\mathbb{R}_+$ to the jumps $J_1, \ldots, J_5$ of a CRM. *Right* tail of the Lévy measure $N(\cdot)$ of the superposed gamma process with $\eta \in \{1, \ldots, 10\}$, $\eta = 1$ for the *lowest curve*, $\eta = 10$ for the *highest curve*

$$N(v) = a E_\eta(v), \text{ where } E_\eta(v) = \sum_{k=1}^{\eta} E_1(kv) \text{ and } E_1(v) = \int_v^\infty u^{-1} e^{-u} du, \quad (3)$$

and where the function E_1 denotes the *exponential integral function*.

Since it is impossible to sample an infinite number of jumps, approximate simulation of $\tilde{\mu}$ is in order. This becomes a question of determining the number M of jumps to sample leading to the truncation $\tilde{\mu}_M$ and truncation error T_M as follow

$$\tilde{\mu}_M = \sum_{i=1}^{M} J_i \delta_{Z_i}, \quad T_M = \sum_{i=M+1}^{\infty} J_i. \quad (4)$$

The Ferguson and Klass representation has the key advantage of generating the jumps in decreasing order implicitly minimizing such an approximation error. However, a precise evaluation of T_M, for example in expectation, is a daunting task due to the non independence of the jumps in the Ferguson and Klass representation. The algorithm is summarized in Algorithm 1.

Then, the natural path to determining the truncation level M would be the evaluation of the Ferguson and Klass tail sum

$$\sum_{i=M+1}^{\infty} N^{-1}(\xi_i). \quad (5)$$

Brix ([5], Theorem A.1) provided an upper bound in probablity for (5) in the generalized gamma case. In Proposition 1 of Sect. 3 we derive also an upper bound for the tail sum of the superposed gamma process. However, both bounds are far from sharp and therefore of little practical use as highlighted in Sect. 3. This motivates the idea

of looking for a different route and our proposal consists in the moment-matching technique detailed in the next section.

From Fig. 1 it is apparent that increasing η leads to larger jumps which in turn leads to the need of a higher truncation level in order to match a given precision level. This is not surprising given the CRM at hand can be thought of as a superposition of η gamma CMRs. Such an intuition is made precise in the next section.

Algorithm 1 Ferguson & Klass algorithm

1: sample $\xi_i \sim$ PP for $i = 1, \ldots, M$
2: define $J_i = N^{-1}(\xi_i)$ for $i = 1, \ldots, M$
3: sample $Z_i \sim P_0$ for $i = 1, \ldots, M$
4: approximate $\tilde{\mu}$ by $\sum_{i=1}^{M} J_i \delta_{Z_i}$

---

3 Truncation Error of the Superposed Gamma Process

3.1 Bound in Probability

We provide an evaluation in probability of the truncation error T_M in (4).

Proposition 1 *Let $(\xi_j)_{j \geq 1}$ be the jump times for a homogeneous Poisson process on $\mathbb{R}^+$ with unit intensity. Then for any $\epsilon \in (0, 1)$, the tail sum of the superposed gamma process (4) satisfies*

$$\mathbb{P}\left(T_M \leq t_M^\epsilon\right) \geq 1 - \epsilon, \; for \; t_M^\epsilon = \frac{C}{(\eta!)^{1/\eta}} e^{1 - \frac{M}{C}}, \; where \; C = \frac{2ea\eta}{\epsilon}.$$

Proof The proof follows along the same lines as the proof of Theorem A.1. by Brix [5] for the generalized gamma process and Proposition 4 by Arbel and Prünster [1] for the stable-beta process. Let q_j denote the $\epsilon 2^{M-j}$ quantile, for $j = M + 1, M + 2, \ldots$, of a gamma distribution with mean and variance equal to j. Then

$$\mathbb{P}\left(\sum_{j=M+1}^{\infty} N^{-1}(\xi_j) \leq \sum_{j=M+1}^{\infty} N^{-1}(q_j) \right) \geq 1 - \epsilon.$$

Denote $\tilde{t}_M^\epsilon = \sum_{j=M+1}^{\infty} N^{-1}(q_j) = \sum_{j=M+1}^{\infty} E_\eta^{-1}(q_j/a)$, and let us upper bound E_η^{-1}. By using $E_1(u) \leq 1 - \log(u)$, one gets

$$E_\eta(u) = \sum_{l=1}^{\eta} E_1(lu) \leq \eta - \sum_{l=1}^{\eta} \log(lu) = \eta - \log\left(\eta! u^\eta\right),$$

which can be inverted to obtain

$$E_\eta^{-1}(x) \leq \frac{1}{(\eta!)^{1/\eta}} e^{1-\frac{x}{\eta}}.$$

Additionally, since the quantiles satisfy $q_j \geq \frac{\epsilon}{2e} j$, we can conclude that

$$\tilde{t}_M^\epsilon \leq \frac{1}{(\eta!)^{1/\eta}} \sum_{j=M+1}^\infty e^{1-\frac{q_j}{a\eta}} \leq \frac{1}{(\eta!)^{1/\eta}} \sum_{j=M+1}^\infty e^{1-\frac{\epsilon j}{2ea\eta}} \leq \frac{2ea\eta}{\epsilon(\eta!)^{1/\eta}} e^{1-\frac{\epsilon M}{2ea\eta}}. \quad \square$$

Remark It is interesting to note that the bound t_M^ϵ for the superposed gamma process is equal to its counterpart for the beta process with concentration parameter c set to η, all else things being equal (total mass parameter a and threshold ϵ). See Proposition 4 in [1]. This finding provides a nice connection between both processes otherwise seemingly unrelated.

The bound t_M^ϵ obtained in Proposition 1 is exponentially decreasing with M, which is reminiscent of the results obtained by Brix [5] and Arbel and Prünster [1], respectively, for the generalized gamma process and the stable-beta process with no stable component. As already pointed out by these authors, the bound t_M^ϵ is very conservative due to a crude lower bound on the quantiles q_j (notation of the proof). The left panel of Fig. 2 displays this bound t_M^ϵ, while the right panel illustrates the truncation level M (in log-scale) required in order to guarantee with 95% probability an upper bound on T_M of $t_{max} \in \{1, 10, 100\}$, for varying values of η. Inspection of the plots demonstrates the rapid increase with η of the number of jumps needed in order to assess a given bound in probability.

As suggested by a Referee, a possible strategy for improving the result in Proposition 1 is to rely on concentration of measure techniques. This will be the object of future investigations. A numerical strategy to improve the approximation consists in
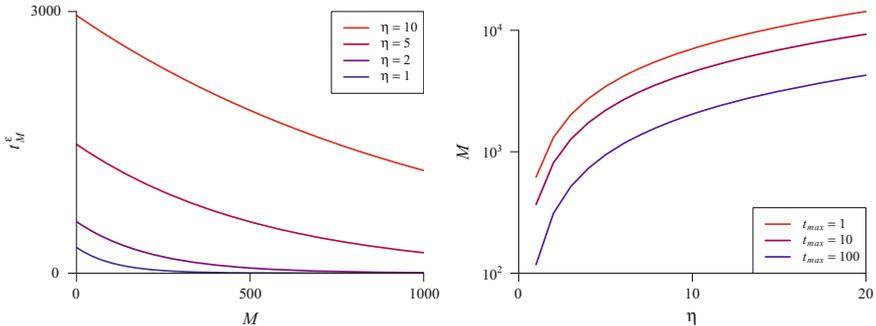

Fig. 2 *Left* variation of $M \mapsto t_M^\epsilon$ for $\eta \in \{1, 2, 5, 10\}$. *Right* variation of the threshold function $\eta \mapsto M$ needed to match an error bound of $t_{max} \in \{1, 10, 100\}$ with $\eta \in \{1, \ldots, 20\}$, log scale on y-axis

directly calculating the quantiles q_j (instead of resorting to the lower bound), thus loosing the closed form expression of the bound.

3.2 Moment-Matching Criterion

We first concisely recall the moment-matching methodology introduced by Arbel and Prünster [1] and then tailor it to the superposed gamma process. We assess the quality of approximation of the Ferguson & Klass algorithm by comparing the actual distribution of the random total mass $\tilde{\mu}(\mathbb{X}) = \sum_{i=1}^{\infty} J_i$ with its empirical distribution (obtained by the sampled trajectories). Motivated by the fact that the first moments carry much information about a distribution, the comparison is made by comparing theoretical and empirical moments of $\tilde{\mu}(\mathbb{X})$. As measure of discrepancy, we use the mean squared error between theoretical and empirical moments. We refer to [1] for illustrations of this moment-matching criterion on the generalized gamma process and the stable-beta process.

In order to apply this methodology also to the superposed gamma process, we need to derive its theoretical moments. The n-th (raw) moment of the random total mass is defined as

$$m_n = \mathbb{E}\big[\tilde{\mu}^n(\mathbb{X})\big].$$

For general homogeneous CRMs, it takes on the form (see, e.g., Proposition 1 in [1])

$$m_n = \sum_{(*)} \binom{n}{k_1 \cdots k_n} \prod_{i=1}^{n} \big(\kappa_i / i!\big)^{k_i}, \tag{6}$$

where the sum $(*)$ is over all n-tuples of nonnegative integers $(k_1, \ldots, k_n)$ satisfying the constraint $k_1 + 2k_2 + \cdots + nk_n = n$ and where κ_i is the ith cumulant defined by

$$\kappa_i = a \int_0^{\infty} v^i \rho(\mathrm{d}v).$$

In the case of the superposed gamma process, simple algebra leads to the following expression for the cumulants

$$\kappa_i = a(i-1)!\zeta_\eta(i) \tag{7}$$

which are in terms of the *incomplete Euler–Riemann zeta function* $\zeta_\eta(i) = \sum_{l=1}^{\eta} \frac{1}{l^i}$. Hence the moment-matching methodology introduced by Arbel and Prünster [1] can be readily applied by resorting (6) and (7).

References

1. Arbel, J., Prünster, I.: A moment-matching Ferguson and Klass algorithm. Stat. Comput. **27**(1), 3–17 (2017). doi:10.1007/s11222-016-9676-8, arxiv:1606.02566
2. Arbel, J., Lijoi, A., Nipoti, B.: Full Bayesian inference with hazard mixture models. Comput. Stat. Data Anal. **93**, 359–372 (2016)
3. Argiento, R., Bianchini, I., Guglielmi, A.: A blocked Gibbs sampler for NGG-mixture models via a priori truncation. Stat. Comput. **26**(3), 641–661 (2016)
4. Barrios, E., Lijoi, A., Nieto-Barajas, L.E., Prünster, I.: Modeling with normalized random measure mixture models. Stat. Sci. **28**(3), 313–334 (2013)
5. Brix, A.: Generalized gamma measures and shot-noise Cox processes. Adv. Appl. Prob. **31**, 929–953 (1999)
6. Caron, F., Fox, E.B.: Sparse graphs using exchangeable random measures (2015). arXiv:14011137
7. De Blasi, P., Favaro, S., Muliere, P.: A class of neutral to the right priors induced by superposition of beta processes. J. Stat. Plan. Inference **140**(6), 1563–1575 (2010)
8. Epifani, I., Lijoi, A., Prünster, I.: Exponential functionals and means of neutral-to-the-right priors. Biometrika **90**(4), 791–808 (2003)
9. Ferguson, T.S.: A Bayesian analysis of some nonparametric problems. Ann. Stat. **1**(2), 209–230 (1973)
10. Ferguson, T.S.: Prior distributions on spaces of probability measures. Ann. Stat. **2**(4), 615–629 (1974)
11. Ferguson, T.S., Klass, M.J.: A representation of independent increment processes without Gaussian components. Ann. Math. Stat. **43**(5), 1634–1643 (1972)
12. Gallant, A.R., Tauchen, G.: Which moments to match? Econometr. Theor. **12**, 657–681 (1996)
13. Griffin, J.E., Walker, S.G.: Posterior simulation of normalized random measure mixtures. J. Comput. Graph. Stat. **20**(1), 241–259 (2011)
14. Hjort, N.L.: Nonparametric bayes estimators based on beta processes in models for life history data. Ann. Stat. **18**(3), 1259–1294 (1990)
15. Jordan, M.I.: Hierarchical models, nested models and completely random measures. Frontiers of Statistical Decision Making and Bayesian Analysis: In Honor of James O Berger, pp. 207–218. Springer, New York (2010)
16. Kingman, J.: Completely random measures. Pac. J. Math. **21**(1), 59–78 (1967)
17. Lijoi, A., Nipoti, B.: A class of hazard rate mixtures for combining survival data from different experiments. J. Am. Stat. Assoc. **109**(506), 802–814 (2014)
18. Lijoi, A., Prünster, I.: Models beyond the Dirichlet process. In: Hjort, N.L., Holmes, C.C., Müller, P., Walker, S.G. (eds.) Bayesian Nonparametrics, pp. 80–136. Cambridge University Press, Cambridge (2010)
19. Lijoi, A., Mena, R.H., Prünster, I.: Bayesian nonparametric analysis for a generalized Dirichlet process prior. Stat. Inference Stoch. Process. **8**(3), 283–309 (2005)
20. Lijoi, A., Mena, R.H., Prünster, I.: Bayesian nonparametric estimation of the probability of discovering new species. Biometrika **94**(4), 769–786 (2007)
21. McFadden, D.: A method of simulated moments for estimation of discrete response models without numerical integration. Econometrica **57**(5), 995–1026 (1989)
22. Nieto-Barajas, L.E.: Bayesian semiparametric analysis of short- and long-term hazard ratios with covariates. Comput. Stat. Data Anal. **71**, 477–490 (2014)

23. Nieto-Barajas, L.E., Prünster, I.: A sensitivity analysis for Bayesian nonparametric density estimators. Stat. Sin. **19**(2), 685 (2009)
24. Nieto-Barajas, L.E., Walker, S.G.: Markov beta and gamma processes for modelling hazard rates. Scand. J. Stat. **29**(3), 413–424 (2002)
25. Nieto-Barajas, L.E., Walker, S.G.: Bayesian nonparametric survival analysis via Lévy driven Markov processes. Stat. Sin. **14**(4), 1127–1146 (2004)
26. Nieto-Barajas, L.E., Prünster, I., Walker, S.G.: Normalized random measures driven by increasing additive processes. Ann. Stat. **32**(6), 2343–2360 (2004)
27. Regazzini, E., Lijoi, A., Prünster, I.: Distributional results for means of normalized random measures with independent increments. Ann. Stat. **31**(2), 560–585 (2003)
28. Ren, B., Bacallado, S., Favaro, S., Holmes, S., Trippa, L.: Bayesian nonparametric ordination for the analysis of microbial communities. J. Am. Stat. Assoc. (2017)
29. Teh, Y.W., Gorur, D.: Indian buffet processes with power-law behavior. In: Advances in Neural Information Processing Systems, pp. 1838–1846 (2009)
30. Walker, S.G., Damien, P.: Representations of Lévy processes without Gaussian components. Biometrika **87**(2), 477–483 (2000)

Springer