Bayesian Analysis (2023) **18**, Number 1, pp. 253–286

On Quasi-Infinitely Divisible Random Measures*

Riccardo Passeggeri†

Abstract. Quasi-infinitely divisible (QID) distributions have been recently introduced. A random variable X is QID if and only if there exist two infinitely divisible random variables Y and Z s.t. $X + Y \stackrel{d}{=} Z$ and Y is independent of X. In this work, we introduce QID completely random measures (CRMs) and show that a certain family of QID CRMs is dense in the space of all CRMs with respect to convergence in distribution. We further demonstrate that the elements of this family possess a Lévy-Khintchine formulation and that there exists a one to one correspondence between their law and certain characteristic pairs. We prove the same results also for the class of point processes with independent increments. In the second part of the paper, we show the relevance of these results in a general Bayesian nonparametric framework suitable for topic modelling, and provide a truncation error analysis.

Keywords: infinite divisibility, completely random measure, dense class, nonparametric Bayesian analysis, automatic conjugacy, truncation error analysis.

MSC2020 subject classifications: 60E07, 60G57, 60A10, 62F15.

1 Introduction

1.1 Motivation

Imagine that we want to understand the topics of a corpus of documents, but we only have access to some of the documents. How do we proceed? First, a topic can be considered as a certain collection of words, and thus by possessing only some of the documents of the corpus we only observe certain topics. Since we do not observe all the topics, a possible solution to our inferential problem is to model each topic as having a frequency. The higher is the frequency with which a topic appears in the observed documents, the higher is the probability that topic will reappear again in the corpus. Thus, the objective reduces to learn the content of the entire corpus by the topics and their respective frequencies exhibited by each of the observed documents.

Hence, the prior can be represented as

$$\Theta = \sum_{j=1}^{K} \theta_k \delta_{\psi_k},$$

where the cardinality K may be either finite or infinite and where (ψ_k, θ_k) is a pair consisting of the topic ψ_k, which belongs to some space Ψ of topics, and its frequency

*The research developed in this paper is supported by the EPSRC (award ref. 1643696) at Imperial College London and by the Fondation Sciences Mathématiques de Paris (FSMP) fellowship, held at LPSM (Sorbonne University).

†Department of Mathematics, Imperial College London, riccardo.passeggeri14@imperial.ac.uk

 https://doi.org/10.1214/21-BA1289

(or rate) θ_k. Assume that we posses m documents. These observed documents can be represented by $X_1, \dots, X_m$, where

$$X_l = \sum_{j=1}^{K_l} x_{l,k} \delta_{\psi_k},$$

and where $x_{l,k}$ represents the frequency of the ψ_k topic in the l document and K_l is a.s. finite, for $l = 1, \dots, m$. The topics ψ_k's are in general random because we do not know them a priori. However, in some cases we might know some of the topics a priori, for example if we know the title of the corpus or we know that the corpus consists of articles of the *Financial Times*. Thus, in these cases some of the ψ_k's are *fixed*.

From a mathematical point of view, Θ and $X_1, \dots, X_m$ are random measures, namely random elements whose values are measures. If in addition we assume that the presence and the frequency of a topic is independent from the presence and the frequency of another topic, then we talk about completely random measures.

Completely random measures (CRMs), also known as independently scattered random measures or random measures with independent increments, and their normalization have been vastly used in nonparametric Bayesian analysis. In Lijoi and Prünster (2010), the authors provide a unifying treatment of nonparametric Bayesian analysis modeling under the general CRM framework, while Regazzini et al. (2003) and James et al. (2009), provide general results for normalized CRMs, to name a few. In many cases, including the example just presented, the number of latent traits in a data set is expected to increase as the size of the data increases, *e.g.* the number of different topics observed is expected to increases as m increases. CRMs possess this desirable property due to their infinite dimensionality, in other words by setting K equal to infinite there are countably infinite many latent topics to be discovered.

CRMs have a long history. In 1967, Kingman Kingman (1993) proved a very appealing and useful representation theorem for all CRMs. He showed that any CRM ξ is almost surely given by the sum of three components: one deterministic, one concentrated on a fixed set of atoms, and one concentrated on a random set of atoms. He further showed that the last component, which he called the ordinary component, is fully determined by a Poisson point process: $\xi_{ord}(B) = \int_{(0,\infty)} x\eta(B \times dx)$, where η is a Poisson point process on $S \times (0, \infty)$. The Poisson point process is the prime example of infinitely divisible CRM.

In many cases, the fixed component is left out from the analysis, even though it has a specific and unique modelling role, because it does not have certain useful properties which are possessed by the ordinary component, like having an explicit formulation for the characteristic function (called the Lévy-Khintchine formulation). Moreover, the infinite dimensionality of CRMs typically poses a number of practical challenges regarding posterior inference and estimation, including the need to derive ad-hoc algorithms (since most of the algorithms requires finite dimensionality of the CRMs, see Lee et al. (2019) and references therein).

The goal of this paper is to provide a general recipe to obtain finite dimensional CRMs which possess many useful properties, including the Lévy-Khintchine formulation, and which approximate any CRM (with any Lévy measure and any fixed component).

1.2 Contributions

Infinitely divisible (ID) distributions have an even longer history than the one of CRMs which goes back to the work of P. Lévy, A. N. Kolmogorov and B. de Finetti, among others. They constitute one of the most studied classes of probability distributions. One of their most attractive properties is that their characteristic function has an explicit formulation, called the Lévy-Khintchine formulation, written in terms of three mathematical objects. These are the drift, which is a real valued constant, the Gaussian component, which is a non-negative constant, and the Lévy measure, which is a measure on $\mathbb{R}$ satisfying an integrability condition and with no mass at $\{0\}$. Gaussian and Poisson distributions are examples of this class.

In 2018, in Lindner et al. (2018) Sato, Lindner and Pan introduced the class of quasi-infinitely divisible (QID) distributions. A random variable X is QID (namely has a QID distribution) if and only if there exist two ID random variables Y and Z s.t. $X+Y \stackrel{d}{=} Z$ and Y is independent of X. QID distributions are like ID distributions except for the fact that the Lévy measure is now allowed to take negative values. In other words, a QID distribution has a Lévy-Khintchine formulation which is uniquely determined by a drift, a Gaussian component and by a 'signed measure' (more precisely a real valued set function) called the quasi-Lévy measure. Any ID distribution is QID, but the converse is not always true.

In Lindner et al. (2018), the authors show that QID distributions are dense in the space of all probability distributions with respect to weak convergence and that distributions concentrated on the integers (or any shift and dilation of them) are QID if and only if their characteristic functions have no zeros, among other results. Further theoretical results have been achieved in Berger (2019); Khartov (2019); Passeggeri (2020c,b,a). In Passeggeri (2020c), the QID framework is extended to real-valued random noises and stochastic processes. QID distributions have already shown to have an impact in various fields: from mathematical physics, see Chhaiba et al. (2016) and Demni and Mouayn (2015), to number theory, see Nakamura (2015), and to insurance mathematics, see Zhang et al. (2014).

The first main contribution of this paper is the denseness result for QID random measures. We prove that a certain class of QID completely random measures (CRMs), which we denote by $\mathcal{A}$, is dense with respect to convergence in distribution (precisely in both weak and vague convergence) in the space of all CRMs. This result extends the denseness result in Lindner et al. (2018) to the infinite dimensional setting of CRMs.

The class $\mathcal{A}$ has quite remarkable features. First, it has finite fixed atoms. Second, these random measures are almost surely finite and even more their ordinary component has finite Lévy measure. Third, for the elements of this class, we are able to show an explicit spectral representation, namely the Lévy-Khintchine formulation, and prove that there exists a unique one-to-one correspondence between them and pairs of deterministic measures satisfying certain conditions, which we call characteristic pairs. We prove all these results also for the class of point processes with independent increments, of which the Poisson point process is an example.

With these results this paper shows that the fixed component of a CRM, which has been left out in Kingman's analysis and in the theory of CRMs in general, have nice representation results as the widely studied ordinary component. Thus, not only there is no real need of leaving out of the analysis the fixed component (as Kingman graphically says, fixed atoms can be removed by simple surgery), but this might also be dangerous since in many applications the fixed component has an irreplaceable role. This will also appear evident in the Bayesian setting we discuss in this work (see also Broderick et al. (2018)).

The second main contribution of the paper is the investigation of the impact of these results in the nonparametric Bayesian statistical framework presented by Broderick, Wilson and Jordan in Broderick et al. (2018) based on CRMs (see also Campbell et al. (2019)). In particular, we consider priors to be given by elements in $\mathcal{A}$ (with quasi-Lévy measure having a particular structure). First, we show that they are dense in the space of priors considered in Broderick et al. (2018) and Campbell et al. (2019) with respect to convergence in distribution, thus showing also that our denseness result is flexible enough to adjust to various assumptions/settings. Second, we present explicit formulations for their posterior distributions. Third, when focusing on point processes, we prove automatic conjugacy for all the elements of $\mathcal{A}$ under the only condition that the characteristic function of the posterior distribution has no zeros. This condition is new and different from the usual condition based on the exponential structure of the likelihood and of the prior.

We remark that the choice of the setting in Broderick et al. (2018) is not *ad hoc*, the general nature of our results allow them to be applied to more general settings and to answer more general questions.

In the last section we present a truncation error analysis. Building on the class of CRMs in $\mathcal{A}$, we provide a truncation procedure for any CRM ξ. The truncated random measure ξ_n and the tail measure $\xi - \xi_n$ are independent CRMs. Moreover, the ξ_n is composed by the atoms of ξ with weight greater than $1/n$ (and in some cases lying in a bounded region). Our truncation procedure, which can be seen as a generalization of the ε-approximation of Argiento et al. (2016), is remarkably not of the forms discussed in Arbel and Prünster (2017); Campbell et al. (2019); Nguyen et al. (2021); Lee et al. (2019), because it is not obtained by truncating a series representation of the CRM, with stochastically decreasing weights, or by considering a finite measure with n atoms and iid weights converging in distribution to the CRM as n tends to infinity. Moreover, we present an upper bound for the L_1 error on the marginal likelihood when truncated CRMs are used in a general hierarchical Bayesian setting. Our finite-dimensional approximation is more general than the ones presented in Campbell et al. (2019); Nguyen et al. (2021); Lee et al. (2019) because it applies to any CRM without any assumption.

The paper is structured as follows. Section 2 concerns with the notations and some preliminaries. In Section 3 we provide the denseness results for CRMs and in Subsection 3.2 the one for point processes with independent increments. In Section 4, we show various properties for the classes of QID CRMs and QID point processes presented in

Section 3. In particular we present the Lévy-Khintchine formulation and the one-to-one correspondence of these random measures with their unique characteristic pair. In Section 5, we present the Bayesian setting and the relative results: computation of the posterior, convergence results for the posterior, and automatic conjugacy. In Section 6 we present the truncation error analysis.

2 Notation and preliminaries

By a measure on a measurable space $(V, \mathcal{G})$ we always mean a positive measure on $(V, \mathcal{G})$, *i.e.* a $[0, \infty]$-valued σ-additive set function on $\mathcal{G}$ that assigns the value 0 to the empty set. For a non-empty set V, by $\mathcal{B}(V)$ we mean the Borel σ-algebra of V. The law and the characteristic function of a random variable X will be denoted by $\mathcal{L}(X)$ and by $\hat{\mathcal{L}}(X)$, respectively. For two measurable spaces $(V, \mathcal{G})$ and $(T, \mathcal{F})$, we denote by $\mathcal{G} \otimes \mathcal{F}$ the product σ-algebra of $\mathcal{G}$ and $\mathcal{F}$, and by $\mathcal{G} \times \mathcal{F}$ their Cartesian product. Let us recall some definitions.

Definition 2.1 (extended signed measure)**.** Given a measurable space $(V, \mathcal{G})$, that is, a set V with a σ-algebra $\mathcal{G}$ on it, an extended signed measure is a function $\mu : \mathcal{G} \to \mathbb{R} \cup \{\infty, -\infty\}$ s.t. $\mu(\emptyset) = 0$ and μ is σ-additive, that is, it satisfies the equality $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$ where the series on the right must converge in $\mathbb{R} \cup \{\infty, -\infty\}$ absolutely (namely the value of the series is independent of the order of its elements), for any sequence $A_1, A_2, \dots$ of disjoint sets in $\mathcal{G}$.

As a consequence any extended signed measure can take plus or minus infinity as value, but not both. In this work, we use the term 'signed measure' for an extended signed measure. Further, the *total variation* of a signed measure μ is defined as the measure $|\mu| : \mathcal{G} \to [0, \infty]$ defined by $|\mu|(A) := \sup \sum_{j=1}^{\infty} |\mu(A_j)|$, where the supremum is taken over all the partitions $\{A_j\}$ of $A \in \mathcal{G}$. The total variation $|\mu|$ is finite if and only if μ is finite. Let us recall the definition of a signed bimeasure.

Definition 2.2 (Signed bimeasure)**.** Let $(V, \mathcal{G})$ and (T, Γ) be two measurable spaces. A signed bimeasure is a function $M : \mathcal{G} \times \Gamma \to [-\infty, \infty]$ such that:

(i) the function $A \to M(A, B)$ is a signed measure on $\mathcal{G}$ for every $B \in \Gamma$,

(ii) the function $B \to M(A, B)$ is a signed measure on Γ for every $A \in \mathcal{G}$.

Let S be a separable and complete metric space with Borel σ-algebra $\mathbf{S}$ and let $\hat{\mathbf{S}}$ be the ring composed by bounded Borel sets in S. The triplet $(S, \mathbf{S}, \hat{\mathbf{S}})$ is called localised Borel space (see page 19 in Kallenberg (2017)).

Definition 2.3 (random measure)**.** A random measure ξ on S, with underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$, is a function $\Omega \times \mathbf{S} \to [0, \infty]$, such that $\xi(\omega, B)$ is a $\mathcal{F}$-measurable in $\omega \in \Omega$ for fixed B and a locally finite measure in $B \in \mathbf{S}$ for fixed ω.

Definition 2.4 (completely random measure)**.** A completely random measure (CRM) ξ is a random measure s.t. for any disjoint $B_1, B_2, \dots, B_k \in \mathbf{S}$, $k \in \mathbb{N}$, the random variables $\xi(B_1), \xi(B_2), \dots, \xi(B_k)$ are independent. CRMs are also called independently scattered random measure or random measure with independent increments.

Definition 2.5 (diffuse random measure)**.** Using the notation of the previous definition, we say that a random measure ξ on S is diffuse if $\xi(\omega, B)$ is a locally finite diffuse measure in $B \in \mathbf{S}$ for fixed ω (namely $\xi(\omega, \{x\}) = 0$ for every $x \in S$).

Remark 2.6. The term *finite* for random measures stands for *a.s. finite.* Thus, for a *finite random measure* we mean an *a.s. finite random measure.*

Definition 2.7 (fixed atoms, atomless random measure)**.** For a random measure ξ on a Polish space V, $x \in V$ is a fixed atom of ξ if and only if $\mathbb{P}(\xi(\{x\}) > 0) > 0$. Further, a random measure ξ is called atomless if it has no fixed atoms, namely if $\xi(\{x\}) \overset{a.s.}{=} 0$ for every $x \in V$.

The atomless condition is for random measures what the continuity in probability is for continuous time stochastic processes. We remark that an atomless random measure is not necessarily a diffuse random measure (see Corollary 12.11 in Kallenberg (2002)). For example, think of a Poisson point process with $\mathbb{E}[\xi(s)] \equiv 0$, like the homogeneous Poisson point process, which has no fixed atoms but it is not diffuse. Next, we introduce the QID distribution (see Lindner et al. (2018)).

Definition 2.8. Let $\mathcal{B}_r(\mathbb{R}) := \{B \in \mathcal{B}(\mathbb{R}) | B \cap (-r, r) = \emptyset\}$ for $r > 0$ and $\mathcal{B}_0(\mathbb{R}) := \bigcup_{r>0} \mathcal{B}_r(\mathbb{R})$ be the class of all Borel sets that are bounded away from zero. Let $\nu : \mathcal{B}_0(\mathbb{R}) \to \mathbb{R}$ be a function such that $\nu_{|\mathcal{B}_r(\mathbb{R})}$ is a finite signed measure for each $r > 0$ and denote the total variation, positive and negative part of $\nu_{|\mathcal{B}_r(\mathbb{R})}$ by $|\nu_{|\mathcal{B}_r(\mathbb{R})}|$, $\nu^+_{|\mathcal{B}_r(\mathbb{R})}$ and $\nu^-_{|\mathcal{B}_r(\mathbb{R})}$ respectively. Then the total variation $|\nu|$, the positive part ν^+ and the negative part ν^- of ν are defined to be the unique measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ satisfying

$$|\nu|(\{0\}) = \nu^+(\{0\}) = \nu^-(\{0\}) = 0$$

$$\text{and} \quad |\nu|(A) = |\nu_{|\mathcal{B}_r(\mathbb{R})}(A)|, \ \nu^+(A) = \nu^+_{|\mathcal{B}_r(\mathbb{R})}(A), \ \nu^-(A) = \nu^-_{|\mathcal{B}_r(\mathbb{R})}(A),$$

for $A \in \mathcal{B}_r(\mathbb{R})$, for some $r > 0$.

Lemma 2.14 in Passeggeri (2020c) shows the uniqueness of $|\nu|$, ν^+ and ν^-.

Definition 2.9 (quasi-Lévy type measure, quasi-Lévy measure, QID distribution)**.** A quasi-Lévy type measure is a function $\nu : \mathcal{B}_0(\mathbb{R}) \to \mathbb{R}$ satisfying the condition in Definition 2.8 and such that its total variation $|\nu|$ satisfies $\int_{\mathbb{R}} (1 \wedge x^2)|\nu|(dx) < \infty$. Let μ be a probability distribution on $\mathbb{R}$. We say that μ is quasi-infinitely divisible if its characteristic function has a representation

$$\hat{\mu}(\theta) = \exp\left(i\theta\gamma - \frac{\theta^2}{2}a + \int_{\mathbb{R}} \left(e^{i\theta x} - 1 - i\theta\tau(x)\right)\nu(dx)\right),$$

where $a, \gamma \in \mathbb{R}$ and ν is a quasi-Lévy type measure. The characteristic triplet (γ, a, ν) of μ is unique, and a and γ are called the Gaussian component and the drift of μ, respectively. A quasi-Lévy type measure ν is called quasi-Lévy measure, if additionally there exist a quasi-infinitely divisible distribution μ and some $a, \gamma \in \mathbb{R}$ such that (γ, a, ν) is the characteristic triplet of μ. We call ν the quasi-Lévy measure of μ.

As pointed out in Example 2.9 of Lindner et al. (2018), a quasi-Lévy measure is always a quasi-Lévy type measure, while the converse is not true. In this work the characteristic triplet have always the same order: drift, Gaussian component, and (quasi) Lévy measure.

Definition 2.10 (QID random measure)**.** Let Λ be a random measure. If $\Lambda(A)$ is a QID random variable, for every $A \in \mathbf{S}$, then we call Λ a QID random measure.

3 The denseness results

In this section we present the denseness with respect to convergence in distribution of QID CRMs in the space of all CRMs and of QID point processes with independent increments in the space of all point processes with independent increments. Let us start with some preliminaries. Recall that S is a separable and complete metric space with Borel σ-algebra $\mathbf{S}$ and $\hat{\mathbf{S}}$ is the ring composed by bounded Borel sets in S. Let $\hat{C}_S$ be the space of all bounded continuous functions $f : S \to \mathbb{R}_+$ with bounded support. Let $\mathcal{M}_S$ be the space of locally finite measures, namely $\mu \in \mathcal{M}_S$ if $\mu(B) < \infty$ for every $B \in \hat{\mathbf{S}}$. The space $\mathcal{M}_S$ might be endowed with the vague topology, denoted by $\mathbf{B}_{\mathcal{M}_S}$, generated by the integration maps $\pi_f : \mu \mapsto \int f(x)\mu(dx)$, for all $f \in \hat{C}_S$. The vague topology is the coarsest topology making all π_f continuous. The measurable space $(\mathcal{M}_s, \mathbf{B}_{\mathcal{M}_S})$ is a Polish space. The associated notion of vague convergence denoted by $\mu_n \xrightarrow{v} \mu$ is defined by the condition $\int f(x)\mu_n(dx) \to \int f(x)\mu(dx)$ for all $f \in \hat{C}_S$.

An equivalent definition of random measure (see Definition 2.3) is the following: a random measure ξ is a measurable mapping from $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\mathcal{M}_S, \mathcal{B}_{\mathcal{M}_S})$, where $\mathcal{B}_{\mathcal{M}_S}$ is the topology generated by all projection maps $\pi_B : \mu \mapsto \mu(B)$ with $B \in \mathbf{S}$, or, equivalently, by all integration maps π_f with measurable $f \geq 0$. From Lemma 4.1 in Kallenberg (1983) or Theorem 4.2 in Kallenberg (2017), we know that $\mathcal{B}_{\mathcal{M}_S}$ and $\mathbf{B}_{\mathcal{M}_S}$ coincide. Hence it is equivalent to consider a random measure as a measurable mapping from $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\mathcal{M}_S, \mathbf{B}_{\mathcal{M}_S})$ or to $(\mathcal{M}_S, \mathcal{B}_{\mathcal{M}_S})$.

The convergence in distribution of ξ_n to ξ means that $\mathbb{E}[g(\xi_n)] \to \mathbb{E}[g(\xi)]$ for every real valued bounded continuous function g on $\mathcal{M}_S$, or equivalently that $\mathcal{L}(\xi_n) \xrightarrow{w} \mathcal{L}(\xi)$, where for any bounded measures ν_n and ν, the weak convergence $\nu_n \xrightarrow{w} \nu$ stands for $\int g(y)\nu_n(dy) \to \int g(y)\nu(dy)$ for all g as above. We write $\xi_n \xrightarrow{vd} \xi$ to stress that the convergence of distribution is for random measures considered as random elements in the space $\mathcal{M}_S$ with vague topology. As mentioned in the previous section, in this setting an atom of a random measure ξ is an element $s \in S$ such that $\mathbb{P}(\xi(\{s\}) > 0) > 0$.

3.1 The denseness result for QID CRMs

In this section we present the denseness of QID CRMs in the space of all CRMs. We start with following denseness result which extends Theorem 4.1 in Lindner et al. (2018).

Theorem 3.1. *Let A be a connected interval of the real line. The class of QID distributions with finite quasi-Lévy measure, zero Gaussian component and with support on A is dense in the class of probability distributions with support on A with respect to weak convergence.*

Proof. The arguments of this proof extend the arguments in the proof of Theorem 4.1 in Lindner et al. (2018). First, we prove the result when A is bounded. Let A be a finite closed interval, thus $A = [k, c]$ for some $k, c \in \mathbb{R}$. Let μ be a probability distribution with support $[k, c]$. For $n \in \mathbb{N}$, let $b_{j,n} = k + (c-k)j/2n^2$, $j \in \{0, \dots, 2n^2\}$ and define the discrete distribution μ_n concentrated on the lattice $\{b_{0,n}, \dots, b_{2n^2,n}\}$ by

$$\mu_n(\{b_{j,n}\}) = \begin{cases} \mu((-\infty, b_{0,n}]), & j = 0, \\ \mu((b_{j-1,n}, b_{j,n}]), & j = 1, \dots, 2n^2 - 1, \\ \mu((b_{2n^2-1,n}, \infty)), & j = 2n^2. \end{cases} \tag{1}$$

Then, $\mu_n \xrightarrow{w} \mu$ as $n \to \infty$. Observe that μ_n is the probability distribution of a random variable with values on $\{b_{0,n}, \dots, b_{2n^2,n}\} \subset [k, c]$. It remains to prove that each μ_n is a weak limit of QID distributions with finite quasi-Lévy measure, zero Gaussian component and with support on $[k, c]$. W.l.o.g. assume that the approximating sequence of distributions σ is such that $\sigma(\{b_{j,n}\}) > 0$ for every $j \in \{0, \dots, 2n^2\}$. Assume that the characteristic function $\hat{\sigma}$ has zeros (in the other case we can directly use Corollary 3.10 in Lindner et al. (2018) to conclude). Let X be a random variable with distribution σ and define $Y = \frac{(X-k)2n^2}{c-k}$. Then, Y is concentrated on $\{0, \dots, 2n^2\}$ with masses $a_j = \mathbb{P}(Y = j) > 0$ for $j = 0, \dots, 2n^2$, and its characteristic function has zeroes. Then, the polynomial $f(w) = \sum_{j=0}^{2n^2} a_j w^j$ has zeroes on the unit circle. Factorizing, we obtain $f(w) = a_{2n^2} \prod_{j=1}^{2n^2} (w - \xi_j)$, where ξ_j, $j = 1, \dots, 2n^2$, denote the complex roots. Let $f_h(w) = a_{2n^2} \prod_{j=1}^{2n^2} (w - \xi_j - h)$, where $w \in \mathbb{C}$ and $h > 0$. Then, for small enough h, f_h is a polynomial with real coefficients, namely $f_h(w) = \sum_{j=0}^{2n^2} a_{h,j} w^j$ with $a_{h,j} \in \mathbb{R}$. Observe that for small enough h, $a_{h,j}$ and a_j will be close, so $a_{h,j} > 0$. Now, let Z_h be a random variable with distribution $\sigma_h = \left(\sum_{j=0}^{2n^2} a_{h,j}\right)^{-1} \sum_{j=0}^{2n^2} a_{h,j}\delta_j$ and let $X_h = \frac{Z_h(c-k)}{2n^2} + k$. Observe that, for every $h > 0$, X_h is random variable with values on the lattice $\{b_{0,n}, \dots, b_{2n^2,n}\}$ and its characteristic function has no zeros, and that $X_h \xrightarrow{d} X$ as $h \searrow 0$. Finally, by Corollary 3.10 in Lindner et al. (2018) we know that X_h is QID with finite quasi-Lévy measure and zero Gaussian component.

Observe that if A is a bounded open interval, say $A = (k', c')$ for some $c, k \in \mathbb{R}$, then the above arguments apply. Let μ be a probability distribution with support (k', c'). For any $n \in \mathbb{N}$ let $k'_n = k' + \frac{(c'-k')}{2n^2}$ and $c'_n = c' - \frac{(c'-k')}{2n^2}$ and let $b_{j,n} = k'_n + (c'_n - k'_n)j/2n^2$, $j \in \{0, \dots, 2n^2\}$ and define the discrete distribution μ_n concentrated on the lattice $\{b_{0,n}, \dots, b_{2n^2,n}\}$ as in (1). Then, $\mu_n \xrightarrow{w} \mu$ as $n \to \infty$ and, applying the same reaming arguments (in which n is fixed) for k'_n and c'_n instead of k and c, we obtain the result for A bounded and open.

Let now A be an unbounded interval of the form $A = [k, \infty)$ for some $k \in \mathbb{R}$. Let μ be a probability distribution with support on $[k, \infty)$. For $n \in \mathbb{N}$, let $b_{j,n} = k + j/n$, $j \in \{0, \dots, 2n^2\}$ and define the discrete distribution μ_n concentrated on the lattice $\{b_{0,n}, \dots, b_{2n^2,n}\}$ as in (1). Then, $\mu_n \xrightarrow{w} \mu$ as $n \to \infty$. Using the notation above, let X be a random variable with distribution σ and define $Y = (X-k)n$. Then,

Y is concentrated on $\{0, \ldots, 2n^2\}$ with masses a_j and its characteristic function has zeroes by assumption. We proceed as before. Thus, for small enough h, we obtain a polynomial with real coefficients f_h, namely $f_h(w) = \sum_{j=0}^{2n^2} a_{h,j} w^j$ with $a_{h,j} \in \mathbb{R}$ and $a_{h,j} > 0$, for small enough h. Then, let Z_h be a random variable with distribution $\sigma_h = \left(\sum_{j=0}^{2n^2} a_{h,j}\right)^{-1} \sum_{j=0}^{2n^2} a_{h,j}\delta_j$ and let $X_h = \frac{Z_h}{n} + k$. Then, X_h is random variables with support on $\{b_{0,n}, \ldots, b_{2n^2,n}\} \subset [k, \infty)$ and its characteristic function has no zeros, and that $X_h \xrightarrow{d} X$ as $h \searrow 0$. Hence, by Corollary 3.10 in Lindner et al. (2018) we obtain the result. Similarly we obtain the result for (k', ∞), for $(-\infty, c]$ and for $(-\infty, c')$, where $k', c, c' \in \mathbb{R}$. $\square$

Recall that the Lévy-Prokhorov metric (or better just Lévy metric since we work on $\mathbb{R}$) for two probability distributions F and G on $\mathbb{R}$ is defined as

$$\rho(F, G) := \inf \{\varepsilon > 0 \,|\, F(x - \varepsilon) - \varepsilon \leq G(x) \leq F(x + \varepsilon) + \varepsilon \text{ for all } x \in \mathbb{R}\}.$$

Lemma 3.2. *Let F and G be any two probability distributions on $\mathbb{R}$ and let $F_c(x) := F(\frac{x}{c})$ and $G_c(x) := G(\frac{x}{c})$ where $c \in \mathbb{R} \setminus \{0\}$. For every positive constant $c \leq 1$ we have that $\rho(F_c, G_c) \leq \rho(F, G)$.*

Proof. Let c be any positive constant $c \leq 1$. Observe that $F_c(x-\varepsilon) = F(\frac{x-\varepsilon}{c}) \leq F(\frac{x}{c}-\varepsilon)$ and similarly we have that $F_c(x + \varepsilon) \geq F(\frac{x}{c} + \varepsilon)$. This implies that if $\varepsilon > 0$ satisfies $F(x - \varepsilon) - \varepsilon \leq G(x) \leq F(x + \varepsilon) + \varepsilon$ for all $x \in \mathbb{R}$, then it also satisfies $F_c(x - \varepsilon) - \varepsilon \leq G_c(x) \leq F_c(x + \varepsilon) + \varepsilon$ for all $x \in \mathbb{R}$. Then, we have

$$\begin{aligned}\rho(F_c, G_c) &= \inf \{\varepsilon > 0 \,|\, F_c(x - \varepsilon) - \varepsilon \leq G_c(x) \leq F_c(x + \varepsilon) + \varepsilon \text{ for all } x \in \mathbb{R}\} \\ &\leq \inf \{\varepsilon > 0 \,|\, F(x - \varepsilon) - \varepsilon \leq G(x) \leq F(x + \varepsilon) + \varepsilon \text{ for all } x \in \mathbb{R}\} = \rho(F, G). \qquad \square\end{aligned}$$

Observe that for two real valued random variables X and Y the above lemma affirms that for any $0 < c \leq 1$ we have that $\rho(cX, cY) \leq \rho(X, Y)$. Moreover, from condition 3) of the section "Lévy metric" in Zolotarev (2001) (page 405) given any probability distributions on $\mathbb{R}$ $F_1, \ldots, F_k, G_1, \ldots, G_k$, where $k \in \mathbb{N}$, we have that

$$\rho(F_1 * \cdots * F_k, G_1 * \cdots * G_k) \leq \sum_{j=1}^{k} \rho(F_j, G_j). \tag{2}$$

For the next two results denote by S_n the sequence of bounded sets (*i.e.* $S_n \in \hat{\mathbf{S}}$) s.t. $S_n \uparrow S$. Notice that such sequence exists by the definition of $\hat{\mathbf{S}}$, see page 19 in Kallenberg (2017).

Proposition 3.3. *Consider an atomless CRM α with corresponding unique pair (γ, F). Let $\gamma_n(A) = \gamma(S_n \cap A)$ and let $F_n(C) = F(C \cap (S_n \times (\frac{1}{n}, \infty)))$, for every $A \in \mathbf{S}$, $C \in \mathbf{S} \otimes \mathcal{B}((0, \infty))$ and $n \in \mathbb{N}$. Then, γ_n and F_n are finite measures and there exists a sequence of atomless finite CRMs α_n with pair (γ_n, F_n) s.t. $\alpha_n \xrightarrow{d} \alpha$.*

Proof. From Kingman's representation theorem (see Kingman (1993) and see also Corollary 12.11 in Kallenberg (2002) and Corollary 3.21 in Kallenberg (2017)), we have that every atomless CRM α has the following representation:

$$\alpha = \gamma + \int_0^\infty \int_S x\delta_s \eta(ds\, dx), \quad \text{a.s.} \tag{3}$$

for some non-random diffuse measure $\gamma \in \mathcal{M}_S$ and a Poisson process η on $S \times (0, \infty)$ with intensity F satisfying

$$\int_0^\infty (1 \wedge x) F(A \times dx) < \infty, \tag{4}$$

for every $A \in \hat{\mathbf{S}}$. In particular, for every $B \in \mathbf{S}$ we have that $\alpha(B) < \infty$ if and only if $\gamma(B) < \infty$ and condition (4) holds for $B \in \mathbf{S}$ (see Corollary 12.11 in Kallenberg (2002)). Further, notice that the above formulation implies that for every $A \in \mathbf{S}$ and $f \in \hat{C}_S$

$$\alpha(A) = \gamma(A) + \int_0^\infty x\eta(A \times dx) \;\text{ and }\; \alpha f = \gamma f + \int_0^\infty \int_S x f(s) \eta(ds\, dx), \;\text{ a.s..}$$

Moreover, the unique one to one correspondence between α and (γ, F) is shown in Theorem 3.20 of Kallenberg (2017). It is possible to see that γ_n and F_n are measures on $\mathbf{S}$ and on $\mathbf{S} \otimes \mathcal{B}((0, \infty))$, respectively. In particular, since $\alpha(A) < \infty$ for every $A \in \hat{\mathbf{S}}$ then $\gamma(S_n) < \infty$ and

$$\int_0^\infty (1 \wedge x) F(S_n \times dx) < \infty \Rightarrow F(S_n \times (\frac{1}{n}, \infty)) < \infty,$$

for every $n \in \mathbb{N}$. Thus, γ_n and F_n are finite measures, for every $n \in \mathbb{N}$. Now, for every $n \in \mathbb{N}$, let η_n be a Poisson process on $S \times (0, \infty)$ with intensity F_n and let

$$\alpha_n = \gamma_n + \int_0^\infty \int_S x\delta_s \eta_n(ds\, dx).$$

Then, we have that α_n is an atomless CRM and since γ_n and F_n are finite then α_n is finite, for every $n \in \mathbb{N}$ (see Corollary 12.11 in Kallenberg (2002)).

Concerning the stated convergence we have the following. From Lemma 12.2 in Kallenberg (2002) (or from Lemma 3.1 in Kallenberg (2017)) we have that for every $f \in \hat{C}_S$

$$-\log \mathbb{E}\left[\exp\left(-\int f(s)\alpha(ds)\right)\right] = \gamma f + \int_0^\infty \int_S 1 - e^{-x\delta_s f} F(ds\, dx).$$

Hence, by assumption we have that for every $f \in \hat{C}_S$

$$-\log \mathbb{E}\left[\exp\left(-\int f(s)\alpha(ds)\right)\right] + \log \mathbb{E}\left[\exp\left(-\int f(s)\alpha_n(ds)\right)\right]$$

$$= \int_S f(s)\gamma(ds) + \int_0^\infty \int_S 1 - e^{-xf(s)} F(ds\, dx) - \int_{S_n} f(s)\gamma(ds) + \int_{\frac{1}{n}}^\infty \int_{S_n} 1 - e^{-xf(s)} F(ds\, dx)$$

$$= \int_{S \setminus S_n} f(s)\gamma(ds) + \int_0^{\frac{1}{n}} \int_{S_n} 1 - e^{-xf(s)} F(ds\, dx) + \int_0^\infty \int_{S \setminus S_n} 1 - e^{-xf(s)} F(ds\, dx) \to 0,$$

as $n \to \infty$. Then, by point (iii) in Theorem 4.11 in Kallenberg (2017) (see also Lemma 4.24 in Kallenberg (2017)) we obtain that $\alpha_n \xrightarrow{d} \alpha$, as $n \to \infty$. □

From Theorem 7.1 in Kallenberg (1983) we know that a CRM ξ has the following representation

$$\xi \overset{a.s.}{=} \alpha + \sum_{j=1}^{K} \beta_j \delta_{s_j}$$

with $K \in \mathbb{N} \cup \{0\} \cup \{\infty\}$, where $\{s_j : j \geq 1\}$ is the set of fixed atoms of ξ in S, α is an atomless CRM, and β_j, $j \geq 1$, are $\mathbb{R}_+$-valued random variables, which are mutually independent and independent of α. We call $\sum_{j=1}^{K} \beta_j \delta_{s_j}$ the fixed component of ξ. We remark that in the Kingman's representation α is the sum of a deterministic and a ordinary component (see eq. (3)).

Denote by $\mathcal{A}$ the following class of QID CRMs: a CRM ξ belongs to $\mathcal{A}$ if $\xi \overset{a.s.}{=} \alpha + \sum_{j=1}^{K} \beta_j \delta_{s_j}$, where α is an atomless CRM with finite Lévy measure, $\{s_j : j = 1, \dots, K\}$ is a finite set of fixed atoms in S, and β_j, $j \geq 1$, are $\mathbb{R}_+$-valued QID random variables with finite quasi-Lévy measure and zero Gaussian component and that are mutually independent and independent of α.

Since any atomless random measure with independent increments is ID, α is ID. Observe that, in contrast with general CRMs, the elements of $\mathcal{A}$ have that the atomless random measure α has finite Lévy measure, the number of fixed atoms K is finite, and β_j, $j = 1, \dots, K$, are $\mathbb{R}_+$-valued QID random variables with finite quasi-Lévy measure and zero Gaussian component. Thus, the elements of $\mathcal{A}$ are almost surely finite on $\mathbf{S}$. Hence, $\mathcal{A}$ is strictly smaller than the class of QID CRMs, which in turn is strictly smaller than the class of all CRMs. We are ready to present the main result of this section.

Theorem 3.4. *$\mathcal{A}$ is dense in the space of all CRMs with respect to the convergence in distribution.*

Proof. From Theorem 7.1 in Kallenberg (1983) we know that any CRM has the following unique representation

$$\xi \overset{a.s.}{=} \alpha + \sum_{j=1}^{K} \beta_j \delta_{s_j} \tag{5}$$

with $K \leq \infty$, where $\{s_j : j \geq 1\}$ is the set of fixed atoms of ξ, α is a random measure without fixed atoms with independent increments (hence, α is an atomless ID random measure), and β_j, $j \geq 1$, are $\mathbb{R}_+$-valued random variables, which are mutually independent and independent of α.

From Theorem 3.1 with $A = [0, \infty)$, we know that for each β_j there exists a sequence of non-negative QID random variable with zero Gaussian component and finite Lévy measure that converges in distribution to β_j, for every $j \in \mathbb{N}$. Denote by $\beta_{n,j}$ such a sequence. Denote by S_n the sequence of bounded sets s.t. $S_n \uparrow S$ and by (γ, F) be the pair associated to α. Let $\gamma_n(A) = \gamma(S_n \cap A)$ and $F_n(C) = F(C \cap (S_n \times (\frac{1}{n}, \infty)))$, for every

$A \in \mathbf{S}$, $C \in \mathbf{S} \otimes \mathcal{B}((0,\infty))$ and $n \in \mathbb{N}$, as in Proposition 3.3. Then, by Proposition 3.3 there exists a sequence of finite CRMs α_n with pair (γ_n, F_n) s.t. $\alpha_n \xrightarrow{d} \alpha$.

The first step is to show the existence of random measures $\xi_n \in \mathcal{A}$ with ID atomless random measure equal in distribution to α_n, with fixed atoms in $\{s_j : j \geq 1\}$, and weights equal in distributions to $\beta_{n,j}$. The existence is not immediate because we do not know whether the $\beta_{n,j}$ are mutually independent and independent of α_n in the underlying probability space of ξ. This is a classical problem in probability and the solution lies in the construction of a probability space under which these conditions are satisfied, which is given by the 'product' of the probability spaces.

For the sake of clarity and completeness let us write here the arguments. Fix $n \in \mathbb{N}$. Denote the underlying probability spaces of α_n by $(\Omega, \mathcal{F}, \mathbb{P})$ and of the random variable $\beta_{n,j}$ by $(\Omega_j, \mathcal{F}_j, \mathbb{P}_j)$, for $j = 1, \dots, n$. Consider the probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ where $\Omega' = \Omega \times \Omega_1 \times \cdots \times \Omega_n$, $\mathcal{F}' = \mathcal{F} \otimes \mathcal{F}_1 \otimes \cdots \otimes \mathcal{F}_n$ and $\mathbb{P}'$ is the product probability measure of $\mathbb{P}, \mathbb{P}_1, \dots, \mathbb{P}_n$. Let $\alpha'_n(\cdot)(\omega, \omega_1, \dots, \omega_n) := \alpha_n(\cdot)(\omega)$ and let $\beta'_{n,j}(\omega, \omega_1, \dots, \omega_n) := \beta_{n,j}(\omega_j)$, where $j = 1, \dots, n$, for every $(\omega, \omega_1, \dots, \omega_n) \in \Omega'$. Observe that for every $B_1, \dots, B_k \in \mathbf{S}$ and $x_1, \dots, x_k, x_1^{(1)}, \dots, x_k^{(1)}, \dots, x_1^{(n)}, \dots, x_k^{(n)} \in \mathbb{R}_+$ we have that

$$\mathbb{P}'\Big(\alpha'_n(B_1) < x_1, \dots, \alpha'_n(B_k) < x_k, \delta_{s_1}(B_1)\beta'_{n,1} < x_1^{(1)}, \dots, \delta_{s_1}(B_k)\beta'_{n,1} < x_k^{(1)},$$

$$\dots, \delta_{s_n}(B_1)\beta'_{n,n} < x_1^{(n)}, \dots, \delta_{s_n}(B_k)\beta'_{n,n} < x_k^{(n)}\Big) = \mathbb{P}\Big(\alpha(B_1) < x_1, \dots, \alpha(B_k) < x_k\Big),$$

$$\prod_{l=1}^{n} \mathbb{P}_l\Big(\delta_{s_l}(B_1)\beta'_{n,l} < x_1^{(l)}, \dots, \delta_{s_l}(B_k)\beta'_{n,l} < x_k^{(l)}\Big).$$

Now, let

$$\xi_n(\cdot)(\omega') := \alpha'_n(\cdot)(\omega') + \sum_{j=1}^{n} \beta'_{n,j}(\omega')\delta_{s_j}(\cdot), \quad \forall \omega' \in \Omega', \tag{6}$$

where $s_1, \dots, s_n$ are the same as the ones in (5). It is possible to see that, for every $\omega' \in \Omega'$, $\xi_n(\cdot)(\omega')$ is a measure because it is the sum of measures and that, for every $B \in \mathbf{S}$, $\xi_n(B)(\cdot)$ is a measurable function because it is the sum of measurable functions. Thus, ξ_n is a random measure on S and from its definition it is possible to see that it belongs to $\mathcal{A}$.

Since $\beta_{n,j} \xrightarrow{d} \beta_j$ we can choose a subsequence of $\beta_{n,j}$, which by abuse of notation we denote by $\beta_{n,j}$, such that $\rho(\beta_{n,j}, \beta_j) < \frac{1}{n^2}$ for every $j = 1, \dots, n$ and $n \in \mathbb{N}$. From the above arguments there exists a sequence of random measures in $\mathcal{A}$ (with possibly different underlying probability spaces) such that $\xi_n = \alpha'_n + \sum_{j=1}^{n} \beta'_{n,j}\delta_{s_j}$. Thus, using that $\beta'_{n,j} \stackrel{d}{=} \beta_{n,j}$ we obtain that $\rho(\beta'_{n,j}, \beta_j) < \frac{1}{n^2}$ for every $j = 1, \dots, n$ and $n \in \mathbb{N}$.

Now, we need to show that $\xi_n \xrightarrow{vd} \xi$. From Theorem 4.11 in Kallenberg (2017), it is sufficient to show that $\int f(x)\xi_n(dx) \xrightarrow{d} \int f(x)\xi(dx)$ for all $f \in \hat{C}_S$. Since $\alpha'_n \stackrel{d}{=} \alpha_n$ for every $n \in \mathbb{N}$ and $\alpha_n \xrightarrow{d} \alpha$ for every $\omega \in \Omega$ then $\alpha'_n \xrightarrow{d} \alpha$. Further, since α'_n

and α are independent of the corresponding fixed component, this reduces the goal to prove that $\sum_{j=1}^{n} f(s_j)\beta'_{n,j} \xrightarrow{d} \sum_{j=1}^{\infty} f(s_j)\beta_j$ for all $f \in \hat{C}_S$. Let $f \in \hat{C}_S$, hence, f is bounded and has bounded support, and by denoting B the support of f we have that $B \in \hat{\mathbf{S}}$ and so that almost surely $\xi_n(B) < \infty$, $n \in \mathbb{N}$, and $\xi(B) < \infty$. Thus, for each $n \in \mathbb{N}$, $\sum_{j=1}^{n} f(s_j)\beta'_{n,j} < \infty$ a.s. and $\sum_{j=1}^{\infty} f(s_j)\beta_j < \infty$ a.s.. Moreover, notice that it is sufficient to prove the result for any $f \in \hat{C}_S$ with $f(s) \leq 1$ for every $s \in S$. Indeed, consider any $f \in \hat{C}_S$ and let $\bar{C} \in \mathbb{R}_+$ be its bound, then $\sum_{j=1}^{n} f(s_j)\beta'_{n,j} = \bar{C}\sum_{j=1}^{n} \frac{f(s_j)}{\bar{C}}\beta'_{n,j}$ and so if $\sum_{j=1}^{n} \frac{f(s_j)}{\bar{C}}\beta'_{n,j} \xrightarrow{d} \sum_{j=1}^{\infty} \frac{f(s_j)}{\bar{C}}\beta_j$ then $\sum_{j=1}^{n} f(s_j)\beta'_{n,j} \xrightarrow{d} \sum_{j=1}^{\infty} f(s_j)\beta_j$. Now, consider any $f \in \hat{C}_S$ with $f(s) \leq 1$ for every $s \in S$. By the triangular inequality we have that

$$\rho\Bigg(\sum_{j=1}^{n} f(s_j)\beta'_{n,j}, \sum_{j=1}^{\infty} f(s_j)\beta_j\Bigg)$$
$$\leq \rho\Bigg(\sum_{j=1}^{n} f(s_j)\beta'_{n,j}, \sum_{j=1}^{n} f(s_j)\beta_j\Bigg) + \rho\Bigg(\sum_{j=1}^{n} f(s_j)\beta_j, \sum_{j=1}^{\infty} f(s_j)\beta_j\Bigg).$$

The last element converges to zero as $n \to \infty$ because $\sum_{j=1}^{n} f(s_j)\beta_j \xrightarrow{a.s.} \sum_{j=1}^{\infty} f(s_j)\beta_j$ as $n \to \infty$. For the other element, by (2) and by Lemma 3.2 we obtain that

$$\rho\Bigg(\sum_{j=1}^{n} f(s_j)\beta'_{n,j}, \sum_{j=1}^{n} f(s_j)\beta_j\Bigg) \leq \sum_{j=1}^{n} \rho\left(f(s_j)\beta'_{n,j}, f(s_j)\beta_j\right) \leq \sum_{j=1}^{n} \rho\left(\beta'_{n,j}, \beta_j\right) < \frac{1}{n}.$$

Thus, we have that $\sum_{j=1}^{n} f(s_j)\beta'_{n,j} \xrightarrow{d} \sum_{j=1}^{\infty} f(s_j)\beta_j$ as $n \to \infty$, which concludes the proof. $\square$

Remark 3.5. We could alternatively consider an almost sure equality in (6) and then use the existence and uniqueness results for random measures (see Theorem 2.15 and Corollary 2.16 in Kallenberg (2002)) to obtain a random measure almost surely equal to ξ_n. In addition, by the Kolmogorov extension theorem the same arguments of the first part of the above proof hold for the case of n 'equal' to infinity, namely $\xi_n = \alpha'_n + \sum_{j=1}^{\infty} \beta'_{n,j}$. Further, we point out that if ξ is such that the number of fixed atoms in any bounded set (*i.e.* in any $B \in \hat{\mathbf{S}}$) is finite then the number of fixed atoms in the support of every $f \in \hat{C}_S$ is finite, namely $\{s_j : j \geq 1\} \cap \operatorname{supp}(f)$ has finite cardinality, and so the stated result follows directly from the mutual independence of the $\beta'_{n,j}$, $j = 1, \ldots, n$, from the fact that $\beta'_{n,j} \xrightarrow{d} \beta_j$ as $n \to \infty$, for every $j = 1, \ldots, n$ and $n \in \mathbb{N}$, and from the continuous mapping theorem.

Remark 3.6. Let $\mathcal{A}_\infty$ be a class of random measures like $\mathcal{A}$, but such that the ID component is not necessarily finite, *i.e.* the 'α' is not necessarily finite. Then, trivially $\mathcal{A}_\infty$ is dense in the space of all CRMs w.r.t. the convergence in distribution. Indeed, let $\xi = \alpha + \sum_{j=1}^{K} \beta_j \delta_{s_j}$ be any CRM on S. If we know the ID component of ξ, *i.e.* α, and for modelling/theoretical reasons we can take an approximating sequence of *unbounded* ξ_n, then we can define the ξ_n s.t. $\xi_n(\cdot)(\omega') := \tilde{\alpha}'_n(\cdot)(\omega') + \sum_{j=1}^{n} \beta'_{n,j}(\omega')\delta_{s_j}(\cdot)$, $\forall \omega' \in \Omega'$, where $\tilde{\alpha}'_n(\cdot)(\omega, \omega_1, \ldots, \omega_n) := \alpha(\cdot)(\omega)$. Then, $\xi_n \in \mathcal{A}_\infty$ and from the arguments of the proof of Theorem 3.4 it is possible to see that $\xi_n \xrightarrow{d} \xi$.

It is possible to consider also the set of bounded measures, denoted by $\hat{\mathcal{M}}_S$, which can be endowed with the vague topology, as for $\mathcal{M}_S$, but also with the weak topology. The weak topology on $\hat{\mathcal{M}}_S$ is the topology generated by the integration maps π_f for all bounded continuous functions. Then, for random measures $\xi, \xi_1, \xi_2, \ldots$ considered as random elements in $\hat{\mathcal{M}}_S$, endowed with the weak topology, we will denote by $\xi_n \overset{wd}{\to} \xi$ the convergence in distribution. Recall that QID CRMs in $\mathcal{A}$ are a.s. bounded, and so they are random measures in this setting as well. We are now ready to present our next result, which is similar to Theorem 3.4, but applies to $\hat{\mathcal{M}}_S$ and involves both the vague and the weak topology.

Theorem 3.7. *$\mathcal{A}$ is dense in the space of all CRMs, considered as random elements in $\hat{\mathcal{M}}_S$, endowed with either the vague topology or the weak topology, with respect to the convergence in distribution.*

Proof. Consider first the case of $\hat{\mathcal{M}}_S$ endowed with the vague topology. Then, by the same arguments as the ones used in the proof of Theorem 3.4 we obtain the result. For the weak topology case, by the same arguments as the ones used in the proof of Theorem 3.4 we have that $\xi_n \overset{vd}{\to} \xi$. Hence, according to Theorem 4.19 in Kallenberg (2017) it remains to prove that $\xi_n(S) \overset{d}{\to} \xi(S)$, namely that $\alpha'_n(S) + \sum_{j=1}^{n} \beta'_{n,j} \overset{d}{\to} \alpha(S) + \sum_{j=1}^{\infty} \beta_j$. However, this has been proved in the proof of Theorem 3.4 – indeed, consider $f \equiv 1$ and notice that $\xi_n(S)$ and $\xi(S)$ are a.s. finite since ξ_n and ξ are almost surely bounded. Thus, the proof is complete. □

3.2 The denseness result for QID point processes

In this subsection we answer positively the following question: given any point process with independent increments is it possible to find a sequence of QID point processes with independent increments which converges in distribution to it?

Thus, in this subsection we restrict our focus to point processes with independent increments and check that the denseness result holds. There are two main reasons for doing this. First, the class of point processes with independent increments represents one of the most studied class of completely random measures due to their nice theoretical properties and their importance in applications. Second, we have an explicit formulation for the quasi-Lévy measure and the drift of QID random variables supported on finite subsets of $\mathbb{N} \cup \{0\}$ (see Theorem 3.9 in Lindner et al. (2018)). Let us first show the denseness result for random variables supported on $\mathbb{N} \cup \{0\}$.

Proposition 3.8. *The class of QID distributions supported on finite subsets of $\mathbb{N} \cup \{0\}$ is dense in the class of probability distributions with support on $\mathbb{N} \cup \{0\}$ with respect to weak convergence.*

Proof. Let μ be a probability distribution with support on $\mathbb{N}$. For $n \in \mathbb{N}$, define the discrete distribution μ_n concentrated on the lattice $\{0, \ldots, 2n\}$ by

$$\mu_n(\{j\}) = \mu(\{j\}), \quad j \in \{0, \ldots, 2n\}.$$

Then, $\mu_n \xrightarrow{w} \mu$ as $n \to \infty$. It remains to prove that each μ_n is a weak limit of QID distributions with support on $\{0, \dots, 2n\}$. W.l.o.g. assume that the approximating sequence of distributions σ is such that $\sigma(\{j\}) > 0$ for every $j \in \{0, \dots, 2n\}$. Assume that the characteristic function $\hat{\sigma}$ has zeros (in the other case we can directly use Theorem 3.9 in Lindner et al. (2018) to conclude). Let X be a random variable with distribution σ and let $a_j = \mathbb{P}(X = j) > 0$ for $j = 0, \dots, 2n$. Then, the polynomial $f(w) = \sum_{j=0}^{2n} a_j w^j$ has zeroes on the unit circle. Factorizing, we obtain $f(w) = a_{2n} \prod_{j=1}^{2n}(w - \xi_j)$, where ξ_j, $j = 1, \dots, 2n$, denote the complex roots. Let $f_h(w) = a_{2n} \prod_{j=1}^{2n}(w - \xi_j - h)$, where $w \in \mathbb{C}$ and $h > 0$. Then, for small enough h, f_h is a polynomial with real coefficients, namely $f_h(w) = \sum_{j=0}^{2n} a_{h,j} w^j$ with $a_{h,j} \in \mathbb{R}$. Observe that for small enough h, $a_{h,j}$ and a_j will be close, so $a_{h,j} > 0$. Now, let X_h be a random variable with distribution $\sigma_h = \left(\sum_{j=0}^{2n} a_{h,j}\right)^{-1} \sum_{j=0}^{2n} a_{h,j}\delta_j$. We conclude by noticing that, for every $h > 0$, X_h is random variable with values on the lattice $\{0, \dots, 2n\}$ and its characteristic function has no zeros (thus it is QID by Theorem 3.9 in Lindner et al. (2018)), and that $X_h \xrightarrow{d} X$ as $h \searrow 0$. $\square$

From Corollary 3.21 in Kallenberg (2017), for an atomless point process with independent increments the corresponding unique pair, which we denote by (γ, F), is such that $\gamma = 0$ and F is restricted to $S \times \mathbb{N}$. Let $\mathcal{A}'$ be the set of all the point processes in $\mathcal{A}$. In other words, let $\mathcal{A}'$ be composed by CRMs of the form $\xi \overset{a.s.}{=} \alpha + \sum_{j=1}^{K} \beta_j \delta_{s_j}$, with α an atomless point process with independent increments and finite Lévy measure, $\{s_j : j = 1, \dots, K\}$ a finite set of fixed atoms in S, and β_j, $j \geq 1$, QID random variables concentrated on finite subsets of $\mathbb{N} \cup \{0\}$ and that are mutually independent and independent of α. Obviously, we have $\mathcal{A}' \subsetneq \mathcal{A}$.

Theorem 3.9. *$\mathcal{A}'$ is dense in the space of all point processes with independent increments with respect to the convergence in distribution.*

Proof. It follows from the same arguments as the ones used in the proof of Theorem 3.4. In particular, now we need to use Proposition 3.8 instead of Theorem 3.1. Further, now $\gamma_n = \gamma = 0$ and F_n and F are concentrated on $S \times \mathbb{N}$. Then, following the same arguments as the ones used in the proof of Theorem 3.4 we obtain the result. $\square$

We conclude this subsection with the denseness result for finite point processes (for which the weak topology might also be used), namely the equivalent of Theorem 3.7 for point processes with independent increments.

Proposition 3.10. *$\mathcal{A}'$ is dense in the space of point processes with independent increments, considered as random elements in $\hat{\mathcal{M}}_S$, endowed with either the vague topology or with the weak topology, with respect to the convergence in distribution.*

Proof. It follows from the same arguments as the ones used in the proof of Theorem 3.7, with Theorem 3.9 instead of Theorem 3.4. $\square$

4 Properties of the dense classes

4.1 Properties of the dense class $\mathcal{A}$

In this section we explore some of the properties of the random measures in $\mathcal{A}$, with a particular focus on their spectral representations. Consider the same notation as in the previous section. Let $\xi \in \mathcal{A}$ and so $\xi \overset{a.s.}{=} \alpha + \sum_{j=1}^{n} \delta_{s_j}\beta_j$ for some atomless CRM with finite Lévy measure α, some finite set of fixed atoms in S $\{s_j : j = 1, \dots, K\}$, and some $\mathbb{R}_+$-valued QID random variables β_j, $j \geq 1$, with finite quasi-Lévy measure and zero Gaussian component and that are mutually independent and independent of α. Using Theorem 12.10 and Corollary 12.11 in Kallenberg (2002) we have that

$$\hat{\mathcal{L}}(\alpha(A))(\theta) = \exp\left(i\theta\gamma^{(1)}(A) + \int_0^{\infty}(e^{i\theta x} - 1)F_A^{(1)}(dx)\right), \tag{7}$$

for every $\theta \in \mathbb{R}$ and $A \in \mathbf{S}$, where γ is a finite diffuse measure on $\mathbf{S}$ and $F^{(1)}$ is a finite measure on $\mathbf{S} \otimes \mathcal{B}((0,\infty))$ with diffuse projections onto S. Observe that we can extend $F^{(1)}$ to a finite measure on $\mathbf{S} \otimes \mathcal{B}(\mathbb{R})$, by assigning value zero outside $\mathbf{S} \otimes \mathcal{B}((0,\infty))$; by abuse of notation, we call this finite measure $F^{(1)}$.

Further, with centering function equal zero (as in (7)), denote by c_j and b_j the drift and the quasi-Lévy measure of β_j, for $j = 1, \dots, n$. Notice that we can use such centering function because the β_j's have finite quasi-Lévy measure. Then, the Lévy-Khintchine formulation of $\sum_{j=1}^{n} \delta_{s_j}\beta_j$ is given by

$$\hat{\mathcal{L}}\left(\sum_{j=1}^{n} \delta_{s_j}(A)\beta_j\right)(\theta) = \exp\left(i\theta\gamma^{(2)}(A) + \int_{\mathbb{R}}(e^{i\theta x} - 1)F_A^{(2)}(dx)\right),$$

for every $\theta \in \mathbb{R}$ and $A \in \mathbf{S}$, where $\gamma^{(2)}(A) = \sum_{j=1}^{n} \delta_{s_j}(A)c_j$ and $F_A^{(2)}(\cdot) = \sum_{j=1}^{n} \delta_{s_j}(A)b_j(\cdot)$. Then, ξ has the following characteristic function

$$\hat{\mathcal{L}}(\xi(A))(\theta) = \exp\left(i\theta\nu_0(A) + \int_{\mathbb{R}}(e^{i\theta x} - 1)F_A(dx)\right), \tag{8}$$

for every $\theta \in \mathbb{R}$ and $A \in \mathbf{S}$, where $\nu_0(A) = \gamma^{(1)}(A)+\gamma^{(2)}(A)$ and $F_A(\cdot) = F_A^{(1)}(\cdot)+F_A^{(2)}(\cdot)$.

Proposition 4.1. *Let $\xi \in \mathcal{A}$ and adopt the notation above. Then, F extends uniquely to a finite signed measure on $\mathbf{S} \otimes \mathcal{B}(\mathbb{R})$.*

Proof. Consider the notations above. For the first statement we need to show that F is a finite signed measure on $\mathbf{S} \otimes \mathcal{B}(\mathbb{R})$. Since $F^{(1)}$ is a finite measure on $\mathbf{S} \otimes \mathcal{B}((0,\infty))$, it remains to show that $F^{(2)}$ is a finite signed measure on $\mathbf{S} \otimes \mathcal{B}(\mathbb{R})$. We know that $F_A^{(2)}(\cdot) = \sum_{j=1}^{n} \delta_{s_j}(A)b_j(\cdot)$ where $b_j(\cdot)$ are finite signed measures on $\mathcal{B}(\mathbb{R})$. It is possible to see that $F^{(2)}$ is a bimeasure on $\mathbf{S} \times \mathcal{B}(\mathbb{R})$ and that

$$\sup_I \sum_{i \in I} |F_{A_i}^{(2)}(B_i)| = \sum_{j=1}^{n} |b_j|(\mathbb{R}) < \infty,$$

where the supremum is taken over all the finite families of disjoints elements of $\mathbf{S}\times\mathcal{B}(\mathbb{R})$. Then, by Theorem 5.18 in Passeggeri (2020c) (see also Theorem 4 in Horowitz (1977)) $F^{(2)}$ extends to a finite signed measure on $\mathbf{S}\otimes\mathcal{B}(\mathbb{R})$. Thus, F is a finite signed measure on $\mathbf{S}\otimes\mathcal{B}(\mathbb{R})$. ☐

Following the notation of the ID case (see Kallenberg (2017) page 89), we call F the *quasi-Lévy measure* of ξ. Observe that in the ID case the Lévy measure, which we denote by F_{ID}, is a σ-finite measure on $\mathbf{S}\otimes\mathcal{B}((0,\infty))$ such that $A\mapsto F_{ID}(A\times(0,\infty))$ is not necessarily σ-finite (see Kingman (1993) pages 82–83), while our quasi-Lévy measure is a finite signed measure. In the following result we show the existence of a unique correspondence between any element in $\mathcal{A}$ and a characteristic pair.

Theorem 4.2. *Let $\xi\in\mathcal{A}$. Then, there exists a pair (ν_0,F) s.t. (8) holds, where ν_0 and F are a finite signed measure on $\mathbf{S}$ and $\mathbf{S}\otimes\mathcal{B}(\mathbb{R})$, respectively, s.t. for every $A\in\mathbf{S}$ and $B\in\mathcal{B}(\mathbb{R})$:*
(i) $\nu_0(A)=\gamma(A)+\sum_{j=1}^{n}\delta_{s_j}(A)c_j$*, for some diffuse finite measure γ on $\mathbf{S}$, $c_1,\dots,c_n\in\mathbb{R}$, and finitely many atoms $s_1,\dots,s_n\in S$,*
(ii) $F(A\times B)=\tilde{G}(A\times B)+\sum_{j=1}^{n}\delta_{s_j}(A)b_j(B)$*, for some finite measure $\tilde{G}$ on $\mathbf{S}\otimes\mathcal{B}(\mathbb{R})$, which is the extension by zero of some measure G on $\mathbf{S}\otimes\mathcal{B}((0,\infty))$ with diffuse projections onto S, and for some finite signed measures b_j's on $\mathcal{B}(\mathbb{R})$, such that $\exp(b_1),\dots,\exp(b_n)$ are measures.*

Conversely, for every such pair (ν_0,F) there exists a unique random measure $\xi\in\mathcal{A}$ s.t. (8) holds.

Proof. Concerning the atomless component of ξ, from Corollary 12.11 in Kallenberg (2002) and Theorem 3.20 in Kallenberg (2017) we know that there exists a one to one correspondence between an ID atomless random measure with independent increments and a characteristic pair, composed by a diffuse measure on $\mathbf{S}$ and a measure on $\mathbf{S}\otimes\mathcal{B}(\mathbb{R})$ with diffuse projections onto S. In our case we note that the components of the characteristic pair are finite measures by definition. For the fixed component of ξ, by Theorem 4.3.4 in Cuppens (1975) we know that a characteristic triplet where the Gaussian component is zero and the quasi-Lévy measure is finite is the characteristic triplet of a QID random variable if and only if the exponential of the finite quasi-Lévy measure is a measure. Then, by the definition of ξ and by the discussion and the computations at the beginning of this section on the characteristic functions of the components of ξ, we immediately obtain the result. Notice that for the converse direction we need also to show the independence of the fixed and atomless components, but this follows immediately from the linear structure of ν_0 and F. ☐

Remark 4.3. Notation: instead of using the characteristic pair we could have equivalently used the characteristic set $(\{s_j\}_{j=1}^{n},\gamma,\{c_j\}_{j=1}^{n},G,\{b_j\}_{j=1}^{n})$, with the above structure, in order to have a more explicit one to one identification with $\xi\in\mathcal{A}$.

In the following result we present the Laplace transform of $\xi\in\mathcal{A}$ when the supports of the quasi-Lévy measure of the random variables of its fixed component lie in $\mathbb{R}_+$, namely when $\mathrm{supp}(\cup_{j=1}^{n}b_j)\subset\mathbb{R}_+$.

Corollary 4.4. *Let $\xi \in \mathcal{A}$. If* $\mathrm{supp}(\cup_{j=1}^{n} b_j) \subset \mathbb{R}_+$*, then* $\mathrm{supp}(F_A) \subset \mathbb{R}_+$ *for every* $A \in \mathbf{S}$*,* $c_j = \inf(\mathrm{supp}(\mathcal{L}(\beta_j)))$ *for every* $j = 1, \ldots, n$*,* $\mathrm{supp}(\nu_0) \subset \mathbb{R}_+$*. Further, for every* $u \geq 0$ *and* $A \in \mathbf{S}$

$$\mathbb{E}\left[e^{-u\xi(A)}\right] = \exp\left(-u\nu_0(A) + \int_0^{\infty} (e^{-ux} - 1)F_A(dx)\right). \tag{9}$$

Proof. It follows from the structure of $\xi \in \mathcal{A}$, Theorem 4.2, and Proposition 5.1 in Lindner et al. (2018). ☐

We have the restriction $\mathrm{supp}(\cup_{j=1}^{n} b_j) \subset \mathbb{R}_+$, because the Laplace transform representation of general non-negative QID random variables is still an open question.

Example 4.5. Consider a distribution μ with $\mu(\{x\}) > 1/2$, for some $x \in \mathbb{R}$, then μ is QID with no Gaussian component and (explicit) finite quasi-Lévy measure (see Theorem 3.1 in Lindner et al. (2018)). Moreover, let μ be the distribution on $[x, \infty)$, for some $x \in \mathbb{R}$, and let $\mu(\{x\}) > 1/2$. Let $b(\cdot)$ be its finite quasi-Lévy measure. Then, it is possible to derive from Theorem 3.1 in Lindner et al. (2018), using Lemma 24.1 in Sato (1999), that $\mathrm{supp}(b) \subset \mathbb{R}_+$. For example, given a continuous random variable X then $\max(X, \mathrm{median}(X) + \varepsilon)$ is an example of such distribution, for any $\varepsilon > 0$. Hence, given a Normal random variable $X \sim N(\mu, \sigma)$, where $\mu \in \mathbb{R}$ and $\sigma > 0$, we have that $\max(N, c)$ is QID without Gaussian component and without being ID, for any $c > \mu$. Similar results holds for any distribution. Therefore, a CRM ξ with a fixed atomic component composed by such distributions, with a finite deterministic component, and with an ordinary component with finite Lévy measure, belongs to $\mathcal{A}$ and has $\mathrm{supp}(\cup_{j=1}^{n} b_j) \subset \mathbb{R}_+$.

Example 4.6. Thanks to Lemma 24.1 in Sato (1999) and the fact that a finite sum of independent QID random variables is QID and its quasi-Lévy measure is the sum of the quasi-Lévy measures of the summands, a finite sum of independent CRMs for which Corollary 4.4 applies is itself a CRM for which Corollary 4.4 applies, namely it belongs to $\mathcal{A}$ and has $\mathrm{supp}(\cup_{j=1}^{n} b_j) \subset \mathbb{R}_+$.

Remark 4.7. The results presented in this section also hold for $\xi \in \mathcal{A}_\infty$ (see Remark 3.6). Concerning Proposition 4.1, $F^{(1)}$ is now a σ-finite measure on $\mathbf{S} \otimes \mathcal{B}(\mathbb{R})$ (see Corollary 3.21 in Kallenberg (2017)), and $F^{(2)}$ is a finite signed measure on $\mathbf{S} \otimes \mathcal{B}(\mathbb{R})$ as shown in the proof of Proposition 4.1. Thus, in this case $F = F^{(1)} + F^{(2)}$ is a signed measure on $\mathbf{S} \otimes \mathcal{B}(\mathbb{R})$ with values in $(-\infty, \infty]$. Concerning Theorem 4.2, the same statement holds except that A is taken in $\hat{\mathbf{S}}$ (to avoid infinite values) and G (and so $\tilde{G}$) is now a σ-finite measure. Concerning Corollary 4.4, the same statement holds except that A is taken in $\hat{\mathbf{S}}$.

4.2 Properties of the dense class $\mathcal{A}'$

Since $\mathcal{A}' \subsetneq \mathcal{A}$ all the results presented in the previous section hold for $\xi \in \mathcal{A}'$. In this subsection, we show that even better results hold for the elements in $\mathcal{A}'$, which is the structure of these random measures. To facilitate the presentation of the results, let us recall Theorem 3.9 in Lindner et al. (2018).

Theorem 4.8 (Theorem 3.9 in Lindner et al. (2018))**.** *Let μ be a discrete distribution concentrated on $\{0, 1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$, i.e., $\mu = \sum_{j=0}^{n} a_j \delta_j$, where $a_0, \ldots, a_{n-1} \geq 0$, $a_n > 0$, and $a_0 + \cdots + a_n = 1$. Then the following are equivalent:*

- (i) *μ is quasi-infinitely divisible.*
- (ii) *The characteristic function of μ has no zeroes.*
- (iii) *The polynomial $w \mapsto \sum_{j=0}^{n} a_j w^j$ in the complex variable w has no roots on the unit circle, i.e. $\sum_{j=0}^{n} a_j w^j \neq 0$, for all $w \in \mathbb{C}$ with $|w| = 1$.*

Further, if one of the equivalent conditions (i)–(iii) holds, then the quasi-Lévy measure of μ is finite and concentrated on $\mathbb{Z}$, the drift lies in $\{0, 1 \ldots, n\}$, and the Gaussian component of μ is 0. More precisely, if $\xi_1, \ldots, \xi_n$ denote the n complex roots of $w \mapsto \sum_{j=0}^{n} a_j w^j$, counted with multiplicity, then the quasi-Lévy measure of μ is given by

$$\nu = -\sum_{m=1}^{\infty} m^{-1} \Bigg(\sum_{j:|\xi_j|<1} \xi_j^m \Bigg) \delta_{-m} - \sum_{m=1}^{\infty} m^{-1} \Bigg(\sum_{j:|\xi_j|>1} \xi_j^{-m} \Bigg) \delta_m, \tag{10}$$

and the drift is equal to the number of those zeroes of this polynomial which lie inside the unit circle (counted with multiplicity), i.e., have modulus less than 1.

In the following theorem we adopt the following notation. Let $\xi \in \mathcal{A}'$. We denote by α its atomless component and by β_j, $j = 1, \ldots, n$ the QID random variables of its fixed component, *i.e.* $\xi \overset{a.s.}{=} \alpha + \sum_{j=1}^{n} \beta_j \delta_{s_j}$. Further, for every $j = 1, \ldots, n$, we denote the law of β_j by $\sum_{l=0}^{k_j} a_{j,l} \delta_l$, namely $\mathcal{L}(\beta_j) = \sum_{l=0}^{k_j} a_{j,l} \delta_l$ and denote by $\zeta_{j,1}, \ldots, \zeta_{j,k_j}$ the k_j complex roots of $w \mapsto \sum_{l=0}^{k_j} a_{j,l} w^l$. Finally, we denote by b_j the quasi-Lévy measure of β_j, *i.e.*

$$b_j = -\sum_{m=1}^{\infty} m^{-1} \Bigg(\sum_{l:|\zeta_{j,l}|<1} \zeta_{j,l}^m \Bigg) \delta_{-m} - \sum_{m=1}^{\infty} m^{-1} \Bigg(\sum_{l:|\zeta_{j,l}|>1} \zeta_{j,l}^{-m} \Bigg) \delta_m, \tag{11}$$

and by c_j its drift, *i.e.* $c_j = \#\{|\zeta_{j,l}| < 1, l = 1, \ldots, k_j\}$.

Theorem 4.9. *Let $\xi \in \mathcal{A}'$. Then, there exists a pair (ν_0, F) s.t. (8) holds, where ν_0 and F are a finite signed measure on $\mathbf{S}$ and $\mathbf{S} \otimes \mathcal{B}(\mathbb{R})$, respectively, s.t. for every $A \in \mathbf{S}$ and $B \in \mathcal{B}(\mathbb{R})$:*

- (i) *$\nu_0(A) = \sum_{j=1}^{n} \delta_{s_j}(A) c_j$, where $n \in \mathbb{N}$, $s_j \in S$ is an atom, and $c_j = \#\{|\zeta_{j,l}| < 1, l = 1, \ldots, k_j\}$, for $j = 1, \ldots, n$,*
- (ii) *$F(A \times B) = \tilde{G}(A \times B) + \sum_{j=1}^{n} \delta_{s_j}(A) b_j(B)$, where $\tilde{G}$ is a finite measure on $\mathbf{S} \otimes \mathcal{B}(\mathbb{R})$ restricted on $S \times \mathbb{N}$ and with diffuse projections onto S, and where b_j satisfies (11), for $j = 1, \ldots, n$.*

Conversely, for every such pair (ν_0, F), where $\zeta_{j,1}, \ldots, \zeta_{j,k_j}$ denote the k_j complex roots of some polynomial $w \mapsto \sum_{l=0}^{k_j} a_{j,l} w^l$ for $j = 1, \ldots, n$, there exists a unique random measure $\xi \in \mathcal{A}$ s.t. (8) holds.

Proof. It follows from the same arguments as the one used in Theorem 4.2 and from Theorem 4.8. In particular, the first direction is trivial. For the other direction, we have the following. As mentioned in the proof of Theorem 4.2, we have a one-to-one correspondence for the atomless part of ξ and its characteristic pair. Concerning the fixed component, let us assume that there exist c_j and b_j which are functions of some complex roots of some complex polynomial $w \mapsto \sum_{l=0}^{k_j} a_{j,l} w^l$ with no roots in the unite circle, where $k_j \in \mathbb{N}$, $a_0, \dots, a_{k_j-1} \geq 0$, $a_{k_j} > 0$, and $a_0 + \dots + a_{k_j} = 1$. Then, by Theorem 4.8 there exists a QID probability distribution $\mathcal{L}(\beta_j) = \sum_{l=0}^{k_j} a_{j,l}\delta_l$. Since this holds for every $j = 1, \dots, n$ then from the set of atoms $s_1, \dots, s_n \in S$ we obtain the fixed component $\sum_{j=1}^{n} \delta_{s_j}\beta_j$ of a random measure in $\mathcal{A}'$. $\square$

The content of Remarks 4.3 and 4.7 holds *mutatis mutandis* here. In addition, we refer to Nehring et al. (2013) for further properties of certain subclasses of point processes with quasi-Lévy measures.

Example 4.10. We explore now an example of $\xi \in \mathcal{A}'$ for which Corollary 4.4 applies. Using Theorem 4.8 it is possible to see that such example is provided by a point process with independent increments ξ with $\mathcal{L}(\beta_j) = \sum_{l=0}^{k_j} a_l \delta_l$, where $a_0, \dots, a_{k_j-1} \geq 0$, $a_{k_j} > 0$, and such that $\zeta_{j,1}, \dots, \zeta_{j,k_j}$ satisfy $|\zeta_{j,l}| > 1$ for every $l = 1, \dots, k_j$ and for every $j = 1, \dots, n$. For example, thanks to the Eneström-Kakeya Theorem, the latter condition is satisfied in the case $a_0 > a_1 > \dots > a_{k_j}$ for every $j = 1, \dots, n$ or more generally in the case $a_0 = a_{h_j} = 0$ and $a_{h_j+1} > \dots > a_{k_j}$, for some $h_j = 0, \dots, k_j - 1$. This theorem states that the (complex) solutions of the equation $\sum_{j=0}^{n} a_j w^j = 0$ are in modulus bounded below by $\min\limits_{j=0,\dots,n}(a_j/a_{j+1})$. For the case $a_0 = a_{h_j} = 0$ and $a_{h_j+1} > \dots > a_{k_j}$ we used the fact that in this case $\sum_{j=0}^{n} a_j w^j = w^{h_j+1} \sum_{j=h_j+1}^{n} a_j w^{j-h_j-1}$ and so the roots are given by 0 and by the roots of $\sum_{j=h_j+1}^{n} a_j w^{j-h_j-1}$, which since $a_{h_j+1} > \dots > a_{k_j}$ we know to be in modulus strictly bounded below by 1. The root 0 does not affect the form of the quasi-Lévy measure as it is possible to see from (11). Thus, for example a CRM with $\mathcal{L}(\beta_1) = \frac{1}{3}\delta_{25} + \frac{1}{4}\delta_{26} + \frac{1}{5}\delta_{27} + \frac{1}{6}\delta_{28} + \frac{1}{20}\delta_{29}$, with a finite deterministic component, and with an ordinary component with finite Lévy measure, belongs to $\mathcal{A}'$ and has $\mathrm{supp}(b_1) \subset \mathbb{N}$.

Example 4.11. Following Example 4.6, it is possible to obtain that a finite sum of independent multinomial distributions of the form explored in Example 4.10 is QID with no Gaussian component and such that its quasi-Lévy measures have support in $\mathbb{N}$.

We remark that using Theorems 8.1 and 8.8 in Lindner et al. (2018) it is possible to build examples with distributions on $\mathbb{N} \cup \{0\}$ with unbounded support.

5 A nonparametric Bayesian setting

In this section we show how the results presented in Sections 3 and 4 apply to a particular class of nonparametric prior distributions. The framework is the one of the paper by Broderick, Wilson and Jordan Broderick et al. (2018), which is further explored in subsequent papers, see Campbell et al. (2019) among others. The authors analyse

Bayesian nonparametric priors and likelihood functions based on CRMs. In particular, they let the prior be modelled as:

$$\Theta := \sum_{k=1}^{K} \theta_k \delta_{\psi_k},$$

where K may be either finite or infinite and where (θ_k, ψ_k) is a pair consisting of the frequency (or rate) of the k-th trait together with its trait ψ_k, which belongs to some space Ψ of traits. Notice that $\psi_1, \ldots, \psi_K$ include both the fixed and non-fixed atoms. This representation follows Kingman's representation of a CRM without deterministic component, see Kingman (1993). The data point for the m-th individual is modelled as:

$$X_m := \sum_{k=1}^{K_m} x_{m,k} \delta_{\psi_k},$$

where $x_{m,k}$ represents the degree to which the m-th data point belongs to the trait ψ_k.

This setting can be applied to many real world applications such as topic modelling, as seen in the introduction. In topic modelling (see Broderick et al. (2018); Campbell et al. (2019); Teh et al. (2006)), we have that ψ_k represents a topic; that is, ψ_k is a distribution over words in a vocabulary. Further, θ_k might represent the frequency with which the topic ψ_k occurs in a corpus of documents. Alternatively to what mentioned in the introduction, $x_{j,k}$ might represent the number of words in topic ψ_k that occur in the jth document. So the jth document has a total length of $\sum_{k=1}^{K} x_{j,k}$ words. In this case, the actual observation consists of the words in each m documents, and the topics of the whole corpus of documents are latent.

From a mathematical (and formal) point of view Θ and X_m are defined as CRMs. In particular, for the data X_m, we let $x_{m,k}$ be drawn according to some distribution H that takes θ_k as a parameter and have support on $\mathbb{N} \cup \{0\}$, that is $x_{m,k}|\theta_k \stackrel{iid}{\sim} h(\cdot|\theta_k)$, independently across m and k. We assume that $X_1, \ldots, X_m$ are i.i.d. conditional on Θ. Moreover, Broderick et al. (2018) consider the following assumptions for Θ and X_m:

Assumption A00. the atomless component of Θ has characteristic pair (γ, μ) s.t. $\gamma = 0$ and $\mu(d\theta \times d\psi) = \nu(d\theta) \cdot G(d\psi)$, where ν is any σ-finite measure on $\mathbb{R}_+$ and G is a proper distribution on Ψ with no atoms.

Assumptions A0, A1, and A2. Θ has a finite number of fixed atoms, $\nu(\mathbb{R}_+) = \infty$, and $\sum_{x=1}^{\infty} \int_{\mathbb{R}_+} h(x|\theta)\nu(d\theta) < \infty$, respectively.

We remark that by Assumption A00 we have that the location of the non-fixed atoms ψ and the frequencies θ_k are stochastically independent. We call ν the *weight rate measure* of Θ. Moreover, the assumptions A0, A1 and A2 comes from a modelling need. By assuming A0 we are saying that we initially know certain traits, by A1 that there are countably infinite possible traits, and by A2 that the amount of information from finitely represented data is finite (because by A2 the number of non-fixed atoms is finite).

The first main result in Broderick et al. (2018) is Theorem 3.1, which shows explicit formulations for the posterior distribution $\Theta|X_1$, and it is extended in Corollary 3.2 to the posterior $\Theta|X_{1:m}$ (see also James (2017)). In the following result we are going to show that similar results hold for any random measure in $\mathcal{A}$ without assuming A0, A1 or A2. Notice that we can write $\Theta = \sum_{k=1}^{K} \theta_k \delta_{\psi_k}$, where $K = K_{fix} + K_{ord}$, namely K is the sum of the fixed and non-fixed atoms, thus K is random. Following the notation of Broderick et al. (2018), we denote the fixed component of Θ by $\Theta_{fix} = \sum_{k=1}^{K_{fix}} \theta_{fix,k} \delta_{\psi_{fix,k}}$ and the law of $\theta_{fix,k}$ by $F_{fix,k} := \mathcal{L}(\Theta(\{\psi_{fix,k}\}))$.

Proposition 5.1. *Let $\Theta \in \mathcal{A}$ satisfy A00. Write $\Theta = \sum_{k=1}^{K} \theta_k \delta_{\psi_k}$, and let $X_1, \ldots, X_m$ be generated conditional on Θ according to $X_1 := \sum_{k=1}^{K} x_{1,k} \delta_{\psi_k}$ with $x_{1,k}|\theta_k \overset{iid}{\sim} h(\cdot|\theta_k)$ for proper, discrete probability mass function h. It is enough to make the assumption for X_1 since the $X_1, \ldots, X_m$ are i.i.d. conditional on Θ.*

Then let Θ_{post} be a random measure with the distribution of $\Theta|X_{1:m}$, where $X_{1:m} = (X_1, \ldots, X_m)$. Θ_{post} is a CRM with three parts.

1. *For each $k \in [K_{fix}]$, Θ_{post} has a fixed atom at $\psi_{fix,k}$ with weight $\theta_{post,fix,k}$ distributed according to the finite-dimensional posterior $F_{post,fix,k}(d\theta)$ that comes from prior $F_{fix,k}$, likelihood h, and observation $X(\psi_{fix,k})$. Moreover, $F_{fix,k}$ is QID with no Gaussian component and finite quasi-Lévy measure, and $F_{post,fix,k}(d\theta) \propto F_{fix,k}(d\theta) \prod_{j=1}^{m} h(x_{fix,j,k}|\theta)$.*

2. *Let $\{\psi_{new,k} : k \in [K_{new}]\}$ be the union of atom locations across $X_1, X_2, \ldots, X_m$ minus the fixed locations in the prior of Θ. K_{new} is finite. Let $x_{new,j,k}$ be the weight of the atom in X_j located at $\psi_{new,k}$, for some $j = 1, \ldots, m$. Then Θ_{post} has a fixed atom at $x_{new,k}$ with random weight $\theta_{post,new,k}$, whose distribution $F_{post,new,k}(d\theta) \propto \nu(d\theta) \prod_{j=1}^{m} h(x_{new,j,k}|\theta)$.*

3. *The ordinary component of Θ_{post} has finite weight rate measure $\nu_{post,m}(d\theta) := \nu(d\theta) h(0|\theta)^m$.*

Remark 5.2. Observe that since $\Theta \in \mathcal{A}$ then it has finite fixed atoms so assumption A0 is satisfied. Moreover, since ν is also finite and $h(x|\theta) \leq 1$, then assumption A2 is also satisfied. The only difference with Theorem 3.1 and Corollary 3.2 in Broderick et al. (2018) is that we do not necessarily satisfy assumption A1. However, A1 is a modelling assumption rather than a technical one. Indeed, the proof of this result follows from similar arguments as the one used in the proof of Theorem 3.1 and Corollary 3.2 in Broderick et al. (2018).

Proof. To lighten the notation let $X := X_1$. Let us first prove the result for $\Theta|X$. Any fixed atom $\theta_{fix,k} \delta_{\psi_{fix,k}}$ in the prior is independent of the other fixed atoms and of the ordinary component. Thus, all of X except $x_{fix,k} := X(\{\psi_{fix,k}\})$ is independent of $\theta_{fix,k}$. Thus, $\Theta|X$ has a fixed atom at $\psi_{fix,k}$ and $\mathcal{L}(\theta_{post,fix,k}) \propto F_{fix,k}(d\theta) h(x_{fix,k}|\theta)$. Recall that since G is continuous, all the fixed and non-fixed atoms of Θ are at a.s. distinct locations. Observe that by letting $\Psi_{fix} := \{\psi_{fix,1}, \ldots, \psi_{fix,K_{fix}}\}$ we can define the fixed and ordinary component of X by $X_{fix}(A) := X(A \cap \Psi_{fix})$ and $X_{ord}(A) := X(A \cap (\Psi \setminus \Psi_{fix}))$, respectively.

Let $x \in \mathbb{N} \cup \{0\}$ and let $\{\psi_{new,x,1}, \ldots, \psi_{new,x,K_{new,x}}\}$ be all the locations of atoms in X_{ord} of size x, which is finite and it is a subset of the locations of atoms of Θ_{ord}. Further, let $\theta_{new,x,k} := \Theta(\{\psi_{new,x,k}\})$. Observe that the values $\{\theta_{new,x,k}\}_{k=1}^{K_{new,x}}$ are generated from a thinned Poisson point process with rate measure (also known as intensity measure) $\nu_x(d\theta) = \nu(d\theta)h(x|\theta)$, this is due to the $h(x|\theta)$-thinning of the Poisson point process $\{\theta_{ord,k}\}_{k=1}^{K_{ord}}$ which has rate measure ν. Moreover, given that $\nu_x(\mathbb{R}_+) < \infty$, we have that $\mathcal{L}(\theta_{new,x,k}) \propto \nu(d\theta)h(x|\theta)$. Finally, observe that there is a possibility that atoms in Θ_{ord} are not observed in X_{ord}, this happens when the likelihood draw returns a zero. These atom weights form a Poisson point process with rate measure $\nu(d\theta)h(0|\theta)$.

Considering $\Theta|X_1$ as the new prior we obtain the formulation for the posterior $\Theta|X_1, X_2$ by induction and by observing that the assumptions are still satisfied by $\Theta|X_1$. Then, by induction we conclude the proof. $\square$

In the next result, we show that random measures in $\mathcal{A}$ satisfying A00 are dense in the space of all CRMs satisfying A0, A1 and A2, namely all the random measures considered in Broderick et al. (2018) (and in Campbell et al. (2019)). Further, we show how this result translates into a convergence for the ordinary component of the respective posteriors.

Proposition 5.3. *Consider any random measure Θ satisfying A00, A0, A1 and A2. Then, there exists a sequence of random measures $(\Theta_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ and satisfying A00 such that $\Theta_n \xrightarrow{d} \Theta$, as $n \to \infty$. Further, $\Theta_{n,post,ord} \xrightarrow{d} \Theta_{post,ord}$, as $n \to \infty$.*

Proof. The first part of this proof consists in realising that the arguments in the proof of Proposition 3.3 and Theorem 3.4 can be adapted to the present case. Denote by $\mu(d\theta\times d\psi) = \nu(d\theta)\cdot G(d\psi)$ the Lévy measure of Θ. Following the proofs of Proposition 3.3 and Theorem 3.4 it is possible to see that the approximating sequence Θ_n should have Lévy measure $\nu_n(d\theta)\cdot G_n(d\psi)$ where $\nu_n(d\theta) := \nu((\frac{1}{n}, \infty)\cap d\theta)$ and $G_n(d\theta) := G(S_n \cap d\psi)$. However, given the assumptions on μ, namely that G is a finite measure, we can (and we do) take the Lévy measure of Θ_n to be given by $\mu_n(d\theta \times d\psi) := \nu_n(d\theta) \cdot G(d\psi)$. Then, applying the same arguments as the one used in the proof of Proposition 3.3 and Theorem 3.4, we obtain that the ordinary component of Θ_n converges in distribution to the one of Θ. The convergence of the fixed component follows directly from Theorem 3.4. Since μ_n is finite, we have that Θ_n is in $\mathcal{A}$ and that it satisfies A00.

For the convergence of the posteriors, consider Θ_n with its respective data points $X_{n,1}, \ldots, X_{n,m}$, which are defined conditional on Θ_n as in Proposition 5.1 and belong to some probability spaces possibly different from the one of the other data points. From Proposition 5.1 we know that $\Theta_{n,post}$ has finite weight rate measure $\nu_{n,post,m}(d\theta) := \nu_n(d\theta)h(0|\theta)^m$, while from Corollary 3.2 in Broderick et al. (2018) we know that Θ_{post} has finite weight rate measure $\nu_{post,m}(d\theta) := \nu(d\theta)h(0|\theta)^m$. Since $\nu_{n,post,m}(\cdot) = \nu_{post,m}((\frac{1}{n}, \infty) \cap \cdot)$ we obtain the result by Proposition 3.3. $\square$

We summarise our findings so far in words. First, we obtain an explicit expression for the posterior of any random measure in $\mathcal{A}$ satisfying A00. Second, such random measures are dense with respect to convergence in distribution in the space of all priors considered in Broderick et al. (2018). Third, when approximating in distribution such a prior, call it Θ, the ordinary component of the posteriors of these approximating random measures converge to the one of Θ.

Example 5.4. Consider the beta process without fixed atomic component as a prior Θ. The beta process has an ordinary component whose weight rate measure has a beta distribution kernel,

$$\nu(d\theta) = \gamma\theta^{-\alpha-1}(1-\theta)^{c+\alpha-1}d\theta,$$

with support on $(0,1]$ and γ, c, and α are three fixed hyperparameters. Consider Θ_n to be a CRM without deterministic and fixed components and let its ordinary component having weight rate measure

$$\nu_n(d\theta) = \gamma\theta^{-\alpha-1}(1-\theta)^{c+\alpha-1}\mathbf{1}_{(1/n,1]}(\theta)d\theta.$$

Then, $\Theta_n \in \mathcal{A}$ and $\Theta_n \xrightarrow{d} \Theta$, as $n \to \infty$. Usually the beta process is paired with the Bernoulli process likelihood, namely given Θ we draw $X_m = \sum_{k=1}^{K} x_{m,k}\delta_{\psi_k}$ with $x_{m,k} \overset{ind}{\sim} \text{Bern}(x|\theta_k)$, where $\text{Bern}(x|\theta_k)$ stands for a Bernoulli random variable with parameter θ_k. Notice that $\mathbb{P}(x_{m,k}=0|\theta_k=0)=1$ and that the marginal distribution of the $X_{1:N}$, drawn given Θ, is called an Indian buffet process Griffiths and Ghahramani (2005); James (2017). Further, consider the Bernoulli process likelihood also for Θ_n, namely given Θ_n we draw $X_{n,m} = \sum_{k=1}^{K} x_{n,m,k}\delta_{\psi_k}$ with $x_{n,m,k} \overset{ind}{\sim} \text{Bern}(x|\theta_{n,k})$, then the marginal distribution of the $X_{n,1:N}$, drawn given Θ_n, is a truncation of the Indian buffet process.

Example 5.5. Consider the priors Θ and Θ_n of the previous example and instead of the Bernoulli process likelihood we consider the negative binomial process likelihood. In this case, $x_k \overset{ind}{\sim} \text{NegBin}(x|r,\theta_k)$ for some fixed hyperparameter $r>0$, and notice that $\mathbb{P}(x_k=0|r,\theta_k=0)=1$. Then, $\Theta_{n,post,fix}$, namely the fixed atomic component of the posterior of Θ_n, has a fixed atom at $\psi_{n,new,k}$ whose weight $\theta_{n,post,new,k}$ has distribution

$$\begin{aligned}
F_{n,post,new,k}(d\theta) &\propto \nu_n(d\theta)\cdot h(x_{n,new,k}|\theta)\\
&= \gamma\theta^{-\alpha-1}(1-\theta)^{c+\alpha-1}\mathbf{1}_{(1/n,1]}(\theta)d\theta\cdot\theta^{x_{n,new,k}}(1-\theta)^{r}\\
&\propto Beta(\theta|-\alpha+x_{n,new,k}, c+\alpha+r)\mathbf{1}_{(1/n,1]}(\theta)d\theta,
\end{aligned}$$

while the ordinary component $\Theta_{n,post,ord}$ has rate measure

$$\begin{aligned}
\nu_n(d\theta)\cdot h(0|\theta) &= \gamma\theta^{-\alpha-1}(1-\theta)^{c+\alpha-1}\mathbf{1}_{(1/n,1]}(\theta)d\theta\cdot(1-\theta)^{r}\\
&= \gamma\theta^{-\alpha-1}(1-\theta)^{c+r+\alpha-1}\mathbf{1}_{(1/n,1]}(\theta)d\theta.
\end{aligned}$$

Thus, the posterior has fixed atoms that are beta distributed and the ordinary part has a beta distribution kernel (truncated at $\theta \in (1/n,1]$). Thus, we have shown that the beta process truncated at $\theta \in (1/n,1]$ is, in fact, conjugate to the negative binomial process, and we generalised the result in Broderick et al. (2015) (see also Example 3.3 in Broderick et al. (2018)). Finally, observe that $\Theta_{n,post,ord} \xrightarrow{d} \Theta_{post,ord}$, as $n \to \infty$, for any likelihood function.

In the same spirit as in the previous two examples it is possible to exploit the automatic conjugacy of the exponential CRMs (see Broderick et al. (2018)) to obtain automatic conjugacy for truncated exponential CRMs. In the following example we consider CRMs with only fixed atomic component.

Example 5.6. Consider $\beta_1, \beta_2, \dots$ to be Normal random variables with $\beta_j \sim N(\mu_j, \sigma)$, where $\mu_j \in [0, \infty)$ and $\sigma \in (0, \infty)$, for every $j \in \mathbb{N}$. Let $\Theta = \sum_{j=1}^{\infty} \max(\beta_j, \mu_j)\delta_{s_j}$, for some points $s_1, s_2, \dots$ in S. Let $\Theta_n = \sum_{j=1}^{n} \max\big(\beta_j, \mu_j + \frac{1}{n^2}\big)\delta_{s_j}$. Then, $\Theta_n \in \mathcal{A}$, and since $\rho\big(\max\big(\beta_j, \mu_j\big), \max\big(\beta_j, \mu_j + \frac{1}{n^2}\big)\big) \leq \frac{1}{n^2}$ we have that $\Theta_n \xrightarrow{d} \Theta$, as $n \to \infty$. In particular, for every n and $A \in \mathbf{S}$ we have that

$$\mathbb{E}\left[e^{-u\Theta_n(A)}\right] = \prod_{j=1}^{n} \exp\left(\delta_{s_j}(A)\left(-u\big(\mu_j + \frac{1}{n^2}\big) + \int_0^{\infty}(e^{-ux} - 1)b_n(dx)\right)\right),$$

where b_n is a finite signed measure with support in $[0, \infty)$, in particular let $\Phi(dx)$ be the Gaussian measure with mean $-\frac{1}{n^2}$ and variance σ^2, let $p_n = \Phi((-\infty, 0))$, and let $\Phi_{(0,\infty)}$ the truncation of Φ on $(0, \infty)$ then

$$b_n = \sum_{m=1}^{\infty} \frac{1}{m}(-1)^{m+1} p_n^{-m} \Phi_{(0,\infty)}^{*m}$$

and $b_n\{0\} = 0$. The notation $\Phi_{(0,\infty)}^{*m}$ stands for $\Phi_{(0,\infty)} * \cdots * \Phi_{(0,\infty)}$ (m times). We stress that b_n does not depend j. Now, consider a Poisson likelihood process, namely given Θ_n we draw $X = \sum_{k=1}^{K} x_k \delta_{\psi_k}$ with $x_k \overset{ind}{\sim} \text{Poisson}(x|\theta_k)$ that is $h(x|\theta_k) = \frac{1}{x!}\theta_k^x e^{-\theta_k}$. Observe that $h(x|\theta)$ is defined only for $\theta > 0$ and so $\text{supp}(X) \subset \text{supp}(\Theta)$ a.s.. Then, the posterior has only fixed atomic component where for each atom $\psi_{fix,k}$, letting $x_{n,post,fix,k} := X(\{\psi_{fix,k}\})$, the weight $\theta_{n,post,fix,k}$ has distribution

$$\begin{aligned} F_{n,post,fix,k}(d\theta) &\propto F_{n,fix,k}(d\theta) \cdot h(x_{n,post,fix,k}|\theta) \\ &= \gamma\theta^{-\alpha-1}(1-\theta)^{c+\alpha-1}\mathbf{1}_{(1/n,1]}(\theta)d\theta \cdot \theta^{x_{n,new,k}}(1-\theta)^r. \end{aligned}$$

From the results and examples of this section, Θ_n can be seen as a random truncation of Θ, and this is because the number of non-fixed atoms of the approximating priors Θ_n is a.s. finite for every $n \in \mathbb{N}$, while Θ has countably infinite many of them, and $\Theta_n \xrightarrow{d} \Theta$. We discuss the properties of the truncation procedure in more detail in Section 6. In the next result we show that, under certain conditions, we have automatic conjugacy for random measures in $\mathcal{A}'$ satisfying A00.

Proposition 5.7. *Let $\Theta \in \mathcal{A}'$ satisfying A00 and with weight rate measure having finite support. Let X be generated conditional on Θ according to $X := \sum_{k=1}^{K} x_k\delta_{\psi_k}$ with $x_k \overset{indep}{\sim} h(x|\theta_k)$ for proper, discrete probability mass function h. Assume that the characteristic functions of the random variables of the fixed component of Θ_{post} have no zeros, namely assume that*

$$\int_0^{\infty} e^{iz\theta}h(x|\theta)F_{fix,k}(d\theta) \neq 0 \quad \textit{and} \quad \int_0^{\infty} e^{iz\theta}h(x|\theta)\nu(d\theta) \neq 0, \tag{12}$$

for every $z \in \mathbb{R}$, $k \in [K_{fix}]$, and $x \in \mathbb{N}$ such that $h(x|\theta) > 0$ for some θ. Then, $\Theta_{post} \in \mathcal{A}'$, satisfies A00 and has weight rate measure with finite support.

Proof. Assumption (12) implies that the characteristic functions of $F_{post,fix,k}$ and of $F_{post,new,j}$ have no zeros. Further, they are also supported on a finite subset of $\mathbb{N} \cup \{0\}$. Then, by Theorem 4.8 we obtain the result. $\square$

Remark 5.8. Let Θ and X be as in Proposition 5.7. Notice that we can write $F_{fix,k} = \sum_{j=0}^{n^{(k)}} a_j^{(k)} \delta_j$, where $a_0^{(k)}, \dots, a_{n-1}^{(k)} \geq 0$, $a_n^{(k)} > 0$, and $a_0^{(k)} + \cdots + a_n^{(k)} = 1$, for $k \in K_{fix}$. Further, we can write $\nu = \sum_{j=1}^{K_\nu} b_j \delta_j$, where $K_\nu \in \mathbb{N}$ indicates the highest value in $\operatorname{supp}(\nu)$, $b_1, \dots, b_{K_\nu - 1} \geq 0$ and $b_{K_\nu} > 0$. Assumption (12) can be rewritten as: assume that

$$\sum_{j=0}^{n^{(k)}} e^{izj} h(x|j) a_j^{(k)} \neq 0 \quad \text{and} \quad \sum_{j=1}^{K_\nu} e^{izj} h(x|j) b_j \neq 0,$$

for every $z \in \mathbb{R}$, $k \in [K_{fix}]$, and $x \in \mathbb{N}$ such that $h(x|\theta) > 0$ for some θ Moreover, by Theorem 4.8 this assumption (and so assumption (12)) is equivalent to the following assumption: Assume that the polynomials $w \mapsto \sum_{j=0}^{n^{(k)}} h(x|j) a_j^{(k)} w^j$ and $w \mapsto \sum_{j=1}^{K_\nu} h(x|j) b_j w^j$ in the complex variable w have no roots on the unit circle, for every $k \in [K_{fix}]$ and $x \in \mathbb{N}$ such that $h(x|\theta) > 0$ for some θ.

Example 5.9. In this example we build on Example 4.10 and use the notation of Remark 5.8. Consider any point process with independent increments Θ satisfying A00, with weight rate measure having bounded support and with $a_0^{(k)} < \ldots < a_n^{(k)}$, for every $k \in K_{fix}$, and $b_1 < \ldots < b_{K_\nu}$. Then, by the Eneström-Kakeya Theorem $\Theta \in \mathcal{A}'$. Without loss of generality let $K_\nu \geq \max_{k=1,\dots,K_{fix}} n^{(k)}$. Consider a Bernoulli process likelihood, namely $x_k \sim \mathrm{Bern}(x|\frac{\theta_k}{K_\nu})$. Then, for each atom $\psi_{fix,k}$, letting $x_{post,fix,k} := X(\{\psi_{fix,k}\})$, the weight $\theta_{post,fix,k}$ has distribution

$$F_{post,fix,k}(d\theta) \propto F_{fix,k}(d\theta) \cdot h(x_{post,fix,k}|\theta) = \sum_{j=0}^{n^{(k)}} a_j^{(k)} \delta_j d\theta \cdot \frac{\theta}{K_\nu}.$$

Moreover, the posterior of Θ, has a fixed atom at $\psi_{new,k}$ whose weight $\theta_{post,new,k}$ has distribution

$$F_{post,new,k}(d\theta) \propto \nu(d\theta) h(x_{new,j,k}|\theta) = \sum_{j=1}^{K_\nu} b_j \delta_j d\theta \cdot \frac{\theta}{K_\nu}.$$

Thus, any fixed atom of Θ_{post} has distribution of the form $\sum_{j=0}^{\tilde{n}^{(k)}} \tilde{a}_j^{(k)} \delta_j$, where $0 \leq \tilde{a}_0^{(k)} < \ldots < \tilde{a}_{n-1}^{(k)} < a_n^{(k)}$, for $k \in K_{fix} + K_{new}$, and so its characteristic function has no zeros. Moreover, the ordinary component of the posterior has weight rate measure $\sum_{j=1}^{K_\nu} b_j \delta_j d\theta \cdot \frac{1-\theta}{K_\nu}$. Therefore, by Proposition 5.7 Θ_{post} belongs to $\mathcal{A}'$, satisfies $A00$, has weight rate measure with bounded support.

Remark 5.10. The results presented in this section hold also if the weight rate measure is infinite, namely $\nu(\mathbb{R}_+) = \infty$ (under the additional assumptions A1 and A2). In particular, the equivalent of Proposition 5.1 would be identical to Corollary 3.2 in Broderick et al. (2018) except for the result of point 1, because here we additionally know that $F_{fix,k}$ is QID with no Gaussian component and finite quasi-Lévy measure. Further, the equivalent of Proposition 5.3 would follows from the arguments presented taking into consideration Remark 3.6. The equivalent of Proposition 5.7 is more subtle and it is presented below.

Now, let $\mathcal{A}''$ be the class of QID CRMs of the form: $\xi \stackrel{a.s.}{=} \alpha + \sum_{j=1}^{K} \beta_j \delta_{s_j}$, with α an atomless point process with independent increments and finite Lévy measure, $\{s_j : j = 1, \dots, K\}$ a finite set of fixed atoms in S, and β_j, $j \geq 1$, $\mathbb{N} \cup \{0\}$-valued QID random variables that are mutually independent and independent of α. Let $\mathcal{A}''_\infty$ indicate the set of random measures like in $\mathcal{A}''$ but with α being any atomless point process with independent increments (hence with a possibly infinite Lévy measure).

As a side comment, we remark that is possible to see that similar results to Theorem 4.2 and Theorem 4.9 hold for the elements in $\mathcal{A}''$. In this case we would even know the structure of the Lévy-Khintchine representation in more details thanks to Theorem 8.1 in Lindner et al. (2018).

Proposition 5.11. *Let $\Theta \in \mathcal{A}''_\infty$ and assume A00, A0, A1 and A2. Let X be generated conditional on Θ according to $X := \sum_{k=1}^{\infty} x_k \delta_{\psi_k}$ with $x_k \stackrel{indep}{\sim} h(x|\theta_k)$ for proper, discrete probability mass function h. Assume that the characteristic functions of the random variables of the fixed component of Θ_{post} have no zeros, namely assume that*

$$\int_0^\infty e^{iz\theta} h(x|\theta) F_{fix,k}(d\theta) \neq 0 \quad \textit{and} \quad \int_0^\infty e^{iz\theta} h(x|\theta) \nu(d\theta) \neq 0, \tag{13}$$

for every $z \in \mathbb{R}$, $k \in [K_{fix}]$, and $x \in \mathbb{N}$ such that $h(x|\theta) > 0$ for some θ. Then, $\Theta_{post} \in \mathcal{A}''_\infty$ and satisfies A00, A0, A1 and A2.

Proof. Assumption (13) implies that the characteristic functions of $F_{post,fix,k}$ and of $F_{post,new,j}$ have no zeros. Further, they are also supported on $\mathbb{N} \cup \{0\}$. Then, by Theorem 8.1 in Lindner et al. (2018) we obtain the result. ☐

Notice that, thanks to Theorem 8.1 in Lindner et al. (2018), $\int_0^\infty e^{iz\theta} F_{fix,k}(d\theta) \neq 0$ for every $z \in \mathbb{R}$ and $k \in [K_{fix}]$, and that, given any $x \in \mathbb{N}$, if $h(x|\theta)$ is QID then $\int_0^\infty e^{iz\theta} h(x|\theta)(d\theta) \neq 0$ for every $z \in \mathbb{R}$. Further, assumption (13) can be rewritten more explicitly as done for assumption (12) in Remark 5.8.

The condition for automatic conjugacy in Propositions 5.7 and 5.11 is the absence of zeros in the characteristic function of the posterior, which in turn is a condition on the prior and the likelihood function. To the best of our knowledge this is the first time such condition for automatic conjugacy is explored, which usually relies on the exponential structure of the probability density function of the prior and of the likelihood function, see Broderick et al. (2018). We remark that other automatic conjugacy results can be obtained using properties of QID distributions, for example using the properties that a distribution with an atom with weight greater than 1/2 is QID (see Example 4.5).

6 Truncation analysis

In this section we investigate the properties of the truncation procedure considered in this paper.

Proposition 6.1. *Let ξ be a CRM. Thus, $\xi \overset{a.s.}{=} \sum_{j=1}^{K} \beta_j \delta_{s_j} + \gamma + \int_0^\infty \int_S x \delta_s \eta(ds\, dx)$ for some $K \in \mathbb{N} \cup \{0\} \cup \{\infty\}$, β_j's random variables, s_j's in S, deterministic measure γ, and Poisson random measure η. Let ξ_n be a CRM such that*

$$\xi_n \overset{a.s.}{=} \sum_{j=1}^{\min(n,K)} \beta_{n,j} \delta_{s_j} + \gamma_n + \int_{1/n}^{\infty} \int_{S_n} x \delta_s \eta'(ds\, dx),$$

where $\gamma_n = \gamma(\cdot \cap S_n)$, $\eta' \overset{d}{=} \eta$, and $\beta_{n,j}$'s are QID with no Gaussian component and finite quasi-Lévy measure with $\rho(\beta_{n,j}, \beta_j) < (1/n)^2$ if K is infinite, and $\beta_{n,j} \xrightarrow{d} \beta_j$, if K is finite. Then, $\xi_n \in \mathcal{A}$ and $\xi_n \xrightarrow{d} \xi$ as $n \to \infty$. If ξ has only ordinary component and if we choose $\eta' \overset{a.s.}{=} \eta$ then $\xi - \xi_n$ is a CRM independent of ξ_n and

$$\xi - \xi_n \overset{a.s.}{=} \int_0^\infty \int_{S \setminus S_n} x \delta_s \eta(ds\, dx) + \int_{1/n}^{\infty} \int_{S_n} x \delta_s \eta(ds\, dx), \tag{14}$$

and if in addition ξ satisfies A00 then

$$\xi - \xi_n \overset{a.s.}{=} \int_0^{1/n} \int_S x \delta_s \eta(ds\, dx). \tag{15}$$

Proof. Let F be the Lévy measure of ξ, namely $F = \mathbb{E}\eta$. Recall that for any measurable function $f \geq 0$ on S we have that $\int f dF = 0 \Leftrightarrow \int f d\eta \overset{a.s.}{=} 0$. Let $\tilde{\eta}(\cdot) \overset{a.s.}{=} \eta'(\cdot \cap (S_n \times (1/n, \infty)))$ and observe that $\tilde{\eta}$ is a Poisson random measure with intensity $\mathbb{E}\tilde{\eta}(\cdot) = F(\cdot \cap (S_n \times (1/n, \infty)))$. Let F_n be the Lévy measure of ξ_n, thus we have $F_n(\cdot) = \mathbb{E}\tilde{\eta}(\cdot) = \mathbb{E}\eta'(\cdot \cap (S_n \times (1/n, \infty))) = \mathbb{E}\eta(\cdot \cap (S_n \times (1/n, \infty))) = F(\cdot \cap (S_n \times (1/n, \infty)))$. Then, we obtain that $\xi_n \in \mathcal{A}$ and by the arguments in the proofs of Proposition 3.3 and Theorem 3.4 we obtain that $\xi_n \xrightarrow{d} \xi$ as $n \to \infty$. The independence comes from the independence of the increments of η and the equations (14) and (15) are easily obtained. □

Proposition 6.1 can be seen as an equivalent of Proposition 5.1 and Proposition 5.2 in Lee et al. (2019) applied to our truncation procedure. However, our result is more general because it applies to any CRM (with only ordinary component) without any further assumption. Our truncation procedure is a generalization of the ε-approximation developed in Argiento et al. (2016) because it is not restricted to CRMs with only ordinary component and with Lévy measure satisfying A00.

We talk about truncation method because ξ_n is a CRM with n fixed atoms and having the ordinary component composed by atoms with weights greater than $1/n$ and lying in S_n. Moreover, if $\eta' = \eta$, which is the usual assumption in truncation analysis and in series representation (see Lee et al. (2019) and Campbell et al. (2019)), then ξ_n is composed by the atoms of ξ which are located in S_n and have weight greater than $1/n$.

Thus, differently from Campbell et al. (2019), in the present truncation procedure the number of non-fixed atoms of the approximating CRM ξ_n is not arbitrarily fixed. Moreover, when they are considered as priors in the Bayesian setting of Section 5 they possess explicit posterior formulations and certain denseness and automatic conjugacy properties.

In comparison with usual truncation procedures (see Argiento et al. (2016); Campbell et al. (2019); Lee et al. (2019); Nguyen et al. (2021)) our procedure takes into account the fixed atomic component in a non-trivial way. In particular, by Theorem 4.2 the fixed atomic component (and so the whole CRM) has a simple Lévy-Khintchine formulation, and in some cases it also has an explicit Laplace transform (see Corollary 4.4), which apart from providing flexibility and improving computability it is also useful in nonparametric Bayesian spectral estimation, see Tobar (2018) and Meier et al. (2020), where the Fourier transform of the law of processes and/or random measures plays a key role, and in moment-matching criterion for quantifying approximations, see Arbel and Prünster (2017).

Building on Proposition 6.1, it is possible to obtain an upper bound for the L_1 error on the marginal likelihood when truncated CRMs are used for hierarchical Bayesian models, as in Campbell et al. (2019), Lee et al. (2019). In particular, consider $\Theta := \sum_k \theta_k \delta_{\psi_k}$ to be a CRM with only ordinary component and satisfying $A00$ (we denote by ν its weight rate measure), and let $h(\cdot|\theta)$ be a proper probability mass function on $\mathbb{N}\cup\{0\}$ for all θ in the support of ν. Denote by η its Poisson random measure. Consider a collection of conditionally independent observations $X_{1:N} := \{X_m\}_{m=1}^N$ given Θ defined by $X_m := \sum_k x_{m,k}\delta_{\psi_k}$ with $x_{m,k} \sim h(x|\theta_k)$ independently across k and iid across m. Further, define the observed data $Y_m|X_m \sim f(\cdot|X_m)$ for a conditional density f with respect to a measure κ on some space. Let $\Theta_n \in \mathcal{A}$ be an element of the approximating sequence of CRMs driven by a Poisson random measure η_n everywhere equals to η on $(1/n, \infty) \times S$ and almost surely equals to zero otherwise. Define $Z_{1:N}$ and $W_{1:N}$ for Θ_n analogous to the definitions of $X_{1:N}$ and $Y_{1:N}$ for Θ, and let $p_{N,\infty}$ and $p_{N,n}$ be the marginal densities (with respect to κ) of the final observations $Y_{1:N}$ and $W_{1:N}$.

Differently from the setting of Section 5 here Θ has only ordinary component, and the Poisson random measures of Θ_n are specified in more detail (in Section 5 the specification is only on the weight rate measure). We stress that $h(\cdot|\theta)$ is defined for all θ in the support of ν meaning that $h(\cdot|0)$ is not defined (or equivalently $h(\cdot|0) \equiv 0$) and so the $\mathrm{supp}(X_{1:N}) \subset \mathrm{supp}(\Theta)$ everywhere (as in Section 5).

Proposition 6.2. *We have the bound* $\frac{1}{2}\|p_{N,\infty} - p_{N,n}\|_1 \leq 1 - e^{\int_0^{1/n}(1-h(0|\theta)^N)\nu(d\theta)}$*, and if $A2$ is satisfied then* $\|p_{N,\infty} - p_{N,n}\|_1 \to 0$ *as* $n \to \infty$.

Proof. We follow similar steps as the ones in the proof of Theorem 4.3 in Campbell et al. (2019). By Lemma 4.1 in Campbell et al. (2019) (and Lemma D.1 in its supplementary material) we have that $\frac{1}{2}\|p_{N,\infty} - p_{N,n}\|_1 \leq 1 - \mathbb{P}(\mathrm{supp}(X_{1:N}) \subset \mathrm{supp}(\Theta_n))$. Observe that $\mathbb{P}(\mathrm{supp}(X_{1:N}) \subset \mathrm{supp}(\Theta_n)) = \mathbb{P}(\mathrm{supp}(X_{1:N}) \cap \mathrm{supp}(\Theta - \Theta_n) = \emptyset)$. By construction, if $\theta_k(\omega) > 0$ then $x_{m,k}(\omega) \geq 0$ and if $\theta_k(\omega) = 0$ then $x_{m,k}(\omega) = 0$. Further, notice that $\Theta h(0|\theta)^N$ is a thinned CRM such that $\mathbb{P}(\mathrm{supp}(X_{1:N}) \cap \mathrm{supp}(\Theta h(0|\theta)^N) = \emptyset) = 1$, and since $\Theta - \Theta_n$ is everywhere composed by the atoms of Θ with weights less than or equal to $1/n$, we have that $\mathbb{P}(\mathrm{supp}(X_{1:N}) \cap \mathrm{supp}((\Theta - \Theta_n)h(0|\theta)^N) = \emptyset) = 1$. Thus, we are left with the CRM $(\Theta - \Theta_n)(1 - h(0|\theta)^N)$ and using the fact that a Poisson process with measure $\mu(d\theta)$ has no atoms with probability $e^{-\int \mu(d\theta)}$ we have

$$\mathbb{P}(\mathrm{supp}(X_{1:N}) \subset \mathrm{supp}(\Theta_n)) = e^{\int_0^{1/n}(1-h(0|\theta)^N)\nu(d\theta)}.$$

If $A2$ is satisfied, namely $\int_0^{1/n}(1 - h(0|\theta)^N)\nu(d\theta) < \infty$, then $e^{\int_0^{1/n}(1-h(0|\theta)^N)\nu(d\theta)} \to 1$ as $n \to \infty$. □

It would be interesting to extend the previous proposition to the case of Θ with a fixed component too. However, adding countably many fixed points makes the bound developed in Lemma 4.1 in Campbell et al. (2019) not useful because $\mathbb{P}(\mathrm{supp}(X_{1:N}) \subset \mathrm{supp}(\Theta_n)) = 0$ and this is true in our case as well as in Campbell et al. (2019). On the other hand, since the elements of $\mathcal{A}$ have an explicit Lévy-Khintchine formulation it might be possible to quantify the approximation by a moment-matching criterion, extending Arbel and Prünster (2017) to any CRMs.

7 Discussion

In this work we first prove a denseness result for a class of QID CRMs, which we call $\mathcal{A}$, in the space of all CRMs. The elements of this class have Lévy-Khintchine formulations, in some cases even explicit Laplace transform, and their law are uniquely determined by characteristic pairs. These results allow the fixed atomic component of (asymptotically) all CRMs, which is usually disregarded in the analysis of CRMs and of their truncations, to be fully accessible.

In the proof of the denseness result we choose a certain approximation for the ordinary and the deterministic components (see Proposition 3.3). The denseness result and the whole paper up to Section 6 could be rewritten using another approximation for these components and the nature of the results and of the examples would be the same. We have chosen such approximation due to its simplicity and its capacity to adapt to the deterministic and ordinary components of any CRM. Moreover, such approximation leads to extremely nice truncation properties, presented in Section 6. Remarkably, our truncation procedure is different from the ones surveyed and analysed in Arbel and Prünster (2017); Campbell et al. (2019); Nguyen et al. (2021); Lee et al. (2019), it is more general than Argiento et al. (2016); Campbell et al. (2019); Nguyen et al. (2021); Lee et al. (2019) since it does not impose any assumption on the Lévy measure of the CRM, and in addition it allows for a non-trivial consideration of the atomic component.

We show the relevance of our results using the (hierarchical) nonparametric Bayesian settings of Broderick et al. (2018) and of Campbell et al. (2019) and by means of various examples. In particular, we show that the elements of $\mathcal{A}$ with homogeneous Lévy measure when considered as priors have an explicit posterior distribution and possess certain denseness and automatic conjugacy properties. Concerning the truncation analysis, we provide an upper bound in the total variation distance between the data distributions induced by the full and truncated priors in a general hierarchical Bayesian model.

Due to the general nature of our results, their applicability is not limited to these settings, but they can be applied to different frameworks to answer numerous questions. A first research question concerns the extension of the results presented in this paper to the framework of normalized CRMs. The same research question applies to the class of correlated random measures (see Ranganath and Blei (2018)) and of dependent normalized CRMs (see Camerlenghi et al. (2019); Lijoi et al. (2014)). A second research question concerns the connections between normalized QID CRMs and the Dirichlet process and/or the Pitman-Yor process (see Arbel et al. (2019); Lijoi et al. (2020); Pitman (2002)). Further, it would be interesting to explore the role of normalised CRMs taking values in the space of QID distributions, given that QID distributions are dense in the space of distributions.

Normalized CRMs are vastly used in Bayesian analysis and answering these questions might shed new lights on their theoretical properties and their applicability. We leave these questions to further research.

References

Arbel, J., Blasi, P. D., and Prünster, I. (2019). "Stochastic Approximations to the Pitman–Yor Process." *Bayesian Analysis*, 14(4): 1201–1219. MR4136558. doi: `https://doi.org/10.1214/18-BA1127`. 283

Arbel, J. and Prünster, I. (2017). "A moment-matching Ferguson and Klass algorithm." *Statistics and Computing*, 27(1): 3–17. MR3598905. doi: `https://doi.org/10.1007/s11222-016-9676-8`. 256, 281, 282

Argiento, R., Bianchini, I., and Guglielmi, A. (2016). "Posterior sampling from ε-approximation of normalized completely random measure mixtures." *Electronic Journal of Statistics*, 10(2): 3516–3547. MR3572858. doi: `https://doi.org/10.1214/16-EJS1168`. 256, 280, 281, 282

Berger, D. (2019). "On quasi-infinitely divisible distributions with a point mass." *Mathematische Nachrichten*, 292(8): 1674–1684. MR3994295. doi: `https://doi.org/10.1002/mana.201800073`. 255

Broderick, T., Mackey, L., Paisley, J., and Jordan, M. I. (2015). "Combinatorial Clustering and the Beta Negative Binomial Process." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(2): 290–306. doi: `https://doi.org/10.1109/TPAMI.2014.2318721`. 276

Broderick, T., Wilson, A. C., and Jordan, M. I. (2018). "Posteriors, conjugacy, and exponential families for completely random measures." *Bernoulli*, 24(4B): 3181–3221. MR3788171. doi: `https://doi.org/10.3150/16-BEJ855`. 256, 272, 273, 274, 275, 276, 277, 279, 282

Camerlenghi, F., Lijoi, A., Orbanz, P., and Prünster, I. (2019). "Distribution theory for hierarchical processes." *The Annals of Statistics*, 47(1): 67–92. MR3909927. doi: `https://doi.org/10.1214/17-AOS1678`. 283

Campbell, T., Huggins, J. H., How, J. P., and Broderick, T. (2019). "Truncated random measures." *Bernoulli*, 25(2): 1256–1288. MR3920372. doi: https://doi.org/10.3150/18-bej1020. 256, 272, 273, 275, 280, 281, 282

Chhaiba, H., Demni, N., and Mouayn, Z. (2016). "Analysis of generalized negative binomial distributions attached to hyperbolic Landau levels." *Journal of Mathematical Physics*, 57(7): 072103. MR3522606. doi: https://doi.org/10.1063/1.4958724. 255

Cuppens, R. (1975). *Decomposition of Multivariate Probabilities*. Probability and Mathematical Statistics: A Series of Monographs and Textbooks. Academic Press. doi: https://doi.org/10.1016/B978-0-12-199450-1.50001-7. 269

Demni, N. and Mouayn, Z. (2015). "Analysis of generalized Poisson distributions associated with higher Landau levels." *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 18(04): 1550028. MR3447229. doi: https://doi.org/10.1142/S0219025715500289. 255

Griffiths, T. L. and Ghahramani, Z. (2005). "Infinite Latent Feature Models and the Indian Buffet Process." In *In NIPS*, 475–482. MIT Press. 276

Horowitz, J. (1977). "Une remarque sur les bimesures." In Dellacherie, C., Meyer, P. A., and Weil, M. (eds.), *Séminaire de Probabilités XI*, 59–64. Berlin, Heidelberg: Springer Berlin Heidelberg. MR0450494. doi: https://doi.org/10.1007/BFb0087188. 269

James, L. F. (2017). "Bayesian Poisson calculus for latent feature modeling via generalized Indian Buffet Process priors." *The Annals of Statistics*, 45(5): 2016–2045. MR3718160. doi: https://doi.org/10.1214/16-AOS1517. 274, 276

James, L. F., Lijoi, A., and Prünster, I. (2009). "Posterior Analysis for Normalized Random Measures with Independent Increments." *Scandinavian Journal of Statistics*, 36(1): 76–97. MR2508332. doi: https://doi.org/10.1111/j.1467-9469.2008.00609.x. 254

Kallenberg, O. (1983). *Random Measures*. 3rd edn. Akademie-Verlag, Berlin. MR0431372. 259, 263

Kallenberg, O. (2002). *Foundations of Modern Probability*. Probability and Its Applications. Springer New York. MR1876169. doi: https://doi.org/10.1007/978-1-4757-4015-8. 258, 261, 262, 265, 268, 269

Kallenberg, O. (2017). *Random Measures, Theory and Applications*. Probability Theory and Stochastic Modelling series. Springer. MR3642325. doi: https://doi.org/10.1007/978-3-319-41598-7. 257, 259, 261, 262, 263, 264, 266, 267, 269, 270

Khartov, A. (2019). "Compactness criteria for quasi-infinitely divisible distributions on the integers." *Statistics and Probability Letters*, 153: 1–6. MR3957474. doi: https://doi.org/10.1016/j.spl.2019.05.008. 255

Kingman, J. (1993). *Poisson Processes*. Oxford University Press. MR1207584. 254, 261, 269, 273

Lee, J., Miscouridou, X., and Caron, F. (2019). "A unified construction for series representations and finite approximations of completely random measures." URL https://arxiv.org/abs/1905.10733. 254, 256, 280, 281, 282

Lijoi, A., Nipoti, B., and Prünster, I. (2014). "Bayesian inference with dependent normalized completely random measures." *Bernoulli*, 20(3): 1260–1291. MR3217444. doi: https://doi.org/10.3150/13-BEJ521. 283

Lijoi, A. and Prünster, I. (2010). *Models beyond the Dirichlet process*, 80–136. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press. MR2730661. doi: https://doi.org/10.1017/CBO9780511802478.004. 254

Lijoi, A., Prünster, I., and Rigon, T. (2020). "The Pitman–Yor multinomial process for mixture modelling." *Biometrika*, 107(4): 891–906. MR4186494. doi: https://doi.org/10.1093/biomet/asaa030. 283

Lindner, A., Pan, L., and Sato, K. (2018). "On quasi-infinitely divisible distribution." *Transactions of the American Mathematical Society*, 370: 8483–8520. MR3864385. doi: https://doi.org/10.1090/tran/7249. 255, 258, 259, 260, 261, 266, 267, 270, 271, 272, 279

Meier, A., Kirch, C., and Meyer, R. (2020). "Bayesian nonparametric analysis of multivariate time series: A matrix Gamma Process approach." *Journal of Multivariate Analysis*, 175: 104560. MR4028661. doi: https://doi.org/10.1016/j.jmva.2019.104560. 281

Nakamura, T. (2015). "A complete Riemann zeta distribution and the Riemann hypothesis." *Bernoulli*, 21(1): 604–617. MR3322332. doi: https://doi.org/10.3150/13-BEJ581. 255

Nehring, B., Poghosyan, S., and Zessin, H. (2013). "On the construction of point processes in statistical mechanics." *Journal of Mathematical Physics*, 54(6): 063302. MR3112542. doi: https://doi.org/10.1063/1.4807724. 272

Nguyen, T. D., Huggins, J., Masoero, L., Mackey, L., and Broderick, T. (2021). "Independent finite approximations for Bayesian nonparametric inference." URL https://arxiv.org/abs/2009.10780. 256, 281, 282

Passeggeri, R. (2020a). "A density property for stochastic processes." URL https://arxiv.org/abs/2010.07752. 255

Passeggeri, R. (2020b). "On the extension and kernels of signed bimeasures and their role in stochastic integration." URL https://arxiv.org/abs/2009.10657. 255

Passeggeri, R. (2020c). "Spectral representations of quasi-infinitely divisible processes." *Stochastic Processes and their Applications*, 130(3): 1735–1791. MR4058288. doi: https://doi.org/10.1016/j.spa.2019.05.014. 255, 258, 269

Pitman, J. (2002). *Combinatorial Stochastic Processes*. École d'Été de Probabilités de Saint-Flour. Springer-Verlag Berlin Heidelberg. MR2245368. doi: https://doi.org/10.1007/b11601500. 283

Ranganath, R. and Blei, D. M. (2018). "Correlated Random Measures." *Journal of the American Statistical Association*, 113(521): 417–430. MR3803475. doi: https://doi.org/10.1080/01621459.2016.1260468. 283

Regazzini, E., Lijoi, A., and Prünster, I. (2003). "Distributional results for means of normalized random measures with independent increments." *The Annals of Statistics*, 31(2): 560–585. MR1983542. doi: https://doi.org/10.1214/aos/1051027881. 254

Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, UK, 1st edition. MR1739520. 270

Teh, Y. W., Jordan, M. I., Beal, M. J., and Blei, D. M. (2006). "Hierarchical Dirichlet Processes." *Journal of the American Statistical Association*, 101(476): 1566–1581. MR2279480. doi: https://doi.org/10.1198/016214506000000302. 273

Tobar, F. (2018). "Bayesian Nonparametric Spectral Estimation." In *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc. 281

Zhang, H., Liu, Y., and Li, B. (2014). "Notes on discrete compound Poisson model with applications to risk theory." *Insurance: Mathematics and Economics*, 59: 325–336. MR3283233. doi: https://doi.org/10.1016/j.insmatheco.2014.09.012. 255

Zolotarev, V. (2001). "Lévy metric." In Hazewinkel, M. (ed.), *Encyclopedia of Mathematics*, volume 5. Kluwer Academic Publishers. MR1093240. doi: https://doi.org/10.1007/978-94-009-6000-8. 261

Acknowledgments

The author would like to thank Almut Veraart, Fabio Bernasconi, Ismael Castillo and Antonio Lijoi for useful discussions. We also thank two anonymous Referees, the Associate Editor, and the Editor-in-Chief for their numerous constructive comments, which improved the presentation and the content of the work.