Integrability conditions for Compound Random Measures

Alan Riva Palacio[a], Fabrizio Leisen[a,*]

[a] *School of Mathematics, Statistics and Actuarial Sciences, University of Kent, UK*

Abstract

Compound random measures (CoRM's) are a flexible and tractable framework for vectors of completely random measure. In this paper, we provide conditions to guarantee the existence of a CoRM. Furthermore, we prove some interesting properties of CoRM's when exponential scores and regularly varying Lévy intensities are considered.

Keywords: Bayesian non-parametrics, Multivariate Lévy measure, Partial exchangeability, Exchangeable Partition Probability function

1. Introduction

Recently, a growing literature in Bayesian non-parametrics (BNP) proposed new priors which can take into account different features of the data, such as *partial exchangeability*, see De Finetti (1938). In this case, one would like to consider different densities for different groups instead of a single common density for all the data. After the seminal paper of MacEachern (1999), the problem of modeling a finite number of dependent densities has become an active area of research in Bayesian non-parametrics. A common approach is to construct BNP priors based on functions of *Completely Random Measures* (CRM's), see Lijoi and Prünster (2010). For example, special attention has been given to the normalization of CRM's starting with the work of Regazzini, Lijoi and Prünster (2003). Roughly speaking, a CRM is a generalization of a subordinator, that is a process with independent increments and almost surely increasing paths; for a full account of CRM's see Kingman (1993). This property is very helpful to derive the Laplace functional transform which is the basis to derive some analytical quantities of interest such as posterior and predictive distributions or the Exchangeable Partition Probability Function (EPPF), see James, Lijoi and Prünster (2009). To build more flexible priors in possibly higher dimensional spaces, vectors of dependent CRM's

*Corresponding author at: School of Mathematics, Statistics and Actuarial Sciences Cornwallis Building, University of Kent, Canterbury, CT2 7NF, UK.

Email addresses: ar515@kent.ac.uk (Alan Riva Palacio), fabrizio.leisen@gmail.com (Fabrizio Leisen)

are constructed for example in Leisen and Lijoi (2011), Leisen, Lijoi and Spano (2013) and Zhu and Leisen (2015) where respectively they build vectors of Poisson-Dirichlet and Dirichlet processes. These papers deal with the Lévy-Copula approach introduced in Cont and Tankov (2004) to induce dependence among the components of the vector. In a similar fashion, Lijoi, Nipoti and Prünster (2014a) introduce a vector of random probability measures where the dependence arises by virtue of a suitable construction of the Poisson random measures underlying the CRM's; furthermore, in the framework of survival analysis, Lijoi and Nipoti (2014) introduce a new class of vectors of random hazard rate functions that are expressed as kernel mixtures of dependent CRM's. Camerlenghi, Lijoi and Prünster (2017) focus on partial exchangeable models which arise from hierarchical specifications of CRM's.

Compound random measures (CoRM's), introduced by Griffin and Leisen (2017), are a flexible and tractable framework for many dependent random measures including many of the superposition and Lévy copula approaches. They have recently been applied to modeling graphs for overlapping communities by Todeschini, Miscouridou and Caron (2017). Griffin and Leisen (2017) and Griffin and Leisen (2018) described posterior sampling methods for a particular class of normalized compound random measure mixtures which exploits a representation of the Laplace transform of the CoRM through a univariate integral of a moment generating function.

In this paper we aim to provide explicit existence conditions for CoRM's in order to guarantee the existence of the marginal Lévy intensities. On the other hand, we prove that the resulting CoRM is well posed in the sense that it satisfies the usual integrability condition for multivariate Lévy processes. Furthermore, this paper provides an interesting result for CoRM's when regularly varying Lévy intensities are considered. The paper closes highlighting a representation on the multivariate Lévy intensity of a CoRM when the score distribution is the result of marginal independent and identically distributed exponential scores.

The outline of the paper is as follows. Section 2 will set the scene by introducing the basic definitions which are required in the CoRM setting. Section 3 is devoted to prove our main results. Section 4 deals with CoRM's built with regularly varying Lévy intensities and exponential scores. Section 5 concludes.

2. Preliminaries

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathbb{X}$ a Polish space with corresponding Borel σ-algebra $\mathcal{X}$. We denote by $\mathbb{M}_{\mathbb{X}}$ the space of boundedly finite measures on the measurable space $(\mathbb{X}, \mathcal{X})$ and by $\mathcal{M}_{\mathbb{X}}$ the associated Borel σ-algebra, see Appendix 2 in Daley and Vere-Jones (2003) for technical details.

Definition 1. *A random measure μ on $\mathbb{X}$ is called a completely random measure (CRM) if for any $n > 1$ and disjoint sets $A_1, \ldots, A_n \in \mathcal{X}$ the random variables $\mu(A_1), \ldots, \mu(A_n)$ are mutually independent.*

In the following we consider only CRM's without fixed jumps, namely CRM's of the form $\mu = \sum_{i=1}^{\infty} w_i \delta_{u_i}$ for collections of random variables $\{w_i\}_{i=1}^{\infty}$ in $\mathbb{R}^+$ and $\{u_i\}_{i=1}^{\infty}$ in $\mathbb{X}$. Such CRM's can be characterized by their Laplace transform

$$\mathbb{E}\left[e^{-\mu(f)}\right] = e^{-\int_{\mathbb{R}^+ \times \mathbb{X}} (1 - e^{-sf(x)}) \tilde{\nu}(\mathrm{d}s, \mathrm{d}x)}$$

where $\mu(f) = \int_{\mathbb{X}} f(x)\mu(\mathrm{d}x)$, $f : \mathbb{X} \to \mathbb{R}^+$ is such that $\mu(f) < \infty$ and $\tilde{\nu}(\mathrm{d}s, \mathrm{d}x)$ is a measure in $(\mathbb{R}^+ \times \mathbb{X}, \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{X})$ such that

$$\int_{\mathbb{R}^+ \times X} \min\{1, s\} \tilde{\nu}(\mathrm{d}s, \mathrm{d}x) < \infty \tag{1}$$

for any bounded set $X \in \mathcal{X}$. A measure $\tilde{\nu}$ satisfying the condition displayed in equation (12) is called the *Lévy intensity* of μ. We say that $\tilde{\nu}$ is homogeneous when $\tilde{\nu}(\mathrm{d}s, \mathrm{d}x) = \rho(\mathrm{d}s)\alpha(\mathrm{d}x)$ with ρ a measure in $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ and α a measure in $(\mathbb{X}, \mathcal{X})$. The notion of a completely random measure can be generalized to higher dimensions in a similar fashion to Definition 1, see for instance Griffin and Leisen (2017). As a result, we have a representation in terms of a Laplace functional transform. Precisely, for a vector of completely random measures $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)$ on $\mathbb{X}$ we have that

$$\mathbb{E}\left[e^{-\mu_1(f_1) - \cdots - \mu_d(f_d)}\right] = e^{-\int_{(\mathbb{R}^+)^d \times \mathbb{X}} \left(1 - e^{-s_1 f_1(x) - \cdots - s_d f_d(x)}\right) \tilde{\nu}_d(\mathrm{d}\boldsymbol{s}, \mathrm{d}x)}$$

with $f_j : \mathbb{X} \to \mathbb{R}^+$, $j \in \{1, \ldots, d\}$ such that $\mu_j(f_j) < \infty$, where for $g : \mathbb{X} \to \mathbb{R}^+$ we have $\mu_j(g) = \int_{\mathbb{X}} g(x)\mu_j(\mathrm{d}x)$. The measure $\tilde{\nu}_d$ in $\left((\mathbb{R}^+)^d \times \mathbb{X}, \mathcal{B}((\mathbb{R}^+)^d \otimes \mathcal{X}\right)$ must be such that

$$\int_{(\mathbb{R}^+)^d \times \mathbb{X}} \min\{1, \|\boldsymbol{s}\|\} \tilde{\nu}_d(\mathrm{d}\boldsymbol{s}, \mathrm{d}x) < \infty \tag{2}$$

for any bounded set $X \in \mathcal{X}$; we call $\tilde{\nu}_d$ a multivariate Lévy intensity. We set the notation

$$\nu_j(A, X) = \int_{(\mathbb{R}^+)^{d-1}} \tilde{\nu}_d(\mathrm{d}s_1, \ldots, \mathrm{d}s_{j-1}, A, \mathrm{d}s_{j+1}, \ldots, \mathrm{d}s_d, X)$$

with $j \in \{1, \ldots, d\}$ and $A \in \mathcal{B}(\mathbb{R}^+)$. We call ν_j the j-th marginal of the d-variate Lévy intensity $\tilde{\nu}_d$; it

follows that for each $j \in \{1, \ldots, \mathrm{d}\}$, μ_j has Lévy intensity ν_j. In this framework we can define the concept of *Compound Random Measure* (CoRM). The following definition differs from the one in Griffin and Leisen (2017) since it takes into account the *inhomogeneous* case, where the locations and associated weights in the CRM are not independent.

Definition 2. *A Compound Random Measure (CoRM) is a vector of CRM's whose Lévy intensity is given by*

$$\tilde{\nu}_d(\mathrm{d}\boldsymbol{s}, \mathrm{d}x) = \int_{\mathbb{R}^+} z^{-d} h\left(\frac{s_1}{z}, \ldots, \frac{s_d}{z}\right) \mathrm{d}\boldsymbol{s}\nu^\star(\mathrm{d}z, \mathrm{d}x) \tag{3}$$

where h, the score distribution, is a d-variate probability density function and, $\nu^\star$, the directing Lévy measure, is a Lévy intensity.

By performing a simple change of variable we note that

$$\int_{(\mathbb{R}^+)^d} z^{-d} h\left(\frac{s_1}{z}, \ldots, \frac{s_d}{z}\right) \mathrm{d}\boldsymbol{s} = 1.$$

Therefore, $z^{-d} h\left(\frac{s_1}{z}, \ldots, \frac{s_d}{z}\right)$ can be seen as the density of a distribution function $H\left(\frac{\mathrm{d}s_1}{z}, \ldots, \frac{\mathrm{d}s_d}{z}\right)$. This allows to write the multivariate Lévy intensity in equation (3) as

$$\tilde{\nu}_d(\mathrm{d}\boldsymbol{s}, \mathrm{d}x) = \int_{\mathbb{R}^+} H\left(\frac{\mathrm{d}s_1}{z}, \ldots, \frac{\mathrm{d}s_d}{z}\right) \nu^\star(\mathrm{d}z, \mathrm{d}x). \tag{4}$$

To write the Lévy intensity of a CoRM in terms of distribution functions rather than densities will be convenient for the results proved in the next section.

3. Integrability conditions

The specification of a CoRM needs the initial choice of a score distribution and a directing Lévy measure. Although this sounds straightforward, it is necessary to check that theses choices lead to a well defined CoRM. Otherwise, the risk is to perform a Bayesian statistical analysis based on an ill-posed prior. In this section we look at two important aspects of Definition 2:

1. we provide conditions on the score distribution and the directing Lévy measure for the existence of the marginal Lévy intensities of a CoRM, see Theorem 1 and Corollary 1,
2. we provide conditions on the score distribution and the directing Lévy measure for the existence of the multivariate Lévy intensity of a CoRM, see Theorem 2.

Summing up, Theorem 1 and Corollary 1 are focusing on the marginal existence of a CoRM. On the other hand, Theorem 2 focuses on the global existence of a CoRM. The proofs of the theorems can be found in the supplementary material.

Theorem 1. *Let H be a d-variate score distribution and $\nu^\star$ a directing Lévy measure defining a measure $\tilde{\nu}_d$ as in* (4) *with corresponding marginals ν_j for $j \in \{1, \ldots, \mathrm{d}\}$. Let X be a bounded set in $\mathcal{X}$, then the measure ν_j satisfies the integrability condition (12) if and only if*

$$\int_{(0,1)\times X} \mathbb{P}\left[S_j \geq \frac{1}{z}\right] \nu^\star(\mathrm{d}z, \mathrm{d}x) < \infty \tag{5}$$

and

$$\int_{[1,\infty)\times X} \mathbb{P}\left[S_j \leq \frac{1}{z}\right] z\nu^\star(\mathrm{d}z, \mathrm{d}x) < \infty. \tag{6}$$

Furthermore if the marginal score H_j satisfies that

$$1 - H_j\left(\frac{1}{z}\right) \leq z \;\forall\; z \in (0, \epsilon) \text{ for some } \epsilon > 0 \tag{7}$$

and

$$\lim_{z\to\infty} zH_j\left(\frac{1}{z}\right) < \infty \tag{8}$$

then conditions (15), (16) are satisfied with an arbitrary choice of the directing Lévy measure $\nu^\star$.

As set in Definition 2, we usually work with CoRM's given by a score with a probability density; in such case the following corollary to Theorem 1 follows.

Corollary 1. *If S_j has probability density function h_j then conditions (17), (18) reduce to*

$$\lim_{z\to 0} \frac{h_j\left(\frac{1}{z}\right)}{z^2} < 1 \tag{9}$$

and

$$\lim_{\epsilon\to 0} h_j(\epsilon) < \infty. \tag{10}$$

The previous results concerned conditions for the marginals of a CoRM to be well defined, now we focus on such a result for the CoRM. For a score density function h and directing Lévy measure $\nu^\star$ to properly define a CoRM we need to check the condition (2) which takes the form

$$\int_{\mathbb{R}^+\times X}\int_{(\mathbb{R}^+)^d}\min\{1,\|\boldsymbol{s}\|\}h\left(\frac{s_1}{z},\ldots,\frac{s_d}{z}\right)\frac{\mathrm{d}\boldsymbol{s}}{z^d}\nu^\star(\mathrm{d}z,\mathrm{d}x)<\infty \tag{11}$$

for bounded set $X\in\mathcal{X}$. As stated at the beginning of this section, in the next theorem we provide conditions on the score distribution and the directing Lévy measure for the existence of the multivariate Lévy intensity of a CoRM. This is equivalent to provide conditions such that the latter inequality holds true.

Theorem 2. *Consider a CoRM which satisfies conditions (15) and (16) for each marginal ν_j, $j\in\{1,\ldots,\mathrm{d}\}$, then the integrability condition (11) is satisfied.*

We conclude this section by providing two examples of the use of the previous results when considering Gamma and Beta distributed scored distributions.

Example 1: Gamma scores

We consider the marginal gamma score case. Let h be the d-variate probability density of the score distribution; for $j\in\{1,\ldots,d\}$ we denote the j-th marginal density h_j and let it correspond to a Gamma distribution with shape and rate parameters $\alpha_j,\ \beta_j$, i.e.

$$h_j(s)=\frac{\beta_j^{\alpha_j}s^{\alpha_j-1}e^{-\beta_j s}}{\Gamma(\alpha_j)}\mathbb{1}_{\{s\in(0,\infty)\}}.$$

We check the constraints (17), (18) by making use of Corollary 1 as we have probability densities. To check (24) we see that

$$\lim_{s\to0}\frac{h_j\left(\frac{1}{s}\right)}{s^2}=\lim_{s\to0}\frac{\beta_j^{\alpha_j}e^{-\frac{\beta_j}{s}}}{\Gamma(\alpha_j)s^{\alpha_j+1}}=0$$

and constraint (25) is satisfied for arbitrary Lévy directing measure $\nu^\star$ whenever $\alpha_j\geq1$, as in the examples presented in Griffin and Leisen (2017). However for $\alpha_j<1$ the associated CoRM will be well posed depending on the choice of $\nu^\star$. If for example we take the directing Lévy measure to be the σ-stable, i.e.

$$\nu^\star(\mathrm{d}z,\mathrm{d}x)=\frac{\sigma}{z^{1+\sigma}}\mathrm{d}z\mathrm{d}x$$

then constraint (15) in Theorem 1 can be reduced to

$$\int_0^1 \int_{\frac{1}{z}}^{\infty} h_j(s) \frac{\sigma}{z^{1+\sigma}} \mathrm{d}s\mathrm{d}z = \int_1^{\infty} \int_{\frac{1}{s}}^1 h_j(s) \frac{\sigma}{z^{\sigma+1}} \mathrm{d}z\mathrm{d}s = \int_1^{\infty} h_j(s)(s^{\sigma}-1)\mathrm{d}s < \infty,$$

which is always satisfied since h_j is a Gamma density. On the other hand, condition (16) in Theorem 1 becomes

$$\int_1^{\infty} \int_0^{\frac{1}{z}} h_j(s) \frac{\sigma z}{z^{1+\sigma}} \mathrm{d}s\mathrm{d}z = \int_0^1 \int_1^{\frac{1}{s}} h_j(s) \frac{\sigma}{z^{\sigma}} \mathrm{d}z\mathrm{d}s = \int_0^1 h_j(s)\sigma(s^{\sigma-1}-1)\mathrm{d}s < \infty$$

which is not satisfied when $\alpha_j + \sigma < 1$.

Example 2: Beta scores

In the setting as above, if the marginal scores are Beta distributed, i.e.

$$h_j(s) = \frac{s^{\alpha_j - 1}(1-s)^{\beta_j - 1}}{B(\alpha_j, \beta_j)} \mathbb{1}_{\{s \in (0,1)\}}$$

then constraint (17) becomes

$$\lim_{s \to 0} \frac{h_j\left(\frac{1}{s}\right)}{s^2} = \lim_{s \to 0} \frac{(s-1)^{\beta_j - 1}}{s^{\alpha_j + \beta_j} B(\alpha_j, \beta_j)} = 0,$$

so it is always satisfied; and condition (18) is satisfied whenever $\alpha_j \geq 1$. We consider again a σ-stable Lévy intensity for $\nu^{\star}$ when $\alpha_j < 1$. Proceeding as in the previous example, constraint (15) becomes

$$\int_1^{\infty} \int_{\frac{1}{s}}^1 h_j(s) \frac{\sigma}{z^{\sigma+1}} \mathrm{d}z\mathrm{d}s = 0 < \infty$$

so it always holds; and constraint (16) becomes

$$\int_0^1 h_j(s)\sigma(s^{\sigma-1}-1)\mathrm{d}s < \infty,$$

which does not hold if $\alpha_j + \sigma < 1$.

4. Other interesting properties

The aim of this section is to investigate two interesting properties of CoRM's. First, we focus on CoRM's which arise from regurlarly varying directing Lévy measures. This result is motivated by the recent papers of Caron and Fox (2017) and Todeschini, Miscouridou and Caron (2017) which made use of regularly

varying Lévy measures to construct sparse random graphs. Second, we provide an explicit expression of the multivariate Lévy intensity of a CoRM with independent exponential scores. This result is interesting when compared with Theorem 3.2 in Zhu and Leisen (2015) and Corollary 2 in Griffin and Leisen (2017) which provide, respectively, the Lévy copula representation and the Laplace exponent of CoRM's with independent exponential scores. The proofs of the theorems can be found in the supplementary material.

Both results deal with a d-variate CoRM given by an homogeneous directing Lévy intensity $\nu^\star(\mathrm{d}z, \mathrm{d}x) = \rho^\star(\mathrm{d}z)\alpha(\mathrm{d}x)$. Therefore, the corresponding marginals can be written as $\nu_j(\mathrm{d}s, \mathrm{d}x) = \rho_j(\mathrm{d}s)\alpha(\mathrm{d}x)$ with ρ_j a measure in $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$.

4.1. Regularly varying directing Lévy measure

In this section we focus on CoRM's given by a directing Lévy measure that is regularly varying. We recall that a real valued function L is slowly varying if $\lim_{t\to\infty} L(at)/L(t) = 1 \ \forall a > 0$.

Definition 3. *An homogeneous Lévy measure $\rho^\star(\mathrm{d}z)\alpha(\mathrm{d}x)$ in $\mathbb{R}^+ \times \mathbb{X}$ is said to be regularly varying if the tail integral $U^\star(y) = \int_y^\infty \rho^\star(\mathrm{d}s)$ is a regularly varying function, i.e. it satisfies*

$$U^\star(y) = L\left(\frac{1}{y}\right)\frac{1}{y^\sigma}$$

for some $\sigma \in [0, 1)$ which we call the index and L a slowly varying function.

The following Theorem highlights an interesting link between the directing Lévy measure and the marginal Lévy intensities in terms of the regularly varying property.

Theorem 3. *Consider a CoRM with an homogeneous directing Lévy measure $\rho^\star(\mathrm{d}s)\alpha(\mathrm{d}x)$ such that the conditions of Theorem 2 are satisfied. If $\rho^\star$ is regularly varying with tail integral U then the marginals ρ_j, $j = 1, \ldots, d$, are regularly varying.*

Example 3: σ-stable directing Lévy measure

Consider a σ-stable directing Lévy measure

$$\nu^\star(\mathrm{d}s, \mathrm{d}x) = \frac{\sigma}{\Gamma(1-\sigma)s^{\sigma+1}}\mathrm{d}s\mathrm{d}x.$$

The related tail integral is

$$U^\star(y) = \frac{1}{\Gamma(1-\sigma)y^\sigma}$$

which is a regularly varying function with index σ and slowly varying function $L(y) = \frac{1}{\Gamma(1-\sigma)}$. We see that the regularly varying tail integrals related to the CoRM construction arise as a factor of U, namely $U_j(y) = \mathbb{E}\left[S_j^\sigma\right] U^\star(y)$ and from Theorem 3 the associated marginal tail integrals are regularly varying.

4.2. Independent Exponential scores

Consider a d-variate CoRM given by an homogeneous directing Lévy measure $\nu^\star(\mathrm{d}z, \mathrm{d}x) = \rho^\star(\mathrm{d}z)\alpha(\mathrm{d}x)$ and a score distribution corresponding to d independent standard exponential distributions, i.e.

$$h(s_1, \ldots, s_d) = \prod_{i=1}^{d} e^{-s_i}.$$

We observe that each associated marginal takes the form $\nu_j(\mathrm{d}s, \mathrm{d}x) = f(s)\mathrm{d}s\alpha(\mathrm{d}x)$, where $f(s) = \int_0^\infty z^{-1}e^{-\frac{s}{z}}\rho^\star(\mathrm{d}z)$. The following Theorem provides a characterization for this class of CoRM's.

Theorem 4. *Consider a CoRM as described above; the corresponding d-variate Lévy intensity* $\tilde{\nu}_d(\mathrm{d}\boldsymbol{s}, \mathrm{d}x) = \tilde{\rho}_d(\boldsymbol{s})\mathrm{d}\boldsymbol{s}\alpha(\mathrm{d}x)$ *is such that*

$$\tilde{\rho}_d(\boldsymbol{s}) = (-1)^{d-1}\frac{\partial^{d-1}}{\partial s^{d-1}} f(s)\Big|_{s=s_1+\ldots+s_d}.$$

5. Conclusions

In this paper, we proved some integrability condition for Compound Random Measures. The new findings can be useful to Statisticians which aim to use vectors of dependent completely random measures which arise from CoRM's with directing Lévy measure and score distribution that have not been considered so far. Furthermore, in the homogeneous case, we proved that the marginal Lévy intensities are regularly varying whenever a regularly varying directing Lévy measure is considered. Finally, we provide a representation of the homogeneous CoRM's when the score distribution is the result of independent and identically distributed exponential distributions.

Acknowledgments

Fabrizio Leisen was supported by the European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement no: 630677.

References

Camerlenghi, F., Lijoi, A. and Prünster, I. (2017). 'Bayesian prediction with multiple-samples information', Journal of Multivariate Analysis, **156**, 18-28.

Caron, F. and Fox, E. (2017). 'Sparse graphs using exchangeable random measures', forthcoming in the Journal of the Royal Statistical Society-Series B.

Cont, R., and Tankov, P., (2004). *Financial modelling with jump processes*, Boca Raton, FL: Chapman & Hall/CRC.

Daley, D.J., and Vere-Jones, D., (2003), *An introduction to the theory of point processes. Vol. 1.*, New York: Springer.

De Finetti, B. (1938). 'Sur la condition de "equivalence partielle"', Colloque consacré à la théorie des probabilités, Vol. VI, Université de Genève, Hermann et C. ie, Paris.

Griffin J. E. and Leisen F. (2017). 'Compound random measures and their use in Bayesian nonparametrics', *Journal of the Royal Statistical Society - Series B*, **79**, 525-545.

Griffin J. E. and Leisen F. (2018). 'Modelling and computation using NCoRM mixtures for density regression ', forthcoming in *Bayesian Analysis*.

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.

Kingman, J.F.C., (1993), *Poisson processes*, Oxford: Oxford University Press.

Leisen, F., and Lijoi, A., (2011). 'Vectors of Poisson-Dirichlet processes', *J. Multivariate Anal.*, **102**, 482–495.

Leisen, F., Lijoi, A., and Spano, D., (2013). 'A Vector of Dirichlet processes', *Electronic Journal of Statistics*, **7**, 62–90.

Lijoi, A. and Nipoti, B. (2014). 'A class of hazard rate mixtures for combining survival data from different experiments', *Journal of the American Statistical Association*, **109**, 802-814.

Lijoi, A., Nipoti, B., and Prünster, I. (2014). 'Bayesian inference with dependent normalized completely random measures', *Bernoulli*, **20**, 1260-1291.

A. Lijoi and I. Prünster (2010). Beyond the Dirichlet process. *Bayesian Nonparametrics* (Holmes, C.C., Hjort, N.L., Müller, P. and Walker, S.G., Eds.), 80–136, Cambridge University Press, Cambridge.

MacEachern, S.N., (1999). 'Dependent nonparametric processes'. *ASA Proceedings of the Section on Bayesian Statistical Science*, 50–55.

E. Regazzini, A. Lijoi, I. Prünster (2003). Distributional results for means of normalized random measures with independent increments. *Ann. Statist.* **31**, 560–585.

Todeschini A., Miscouridou X. and Caron F. (2017). 'Exchangeable random measures for sparse and modular graphs with overlapping communities', *arXiv: 1602.02114v2* .

Zhu W. and Leisen F. (2015). 'A multivariate extension of a vector of Poisson-Dirichlet processes', *Journal of Nonparametric Statistics*, **27**, 89–105.

6. Supplemetary Material: Proofs

First, we recall the integrability condition for a Lévy measure:

$$\int_{\mathbb{R}^+ \times X} \min\{1, s\} \tilde{\nu}(\mathrm{d}s, \mathrm{d}x) < \infty \tag{12}$$

see Section 2 of the main document.

Let H and $\nu^\star$ be, respectively, a score distribution and a directing Lévy measure which define a CoRM. We denote with H_j, $j \in \{1, \dots, \mathrm{d}\}$, the j-th marginal of a d-dimensional score distribution H. A simple change of variable leads to the j-th marginal of a CoRM, namely

$$\nu_j(A, X) = \int_{\mathbb{R}^+ \times X} \int_{A/z} H_j(\mathrm{d}s) \nu^\star(\mathrm{d}z, \mathrm{d}x) = \int_{\mathbb{R}^+} \int_{A \times X} \nu^\star\left(\frac{\mathrm{d}z}{s}, \mathrm{d}x\right) H_j(\mathrm{d}s). \tag{13}$$

We can see the formula above as a mean. Let S_j be a random variable with distribution H_j, then

$$\nu_j(A, X) = \mathbb{E}\left[\nu^\star\left(\frac{A}{S_j}, X\right)\right] \tag{14}$$

for $A \in \mathcal{B}(\mathbb{R}^+)$. We use the last identity to give conditions for the marginal intensity ν_j to be a proper Lévy intensity, i.e. a measure that satisfies the condition displayed in equation (12). As stated at the beginning of this section, the next theorem (and its corollary) provides conditions on the score distribution and the directing Lévy measure for the existence of the marginal Lévy intensities of a CoRM.

Proof of Theorem 1

We recall the conditions stated in the first part of the theorem:

$$\int_{(0,1) \times X} \mathbb{P}\left[S_j \geq \frac{1}{z}\right] \nu^\star(\mathrm{d}z, \mathrm{d}x) < \infty \tag{15}$$

and

$$\int_{[1,\infty) \times X} \mathbb{P}\left[S_j \leq \frac{1}{z}\right] z \nu^\star(\mathrm{d}z, \mathrm{d}x) < \infty, \tag{16}$$

and the conditions stated in the second part of the Theorem:

$$1 - H_j\left(\frac{1}{z}\right) \leq z \;\forall\; z \in (0, \epsilon) \text{ for some } \epsilon > 0 \tag{17}$$

and

$$\lim_{z \to \infty} z H_j\left(\frac{1}{z}\right) < \infty. \tag{18}$$

Proof. We recall that $\nu^\star$ satisfies (12) since it is a Lévy intensity. Using (14), condition (12) for ν_j becomes

$$\begin{aligned}&\mathbb{E}\left[\int_{\mathbb{R}^+\times X}\min\{1,z\}\nu^\star\left(\frac{\mathrm{d}z}{S_j},X\right)\right]\\&=\mathbb{E}\left[\int_{\left(0,\frac{1}{S_j}\right)\times X}z\nu^\star(\mathrm{d}z,\mathrm{d}x)\right]+\mathbb{E}\left[\int_{\left[\frac{1}{S_j},\infty\right)\times X}\nu^\star(\mathrm{d}z,\mathrm{d}x)\right]<\infty.\end{aligned}\tag{19}$$

Therefore, ν_j satisfies (12) if and only if

$$\mathbb{E}\left[\int_{\left(0,\frac{1}{S_j}\right)\times X}z\nu^\star(\mathrm{d}z,\mathrm{d}x)\right]<\infty\tag{20}$$

and

$$\mathbb{E}\left[\int_{\left[\frac{1}{S_j},\infty\right)\times X}\nu^\star(\mathrm{d}z,\mathrm{d}x)\right]<\infty.\tag{21}$$

The former can be decomposed using the Fubini-Tonelli theorem in

$$\begin{aligned}&\mathbb{E}\left[\int_{\left(0,\frac{1}{S_j}\right)\times X}z\nu^\star(\mathrm{d}z,\mathrm{d}x)\right]=\int_{\mathbb{R}^+\times X}\mathbb{P}\left[S_j\le\frac{1}{z}\right]z\nu^\star(\mathrm{d}z,\mathrm{d}x)\\&=\int_{(0,1)\times X}\mathbb{P}\left[S_j\le\frac{1}{z}\right]z\nu^\star(\mathrm{d}z,\mathrm{d}x)+\int_{[1,\infty)\times X}\mathbb{P}\left[S_j\le\frac{1}{z}\right]z\nu^\star(\mathrm{d}z,\mathrm{d}x).\end{aligned}$$

Condition (16) ensures that the second term of the above equation is finite. It is easy to see that the first term is finite as well. Indeed,

$$\int_{(0,1)\times X}\mathbb{P}\left[S_j\le\frac{1}{z}\right]z\nu^\star(\mathrm{d}z,\mathrm{d}x)\le\int_{(0,1)\times X}z\nu^\star(\mathrm{d}z,\mathrm{d}x)<\infty.$$

On the other hand, the second term in (19) can be decomposed in

$$\begin{aligned}&\mathbb{E}\left[\int_{\left[\frac{1}{S_j},\infty\right)\times X}\nu^\star(\mathrm{d}z,\mathrm{d}x)\right]=\int_{\mathbb{R}^+\times X}\mathbb{P}\left[\frac{1}{z}\le S_j\right]\nu^\star(\mathrm{d}z,\mathrm{d}x)\\&=\int_{(0,1)\times X}\mathbb{P}\left[S_j\ge\frac{1}{z}\right]\nu^\star(\mathrm{d}z,\mathrm{d}x)+\int_{[1,\infty)\times X}\mathbb{P}\left[S_j\ge\frac{1}{z}\right]\nu^\star(\mathrm{d}z,\mathrm{d}x).\end{aligned}$$

Condition (15) ensures that the first term of the above equation is finite. It is easy to see that the second

term is finite as well. Indeed,

$$\int_{[1,\infty)\times X} \mathbb{P}\left[S_j \geq \frac{1}{z}\right] \nu^\star(\mathrm{d}z, \mathrm{d}x) \leq \int_{[1,\infty)\times X} \nu^\star(\mathrm{d}z, \mathrm{d}x) < \infty.$$

Therefore, the first part of the theorem follows from (19), (20) and (21).

For the remaining part of the Theorem we use that (12) is attained when considering the directing Lévy measure $\nu^\star$. Indeed, if

$$\lim_{z\to\infty} z\mathbb{P}\left[S_j \leq \frac{1}{z}\right] < \infty \tag{22}$$

then as $\nu^\star$ is a Lévy intensity

$$\int_{[1,\infty)\times X} \mathbb{P}\left[S_j \leq \frac{1}{z}\right] z\nu^\star(\mathrm{d}z, \mathrm{d}x) < \infty. \tag{23}$$

so (20) holds. And if there exists $\epsilon > 0$ such that $1 - H_j\left(\frac{1}{z}\right) \leq z \;\forall\; z \in (0, \epsilon)$ then

$$\int_{(0,1)\times X} \mathbb{P}\left[\frac{1}{z} \leq S_j\right] \nu^\star(\mathrm{d}z, \mathrm{d}x) < \int_{(0,1)\times X} z\nu^\star(\mathrm{d}z, \mathrm{d}x) < \infty,$$

so (21) also holds. From the first part of the theorem the CoRM marginal ν_j satisfies the integrability conditions for arbitrary $\nu^\star$. □

Proof of Corollary 1

We recall the conditions stated in Corollary 1:

$$\lim_{z\to 0} \frac{h_j\left(\frac{1}{z}\right)}{z^2} < 1 \tag{24}$$

and

$$\lim_{\epsilon\to 0} h_j(\epsilon) < \infty. \tag{25}$$

Proof. We define $f(z) = z - (1 - H_j\left(\frac{1}{z}\right))$ and observe that $f(0^+) = 0$ so the existence of $f'(0^+) > 0$ implies (17). As S_j has a probability density we get that $f'(0^+)$ exists and (17) is equivalent to $f'(0^+) > 0$ which we

write as

$$\lim_{z\to 0}\frac{h_j\left(\frac{1}{z}\right)}{z^2}<1.$$

Using the fundamental theorem of calculus we see that (18) reduces to

$$\lim_{z\to\infty} z\mathbb{P}\left[S_j\leq\frac{1}{z}\right]=\lim_{\epsilon\to 0}h_j(\epsilon)<\infty$$

which is satisfied when h_j is continuous at zero. □

6.1. Proof of Theorem 2

Proof. Denote $P_j=\{\boldsymbol{s}\in(\mathbb{R}^+)^d\,:\,\max\{s_1,\dots,s_d\}=s_j\}$ for $j\in\{1,\dots,\mathrm{d}\}$; then, by using (13) and the fact that each ν_j is a Lévy intensity we get that for any bounded set X in $\mathcal{X}$

$$\begin{aligned}
&\int_{\mathbb{R}^+\times X}\int_{(\mathbb{R}^+)^d}\min\{1,\|\boldsymbol{s}\|\}h\left(\frac{s_1}{z},\dots,\frac{s_d}{z}\right)\frac{\mathrm{d}\boldsymbol{s}}{z^d}\nu^\star(\mathrm{dz},\mathrm{d}x)\\
&=\sum_{j=1}^d\int_{\mathbb{R}^+\times X}\int_{P_j}\min\{1,\|\boldsymbol{s}\|\}h\left(\frac{s_1}{z},\dots,\frac{s_d}{z}\right)\frac{\mathrm{d}\boldsymbol{s}}{z^d}\nu^\star(\mathrm{dz},\mathrm{d}x)\\
&\leq\sum_{j=1}^d\int_{\mathbb{R}^+\times X}\int_{P_j}\min\{1,\sqrt{d}s_j\}h\left(\frac{s_1}{z},\dots,\frac{s_d}{z}\right)\frac{\mathrm{d}\boldsymbol{s}}{z^d}\nu^\star(\mathrm{dz},\mathrm{d}x)\\
&\leq\sum_{j=1}^d\int_{\mathbb{R}^+\times X}\int_{(\mathbb{R}^+)^d}\min\{1,\sqrt{d}s_j\}h\left(\frac{s_1}{z},\dots,\frac{s_d}{z}\right)\frac{\mathrm{d}\boldsymbol{s}}{z^d}\nu^\star(\mathrm{dz},\mathrm{d}x)\\
&=\sum_{j=1}^d\int_{\mathbb{R}^+\times X}\int_{\mathbb{R}^+}\min\{1,\sqrt{d}s\}h_j\left(\frac{s}{z}\right)\frac{\mathrm{d}s}{z}\nu^\star(\mathrm{dz},\mathrm{d}x)\\
&=\sum_{j=1}^d\int_{\mathbb{R}^+\times X}\min\{1,\sqrt{d}s\}\nu_j(\mathrm{d}s,\mathrm{d}x)\\
&\leq\sum_{j=1}^d\int_{\mathbb{R}^+\times X}\min\{\sqrt{d},\sqrt{d}s\}\nu_j(\mathrm{d}s,\mathrm{d}x)\\
&=\sqrt{d}\sum_{j=1}^d\int_{\mathbb{R}^+\times X}\min\{1,s\}\nu_j(\mathrm{d}s,\mathrm{d}x)<\infty.
\end{aligned}$$

□

Proof of Theorem 3

We recall that for the case at hand

$$U^\star(y) = L\left(\frac{1}{y}\right)\frac{1}{y^\sigma} \tag{26}$$

is a tail integral.

Proof. We note that equation (14) implies that

$$\rho_j(A) = \mathbb{E}\left[\rho^\star\left(\frac{A}{S_j}\right)\right]$$

It follows that the marginals of the CoRM are given by

$$\begin{aligned} U_j(y) &= \rho_j\left((y,\infty)\right) \\ &= \mathbb{E}\left[U^\star\left(\frac{y}{S_j}\right)\right] \\ &= \mathbb{E}\left[L\left(\frac{S_j}{y}\right)\left(\frac{S_j}{y}\right)^\sigma\right] \\ &= \mathbb{E}\left[L\left(\frac{S_j}{y}\right)S_j^\sigma\right]\frac{1}{y^\sigma}. \end{aligned}$$

Hence, it is enough to check if the function $l(z) = \mathbb{E}\left[L(S_j z)S_j^\sigma\right]$ is slowly varying for L a slowly varying function. Let $a > 0$, we need to check

$$\lim_{t\to\infty}\frac{l(at)}{l(t)} = \lim_{t\to\infty}\frac{\mathbb{E}\left[L(atS_j)S_j^\sigma\right]}{\mathbb{E}\left[L(tS_j)S_j^\sigma\right]} = 1.$$

For a fixed $\epsilon > 0$ we can choose t_0 such that $\forall u > t_0$

$$|L(au)/L(u) - 1| < \frac{\epsilon}{2},$$

since L is slowly varying. Then for $t > t_0$

$$\begin{aligned}
&\left|\frac{\mathbb{E}\left[L(atS_j)S_j^\sigma\right]}{\mathbb{E}\left[L(tS_j)S_j^\sigma\right]} - 1\right| = \left|\frac{\mathbb{E}\left[S_j^\sigma\left(L(atS_j) - L(tS_j)\right)\right]}{\mathbb{E}\left[L(tS_j)S_j^\sigma\right]}\right| \\
&\leq \frac{\mathbb{E}\left[\mathbb{1}_{\left\{S_j > \frac{t_0}{t}\right\}} S_j^\sigma \left|L(atS_j) - L(tS_j)\right|\right]}{\mathbb{E}\left[L(tS_j)S_j^\sigma\right]} + \frac{\mathbb{E}\left[\mathbb{1}_{\left\{S_j \leq \frac{t_0}{t}\right\}} S_j^\sigma \left|L(atS_j) - L(tS_j)\right|\right]}{\mathbb{E}\left[L(tS_j)S_j^\sigma\right]} \\
&< \frac{\mathbb{E}\left[\mathbb{1}_{\left\{S_j > \frac{t_0}{t}\right\}} S_j^\sigma \frac{\epsilon}{2} L(tS_j)\right]}{\mathbb{E}\left[L(tS_j)S_j^\sigma\right]} + \frac{\mathbb{E}\left[\mathbb{1}_{\left\{S_j \leq \frac{t_0}{t}\right\}} S_j^\sigma \left|L(atS_j) - L(tS_j)\right|\right]}{\mathbb{E}\left[L(tS_j)S_j^\sigma\right]} \\
&< \frac{\epsilon}{2} + \frac{\mathbb{E}\left[\mathbb{1}_{\left\{S_j \leq \frac{t_0}{t}\right\}} S_j^\sigma \left|L(atS_j) - L(tS_j)\right|\right]}{\mathbb{E}\left[L(tS_j)S_j^\sigma\right]} \\
&= \frac{\epsilon}{2} + \frac{\mathbb{E}\left[\mathbb{1}_{\left\{S_j \leq \frac{t_0}{t}\right\}} S_j^\sigma t^\sigma \left|L(atS_j) - L(tS_j)\right|\right]}{U_j(\frac{1}{t})} \\
&= \frac{\epsilon}{2} + \frac{\int_{(0,\frac{t_0}{t}]} s^\sigma t^\sigma \left|L(ats) - L(ts)\right| H_j(\mathrm{d}s)}{U_j(\frac{1}{t})} \\
&= \frac{\epsilon}{2} + \frac{\int_{(0,t_0]} u^\sigma \left|L(au) - L(u)\right| H_j\left(\frac{\mathrm{d}u}{t}\right)}{U_j(\frac{1}{t})} \qquad (27)
\end{aligned}$$

We observe that $\lim_{x\to 0} U_j(x) = \infty$. since U_j is a tail integral. From (26) it follows that $\lim_{x\to 0} x^\sigma \left(L(ax) - L(x)\right) = 0$. Hence, the function $g(x) = x^\sigma \left(L(ax) - L(x)\right)$ is bounded in $[0, t_0]$ by a constant K_{1,t_0}. Finally we observe that for $t > t_0$

$$\int_{(0,t_0]} H_j\left(\frac{\mathrm{d}u}{t}\right) < \int_{(0,1]} H_j\left(\mathrm{d}u\right) \leq 1.$$

We set $t_1 > t_0$ such that for $u > t_1$

$$\frac{2K_{1,t_0}}{\epsilon} < U_j(1/u).$$

Choosing $t > t_1$ we get

$$\frac{\int_{(0,t_0]} u^\sigma \left|L(au) - L(u)\right| H_j\left(\frac{\mathrm{d}u}{t}\right)}{U_j(\frac{1}{t})} < \frac{K_{1,t_0}}{U_j(\frac{1}{t})} < \frac{\epsilon}{2}$$

It follows from (27) that

$$\left|\frac{\mathbb{E}\left[L(atS_j)S_j^\sigma\right]}{\mathbb{E}\left[L(tS_j)S_j^\sigma\right]} - 1\right| < \epsilon.$$

Consequently, l defined above is slowly varying, implying that the marginal tail integral U_j is regularly varying. □

6.2. Proof of Theorem 4

Proof. It is straightforward to see that

$$\tilde{\rho}_d(\boldsymbol{s}) = \int_0^\infty z^{-d} h\left(\frac{s_1}{z}, \cdots, \frac{s_d}{z}\right) \rho^\star(\mathrm{d}z) = \int_0^\infty z^{-d} e^{-\frac{s_1+\ldots+s_d}{z}} \rho^\star(\mathrm{d}z).$$

From Example 1 in the main document, we know that, for arbitrary $\rho^\star$ and $d \in \mathbb{N} \setminus \{0\}$, the previous integral is finite. Therefore for $s \neq 0$

$$\int_0^\infty \left| \frac{\partial^j}{\partial s^j} z^{-1} e^{-\frac{s}{z}} \right| \rho^\star(\mathrm{d}z) = \tilde{\rho}_{j+1}(s, 0, \ldots, 0) < \infty$$

and this concludes the proof since using the Dominated Convergence Theorem we can take the derivative under the integral sign as follows

$$\begin{aligned}(-1)^{d-1} \frac{\partial^{d-1}}{\partial s^{d-1}} f(s)\Big|_{s=s_1+\ldots+s_d} &= (-1)^{d-1} \int_0^\infty \frac{\partial^{d-1}}{\partial s^{d-1}} \left(z^{-1} e^{-\frac{s}{z}}\right) \rho^\star(\mathrm{d}z)\Big|_{s=s_1+\ldots+s_d} \\ &= \int_0^\infty z^{-d} e^{-\frac{s_1+\ldots+s_d}{z}} \rho^\star(\mathrm{d}z).\end{aligned}$$

□