Submitted to the Annals of Statistics

DISTRIBUTION THEORY FOR HIERARCHICAL PROCESSES: SUPPLEMENTARY MATERIAL

By Federico Camerlenghi*, Antonio Lijoi*,† Peter Orbanz and Igor Prünster†

University of Milano–Bicocca, Bocconi University, Columbia University and Bocconi University

This appendix displays all detailed proofs of the theorems in the main manuscript "Distribution theory for hierarchical processes" and a specialization of the Blackwell-MacQueen Pólya urn scheme to the hierarchical Pitman–Yor process.

APPENDIX A: PROOFS

A.1. Proof of Theorem 1. For the sake of notational simplicity set

$$\mathcal{I}_q = \int_0^\infty u \mathrm{e}^{-c\psi(u)} \tau_q(u) \, \mathrm{d}u, \qquad \mathcal{I}_{q,0} = \int_0^\infty u \mathrm{e}^{-c_0\psi_0(u)} \tau_{q,0}(u) \, \mathrm{d}u.$$

From [2, Proposition 1] one has $\mathrm{Var}(\tilde{p}_0(A)) = c_0 P_0(A)(1 - P_0(A)) \, \mathcal{I}_{2,0}$ and $\mathrm{Var}(\tilde{p}_i(A) \,|\, \tilde{p}_0) = c\tilde{p}_0(A)(1 - \tilde{p}_0(A)) \, \mathcal{I}_2$. Hence, for any $i = 1, \ldots, d$,

$$\begin{aligned}\mathrm{Var}(\tilde{p}_i(A)) &= \mathbb{E}\,\mathrm{Var}(\tilde{p}_i(A) \,|\, \tilde{p}_0) + \mathrm{Var}(\tilde{p}_0(A)) \\ &= c_0 P_0(A)(1 - P_0(A))\{cc_0 \, \mathcal{I}_2 \mathcal{I}_{1,0} + \mathcal{I}_{2,0}\}.\end{aligned}$$

Moreover, for any $i \neq j$,

$$\begin{aligned}\mathrm{Cov}(\tilde{p}_i(A), \tilde{p}_j(A)) &= \mathbb{E}\,\mathbb{E}[\tilde{p}_i(A)\tilde{p}_j(A)|\tilde{p}_0] - (\mathbb{E}\,\mathbb{E}[\tilde{p}_i(A)|\tilde{p}_0])\,(\mathbb{E}\,\mathbb{E}[\tilde{p}_j(A)|\tilde{p}_0]) \\ &= \mathbb{E}[\mathbb{E}[\tilde{p}_i(A)|\tilde{p}_0]\mathbb{E}[\tilde{p}_j(A)|\tilde{p}_0]] - (\mathbb{E}[\tilde{p}_0(A)])^2 \\ &= \mathbb{E}[\tilde{p}_0^2(A)] - (\mathbb{E}[\tilde{p}_0(A)])^2 = \mathrm{Var}(\tilde{p}_0(A)),\end{aligned}$$

and the result follows. □

A.2. Proof of Theorem 2. The same line of reasoning in the proof of Theorem 1 may be used to establish that $\mathrm{Cov}(\tilde{p}_i(A), \tilde{p}_j(A)) = \mathrm{Var}(\tilde{p}_0(A))$. Moreover, by the definition of Pitman–Yor process in Section 2, one has

$$\mathrm{Var}(\tilde{p}_0(A)) = P_0(A)(1 - P_0(A)) \, \frac{\sigma_0^2(1-\sigma_0)}{\theta_0(\theta_0+1)\Gamma(\theta_0/\sigma_0)} \int_0^\infty u^{\theta_0+\sigma_0-1} \mathrm{e}^{-u^{\sigma_0}} \, \mathrm{d}u$$

*Also affiliated to Collegio Carlo Alberto, Moncalieri, Italy.

†Supported by the European Research Council (ERC) through StG "N-BNP" 306406

$$= P_0(A)(1 - P_0(A))\,\frac{1-\sigma_0}{\theta_0+1}$$

Furthermore, $\mathrm{Var}(\tilde p_i(A)) = \mathbb{E}\,\mathrm{Var}(\tilde p_i(A)\,|\,\tilde p_0) + \mathrm{Var}(\tilde p_0(A))$ holds and one obtains

$$\begin{aligned}\mathrm{Var}(\tilde p_i(A)) &= \frac{1-\sigma}{\theta+1}\,\mathbb{E}\,\tilde p_0(A)(1-\tilde p_0(A)) + \frac{1-\sigma_0}{\theta_0+1}P_0(A)(1-P_0(A))\\ &= \frac{P_0(A)(1-P_0(A))}{\theta_0+1}\left\{(1-\sigma_0) + (\theta_0+\sigma_0)\frac{1-\sigma}{\theta+1}\right\},\end{aligned}$$

from which (10) easily follows. □

A.3. Proof of Theorem 3. In order to prove Theorem 3, we first need to display a useful technical lemma, which follows from a formula by Faà di Bruno. See, e.g., [1]. According to it, the n-th order derivative of $\mathrm{e}^{-m\psi(u)}$, with $\psi(u) = \int_0^\infty [1-\mathrm{e}^{-uv}]\,\rho(v)\,\mathrm{d}v$, is given by

$$\frac{\mathrm{d}^n}{\mathrm{d}u^n}\mathrm{e}^{-m\psi(u)} = \sum_{\pi} \frac{\mathrm{d}^{|\pi|}}{\mathrm{d}x^{|\pi|}}(\mathrm{e}^{-mx})\Big|_{x=\psi(u)} \prod_{B\in\pi} \frac{\mathrm{d}^{|B|}}{\mathrm{d}u^{|B|}}\psi(u) \tag{A.1}$$

where $m \in \mathbb{R}$, and the sum is extended over all partitions π of $[n] = \{1,\dots,n\}$. Clearly (A.1) can be rewritten as

$$\frac{\mathrm{d}^n}{\mathrm{d}u^n}\mathrm{e}^{-m\psi(u)} = \sum_{i=1}^n (-m)^i \mathrm{e}^{-m\psi(u)} \sum_{\pi:|\pi|=i}\prod_{B\in\pi}\frac{\mathrm{d}^{|B|}}{\mathrm{d}u^{|B|}}\psi(u).$$

Since to each unordered partition π of size i there correspond $i!$ ordered partitions of the set $[n]$ into i components, which are obtained by permuting the elements of π in all the possible ways, one has

$$\begin{aligned}\frac{\mathrm{d}^n}{\mathrm{d}u^n}\mathrm{e}^{-m\psi(u)} &= \sum_{i=1}^n (-m)^i \mathrm{e}^{-m\psi(u)}\frac{1}{i!}\sum_{\pi:|\pi|=i}^{\mathrm{o}}\ \prod_{B\in\pi}\frac{\mathrm{d}^{|B|}}{\mathrm{d}u^{|B|}}\psi(u)\\ &= \sum_{i=1}^n (-m)^i \mathrm{e}^{-m\psi(u)}\frac{1}{i!}\sum_{(*)}\binom{n}{q_1,\cdots,q_i}\frac{\mathrm{d}^{q_1}}{\mathrm{d}u^{q_1}}\psi(u)\cdots\frac{\mathrm{d}^{q_i}}{\mathrm{d}u^{q_i}}\psi(u)\end{aligned}$$

where $\sum^{\mathrm{o}}$ is the sum over the ordered partitions of the set $[n]$ while the sum $(*)$ runs over all vectors $(q_1,\cdots,q_i)$ of positive integers such that $\sum_{j=1}^i q_j = n$. The second equality follows upon noting that the derivative of ψ depends only on the number of elements within each component of the partition π and that the number of partitions π of the set $[n]$ containing i elements $(B_1,\ \cdots, B_i)$, with $(|B_1|,\cdots,|B_i|) = (q_1,\cdots,q_i)$, equals the multinomial coefficient above. In view of these remarks, one has

LEMMA 1. *If* $\tau_q(u) = \int_0^\infty v^q\, \mathrm{e}^{-uv}\, \rho(v)\,\mathrm{d}v$ *and*

$$\xi_{n,i}(u) = \sum_{(*)} \frac{1}{i!} \binom{n}{q_1, \cdots, q_i} \tau_{q_1}(u) \cdots \tau_{q_i}(u), \tag{A.2}$$

the following relation holds

$$(-1)^n \frac{\mathrm{d}^n}{\mathrm{d}u^n} e^{-m\psi(u)} = e^{-m\psi(u)} \sum_{i=1}^n m^i \xi_{n,i}(u) \tag{A.3}$$

In view of this, for any $x_1 \neq \cdots \neq x_k$ in $\mathbb{X}$ and $k \geq 1$, one can evaluate

$$M_{\boldsymbol{n}_1,\ldots,\boldsymbol{n}_d}(\mathrm{d}x_1, \ldots, \mathrm{d}x_k) = \mathbb{E} \prod_{j=1}^k \prod_{i=1}^d \tilde{p}_i^{n_{i,j}}(\mathrm{d}x_j)$$

where $\boldsymbol{n}_i = (n_{i,1}, \ldots, n_{i,k})$, for $i = 1, \ldots, d$, and it is apparent that there may be $n_{i,j}$'s equal to zero, though $\sum_{i=1}^d n_{i,j} \geq 1$. Indeed, if one sets $A_{j,\varepsilon} = B(x_j;\varepsilon)$ a ball of radius ε around x_j, with $\varepsilon > 0$ small enough so that $A_{i,\varepsilon} \cap A_{j,\varepsilon} = \varnothing$ for any $i \neq j$, then

$$\begin{aligned}
M_{\boldsymbol{n}_1,\ldots,\boldsymbol{n}_d}(A_{1,\varepsilon} \times \cdots \times A_{k,\varepsilon}) &= \mathbb{E} \prod_{i=1}^d \mathbb{E}\Big[\prod_{j=1}^k \tilde{p}_i^{n_{i,j}}(A_{j,\varepsilon}) \Big| \tilde{p}_0 \Big] \\
&= \mathbb{E} \prod_{i=1}^d \frac{1}{\Gamma(N_i)} \int_0^\infty u^{N_i - 1} \mathbb{E}[\mathrm{e}^{-u\tilde{\mu}_i(\mathbb{X}_\varepsilon^*)} | \tilde{p}_0] \\
&\qquad \times \prod_{j=1}^k \mathbb{E}[\mathrm{e}^{-u\tilde{\mu}_i(A_{j,\varepsilon})} \tilde{\mu}_i^{n_{i,j}}(A_{j,\varepsilon}) | \tilde{p}_0]\,\mathrm{d}u \\
&= \mathbb{E} \prod_{i=1}^d \frac{1}{\Gamma(N_i)} \int_0^\infty u^{N_i - 1} \mathrm{e}^{-c\psi(u)\tilde{p}_0(\mathbb{X}_\varepsilon^*)} \\
&\qquad \times \prod_{j=1}^k \Big((-1)^{n_{i,j}} \frac{\mathrm{d}^{n_{i,j}}}{\mathrm{d}u^{n_{i,j}}} e^{-c\psi(u)\tilde{p}_0(A_{j,\varepsilon})} \Big)\,\mathrm{d}u
\end{aligned}$$

where $\mathbb{X}_\varepsilon^* = \mathbb{X} \backslash (\cup_{j=1}^k A_{j,\varepsilon})$. If $\boldsymbol{\ell} = (\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_d)$ where each $\boldsymbol{\ell}_i = (\ell_{i,1}, \ldots, \ell_{i,k}) \in \times_{j=1}^k \{1, \ldots, n_{i,j}\}$, by virtue of (A.3) one obtains

$$\begin{aligned}
&M_{\boldsymbol{n}_1,\ldots,\boldsymbol{n}_d}(A_{1,\varepsilon} \times \cdots \times A_{k,\varepsilon}) \\
&\quad = \mathbb{E} \prod_{i=1}^d \frac{1}{\Gamma(N_i)} \int_0^\infty u^{N_i - 1}\, \mathrm{e}^{-c\,\psi(u)} \prod_{j=1}^k \sum_{\ell_{i,j}=1}^{n_{i,j}} c^{\ell_{i,j}} \tilde{p}_0^{\ell_{i,j}}(A_{j\varepsilon})\, \xi_{n_{i,j},\ell_{i,j}}(u)\,\mathrm{d}u
\end{aligned}$$

$$= \sum_{\boldsymbol{\ell}} \mathbb{E} \prod_{j=1}^{k} \tilde{p}_0^{\bar{\ell}_{\bullet j}}(A_{j,\varepsilon}) \prod_{i=1}^{d} \Big(\frac{c^{\bar{\ell}_{i\bullet}}}{\Gamma(N_i)} \int_0^\infty u^{N_i-1} e^{-c\psi(u)} \prod_{j=1}^{k} \xi_{n_{i,j},\ell_{i,j}}(u) \mathrm{d}u \Big)$$

and we further agree that, whenever $n_{i,j} = 0$, we have $\ell_{i,j} = 0$ and $\sum_{\ell_{i,j}=1}^{0} \equiv 1$. If we set

$$M^0_{\bar{\ell}_{\bullet 1},\dots,\bar{\ell}_{\bullet k}}(A_{1,\varepsilon},\dots,A_{k,\varepsilon}) = \mathbb{E} \prod_{j=1}^{k} \tilde{p}_0^{\bar{\ell}_{\bullet j}}(A_{j,\varepsilon})$$

as $\varepsilon \to 0$ the non–atomicity of P_0 implies

$$\begin{aligned} M^0_{\bar{\ell}_{\bullet 1},\dots,\bar{\ell}_{\bullet k}}(\mathrm{d}x_1,\dots,\mathrm{d}x_k) &= \mathbb{E} \prod_{j=1}^{k} \tilde{p}_0^{\ell_j}(\mathrm{d}x_j) = \prod_{j=1}^{k} P_0(\mathrm{d}x_j) \\ &\quad \times \frac{c_0^k}{\Gamma(|\boldsymbol{\ell}|)} \int_0^\infty u^{|\boldsymbol{\ell}|-1} \, \mathrm{e}^{-c_0\psi_0(u)} \prod_{j=1}^{k} \tau_{\bar{\ell}_{\bullet j},0}(u) \, \mathrm{d}u \\ &= \Big(\prod_{j=1}^{k} P_0(\mathrm{d}x_j) \Big) \, \Phi^{(|\boldsymbol{\ell}|)}_{k,0}(\bar{\ell}_{\bullet 1},\dots,\bar{\ell}_{\bullet k}). \end{aligned}$$

Then

$$\begin{aligned} M_{\boldsymbol{n}_1,\dots,\boldsymbol{n}_d}(\mathrm{d}x_1,\dots,\mathrm{d}x_k) &= \Big(\prod_{j=1}^{k} P_0(\mathrm{d}x_j) \Big) \sum_{\boldsymbol{\ell}} \Phi^{(|\boldsymbol{\ell}|)}_{k,0}(\bar{\ell}_{\bullet 1},\dots,\bar{\ell}_{\bullet k}) \\ &\quad \times \prod_{i=1}^{d} \Big(\frac{c^{\bar{\ell}_{i\bullet}}}{\Gamma(N_i)} \int_0^\infty u^{N_i-1} e^{-c\psi(u)} \prod_{j=1}^{k} \xi_{n_{i,j},\ell_{i,j}}(u) \mathrm{d}u \Big) \end{aligned}$$

and the result follows by virtue of (A.2) and by noting that

$$\Pi_k^{(n)}(\boldsymbol{n}_1,\dots,\boldsymbol{n}_d) = \int_{\mathbb{X}^k} M_{\boldsymbol{n}_1,\dots,\boldsymbol{n}_d}(\mathrm{d}x_1,\dots,\mathrm{d}x_k).$$

□

A.4. Proof of Theorem 4. This can be deduced by working along the same lines as for the Proof of Theorem 3. One only needs to take into account the change of measure (8) in order to work directly with CRMs. Indeed, using the same notation, one has

$$M_{\boldsymbol{n}_1,\dots,\boldsymbol{n}_d}(A_{1,\varepsilon} \times \cdots \times A_{k,\varepsilon})$$

$$
\begin{aligned}
&= \frac{\Gamma^d(\theta+1)}{\Gamma^d\Big(\frac{\theta}{\sigma}+1\Big)}\,\mathbb{E}\,\prod_{i=1}^{d}\frac{1}{\Gamma(\theta+N_i)}\int_0^\infty u^{\theta+N_i-1}\,\mathrm{e}^{-\psi(u)}\\
&\qquad\times\prod_{j=1}^{k}\sum_{\ell_{i,j}=1}^{n_{i,j}}\tilde p_0^{\ell_{i,j}}(A_{j\varepsilon})\,\xi_{n_{i,j},\ell_{i,j}}(u)\,\mathrm{d}u\\
&= \frac{\Gamma^d(\theta+1)}{\Gamma^d\Big(\frac{\theta}{\sigma}+1\Big)}\sum_{\boldsymbol{\ell}}\Big(\mathbb{E}\prod_{j=1}^{k}\tilde p_0^{\bar\ell_{\bullet j}}(A_{j,\varepsilon})\Big)\\
&\qquad\times\prod_{i=1}^{d}\Big(\frac{1}{\Gamma(N_i)}\int_0^\infty u^{N_i-1}e^{-\psi(u)}\prod_{j=1}^{k}\xi_{n_{i,j},\ell_{i,j}}(u)\mathrm{d}u\Big)
\end{aligned}
$$

where $\psi(u)=u^\sigma$ and

$$
\begin{aligned}
\xi_{n_{i,j},\ell_{i,j}}(u) &= \frac{\sigma^{\ell_{i,j}}}{u^{n_{i,j}-\ell_{i,j}\sigma}\,\ell_{i,j}!}\sum_{\boldsymbol{q}}\binom{n_{i,j}}{q_{i,1}\ \cdots\ q_{i,\ell_{i,j}}}\prod_{r=1}^{\ell_{i,j}}(1-\sigma)_{q_{i,r}-1}\\
&= \frac{1}{u^{n_{i,j}-\ell_{i,j}\sigma}}\,\mathscr{C}(n_{i,j},\ell_{i,j};\sigma).
\end{aligned}
$$

The result now easily follows upon noting that, as $\varepsilon\downarrow 0$,

$$
\mathbb{E}\prod_{j=1}^{k}\tilde p_0^{\bar\ell_{\bullet j}}(A_{j,\varepsilon}) = \prod_{j=1}^{k}P_0(A_{j\varepsilon})\,\frac{\prod_{i=1}^{k-1}(\theta_0+i\sigma_0)}{(\theta_0+1)_{|\boldsymbol{\ell}|-1}}\,\prod_{j=1}^{k}(1-\sigma_0)_{\bar\ell_{\bullet j}-1}+\lambda_{k,\varepsilon}
$$

where $\lambda_{k,\varepsilon}=\mathrm{o}\Big(\prod_{j=1}^k P_0(A_{j,\varepsilon})\Big)$. □

A.5. Proof of Theorem 5. Let $\bar N_0\equiv 0$ and, for each $i=1,\dots,d$, $\bar N_i=\sum_{j=1}^i N_j$. Further suppose $\tilde\pi_1,\dots,\tilde\pi_d$ denote independent random partitions of $\mathbb{N}$ such that the restriction $\tilde\pi_{i,N_i}$ of $\tilde\pi_i$ to $[N_i]=\{\bar N_{i-1}+1,\dots,\bar N_i\}$ has probability distribution $\Phi_{\zeta_i,i}^{(N_i)}$ as defined in (13). Furthermore, $\tilde\pi_0$ is a random partition of $\mathbb{N}$ such that, conditional on $(\tilde\pi_{1,N_1},\dots,\tilde\pi_{d,N_d})$, its restriction $\tilde\pi_{0,h}$ to $[h]$ has probability distribution $\Phi_{k,0}^{(h)}$ in (11), where $h=\sum_{i=1}^d|\tilde\pi_{i,N_i}|$ with $|\tilde\pi_{i,N_i}|$ the number of blocks in the partition $\tilde\pi_{i,N_i}$. Relying on Theorem 3 we have

$$
\Pi_k^{(N)}(\boldsymbol{n}_1,\dots,\boldsymbol{n}_d)=\sum_{h=k}^{N}\prod_{i=1}^{d}\sum_{(*)}\mathbb{P}[\tilde\pi_{i,N_i}=\{B_{i,1},\dots,B_{i,\zeta_i}\}]
$$

$$\times \sum_{(**)} \mathbb{P}\Big[\tilde\pi_{0,h} = \{C_1, \dots, C_k\} \,\Big|\, \bigcap_{i=1}^{d} \Big\{\tilde\pi_{i,N_i} = \{B_{i,1}, \dots, B_{i,\zeta_i}\}\Big\}\Big]$$

where the sums are taken over all partitions such that $\sum_{i=1}^d \zeta_i = h$, for any $h \in \{k \dots, N\}$, and

$$\sum_{\{t:\, \bar\zeta_{i-1}+t\in C_j\}\cap\{1,\dots,\zeta_i\}} \mathrm{card}(B_{i,t}) = n_{i,j}$$

where $\bar\zeta_0 \equiv 0$ and $\bar\zeta_i = \sum_{r=1}^i \zeta_r$, for each $i = 1, \dots, d$. Note that one may have $\{t : \bar\zeta_{i-1} + t \in C_j\} \cap \{1, \dots, \zeta_i\} = \varnothing$, in which case $n_{i,j} = 0$, and in the Chinese restaurant franchise terminology it would mean that the j–th dish is not served in restaurant i. According to this notation, one has $K'_{i,N_i} = |\tilde\pi_{i,N_i}|$ and $K_{0,h} = |\tilde\pi_{0,h}|$ and

$$\mathbb{P}[K_N = k] = \mathbb{P}\Big[\bigcup_{t=k}^{N} \bigcup_{(t_1,\dots,t_d)\in\Delta_{d,t}} \{K'_{1,N_1} = t_1,\, \dots,\, K'_{d,N_d} = t_d,\, K_{0,t} = k\}\Big].$$

The result, then, follows from independence of $K'_{1,N_1}, \dots, K'_{d,N_d}$ and from the fact that

$$\mathbb{P}[K_{0,t} = k \,|\, K'_{1,N_1} = t_1, \dots, K'_{d,N_d} = t_d] = \mathbb{P}[K_{0,t} = k]\, \mathbb{1}_{\Delta_{d,t}}(t_1, \dots, t_d).$$

□

A.6. Proof of Theorem 6. The proof follows exactly the same lines as the proof of Theorem 5, the only difference being that K'_{i,N_i} are the number of blocks corresponding to independent random partitions from $\mathrm{PY}(\sigma, \theta; H_0)$ processes, for some diffuse probability measure H_0, and $K_{0,t}$ is the number of blocks of a random partition from a $\mathrm{PY}(\sigma_0, \theta_0; P_0)$. □

A.7. Proof of Theorem 7. First note that the combination of partial exchangeability, which entails conditional independence across samples, and the hierarchical structure of $(\tilde p_1, \dots, \tilde p_d)$, which implies the sharing of atoms across samples, lead to $K_n = K_{0,\xi(\boldsymbol{N})}$ almost surely. Moreover, in view of the growth assumptions, let

$$\frac{K_{0,n}}{\lambda_0(n)} \xrightarrow{\text{a.s.}} M_0, \qquad \frac{K'_{i,n}}{\lambda(n)} \xrightarrow{\text{a.s.}} M'_i \quad (i = 1, \dots, d)$$

as $n \to \infty$, where M_0 and the M_i's are positive and finite random variables. Moreover, one has $\xi(\boldsymbol{N}) = \sum_{i=1}^d K'_{i,N^*}$ and

$$\frac{\xi(\boldsymbol{N})}{\lambda(N^*)} \xrightarrow{\text{a.s.}} \sum_{i=1}^d M_i' = \eta$$

as $N^* \to \infty$. Exploiting (H1) and recalling that $N = dN^*$, one then obtains

$$\frac{K_{0,\xi(\boldsymbol{N})}}{K_{0,\eta\lambda(N^*)}} = \frac{\lambda_0(\xi(\boldsymbol{N}))}{\lambda_0(\eta\lambda(N^*))} \frac{K_{0,\xi(\boldsymbol{N})}/\lambda_0(\xi(\boldsymbol{N}))}{K_{0,\eta\lambda(N^*)}/\lambda_0(\eta\lambda(N^*))} \xrightarrow{\text{a.s.}} 1,$$

which entails, as $N^* \to \infty$,

$$\frac{K_N}{\lambda_0(\eta\lambda(N^*))} = \frac{K_{0,\xi(\boldsymbol{N})}}{K_{0,\eta\lambda(N^*)}} \frac{K_{0,\eta\lambda(N^*)}}{\lambda_0(\eta\lambda(N^*))} \xrightarrow{\text{a.s.}} M_0$$

If $(\tilde p_1, \dots, \tilde p_d)$ is a vector of hierarchical Dirichlet processes, then $\lambda_0(n) = \lambda(n) = \log n$ and (H1) holds true. Moreover, $\eta = dc$ and

$$\frac{\lambda_0(\eta\lambda(N))}{\lambda_0(\lambda(N))} \to 1$$

as $N \to \infty$, and the result follows. □

A.8. Proof of Theorem 8. In this case, $\lambda_0(n) = n^{\sigma_0}$ and $\lambda(n) = n^{\sigma}$. Hence (H1) holds and one may work out the proof along the same lines as for Theorem 7. □

A.9. Proof of Theorem 9. The technique introduced for proving Theorem 3 can be suitably adapted and extended to establish also the posterior characterization. Our goal is the determination of the posterior Laplace functional of $\tilde\mu_0$

$$\mathbb{E}\Big[\mathrm{e}^{-\tilde\mu_0(f)} \,\Big|\, \boldsymbol{X}\Big] = \lim_{\varepsilon\downarrow 0} \frac{\mathbb{E}\,\mathrm{e}^{-\tilde\mu_0(f)} \prod_{j=1}^k \prod_{i=1}^d \tilde p_i^{n_{i,j}}(A_{j,\varepsilon})}{\mathbb{E}\prod_{j=1}^k \prod_{i=1}^d \tilde p_i^{n_{i,j}}(A_{j,\varepsilon})} \tag{A.4}$$

for any measurable $f : \mathbb{X} \to \mathbb{R}^+$, where the $A_{j,\varepsilon}$ notation is the same used in the Proofs of Theorems 3–4. The denominator is $M_{\boldsymbol{n}_1,\dots,\boldsymbol{n}_d}(A_{1,\varepsilon} \times \cdots \times A_{k,\varepsilon})$ that has been defined in the proof of Theorem 3 and, as $\varepsilon \downarrow 0$, equals

$$\prod_{j=1}^k P_0(A_{j\varepsilon}) \sum_{\boldsymbol{\ell}} \sum_{\boldsymbol{q}} \Phi_{k,0}^{(|\boldsymbol{\ell}|)}(\bar\ell_{\bullet 1}, \cdots, \bar\ell_{\bullet k}) \tag{A.5}$$

$$\times \prod_{i=1}^{d}\prod_{j=1}^{k} \frac{1}{\ell_{i,j}!}\binom{n_{i,j}}{q_{i,j,1},\cdots,q_{i,j,\ell_{i,j}}}\,\Phi^{(N_i)}_{\bar\ell_{i\bullet},i}(\boldsymbol{q}_{i,1},\ldots,\boldsymbol{q}_{i,k}) + \lambda_{k,\varepsilon}$$

where $\lambda_{k,\varepsilon} = \mathrm{o}\Big(\prod_{j=1}^{k} P_0(A_{j,\varepsilon})\Big)$. In a similar fashion one can proceed to determine the numerator of (A.4) and

$$\begin{aligned}
\mathbb{E}\,\mathrm{e}^{-\tilde\mu_0(f)}\prod_{j=1}^{k}\prod_{i=1}^{d}\tilde p_i^{n_{i,j}}(A_{j,\varepsilon}) &= \mathbb{E}\,\mathrm{e}^{-\tilde\mu_0(f)}\prod_{i=1}^{d}\mathbb{E}\Big[\tilde p_i^{n_{i,j}}(A_{j,\varepsilon})\,\Big|\tilde\mu_0\Big]\\
&= \sum_{\boldsymbol{\ell}}\prod_{i=1}^{d}\frac{c^{\bar\ell_{i\bullet}}}{\Gamma(N_i)}\int_0^\infty u^{N_i-1}\,\mathrm{e}^{-c\,\psi(u)}\prod_{j=1}^{k}\xi_{n_{i,j},\ell_{i,j}}(u)\,\mathrm{d}u\\
&\quad\times\Big(\mathbb{E}\mathrm{e}^{-\tilde\mu_0(f)}\prod_{j=1}^{k}\tilde p_0^{\bar\ell_{\bullet j}}(A_{j,\varepsilon})\Big)
\end{aligned}$$

It can now be seen that, with $\mathbb{X}^*_\varepsilon = \mathbb{X}\setminus(\cup_{j=1}^{k}A_{j,\varepsilon})$ and as $\varepsilon\downarrow 0$,

$$\begin{aligned}
\mathbb{E}\mathrm{e}^{-\tilde\mu_0(f)}\prod_{j=1}^{k}\tilde p_0^{\bar\ell_{\bullet j}}(A_{j,\varepsilon}) &= \frac{1}{\Gamma(|\boldsymbol{\ell}|)}\int_0^\infty u^{|\boldsymbol{\ell}|-1}\Big(\mathbb{E}\mathrm{e}^{-\tilde\mu_0((f+u)\mathbb{1}_{\mathbb{X}^*_\varepsilon})}\Big)\\
&\quad\times\Big(\prod_{j=1}^{k}\mathbb{E}\mathrm{e}^{-\tilde\mu_0((f+u)\mathbb{1}_{A_{j,\varepsilon}})}\,\tilde\mu_0^{\bar\ell_{\bullet j}}(A_{j,\varepsilon})\Big)\,\mathrm{d}u\\
&= \frac{\prod_{j=1}^{k}c_0\,P_0(A_{j,\varepsilon})}{\Gamma(|\boldsymbol{\ell}|)}\int_0^\infty u^{|\boldsymbol{\ell}|-1}\,\mathrm{e}^{-c_0\,\psi_0(f+u)}\\
&\quad\times\prod_{j=1}^{k}\tau_{\bar\ell_{\bullet j}}(u+f(X_j^*))\,\mathrm{d}u+\lambda_{k,\varepsilon}
\end{aligned}$$

Hence, one has

$$\begin{aligned}
\mathbb{E}\,\mathrm{e}^{-\tilde\mu_0(f)}\prod_{j=1}^{k}\prod_{i=1}^{d}\tilde p_i^{n_{i,j}}(A_{j,\varepsilon}) &= \prod_{j=1}^{k}P_0(A_{j,\varepsilon})\\
&\quad\times\sum_{\boldsymbol{\ell},\boldsymbol{q}}\prod_{i=1}^{d}\prod_{j=1}^{k}\frac{1}{\ell_{i,j}!}\binom{n_{i,j}}{q_{i,j,1}\ \cdots\ q_{i,j,\ell_{i,j}}}\,\Phi_{\bar\ell_{i\bullet},i}(\boldsymbol{q}_{i,1},\ldots,\boldsymbol{q}_{i,k})\\
&\quad\times\frac{c_0^{|\boldsymbol{\ell}|}}{\Gamma(|\boldsymbol{\ell}|)}\int_0^\infty u^{|\boldsymbol{\ell}|-1}\,\mathrm{e}^{-c_0\,\psi_0(f+u)}
\end{aligned}$$

$$\times \prod_{j=1}^{k} \tau_{\bar{\ell}_{\bullet j}}(u + f(X_j^*))\, \mathrm{d}u + \lambda_{k,\varepsilon}.$$

To sum up, if we further condition on $\boldsymbol{T}$, one has

$$\mathbb{E}\Big[\mathrm{e}^{-\tilde{\mu}_0(f)} \,\Big|\, \boldsymbol{X}, \boldsymbol{T}\Big] = \frac{\int_0^\infty u^{|\boldsymbol{\ell}|-1}\, \mathrm{e}^{-c_0\, \psi_0(f+u)} \prod_{j=1}^k \tau_{\bar{\ell}_{\bullet j}}(u + f(X_j^*))\, \mathrm{d}u}{\Phi_{k,0}^{(|\boldsymbol{\ell}|)}(\bar{\ell}_{\bullet 1}, \ldots, \bar{\ell}_{\bullet k})}$$

and the result follows, as it is easy to check that the normalizing constant of the density $f_0(\,\cdot\,|\boldsymbol{X}, \boldsymbol{T})$ in (20) is $\Phi_{k,0}^{(|\boldsymbol{\ell}|)}(\bar{\ell}_{\bullet 1}, \ldots, \bar{\ell}_{\bullet k})$. □

A.10. Proof of Theorem 10. From elementary properties of conditional expectations, one has

$$\mathbb{E}\Big[\mathrm{e}^{-\sum_{i=1}^d \tilde{\mu}_i(f_i)} \,\Big|\, \boldsymbol{X}, \boldsymbol{T}\Big] = \mathbb{E}\Big[\,\mathbb{E}\Big[\mathrm{e}^{-\sum_{i=1}^d \tilde{\mu}_i(f_i)} \,\Big|\, \boldsymbol{X}, \boldsymbol{T}, \tilde{\mu}_0\Big] \,\Big|\, \boldsymbol{X}, \boldsymbol{T}\Big]$$

for any choice of measurable functions $f_i : \mathbb{X} \to \mathbb{R}^+$, for $i = 1, \ldots, d$. The proof, then, boils down to determining posterior Laplace functional

$$\mathbb{E}\Big[\mathrm{e}^{-\sum_{i=1}^d \tilde{\mu}_i(f_i)} \,\Big|\, \boldsymbol{X}, \boldsymbol{T}, \tilde{\mu}_0\Big] =$$

$$\lim_{\varepsilon \downarrow 0} \frac{\mathbb{E}\Big[\mathrm{e}^{-\sum_{i=1}^d \tilde{\mu}_i(f_i)} \prod_{i=1}^d \prod_{j=1}^k \tilde{p}_i^{n_{i,j}}(A_{j,\varepsilon}) \,\Big|\, \boldsymbol{T}, \tilde{\mu}_0\Big]}{\mathbb{E}\Big[\prod_{i=1}^d \prod_{j=1}^k \tilde{p}_i^{n_{i,j}}(A_{j,\varepsilon}) \,\Big|\, \boldsymbol{T}, \tilde{\mu}_0\Big]} \tag{A.6}$$

As far as the denominator is concerned, it plainly equals

$$\Big(\prod_{i=1}^d \prod_{j=1}^k \tilde{p}_0^{\ell_{i,j}}(A_{j,\varepsilon})\Big) \prod_{i=1}^d \frac{c^{\bar{\ell}_{i\bullet}}}{\Gamma(N_i)} \int_0^\infty u^{N_i - 1}\, \mathrm{e}^{-c\,\psi(u)} \prod_{j=1}^k \prod_{t=1}^{\ell_{i,j}} \tau_{q_{i,j,t}}(u)\, \mathrm{d}u$$

As for the numerator, set

$$\tilde{\psi}(f) = \int_{\mathbb{X}} \int_0^\infty [1 - \mathrm{e}^{-s f(x)}]\, \rho(s)\, \mathrm{d}s\, \tilde{p}_0(\mathrm{d}x)$$

and note that

$$\mathbb{E}\left[\mathrm{e}^{-\sum_{i=1}^d \tilde{\mu}_i(f_i)} \prod_{i=1}^d \prod_{j=1}^k \tilde{p}_i^{n_{i,j}}(A_{j,\varepsilon}) \,\Big|\, \boldsymbol{T}, \tilde{\mu}_0\right]$$

$$= \prod_{i=1}^d \mathbb{E}\Big[\mathrm{e}^{-\tilde{\mu}_i(f_i)} \prod_{j=1}^k \tilde{p}_i^{n_{i,j}}(A_{j,\varepsilon}) \,\Big|\, \boldsymbol{T}, \tilde{\mu}_0\Big]$$

$$= \Big(\prod_{i=1}^d \prod_{j=1}^k \tilde p_0^{\ell_{i,j}}(A_{j,\varepsilon})\Big)$$
$$\times \prod_{i=1}^d \frac{c^{\bar\ell_{i\bullet}}}{\Gamma(N_i)} \int_0^\infty u^{N_i-1}\,\mathrm{e}^{-c\tilde\psi(f+u)} \prod_{j=1}^k \prod_{t=1}^{\ell_{i,j}} \tau_{q_{i,j,t}}(u+f_i(X_j^*))\,\mathrm{d}u$$

To sum up, the right–hand–side of (A.6) equals

$$\prod_{i=1}^d \frac{\int_0^\infty u^{N_i-1}\mathrm{e}^{-c\tilde\psi(f_i+u)} \prod_{j=1}^k \prod_{t=1}^{\ell_{i,j}} \tau_{q_{i,j,t}}(u+f_i(X_j^*))\mathrm{d}u}{\int_0^\infty u^{N_i-1}\mathrm{e}^{-c\psi(u)} \prod_{j=1}^k \prod_{t=1}^{\ell_{i,j}} \tau_{q_{i,j,t}}(u)\mathrm{d}u}.$$

which entails

$$\mathbb{E}\Big[\mathrm{e}^{-\sum_{i=1}^d \tilde\mu_i(f_i)}\,\Big|\,\boldsymbol{X},\boldsymbol{T},\boldsymbol{U},\tilde\mu_0\Big] = \prod_{i=1}^d \prod_{j=1}^k \prod_{t=1}^{\ell_{i,j}} \frac{\tau_{q_{i,j,t}}(U_i+f_i(X_j^*))}{\tau_{q_{i,j,t}}(U_i)}$$
$$\times \prod_{i=1}^d \exp\Big\{-c\int_{\mathbb{X}\times\mathbb{R}^+} (1-e^{-sf_i(x)})\mathrm{e}^{-sU_i}\,\rho(s)\,\mathrm{d}s\,\tilde p_0(\mathrm{d}x)\Big\}$$

and the conclusion follows. □

A.11. Proof of Theorem 11. The proof follows the similar lines as that of Theorem 9, the main difference being the polynomial tilting that defines the Pitman–Yor process in (8). Let $\tilde\mu_0'$ and μ_i' denote stable CRMs with respective parameters σ_0 and σ in $(0,1)$. Moreover, the base measure $\tilde\mu_0'$ is P_0 whereas the $\tilde\mu_i'$ are, conditional on $\tilde\mu_0'$, independent and identically distributed CRMs with base measure $\tilde p_0' = \tilde\mu_0'/\tilde\mu_0'(\mathbb{X})$. When determining the posterior Laplace functional transform as in (A.4), one has

$$\text{(A.7)}\quad \mathbb{E}\,\mathrm{e}^{-\tilde\mu_0(f)} \prod_{j=1}^k \prod_{i=1}^d \tilde p_i^{n_{i,j}}(A_{j,\varepsilon})$$
$$= \mathbb{E}\Big[\mathrm{e}^{-\tilde\mu_0(f)} \prod_{i=1}^d \mathbb{E}\Big[(\tilde\mu_i(\mathbb{X}))^{-N_i} \prod_{j=1}^k (\tilde\mu_i(A_{j,\varepsilon}))^{n_{i,j}}\,\Big|\,\tilde\mu_0\Big]\Big].$$

It can now be seen that, for any $i=1,\ldots,d$,

$$\mathbb{E}\Big[(\tilde\mu_i(\mathbb{X}))^{-N_i} \prod_{j=1}^k (\tilde\mu_i(A_{j,\varepsilon}))^{n_{i,j}}\,\Big|\,\tilde\mu_0\Big]$$

$$= \frac{\Gamma(\theta+1)}{\Gamma\Big(\frac{\theta}{\sigma}+1\Big)}\,\mathbb{E}\Big[(\tilde\mu_i'(\mathbb{X}))^{-\theta-N_i}\,\prod_{j=1}^k(\tilde\mu_i'(A_{j,\varepsilon}))^{n_{i,j}}\,\Big|\,\tilde\mu_0\Big]$$

$$= \frac{1}{(\theta+1)_{N_i-1}}\,\frac{1}{\Gamma\Big(\frac{\theta}{\sigma}+1\Big)}$$

$$\times\int_0^\infty v^{\theta+N_i-1}\mathbb{E}\Big[\mathrm{e}^{-v\tilde\mu_i'(\mathbb{X})}\,\prod_{j=1}^k(\tilde\mu_i'(A_{j,\varepsilon}))^{n_{i,j}}\,\Big|\,\tilde\mu_0\Big]\,\mathrm{d}v$$

$$= \frac{1}{(\theta+1)_{N_i-1}}\,\frac{1}{\Gamma\Big(\frac{\theta}{\sigma}+1\Big)}$$

$$\times\int_0^\infty v^{\theta+N_i-1}\,\mathrm{e}^{-v^\sigma}\prod_{j=1}^k\sum_{\ell_{i,j}=1}^{n_{i,j}}\tilde p_0^{\ell_{i,j}}(A_{j,\varepsilon})\,\xi_{n_{i,j},\ell_{i,j}}(v)\,\mathrm{d}v.$$

It is clear that, in order to evaluate (A.7), one may note that as $\varepsilon\downarrow 0$

$$\mathbb{E}\,\mathrm{e}^{-\tilde\mu_0(f)}\prod_{j=1}^k\tilde p_0^{\bar\ell_{\bullet j}}(A_{j,\varepsilon})$$

$$= \frac{\Gamma(\theta_0+1)}{\Gamma\Big(\frac{\theta_0}{\sigma_0}+1\Big)}\,\mathbb{E}(\tilde\mu_0'(\mathbb{X}))^{-\theta_0-|\boldsymbol{\ell}|}\,\mathrm{e}^{-\tilde\mu_0'(f)}\,\prod_{j=1}^k(\tilde\mu_0'(A_{j,\varepsilon}))^{\bar\ell_{\bullet j}}$$

$$= \frac{1}{(\theta_0+1)_{|\boldsymbol{\ell}|-1}}\,\frac{\sigma_0^k}{\Gamma\Big(\frac{\theta_0}{\sigma_0}+1\Big)}\Big(\prod_{j=1}^k P_0(A_{j,\varepsilon})\,(1-\sigma_0)_{\bar\ell_{\bullet j}-1}\Big)\int_0^\infty v_0^{\theta_0+|\boldsymbol{\ell}|-1}$$

$$\times\;\mathrm{e}^{-\int_0^\infty(v_0+f(x))^{\sigma_0}\,P_0(\mathrm{d}x)}\,\prod_{j=1}^k\frac{1}{(v_0+f(X_j^*))^{\bar\ell_{\bullet j}-\sigma_0}}\,\mathrm{d}v_0+\lambda_{k,\varepsilon}$$

where $\lambda_{k,\varepsilon}=\mathrm{o}(\prod_{j=1}^k P_0(A_{j,\varepsilon}))$. If one further conditions on $\boldsymbol{T}$, it can be seen that

$$\mathbb{E}\left[\mathrm{e}^{-\tilde\mu_0(f)}\,\prod_{j=1}^k\prod_{i=1}^d\tilde p_i^{n_{i,j}}(A_{j,\varepsilon})\,\Big|\,\boldsymbol{T}\right]$$

$$= \frac{1}{(\theta_0+1)_{|\boldsymbol{\ell}|-1}}\,\frac{\sigma_0^k}{\Gamma\Big(\frac{\theta_0}{\sigma_0}+1\Big)}\,\prod_{i=1}^d\frac{\prod_{r=1}^{\bar\ell_{i\bullet}-1}(\theta+r\sigma)}{(\theta+1)_{N_i-1}}$$

$$\times \prod_{j=1}^{k} P_0(A_{j,\varepsilon})(1-\sigma_0)_{\bar{\ell}_{\bullet j}-1} \prod_{i=1}^{d} \frac{\mathscr{C}(n_{i,j},\ell_{i,j};\sigma)}{\sigma^{\ell_{i,j}}}$$

$$\times \int_0^\infty v_0^{\theta_0+k\sigma_0-1}\,\mathrm{e}^{-v_0^{\sigma_0}}\,\mathrm{e}^{-\int_0^\infty [(v_0+f(x))^{\sigma_0}-v_0^{\sigma_0})\,P_0(\mathrm{d}x)}$$

$$\times \Big(\prod_{j=1}^{k} \mathbb{E}\,\mathrm{e}^{-I_j f(X_j^*)}\Big)\,\mathrm{d}v_0$$

The proof is, thus, completed by using Theorem 4, which identifies

$$\mathbb{E}\Big[\prod_{i=1}^{d}\prod_{j=1}^{k} \tilde{p}_i^{n_{i,j}}(A_{j,\varepsilon})\,\Big|\,\boldsymbol{T}\Big],$$

as $\varepsilon \downarrow 0$. □

A.12. Proof of Theorem 12. As in the proof of Theorem 10, we aim at determining

$$\mathbb{E}\Big[\mathrm{e}^{-\sum_{i=1}^{d}\tilde{\mu}_i(f_i)}\,\Big|\,\boldsymbol{X},\boldsymbol{T}\Big] = \mathbb{E}\Big[\mathbb{E}\Big[\mathrm{e}^{-\sum_{i=1}^{d}\tilde{\mu}_i(f_i)}\,\Big|\,\boldsymbol{X},\boldsymbol{T},\tilde{\mu}_0\Big]\,\Big|\,\boldsymbol{X},\boldsymbol{T}\Big]$$

for any collection of non–negative and measurable functions $f_1,\ldots,f_d$ defined on $\mathbb{X}$. See (A.6). This is achieved through the following

$$\mathbb{E}\Big[\mathrm{e}^{-\sum_{i=1}^{d}\tilde{\mu}_i(f_i)}\,\Big|\,\boldsymbol{X},\boldsymbol{T},\tilde{\mu}_0\Big] = \prod_{i=1}^{d}\mathbb{E}\Big[\mathrm{e}^{-\sum_{i=1}^{d}\tilde{\mu}_i(f_i)}\prod_{j=1}^{k}\tilde{p}_i^{n_{i,j}}(A_{j,\varepsilon})\,\Big|\,\boldsymbol{T},\tilde{\mu}_0\Big].$$

For any $i=1,\ldots,d$, one has

$$\begin{aligned}
&\mathbb{E}\Big[\mathrm{e}^{-\sum_{i=1}^{d}\tilde{\mu}_i(f_i)}\prod_{j=1}^{k}\tilde{p}_i^{n_{i,j}}(A_{j,\varepsilon})\,\Big|\,\boldsymbol{T},\tilde{\mu}_0\Big]\\
&\quad= \frac{1}{(\theta+1)_{N_i-1}}\,\frac{1}{\Gamma\Big(\frac{\theta}{\sigma}+1\Big)}\\
&\qquad\times \int_0^\infty v^{\theta+N_i-1}\,\mathbb{E}\Big[\mathrm{e}^{-\tilde{\mu}_i'(v+f_i)}\prod_{j=1}^{k}(\tilde{\mu}_i'(A_{j,\varepsilon}))^{n_{i,j}}\,\Big|\,\boldsymbol{T},\tilde{\mu}_0\Big]\,\mathrm{d}v\\
&\quad= \frac{1}{(\theta+1)_{N_i-1}}\,\frac{1}{\Gamma\Big(\frac{\theta}{\sigma}+1\Big)}\prod_{j=1}^{k}\tilde{p}_0^{\ell_{i,j}}(A_{j,\varepsilon})\,\mathscr{C}(n_{i,j},\ell_{i,j};\sigma)
\end{aligned}$$

$$\times \int_0^\infty v^{\theta+N_i-1}\,\mathrm{e}^{-\int_0^\infty [v+f_i(x)]^\sigma \tilde p_0(\mathrm{d}x)} \prod_{j=1}^k \frac{1}{(v+f_i(X_j^*))^{n_{i,j}-\ell_{i,j}\sigma}}\,\mathrm{d}v$$

$$= \frac{1}{(\theta+1)_{N_i-1}}\,\frac{1}{\Gamma\Big(\frac{\theta}{\sigma}+1\Big)}\prod_{j=1}^k \tilde p_0^{\ell_{i,j}}(A_{j,\varepsilon})\,\mathscr{C}(n_{i,j},\ell_{i,j};\sigma)$$

$$\times \int_0^\infty v^{\theta+\bar\ell_{i\bullet}\sigma-1}\,\mathrm{e}^{-v^\sigma}\,\mathrm{e}^{-\int_{\mathbb{X}}[(v+f_i(x))^\sigma - v^\sigma]\tilde p_0(\mathrm{d}x)}$$

$$\times \prod_{j=1}^k \frac{1}{\Big(1+\frac{f_i(X_j^*)}{v}\Big)^{n_{i,j}-\ell_{i,j}\sigma}}\,\mathrm{d}v$$

$$= \frac{\prod_{r=1}^{\bar\ell_{i\bullet}-1}(\theta+r\sigma)}{(\theta+1)_{N_i-1}}\prod_{j=1}^k \tilde p_0^{\ell_{i,j}}(A_{j,\varepsilon})\,\frac{\mathscr{C}(n_{i,j},\ell_{i,j};\sigma)}{\sigma^{\ell_{i,j}}}$$

$$\times \int_0^\infty h_i(v)\,\Big(\mathbb{E}_v\,\mathrm{e}^{-\tilde\mu_i^*(f_i)}\Big)\,\Big(\prod_{j=1}^k \mathbb{E}_v \mathrm{e}^{-H_{i,j}f_i(X_j^*)}\Big)\,\mathrm{d}v$$

where h_i is the density function of a random variable V_i such that $V_i^\sigma \sim \mathrm{Ga}(\bar\ell_{i\bullet}+\theta/\sigma;1)$ and $\tilde\mu_i^*$ is, conditionally on V_i and on $\tilde\mu_0$, a generalized gamma CRM with parameters (V_i,σ) and base measure $\tilde p_0 = \tilde\mu_0/\tilde\mu_0(\mathbb{X})$. Moreover, the $H_{i,j}$'s are independent gamma random variables whose parameters are $n_{i,j}-\ell_{i,j}\sigma$ and V_i. If one, now, notes that

$$\mathbb{E}\Big[\prod_{i=1}^d\prod_{j=1}^k \tilde p_i^{n_{i,j}}(A_{j,\varepsilon})\,\Big|\,\boldsymbol{T},\tilde\mu_0\Big] =$$

$$\frac{\prod_{r=1}^{\bar\ell_{i\bullet}-1}(\theta+r\sigma)}{(\theta+1)_{N_i-1}}\prod_{j=1}^k \tilde p_0^{\ell_{i,j}}(A_{j,\varepsilon})\,\frac{\mathscr{C}(n_{i,j},\ell_{i,j};\sigma)}{\sigma^{\ell_{i,j}}}$$

which, combined with the result proved in Theorem 11, yields the result. □

A.13. Proof of Theorem 13. Let $\tilde\mu_\sigma$ denote a σ–stable CRM and, as in (9), $\tilde\mu_{\sigma,\theta}$ indicates a random measure whose normalization yields a Pitman–Yor process with parameters (σ,θ). From Theorem 11, one has that

$$\mathbb{E}\Big[\mathrm{e}^{-\eta_0^*(f)}\Big] = \frac{\sigma_0}{\Gamma\Big(\frac{\theta_0}{\sigma_0}+k\Big)}$$

$$\times \int_0^\infty v^{\theta_0+k\sigma_0-1}\,\mathrm{e}^{-v^{\sigma_0}}\,\mathrm{e}^{-\int_{\mathbb{X}}[(v+f(x))^{\sigma_0}-v^{\sigma_0}]\,P_0(\mathrm{d}x)}\,\mathrm{d}v$$
$$= \frac{\sigma_0}{\Gamma\Big(\frac{\theta_0}{\sigma_0}+k\Big)}\int_0^\infty v^{\theta_0+k\sigma_0-1}\Big(\mathbb{E}\,\mathrm{e}^{-v\tilde\mu_{\sigma_0}(\mathbb{X})-\tilde\mu_{\sigma_0}(f)}\Big)\,\mathrm{d}v$$
$$= \frac{\sigma_0\,\Gamma(\theta_0+k\sigma_0)}{\Gamma\Big(\frac{\theta_0}{\sigma_0}+k\Big)}\Big(\mathbb{E}\,\mathrm{e}^{-\tilde\mu_{\sigma_0}(f)}\{\tilde\mu_{\sigma_0}(\mathbb{X})\}^{-\theta_0-k\sigma_0}\Big)$$
$$= \mathbb{E}\,\mathrm{e}^{-\tilde\mu_{\sigma_0,\theta_0+k\sigma_0}(f)}$$

Hence, one can conclude that $\eta_0^* \stackrel{\mathrm{d}}{=} \tilde\mu_{\sigma_0,\theta_0+k\sigma_0}$ and

$$\tilde p_0\,|\,(\boldsymbol{X},\boldsymbol{T},V_0) \stackrel{\mathrm{d}}{=} \sum_{j=1}^k W_{j,V_0}\,\delta_{X_j^*} + W_{k+1,V_0}\,\tilde p_{\sigma_0,\theta_0+k\sigma_0}$$

where $\tilde p_{\sigma_0,\theta_0+k\sigma_0} \sim \mathrm{PY}(\sigma_0,\theta_0+k\sigma_0;P_0)$. Moreover $W_{j,V_0} = I_j/(\eta_0^*(\mathbb{X}) + \sum_{i=1}^k I_i)$, for any $j=1,\dots,k$, and $W_{k+1,V_0} = \eta_0^*(\mathbb{X})/(\eta_0^*(\mathbb{X})+\sum_{i=1}^k I_i)$. If f_v is the density of the vector $(W_{1,V_0},\dots,W_{k,V_0})$ on the k–dimensional simplex Δ_k, we would like to determine

$$f(w_1,\dots,w_k) = \int_0^\infty f_v(w_1,\dots,w_k)\,\frac{\sigma_0}{\Gamma\Big(\frac{\theta_0}{\sigma_0}+k\Big)}\,v^{\theta_0+k\sigma_0-1}\,\mathrm{e}^{-v^{\sigma_0}}\,\mathrm{d}v.$$

To this end, we denote by h_v the density function of $\eta_0^*(\mathbb{X})$ and, using independence, the vector $(I_1,\dots,I_k,\eta_0^*(\mathbb{X}))$ has density given by

$$f_v^*(x_1,\dots,x_k,t) = h_v(t)\,\frac{v^{|\boldsymbol{\ell}|-k\sigma_0}}{\prod_{j=1}^k\Gamma(\bar\ell_{\bullet j}-\sigma_0)}\,\mathrm{e}^{-v\sum_{i=1}^k x_i}\prod_{j=1}^k x_j^{\bar\ell_{\bullet j}-\sigma_0-1}.$$

If $W_{V_0} = \sum_{i=1}^k I_j + \eta_0^*(\mathbb{X})$, a simple vector transformation yields a density function for the vector $(W_{1,V_0},\dots,W_{k,V_0},W_{V_0})$

$$f_v(w_1,\dots,w_k,w) = \frac{v^{|\boldsymbol{\ell}|-k\sigma_0}\prod_{j=1}^k w_j^{\bar\ell_{\bullet j}-\sigma_0-1}}{\prod_{j=1}^k\Gamma(\bar\ell_{\bullet j}-\sigma_0)}\,w^{|\boldsymbol{\ell}|-k\sigma_0}\,\mathrm{e}^{-vw|\boldsymbol{w}|}\,h_v(w(1-|\boldsymbol{w}|))$$

where $|\boldsymbol{w}| = \sum_{i=1}^k w_i$. From this, an expression for the density of $(W_{1,V_0},\dots,W_{k,V_0})$ easily follows and it turns out to be equal to

$$f_v(w_1,\dots,w_k) = \frac{v^{|\boldsymbol{\ell}|-k\sigma_0}\prod_{j=1}^k w_j^{\bar\ell_{\bullet j}-\sigma_0-1}}{\prod_{j=1}^k\Gamma(\bar\ell_{\bullet j}-\sigma_0)}\,\frac{1}{(1-|\boldsymbol{w}|)^{|\boldsymbol{\ell}|-k\sigma_0+1}}$$

$$
\begin{aligned}
&\times \int_0^\infty s^{|\boldsymbol{\ell}|-k\sigma_0}\, \mathrm{e}^{-\frac{vs|\boldsymbol{w}|}{1-|\boldsymbol{w}|}}\, h_v(s)\,\mathrm{d}s \\
&= \frac{v^{|\boldsymbol{\ell}|-k\sigma_0}\, \prod_{j=1}^k w_j^{\bar\ell_{\bullet j}-\sigma_0-1}}{\prod_{j=1}^k \Gamma(\bar\ell_{\bullet j}-\sigma_0)}\, \frac{1}{(1-|\boldsymbol{w}|)^{|\boldsymbol{\ell}|-k\sigma_0+1}} \\
&\times \Big(\mathbb{E}(\eta_0^*(\mathbb{X}))^{|\boldsymbol{\ell}|-k\sigma_0}\, \mathrm{e}^{-\frac{v|\boldsymbol{w}|}{1-|\boldsymbol{w}|}\,\eta_0^*(\mathbb{X})}\Big).
\end{aligned}
$$

Since η_0^* is a generalized gamma CRM with parameters (σ_0, V_0) and base measure P_0, its probability distributions $\mathbb{P}^*$ is absolutely continuous with respect to the probability distribution $\mathbb{P}_{\sigma_0}$ of a σ_0–stable CRM and

$$
\frac{\mathrm{d}\mathbb{P}^*}{\mathrm{d}\mathbb{P}_{\sigma_0}}(m) = \exp\{-vm(\mathbb{X}) + v^{\sigma_0}\},
$$

then

$$
\mathbb{E}(\eta_0^*(\mathbb{X}))^{|\boldsymbol{\ell}|-k\sigma_0}\, \mathrm{e}^{-\frac{v|\boldsymbol{w}|}{1-|\boldsymbol{w}|}\,\eta_0^*(\mathbb{X})} = \mathrm{e}^{v^{\sigma_0}}\, \mathbb{E}\left(\tilde\mu_{\sigma_0}(\mathbb{X})\right)^{|\boldsymbol{\ell}|-k\sigma_0}\, \mathrm{e}^{-\frac{v}{1-|\boldsymbol{w}|}\,\tilde\mu_{\sigma_0}(\mathbb{X})}
$$

In view of this, one can now marginalize f_v with respect to v and and obtain a density of $(W_1, \ldots, W_k)$. Indeed, one has

$$
\begin{aligned}
f(w_1,\ldots,w_k) &= \frac{\sigma_0}{\Gamma\Big(\frac{\theta_0}{\sigma_0}+k\Big)} \int_0^\infty v^{\theta_0+k\sigma_0-1}\, \mathrm{e}^{-v^{\sigma_0}}\, f_v(w_1,\ldots,w_k)\,\mathrm{d}v \\
&= \frac{\sigma_0}{\Gamma\Big(\frac{\theta_0}{\sigma_0}+k\Big)}\, \frac{\prod_{j=1}^k w_j^{\bar\ell_{\bullet j}-\sigma_0-1}}{\prod_{j=1}^k \Gamma(\bar\ell_{\bullet j}-\sigma_0)}\, \frac{1}{(1-|\boldsymbol{w}|)^{|\boldsymbol{\ell}|-k\sigma_0+1}} \\
&\quad \times\ \mathbb{E}\left(\tilde\mu_{\sigma_0}(\mathbb{X})\right)^{|\boldsymbol{\ell}|-k\sigma_0} \int_0^\infty v^{\theta_0+|\boldsymbol{\ell}|-1}\, \mathrm{e}^{-\frac{v}{1-|\boldsymbol{w}|}\,\tilde\mu_{\sigma_0}(\mathbb{X})}\mathrm{d}v \\
&= \frac{\sigma_0\,\Gamma(\theta_0+|\boldsymbol{\ell}|)}{\Gamma\Big(\frac{\theta_0}{\sigma_0}+k\Big)}\, \frac{\Big(\mathbb{E}\left(\tilde\mu_{\sigma_0}(\mathbb{X})\right)^{-\theta_0-k\sigma_0}\Big)}{\prod_{j=1}^k \Gamma(\bar\ell_{\bullet j}-\sigma_0)} \\
&\quad \times\ (1-|\boldsymbol{w}|)^{\theta_0+k\sigma_0-1} \prod_{j=1}^k w_j^{\bar\ell_{\bullet j}-\sigma_0-1} \\
&= \frac{\Gamma(\theta_0+|\boldsymbol{\ell}|)}{\Gamma(\theta_0+k\sigma_0)\prod_{j=1}^k \Gamma(\bar\ell_{\bullet j}-\sigma_0)}\, (1-|\boldsymbol{w}|)^{\theta_0+k\sigma_0-1} \prod_{j=1}^k w_j^{\bar\ell_{\bullet j}-\sigma_0-1}
\end{aligned}
$$

and this completes the proof of the posterior characterization of $\tilde p_0$.
Finally, one can proceed in a similar fashion in order to prove (25) □

APPENDIX B: BLACKWELL–MACQUEEN URN SCHEME FOR HIERARCHICAL PITMAN–YOR PROCESSES

Here we specialize the algorithm described in Section 6.1 to hierarchies of Pitman–Yor processes. Hence we are assuming that $\boldsymbol{X}$ are partially exchangeable as in (1), where the prior Q_d is characterized as follows

$$\tilde{p}_i \,|\, \tilde{p}_0 \overset{\text{iid}}{\sim} \text{PY}(\sigma, \theta; \tilde{p}_0) \quad (i = 1, \dots, d), \qquad \tilde{p}_0 \sim \text{PY}(\sigma_0, \theta_0; P_0)$$

The same notations as in Section 6.1 are used. The determination of the full conditionals follows immediately from the augmented pEPPF

$$\begin{aligned}
\Pi_{k+j}^{(n+m)}&(\boldsymbol{n}_1, \cdots, \boldsymbol{n}_d; \boldsymbol{\ell}, \boldsymbol{q}) \\
&= \frac{\prod_{r=1}^{k+j-1}(\theta_0 + r\sigma_0)}{(\theta_0 + 1)_{|\boldsymbol{\ell}|-1}} \prod_{t=1}^{k+j} (1-\sigma_0)_{\bar{\ell}_{\bullet t}-1} \\
&\quad \times \frac{\prod_{r=1}^{\bar{\ell}_{1\bullet}-1}(\theta + r\sigma)}{(\theta+1)_{N_1+m_1-1}} \prod_{v=1}^{k+j} \prod_{t=1}^{\ell_{1,v}} (1-\sigma)_{q_{1,v,t}-1} \\
&\quad \times \frac{\prod_{r=1}^{\bar{\ell}_{2\bullet}-1}(\theta + r\sigma)}{(\theta+1)_{N_2+m_2-1}} \prod_{v=1}^{k+j} \prod_{t=1}^{\ell_{2,v}} (1-\sigma)_{q_{2,v,t}-1}
\end{aligned} \tag{B.1}$$

with the convention $(1-\sigma)_{-1} \equiv 1$. Based on (B.1) one can devise a Gibbs sampler that generates $(T_{i,1}, \dots, T_{i,N_i})$, for $i = 1, 2$, and $(X_{i,N_i+r}, T_{i,N_i+r})$, for $r = 1, \dots, m_i$ and $i = 1, 2$, from their respective full conditionals. Details are provided for $i = 1$, the case $i = 2$ being identical with the appropriate adaptations.

(1) At $t = 0$, start from an initial configuration $X_{i,N_i+1}^{(0)}, \dots, X_{i,N_i+m_i}^{(0)}$ and $T_{i,1}^{(0)}, \dots, T_{i,N_i+m_i}^{(0)}$, for $i = 1, 2$.

(2) At iteration $t \geq 1$

(2.a) With $X_{1,r} = X_h^*$ generate latent variables $T_{1,r}^{(t)}$, for $r = 1, \dots, N_i$, from

$$\mathbb{P}(T_{1,r} = \text{“new”} \,|\, \cdots) \propto w_{h,r} \frac{(\theta + \bar{\ell}_{1\bullet}^{-r}\sigma)}{(|\boldsymbol{\ell}^{-r}| + \theta_0)}$$

$$\mathbb{P}(T_{1,r} = T_{1,h,\kappa}^{*,-r} \,|\, \cdots) \propto (q_{1,h,\kappa}^{-r} - \sigma) \qquad \text{for } \kappa = 1, \dots, \ell_{1,h}^{-r}$$

where $w_{h,r} = \bar{\ell}_{\bullet h}^{-r} - \sigma_0$ if $\bar{\ell}_{\bullet h}^{-r} > 0$ and $w_{h,r} = 1$ otherwise. Moreover, $T_{1,h,1}^{*,-r}, \dots, T_{1,h,\ell_{1,h}^{-r}}^{*,-r}$ are the tables at the first restaurant where the h-th dish is served, after the removal of $T_{1,r}$.

(2.b) For $r = 1, \dots, m_1$, generate $(X^{(t)}_{N_i+r}, T^{(t)}_{N_i+r})$ from the following predictive distributions

$$\mathbb{P}(X_{1,r} = \text{“new”}, T_{1,r} = \text{“new”} | \cdots) = \frac{(\theta_0 + (k + j^{-r})\sigma_0)}{(\theta + N_1 + m_1 - 1)} \frac{(\theta + \bar{\ell}^{-r}_{1\bullet}\sigma)}{(\theta_0 + |\boldsymbol{\ell}^{-r}|)}$$

while, for any $h = 1, \cdots, k + j^{-r}$ and $\kappa = 1, \cdots, \ell^{-r}_{1,h}$,

$$\mathbb{P}(X_{1,r} = X^{*,-r}_h, T_{1,r} = \text{“new”} | \cdots) = \frac{(\bar{\ell}^{-r}_{\bullet h} - \sigma_0)}{(\theta + N_1 + m_1 - 1)} \frac{(\theta + \bar{\ell}^{-r}_{1\bullet}\sigma)}{(\theta_0 + |\boldsymbol{\ell}^{-r}|)}$$

$$\mathbb{P}(X_{1,r} = X^{*,-r}_h, T_{1,r} = T^{*,-r}_{1,h,\kappa} | \cdots) = \frac{q^{-r}_{1,h,\kappa} - \sigma}{\theta + N_1 + m_1 - 1} \mathbb{1}_{\{n^{-r}_{i,h} > 0\}}$$

recalling that $X^{*,-r}_h$, for $h = 1, \dots, k + j^{-r}$ denote the distinct dishes in the whole franchise after the removal of the r–th observation, while the condition $n^{-r}_{i,h} > 0$ entails that the h–th dish is served in the i–th restaurant.

REFERENCES

[1] Hardy M. (2006). Combinatorics of partial derivatives. *Electron. J. Comb.* **13,** 1–13.

[2] James, L.F., Lijoi, A. and Prünster, I. (2006). Conjugacy as a distinctive feature of the Dirichlet process. *Scand. J. Statist.* **33**, 105–120.

Federico Camerlenghi
Department of Economics,
Management and Statistics
University of Milano–Bicocca
Piazza dell'Ateneo Nuovo 1, 20126 Milano, Italy
E-mail: federico.camerlenghi@unimib.it

Antonio Lijoi and Igor Prünster
Department of Decision Sciences and BIDSA
Bocconi University
Via Röntgen 1, 20136 Milano, Italy
E-mail: antonio.lijoi@unibocconi.it; igor@unibocconi.it

Peter Orbanz
Department of Statistics
Columbia University
1255 Amsterdam Avenue 10027 New York, USA
E-mail: porbanz@stat.columbia.edu