Extended Multiobject Tracking with Nonparametric Bayesian Learning of Object Classes and Model Parameters: Supplementary Material

Thomas John Bucco, Bernd Kreidl, Günther Koliander, and Franz Hlawatsch

March 7, 2025

This manuscript provides detailed derivations and background material for the publication "Extended Multiobject Tracking with Nonparametric Bayesian Learning of Object Classes and Model Parameters" [1], hereafter referred to as "main paper." The section numbers and equation labels in this manuscript contain the prefix "S", e.g., (S.1). When sections and equations are referenced without this prefix, they can be found in the main paper. Notation and abbreviations established in the main paper will typically be used without redefining them.

S1 Statistical Independence Properties

We will first list and, in some cases, briefly discuss the independence and conditional independence properties that are inherent in the system model presented in Sections II and III of the main paper. The independence of random vectors $\mathbf{a}$ and $\mathbf{b}$ is expressed by $\mathbf{a} \perp\!\!\!\perp \mathbf{b}$, and the conditional independence of $\mathbf{a}$ and $\mathbf{b}$ given $\mathbf{c}$ is expressed by $\mathbf{a} \perp\!\!\!\perp \mathbf{b} | \mathbf{c}$. The main independence structures are illustrated in the Bayesian networks [2] depicted in Fig. 1. These two Bayesian networks are from different viewpoints, and we will summarize in the following the (conditional) independence properties expressed by each network that are required in our derivations below. We start with the (conditional) independence properties between objects as expressed by Fig. 1(a).

A.1) $\mathbf{p}_c^* \perp\!\!\!\perp \mathbf{c}, \mathbf{P}_{\neg c}^*$, where $\mathbf{P}_{\neg c}^*$ is the matrix consisting of vectors $\mathbf{p}_{c'}^*$ for all $c' \in \mathcal{C} \setminus \{c\}$ with the set $\mathcal{C} \triangleq \{c_1, c_2, \dots, c_I\}$. This is a core assumption of DP models [3, 4].

A.2) $\mathbf{X}_{\mathcal{K}_i,i}, \mathbf{Z}_{\mathcal{K}_i,i} \perp\!\!\!\perp \mathbf{X}_{\mathcal{K}_{i'},i'}, \mathbf{Z}_{\mathcal{K}_{i'},i'} \,\big|\, \mathbf{p}_c^*, \mathbf{c}, \mathbf{P}_{\neg c}^*$, for $i' \neq i$.

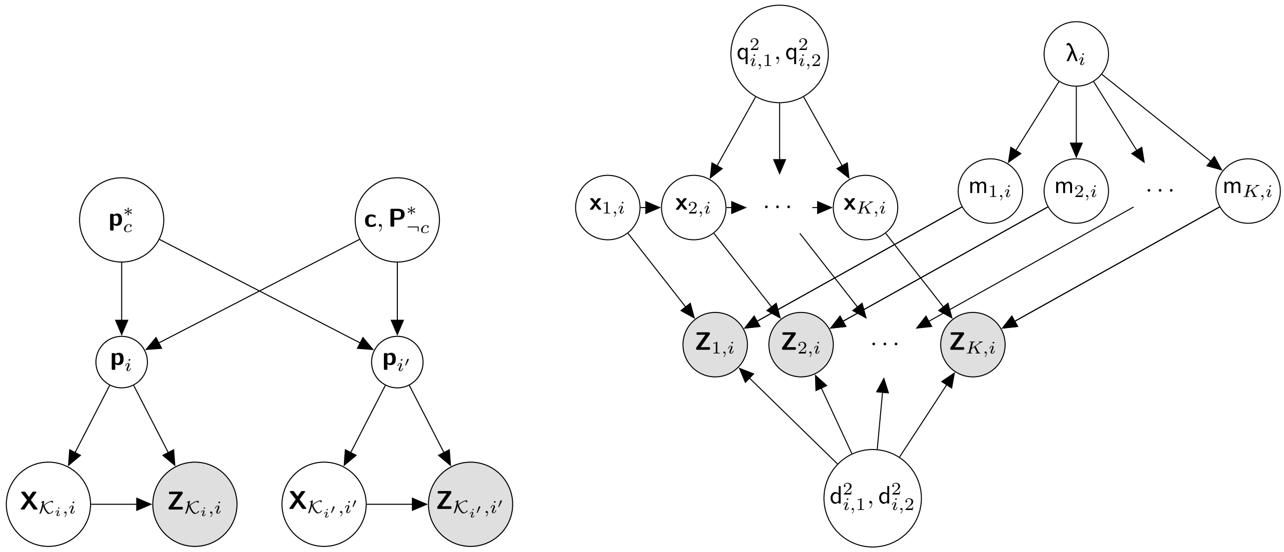
A.3) $\mathbf{X}_{\mathcal{K}_i,i}, \mathbf{Z}_{\mathcal{K}_i,i} \perp\!\!\!\perp \mathbf{p}_c^*, \mathbf{c}, \mathbf{P}_{\neg c}^* \,\big|\, \mathbf{p}_i$.

Next, we list the (conditional) independence properties of the random variables associated with a single object, as expressed by Fig. 1(b).

B.1) $\mathbf{X}_{\mathcal{K}_i,i} \perp\!\!\!\perp \mathsf{d}_{i,1}^2, \mathsf{d}_{i,2}^2, \lambda_i \,\big|\, \mathsf{q}_{i,1}^2, \mathsf{q}_{i,2}^2$.

B.2) $\mathbf{Z}_{\mathcal{K}_i,i} \perp\!\!\!\perp \mathsf{q}_{i,1}^2, \mathsf{q}_{i,2}^2, \lambda_i \,\big|\, \mathbf{X}_{\mathcal{K}_i,i}, \mathbf{m}_{\mathcal{K}_i,i}, \mathsf{d}_{i,1}^2, \mathsf{d}_{i,2}^2$, where $\mathbf{m}_{\mathcal{K}_i,i} \triangleq [\mathsf{m}_{k,i}]_{k \in \mathcal{K}_i}$.

B.3) $\mathbf{m}_{\mathcal{K}_i,i} \perp\!\!\!\perp \mathbf{X}_{\mathcal{K}_i,i}, \mathsf{q}_{i,1}^2, \mathsf{q}_{i,2}^2, \mathsf{d}_{i,1}^2, \mathsf{d}_{i,2}^2 \,\big|\, \lambda_i$.



(a) (b)

Figure 1: Two Bayesian networks representing the (conditional) independence structure inherent in the system model presented in Sections II and III: (a) Structure across objects i and i' for $i' \neq i$, (b) structure of the marginal distribution of object i. In (b), for simplicity of presentation, we assume that $k_{\mathrm{S},i} = 1$ and $k_{\mathrm{E},i} = K$. Note that in some cases, conditioning on the parent nodes results in degenerate distributions (e.g., conditioned on p^*_c, $\mathsf{P}^*_{\neg c}$, and c, the random variable p_i is deterministically equal to $\boldsymbol{p}^*_{c_i}$).

B.4) $\mathsf{Z}_{k,i} \perp\!\!\!\perp \mathsf{Z}_{k',i}, \mathsf{m}_{k',i}, \mathsf{x}_{k',i} \,\big|\, \mathsf{m}_{k,i}, \mathsf{x}_{k,i}, \mathsf{d}^2_{i,1}, \mathsf{d}^2_{i,2}$, for $k' \neq k$.

B.5) $\mathsf{x}_{k,i} \perp\!\!\!\perp \mathsf{X}_{k_{\mathrm{S},i}:k-2,i} \,\big|\, \mathsf{x}_{k-1,i}, \mathsf{q}^2_{i,1}, \mathsf{q}^2_{i,2}$, for $k \geq k_{\mathrm{S},i} + 2$, where $\mathsf{X}_{k_1:k_2,i} \triangleq \big[\mathsf{x}_{k,i}\big]_{k=k_1}^{k_2}$. This is a common assumption in parameter-dependent state-space models [5, 6]. It means that the motion model described in Section II-A constitutes a Markov process, i.e., the state of an object is conditionally independent of all other past object states when given the object state at the previous time step, $\boldsymbol{x}_{k-1,i}$, and the object's driving noise covariance matrix $\boldsymbol{Q}_i = \mathrm{diag}(q^2_{i,1}\mathbf{I}_2, q^2_{i,2}\mathbf{I}_2)$.

B.6) $\mathsf{x}_{k_{\mathrm{S},i},i} \perp\!\!\!\perp \mathsf{q}^2_{i,1}, \mathsf{q}^2_{i,2}$.

B.7) $\mathsf{Z}_{k,i} \perp\!\!\!\perp \mathsf{Z}_{k_{\mathrm{S},i}:k-1,i}, \mathsf{X}_{k_{\mathrm{S},i}:k-1,i}, \mathsf{X}_{k+1:k_{\mathrm{E},i},i} \,\big|\, \mathsf{x}_{k,i}, \mathsf{p}_i$, where $\mathsf{Z}_{k_1:k_2,i} \triangleq \big[\mathsf{Z}_{k,i}\big]_{k=k_1}^{k_2}$. This is a common assumption made in the multiobject tracking literature using parameter-dependent state-space models [5, 6].

Besides these independence relations, we will also make use of the following prior distributions and likelihood functions.

P.1) The base distribution in our DP model, and thus also the prior for any $\boldsymbol{p}_i$ and $\boldsymbol{p}^*_c$, is given by (cf. (16) and (20))

$$f(\boldsymbol{p}_i) = f_{\mathrm{B}}(\boldsymbol{p}_i) = f(q^2_{i,1}, q^2_{i,2}) f(d^2_{i,1}, d^2_{i,2}) f(\lambda_i), \tag{S.1}$$

with (cf. Section III-C)

$$f(q^2_{i,1}, q^2_{i,2}) = \Gamma^{-1}\big(q^2_{i,1}; a_{\mathrm{q},1}, b_{\mathrm{q},1}\big)\, \Gamma^{-1}\big(q^2_{i,2}; a_{\mathrm{q},2}, b_{\mathrm{q},2}\big), \tag{S.2}$$

$$f(d^2_{i,1}, d^2_{i,2}) = \Gamma^{-1}\big(d^2_{i,1}; a_{\mathrm{d},1}, b_{\mathrm{d},1}\big)\, \Gamma^{-1}\big(d^2_{i,2}; a_{\mathrm{d},2}, b_{\mathrm{d},2}\big), \tag{S.3}$$

$$f(\lambda_i) = \Gamma\big(\lambda_i; a_\lambda, b_\lambda\big). \tag{S.4}$$

P.2) The initial state distribution is given as (cf. (1))

$$f(\boldsymbol{x}_{k_{\mathrm{S},i},i}) = \mathcal{N}(\boldsymbol{x}_{k_{\mathrm{S},i},i}; \mathbf{0}_{4\times 1}, \boldsymbol{\Sigma}_{\mathrm{S}}). \tag{S.5}$$

P.3) The state evolution model is given as (cf. (4))

$$f(\boldsymbol{x}_{k,i}|\boldsymbol{x}_{k-1,i}, q_{i,1}^2, q_{i,2}^2) = \mathcal{N}(\boldsymbol{x}_{k,i}; \boldsymbol{F}\boldsymbol{x}_{k-1,i}, \boldsymbol{Q}_i), \tag{S.6}$$

where we recall that $\boldsymbol{Q}_i = \mathrm{diag}(q_{i,1}^2 \mathbf{I}_2, q_{i,2}^2 \mathbf{I}_2)$.

P.4) Given λ_i, the numbers of measurements $\mathsf{m}_{k,i}$ are conditionally iid across $k \in \mathcal{K}_i$ and individually Poisson distributed (cf. Section II-B1 and (5)), i.e.,

$$p(\boldsymbol{m}_{\mathcal{K}_i,i}|\lambda_i) = \prod_{k\in\mathcal{K}_i} \mathrm{Pois}(m_{k,i}; \lambda_i). \tag{S.7}$$

P.5) The measurement model is given as (cf. (10))

$$f(\boldsymbol{Z}_{k,i}|m_{k,i}, \boldsymbol{x}_{k,i}, d_{i,1}^2, d_{i,2}^2) = \prod_{n=1}^{m_{k,i}} \mathcal{N}\big(\boldsymbol{z}_{k,i}^{(n)}; \boldsymbol{H}\boldsymbol{x}_{k,i}, \boldsymbol{R}(\boldsymbol{x}_{k,i})\boldsymbol{D}_i\boldsymbol{R}(\boldsymbol{x}_{k,i})^{\mathrm{T}}\big), \tag{S.8}$$

where $\boldsymbol{D}_i = \mathrm{diag}(d_{i,1}^2, d_{i,2}^2)$.

S2 Joint pdf $f(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i}, \boldsymbol{p}_i)$

As a preparation for deriving expressions of the conditional pdf $f(\boldsymbol{p}_c^*|\boldsymbol{X}, \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*, \boldsymbol{Z})$ in Section S3 and of the conditional pmf $p_i(c;\ell)$ in Section S4, we will first derive some other pdf and pmf expressions. These expressions arise in the course of deriving an expression of the joint pdf $f(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i}, \boldsymbol{p}_i)$, which will itself be used in Section S4.2.

We can factor the joint pdf as

$$f(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i}, \boldsymbol{p}_i) = f(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{p}_i) f(\boldsymbol{p}_i). \tag{S.9}$$

The first factor in (S.9) can be expressed as

$$f(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{p}_i) = \sum_{\boldsymbol{m}\in\mathbb{N}_0^{k_{\mathrm{E},i}-k_{\mathrm{S},i}+1}} f_{\mathsf{X}_{\mathcal{K}_i,i},\mathsf{Z}_{\mathcal{K}_i,i},\mathsf{m}_{\mathcal{K}_i,i}|\mathsf{p}_i}(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i}, \boldsymbol{m}|\boldsymbol{p}_i) \tag{S.10}$$

$$= f\big(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i} = [\mathrm{ncol}(\boldsymbol{Z}_{k,i})]_{k\in\mathcal{K}_i}\big|\boldsymbol{p}_i\big), \tag{S.11}$$

where we recall that $\mathsf{m}_{\mathcal{K}_i,i} = [\mathsf{m}_{k,i}]_{k\in\mathcal{K}_i}$ and $\mathrm{ncol}(\mathsf{Z}_{k,i})$ denotes the number of columns in the matrix $\mathsf{Z}_{k,i}$. We note that the sum in (S.10) collapses because $\mathsf{Z}_{\mathcal{K}_i,i} = \boldsymbol{Z}_{\mathcal{K}_i,i}$ implies $\mathsf{m}_{\mathcal{K}_i,i} = \boldsymbol{m}_{\mathcal{K}_i,i} = [\mathrm{ncol}(\boldsymbol{Z}_{k,i})]_{k\in\mathcal{K}_i}$ (cf. Section II-B4). From this point forward, $\boldsymbol{m}_{\mathcal{K}_i,i}$ is understood to be equal to $[\mathrm{ncol}(\boldsymbol{Z}_{k,i})]_{k\in\mathcal{K}_i}$. Thus, we write (S.11) as

$$f(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{p}_i) = f(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}|\boldsymbol{p}_i). \tag{S.12}$$

Using the chain rule, (S.12) can be further expressed as

$$f(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{p}_i) = f(\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, \boldsymbol{p}_i)\, p(\boldsymbol{m}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{p}_i)\, f(\boldsymbol{X}_{\mathcal{K}_i,i}|\boldsymbol{p}_i). \tag{S.13}$$

Inserting (S.13) and (S.1) into (S.9) then yields

$$\begin{aligned}
f(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i}, \boldsymbol{p}_i) &= f(\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i)\, p(\boldsymbol{m}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i}, q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i) \\
&\quad \times f(\boldsymbol{X}_{\mathcal{K}_i,i}|q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i)\, f(q_{i,1}^2, q_{i,2}^2) f(d_{i,1}^2, d_{i,2}^2)\, f(\lambda_i) \\
&= ABC,
\end{aligned} \tag{S.14}$$

with

$$\begin{aligned}
A &\triangleq f(\boldsymbol{X}_{\mathcal{K}_i,i}|q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i)\, f(q_{i,1}^2, q_{i,2}^2), \\
B &\triangleq f(\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i)\, f(d_{i,1}^2, d_{i,2}^2), \\
C &\triangleq p(\boldsymbol{m}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i}, q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i)\, f(\lambda_i).
\end{aligned}$$

In what follows, we derive expressions of the factors A, B, and C.

S2.1 Factor $A = f(\boldsymbol{X}_{\mathcal{K}_i,i}|q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i)\, f(q_{i,1}^2, q_{i,2}^2)$

Using B.1, the factor A can be simplified to

$$A = f(\boldsymbol{X}_{\mathcal{K}_i,i}|q_{i,1}^2, q_{i,2}^2)\, f(q_{i,1}^2, q_{i,2}^2). \tag{S.15}$$

Using the chain rule, the first factor in (S.15) is given as

$$\begin{aligned}
f(\boldsymbol{X}_{\mathcal{K}_i,i}|q_{i,1}^2, q_{i,2}^2) &= f(\boldsymbol{x}_{k_{\mathrm{S},i},i}|q_{i,1}^2, q_{i,2}^2) \prod_{k=k_{\mathrm{S},i}+1}^{k_{\mathrm{E},i}} f(\boldsymbol{x}_{k,i}|\boldsymbol{X}_{k_{\mathrm{S},i}:k-1,i}, q_{i,1}^2, q_{i,2}^2) \\
&= f(\boldsymbol{x}_{k_{\mathrm{S},i},i}) \prod_{k=k_{\mathrm{S},i}+1}^{k_{\mathrm{E},i}} f(\boldsymbol{x}_{k,i}|\boldsymbol{x}_{k-1,i}, q_{i,1}^2, q_{i,2}^2),
\end{aligned} \tag{S.16}$$

where B.6 and B.5 were used in the last step. Inserting (S.5) and (S.6) into (S.16), we obtain

$$f(\boldsymbol{X}_{\mathcal{K}_i,i}|q_{i,1}^2, q_{i,2}^2) = \mathcal{N}(\boldsymbol{x}_{k_{\mathrm{S},i},i}; \mathbf{0}_{4\times1}, \boldsymbol{\Sigma}_{\mathrm{S}}) \prod_{k=k_{\mathrm{S},i}+1}^{k_{\mathrm{E},i}} \mathcal{N}(\boldsymbol{x}_{k,i}; \boldsymbol{F}\boldsymbol{x}_{k-1,i}, \boldsymbol{Q}_i). \tag{S.17}$$

Substituting $\boldsymbol{u}_{k,i} = \boldsymbol{x}_{k,i} - \boldsymbol{F}\boldsymbol{x}_{k-1,i}$ (cf. (2)) and noting that $\mathcal{N}(\boldsymbol{x}_{k,i}; \boldsymbol{F}\boldsymbol{x}_{k-1,i}, \boldsymbol{Q}_i) = \mathcal{N}(\boldsymbol{u}_{k,i}; \mathbf{0}_{4\times1}, \boldsymbol{Q}_i)$, and recalling that $\boldsymbol{Q}_i = \mathrm{diag}(q_{i,1}^2\mathbf{I}_2, q_{i,2}^2\mathbf{I}_2)$, we further obtain

$$\begin{aligned}
f(\boldsymbol{X}_{\mathcal{K}_i,i}|q_{i,1}^2, q_{i,2}^2) &= \mathcal{N}(\boldsymbol{x}_{k_{\mathrm{S},i},i}; \mathbf{0}_{4\times1}, \boldsymbol{\Sigma}_{\mathrm{S}}) \prod_{k=k_{\mathrm{S},i}+1}^{k_{\mathrm{E},i}} \mathcal{N}(\boldsymbol{u}_{k,i}; \mathbf{0}_{4\times1}, \boldsymbol{Q}_i) \\
&= \mathcal{N}(\boldsymbol{x}_{k_{\mathrm{S},i},i}; \mathbf{0}_{4\times1}, \boldsymbol{\Sigma}_{\mathrm{S}}) \prod_{k=k_{\mathrm{S},i}+1}^{k_{\mathrm{E},i}} \prod_{\kappa=1}^{2} \mathcal{N}(u_{k,i,\kappa}; 0, q_{i,1}^2)\, \mathcal{N}(u_{k,i,\kappa+2}; 0, q_{i,2}^2),
\end{aligned} \tag{S.18}$$

where $u_{k,i,1}$ through $u_{k,i,4}$ are the elements of $\boldsymbol{u}_{k,i} = \boldsymbol{x}_{k,i} - \boldsymbol{F}\boldsymbol{x}_{k-1,i}$. Inserting (S.18) and (S.2) into (S.15) then gives

$$\begin{aligned}
A &= \mathcal{N}(\boldsymbol{x}_{k_{\mathrm{S},i},i}; \mathbf{0}_{4\times1}, \boldsymbol{\Sigma}_{\mathrm{S}}) \left(\prod_{k=k_{\mathrm{S},i}+1}^{k_{\mathrm{E},i}} \prod_{\kappa=1}^{2} \mathcal{N}(u_{k,i,\kappa}; 0, q_{i,1}^2)\, \mathcal{N}(u_{k,i,\kappa+2}; 0, q_{i,2}^2) \right) \\
&\quad \times \Gamma^{-1}(q_{i,1}^2; a_{\mathrm{q},1}, b_{\mathrm{q},1})\, \Gamma^{-1}(q_{i,2}^2; a_{\mathrm{q},2}, b_{\mathrm{q},2}).
\end{aligned} \tag{S.19}$$

Because the inverse gamma distribution is a conjugate prior for a Gaussian likelihood function with unknown variance [7, Section 2.6], this is proportional to the product of two inverse gamma distributions. However, for our subsequent development (see Section S4.2), we require equality and not merely proportionality, .

S2.2 Factor $B = f(\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i)\, f(d_{i,1}^2, d_{i,2}^2)$

For the factor B, we obtain

$$B = f(\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, d_{i,1}^2, d_{i,2}^2)\, f(d_{i,1}^2, d_{i,2}^2), \tag{S.20}$$

where the removal of $q_{i,1}^2$, $q_{i,2}^2$, and λ_i from the condition set of $f(\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i)$ follows from B.2. Using the chain rule, the first factor in (S.20) becomes

$$\begin{aligned} f(\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, d_{i,1}^2, d_{i,2}^2) &= f(\boldsymbol{Z}_{k_{\mathrm{S},i},i}|\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, d_{i,1}^2, d_{i,2}^2) \\ &\quad \times \prod_{k=k_{\mathrm{S},i}+1}^{k_{\mathrm{E},i}} f(\boldsymbol{Z}_{k,i}|\, \boldsymbol{Z}_{k_{\mathrm{S},i}:k-1,i}, \boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, d_{i,1}^2, d_{i,2}^2). \end{aligned} \tag{S.21}$$

Using B.4, we have $f(\boldsymbol{Z}_{k_{\mathrm{S},i},i}|\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, d_{i,1}^2, d_{i,2}^2) = f(\boldsymbol{Z}_{k_{\mathrm{S},i},i}|\boldsymbol{x}_{k_{\mathrm{S},i},i}, m_{k_{\mathrm{S},i},i}, d_{i,1}^2, d_{i,2}^2)$ and $f(\boldsymbol{Z}_{k,i}|\, \boldsymbol{Z}_{k_{\mathrm{S},i}:k-1,i}, \boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, d_{i,1}^2, d_{i,2}^2) = f(\boldsymbol{Z}_{k,i}|\boldsymbol{x}_{k,i}, m_{k,i}, d_{i,1}^2, d_{i,2}^2)$, and thus expression (S.21) simplifies to

$$f(\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, d_{i,1}^2, d_{i,2}^2) = \prod_{k\in\mathcal{K}_i} f(\boldsymbol{Z}_{k,i}|\boldsymbol{x}_{k,i}, m_{k,i}, d_{i,1}^2, d_{i,2}^2). \tag{S.22}$$

Inserting (S.8) into (S.22), we obtain

$$f(\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, d_{i,1}^2, d_{i,2}^2) = \prod_{k\in\mathcal{K}_i} \prod_{n=1}^{m_{k,i}} \mathcal{N}\big(\boldsymbol{z}_{k,i}^{(n)}; \boldsymbol{H}\boldsymbol{x}_{k,i}, \boldsymbol{R}(\boldsymbol{x}_{k,i})\boldsymbol{D}_i\boldsymbol{R}(\boldsymbol{x}_{k,i})^{\mathrm{T}}\big).$$

Substituting $\boldsymbol{v}_{k,i}^{(n)} = \boldsymbol{R}(\boldsymbol{x}_{k,i})^{-1}\big(\boldsymbol{z}_{k,i}^{(n)} - \boldsymbol{H}\boldsymbol{x}_{k,i}\big)$ (cf. (6)) and noting that $\mathcal{N}\big(\boldsymbol{z}_{k,i}^{(n)}; \boldsymbol{H}\boldsymbol{x}_{k,i}, \boldsymbol{R}(\boldsymbol{x}_{k,i})\boldsymbol{D}_i\boldsymbol{R}(\boldsymbol{x}_{k,i})^{\mathrm{T}}\big) = \mathcal{N}\big(\boldsymbol{v}_{k,i}^{(n)}; \boldsymbol{0}_{2\times1}, \boldsymbol{D}_i\big)$, and recalling that $\boldsymbol{D}_i = \mathrm{diag}(d_{i,1}^2, d_{i,2}^2)$, we further obtain

$$f(\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, d_{i,1}^2, d_{i,2}^2) = \prod_{k\in\mathcal{K}_i} \prod_{n=1}^{m_{k,i}} \mathcal{N}\big(\boldsymbol{v}_{k,i}^{(n)}; \boldsymbol{0}_{2\times1}, \boldsymbol{D}_i\big) = \prod_{k\in\mathcal{K}_i} \prod_{n=1}^{m_{k,i}} \prod_{\kappa=1}^{2} \mathcal{N}\big(v_{k,i,\kappa}^{(n)}; 0, d_{i,\kappa}^2\big), \tag{S.23}$$

where $v_{k,i,1}^{(n)}$ and $v_{k,i,2}^{(n)}$ are the elements of $\boldsymbol{v}_{k,i}^{(n)}$. Finally, inserting (S.23) and (S.3) into (S.20) yields

$$B = \left(\prod_{k\in\mathcal{K}_i} \prod_{n=1}^{m_{k,i}} \prod_{\kappa=1}^{2} \mathcal{N}\big(v_{k,i,\kappa}^{(n)}; 0, d_{i,\kappa}^2\big) \right) \Gamma^{-1}(d_{i,1}^2; a_{\mathrm{d},1}, b_{\mathrm{d},1})\, \Gamma^{-1}(d_{i,2}^2; a_{\mathrm{d},2}, b_{\mathrm{d},2}). \tag{S.24}$$

Again, the proportionality to a product of two inverse gamma distributions is not useful at this point (see Section S4.2).

S2.3 Factor $C = p(\boldsymbol{m}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i}, q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i)\, f(\lambda_i)$

Using B.3, the factor C simplifies to

$$C = p(\boldsymbol{m}_{\mathcal{K}_i,i}|\lambda_i)\, f(\lambda_i).$$

By (S.7) and (S.4), we further obtain

$$C = \left(\prod_{k \in \mathcal{K}_i} \mathrm{Pois}(m_{k,i}; \lambda_i) \right) \Gamma\big(\lambda_i; a_\lambda, b_\lambda\big). \tag{S.25}$$

At this point, we could use the fact that the gamma distribution is a conjugate prior for a Poisson likelihood function with unknown rate parameter λ_i [7, Section 2.6]. However, this is not helpful for our subsequent development (see Section S4.2).

S3 Conditional pdf $f(\boldsymbol{p}_c^* | \boldsymbol{X}, \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*, \boldsymbol{Z})$

Next, we will derive expressions (38)–(42) for the conditional pdf $f(\boldsymbol{p}_c^* | \boldsymbol{X}, \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*, \boldsymbol{Z})$ involved in the Gibbs sampler (see Sections IV-C2 and V-B2). Throughout this section, we will use the symbol $\propto$ to designate proportionality up to a constant independent of $\boldsymbol{p}_c^* = \big[q_{c,1}^{*2}\ q_{c,2}^{*2}\ d_{c,1}^{*2}\ d_{c,2}^{*2}\ \lambda_c^*\big]^{\mathrm{T}}$.

Using Bayes' theorem, we have

$$f(\boldsymbol{p}_c^* | \boldsymbol{X}, \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*, \boldsymbol{Z}) \propto f(\boldsymbol{X}, \boldsymbol{Z} | \boldsymbol{p}_c^*, \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*) f(\boldsymbol{p}_c^* | \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*). \tag{S.26}$$

Using A.2, the first factor in (S.26) factorizes according to

$$f(\boldsymbol{X}, \boldsymbol{Z} | \boldsymbol{p}_c^*, \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*) = \prod_{i=1}^{I} f(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i} | \boldsymbol{p}_c^*, \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*). \tag{S.27}$$

Since $\mathbf{p}_i$ is deterministically fixed as $\boldsymbol{p}_i = \boldsymbol{p}_{c_i}$ given $\boldsymbol{p}_c^*$, $\boldsymbol{c}$, and $\boldsymbol{P}_{\neg c}^*$, we can add $\boldsymbol{p}_i$ as a further condition in each factor on the right-hand side, i.e.,

$$f(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i} | \boldsymbol{p}_c^*, \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*) = f(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i} | \boldsymbol{p}_i = \boldsymbol{p}_{c_i}^*, \boldsymbol{p}_c^*, \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*) = f(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i} | \boldsymbol{p}_i = \boldsymbol{p}_{c_i}^*), \tag{S.28}$$

where we used A.3 in the last equality. Inserting (S.28) into (S.27), we obtain further

$$f(\boldsymbol{X}, \boldsymbol{Z} | \boldsymbol{p}_c^*, \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*) = \prod_{i=1}^{I} f(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i} | \boldsymbol{p}_i = \boldsymbol{p}_{c_i}^*) \propto \prod_{i:\, c_i = c} f(\boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{Z}_{\mathcal{K}_i,i} | \boldsymbol{p}_i = \boldsymbol{p}_c^*), \tag{S.29}$$

where we used the fact that the product of all factors with $c_i \neq c$ is a constant independent of $\boldsymbol{p}_c^*$ (recall that our goal is to find $f(\boldsymbol{p}_c^* | \boldsymbol{X}, \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*, \boldsymbol{Z})$, i.e., a pdf in $\boldsymbol{p}_c^*$. In view of expression (S.29), from this point forward, we only consider i such that $c_i = c$ and, hence, $\boldsymbol{p}_i$ is always considered to be $\boldsymbol{p}_c^*$. Inserting (S.13) into (S.29), we obtain further

$$f(\boldsymbol{X}, \boldsymbol{Z} | \boldsymbol{p}_c^*, \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*) = \prod_{i:\, c_i = c} f(\boldsymbol{Z}_{\mathcal{K}_i,i} | \boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, \boldsymbol{p}_i)\, p(\boldsymbol{m}_{\mathcal{K}_i,i} | \boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{p}_i)\, f(\boldsymbol{X}_{\mathcal{K}_i,i} | \boldsymbol{p}_i). \tag{S.30}$$

Next, using A.1, the second factor in (S.26) simplifies as

$$f(\boldsymbol{p}_c^* | \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*) = f(\boldsymbol{p}_c^*).$$

Using P.1, we obtain further

$$f(\boldsymbol{p}_c^* | \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*) = f(q_{c,1}^{*2}, q_{c,2}^{*2}) f(d_{c,1}^{*2}, d_{c,2}^{*2}) f(\lambda_c^*). \tag{S.31}$$

Finally, inserting (S.30) and (S.31) into (S.26) yields

$$
\begin{aligned}
& f(\boldsymbol{p}_c^* \,|\, \boldsymbol{X}, \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*, \boldsymbol{Z}) \\
& \propto \Bigg(\prod_{i:\, c_i = c} f(\boldsymbol{Z}_{\mathcal{K}_i,i} | \boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i)\, p(\boldsymbol{m}_{\mathcal{K}_i,i} | \boldsymbol{X}_{\mathcal{K}_i,i}, q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i) \\
& \qquad \times f(\boldsymbol{X}_{\mathcal{K}_i,i} | q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i) \Bigg) f(q_{c,1}^{*2}, q_{c,2}^{*2})\, f(d_{c,1}^{*2}, d_{c,2}^{*2})\, f(\lambda_c^*) \\
& = A' B' C',
\end{aligned}
\tag{S.32}
$$

with

$$
\begin{aligned}
A' &\triangleq \Bigg(\prod_{i:\, c_i = c} f(\boldsymbol{X}_{\mathcal{K}_i,i} | q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i) \Bigg) f(q_{c,1}^{*2}, q_{c,2}^{*2}), \\
B' &\triangleq \Bigg(\prod_{i:\, c_i = c} f(\boldsymbol{Z}_{\mathcal{K}_i,i} | \boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i) \Bigg) f(d_{c,1}^{*2}, d_{c,2}^{*2}), \\
C' &\triangleq \Bigg(\prod_{i:\, c_i = c} p(\boldsymbol{m}_{\mathcal{K}_i,i} | \boldsymbol{X}_{\mathcal{K}_i,i}, q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i) \Bigg) f(\lambda_c^*).
\end{aligned}
$$

In these expressions, as mentioned previously, $q_{i,1}^2 = q_{c,1}^{*2}$ etc. In what follows, we derive expressions of the factors A', B', and C'.

S3.1 Factor $A' = \big(\prod_{i:\, c_i = c} f(\boldsymbol{X}_{\mathcal{K}_i,i} | q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i) \big) f(q_{c,1}^{*2}, q_{c,2}^{*2})$

Using B.1, the factor A' can be simplified to

$$
A' = \Bigg(\prod_{i:\, c_i = c} f(\boldsymbol{X}_{\mathcal{K}_i,i} | q_{i,1}^2, q_{i,2}^2) \Bigg) f(q_{c,1}^{*2}, q_{c,2}^{*2}).
$$

Inserting (S.18), using P.1, and recalling that $q_{i,1}^2 = q_{c,1}^{*2}$ and $q_{i,2}^2 = q_{c,2}^{*2}$, we obtain

$$
A' \propto \Bigg(\prod_{i:\, c_i = c} \prod_{k = k_{\mathrm{S},i}+1}^{k_{\mathrm{E},i}} \prod_{\kappa=1}^{2} \mathcal{N}(u_{k,i,\kappa}; 0, q_{c,1}^{*2})\, \mathcal{N}(u_{k,i,\kappa+2}; 0, q_{c,2}^{*2}) \Bigg) \Gamma^{-1}(q_{c,1}^{*2}; a_{\mathrm{q},1}, b_{\mathrm{q},1})\, \Gamma^{-1}(q_{c,2}^{*2}; a_{\mathrm{q},2}, b_{\mathrm{q},2}). \tag{S.33}
$$

The product of Gaussian distributions in (S.33) can be interpreted as a likelihood function for the unknown variances $\mathrm{q}_{c,1}^{*2}$ and $\mathrm{q}_{c,2}^{*2}$. Using the fact that the inverse gamma distribution is the conjugate prior for such a likelihood function [7, Section 2.6], we obtain the closed-form expression

$$
A' \propto \Gamma^{-1}(q_{c,1}^{*2}; a'_{\mathrm{q},c,1}, b'_{\mathrm{q},c,1})\, \Gamma^{-1}(q_{c,2}^{*2}; a'_{\mathrm{q},c,2}, b'_{\mathrm{q},c,2}), \tag{S.34}
$$

with parameters

$$
a'_{\mathrm{q},c,\kappa} = a_{\mathrm{q},\kappa} + K(c) - N(c), \quad b'_{\mathrm{q},c,\kappa} = b_{\mathrm{q},\kappa} + \sum_{i:\, c_i = c} \sum_{k = k_{\mathrm{S},i}+1}^{k_{\mathrm{E},i}} \overline{u_{k,i,\kappa}^2}, \quad \kappa = 1, 2, \tag{S.35}
$$

where $K(c)$ is the number of time steps k in which at least one existing object belongs to class c, $N(c)$ is the number of c_i that are equal to c, $\overline{u_{k,i,1}^2} \triangleq (u_{k,i,1}^2 + u_{k,i,2}^2)/2$, and $\overline{u_{k,i,2}^2} \triangleq (u_{k,i,3}^2 + u_{k,i,4}^2)/2$.

S3.2 Factor $B' = \big(\prod_{i\,:\,c_i=c} f(\boldsymbol{Z}_{\mathcal{K}_i,i} | \boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i) \big) f(d_{c,1}^{*2}, d_{c,2}^{*2})$

Using B.2, the factor B' can be simplified to

$$B' = \left(\prod_{i\,:\,c_i=c} f(\boldsymbol{Z}_{\mathcal{K}_i,i} | \boldsymbol{X}_{\mathcal{K}_i,i}, \boldsymbol{m}_{\mathcal{K}_i,i}, d_{i,1}^2, d_{i,2}^2) \right) f(d_{c,1}^{*2}, d_{c,2}^{*2}).$$

Inserting (S.23), using P.1, and recalling that $d_{i,1}^2 = d_{c,1}^{*2}$ and $d_{i,2}^2 = d_{c,2}^{*2}$, we obtain

$$B' = \left(\prod_{i\,:\,c_i=c} \prod_{k\in\mathcal{K}_i} \prod_{n=1}^{m_{k,i}} \prod_{\kappa=1}^{2} \mathcal{N}\big(v_{k,i,\kappa}^{(n)}; 0, d_{c,\kappa}^{*2}\big) \right) \Gamma^{-1}(d_{c,1}^{*2}; a_{\mathrm{d},1}, b_{\mathrm{d},1})\, \Gamma^{-1}(d_{c,2}^{*2}; a_{\mathrm{d},2}, b_{\mathrm{d},2}).$$

Again using the fact that the inverse gamma distribution is a conjugate prior for a Gaussian likelihood function with unknown variance [7, Section 2.6], we obtain

$$B' \propto \Gamma^{-1}(d_{c,1}^{*2}; a'_{\mathrm{d},c,1}, b'_{\mathrm{d},c,1})\, \Gamma^{-1}(d_{c,2}^{*2}; a'_{\mathrm{d},c,2}, b'_{\mathrm{d},c,2}), \tag{S.36}$$

with parameters

$$a'_{\mathrm{d},c,\kappa} = a_{\mathrm{d},\kappa} + \frac{1}{2} \sum_{i:c_i=c} M^{(i)}, \quad b'_{\mathrm{d},c,\kappa} = b_{\mathrm{d},\kappa} + \frac{1}{2} \sum_{i:c_i=c} \sum_{k\in\mathcal{K}_i} \sum_{n=1}^{m_{k,i}} (v_{k,i,\kappa}^{(n)})^2, \quad \kappa = 1, 2, \tag{S.37}$$

where

$$M^{(i)} \triangleq \sum_{k\in\mathcal{K}_i} m_{k,i}.$$

S3.3 Factor $C' = \big(\prod_{i\,:\,c_i=c} p(\boldsymbol{m}_{\mathcal{K}_i,i} | \boldsymbol{X}_{\mathcal{K}_i,i}, q_{i,1}^2, q_{i,2}^2, d_{i,1}^2, d_{i,2}^2, \lambda_i) \big) f(\lambda_c^*)$

Using B.3, the factor C' simplifies to

$$C' = \left(\prod_{i\,:\,c_i=c} p(\boldsymbol{m}_{\mathcal{K}_i,i} | \lambda_i) \right) f(\lambda_c^*).$$

Inserting (S.7), using P.1, and recalling that $\lambda_i = \lambda_c^*$, we further obtain

$$C' = \left(\prod_{i\,:\,c_i=c} \prod_{k\in\mathcal{K}_i} \mathrm{Pois}(m_{k,i}; \lambda_c^*) \right) \Gamma\big(\lambda_c^*; a_\lambda, b_\lambda\big). \tag{S.38}$$

The product of Poisson distributions in (S.38) can be interpreted as a Poisson likelihood function with unknown rate parameter λ_c^*. Using the fact that the gamma distribution is a conjugate prior for such a likelihood function [7, Section 2.6], it follows that

$$C' \propto \Gamma(\lambda_c^*; a'_{\lambda,c}, b'_{\lambda,c}), \tag{S.39}$$

with parameters

$$a'_{\lambda,c} = a_\lambda + \sum_{i:c_i=c} M^{(i)}, \qquad b'_{\lambda,c} = b_\lambda + N(c). \tag{S.40}$$

S3.4 Final Expression of $f(\boldsymbol{p}_c^* | \boldsymbol{X}, \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*, \boldsymbol{Z})$

Inserting (S.34), (S.36), and (S.39) into (S.32), we finally arrive at

$$\begin{aligned} f(\boldsymbol{p}_c^* | \boldsymbol{X}, \boldsymbol{c}, \boldsymbol{P}_{\neg c}^*, \boldsymbol{Z}) \propto\; & \Gamma^{-1}(q_{c,1}^{*2}; a'_{\mathrm{q},c,1}, b'_{\mathrm{q},c,1})\, \Gamma^{-1}(q_{c,2}^{*2}; a'_{\mathrm{q},c,2}, b'_{\mathrm{q},c,2})\, \Gamma^{-1}(d_{c,1}^{*2}; a'_{\mathrm{d},c,1}, b'_{\mathrm{d},c,1}) \\ & \times \Gamma^{-1}(d_{c,2}^{*2}; a'_{\mathrm{d},c,2}, b'_{\mathrm{d},c,2})\, \Gamma(\lambda_c^*; a'_{\lambda,c}, b'_{\lambda,c}), \end{aligned} \tag{S.41}$$

with the parameters $a'_{\mathrm{q},c,1}$ etc. given by (S.35), (S.37), and (S.40). This result is identical to (38)–(43).

S4 Conditional pmf $p_i(c;\ell)$

We will derive the expressions (36) and (37) for the conditional pmf $p_i(c;\ell)$ involved in the Gibbs sampler (see Sections IV-C1 and V-B1). We recall from (27) that $p_i(c;\ell)$ is the evaluation of $p(c_i\,|\,\boldsymbol{c}_{\neg i},\boldsymbol{P}^*_{\boldsymbol{c}_{\neg i}},\boldsymbol{X},\boldsymbol{Z})$ for specific ℓ-dependent arguments. The central observation now is that the paired matrices $(\mathbf{X}_{\mathcal{K}_i,i},\mathbf{Z}_{\mathcal{K}_i,i})_{i=1}^{I}$ can be interpreted as draws from a DPM, and it is known [8, Eq. (3.6)] that

$$p(c_i\,|\,\boldsymbol{c}_{\neg i},\boldsymbol{P}^*_{\boldsymbol{c}_{\neg i}},\boldsymbol{X},\boldsymbol{Z}) = \begin{cases} C\,\frac{1}{\alpha+I-1}\,N_{c_i}(\boldsymbol{c}_{\neg i})\,f(\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{p}_i\!=\!\boldsymbol{p}^*_{c_i}), & c_i\in\mathcal{C}_{\neg i},\\ C\,\frac{\alpha}{\alpha+I-1}\displaystyle\int_{\mathbb{R}^5_+} f(\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{p}_i)\,f_{\mathrm{B}}(\boldsymbol{p}_i)\,\mathrm{d}\boldsymbol{p}_i, & c_i\notin\mathcal{C}_{\neg i}, \end{cases} \tag{S.42}$$

where C is a normalization constant and $N_{c_i}(\boldsymbol{c}_{\neg i})$ is the number of elements in $\boldsymbol{c}_{\neg i}$ equal to c_i. (Note that the expression for the second case of (S.42) depends on $\boldsymbol{c}_{\neg i}$ and $\boldsymbol{P}^*_{\boldsymbol{c}_{\neg i}}$ only via C.) It remains to find expressions of the pdfs involved in (S.42).

S4.1 Conditional pmf $p(c_i\,|\,\boldsymbol{c}_{\neg i},\boldsymbol{P}^*_{\boldsymbol{c}_{\neg i}},\boldsymbol{X},\boldsymbol{Z})$ for $c_i\in\mathcal{C}_{\neg i}$

The pdf $f(\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{p}_i)$ appearing in the first case ($c_i\in\mathcal{C}_{\neg i}$) in (S.42) can be factored as

$$f(\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{p}_i) = f(\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{p}_i)\,f(\boldsymbol{X}_{\mathcal{K}_i,i}|\boldsymbol{p}_i). \tag{S.43}$$

Using the chain rule, the factor $f(\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{p}_i)$ can be expanded as

$$f(\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{p}_i) = f(\boldsymbol{Z}_{k_{\mathrm{S},i},i}|\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{p}_i)\prod_{k=k_{\mathrm{S},i}+1}^{k_{\mathrm{E},i}} f(\boldsymbol{Z}_{k,i}|\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{Z}_{k_{\mathrm{S},i}:k-1,i},\boldsymbol{p}_i). \tag{S.44}$$

By B.7, we have $f(\boldsymbol{Z}_{k_{\mathrm{S},i},i}|\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{p}_i) = f(\boldsymbol{Z}_{k_{\mathrm{S},i},i}|\boldsymbol{x}_{k_{\mathrm{S},i},i},\boldsymbol{p}_i)$ and $f(\boldsymbol{Z}_{k,i}|\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{Z}_{k_{\mathrm{S},i}:k-1,i},\boldsymbol{p}_i) = f(\boldsymbol{Z}_{k,i}|\boldsymbol{x}_{k,i},\boldsymbol{p}_i)$, and thus expression (S.44) simplifies to

$$f(\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{p}_i) = \prod_{k\in\mathcal{K}_i} f(\boldsymbol{Z}_{k,i}|\boldsymbol{x}_{k,i},\boldsymbol{p}_i). \tag{S.45}$$

For the factor $f(\boldsymbol{X}_{\mathcal{K}_i,i}|\boldsymbol{p}_i)$ in (S.43), B.1 implies

$$f(\boldsymbol{X}_{\mathcal{K}_i,i}|\boldsymbol{p}_i) = f(\boldsymbol{X}_{\mathcal{K}_i,i}|q^2_{i,1},q^2_{i,2}),$$

and using (S.17), we obtain further

$$f(\boldsymbol{X}_{\mathcal{K}_i,i}|\boldsymbol{p}_i) = \mathcal{N}(\boldsymbol{x}_{k_{\mathrm{S},i},i};\mathbf{0}_{4\times1},\boldsymbol{\Sigma}_{\mathrm{S}})\prod_{k=k_{\mathrm{S},i}+1}^{k_{\mathrm{E},i}}\mathcal{N}(\boldsymbol{x}_{k,i};\boldsymbol{F}\boldsymbol{x}_{k-1,i},\boldsymbol{Q}_i). \tag{S.46}$$

We now insert (S.45) and (S.46) into (S.43) and obtain

$$f(\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{p}_i) = \left(\prod_{k\in\mathcal{K}_i} f(\boldsymbol{Z}_{k,i}|\boldsymbol{x}_{k,i},\boldsymbol{p}_i)\right)\mathcal{N}(\boldsymbol{x}_{k_{\mathrm{S},i},i};\mathbf{0}_{4\times1},\boldsymbol{\Sigma}_{\mathrm{S}})\prod_{k'=k_{\mathrm{S},i}+1}^{k_{\mathrm{E},i}}\mathcal{N}(\boldsymbol{x}_{k',i};\boldsymbol{F}\boldsymbol{x}_{k'-1,i},\boldsymbol{Q}^*_{c_i}). \tag{S.47}$$

Finally, inserting (S.47) into the first case in (S.42) gives

$$p(c_i\,|\,\boldsymbol{c}_{\neg i},\boldsymbol{P}^*_{\boldsymbol{c}_{\neg i}},\boldsymbol{X},\boldsymbol{Z}) = C\,\frac{1}{\alpha+I-1}\,N_{c_i}(\boldsymbol{c}_{\neg i})\left(\prod_{k\in\mathcal{K}_i} f(\boldsymbol{Z}_{k,i}\,|\,\boldsymbol{x}_{k,i},\boldsymbol{p}_i\!=\!\boldsymbol{p}^*_{c_i})\right)\mathcal{N}(\boldsymbol{x}_{k_{\mathrm{S},i},i};\boldsymbol{0}_{4\times 1},\boldsymbol{\Sigma}_{\mathrm{S}})$$
$$\times\prod_{k'=k_{\mathrm{S},i}+1}^{k_{\mathrm{E},i}}\mathcal{N}(\boldsymbol{x}_{k',i};\boldsymbol{F}\boldsymbol{x}_{k'-1,i},\boldsymbol{Q}^*_{c_i}),$$

for $c_i\in\mathcal{C}_{\neg i}$. This expression is consistent with (36) (via (27)).

S4.2 Conditional pmf $p(c_i\,|\,\boldsymbol{c}_{\neg i},\boldsymbol{P}^*_{\boldsymbol{c}_{\neg i}},\boldsymbol{X},\boldsymbol{Z})$ for $c_i\notin\mathcal{C}_{\neg i}$

To calculate the integral $\int_{\mathbb{R}^5_+} f(\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{p}_i)\,f_{\mathrm{B}}(\boldsymbol{p}_i)\,\mathrm{d}\boldsymbol{p}_i$ appearing in the second case ($c_i\notin\mathcal{C}_{\neg i}$) in (S.42), we will again use the conjugacy of the prior $f_{\mathrm{B}}(\boldsymbol{p}_i)$. The integrand can be written as

$$f(\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{p}_i)\,f_{\mathrm{B}}(\boldsymbol{p}_i) = f(\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{p}_i)\,f(\boldsymbol{p}_i) = f(\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{Z}_{\mathcal{K}_i,i},\boldsymbol{p}_i),$$

where (20) was used. In (S.14), we expressed $f(\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{Z}_{\mathcal{K}_i,i},\boldsymbol{p}_i)$ as the product of three factors A, B, and C, which are given by (S.19), (S.24), and (S.25), respectively. After inserting these expressions, the result can be organized into five factors that depend on $\boldsymbol{p}_i$ only via, respectively, $q^2_{i,1}$, $q^2_{i,2}$, $d^2_{i,1}$, $d^2_{i,2}$, and λ_i. As a consequence, our integral can be rewritten as the product of five one-dimensional integrals and another factor, according to

$$\int_{\mathbb{R}^5_+} f(\boldsymbol{X}_{\mathcal{K}_i,i},\boldsymbol{Z}_{\mathcal{K}_i,i}|\boldsymbol{p}_i)\,f_{\mathrm{B}}(\boldsymbol{p}_i)\,\mathrm{d}\boldsymbol{p}_i = A_1A_2B_1B_2C_1D, \tag{S.48}$$

where

$$A_1 \triangleq \int_{\mathbb{R}_+}\left(\prod_{k=k_{\mathrm{S},i}+1}^{k_{\mathrm{E},i}}\prod_{\kappa=1}^{2}\mathcal{N}(u_{k,i,\kappa};0,q^2_{i,1})\right)\Gamma^{-1}\big(q^2_{i,1};a_{\mathrm{q},1},b_{\mathrm{q},1}\big)\,\mathrm{d}q^2_{i,1}, \tag{S.49}$$

$$A_2 \triangleq \int_{\mathbb{R}_+}\left(\prod_{k=k_{\mathrm{S},i}+1}^{k_{\mathrm{E},i}}\prod_{\kappa=1}^{2}\mathcal{N}(u_{k,i,\kappa+2};0,q^2_{i,2})\right)\Gamma^{-1}\big(q^2_{i,2};a_{\mathrm{q},2},b_{\mathrm{q},2}\big)\,\mathrm{d}q^2_{i,2}, \tag{S.50}$$

$$A_3 \triangleq \mathcal{N}(\boldsymbol{x}_{k_{\mathrm{S},i},i};\boldsymbol{0}_{4\times 1},\boldsymbol{\Sigma}_{\mathrm{S}}), \tag{S.51}$$

$$B_1 \triangleq \int_{\mathbb{R}_+}\left(\prod_{k\in\mathcal{K}_i}\prod_{n=1}^{m_{k,i}}\mathcal{N}\big(v^{(n)}_{k,i,1};0,d^2_{i,1}\big)\right)\Gamma^{-1}\big(d^2_{i,1};a_{\mathrm{d},1},b_{\mathrm{d},1}\big)\,\mathrm{d}d^2_{i,1}, \tag{S.52}$$

$$B_2 \triangleq \int_{\mathbb{R}_+}\left(\prod_{k\in\mathcal{K}_i}\prod_{n=1}^{m_{k,i}}\mathcal{N}\big(v^{(n)}_{k,i,2};0,d^2_{i,2}\big)\right)\Gamma^{-1}\big(d^2_{i,2};a_{\mathrm{d},2},b_{\mathrm{d},2}\big)\,\mathrm{d}d^2_{i,2}, \tag{S.53}$$

$$C_1 \triangleq \int_{\mathbb{R}_+}\left(\prod_{k\in\mathcal{K}_i}\mathrm{Pois}(m_{k,i};\lambda_i)\right)\Gamma\big(\lambda_i;a_\lambda,b_\lambda\big)\,\mathrm{d}\lambda_i. \tag{S.54}$$

Note that, in particular, the factors A_1, A_2, and A_3 are due to the factor A in (S.19), the factors B_1 and B_2 are due to the factor B in (S.24), and the factor C_1 is due to the factor C in (S.25).

Factors $A_1, A_2, B_1,$ and B_2. The integrals A_1, A_2, B_1, and B_2 all have the same structure, being the product of multiple Gaussian pdfs times an inverse gamma pdf. Thus, we first calculate these integrals in a generic way. While for a single Gaussian pdf this would be the classical compound Gaussian construction of a t-distribution, we did not find a similar construction in the literature for this more general setting of multiple Gaussians, and therefore we provide the details here. In what follows, we consider a generic factor $\tilde{A}$ with observations $y_{k,n}$ for $k \in \tilde{\mathcal{K}} = [\tilde{k}_{\mathrm{S}}, \tilde{k}_{\mathrm{E}}] \subseteq \{1, \dots, K\}$ and $n = 1, 2, \dots, \tilde{m}_k$, variance σ^2, shape parameter a, and scale parameter b, i.e.,

$$\tilde{A} \triangleq \int_{\mathbb{R}_+} \Bigg(\prod_{k \in \tilde{\mathcal{K}}} \prod_{n=1}^{\tilde{m}_k} \mathcal{N}(y_{k,n}; 0, \sigma^2) \Bigg) \Gamma^{-1}(\sigma^2; a, b) \, \mathrm{d}\sigma^2.$$

We obtain

$$\begin{aligned}
\tilde{A} &= \int_{\mathbb{R}_+} \Bigg(\prod_{k \in \tilde{\mathcal{K}}} \prod_{n=1}^{\tilde{m}_k} \frac{1}{\sqrt{2\pi\sigma^2}} \, \mathrm{e}^{-y_{k,n}^2/(2\sigma^2)} \Bigg) \frac{b^a}{\Gamma(a)} \, (\sigma^2)^{-(a+1)} \, \mathrm{e}^{-b/\sigma^2} \mathrm{d}\sigma^2 \\
&= \frac{1}{(2\pi)^{\tilde{M}/2}} \frac{b^a}{\Gamma(a)} \int_{\mathbb{R}_+} (\sigma^2)^{-(a+\tilde{M}/2+1)} \, \exp\Bigg(-\frac{1}{\sigma^2} \Bigg(b + \frac{1}{2} \sum_{k \in \tilde{\mathcal{K}}} \sum_{n=1}^{\tilde{m}_k} y_{k,n}^2 \Bigg) \Bigg) \, \mathrm{d}\sigma^2,
\end{aligned}$$

with $\tilde{M} \triangleq \sum_{k \in \tilde{\mathcal{K}}} \tilde{m}_k$, or more compactly

$$\tilde{A} = \frac{1}{(2\pi)^{\tilde{M}/2}} \frac{b^a}{\Gamma(a)} \int_{\mathbb{R}_+} (\sigma^2)^{-(a+\tilde{M}/2+1)} \, \mathrm{e}^{-(b+\tilde{y}/2)/\sigma^2} \mathrm{d}\sigma^2, \tag{S.55}$$

where we defined

$$\tilde{y} \triangleq \sum_{k \in \tilde{\mathcal{K}}} \sum_{n=1}^{\tilde{m}_k} y_{k,n}^2 \,. \tag{S.56}$$

Multiplying and dividing the integrand in (S.55) by $\frac{\Gamma(a+\tilde{M}/2)}{(b+\tilde{y}/2)^{a+\tilde{M}/2}}$ gives

$$\begin{aligned}
(\sigma^2)^{-(a+\tilde{M}/2+1)} \, \mathrm{e}^{-(b+\tilde{y}/2)/\sigma^2} &= \frac{\Gamma(a+\tilde{M}/2)}{(b+\tilde{y}/2)^{a+\tilde{M}/2}} \frac{(b+\tilde{y}/2)^{a+\tilde{M}/2}}{\Gamma(a+\tilde{M}/2)} \, (\sigma^2)^{-(a+\tilde{M}/2+1)} \, \mathrm{e}^{-(b+\tilde{y}/2)/\sigma^2} \\
&= \frac{\Gamma(a+\tilde{M}/2)}{(b+\tilde{y}/2)^{a+\tilde{M}/2}} \, \Gamma^{-1}(\sigma^2; a+\tilde{M}/2, b+\tilde{y}/2). \qquad \text{(S.57)}
\end{aligned}$$

Inserting (S.57) into (S.55), we obtain

$$\begin{aligned}
\tilde{A} &= \frac{1}{(2\pi)^{\tilde{M}/2}} \frac{b^a}{\Gamma(a)} \frac{\Gamma(a+\tilde{M}/2)}{(b+\tilde{y}/2)^{a+\tilde{M}/2}} \int_{\mathbb{R}_+} \Gamma^{-1}(\sigma^2; a+\tilde{M}/2, b+\tilde{y}/2) \, \mathrm{d}\sigma^2 \\
&= \frac{1}{(2\pi)^{\tilde{M}/2}} \frac{b^a}{\Gamma(a)} \frac{\Gamma(a+\tilde{M}/2)}{(b+\tilde{y}/2)^{a+\tilde{M}/2}}, \qquad \text{(S.58)}
\end{aligned}$$

where we used the fact that integrating a pdf over its support yields 1. Using the identity $\frac{b^a}{(b+\tilde{y}/2)^{a+\tilde{M}/2}} = \frac{1}{b^{\tilde{M}/2}} \big(1 + \frac{1}{2b} \tilde{y}\big)^{-(2a+\tilde{M})/2}$ in (S.58) yields further

$$\tilde{A} = \frac{\Gamma(a+\tilde{M}/2)}{\Gamma(a) \, (2\pi b)^{\tilde{M}/2}} \bigg(1 + \frac{1}{2b} \tilde{y} \bigg)^{-(2a+\tilde{M})/2}. \tag{S.59}$$

Finally, noting that $\tilde{y}$ in (S.56) can be written as $\|\boldsymbol{y}\|_2^2$ where $\boldsymbol{y} \triangleq \big[\boldsymbol{y}_{\tilde{k}_{\mathrm{S}}}^{\mathrm{T}} \; \boldsymbol{y}_{\tilde{k}_{\mathrm{S}}+1}^{\mathrm{T}} \cdots \; \boldsymbol{y}_{\tilde{k}_{\mathrm{E}}}^{\mathrm{T}}\big]^{\mathrm{T}}$ with $\boldsymbol{y}_k \triangleq [y_{k,1} \; y_{k,2} \; \cdots \; y_{k,\tilde{m}_k}]^{\mathrm{T}}$, expression (S.59) becomes

$$\tilde{A} = \frac{\Gamma(a+\tilde{M}/2)}{\Gamma(a) \, (2\pi b)^{\tilde{M}/2}} \bigg(1 + \frac{1}{2b} \|\boldsymbol{y}\|_2^2 \bigg)^{-(2a+\tilde{M})/2} = \mathrm{t}\big(\boldsymbol{y}; 2a, \mathbf{0}_{\tilde{M}\times 1}, \tfrac{b}{a} \mathbf{I}_{\tilde{M}}\big), \tag{S.60}$$

where $\mathrm{t}\big(\boldsymbol{y}; \nu, \boldsymbol{\mu}, \boldsymbol{\Sigma}\big)$ denotes the pdf of a multivariate t-distribution [9] with ν degrees of freedom, location parameter vector $\boldsymbol{\mu} \in \mathbb{R}^{\tilde{M}}$, and $\tilde{M} \times \tilde{M}$ scale matrix $\boldsymbol{\Sigma}$.

We can now specialize the result (S.60) to the factors A_1, A_2, B_1, and B_2 in (S.49)–(S.53). Let $\boldsymbol{u}_{i,\kappa:\kappa+1} \triangleq \big[\boldsymbol{u}^{\mathrm{T}}_{k_{\mathrm{S},i}+1,i,\kappa:\kappa+1}\ \boldsymbol{u}^{\mathrm{T}}_{k_{\mathrm{S},i}+2,i,\kappa:\kappa+1} \cdots \boldsymbol{u}^{\mathrm{T}}_{k_{\mathrm{E},i},i,\kappa:\kappa+1}\big]^{\mathrm{T}}$ with $\boldsymbol{u}_{k,i,\kappa:\kappa+1} \triangleq [u_{k,i,\kappa}\ u_{k,i,\kappa+1}]^{\mathrm{T}}$ for $\kappa = 1,3$. Then, A_1 in (S.49) is obtained from (S.60) by setting $\boldsymbol{y} = \boldsymbol{u}_{i,1:2}$, $a = a_{\mathrm{q},1}$, $\tilde{M} = 2(k_{\mathrm{E},i} - k_{\mathrm{S},i})$, and $b = b_{\mathrm{q},1}$:

$$A_1 = \mathrm{t}\big(\boldsymbol{u}_{i,1:2}; 2a_{\mathrm{q},1}, \mathbf{0}_{2(k_{\mathrm{E},i}-k_{\mathrm{S},i})\times 1}, \tfrac{b_{\mathrm{q},1}}{a_{\mathrm{q},1}}\mathbf{I}_{2(k_{\mathrm{E},i}-k_{\mathrm{S},i})}\big). \tag{S.61}$$

Similarly, A_2 in (S.50) is obtained from (S.60) by setting $\boldsymbol{y} = \boldsymbol{u}_{i,3:4}$, $a = a_{\mathrm{q},2}$, $\tilde{M} = 2(k_{\mathrm{E},i} - k_{\mathrm{S},i})$, and $b = b_{\mathrm{q},2}$:

$$A_2 = \mathrm{t}\big(\boldsymbol{u}_{i,3:4}; 2a_{\mathrm{q},2}, \mathbf{0}_{2(k_{\mathrm{E},i}-k_{\mathrm{S},i})\times 1}, \tfrac{b_{\mathrm{q},2}}{a_{\mathrm{q},2}}\mathbf{I}_{2(k_{\mathrm{E},i}-k_{\mathrm{S},i})}\big). \tag{S.62}$$

Next, let $\boldsymbol{v}_{\mathcal{K}_i,i,\kappa} \triangleq \big[\boldsymbol{v}^{\mathrm{T}}_{k_{\mathrm{S},i},i,\kappa}\ \boldsymbol{v}^{\mathrm{T}}_{k_{\mathrm{S},i}+1,i,\kappa} \cdots \boldsymbol{v}^{\mathrm{T}}_{k_{\mathrm{E},i},i,\kappa}\big]^{\mathrm{T}}$ with $\boldsymbol{v}_{k,i,\kappa} \triangleq \big[v^{(1)}_{k,i,\kappa}\ v^{(2)}_{k,i,\kappa} \cdots v^{(m_{k,i})}_{k,i,\kappa}\big]^{\mathrm{T}}$ for $\kappa = 1,2$. Then, B_1 in (S.52) is obtained from (S.60) by setting $\boldsymbol{y} = \boldsymbol{v}_{\mathcal{K}_i,i,1}$, $a = a_{\mathrm{d},1}$, $\tilde{M} = M^{(i)} = \sum_{k\in\mathcal{K}_i} m_{k,i}$, and $b = b_{\mathrm{d},1}$:

$$B_1 = \mathrm{t}\big(\boldsymbol{v}_{\mathcal{K}_i,i,1}; 2a_{\mathrm{d},1}, \mathbf{0}_{M^{(i)}\times 1}, \tfrac{b_{\mathrm{d},1}}{a_{\mathrm{d},1}}\mathbf{I}_{M^{(i)}}\big), \tag{S.63}$$

and B_2 in (S.53) is obtained from (S.60) by setting $\boldsymbol{y} = \boldsymbol{v}_{\mathcal{K}_i,i,2}$, $a = a_{\mathrm{d},2}$, $\tilde{M} = M^{(i)}$, and $b = b_{\mathrm{d},2}$:

$$B_2 = \mathrm{t}\big(\boldsymbol{v}_{\mathcal{K}_i,i,2}; 2a_{\mathrm{d},2}, \mathbf{0}_{M^{(i)}\times 1}, \tfrac{b_{\mathrm{d},2}}{a_{\mathrm{d},2}}\mathbf{I}_{M^{(i)}}\big). \tag{S.64}$$

Factor C_1. It remains to calculate the factor C_1 in (S.54). We obtain

$$\begin{aligned} C_1 &\triangleq \int_{\mathbb{R}_+} \Bigg(\prod_{k\in\mathcal{K}_i} \mathrm{e}^{-\lambda_i}\frac{\lambda_i^{m_{k,i}}}{m_{k,i}!}\Bigg) \frac{b_\lambda^{a_\lambda}}{\Gamma(a_\lambda)}\,\lambda_i^{a_\lambda-1}\mathrm{e}^{-b_\lambda\lambda_i}\,\mathrm{d}\lambda_i \\ &= \frac{b_\lambda^{a_\lambda}}{\Gamma(a_\lambda)\prod_{k\in\mathcal{K}_i} m_{k,i}!}\int_{\mathbb{R}_+} \lambda_i^{a_\lambda+M^{(i)}-1}\mathrm{e}^{-(k_{\mathrm{E},i}-k_{\mathrm{S},i}+1+b_\lambda)\lambda_i}\,\mathrm{d}\lambda_i, \end{aligned} \tag{S.65}$$

where $|\mathcal{K}_i| = k_{\mathrm{E},i} - k_{\mathrm{S},i} + 1$ was used. Multiplying and dividing the integrand in (S.65) by $\frac{\Gamma(a_\lambda+M^{(i)})}{(k_{\mathrm{E},i}-k_{\mathrm{S},i}+1+b_\lambda)^{a_\lambda+M^{(i)}}}$ results in

$$\begin{aligned} &\lambda_i^{a_\lambda+M^{(i)}-1}\mathrm{e}^{-(k_{\mathrm{E},i}-k_{\mathrm{S},i}+1+b_\lambda)\lambda_i} \\ &\quad = \frac{\Gamma(a_\lambda+M^{(i)})}{(k_{\mathrm{E},i}-k_{\mathrm{S},i}+1+b_\lambda)^{a_\lambda+M^{(i)}}}\,\frac{(k_{\mathrm{E},i}-k_{\mathrm{S},i}+1+b_\lambda)^{a_\lambda+M^{(i)}}}{\Gamma(a_\lambda+M^{(i)})}\,\lambda_i^{a_\lambda+M^{(i)}-1}\mathrm{e}^{-(k_{\mathrm{E},i}-k_{\mathrm{S},i}+1+b_\lambda)\lambda_i} \\ &\quad = \frac{\Gamma(a_\lambda+M^{(i)})}{(k_{\mathrm{E},i}-k_{\mathrm{S},i}+1+b_\lambda)^{a_\lambda+M^{(i)}}}\,\Gamma(\lambda_i; a_\lambda+M^{(i)}, k_{\mathrm{E},i}-k_{\mathrm{S},i}+1+b_\lambda). \end{aligned} \tag{S.66}$$

Inserting (S.66) into (S.65) gives

$$\begin{aligned} C_1 &= \frac{b_\lambda^{a_\lambda}}{\Gamma(a_\lambda)\prod_{k\in\mathcal{K}_i} m_{k,i}!}\,\frac{\Gamma(a_\lambda+M^{(i)})}{(k_{\mathrm{E},i}-k_{\mathrm{S},i}+1+b_\lambda)^{a_\lambda+M^{(i)}}}\int_{\mathbb{R}_+}\Gamma(\lambda_i; a_\lambda+M^{(i)}, k_{\mathrm{E},i}-k_{\mathrm{S},i}+1+b_\lambda)\,\mathrm{d}\lambda_i \\ &= \frac{b_\lambda^{a_\lambda}}{\Gamma(a_\lambda)\prod_{k\in\mathcal{K}_i} m_{k,i}!}\,\frac{\Gamma(a_\lambda+M^{(i)})}{(k_{\mathrm{E},i}-k_{\mathrm{S},i}+1+b_\lambda)^{a_\lambda+M^{(i)}}}, \end{aligned}$$

where we again used the fact that integrating a pdf over its support yields 1. Rearranging terms gives

$$\begin{aligned} C_1 &= \frac{\Gamma(a_\lambda+M^{(i)})\big(\frac{b_\lambda}{k_{\mathrm{E},i}-k_{\mathrm{S},i}+1+b_\lambda}\big)^{a_\lambda}}{\Gamma(a_\lambda)}\prod_{k\in\mathcal{K}_i}\frac{\big(\frac{1}{k_{\mathrm{E},i}-k_{\mathrm{S},i}+1+b_\lambda}\big)^{m_{k,i}}}{m_{k,i}!} \\ &= \mathrm{NM}\big(\boldsymbol{m}_{\mathcal{K}_i,i}; a_\lambda, \tfrac{1}{k_{\mathrm{E},i}-k_{\mathrm{S},i}+1+b_\lambda}\mathbf{1}_{k_{\mathrm{E},i}-k_{\mathrm{S},i}+1\times 1}\big), \end{aligned} \tag{S.67}$$

where $\mathrm{NM}(\boldsymbol{m}_{\mathcal{K}_i,i}; a, \boldsymbol{b})$ denotes the pdf of the negative multinomial distribution [10] with parameters a and $\boldsymbol{b} \in \Delta_{k_{\mathrm{E},i}-k_{\mathrm{S},i}}$, where $\Delta_{k_{\mathrm{E},i}-k_{\mathrm{S},i}}$ is the $(k_{\mathrm{E},i}-k_{\mathrm{S},i})$-dimensional simplex.

Final expression of $p(c_i \,|\, \boldsymbol{c}_{\neg i}, \boldsymbol{P}^*_{\boldsymbol{c}_{\neg i}}, \boldsymbol{X}, \boldsymbol{Z})$ for $c_i \notin \mathcal{C}_{\neg i}$. Inserting (S.61)–(S.64), (S.67), and (S.51) into (S.48) and the result into the second case in (S.42) gives

$$\begin{aligned} p(c_i \,|\, \boldsymbol{c}_{\neg i}, \boldsymbol{P}^*_{\boldsymbol{c}_{\neg i}}, \boldsymbol{X}, \boldsymbol{Z}) = C \frac{\alpha}{\alpha + I - 1} \, & \mathrm{t}\big(\boldsymbol{u}_{i,1:2}; 2a_{\mathrm{q},1}, \mathbf{0}_{2(k_{\mathrm{E},i}-k_{\mathrm{S},i})\times 1}, \tfrac{b_{\mathrm{q},1}}{a_{\mathrm{q},1}} \mathbf{I}_{2(k_{\mathrm{E},i}-k_{\mathrm{S},i})}\big) \\ & \times \mathrm{t}\big(\boldsymbol{u}_{i,3:4}; 2a_{\mathrm{q},2}, \mathbf{0}_{2(k_{\mathrm{E},i}-k_{\mathrm{S},i})\times 1}, \tfrac{b_{\mathrm{q},2}}{a_{\mathrm{q},2}} \mathbf{I}_{2(k_{\mathrm{E},i}-k_{\mathrm{S},i})}\big) \\ & \times \mathrm{t}\big(\boldsymbol{v}_{\mathcal{K}_i,i,1}; 2a_{\mathrm{d},1}, \mathbf{0}_{M^{(i)}\times 1}, \tfrac{b_{\mathrm{d},1}}{a_{\mathrm{d},1}} \mathbf{I}_{M^{(i)}}\big) \, \mathrm{t}\big(\boldsymbol{v}_{\mathcal{K}_i,i,2}; 2a_{\mathrm{d},2}, \mathbf{0}_{M^{(i)}\times 1}, \tfrac{b_{\mathrm{d},2}}{a_{\mathrm{d},2}} \mathbf{I}_{M^{(i)}}\big) \\ & \times \mathrm{NM}\big(\boldsymbol{m}_{\mathcal{K}_i,i}; a_\lambda, \tfrac{1}{k_{\mathrm{E},i}-k_{\mathrm{S},i}+1+b_\lambda} \mathbf{1}_{k_{\mathrm{E},i}-k_{\mathrm{S},i}+1\times 1}\big) \, \mathcal{N}(\boldsymbol{x}_{k_{\mathrm{S},i},i}; \mathbf{0}_{4\times 1}, \boldsymbol{\Sigma}_{\mathrm{S}}), \end{aligned} \tag{S.68}$$

for $c_i \notin \mathcal{C}_{\neg i}$. This expression is consistent with (37) (via (27)).

S5 Proposal pdfs $g_{\text{init}}(\boldsymbol{x}_{k_{\mathrm{S},i},i}; \boldsymbol{p}^*_{c_i}, \boldsymbol{Z}_{k_{\mathrm{S},i},i})$ and $g(\boldsymbol{x}_{k,i}; \boldsymbol{x}_{k-1,i}, \boldsymbol{p}^*_{c_i}, \boldsymbol{Z}_{k,i})$

The object state proposal pdf $g_{\text{init}}(\boldsymbol{x}_{k_{\mathrm{S},i},i}; \boldsymbol{p}^*_{c_i}, \boldsymbol{Z}_{k_{\mathrm{S},i},i})$ used in the initialization of the PGAS particle filter (see Section IV-B2) is given by (30). We restate it here for convenience:

$$g_{\text{init}}(\boldsymbol{x}_{k_{\mathrm{S},i},i}; \boldsymbol{p}^*_{c_i}, \boldsymbol{Z}_{k_{\mathrm{S},i},i}) = \mathcal{N}\big(\boldsymbol{x}_{k_{\mathrm{S},i},i}; \boldsymbol{\mu}^{(g_{\text{init}})}_{k_{\mathrm{S},i},i}, \boldsymbol{\Sigma}^{(g_{\text{init}})}_{k_{\mathrm{S},i},i}\big),$$

where

$$\boldsymbol{\mu}^{(g_{\text{init}})}_{k_{\mathrm{S},i},i} = [\bar{\boldsymbol{z}}^{\mathrm{T}}_{k_{\mathrm{S},i},i} \ 0 \ 0]^{\mathrm{T}},$$

with $\bar{\boldsymbol{z}}_{k_{\mathrm{S},i},i}$ as defined in (22), and

$$\boldsymbol{\Sigma}^{(g_{\text{init}})}_{k_{\mathrm{S},i},i} = \boldsymbol{\Sigma}_{\mathrm{S}} \circ \mathrm{diag}(\mathbf{0}_{2\times 2}, \mathbf{I}_2) + \frac{\overline{d^{*2}_{c_i}}}{m_{k_{\mathrm{S},i},i}} \, \mathrm{diag}(\mathbf{I}_2, \mathbf{0}_{2\times 2}),$$

with $\circ$ denoting the Hadamard/elementwise product and $\overline{d^{*2}_{c}} \triangleq \frac{d^{*2}_{c,1} + d^{*2}_{c,2}}{2}$. These expressions of $\boldsymbol{\mu}^{(g_{\text{init}})}_{k_{\mathrm{S},i},i}$ and $\boldsymbol{\Sigma}^{(g_{\text{init}})}_{k_{\mathrm{S},i},i}$ combine prior and empirical information. More specifically, $\boldsymbol{\mu}^{(g_{\text{init}})}_{k_{\mathrm{S},i},i}$ is the sum of the prior mean of $f(\boldsymbol{x}_{k_{\mathrm{S},i},i})$ (which is $\mathbf{0}_{4\times 1}$ by (1)) and the average of the measurements associated with object i, $\bar{\boldsymbol{z}}_{k_{\mathrm{S},i},i}$. This is inspired by the MMSE estimator for the sequential linear Gaussian case [11, Sec. 12.6], which can be decomposed into analogous prior and empirical terms. Furthermore, the two position-related diagonal elements of $\boldsymbol{\Sigma}^{(g_{\text{init}})}_{k_{\mathrm{S},i},i}$ are given by $\overline{d^{*2}_{c_i}}/m_{k_{\mathrm{S},i},i}$, i.e., the average of the two object extent parameters of object i, $\overline{d^{*2}_{c_i}}$, divided by the number of associated measurements, $m_{k_{\mathrm{S},i},i}$. Thus, the position elements of the state $\boldsymbol{x}_{k_{\mathrm{S},i},i}$ are more tightly distributed around the position elements of the mean $\boldsymbol{\mu}^{(g_{\text{init}})}_{k_{\mathrm{S},i},i}$ when $\overline{d^{*2}_{c_i}}$ is smaller and/or $m_{k_{\mathrm{S},i},i}$ is larger. Finally, the velocity-related diagonal elements of $\boldsymbol{\Sigma}^{(g_{\text{init}})}_{k_{\mathrm{S},i},i}$ equal the corresponding elements of the prior covariance matrix $\boldsymbol{\Sigma}_{\mathrm{S}}$ in (1), since no velocity measurements are available.

The object state proposal pdf $g(\boldsymbol{x}_{k,i}; \boldsymbol{x}_{k-1,i}, \boldsymbol{p}^*_{c_i}, \boldsymbol{Z}_{k,i})$ used by the particle filter for $k = k_{\mathrm{S},i}+1, \dots, k_{\mathrm{E},i}$ (see Section IV-B3) is given by (33), and again restated here for convenience as

$$g(\boldsymbol{x}_{k,i}; \boldsymbol{x}_{k-1,i}, \boldsymbol{p}^*_{c_i}, \boldsymbol{Z}_{k,i}) = \mathcal{N}\big(\boldsymbol{x}_{k,i}; \boldsymbol{\mu}^{(g)}_{k,i}, \boldsymbol{\Sigma}^{(g)}_{k,i}\big),$$

where

$$\boldsymbol{\mu}^{(g)}_{k,i} = \boldsymbol{F}\boldsymbol{x}_{k-1,i} + \rho_{k,i} \boldsymbol{H}^{\mathrm{T}}(\bar{\boldsymbol{z}}_{k,i} - \boldsymbol{F}\boldsymbol{x}_{k-1,i}) \tag{S.69}$$

Algorithm 1 Proposed MCMC algorithm (iteration ℓ)

Input: $\boldsymbol{X}^{(\ell-1)}, \boldsymbol{P}^{(\ell-1)} = \big[\boldsymbol{p}_i^{(\ell-1)}\big]_{i=1}^{I}, \boldsymbol{c}^{(\ell-1)}, \big[\boldsymbol{Z}_{k,0:I}\big]_{k=1}^{K}, \big[\mathcal{K}_i\big]_{i=1}^{I}, R, g_{\text{init}}(\cdot), g(\cdot)$

1: **for all** $i = 1, \ldots, I$ **do**

2: sample $\boldsymbol{X}_{\mathcal{K}_i,i}^{(\ell)}$ using Algorithm 2 with input $\boldsymbol{X}_{\mathcal{K}_i,i}^{(\ell-1)}$, $\boldsymbol{p}_i^{(\ell-1)}$, $\boldsymbol{Z}_{\mathcal{K}_i,i} = [\boldsymbol{Z}_{k,i}]_{k\in\mathcal{K}_i}$, R, $g_{\text{init}}(\cdot)$, $g(\cdot)$

3: **end for**

4: sample $\boldsymbol{P}^{(\ell)}$ and $\boldsymbol{c}^{(\ell)}$ using Algorithm 3 with input $\boldsymbol{X}^{(\ell)}, \boldsymbol{P}^{(\ell-1)}, \boldsymbol{c}^{(\ell-1)}, \boldsymbol{Z} = \big[\boldsymbol{Z}_{k,0:I}\big]_{k=1}^{K}$

Output: $\boldsymbol{X}^{(\ell)}, \boldsymbol{P}^{(\ell)}, \boldsymbol{c}^{(\ell)}$

and

$$\boldsymbol{\Sigma}_{k,i}^{(g)} = \boldsymbol{Q}_{c_i}^{*} - q_{c_i,1}^{*2}\,\rho_{k,i}\,\mathrm{diag}(\mathbf{I}_2\,, \mathbf{0}_{2\times 2})\,, \tag{S.70}$$

with $\rho_{k,i} \triangleq q_{c_i,1}^{*2}\big/\big(q_{c_i,1}^{*2} + \overline{d_{c_i}^{*2}}/m_{k,i}\big)$. These expressions of $\boldsymbol{\mu}_{k,i}^{(g)}$ and $\boldsymbol{\Sigma}_{k,i}^{(g)}$ again combine prior and empirical information and are inspired by the MMSE estimator for the sequential linear Gaussian case [11, Sec. 12.6]. More specifically, the first term on the right-hand side of (S.69) equals the mean of $f(\boldsymbol{x}_{k,i}\,|\,\boldsymbol{x}_{k-1,i}, \boldsymbol{Q}_i)$ in (4) while the second term is a correction term involving the average of the measurements associated with object i, $\bar{\boldsymbol{z}}_{k,i}$ defined in (22), and a scaling factor $\rho_{k,i}$ that depends on $q_{c_i,1}^{*2}$, $\overline{d_{c_i}^{*2}}$, and $m_{k,i}$. The correction term—and, thus, the influence of the empirical information provided by $\bar{\boldsymbol{z}}_{k,i}$ on $\boldsymbol{\mu}_{k,i}^{(g)}$—tends to be larger when $m_{k,i}$ is larger and/or $q_{c_i,1}^{*2}$ is larger and/or $\overline{d_{c_i}^{*2}}$ is smaller. Furthermore, $\boldsymbol{\Sigma}_{k,i}^{(g)}$ in (S.70) equals the covariance matrix in (4) (via the identity $\boldsymbol{Q}_{c_i}^{*} = \boldsymbol{Q}_i$) except that $q_{c_i,1}^{*2}\,\rho_{k,i}$ is subtracted from the position-related diagonal elements. Thus, the position elements of $\boldsymbol{x}_{k,i}$ are more tightly distributed around the position elements of the mean $\boldsymbol{\mu}_{k,i}^{(g)}$ when $\rho_{k,i}$ is larger. Note that $\boldsymbol{\Sigma}_{k,i}^{(g)}$ is guaranteed to be positive definite because $q_{c_i,1}^{*2}\rho_{k,i} = q_{c_i,1}^{*4}\big/\big(q_{c_i,1}^{*2} + \overline{d_{c_i}^{*2}}/m_{k,i}\big)$.

S6 Pseudo-code of the MCMC Algorithm

Pseudo-code for one iteration of the proposed MCMC algorithm described in Section IV is presented in Algorithms 1–3. According to Algorithm 1, the MCMC algorithm consists of sampling the object states (using the PGAS algorithm stated in Algorithm 2) and then sampling the class indices and the class parameters/object parameters (using the Gibbs sampler stated in Algorithm 3). We note that the DA-aware measurement matrices $\boldsymbol{Z}_{k,0:I}$ for each $k = 1, \ldots, K$ used as inputs to Algorithm 1 are obtained according to $\boldsymbol{Z}_{k,0:I} = \pi_k^{-1}(\boldsymbol{Z}_k)$. Furthermore, the new class index value $c_{i,\text{new}}^{(\ell)}$ involved in Algorithm 3 is chosen as the smallest integer that is larger than the maximum class index value that arose in the previous sampling steps $\ell' = 1, \ldots, \ell-1$ or in the current sampling step ℓ so far (for $i' = 1, \ldots, i-1$), i.e.,

$$c_{i,\text{new}}^{(\ell)} \triangleq \max\big\{c_{\text{max}}^{(\ell-1)}, \max_{i'\in\{1,\ldots,i-1\}} c_{i'}^{(\ell)}\big\} + 1\,, \tag{S.71}$$

with $c_{\text{max}}^{(\ell-1)} \triangleq \max_{\ell'\in\{1,\ldots,\ell-1\}, i\in\{1,\ldots,I\}} c_i^{(\ell')}$.

MATLAB code for our algorithm is provided at `https://github.com/tjbucco/EMT-BNP-Learning/`.

S7 Blackwell-MacQueen Sampling Procedure

According to Section VI-A, in the case where $\hat{I}^{(j-1)} < I_{\text{max}}$, the "DAOD parameters" $\boldsymbol{p}_{i,\text{DAOD}}^{(j)}$ for $i = \hat{I}^{(j-1)} + 1, \ldots, I_{\text{max}}$ required by the DAOD algorithm are sampled using the Blackwell-MacQueen construction of the DP.

Algorithm 2 PGAS algorithm: sampling the object states (iteration ℓ)

Input: $\boldsymbol{X}^{(\ell-1)}_{\mathcal{K}_i,i}$, $\boldsymbol{p}^{(\ell-1)}_i$, $\boldsymbol{Z}_{\mathcal{K}_i,i} = [\boldsymbol{Z}_{k,i}]_{k\in\mathcal{K}_i}$, R, $g_{\text{init}}(\cdot)$, $g(\cdot)$

for all $r = 1,\ldots,R-1$ **do**
- sample $\boldsymbol{\xi}^{(\ell,k_{\text{S},i},r)}_{k_{\text{S},i},i}$ from $g_{\text{init}}\big(\boldsymbol{x}_{k_{\text{S},i},i}; \boldsymbol{p}^{(\ell-1)}_i, \boldsymbol{Z}_{k_{\text{S},i},i}\big)$ (see (30))
- set $\boldsymbol{\Xi}^{(\ell,k_{\text{S},i},r)}_{k_{\text{S},i},i} = \boldsymbol{\xi}^{(\ell,k_{\text{S},i},r)}_{k_{\text{S},i},i}$

end for

set $\boldsymbol{\xi}^{(\ell,k_{\text{S},i},R)}_{k_{\text{S},i},i} = \boldsymbol{x}^{(\ell-1)}_{k_{\text{S},i},i}$

set $\boldsymbol{\Xi}^{(\ell,k_{\text{S},i},R)}_{k_{\text{S},i},i} = \boldsymbol{\xi}^{(\ell,k_{\text{S},i},R)}_{k_{\text{S},i},i}$

for all $r = 1,\ldots,R$ **do**
- calculate $\nu^{(\ell,k_{\text{S},i},r)}_i$ according to (25)

end for

for all $k = k_{\text{S},i}+1,\ldots,k_{\text{E},i}$ **do**
- **for all** $r = 1,\ldots,R-1$ **do**
  - sample $\boldsymbol{\Xi}^{(\ell,k,r)}_{k_{\text{S},i}:k-1,i} = \boldsymbol{\Xi}^{(\ell,k-1,r')}_{k_{\text{S},i}:k-1,i}$ with probability $\nu^{(\ell,k-1,r')}_i$, $r' \in \{1,\ldots,R\}$
  - sample $\boldsymbol{\xi}^{(\ell,k,r)}_{k,i}$ from $g\big(\boldsymbol{x}_{k,i}; \boldsymbol{\xi}^{(\ell,k-1,r)}_{k-1,i}, \boldsymbol{p}^{(\ell-1)}_i, \boldsymbol{Z}_{k,i}\big)$ (see (33))
  - set $\boldsymbol{\Xi}^{(\ell,k,r)}_{k_{\text{S},i}:k,i} = \big[\boldsymbol{\Xi}^{(\ell,k,r)}_{k_{\text{S},i}:k-1,i} \;\; \boldsymbol{\xi}^{(\ell,k,r)}_{k,i}\big]$
- **end for**
- **for all** $r = 1,\ldots,R$ **do**
  - set $\tilde{P}^{(\ell,k-1,r,R)}_i \propto \nu^{(\ell,k-1,r)}_i \, f_{\mathbf{x}_{k,i}|\mathbf{x}_{k-1,i},\mathbf{p}_i}\big(\boldsymbol{x}^{(\ell-1)}_{k,i} \big| \boldsymbol{\xi}^{(\ell,k-1,r)}_{k-1,i}, \boldsymbol{p}^{(\ell-1)}_i\big)$
- **end for**
- **for all** $r = 1,\ldots,R$ **do**
  - set $P^{(\ell,k-1,r,R)}_i = \frac{\tilde{P}^{(\ell,k-1,r,R)}_i}{\sum_{r'=1}^{R} \tilde{P}^{(\ell,k-1,r',R)}_i}$
- **end for**
- sample $\boldsymbol{\Xi}^{(\ell,k,R)}_{k_{\text{S},i}:k-1,i} = \boldsymbol{\Xi}^{(\ell,k-1,r')}_{k_{\text{S},i}:k-1,i}$ with probability $P^{(\ell,k-1,r',R)}_i$, $r' \in \{1,\ldots,R\}$
- set $\boldsymbol{\Xi}^{(\ell,k,R)}_{k_{\text{S},i}:k,i} = \big[\boldsymbol{\Xi}^{(\ell,k,R)}_{k_{\text{S},i}:k-1,i} \;\; \boldsymbol{x}^{(\ell-1)}_{k,i}\big]$
- **for all** $r = 1,\ldots,R$ **do**
  - calculate $\nu^{(\ell,k,r)}_i$ according to (26)
- **end for**

end for

sample $\boldsymbol{X}^{(\ell)}_{\mathcal{K}_i,i} = \boldsymbol{\Xi}^{(\ell,k_{\text{E},i},r)}_{k_{\text{S},i}:k_{\text{E},i},i}$ with probability $\nu^{(\ell,k_{\text{E},i},r)}_i$, $r \in \{1,\ldots,R\}$

Output: $\boldsymbol{X}^{(\ell)}_{\mathcal{K}_i,i}$

This sampling procedure is described in the following.

First, class index values $c^{(j)}_{i,\text{DAOD}} \in \mathbb{N}$ are generated for $i = \hat{I}^{(j-1)}+1,\ldots,I_{\max}$. Mores specifically, $c^{(j)}_{i,\text{DAOD}}$ is recursively sampled from the conditional pmf [12, Sec. 2.6.6]

$$p\big(c^{(j)}_{i,\text{DAOD}} \big| c^{(j)}_{1,\text{DAOD}},\ldots,c^{(j)}_{i-1,\text{DAOD}}\big) = \frac{1}{\alpha+i-1}\Bigg(\alpha\,\mathbb{1}\big(c^{(j)}_{i,\text{DAOD}} = c^{(j)}_{i-1,\text{DAOD},\max}+1\big) \\ + \sum_{c=1}^{c^{(j)}_{i-1,\text{DAOD},\max}} N\big(c; c^{(j)}_{1,\text{DAOD}},\ldots,c^{(j)}_{i-1,\text{DAOD}}\big)\,\mathbb{1}\big(c = c^{(j)}_{i,\text{DAOD}}\big)\Bigg).$$

Algorithm 3 Gibbs sampler: sampling the class indices and class/object parameters (iteration ℓ)

Input: $\boldsymbol{X}^{(\ell)}, \boldsymbol{P}^{(\ell-1)} = \big[\boldsymbol{p}_i^{(\ell-1)}\big]_{i=1}^{I}, \boldsymbol{c}^{(\ell-1)}, \boldsymbol{Z} = \big[\boldsymbol{Z}_{k,0:I}\big]_{k=1}^{K}$

for all $i = 1, \ldots, I$ **do**
- set $\boldsymbol{p}^{*(\ell-1)}_{c_i^{(\ell-1)}} = \boldsymbol{p}_i^{(\ell-1)}$ (see (19))

end for

for all $i = 1, \ldots, I$ **do**
- set $c_{i,\mathrm{new}}^{(\ell)}$ according to (S.71)
- sample $c_i^{(\ell)} \in \mathcal{C}_{\neg i}^{(\ell,\ell-1)} \cup \{c_{i,\mathrm{new}}^{(\ell)}\}$ from $p_i(c;\ell)$ (see (36) for $c \in \mathcal{C}_{\neg i}^{(\ell,\ell-1)}$ and (37) for $c = c_{i,\mathrm{new}}^{(\ell)}$)
- if $c_i^{(\ell)} = c_{i,\mathrm{new}}^{(\ell)}$
  - sample $\boldsymbol{p}^{*(\ell-1)}_{c_{i,\mathrm{new}}^{(\ell)}}$ from $f_{\mathbf{p}_c^*|\mathbf{X},\mathbf{c},\mathbf{P}_{\neg c}^*,\mathbf{Z}}\big(\boldsymbol{p}_c^* \,\big|\, \boldsymbol{X}^{(\ell)}, \boldsymbol{c}^{(\ell)}, \boldsymbol{P}^{*(\ell-1)}_{\boldsymbol{c}_{\neg i}^{(\ell,\ell-1)}}, \boldsymbol{Z}\big)$ (see (38))

end for

for all $c \in \big\{c_1^{(\ell)}, \ldots, c_I^{(\ell)}\big\}$ **do**
- sample $\boldsymbol{p}_c^{*(\ell)}$ from $f_{\mathbf{p}_c^*|\mathbf{X},\mathbf{c},\mathbf{P}_{\neg c}^*,\mathbf{Z}}\big(\boldsymbol{p}_c^* \,\big|\, \boldsymbol{X}^{(\ell)}, \boldsymbol{c}^{(\ell)}, \boldsymbol{P}_{\neg c}^{*(\ell,\ell-1)}, \boldsymbol{Z}\big)$ (see (38))

end for

for all $i = 1, \ldots, I$ **do**
- set $\boldsymbol{p}_i^{(\ell)} = \boldsymbol{p}^{*(\ell)}_{c_i^{(\ell)}}$ (see (19))

end for

Output: $\boldsymbol{P}^{(\ell)}, \boldsymbol{c}^{(\ell)}$

Here, the $c_{i,\mathrm{DAOD}}^{(j)}$ for $i = 1, \ldots, \hat{I}^{(j-1)}$ equal the respective estimates $\hat{c}_i^{(j-1)}$ that were previously calculated by the MCMC estimation algorithm; $\mathbb{1}(\cdot)$ is 1 if the function argument is true and 0 otherwise; $c_{i-1,\mathrm{DAOD,max}}^{(j)} \triangleq \max_{i' \in \{1,\ldots,i-1\}} c_{i',\mathrm{DAOD}}^{(j)}$; and $N\big(c; c_{1,\mathrm{DAOD}}^{(j)}, \ldots, c_{i-1,\mathrm{DAOD}}^{(j)}\big)$ is the number of $c_{i',\mathrm{DAOD}}^{(j)}$ with $i' \in \{1, \ldots, i-1\}$ that are equal to c. We note that the above pmf is related to the Chinese restaurant process, which is a Bayesian nonparametric model closely related to the DP [12, Sec. 3.1].

The DAOD parameters $\boldsymbol{p}_{i,\mathrm{DAOD}}^{(j)}$ for $i = \hat{I}^{(j-1)} + 1, \ldots, I_{\max}$ are now obtained as follows. If $c_{i,\mathrm{DAOD}}^{(j)} \in \big\{c_{1,\mathrm{DAOD}}^{(j)}, \ldots, c_{\hat{I}^{(j-1)},\mathrm{DAOD}}^{(j)}\big\}$, i.e., $c_{i,\mathrm{DAOD}}^{(j)}$ is the index of a class for which the MCMC estimation algorithm already calculated a parameter estimate $\hat{\boldsymbol{p}}^{*(j-1)}_{c_{i,\mathrm{DAOD}}^{(j)}}$, then $\boldsymbol{p}_{i,\mathrm{DAOD}}^{(j)} = \hat{\boldsymbol{p}}^{*(j-1)}_{c_{i,\mathrm{DAOD}}^{(j)}}$. In the opposite case, i.e., $c_{i,\mathrm{DAOD}}^{(j)} \notin \big\{c_{1,\mathrm{DAOD}}^{(j)}, \ldots, c_{\hat{I}^{(j-1)},\mathrm{DAOD}}^{(j)}\big\}$, there are two subcases: (i) If $c_{i,\mathrm{DAOD}}^{(j)} = c_{i-1,\mathrm{DAOD,max}}^{(j)} + 1$, then $\boldsymbol{p}_{i,\mathrm{DAOD}}^{(j)}$ is sampled from the base pdf f_{B} in (21); (ii) if $c_{i,\mathrm{DAOD}}^{(j)} \in \big\{\hat{c}_{\max}^{(j-1)}+1, \ldots, c_{i-1,\mathrm{DAOD,max}}^{(j)}\big\}$ with $\hat{c}_{\max}^{(j-1)} \triangleq \max_{i \in \{1,\ldots,\hat{I}^{(j-1)}\}} \hat{c}_i^{(j-1)}$, i.e., $c_{i,\mathrm{DAOD}}^{(j)}$ is the index of a class for which the DAOD parameter was already sampled in the former subcase, then that sample is used.

S8 Further Details of the Simulation Setup

In this section, we provide parameters of our simulation and of the considered methods that are not specified in [1] because of space limitations.

The initial positions of the 15 objects are regularly spaced on a circle of radius $100\,\mathrm{m}$, and the objects move with an initial velocity of $10\,\mathrm{m/s}$ toward the center of the circle. The mean driving noise variances are $\mathbb{E}[\mathrm{q}_{c,1}^{*2}] =$

$100\,\mathrm{m}^2$ and $\mathbb{E}[\mathsf{q}_{c,2}^{*2}] = 50\,\mathrm{m}^2/\mathrm{s}^2$, the mean extent parameters (measurement noise variances) are $\mathbb{E}[\mathsf{d}_{c,1}^{*2}] = 675\,\mathrm{m}^2$ and $\mathbb{E}[\mathsf{d}_{c,2}^{*2}] = 400\,\mathrm{m}^2$, and the mean measurement rate is $\mathbb{E}[\lambda_c] = 6$.

SEP uses values of some of the hyperparameters that differ from our system model because this resulted in better performance. Concretely, SEP uses $a_{\mathrm{d},1} = a_{\mathrm{d},2} = 100$, $b_{\mathrm{d},1} = 66825\,\mathrm{m}^2$, and $b_{\mathrm{d},2} = 39600\,\mathrm{m}^2$, which causes the variances of the extent parameters $\mathsf{d}_{i,1}^2$ and $\mathsf{d}_{i,2}^2$ to be significantly smaller than those of the system model (note, however, that the mean extent parameters $\mathbb{E}[\mathsf{d}_{i,1}^2]$ and $\mathbb{E}[\mathsf{d}_{i,2}^2]$ are still equal to those of the system model).

PMBM uses the fixed driving noise covariance matrices $\boldsymbol{Q}_i = \mathrm{diag}\big(\mathbb{E}[\mathsf{q}_{c,1}^{*2}]\,\mathbf{I}_2, \mathbb{E}[\mathsf{q}_{c,2}^{*2}]\,\mathbf{I}_2\big) = \mathrm{diag}(50\mathrm{m}^2\,\mathbf{I}_2,$ $50\mathrm{m}^2/\mathrm{s}^2\,\mathbf{I}_2)$ and the fixed measurement rates $\lambda_i = \mathbb{E}[\lambda_c] = 6$, for all $i = 1, \ldots, I$. Furthermore, PMBM assumes that the object extent matrix $\mathbf{D}_i$ is time-varying and, at each time step, distributed according to an inverse Wishart distribution [6]. We choose the scale matrix and the number of degrees of freedom of the inverse Wishart distribution as $\mathrm{diag}(50\,b_{\mathrm{d},1}, 50\,b_{\mathrm{d},2}) = \mathrm{diag}(371250\mathrm{m}^2, 220000\mathrm{m}^2)$ and $50\,a_{\mathrm{d},1} + 3 = 603$, respectively. This means that the mean of the object extent matrix is the same as the mean object extent matrix of our model. The remaining parameters for PMBM are set as in [5].

The parameters of the OSPA-GW error [13] are as follows: cutoff parameter 1000, order 1, and parameter α (not to be confused with the DP's concentration hyperparameter α) set to 2. For the OSPA-E error [14], we use cutoff parameter 120, order 1, and $\alpha = 2$.

References

[1] T. J. Bucco, B. Kreidl, G. Koliander, and F. Hlawatsch, "Extended multiobject tracking with nonparametric Bayesian learning of object classes and model parameters," Mar. 2025, submitted.

[2] J. Pearl, "Bayesian networks: A model of self-activated memory for evidential reasoning," in *Proc. COGSCI-1985*, Irvine, CA, USA, Aug. 1985, pp. 329–334.

[3] S. Ghosal and A. Van der Vaart, *Fundamentals of Nonparametric Bayesian Inference*. Cambridge, UK: Cambridge University Press, 2017.

[4] T. S. Ferguson, "A Bayesian analysis of some nonparametric problems," *Ann. Statist.*, vol. 1, no. 2, pp. 209–230, Mar. 1973.

[5] F. Meyer and J. L. Williams, "Scalable detection and tracking of geometric extended objects," *IEEE Trans. Signal Process.*, vol. 69, pp. 6283–6298, Oct. 2021.

[6] K. Granström, M. Fatemi, and L. Svensson, "Poisson multi-Bernoulli mixture conjugate prior for multiple extended target filtering," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 1, pp. 208–225, Feb. 2020.

[7] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis*. New York, NY, USA: Chapman and Hall, 1995.

[8] R. M. Neal, "Markov chain sampling methods for Dirichlet process mixture models," *J. Comput. Graph. Stat.*, vol. 9, no. 2, pp. 249–265, 2000.

[9] M. Roth, "On the multivariate t distribution," Linköping University, The Institute of Technology, Linköping, Sweden, Tech. Rep. 3059, April 2013.

[10] M. Sibuya, I. Yoshimura, and R. Shimizu, "Negative multinomial distribution," *Ann. Inst. Statist. Math.*, vol. 16, no. 1, pp. 409–426, 1964.

[11] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[12] J. Pitman, *Combinatorial Stochastic Processes*. Berlin, Germany: Springer, 2006.

[13] S. Yang, M. Baum, and K. Granström, "Metrics for performance evaluation of elliptic extended object tracking methods," in *Proc. MFI-16*, Baden-Baden, Germany, 2016, pp. 523–528.

[14] A. S. Rahmathullah, Á. F. García-Fernández, and L. Svensson, "Generalized optimal sub-pattern assignment metric," in *Proc. FUSION-17*, Xi'an, China, July 2017, pp. 1–8.