$\bar{f} = 1 - f$ of its current value f:

$$
\begin{aligned}
& f_{ik} \sim \rho(\bar{f}|f)\delta(f_{ik}, \bar{f}) + (1 - \rho(\bar{f}|f))\delta(f_{ik}, f), \\
& \rho(\bar{f}|f) = \min\left\{ \frac{p(f_{ik} = \bar{f}|\mathbf{F}^{-ik}, \mathbf{y}^{(i)}_{1:T_i}, \boldsymbol{\eta}^{(i)}, \theta_{1:K_+^{-i}}, c)}{p(f_{ik} = f|\mathbf{F}^{-ik}, \mathbf{y}^{(i)}_{1:T_i}, \boldsymbol{\eta}^{(i)}, \theta_{1:K_+^{-i}}, c)}, 1 \right\}.
\end{aligned} \tag{15}
$$

To compute likelihoods $p(\mathbf{y}^{(i)}_{1:T_i}|\mathbf{f}_i, \boldsymbol{\eta}^{(i)}, \boldsymbol{\theta})$, we combine $\mathbf{f}_i$ and $\boldsymbol{\eta}^{(i)}$ to construct the transition distributions $\pi^{(i)}_j$ as in equation (10), and marginalize over the possible latent state sequences by applying a forward–backward message passing algorithm for AR-HMMs [see Supplement C.2 of Fox et al. (2014)]. In each sampler iteration, we apply these proposals sequentially to each entry of the feature matrix $\mathbf{F}$, visiting each entry one at a time and retaining any accepted proposals to be used as the fixed $\mathbf{F}^{-ik}$ for subsequent proposals.

6.3. *Sampling state sequences* $\mathbf{z}$. For each sequence i contained in $\mathbf{z}$, we block sample $\mathbf{z}^{(i)}_{1:T_i}$ in one coherent move. This is possible because $\mathbf{f}_i$ defines a finite AR-HMM for each sequence, enabling dynamic programming with auxiliary variables $\boldsymbol{\pi}^{(i)}, \boldsymbol{\theta}$. We compute backward messages $m_{t+1,t}(z^{(i)}_t) \propto p(\mathbf{y}^{(i)}_{t+1:T_i}|z^{(i)}_t, \tilde{\mathbf{y}}^{(i)}_t, \boldsymbol{\pi}^{(i)}, \boldsymbol{\theta})$, and recursively sample each $z^{(i)}_t$:

$$
z^{(i)}_t|z^{(i)}_{t-1}, \mathbf{y}^{(i)}_{1:T_i}, \boldsymbol{\pi}^{(i)}, \boldsymbol{\theta} \sim \pi^{(i)}_{z^{(i)}_{t-1}}(z^{(i)}_t)\mathcal{N}(\mathbf{y}^{(i)}_t; \mathbf{A}_{z^{(i)}_t}\tilde{\mathbf{y}}^{(i)}_t, \Sigma_{z^{(i)}_t})m_{t+1,t}(z^{(i)}_t). \tag{16}
$$

Supplement Algorithm D.3 of Fox et al. (2014) explains backward-filtering, forward-sampling in detail.

6.4. *Sampling auxiliary parameters*: $\boldsymbol{\theta}$ *and* $\boldsymbol{\eta}$. Given fixed features $\mathbf{F}$ and state sequences $\mathbf{z}$, the posterior over auxiliary parameters factorizes neatly:

$$
p(\boldsymbol{\theta}, \boldsymbol{\eta}|\mathbf{F}, \mathbf{z}, \mathbf{y}) = \prod_{k=1}^{K_+} p(\theta_k|\{\mathbf{y}^{(i)}_t : z^{(i)}_t = k\}) \prod_{i=1}^{N} p(\boldsymbol{\eta}^{(i)}|\mathbf{z}^{(i)}, \mathbf{f}_i). \tag{17}
$$

We can thus sample each θ_k and $\boldsymbol{\eta}^{(i)}$ independently, as outlined below.

Transition weights $\boldsymbol{\eta}^{(i)}$. Given state sequence $\mathbf{z}^{(i)}$ and features $\mathbf{f}_i$, sequence i's Markov transition weights $\boldsymbol{\eta}^{(i)}$ have posterior distribution

$$
p(\eta^{(i)}_{jk}|\mathbf{z}^{(i)}, f_{ij} = 1, f_{ik} = 1) \propto \frac{(\eta^{(i)}_{jk})^{n^{(i)}_{jk}+\gamma+\kappa\delta(j,k)-1} e^{-\eta^{(i)}_{jk}}}{[\sum_{k': f_{ik'}=1} \eta^{(i)}_{jk'}]^{n^{(i)}_j}}, \tag{18}
$$

where $n^{(i)}_{jk}$ counts the transitions from state j to k in $z^{(i)}_{1:T_i}$, and $n^{(i)}_j = \sum_k n^{(i)}_{jk}$ counts all transitions out of state j.