expand S^0 to have zero mass on new components $K+1, K+2, \ldots K+J'$. Now, imagine visiting the first batch of data $\mathcal{B}_1$ and performing the E-step. After this step, the only mass on new components will come from the current batch. For each new component j: $S_j^0 = S_j(\mathcal{B}_1)$. If this batch does not assign any mass to component j, then $s_j^0 = 0$. Next, the following M-step (see the main text's Eq. (9)) will completely rewrite $\hat{\lambda}_j = \lambda_0 + 0 = \lambda_0$, reseting to its prior value. Component j will lose all information from the targeted dataset $\mathbf{x}'$ after only one update, becoming useless even though later batches may have highly preferred it.

To avoid this disaster, we choose to retain the "dual" interpretations of the data $\mathbf{x}'$ in S^* throughout the pass, which ensures every brand-new component j always has mass at least $\hat{N}'_j$. Thus, even when the first batch is not assigned at all to component j, we'll have $s_j^* = s_j(\mathbf{x}')$, and the update $\hat{\lambda}_j = \lambda_0 + s_j(\mathbf{x}')$ will retain vital information from our targeted analysis.

At the end of the adoption pass, immediately before the last M-step update to global parameters, we subtract-away all targeted summaries S' from the final global summaries S^*. This ensures that by the end of the adoption pass, both the final global summary and all global factors $q^*(v), q^*(\phi)$ have scale exactly consistent with the dataset $\mathbf{x}$. Under these conditions, the ELBO can be calculated exactly and merges can proceed.

2.4 Multiple birth moves in one pass

As a final note, performing several birth moves during one pass (refining multiple components at once) is definitely possible. We need only to collect several subsampled datasets $\mathbf{x}'_1, \mathbf{x}'_2, \ldots$, discover new components from each one via separate variational analyses, and then adopt all new components into an expanded model. For simplicity we focus on just one birth in the description below. All our experiments perform just one birth per pass, except for the final analysis of 8×8 image patches, where we execute two births per pass.