Incoming links $k_s \in \mathcal{K}_{t_{ij}}$ to a split cluster are independently assigned to the new clusters with equal probability:

$$q_{\text{in}}(\mathcal{K}_{t_{ij}}) = \prod_{k_s \in \mathcal{K}_{t_{ij}}} \left(\frac{1}{2}\right)^{\delta(k_s,t_i)} \left(\frac{1}{2}\right)^{\delta(k_s,t_j)}. \quad (9)$$

The current outgoing link is retained by one of the split clusters, $k_{t_j} = k_{t_{ij}}$. To allow likelihood-based link proposals, we *temporarily* fix the other cluster link as $k_{t_i} = t_i$.

Propose Link We compare two proposals for c_i, the ddCRP prior distribution $q(c_i) = p(c_i \mid \alpha, A)$, and a data-dependent "pseudo-Gibbs" proposal distribution:

$$q(c_i) \propto p(c_i \mid \alpha, A)\Gamma(\mathbf{x}, \mathbf{z}(c_i, \mathbf{c}_{-i}, \mathbf{k})), \quad (10)$$

$$\Gamma(\mathbf{x}, \mathbf{z}(c_i, \mathbf{c}_{-i}, \mathbf{k})) = \begin{cases} \dfrac{p(\mathbf{x}_{\mathbf{z}=m_a} \cup \mathbf{x}_{\mathbf{z}=m_b} \mid \lambda)}{p(\mathbf{x}_{\mathbf{z}=m_a} \mid \lambda) p(\mathbf{x}_{\mathbf{z}=m_b} \mid \lambda)} & \text{if } c_i \text{ merges } m_a, m_b, \\ 1 & \text{otherwise.} \end{cases}$$

The prior proposal, although naïve, can perform reasonably when A is sparse. The pseudo-Gibbs proposal is more sophisticated, as data links are proposed conditioned on both the observations $\mathbf{x}$ and the current state of the sampler. Our experiments in Sec. 4 show it is much more effective.

Merge? Let $c_i = j^*$ denote the new data link sampled according to either the ddCRP prior or Eq. (10). Relative to the reference configuration in which $c_i = i$, this link may either leave the partition structure unchanged, or cause clusters t_i and t_{j^*} to merge into t_{ij^*}. In case of a merge, the new cluster retains the current outgoing link $k_{t_{ij^*}} = k_{t_{j^*}}$, and inherits the incoming links $\mathcal{K}_{t_{ij^*}} = \mathcal{K}_{t_i} \cup \mathcal{K}_{t_{j^*}}$.

If a merge does not occur, but t_{ij} was previously split into t_i and t_j, the outgoing link $k_{t_j} = k_{t_{ij}}$ is kept fixed. For newly created cluster t_i, we then propose a corresponding cluster link k_{t_i} from its full conditional distribution:

$$q_{\text{out}}(k_{t_i}) = p(k_{t_i} \mid \alpha_0, A^0(\mathbf{c}), \mathbf{x}, \mathbf{k}_{-t_i}, \mathbf{c}). \quad (11)$$

Note that the proposal $c_i = j^*$ may leave the original partition unchanged if $c_i = i$ does not cause t_{ij} to split, and $c_i = j^*$ does not result in a merge. In this case, the corresponding cluster links are also left unchanged.

Accept or Reject Combining the two pairs of cases above, our overall proposal distribution equals

$$q(\mathbf{c}^*, \mathbf{k}^* | \mathbf{c}, \mathbf{k}, \mathbf{x}) = \begin{cases} q(c_i^*) q_{\text{in}}(\mathcal{K}^*_{t_{ij}}) & \text{split, merge,} \\ q(c_i^*) & \text{no split, merge,} \\ q(c_i^*) q_{\text{out}}(k^*_{t_i}) q_{\text{in}}(\mathcal{K}^*_{t_{ij}}) & \text{split, no merge,} \\ p(c_i^* \mid \alpha, A) & \text{otherwise.} \end{cases}$$

Here, $\mathbf{c}^*$ and $\mathbf{k}^*$ denote the proposed values, which are then accepted or rejected according to the MH rule. For acceptance ratio derivations and further details, please see the supplemental material. After cycling through all data links $\mathbf{c}$, we use the Gibbs update of Eq. (11) to resample the cluster links $\mathbf{k}$, analogously to a standard ddCRP.

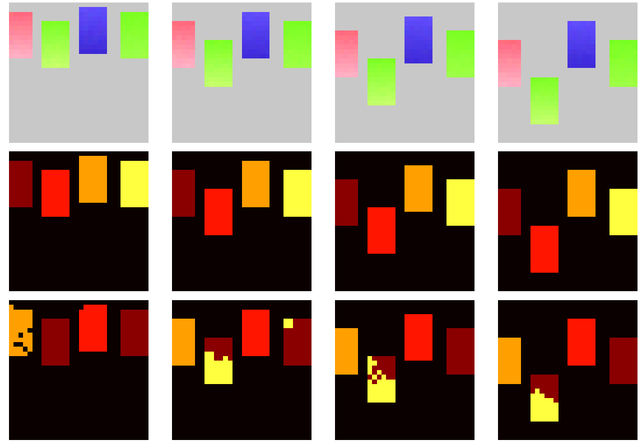

Figure 3: Experiments on synthetic data. *Top:* Ground truth partitions of a toy dataset containing four groups. Each group contains four objects exhibiting motion and color gradients. *Middle:* MAP partitions inferred by an hddCRP using size and optical flow-based cluster affinities. *Bottom:* MAP partitions discovered by a baseline hCRP using only color-based likelihoods.

4 EXPERIMENTS

In this section we present a series of experiments investigating the properties of the hddCRP model and our proposed MCMC inference algorithms. We examine a pair of challenging real-world tasks, video and discourse segmentation. We quantify performance by measuring agreement with held-out human annotations via the Rand index (Rand, 1971) and the WindowDiff metric (Pevzner & Hearst, 2002), demonstrating competitive performance.

To provide intuition, we first compare the hddCRP with the hCRP on a synthetic dataset (Figure 3) with four 30×30 frames (groups). Each frame contains four objects moving from top to bottom at different rates, and object appearances exhibit small color gradients. The hddCRP utilizes data link affinities that allow pixels (data instances) to connect to one of their eight spatial neighbors with equal probability. To exploit the differing motions of the objects, we define optical flow-based cluster affinities (Sun et al., 2010). Letting $w(t)$ denote the spatial positions occupied by cluster t after being warped by optical flow, and $w(s)$ the corresponding support of cluster s, the affinity is defined as $A^0_{ts}(\mathbf{c}) = (w(t) \cap w(s))/(w(t) \cup w(s)), t \neq s$, encouraging clusters to link to other clusters with similar spatial support. Given this affinity function, hddCRP was able to robustly disambiguate the four uniquely moving objects, while the hCRP produced noisy segmentations and consistently confused objects with local similarity but distinct motion.

4.1 VIDEO SEGMENTATION

Likelihood As a preprocessing step, we divide each frame into approximately 1200 superpixels using the