

Reliable and scalable variational inference for Bayesian nonparametric models

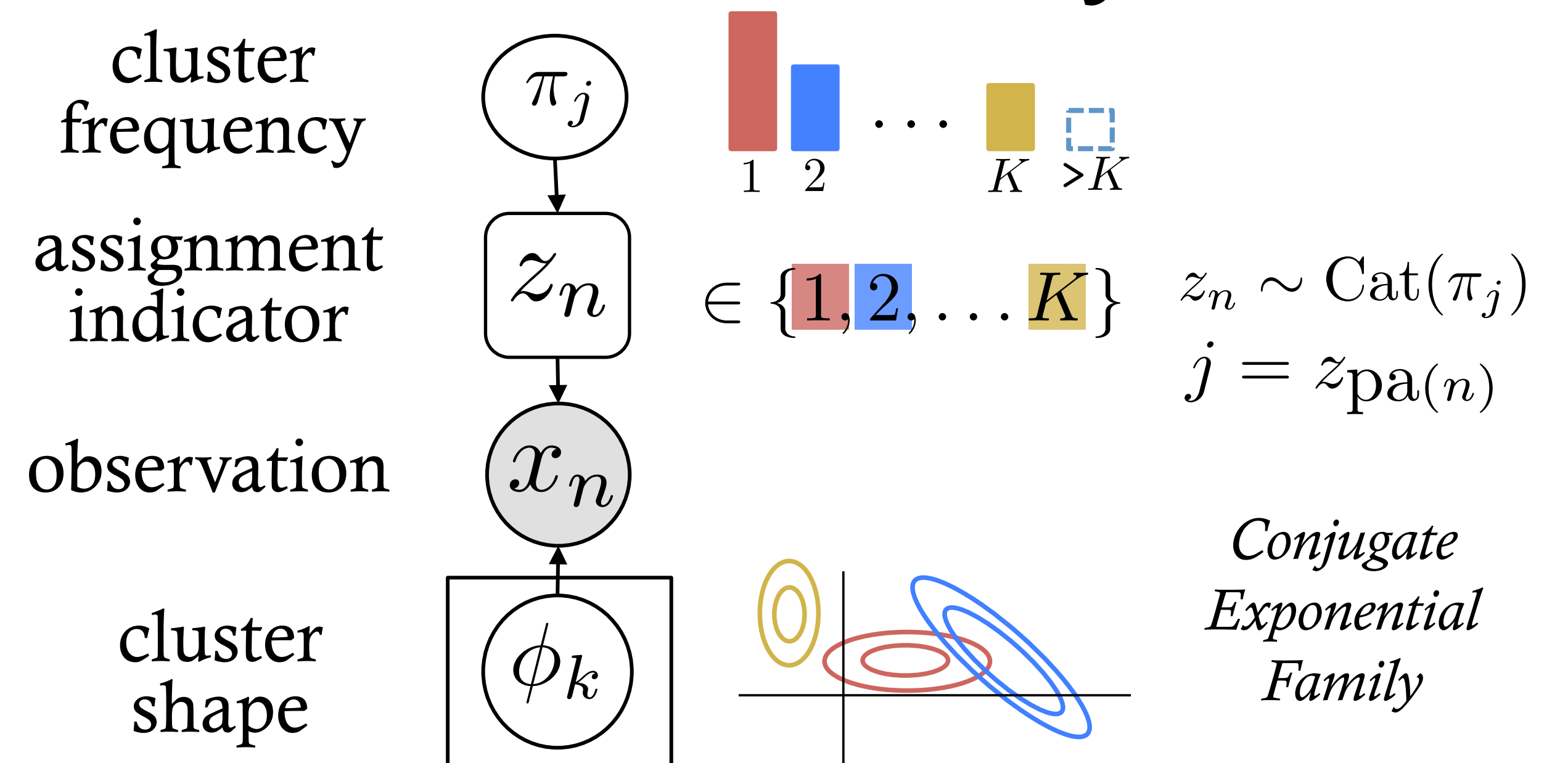
Michael C. Hughes & Erik B. Sudderth, Brown Univ. Computer Science

Goals

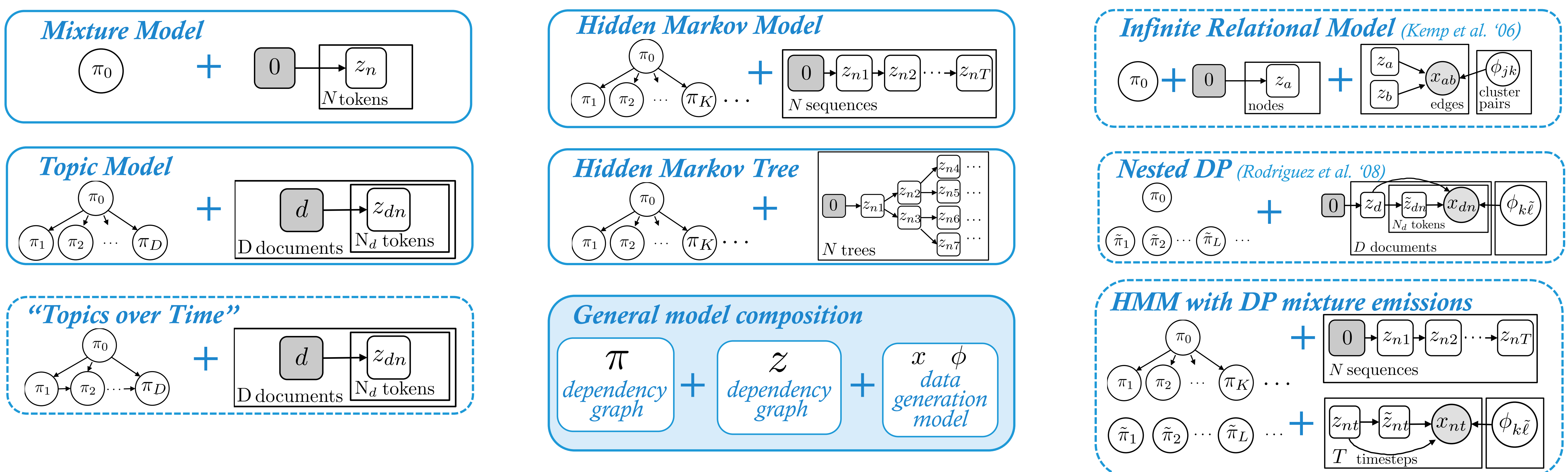
Design inference engine for a broad (but not universal) family of parametric and nonparametric models widely used in clustering applications.

- **scalable** – process big data in batches
- **reliable** – find compact, high-quality set of clusters; avoid local optima
- **expressive** – hierarchical/sequential/spatial structure

Model Family



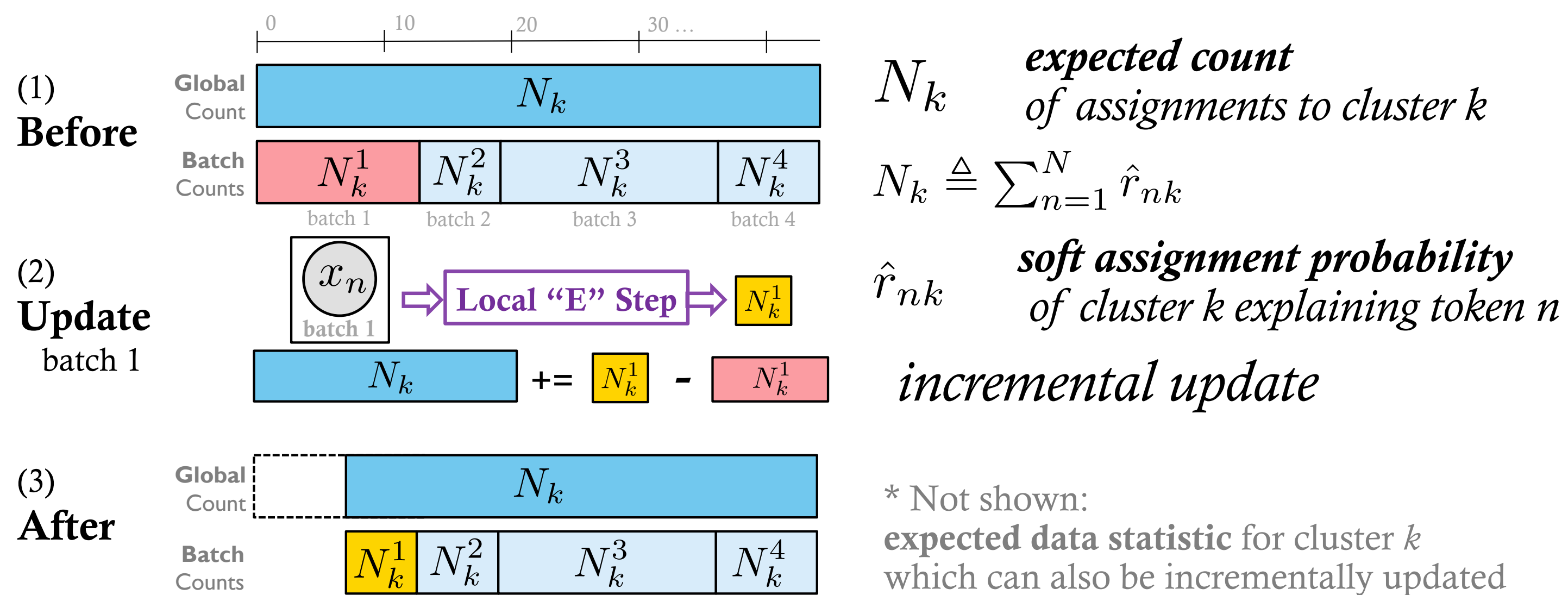
Examples



Scalable Inference

Memoized variational algorithm. (Hughes & Sudderth '13) (Neal & Hinton '99)

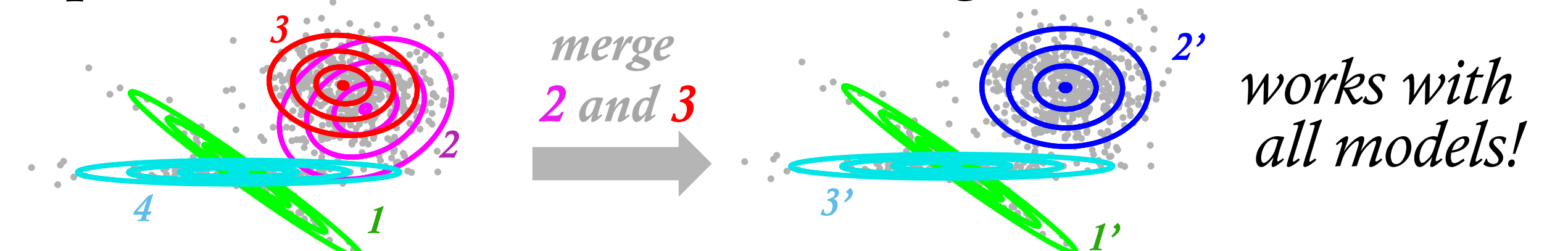
- process huge data as small, fixed batches
- no pesky learning rates
- track exact sufficient statistics for entire dataset



Reliable Inference

Merge moves for simpler models, faster learning.

- Propose candidate model combining two clusters into one



- Accept or reject proposal via whole-dataset objective
- Construct candidate by simple addition rule

$$\mathcal{L}(\text{merge}) < \mathcal{L}(\text{original}) \quad \text{memoized - compute } \mathcal{L} \text{ exactly}$$

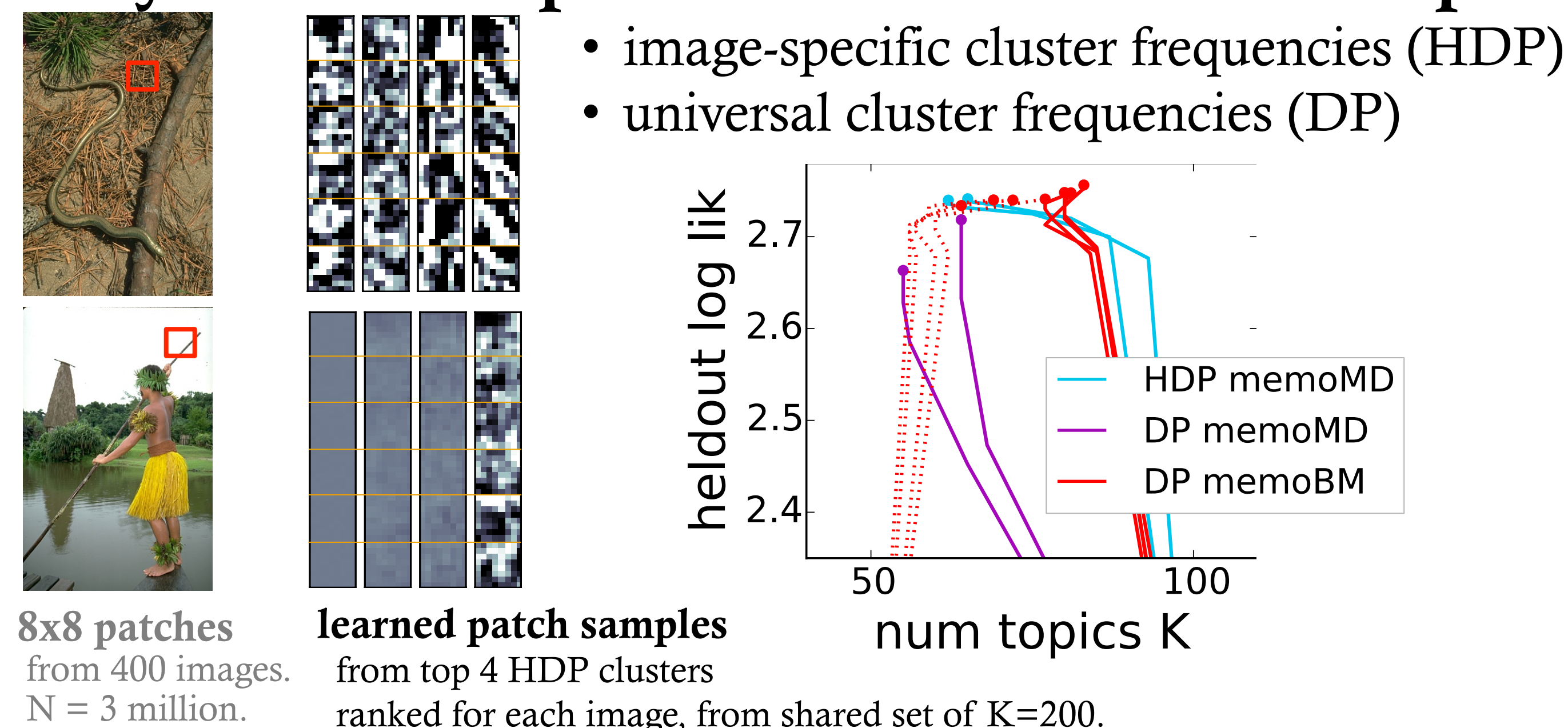
$$\hat{r}_{n2} + \hat{r}_{n3} \rightarrow \hat{r}'_{n2} \quad \text{soft assignments}$$

$$N_2 + N_3 \rightarrow N'_2 \quad \text{statistics follow by additivity}$$

* Birth and delete moves also possible

Image Patches

Easy model comparison on millions of examples.



Topic Models

Merge/delete moves find compact set of clusters.

