Dependent Dirichlet Processes based on Poisson Processes

Dahua Lin
CSAIL, MIT
dhlin@mit.edu

Eric Grimson
CSAIL, MIT
welg@csail.mit.edu

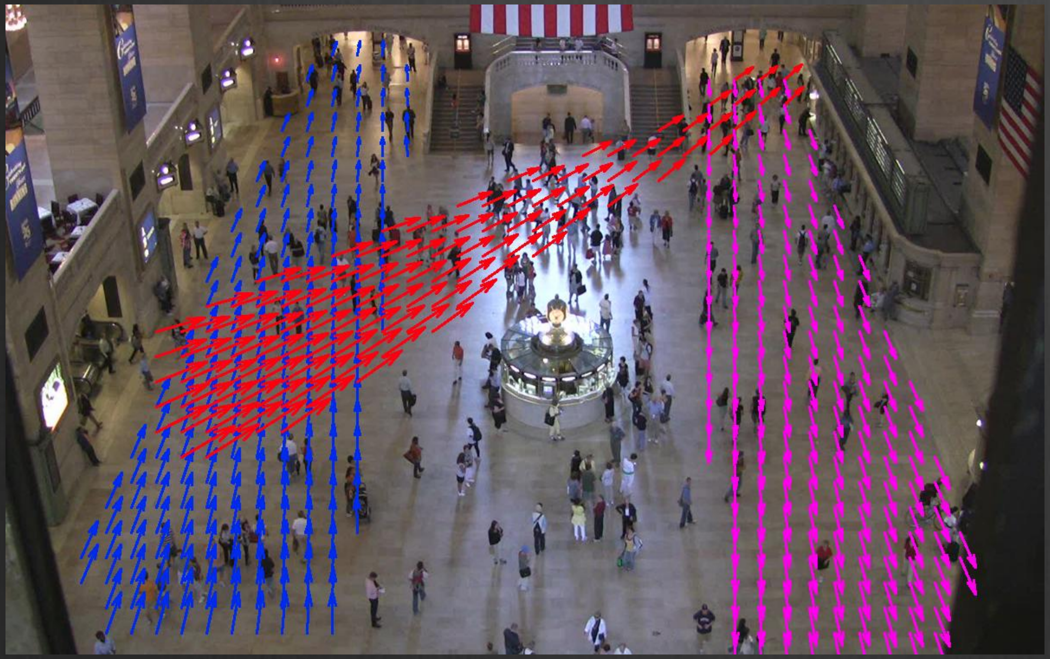
John Fisher
CSAIL, MIT
fisher@csail.mit.edu

Mixture Models : From Static to Dynamic

Document modeling

Financial analysis

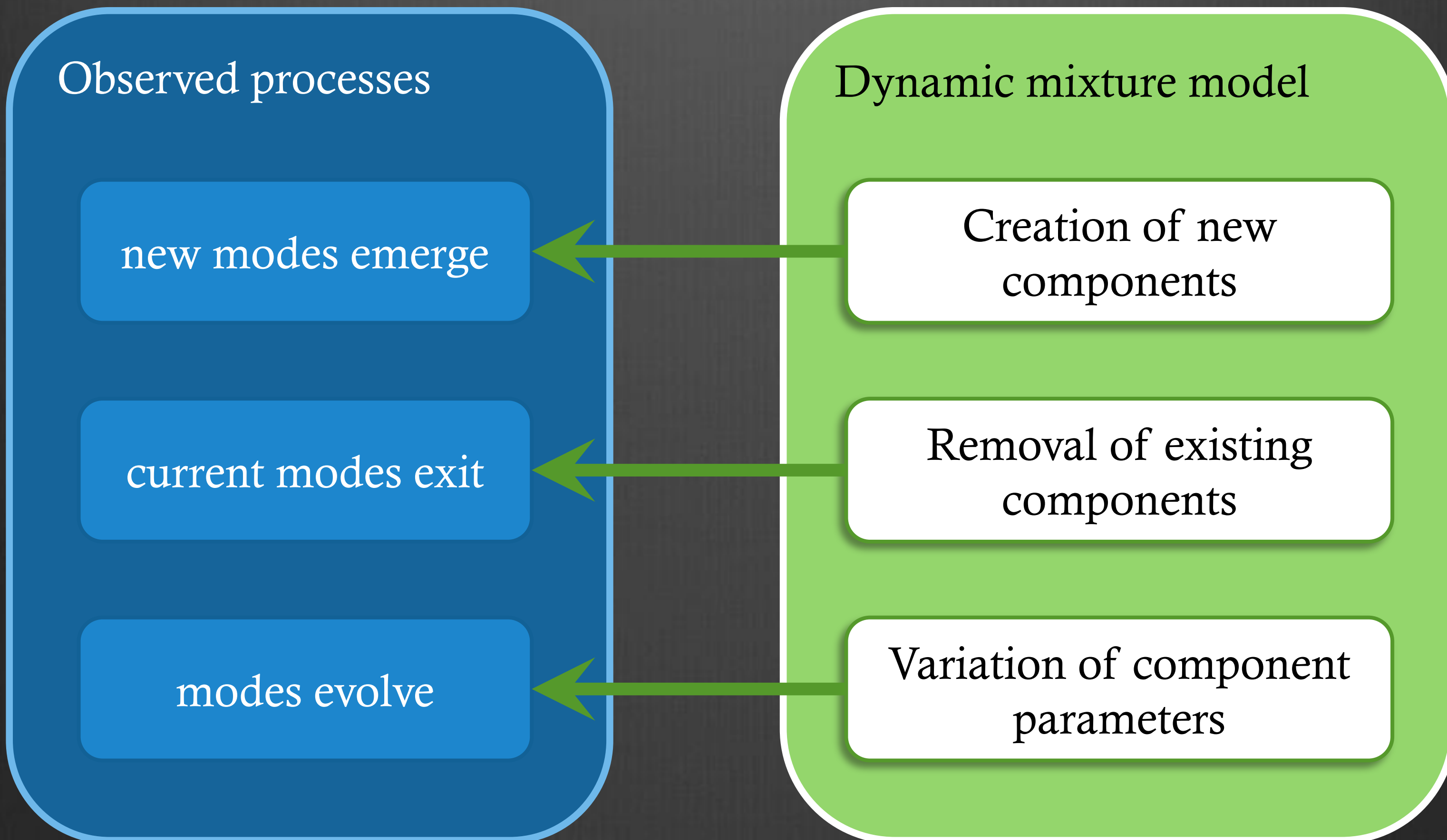
Image understanding



?

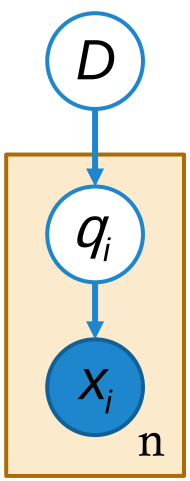
How to do mixture modeling in response to changes?

Dynamic Mixture Models



Dirichlet Processes

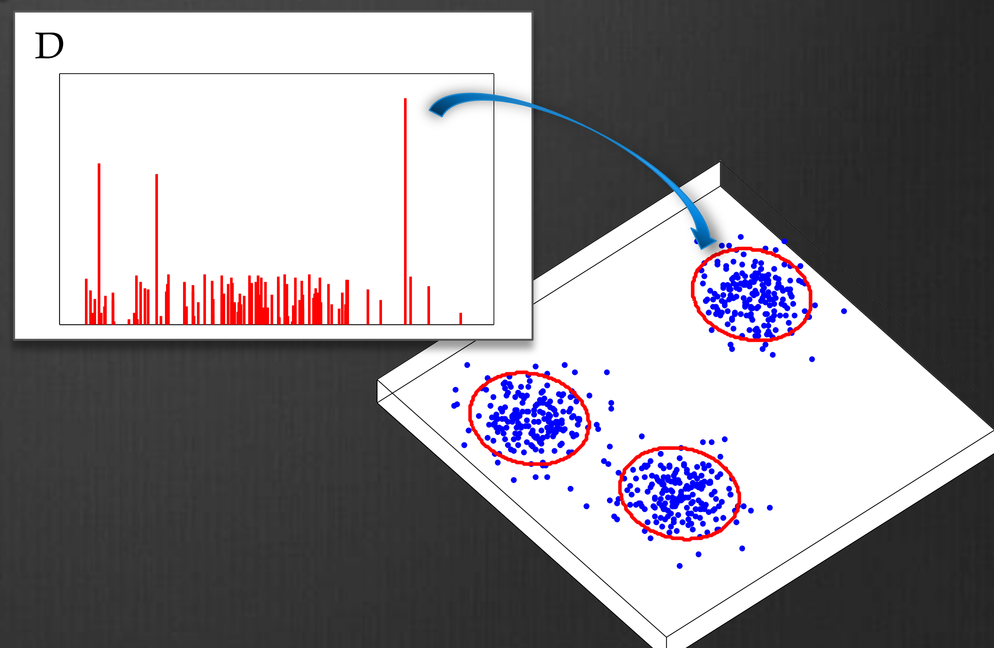
- Mixture Models
  - **Finite mixture model (FMM):** pre-specified number of components. *(see Everitt and Hand, 1981)*
  - **Dirichlet process mixture model (DPMM):** allows indefinite number of mixture components. *(see Rasmussen, 2000, and Neal, 2000)*



$$D \sim DP(\mu)$$

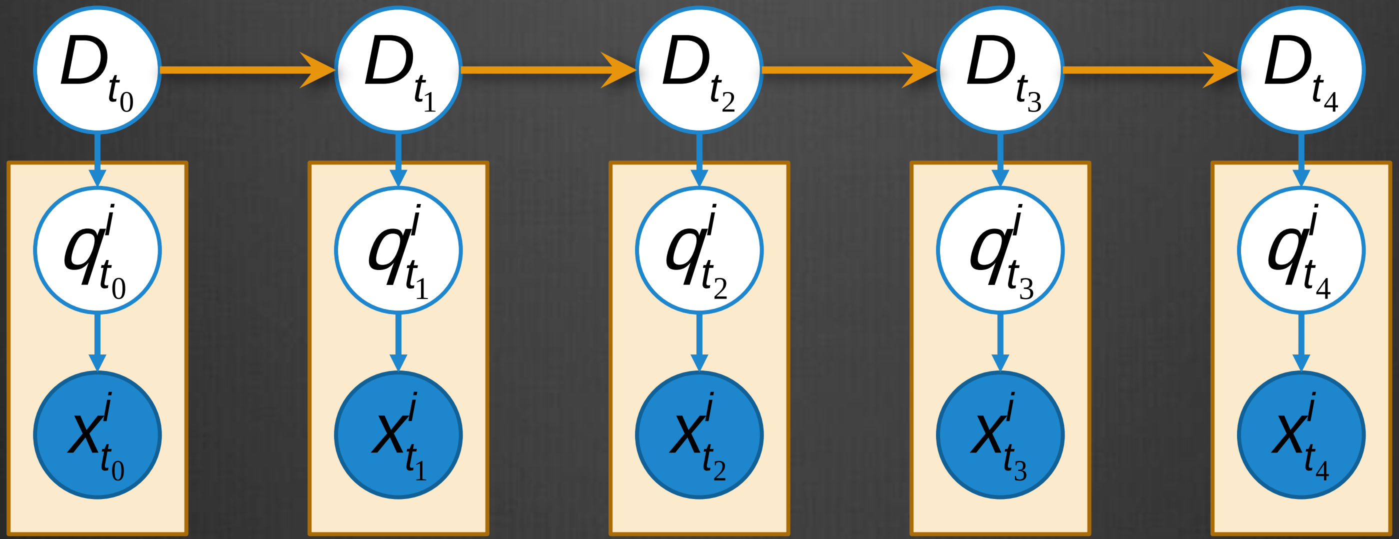
$$\theta_i \sim D$$

$$x_i \sim G(\theta_i)$$



From DP to Dependent DP

- Extend DPMM to model a dynamic process



Central Problem: introduce **dependency** between Dirichlet processes.

Important to maintain the property of being marginal DP.

Why a New Approach?

- Related work
  - Single-p DDP *(McEachern, 99)*
  - Time-sensitive DP *(Zhu and Lafferty, 05)*
  - Hierarchical DP *(Teh et. al, 06)*
  - Dynamic HDP *(Ren et. al, 08)*
  - Generalized Pólya Urn *(Caron et. al, 07)*
  - Recurrent CRP *(Ahmed and Xing, 08)*
  - πDDP *(Griffin and Steel, 06)*
  - Local DP *(Chung and Dunson, 09)*
  - Spatially normalized Gamma processes *(Rao and Teh, 09)*

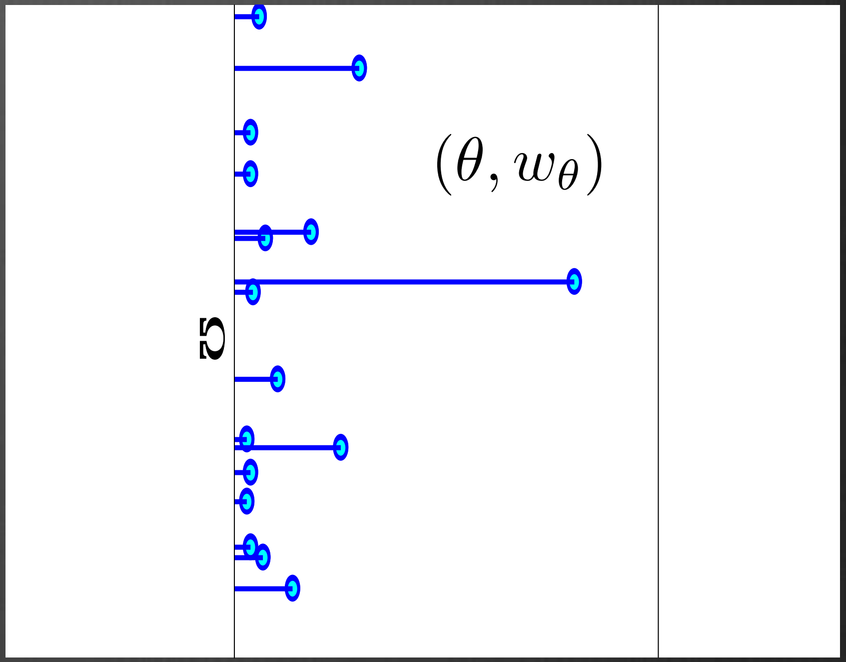
Poisson, Gamma, and Dirichlet

Given a measure space $(\Omega, \mathcal{F}, \mu)$

Poisson process (over $\Omega \times R^+$):

$$\Pi^* \sim \mathrm{PoissonP}(\mu \times \gamma)$$

$$\gamma(dw) = w^{-1} e^{-w} dw$$



Gamma process:

$$G \triangleq \sum_{(\theta, w_\theta) \in \Pi^*} w_\theta \delta_\theta \sim \Gamma P(\mu)$$

Dirichlet process (Normalized Gamma process):

$$D \triangleq G / G(\Omega) \sim DP(\mu)$$

Our Approach

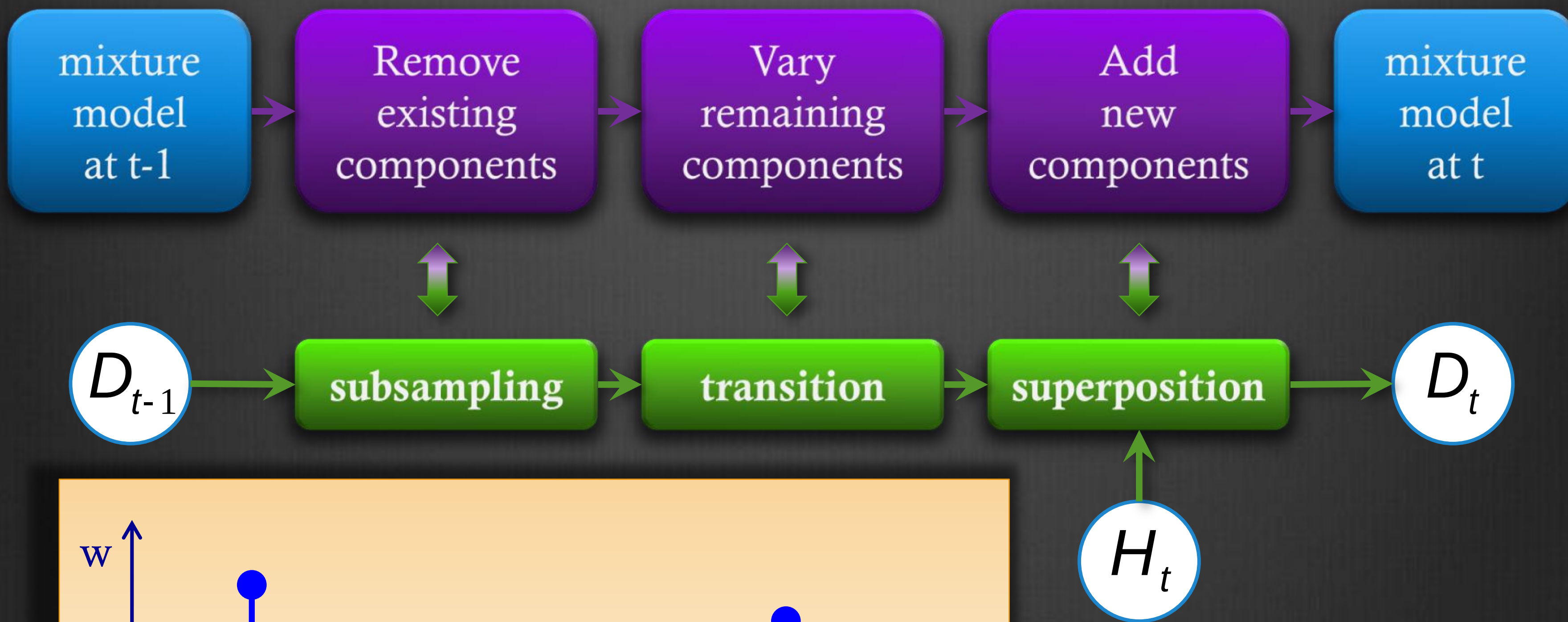
Completely Random Measure

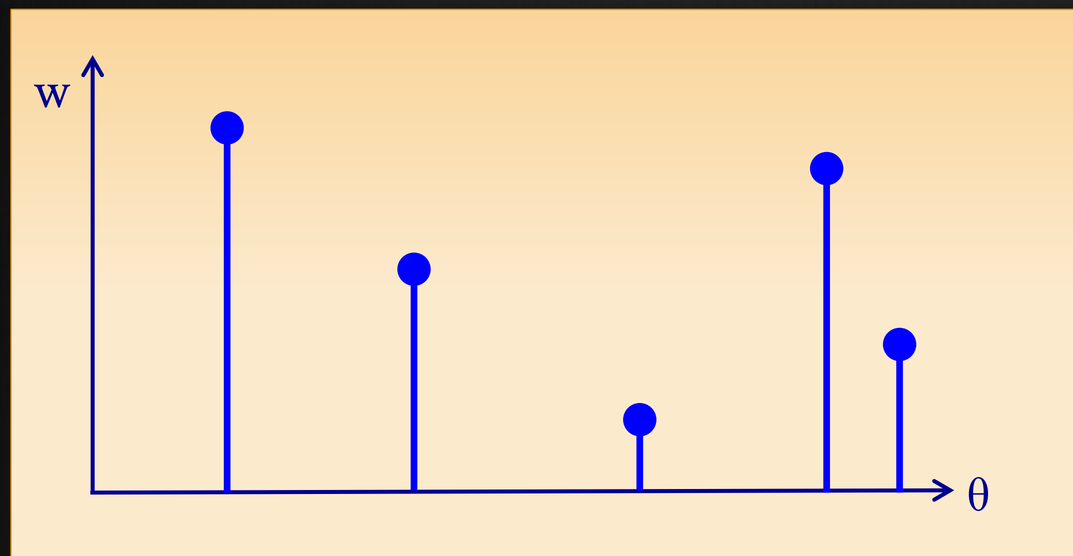
A random measure of which the measure values of disjoint subsets are independent.

Complete Randomness Preserving Operations

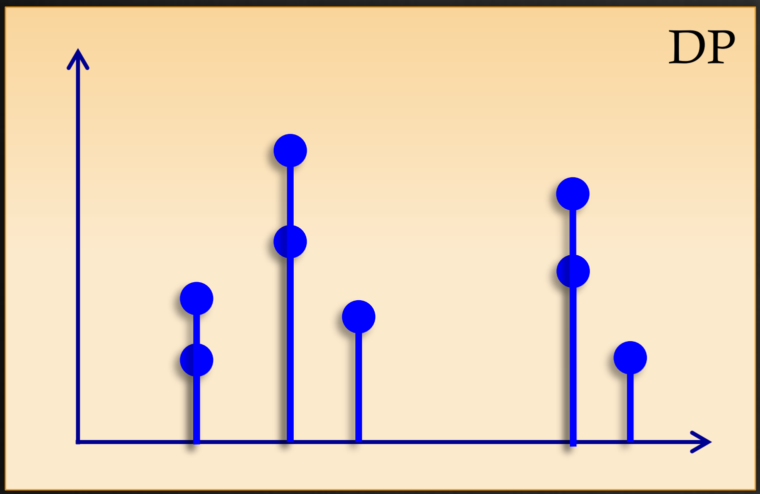
Applying any operations that preserve complete randomness to Poisson processes results in a new Poisson process.

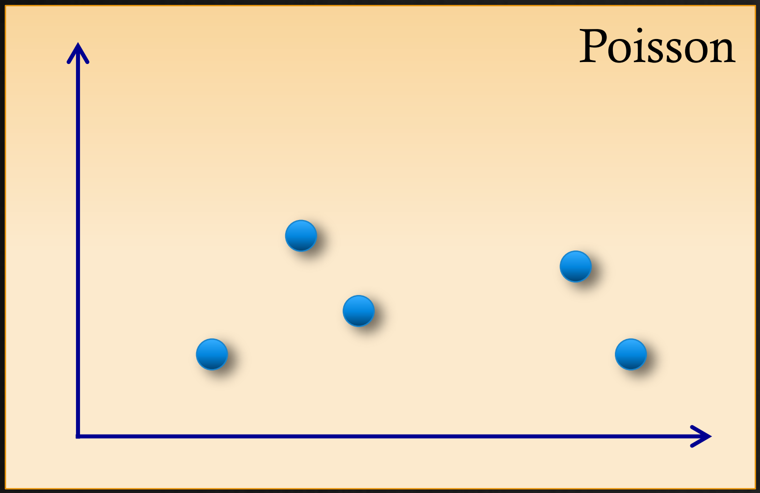
Construct a Chain of DPs





Subsampling





Equivalent operations directly on a DP

$$D = \sum_{\eta} w_\theta \delta_\theta \Rightarrow$$

$$\mathcal{S}_q(D) \triangleq \sum_{z_\eta = 1} w'_\theta \delta_\theta \sim DP(q\mu)$$

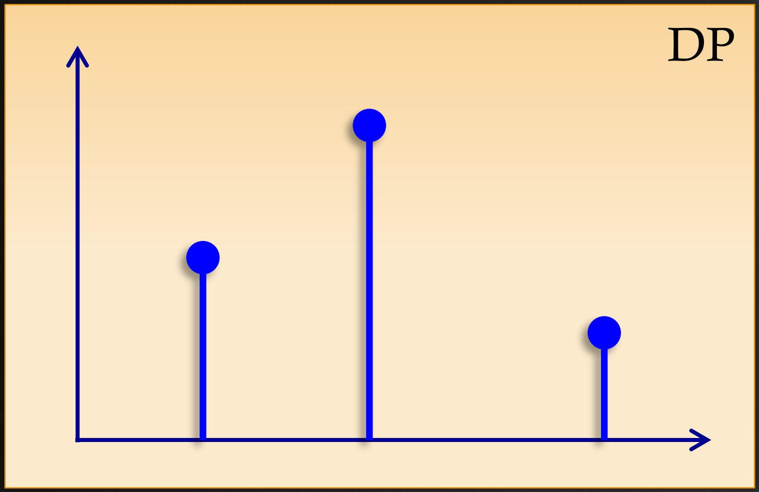
$$w'_\theta = w_\theta / \sum_{z_\eta = 1} w_\theta$$

Subsampling via Independent Bernoulli Trial:

$$\forall\ \eta = (\theta, w_\theta) \in \Pi^*,\ \ z_\eta \sim \text{Bernoulli}(q)$$

$$\mathcal{S}_q(\Pi^*) \triangleq \{\eta \in \Pi^* : z_\eta = 1\} \sim \text{PoissonP}(q\mu)$$

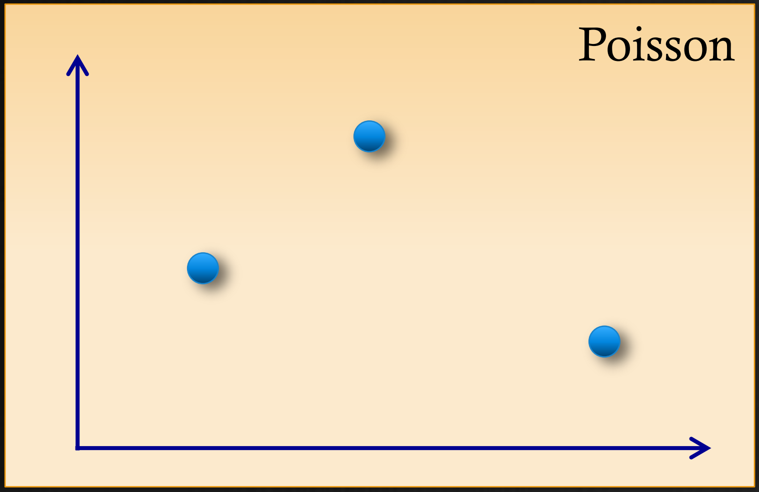
Transition



Equivalent operation directly on a DP

$$D = \sum_{\eta \in \Pi^*} w_\theta \delta_\theta \Rightarrow$$

$$T(D) \triangleq \sum_{\eta \in \Pi^*} w_\theta \delta_{T(\theta)} \sim DP(T\mu)$$



Independent movement of each point

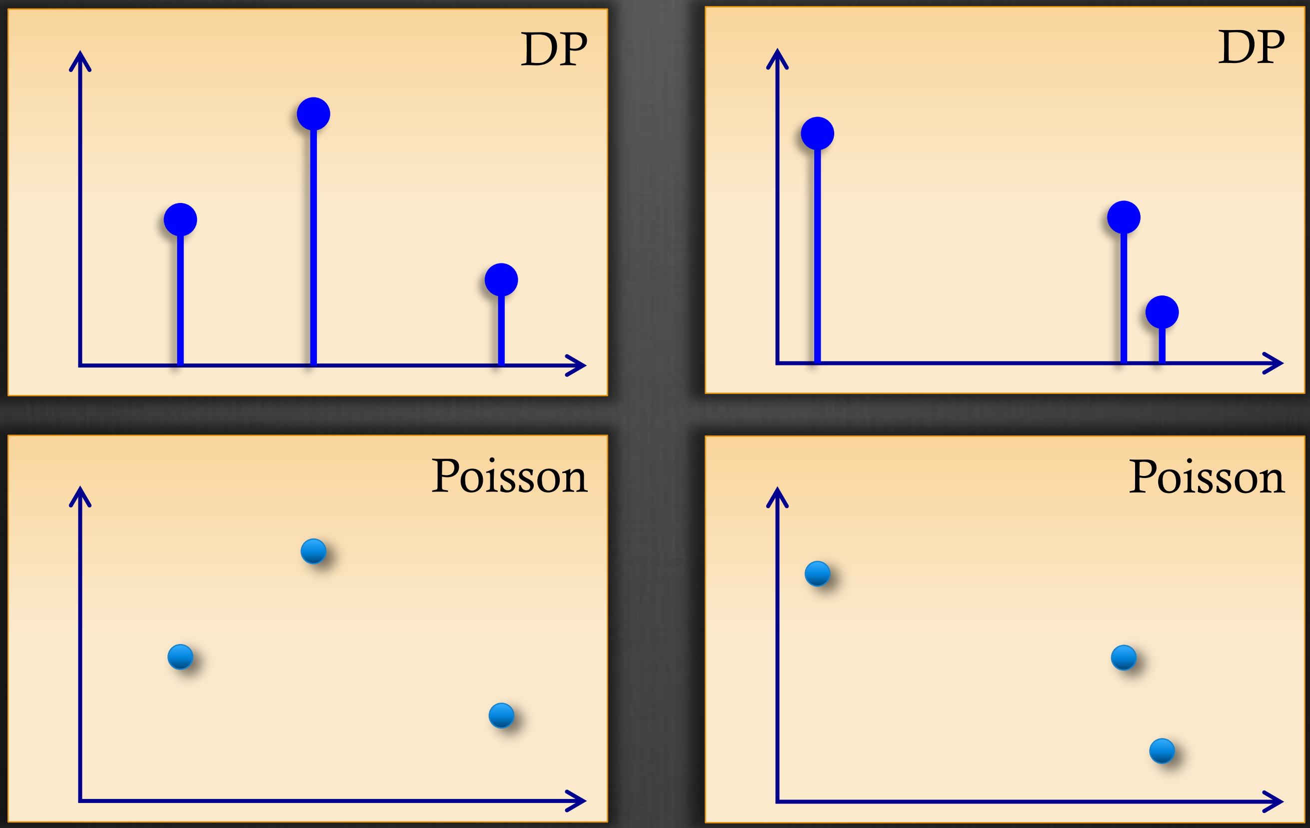
T : probabilistic transition kernel

$$T(\Pi^*) = \{(T(\theta), w_\theta) : (\theta, w_\theta) \in \Pi^*\}$$

$$\sim \text{PoissonP}(T\mu)$$

$$\text{with } (T\mu)(A) = \int_\Omega T(\theta, A)\mu(d\theta), \ \forall\ A \in \mathcal{F}$$

Superposition

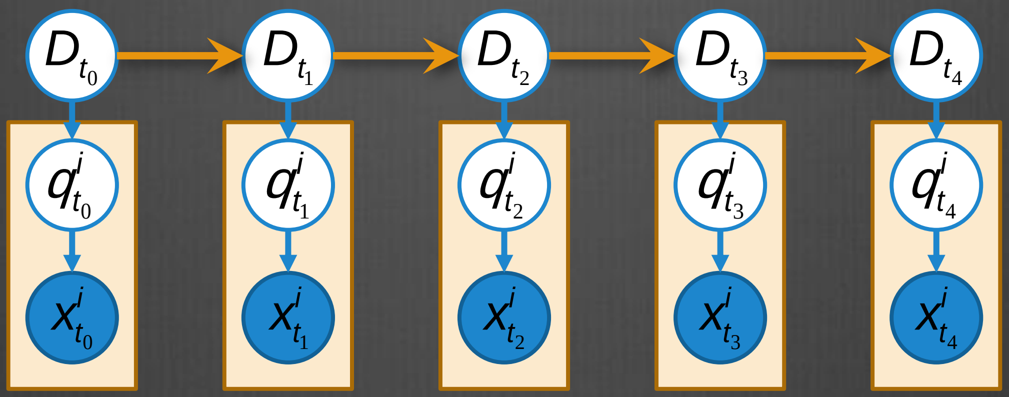


$D_k \sim DP(\mu_k), k = 1, \ldots, m$ be independent ,

$(c_1, \ldots, c_m) \sim \mathrm{Dir}(\mu_1(\Omega), \ldots, \mu_m(\Omega))$

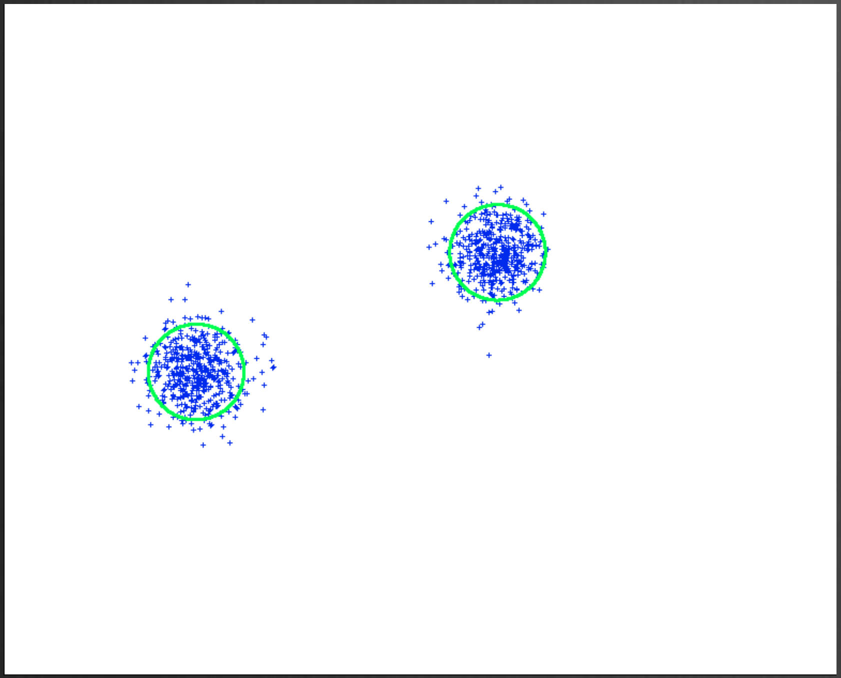
$\Rightarrow c_1 D_1 + \cdots + c_m D_m \sim DP(\mu_1 + \cdots + \mu_m)$

Inference

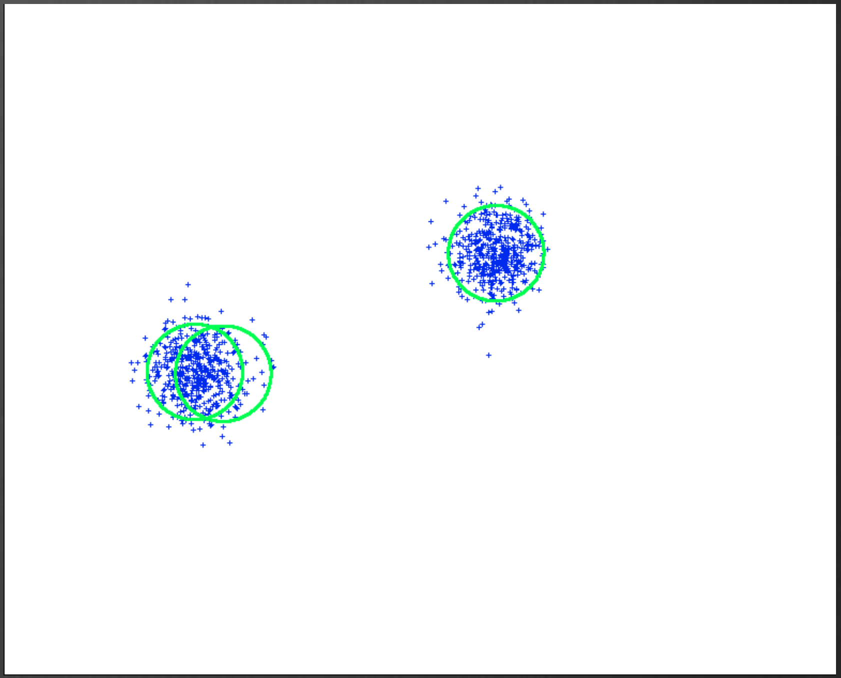


- Combine operations → Markov Chain of DPs → Dynamic mixture model.
- Gibbs sampling algorithm for inference from observations.
  - Sequential sampling (particle filtering, etc.)
  - The sampling in each phase generalizes CRP
- Fundamental difference in theoretical foundation.

Simulation
Compare with FMM

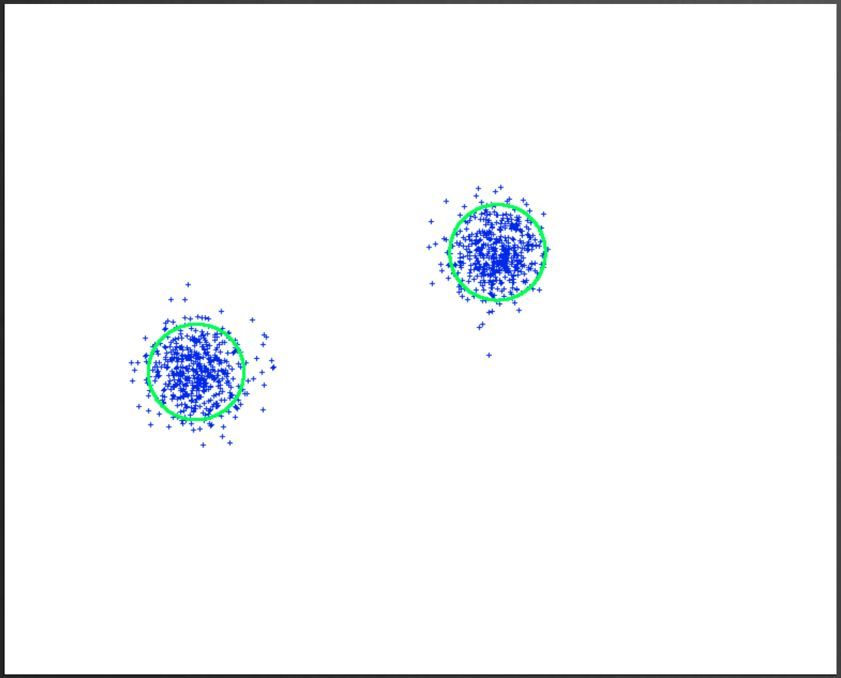


Our approach

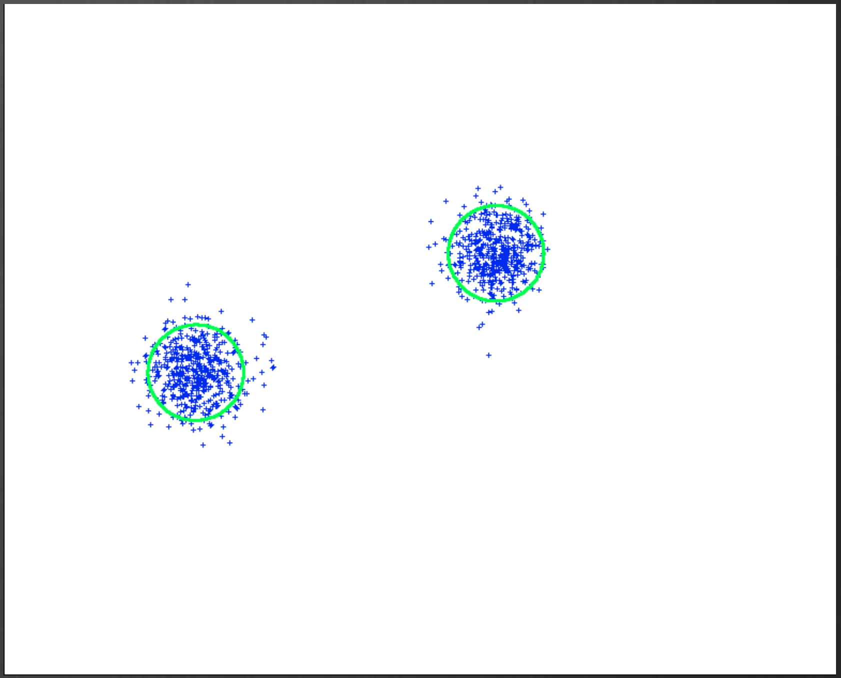


D-FMM with K = 3

Simulation
Compare with Independent DPs



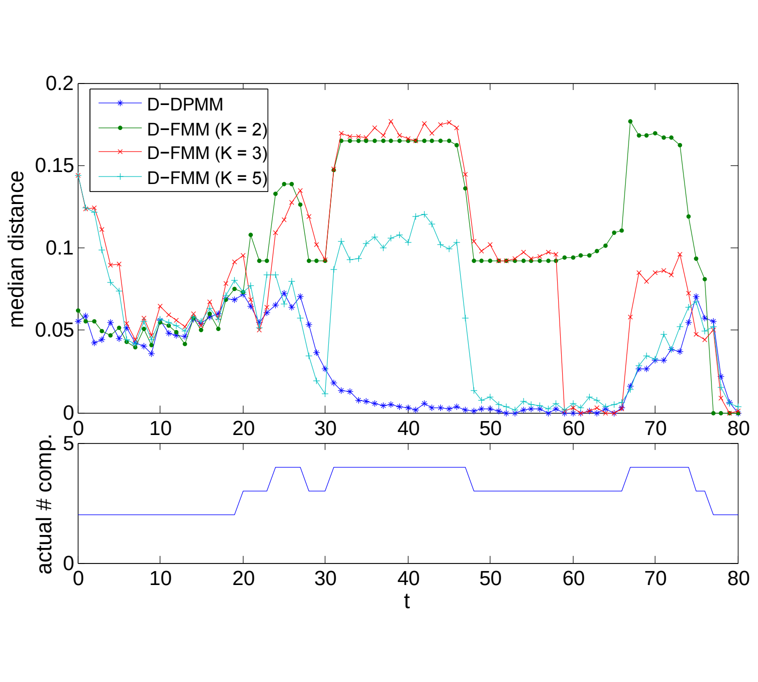
Our approach

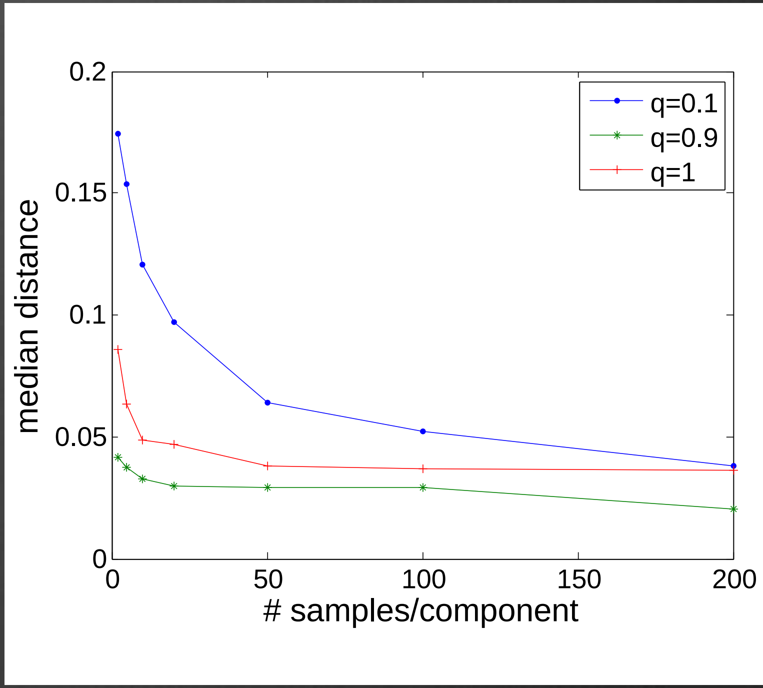


Independent DPs

The estimation based on independent DPs does not maintain component identities across phases, making it difficult when we intend to study the evolution of a particular component.

Empirical Comparison





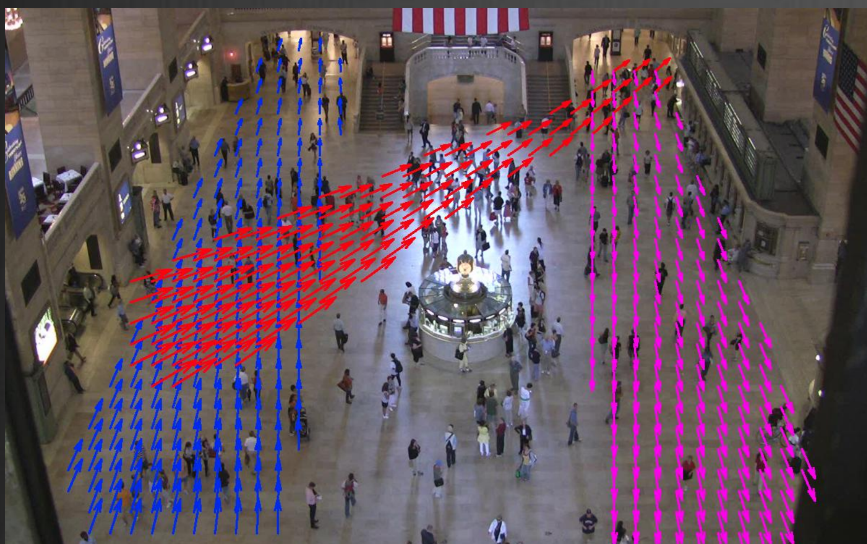
Real Applications

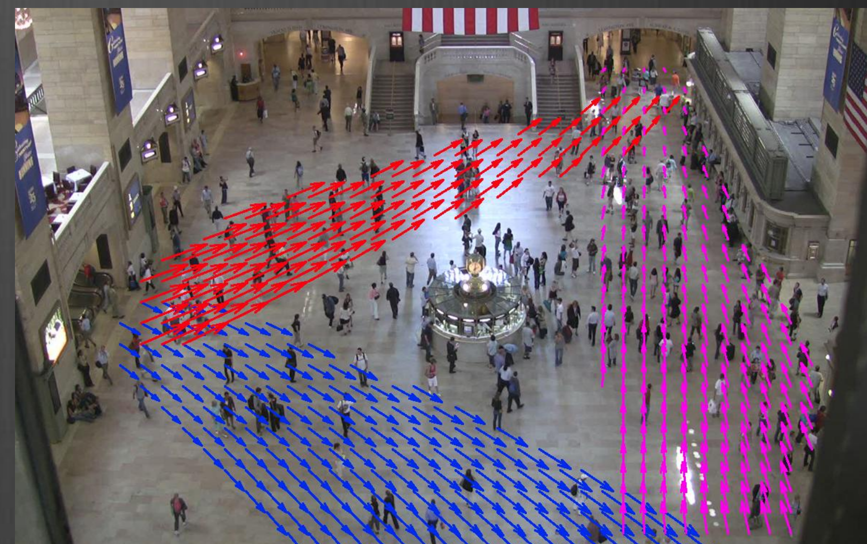
- Modeling people flows in a rail station
- Modeling the evolution of the trends of research topics *(presented in poster)*

Modeling People Flows

- The motion of people in New York Grand Central station.
- Data: 90,000 frames in one hour, divided into 60 phases
- Try to group people tracks into flows depending on their motion patterns

Results





D_1 → D_2 → …… → D_t →

Summary

- Propose a principled methodology to construct dependent Dirichlet processes based on the theoretical connections between Poisson, Gamma and Dirichlet processes.
- Develop a framework of evolving mixture model, which allows creation and removal of mixture components, as well as variation of parameters.
- Derive a Gibbs sampling algorithm for inferring mixture model parameters from observations.
- Test the approach on both synthetic data and real applications.

THANK YOU!

More details are provided at poster W84