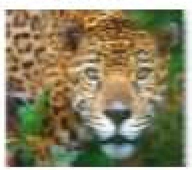
Approximate Inference in Continuous Determinantal Point Processes

Raja Affandi Emily Fox Ben Taskar

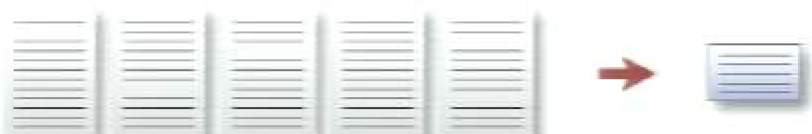
DPPs in ML have focused on *diverse subset selection* from **discrete** sets

- Diverse image search


…

- Document summarization



- Detecting multiple object trajectories

⋮

Ω

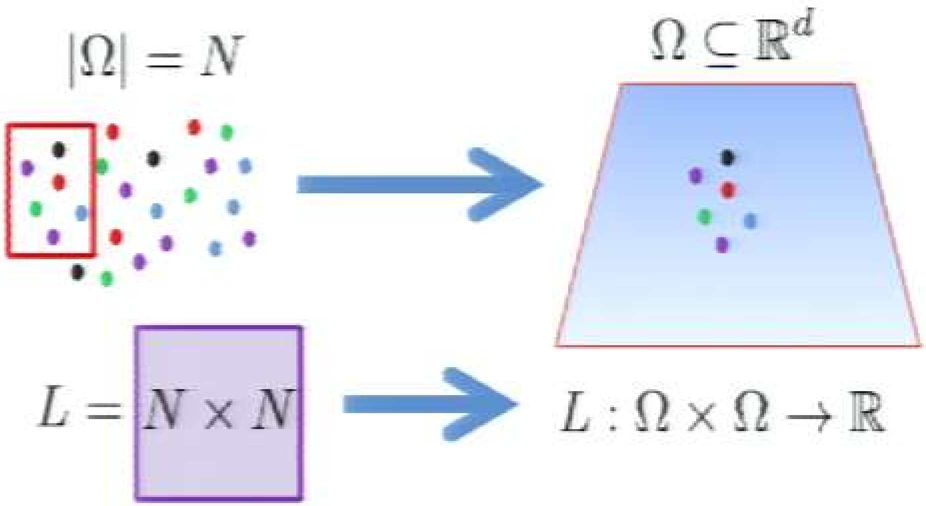
$$\mathcal{P}_L(A \subset \Omega) \propto \det(L_A)$$

Many applications in **continuous** spaces

- Mixture modeling
- Ecological modeling/ spatial stats

⋮

DPPs on Continuous Spaces



In discrete case, sampling is $O(N^3)$. [Hough et. al. , 2006]

- only need eig(L)
- makes DPPs useful in practice

How do we sample in continuous case?
Need to:

1. eigendecompose kernel function
2. sample from densities based on eigenfunctions (**HARD!**)

Method 1: Low Rank Approx. + Dual Samp.

→ Independent approx. samples

Kernel function $L(\mathbf{x}, \mathbf{y}) = \sum_{n=1}^{\infty} \lambda_n \phi_n(\mathbf{x}) \overline{\phi_n}(\mathbf{y})$

Low rank approx. $\tilde{L}(\mathbf{x}, \mathbf{y}) = B(\mathbf{x})^* B(\mathbf{y})$

Nyström Method

Random Fourier Features

Dual kernel matrix $\boxed{C}$ $D \times D$

Amazingly, eig(C) ↔ eig($\tilde{L}$).
So sampling just as in discrete case ... $O(D^3)$.

Method 2: Gibbs Samp. for fixed- sized DPP

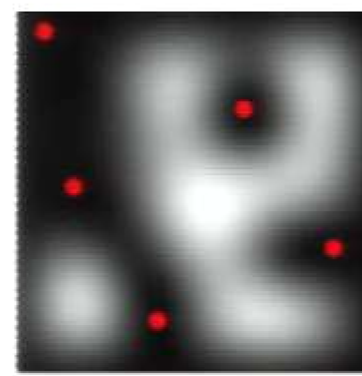
→ Dependent approx. samples

Iteratively sample

$$x_k \sim p\left(x_k \middle| \{x_j\}_{\{j \neq k\}}\right)$$

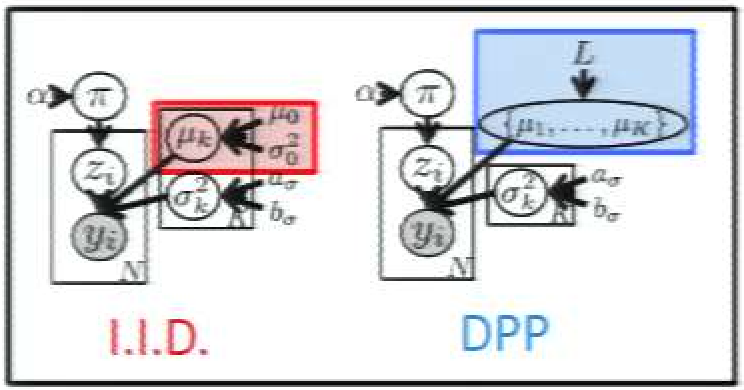
$\parallel$

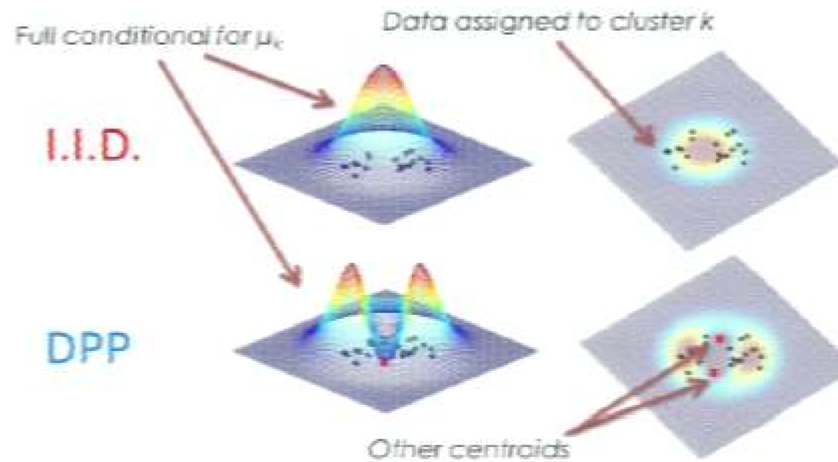
tilted 1-DPP

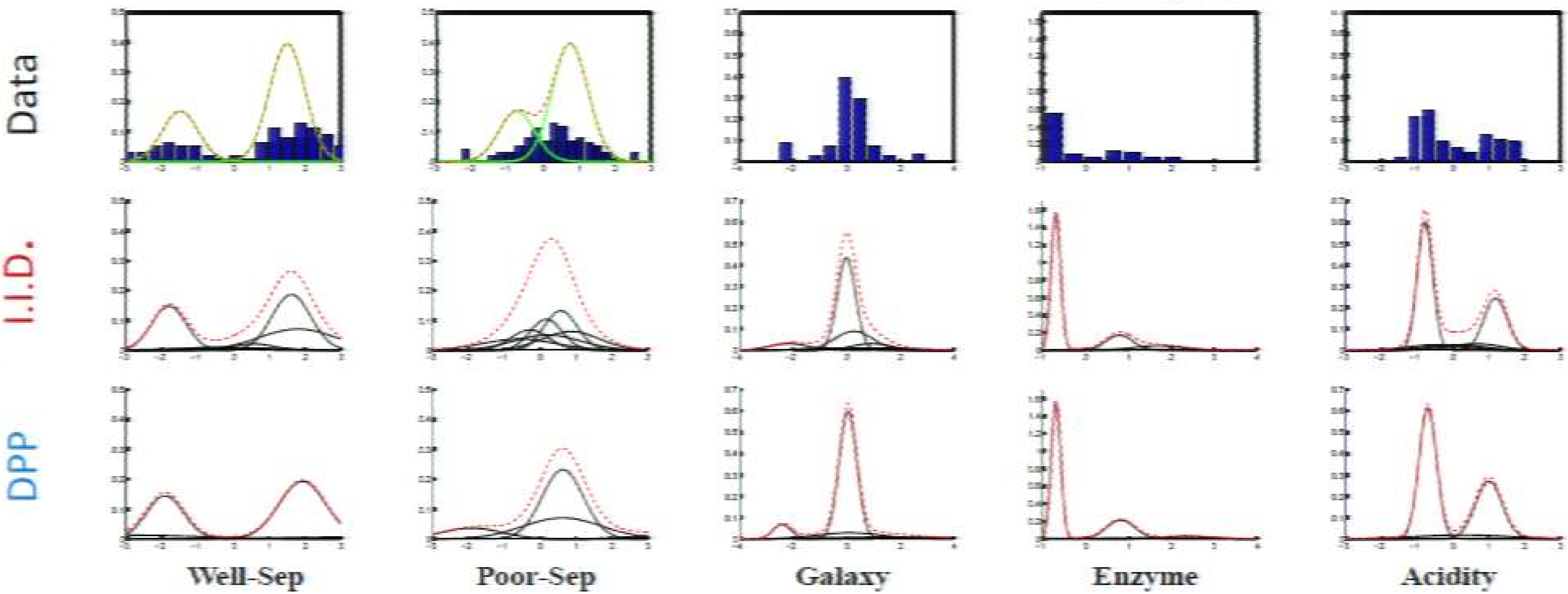


- derived using Schur's determinantal formula

Mixture of Gaussian with DPP Prior







Leads to higher accuracy in classification tasks, as well !!

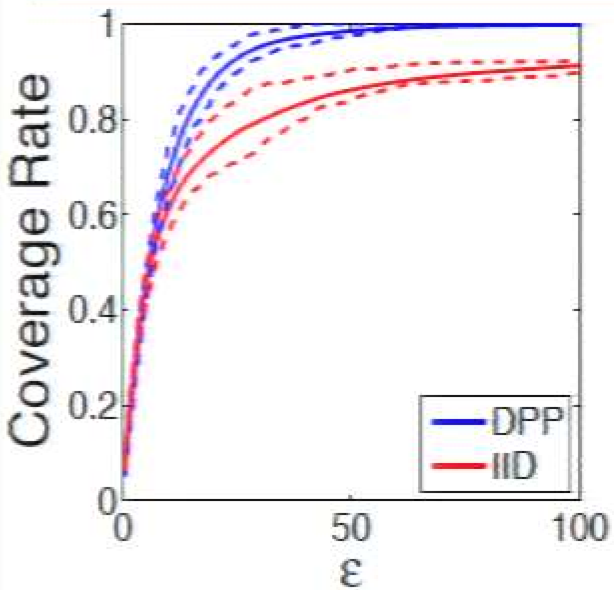
Synthesizing Human Pose

Behavior:
dance

Original

DPP samples

Random samples



In 62 dimensions! Complexity of sampling scales linearly with *d*

Better coverage of space than random sampling