Probabilistic
Graphical
Models

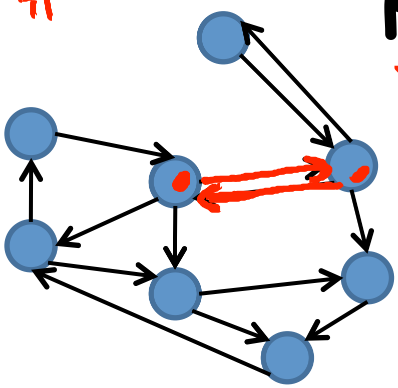
Inference

Sampling Methods

Metropolis-Hastings Algorithm

Reversible Chains

π



$$\boxed{\pi(\boldsymbol{x})T(\boldsymbol{x} \rightarrow \boldsymbol{x}') = \pi(\boldsymbol{x}')T(\boldsymbol{x}' \rightarrow \boldsymbol{x})}$$

detailed balance

Theorem: If detailed balance holds, and T is regular, then T has a unique stationary distribution π

Proof:

$$\sum_{\boldsymbol{x}} \pi(\boldsymbol{x})T(\boldsymbol{x} \rightarrow \boldsymbol{x}') = \sum_{\boldsymbol{x}} \pi(\boldsymbol{x}')T(\boldsymbol{x}' \rightarrow \boldsymbol{x}) = \pi(x') \cdot \underbrace{\sum_{x} T(x' \rightarrow x)}_{=1}$$

$$\longrightarrow \sum_{\boldsymbol{x}} \pi(\boldsymbol{x})T(\boldsymbol{x} \rightarrow \boldsymbol{x}') = \pi(\boldsymbol{x}')$$

definition of π

Daphne Koller

Metropolis Hastings Chain

Proposal distribution $Q(\mathbf{x} \rightarrow \mathbf{x}')$

Acceptance probability: $A(\mathbf{x} \rightarrow \mathbf{x}')$

- At each state $\mathbf{x}$, sample $\mathbf{x}'$ from $Q(\mathbf{x} \rightarrow \mathbf{x}')$
- Accept proposal with probability $A(\mathbf{x} \rightarrow \mathbf{x}')$
  - If proposal accepted, move to $\mathbf{x}'$
  - Otherwise stay at $\mathbf{x}$

$$T(\mathbf{x} \rightarrow \mathbf{x}') = \quad Q(\mathbf{x} \rightarrow \mathbf{x}')\, A(\mathbf{x} \rightarrow \mathbf{x}') \qquad \text{if } x' \neq x$$

$$T(\mathbf{x} \rightarrow \mathbf{x}) = \quad Q(\mathbf{x} \rightarrow \mathbf{x}) + \Sigma_{\mathbf{x}' \neq \mathbf{x}}\, Q(\mathbf{x} \rightarrow \mathbf{x}')\, (1\text{-}A(\mathbf{x} \rightarrow \mathbf{x}'))$$

Daphne Koller

Acceptance Probability

$$\pi(\boldsymbol{x})T(\boldsymbol{x} \to \boldsymbol{x}') = \pi(\boldsymbol{x}')T(\boldsymbol{x}' \to \boldsymbol{x})$$

construct A s.t. holds for Q, π

$x \neq x'$ $$\pi(\boldsymbol{x})Q(\boldsymbol{x} \to \boldsymbol{x}')A(\boldsymbol{x} \to \boldsymbol{x}') = \pi(\boldsymbol{x}')Q(\boldsymbol{x}' \to \boldsymbol{x})A(\boldsymbol{x}' \to \boldsymbol{x})$$

$A(x \to x') = \rho$
$A(x' \to x) = 1$

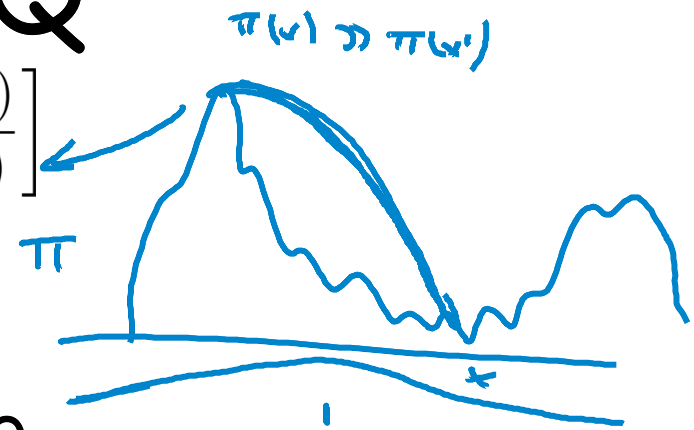
$$\frac{A(\boldsymbol{x} \to \boldsymbol{x}')}{A(\boldsymbol{x}' \to \boldsymbol{x})} = \boxed{\frac{\pi(\boldsymbol{x}')Q(\boldsymbol{x}' \to \boldsymbol{x})}{\pi(\boldsymbol{x})Q(\boldsymbol{x} \to \boldsymbol{x}')} = \rho < 1}$$

$$\boxed{A(\boldsymbol{x} \to \boldsymbol{x}') = \min\left[1, \frac{\pi(\boldsymbol{x}')Q(\boldsymbol{x}' \to \boldsymbol{x})}{\pi(\boldsymbol{x})Q(\boldsymbol{x} \to \boldsymbol{x}')}\right]}$$

Daphne Koller

Choice of Q

$$\mathcal{A}(\boldsymbol{x} \to \boldsymbol{x}') = \min\left[1, \frac{\pi(\boldsymbol{x}')Q(\boldsymbol{x}' \to \boldsymbol{x})}{\pi(\boldsymbol{x})Q(\boldsymbol{x} \to \boldsymbol{x}')}\right]$$

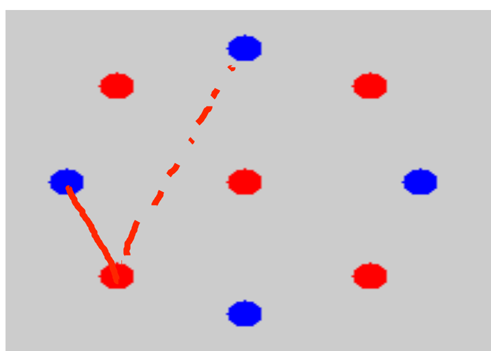


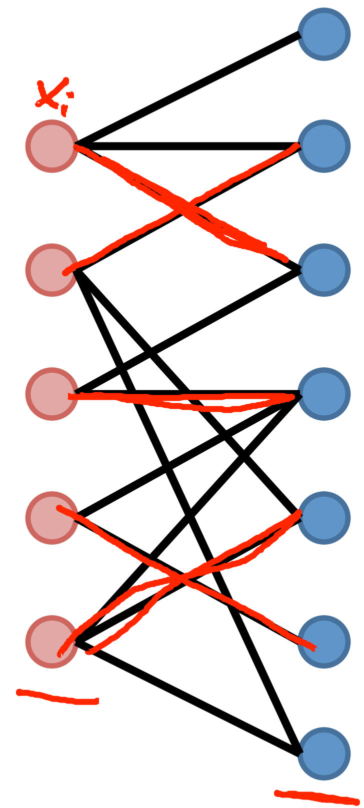
- Q must be reversible:
  - $Q(\mathbf{x} \to \mathbf{x}') > 0 \Rightarrow Q(\mathbf{x}' \to \mathbf{x}) > 0$
- Opposing forces
  - Q should try to spread out, to improve mixing
  - But then acceptance probability often low

MCMC for Matching

$\boxed{X_i = j}$ if i matched to j

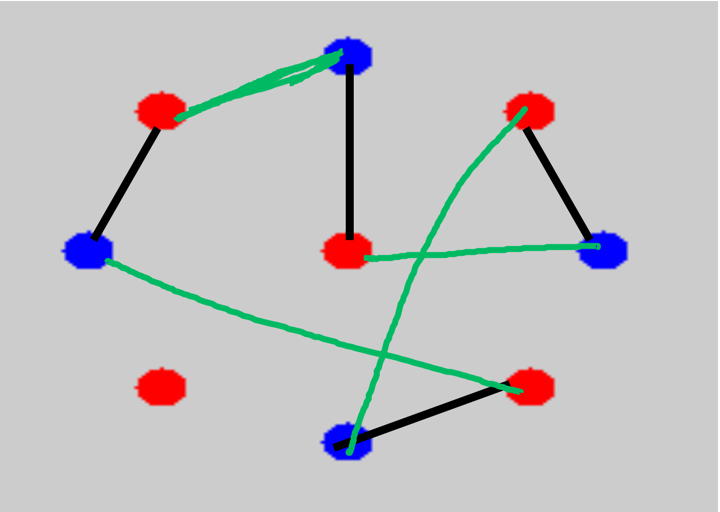
$$P(X_1 = v_1, \ldots, X_4 = v_4) \propto \begin{cases} \exp\left(-\sum_i \mathrm{dist}(i, v_i)\right) & \text{if every } X_i \text{ has different value} \\ 0 & \text{otherwise} \end{cases}$$





Daphne Koller

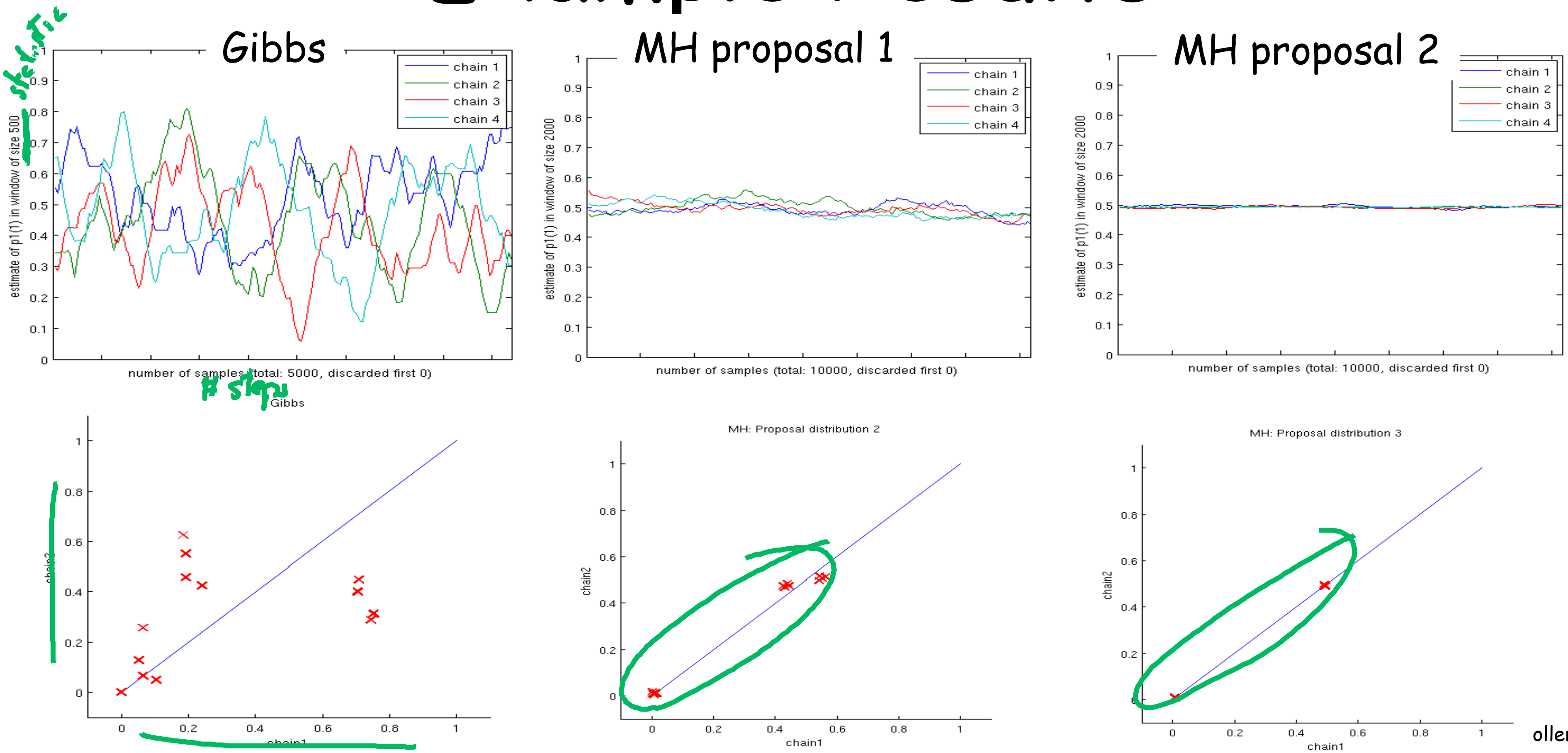
MH for Matching: Augmenting Path



1) randomly pick one variable X_i
2) sample X_i, pretending that all values are available
3) pick the variable whose assignment was taken (conflict), and return to step 2

- When step 2 creates no conflict, modify assignment to flip augmenting path

Example Results



oller

Summary

- MH is a general framework for building Markov chains with a particular stationary distribution
  - Requires a proposal distribution
  - Acceptance computed via detailed balance
- Tremendous flexibility in designing proposal distributions that explore the space quickly
  - But proposal distribution makes a big difference
  - and finding a good one is not always easy