

Theory Working Group

Development and Analysis of
Algorithms for problems traditionally
handled via SMC.

SMC alg \rightarrow MCMC

Outline 1 Oct, 2008

- Perfect Simulation
 - Basic Acceptance/Rejection
- Randomness Recycler
- Examples
 - Hard Core Gas Model
 - Ising Model
- Relation to SST
- Initial Problems?

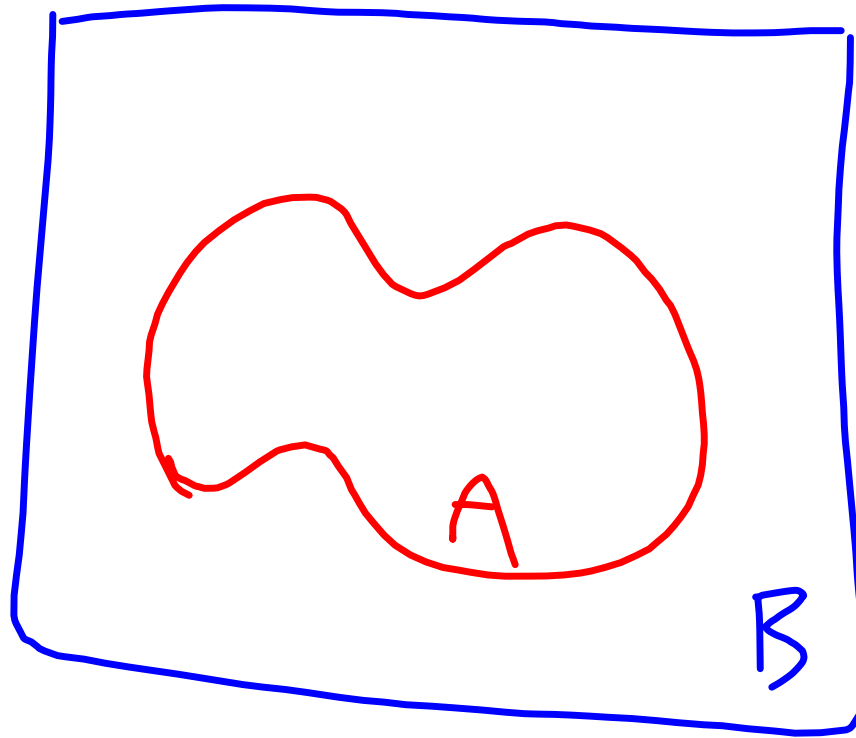
Perfect Simulation

Two salient features

- 1) Draws samples exactly from π .
- 2) Running time is a random variable (usually w/ exponential tails).

Basic Acceptance/Rejection

- Want uniform over A
- Can draw uniform over $B: A \subseteq B$



A/R

1) Repeat

2) Draw $X \leftarrow \text{Unif}(B)$

3) Until $X \in A$.

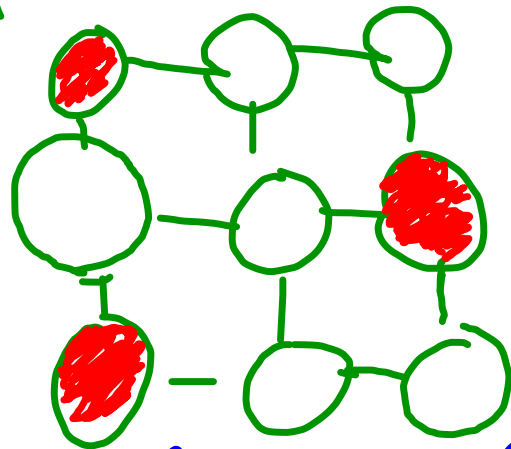
of times through loop $\sim \text{Geo}\left(\frac{\mu(A)}{\mu(B)}\right)$

random time w/
exponential tails

$$E[T] = \frac{\mu(B)}{\mu(A)}$$

Problem: $\frac{\mu(A)}{\mu(B)}$ usually decreases exponentially in size of problem.

Example: Hard Core Gas model



$$w(x) = \lambda^{\#x}$$

$$\pi(x) = \frac{w(x)}{Z_\lambda}$$

No 2 adjacent nodes in set

A/R for ind. sets aka Hard Core Gas..

Graph: $G=(V, E)$

$$Z_{\lambda} \leq (1+\lambda)^{|V|}$$

v in set: $x(v)=1$

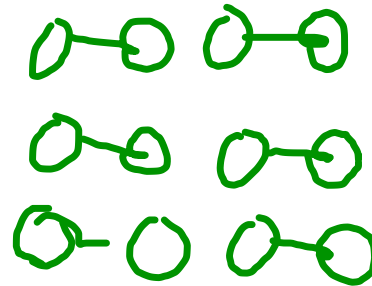
v out of set: $x(v)=0$

$$\Omega = \{0, 1\}^{|V|}$$

AR ALG

- 1) Repeat
- 2) For all $v \in V$
- 3) Draw $x(v)$: $P(x(v)=1) = \frac{\lambda}{1+\lambda}$
- 4) Until $\forall \{i,j\} \in E$ $x(i)x(j)=0$.

Chance of no conflicts in lattice \leq
 $\left(\frac{2\lambda+1}{1+\lambda}\right)^{V/2}$



A solution...

Randomness Recycler

Main Idea: Don't throw away
sample completely when you
reject

RR ALG ingredients

- (a) State space Ex: $\{0,1\}^{|V|} = (\mathcal{X}, \mathcal{F})$
- (b) Dual state space $(\mathcal{X}^*, \mathcal{F}^*)$
- (c) Kernel Δ from $(\mathcal{X}^*, \mathcal{F}^*) \rightarrow (\mathcal{X}, \mathcal{F})$
- (d) π target dist. on $(\mathcal{X}, \mathcal{F})$
- (e) special dual state $x_{\pi}^* \in \mathcal{X}^*$
 $\Delta(x_{\pi}^*, \cdot) = \pi(\cdot)$
- (f) special dual state x_0^* :
 $\Delta(x_0^*, \cdot)$ is easy to sample from

(g) bivariate Markov kernel on
 $X^* \times X$

want to maintain

$$\mathcal{L}(X_t \mid X_0^* = x_0^*, \dots, X_t^* = x_t^*) = \prod (x_t^*, \cdot)$$

RR algorithm

- 1) Start X_0^* at X_0^*
- 2) Draw $X_0 \leftarrow \Gamma(X_{0|1}^*)$
- 3) Repeat
- 4) Take steps in Markov chain
- 5) Until $X_t^* = X_{t+1}^*$
- 6) Output X_t .

Ex: Hard Core Gas Model

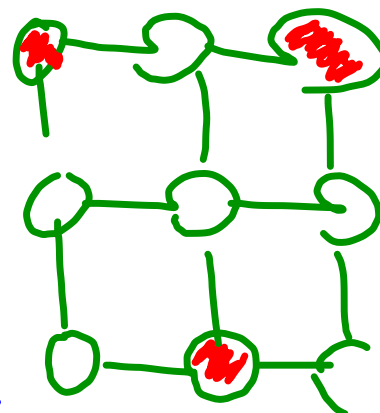
$$\chi^* = 2^E$$

$$\chi_{\text{top}}^* = E, \quad \chi_0^* = \emptyset$$

• 0 0

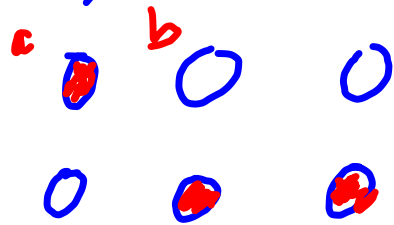
0 • •

0 0 •

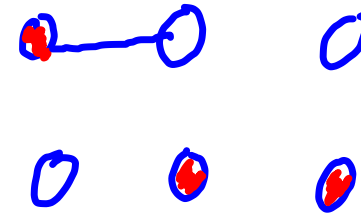


$$\Delta(\chi_t^*, \lambda) = \frac{\lambda^{\#X} \mathbb{1}(\forall e \in X_t^* : x(i) | x(j) = 0)}{Z_{\lambda, X_t^*}}$$

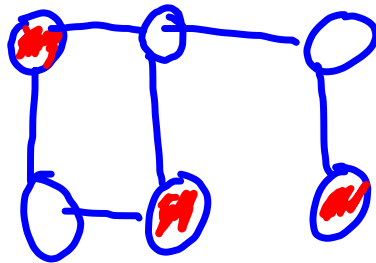
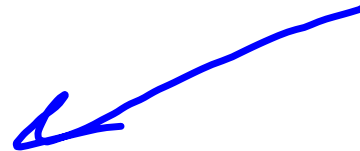
Moving to next x_i^*



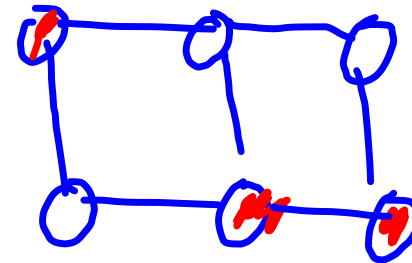
$$X_0^* = \emptyset$$



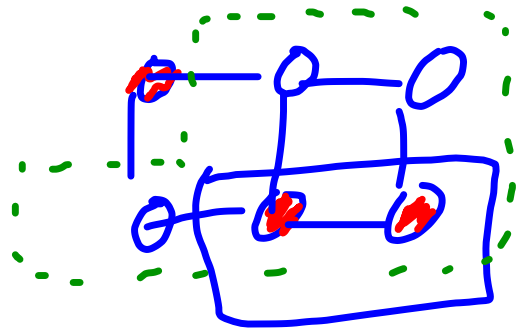
$$X_1^* = \{\{a, b\}\}$$



$$X_2^*$$

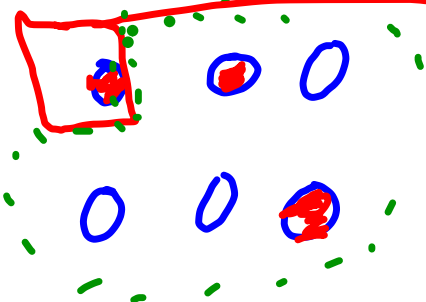


X_2^* is rejected



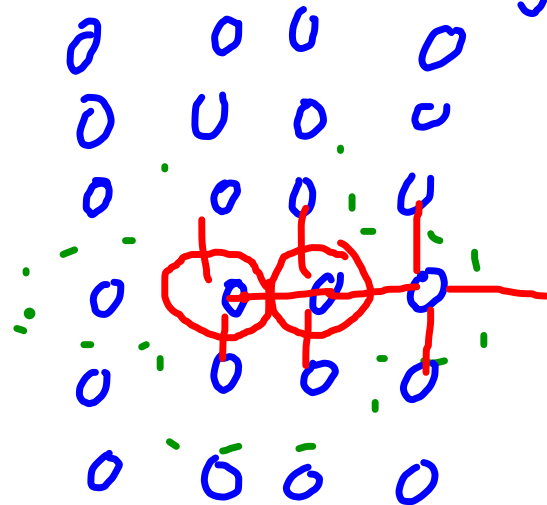
Rejection means $x(i) = x(j) = 1$
for edge $\{i, j\}$.

Recycling comes in here...
Stays some "recycled"



throw away all
edges, redraw
those nodes

Piece of larger graph



max deg Δ

2 things could occur

- Add an edge $+1$
- Delete a bunch of edges
 $2\Delta - 2$

How many edges in X_t^* ?

$$E(X_{t+1}^* | X_t^*) = X_t^* +$$

$$1 \cdot \left(1 - \left(\frac{\lambda}{1+\lambda}\right)^2\right)$$

$$- (2\Delta^2 - 2) \left(\frac{\lambda}{1+\lambda}\right)^2 > X_t^*$$

$$1 - \left(\frac{\lambda}{1+\lambda}\right)^2 > (2\Delta^2 - 2) \left(\frac{\lambda}{1+\lambda}\right)^2$$

$$1 > (2\Delta^2 - 1) \left(\frac{\lambda}{1+\lambda}\right)^2$$

$$\frac{1+\lambda}{\lambda} > \sqrt{2\Delta^2 - 1} \quad 1 + \lambda > \lambda \sqrt{2\Delta^2 - 1}$$

$$1 + \lambda > \lambda \sqrt{2\delta^2 - 1}$$

$$1 > \lambda(\sqrt{2\delta^2 - 1} - 1)$$

$$\boxed{\lambda < \frac{1}{\sqrt{2\delta^2 - 1} - 1}} = O\left(\frac{1}{\delta}\right)$$

$$\boxed{E[\tau]} \leq \frac{|E|}{\left(1 - \left(\frac{\lambda}{1+\lambda}\right)^2 - (2\delta^2 - 1)\left(\frac{\delta}{1+\lambda}\right)^2\right)^2}$$

Another way to create X_t^*

Not to build edge, build nodes

No nodes to begin w /

Add nodes in at each step...

$$\lambda < \frac{4}{3\Delta - 4}$$

Advantages over MCMC

- No mixing time
- $(X_{\#}^*, X)$ is absorbing

Disadvantages

- More difficult to create one of these processes than to create a Gibbs or Metropolis chain.

Hand core qaz model
absorbing $X_t^* = E$
run forward until all edges in
place, stop.