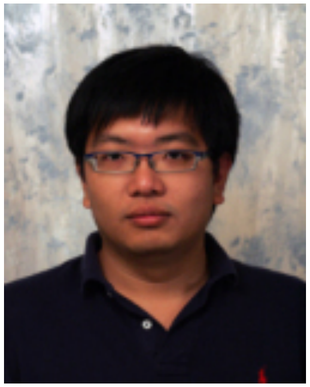
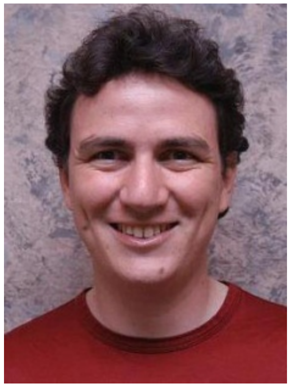
Parallel Gibbs Sampling

From Colored Fields to Thin Junction Trees


Joseph Gonzalez

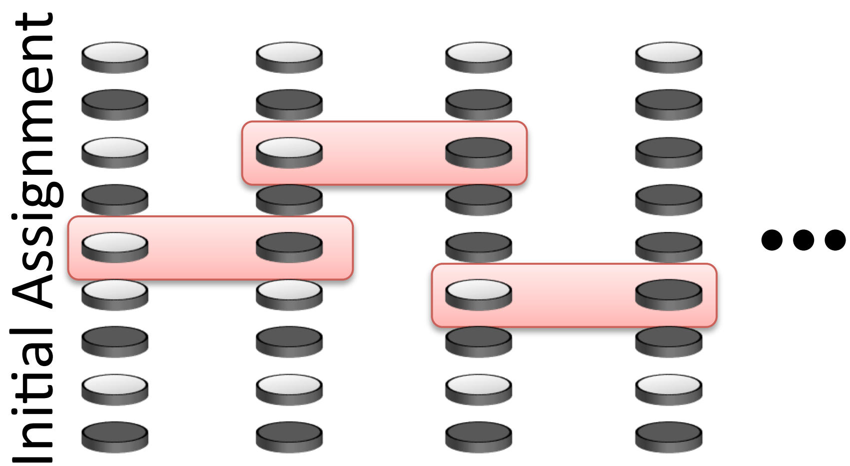

Yucheng Low

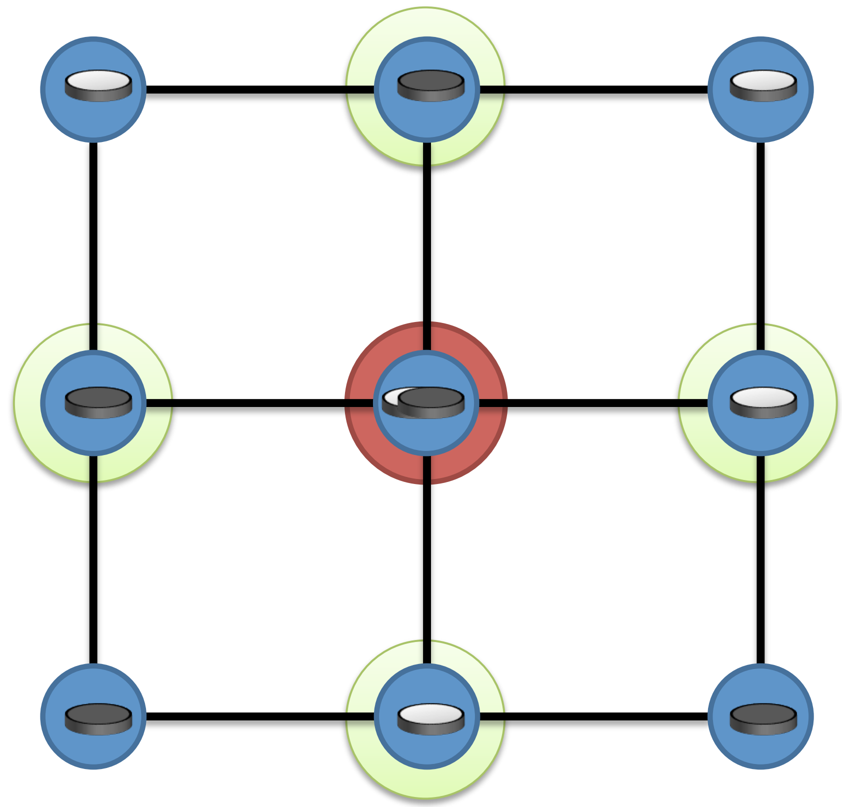

Arthur Gretton


Carlos Guestrin

Gibbs Sampling [Geman & Geman, 1984]

- **Sequentially** for each variable in the model
  - Select variable
  - Construct conditional given adjacent assignments
  - Flip coin and update assignment to variable

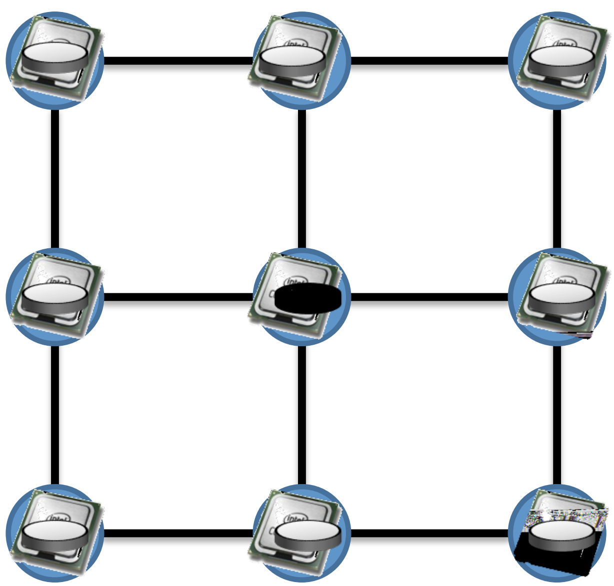




From the original paper on Gibbs Sampling:

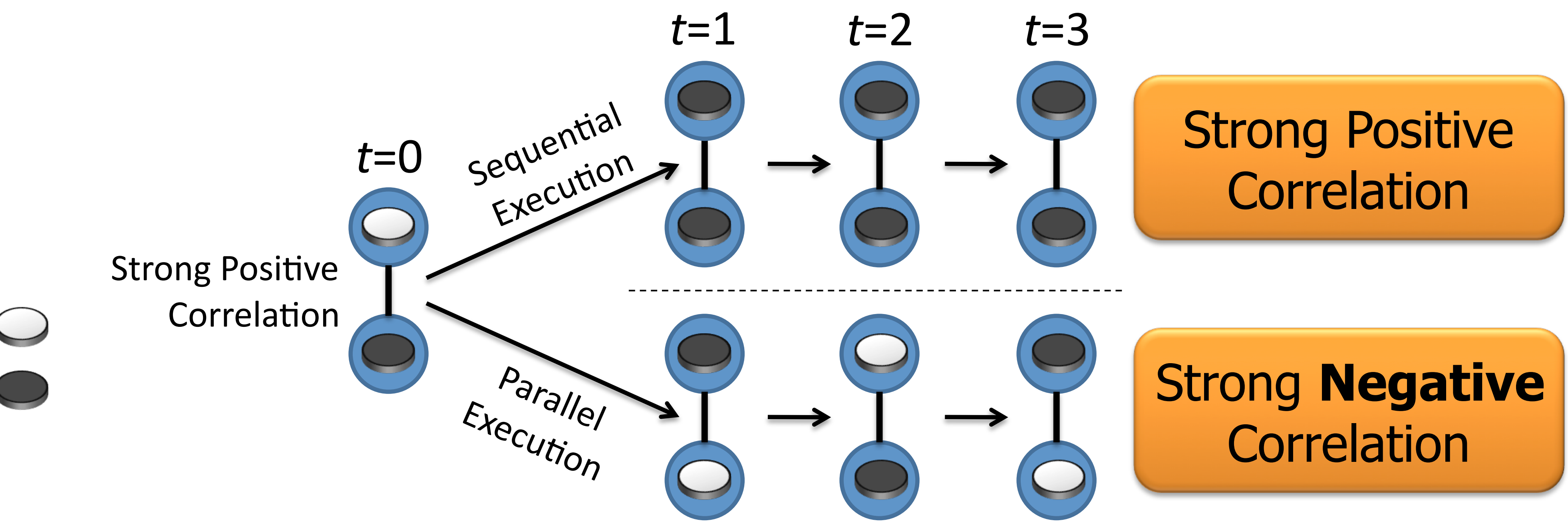
"…the MRF can be divided into collections of [variables] with each collection assigned to an ***independently*** *running* ***asynchronous processor****."*

-- Stuart and Donald Geman, 1984.



Converges to the

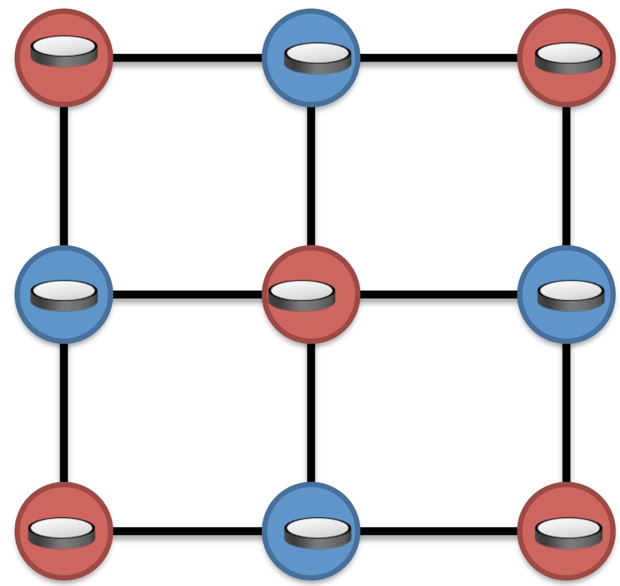
The problem with **Synchronous** Gibbs

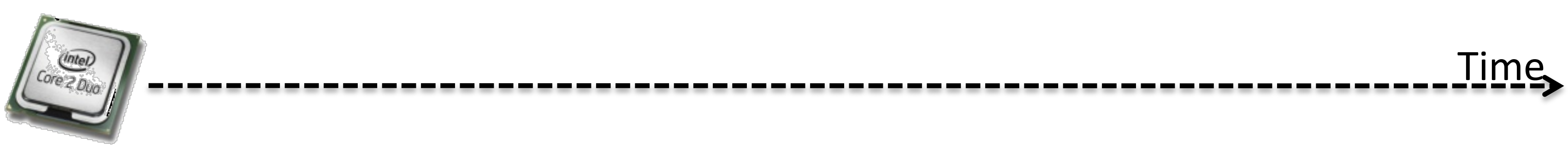


- *Adjacent variables **cannot** be sampled **simultaneously**.*

Chromatic Sampler

- Compute a k-coloring of the graphical model
- Sample all variables with same color in parallel
- Sequential Consistency:





Properties of the **Chromatic** Sampler

- Converges to the correct distribution

- **Quantifiable** acceleration in **mixing**

Time to update
all variables once

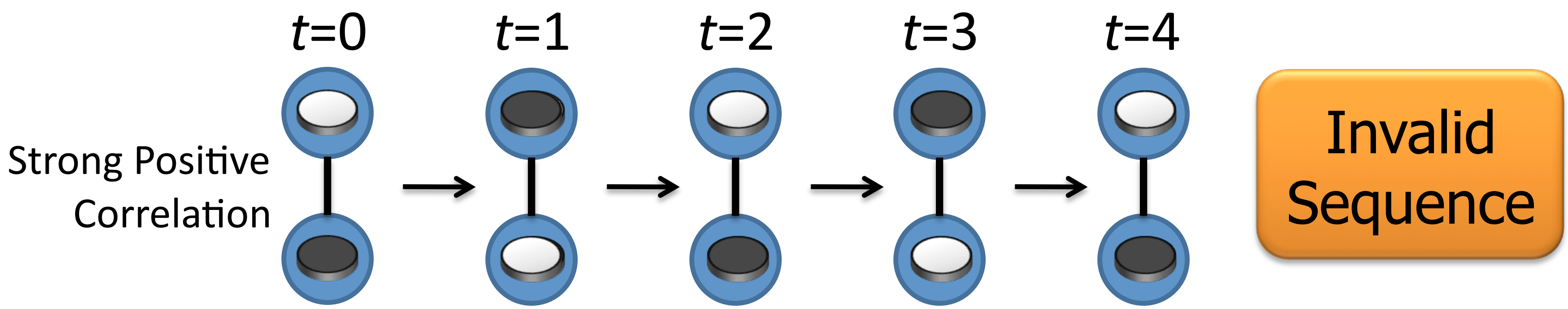
$$O\left(\frac{n}{p} + k\right)$$

\# Variables

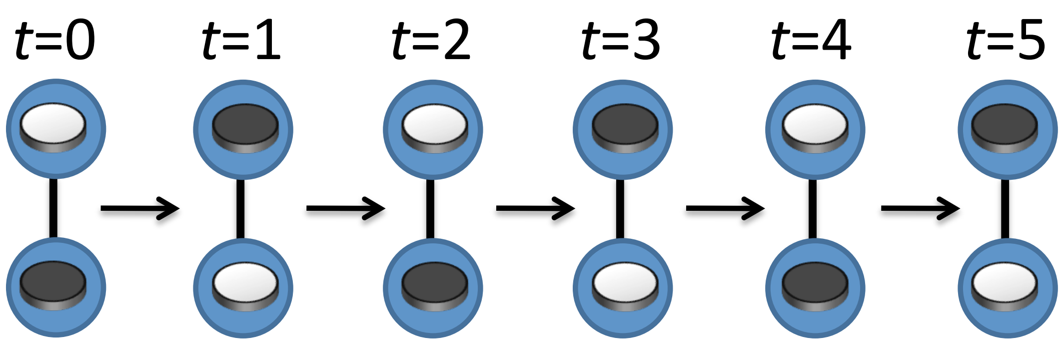
\# Colors

\# Processors

Properties of the *Synchronous* Gibbs Sampler on **2-colorable** models

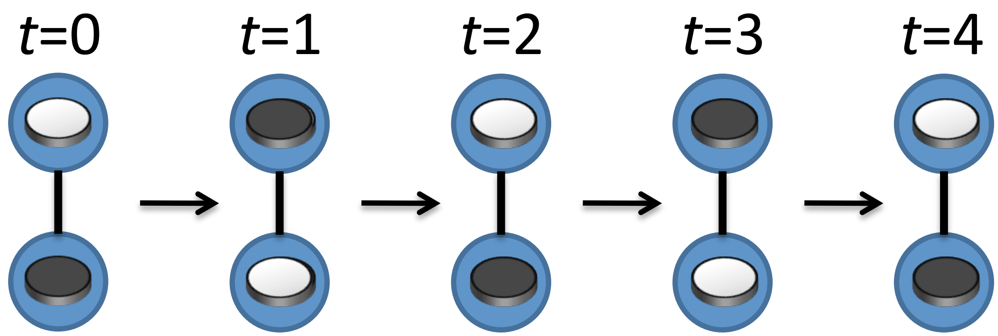


- We can derive two **valid** chains:



Properties of the *Synchronous* Gibbs Sampler on **2-colorable** models

Strong Positive Correlation

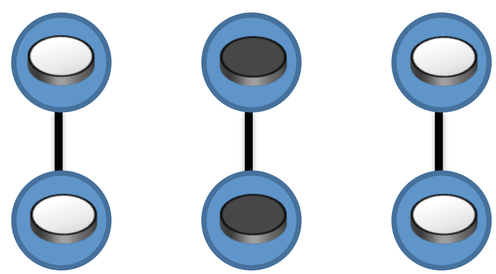


Invalid Sequence

- We can derive two **valid** chains:

Chain 1

Chain 2



Converges to the Correct Distribution

sense learn act

Theoretical Contributions on 2-colorable models

- Stationary distribution of **Synchronous Gibbs**

$$\mathbf{P}_{\text{sync}}\left(X_1, \ldots, X_n\right) = \mathbf{P}\left(X_{\kappa_1}\right)\mathbf{P}\left(X_{\kappa_2}\right)$$

Variables in Color 1

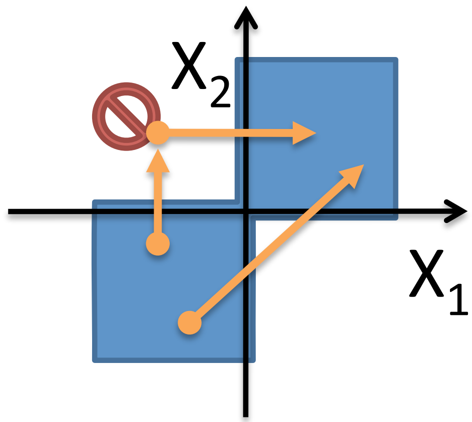
Variables in Color 2

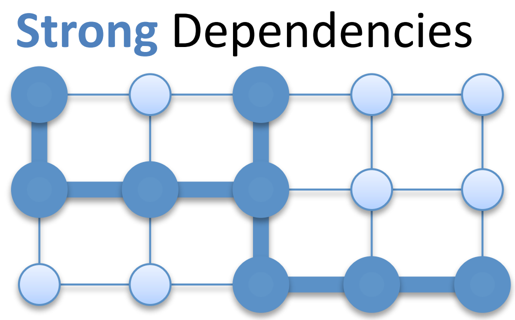
- **Corollary**: Synchronous Gibbs sampler is **correct** for single variable marginals.

$$\mathbf{P}_{\text{sync}}\left(X_i\right) = \mathbf{P}\left(X_i\right)$$

Models With Strong Dependencies

- **Single variable** Gibbs updates tend to mix **slowly:**





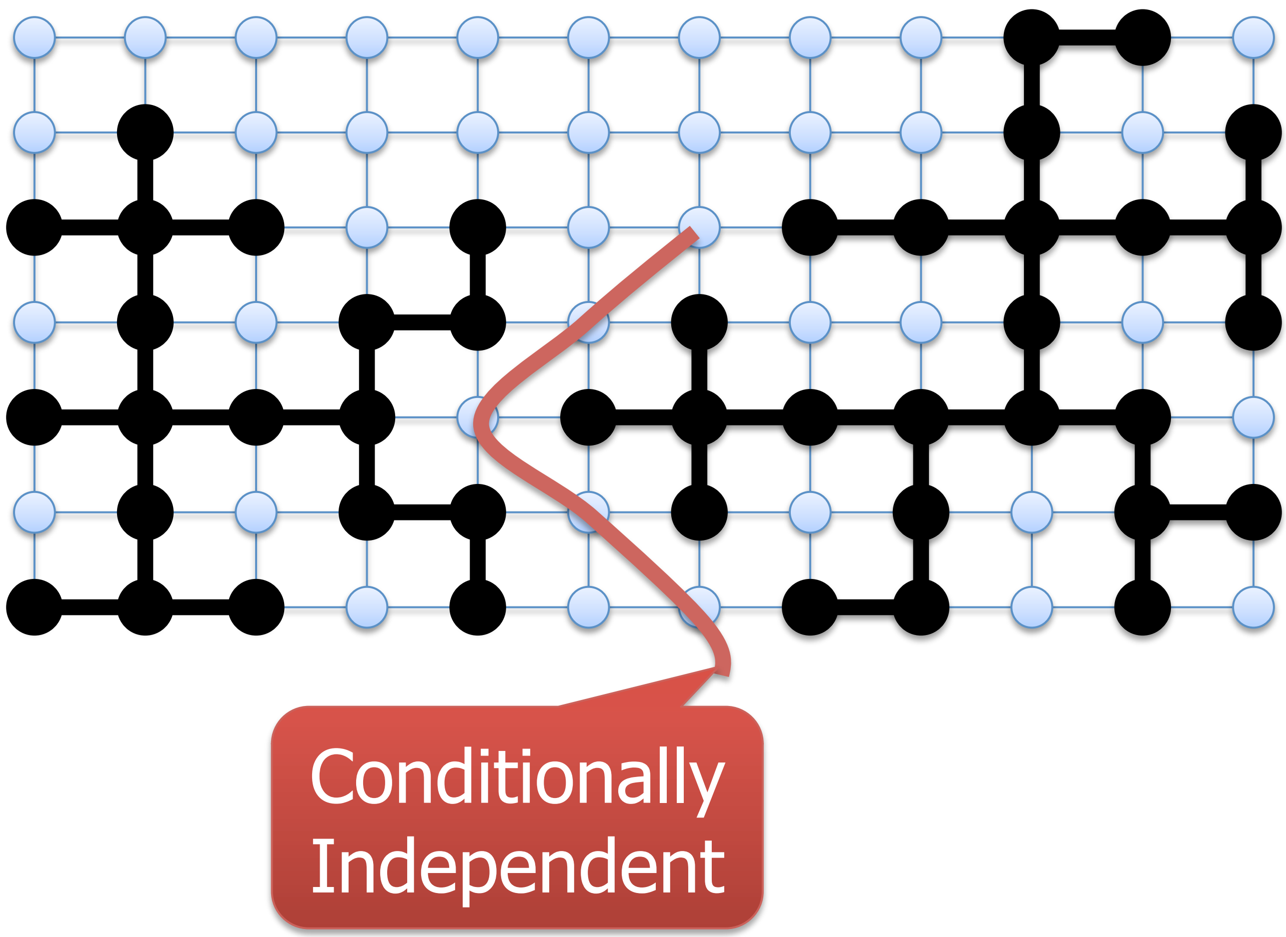
- Ideally we would like to draw joint samples.
  - Blocking

Splash Gibbs Sampler

*An **asynchronous** Gibbs Sampler that **adaptively** addresses **strong dependencies**.*

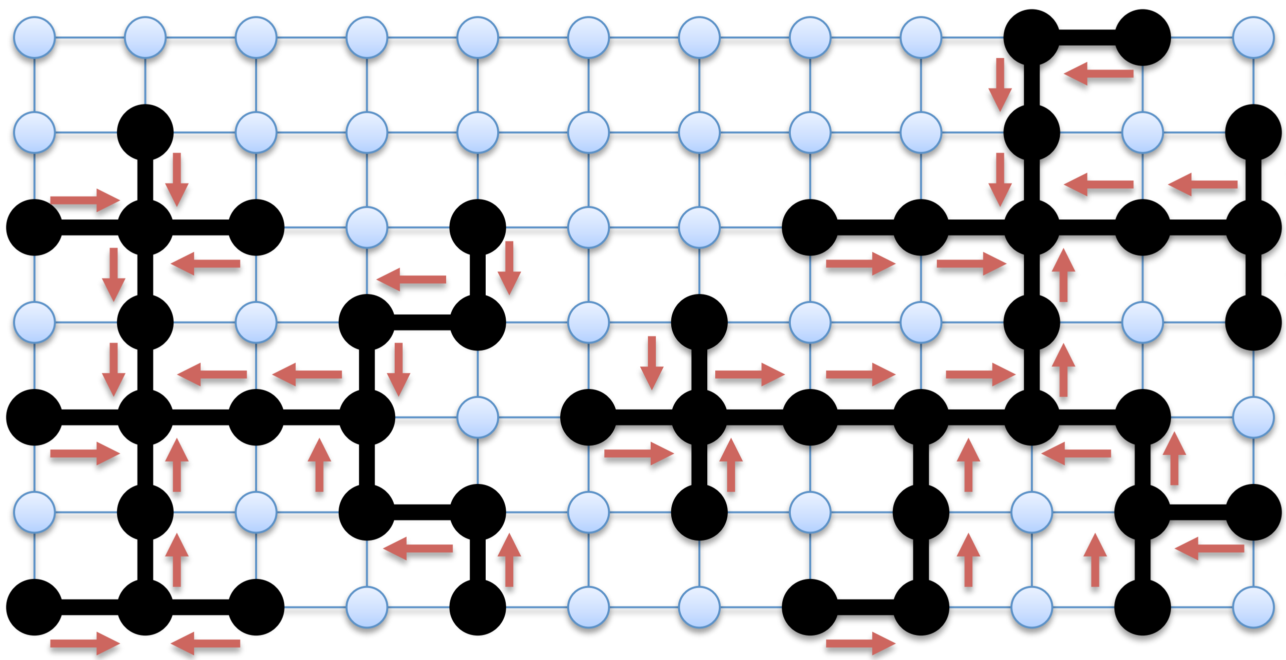
Splash Gibbs Sampler

- **Step 1:** Grow multiple Splashes in parallel:



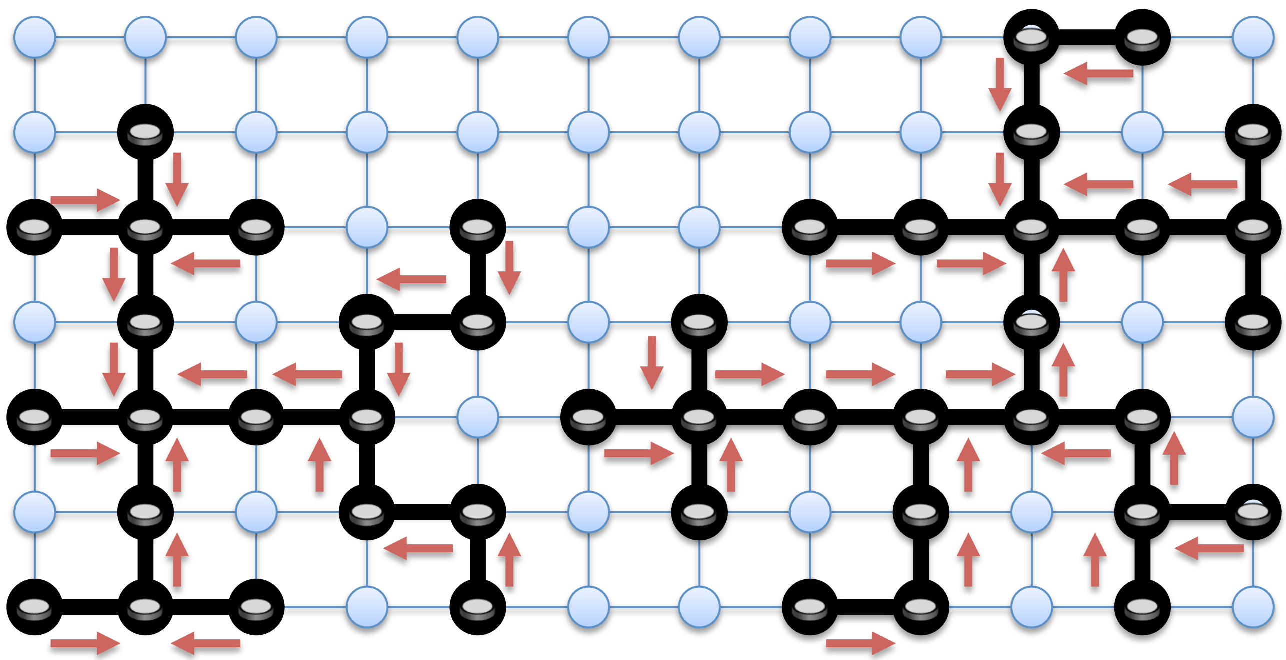
Splash Gibbs Sampler

- **Step 2:** Calibrate the trees in parallel



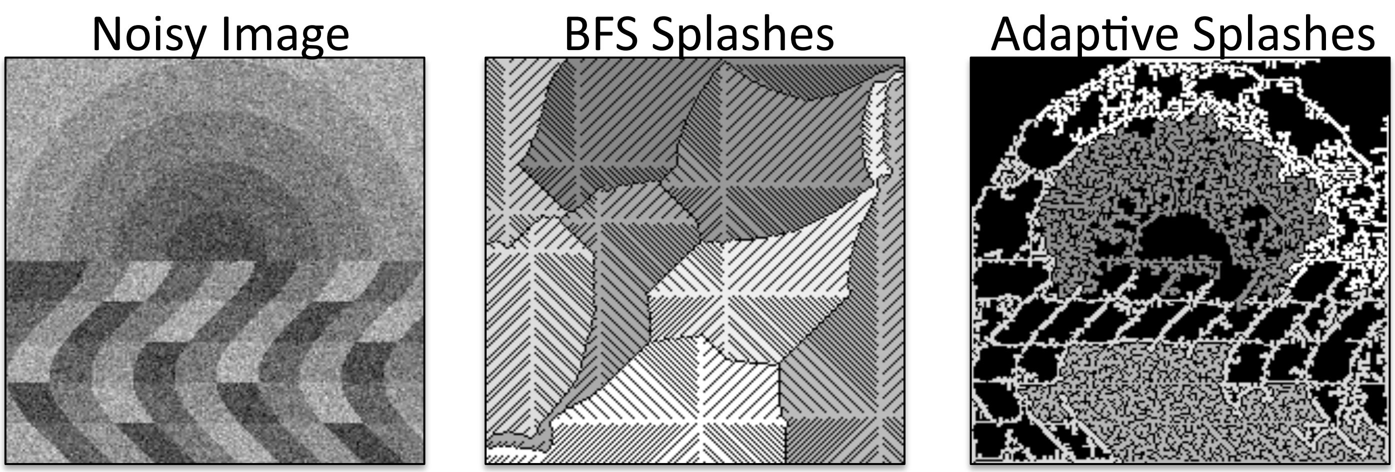
Splash Gibbs Sampler

- **Step 3:** Sample trees in parallel



Adaptively Prioritized Splashes

- Adapt the **shape** of the Splash to span strongly coupled variables:



- Converges to the correct distribution
  - Requires vanishing adaptation

Experimental Results

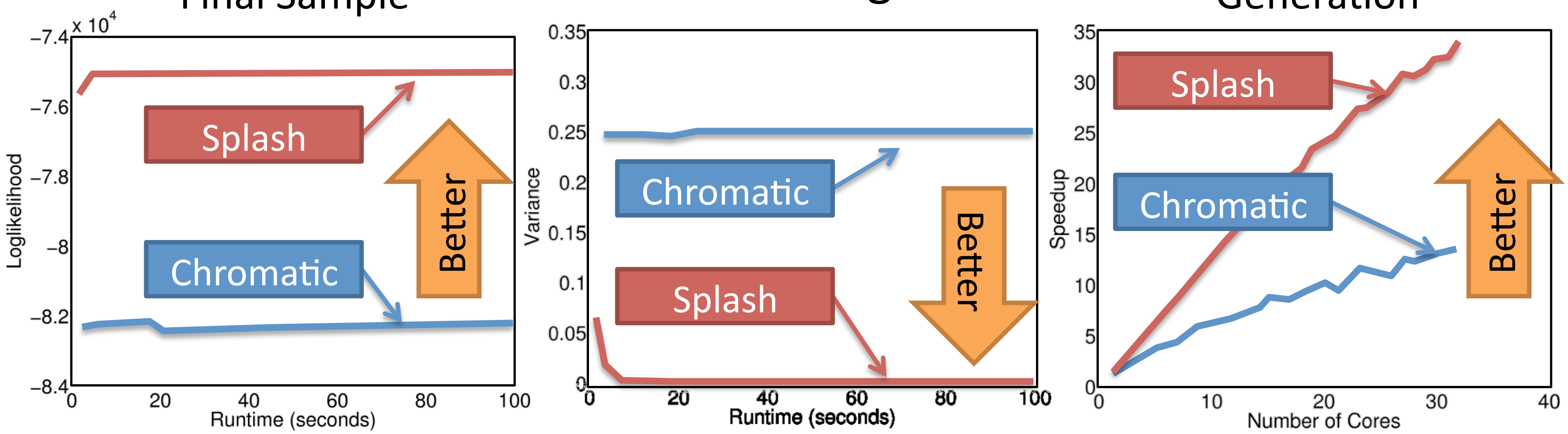
- Markov logic network with strong dependencies

10K Variables 28K Factors

Likelihood Final Sample

"Mixing"

Speedup in Sample Generation



- The *Splash* sampler outperforms the *Chromatic* sampler on models with **strong** dependencies

Conclusions

- **Chromatic Gibbs** sampler for models with *weak dependencies*
  - *Converges to the correct distribution*
  - *Quantifiable improvement in mixing*
- **Theoretical analysis** of the Synchronous Gibbs sampler on *2-colorable models*
  - *Proved marginal convergence on 2-colorable models*
- **Splash Gibbs** sampler for models with *strong dependencies*
  - *Adaptive asynchronous tree construction*
  - *Experimental evaluation demonstrates an improvement in mixing*