Markov Random Fields

- Noncausal model
- Advantages of MRF's
  - Isotropic behavior
  - Only local dependencies
- Disadvantages of MRF's
  - Computing probability is difficult
  - Parameter estimation is difficult
- Key theoretical result: Hammersley-Clifford theorem

Definition of Neighborhood System

- Define

S - set of lattice points

s - a lattice point, $s \in S$

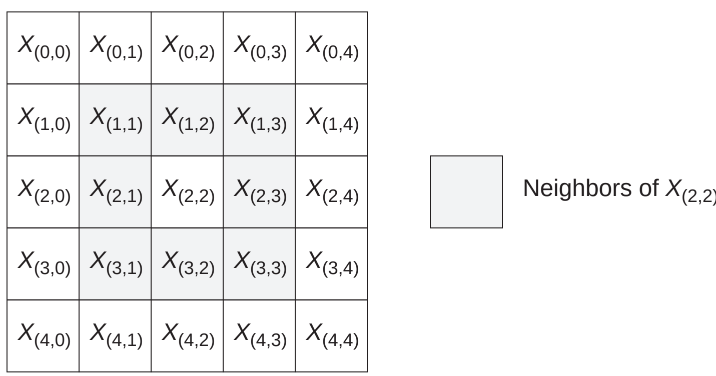
X_s - the value of X at s

$\partial s \subset S$ - the neighboring points of s

- A neighborhood system ∂s must be symmetric

$$r \in \partial s \Rightarrow s \in \partial r \quad \text{also } s \notin \partial s$$

- Example of 8 point neighborhood



Markov Random Field

- Definition: A random object X on the lattice S with neighborhood system ∂s is said to be a Markov random field if for all $s \in S$

$$p(x_s | x_r \text{ for } r \neq s) = p(x_s | x_{\partial s})$$

- Problem: How do we write down the distribution for an MRF?

  Unfortunately

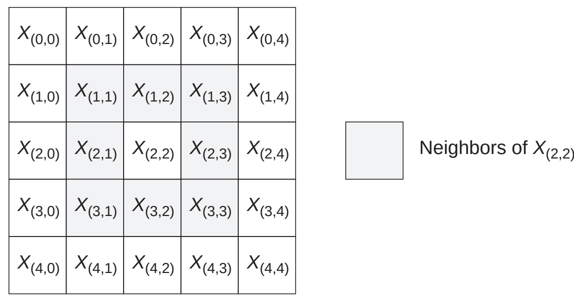
$$p(x) \neq \prod_{s \in S} p(x_s | x_r \text{ for } r \neq s)$$

Definition of Clique

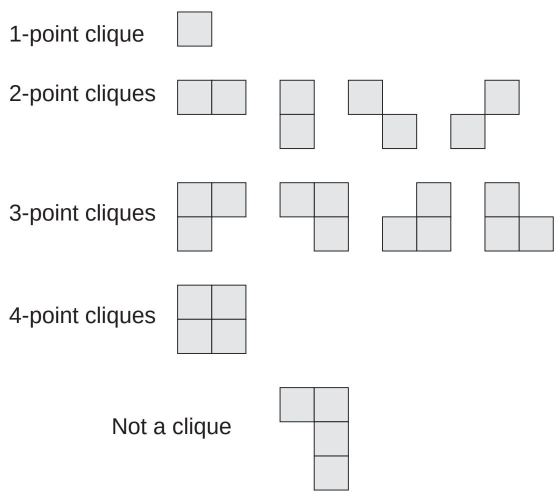
- A clique is a set of points, c, which are all neighbors of each other

$$\forall s, r \in c, \, r \in \partial s$$

- 8 point neighborhood system



- Example of cliques for 8 point neighborhood



Gibbs Distribution

x_c - The value of X at the points in clique c.

$V_c(x_c)$ - A potential function is any function of x_c.

- A (discrete) density is a Gibbs distribution if

$$p(x) = \frac{1}{Z} \exp\left\{ - \sum_{c \in \mathcal{C}} V_c(x_c) \right\}$$

  $\mathcal{C}$ is the set of all cliques

  Z is the normalizing constant for the density.

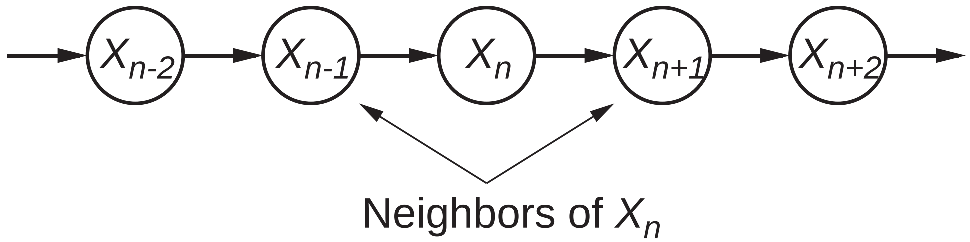
- Z is known as the **partition function**.
- $U(x) = \sum_{c \in \mathcal{C}} V_c(x_c)$ is known as the **energy function**.

Hammersley-Clifford Theorem[1]

$$\left(\begin{array}{c} X \text{ is a Markov random field} \\ \& \\ \forall x,\ P\{X = x\} > 0 \end{array} \right) \Longleftrightarrow \left(\begin{array}{c} P\{X = x\} \text{ has the form} \\ \text{of a Gibbs distribution} \end{array} \right)$$

- Gives you a method for writing the density for a MRF
- Does not give the value of Z, the partition function.
- Positivity, $P\{X = x\} > 0$, is a technical condition which we will generally assume.

Markov Chains are MRF's



- Neighbors of n are $\partial n = \{n-1, n+1\}$
- Cliques have the form $c = \{n-1, n\}$
- Density has the form

$$
\begin{aligned}
p(x) &= p(x_0) \prod_{n=1}^{N} p(x_n | x_{n-1}) \\
&= p(x_0) \exp\left\{ \sum_{n=1}^{N} \log p(x_n | x_{n-1}) \right\}
\end{aligned}
$$

- The potential functions have the form

$$
V(x_n, x_{n-1}) = -\log p(x_n | x_{n-1})
$$

1-D MRF's are Markov Chains

- Let X_n be a 1-D MRF with $\partial n = \{n-1, n+1\}$
- The discrete density has the form of a Gibbs distribution

$$p(x) = p(x_0) \exp\left\{ -\sum_{n=1}^{N} V(x_n, x_{n-1}) \right\}$$

- It may be shown that this is a Markov Chain.
- Transition probabilities may be difficult to compute.

The Ising Model

- First proposed to model 2-D magnetic structures.
- See the work of Peierls for an early treatment[7, 6].
- Kindermann and Snell have a very clear tutorial treatment in [4].
- Lattice geometry
  - S is a rectangular lattice of N pixels.
  - 4-point neighborhood system with cliques $c \in \mathcal{C}$.
  - Assume circular boundary conditions for now.
- Lattice energy
  - Each pixel $X_s \in \{-1, +1\}$ corresponding to north and south poles.
  - Potential of clique $\{r, s\} \in \mathcal{C}$ is $-\frac{J}{2} X_r X_s$.
  - Total energy is

$$u(x) = -\frac{J}{2} \sum_{\{r,s\} \in \mathcal{C}} X_r X_s \ .$$

Physical Basis of Gibbs Distribution

- What is the equilibrium distribution $p_e(x)$?
- Expected energy is
$$\mathcal{E}\{p_e\} = \sum_x p_e(x)\, u(x)$$
- Entropy is
$$\mathcal{H}\{p_e\} = \sum_x -p_e(x)\, \log p_e(x)$$
- First Law of Thermodynamics: Expected energy must be constant.
- Second Law of Thermodynamics: Entropy must be maximized.
$$p_e(x) = \arg \max_{p_e : \mathcal{E}\{p_e\} = \text{const}} \mathcal{H}\{p_e\}$$
- Solution is the Gibbs distribution!
$$p(x) = \frac{1}{z} \exp\left\{ -\frac{1}{kT} u(x) \right\}$$
  - T is tempurature
  - k is Boltzmann's constant

Distribution for Ising Model

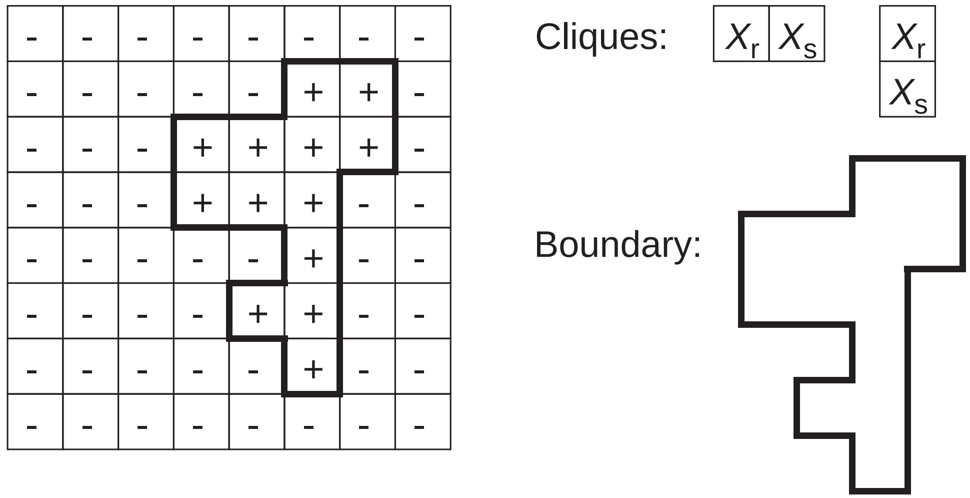
- Equalibrium distribution for Ising model is

$$
\begin{aligned}
p(x) &= \frac{1}{z}\exp\left\{\frac{J}{2kT}\sum_{\{r,s\}\in\mathcal{C}} X_r X_s\right\} \\
&= \frac{1}{z}\exp\left\{\frac{J}{kT}\sum_{\{r,s\}\in\mathcal{C}} \left(\frac{1}{2} - \delta(X_r \neq X_s)\right)\right\} \\
&= \frac{1}{z'}\exp\left\{-\beta\sum_{\{r,s\}\in\mathcal{C}} \delta(X_r \neq X_s)\right\}
\end{aligned}
$$

  where $\beta = \frac{J}{kT}$ is a model parameter and $\delta(X_r \neq X_s)$ is an indicator function for the event $X_r \neq X_s$.

- By the Hammersly-Clifford Theorem, X is a MRF with a 4-point neighborhood.

Interpretation of Ising Model



- Potential functions are given by

$$V(x_r, x_s) = \beta\delta(x_r \neq x_s)$$

- Energy function is given by

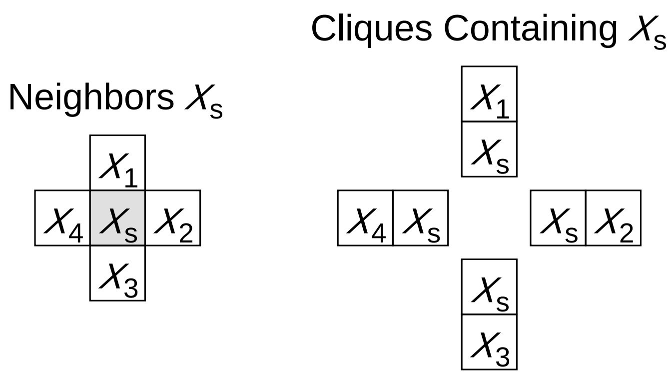
$$\sum_{c\in\mathcal{C}} V_c(x_c) = \beta(\text{Boundary length})$$

- Interpretation of probability density

$$p(x) = \frac{1}{z}\exp\{-\beta(\text{Boundary length})\}$$

- Longer boundaries $\Rightarrow$ less probable

Conditional Probability of a Pixel in Ising Model



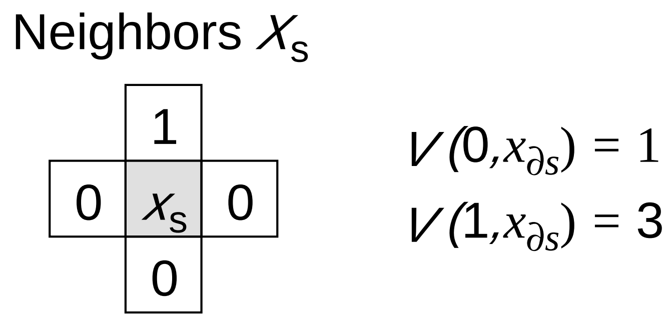
- The probability of a pixel given all other pixels is

$$p(x_s|x_{i\neq s}) = \frac{\frac{1}{Z}\exp\left\{-\sum_{c\in\mathcal{C}} V_c(x_c)\right\}}{\sum_{x_s=0}^{M-1}\frac{1}{Z}\exp\left\{-\sum_{c\in\mathcal{C}} V_c(x_c)\right\}}$$

- Notice: Any term $V_c(x_c)$ which does not include x_s cancels.

$$p(x_s|x_{i\neq s}) = \frac{\exp\left\{-\beta\sum_{i=1}^{4}\delta(x_s\neq x_i)\right\}}{\sum_{x_s=0}^{M-1}\exp\left\{-\beta\sum_{i=1}^{4}\delta(x_s\neq x_i)\right\}}$$

Conditional Probability of a Pixel in Ising Model (Continued)



- Define

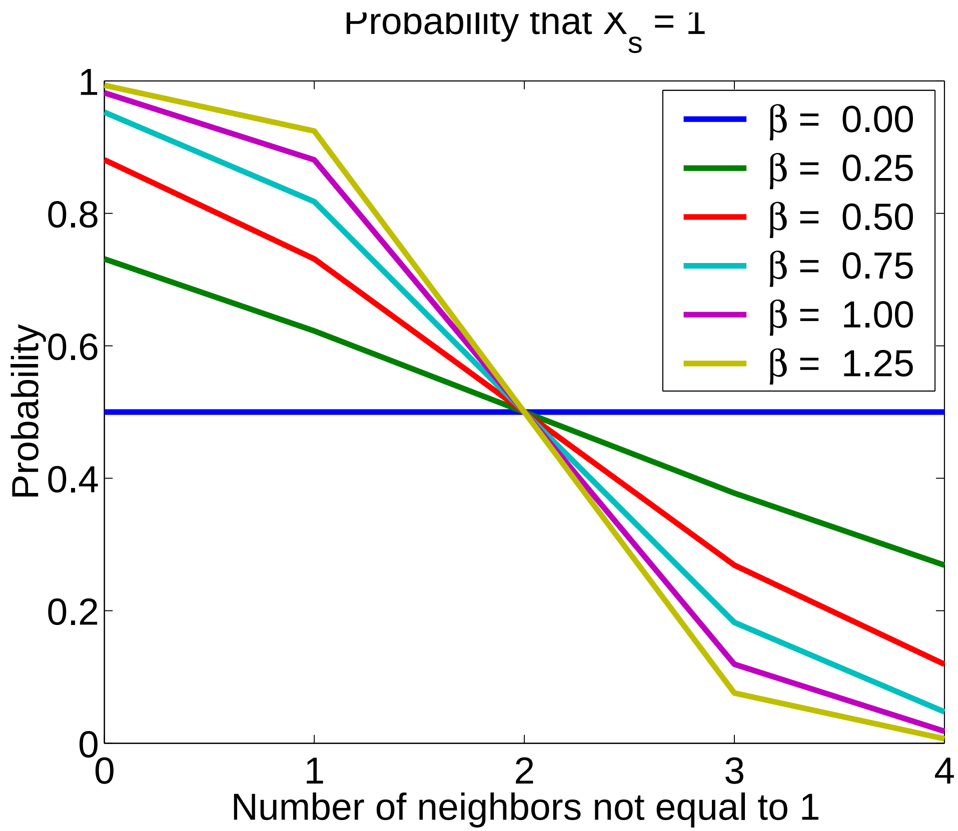
$$v(x_s, \partial x_s) \triangleq \text{\# of horzontal/vertical neighbors} \neq x_s$$

- Then

$$p(x_s|x_{i\neq s}) = \frac{\exp\left\{-\beta v(x_s, \partial x_s)\right\}}{\sum_{x_s'=\{-1,+1\}} \exp\left\{-\beta v(x_s', \partial x_s)\right\}}$$

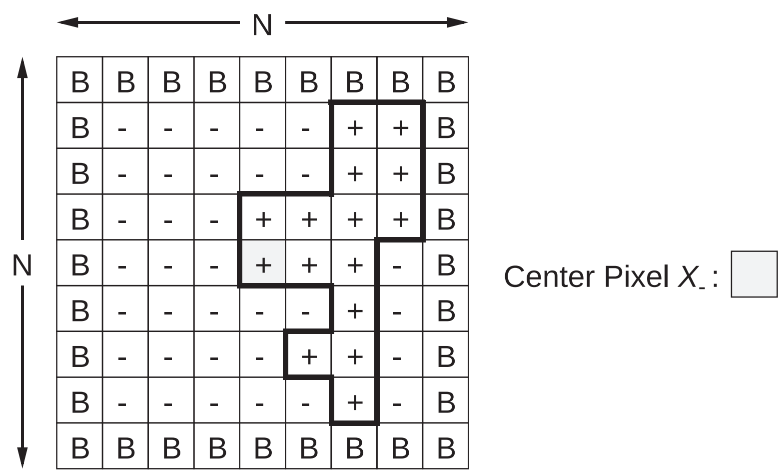
- When $\beta > 0$, X_s is most likely to be the majority neighboring class.

Conditional Distribution Plots



- $P\{X_s = 1 | X_r \text{ for } r \neq s\}$ for different values of β.

Critical Temperature Behavior[7, 6, 4]



- $\frac{1}{\beta}$ is analogous to temperature.
- Peierls showed that for $\beta > \beta_c$

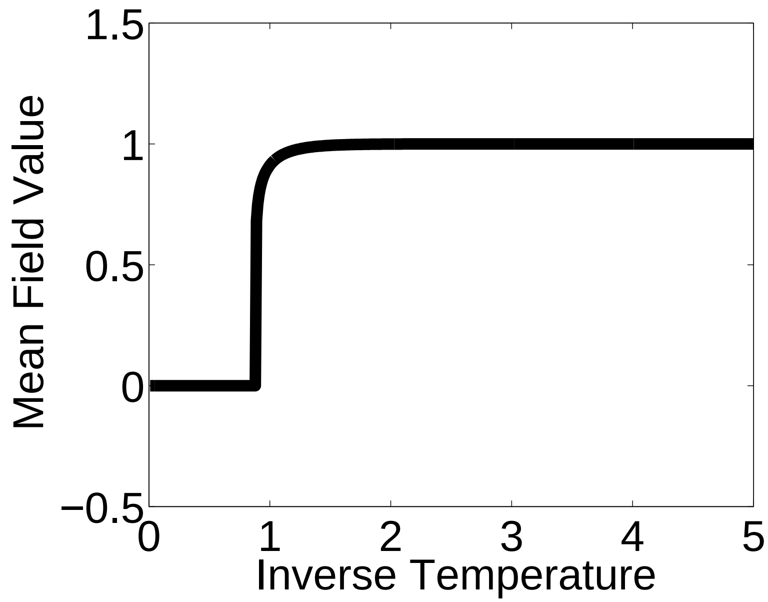
$$\lim_{N\to\infty} P(X_0 = 0 | B = 0) \neq \lim_{N\to\infty} P(X_0 = 0 | B = 1)$$

- The effect of the boundary does not diminish as $N \to \infty$!
- $\beta_c \approx .88$ is known as the critical temperature.
- Very nice proof of critical temperature in [4].

Critical Temperature Analysis[5]

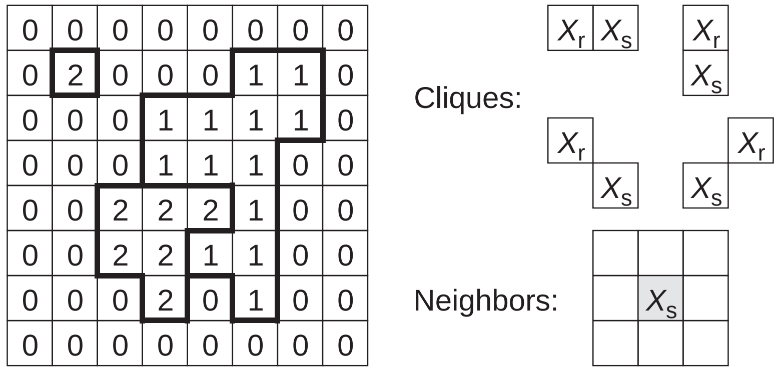
- Amazingly, Onsager was able to compute the following result as $N \rightarrow \infty$.

$$E[X_0|B=1] = \begin{cases} \left(1 - \frac{1}{(\sinh(\beta))^4}\right)^{1/8} & \text{if } \beta > \beta_c \\ 0 & \text{if } \beta < \beta_c \end{cases}$$



- Onsager also computed an analytic expression for $Z(T)$!

M-Level MRF[2]



- Define $\mathcal{C}_1 \triangleq$ (hor./vert. cliques) and $\mathcal{C}_2 \triangleq$ (diag. cliques)
- Then

$$V(x_r, x_s) = \begin{cases} \beta_1 \delta(x_r \neq x_s) & \text{for } \{x_r, x_s\} \in \mathcal{C}_1 \\ \beta_2 \delta(x_r \neq x_s) & \text{for } \{x_r, x_s\} \in \mathcal{C}_2 \end{cases}$$

- Define

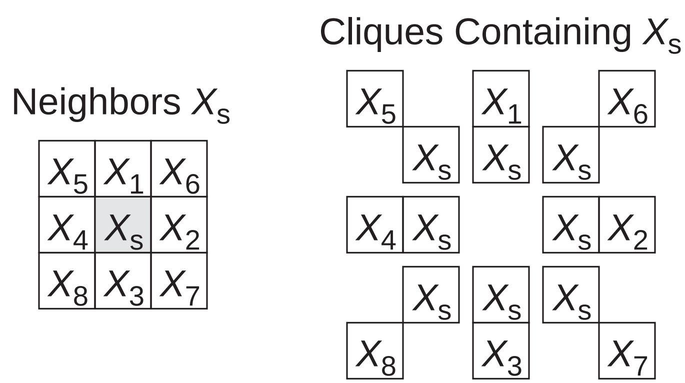
$$t_1(x) \triangleq \sum_{\{s,r\} \in \mathcal{C}_1} \delta(x_r \neq x_s)$$

$$t_2(x) \triangleq \sum_{\{s,r\} \in \mathcal{C}_2} \delta(x_r \neq x_s)$$

- Then the probability is given by

$$p(x) = \frac{1}{Z} \exp\left\{-(\beta_1 t_1(x) + \beta_2 t_2(x))\right\}$$

Conditional Probability of a Pixel



- The probability of a pixel given all other pixels is

$$p(x_s|x_{i\neq s}) = \frac{\frac{1}{Z}\exp\left\{-\Sigma_{c\in\mathcal{C}} V_c(x_c)\right\}}{\Sigma_{x_s=0}^{M-1}\frac{1}{Z}\exp\left\{-\Sigma_{c\in\mathcal{C}} V_c(x_c)\right\}}$$

- Notice: Any term $V_c(x_c)$ which does not include x_s cancels.

$$p(x_s|x_{i\neq s}) = \frac{\exp\left\{-\beta_1\Sigma_{i=1}^{4}\delta(x_s\neq x_i) - \beta_2\Sigma_{i=5}^{8}\delta(x_s\neq x_i)\right\}}{\Sigma_{x_s=0}^{M-1}\exp\left\{-\beta_1\Sigma_{i=1}^{4}\delta(x_s\neq x_i) - \beta_2\Sigma_{i=5}^{8}\delta(x_s\neq x_i)\right\}}$$

Conditional Probability of a Pixel (Continued)

Neighbors X_s

1	1	0
1	X_s	0
0	0	0

$V_1(0,x_{\partial s}) = 2$ $V_2(0,x_{\partial s}) = 1$

$V_1(1,x_{\partial s}) = 2$ $V_2(1,x_{\partial s}) = 3$

- Define

$$\begin{aligned} v_1(x_s, \partial x_s) &\triangleq \text{\# of horz./vert. neighbors} \neq x_s \\ v_2(x_s, \partial x_s) &\triangleq \text{\# of diag. neighbors} \neq x_s \end{aligned}$$

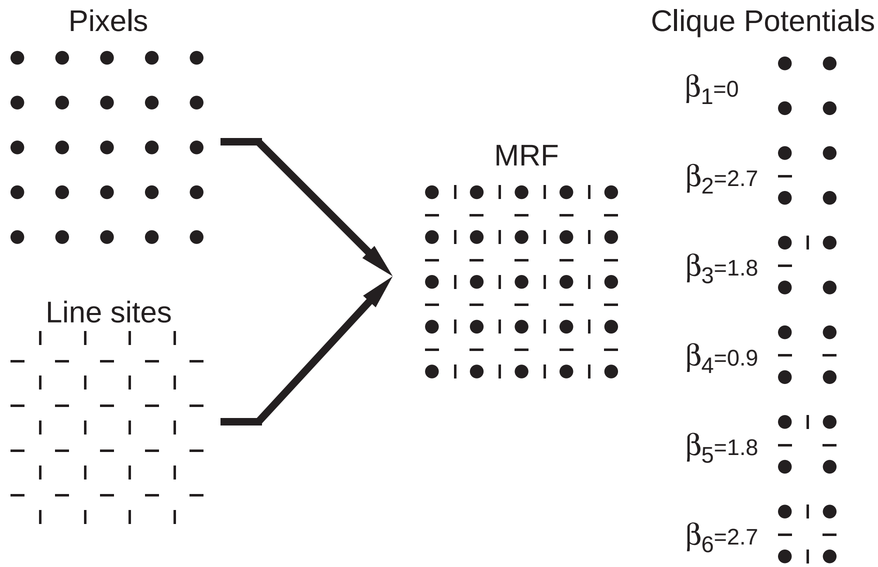
- Then

$$p(x_s | x_{i \neq s}) = \frac{1}{Z'} \exp\left\{ -\beta_1 v_1(x_s, \partial x_s) - \beta_2 v_2(x_s, \partial x_s) \right\}$$

where Z' is an easily computed normalizing constant

- When $\beta_1, \beta_2 > 0$, X_s is most likely to be the majority neighboring class.

Line Process MRF [3]



- Line sites fall between pixels
- The values $\beta_1, \cdots, \beta_2$ determine the potential of line sites
- The potential of pixel values is

$$V(x_s, x_r, l_{r,s}) = \begin{cases} (x_s - x_r)^2 & \text{if } l_{r,s} = 0 \\ 0 & \text{if } l_{r,s} = 1 \end{cases}$$

- The field is
  - Smooth between line sites
  - Discontinuous at line sites

References

[1] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society B*, 36(2):192–236, 1974.

[2] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society B*, 48(3):259–302, 1986.

[3] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-6:721–741, November 1984.

[4] R. Kindermann and J. Snell. *Markov Random Fields and their Applications.* American Mathematical Society, Providence, 1980.

[5] L. Onsager. Crystal statistics i. a two-dimensional model. *Physical Review Letters*, 65:117–149, 1944.

[6] R. E. Peierls. On Ising's model of ferromagnetism. *Proc. Camb. Phil. Soc.*, 32:477–481, 1936.

[7] R. E. Peierls. Statistical theory of adsorption with interaction between the adsorbed atoms. *Proc. Camb. Phil. Soc.*, 32:471–476, 1936.