

Stat 8501 Lecture Notes
Spatial Lattice Processes
Charles J. Geyer
February 26, 2020

1 Introduction

A *spatial lattice process* is a stochastic process with a discrete carrier T , that is, we have random variables $X(t)$ for all $t \in T$ just like in Hoel, Port, and Stone (1986) but since T is discrete there can be no integration or differentiation of the process.

Time series and Markov chains are one-dimensional spatial lattice processes. The only difference is in point of view. We think of time series and Markov chains of evolving from the past to the future. In spatial lattice processes we don't think about order. In a one-dimensional lattice process we treat past and future (or left and right) symmetrically.

When T is two-dimensional or higher, we can't think about order, since two-dimensional and higher space is not ordered (at least there is no order compatible with addition like there is for one-dimensional space).

In two dimensions, there are three regular lattices (that are unchanged by certain translations and rotations), called square, triangular, and hexagonal, but almost all attention focuses on square lattices, because of our familiarity with rectangular coordinates, computer screens having square lattices of pixels, and so forth. A square lattice in d dimensions can be considered to be just \mathbb{Z}^d .

The first spatial lattice process to be intensively studied was the Ising model, and it and the closely related Potts model will be the only ones we will discuss. In physics, it is a model for ferromagnetism and, more generally, for phase transitions. In statistics, it has been used in Bayesian image reconstruction. Although real magnetic materials are three-dimensional, the mathematics of the Ising and Potts models are only well understood in one or two dimensions and only interesting in two dimensions. Thus we restrict our attention to two dimensions. Thus the only lattice we will consider will be \mathbb{Z}^2 .

2 Finite-Dimensional Distributions

If we study processes on infinite lattices, we have to understand how to specify their distributions. As with everything else we have done, we rely on the fact that infinite-dimensional distributions are determined by all of their finite-dimensional marginals.

We will be interested in spatial lattice process that, like the Ising model and Potts model, only have finite state space for the variables X_t (for the Ising model all X_t take the same two values, and for the Potts model all X_t take values in the same small finite set, and the Ising model is the special case where that set has two elements).

We can specify the infinite-dimensional distribution by an unnormalized density function h that is nonnegative. In this case we do not require that h be summable, because over an infinite lattice it won't be. All that is required is that the finite-dimensional marginals exist, and this is guaranteed by having finite state space for the X_t .

3 Hammersley-Clifford Theorem

As we did with spatial point processes, we also define a neighbor relation on the lattice \mathbb{Z}^2 . We write $i \sim j$ if nodes are neighbors. We only discuss the *nearest neighbor* relation, which is the one used for Ising and Potts models. $i \sim j$ if and only if $|i_1 - j_1| + |i_2 - j_2| \leq 1$.

It is helpful for this section to have the notation X_A to denote the vector $(X_i : i \in A)$. The *spatial Markov property* for lattice processes is the following. If A and B are subregions of the lattice such that no node of A is a neighbor of any node in B , then X_A and X_B are conditionally independent given $X_{(A \cup B)^c}$.

As in the case of spatial point process, there is a Hammersley-Clifford theorem for spatial lattice processes. For $A \subset \mathbb{Z}^2$, let $\text{clq}(A)$ denote the set of subsets of A that are cliques (every pair of elements are neighbors). For the nearest neighbor relation on a square lattice, there are only cliques of size one and of size two (single nodes and pairs of neighboring nodes).

Then like what we saw in spatial point processes, the Hammersley-Clifford theorem says the unnormalized density has the form

$$h(x) = \prod_{y \in \text{clq}(x)} \varphi(y)$$

(the proof is similar to the one we gave for spatial point processes, and we won't do this case). If we further assume the model is invariant under

translations and 90° rotations, we obtain that $\varphi(y)$ is the same function for all cliques of size one and the the same function for all cliques of size two, that is, letting E denote the set of all neighbor pairs without double counting (that is, if we have (s, t) we do not also have (t, s)) and also without considering points their own neighbors, we have

$$h(x) = \left(\prod_{t \in \mathbb{Z}^2} \varphi(x_t) \right) \left(\prod_{s, t \in E} \varphi(x_s, x_t) \right) \quad (1)$$

4 Potts Models

To have concrete language, let us call the possible values of X_t “colors” as if we were talking about pixels in an image. We are interested only in state spaces with only a few colors. Let C denote the state space of X_t , the set of colors.

The Potts model is the special case of a model having the spatial Markov property and interactions between variables that depend only on whether the variables have the same color or not, that is,

$$\log \varphi(x, y) = \beta I(x = y)$$

where the I notation does not accord with our previous usage for indicator functions or identity kernels, but has the widely used meaning (one if the argument is true and zero if the argument is false). Here β is an unknown parameter (that can be varied to get different probability distributions). We also write

$$\log \varphi(x) = \alpha(x).$$

Since α is a function whose arguments take C possible values, we can also think of it as a parameter vector whose length is $|C|$, the cardinality of C . Because h is unnormalized, it can be multiplied by an arbitrary constant, which is the same as adding an arbitrary constant to α . Thus one of the values of α considered as a function and one of the components of α considered as a vector is arbitrary. This is similar to the multinomial distribution: if there are k categories then there are only $k - 1$ independent parameters. In fact, if $\beta = 0$, then the Potts models says each X_t has the same multinomial distribution.

5 Simulation and Boundary Conditions

Very little is known exactly about the probability distribution of the Potts model. What is known exactly will be described presently, but in this section we discuss simulation, in particular MCMC.

It is impossible to simulate an infinite number of variables in a finite amount of time, so here we consider simulating finite-dimensional distributions related to the Potts model. These cannot be marginals of the infinite-dimensional distribution, because we do not know how to calculate these marginals (we do not know how to sum out the rest of the variables, infinitely many of them). Instead we do one of the following.

5.1 Free Boundary Conditions

Make a square, finite sublattice. Let $\mathbb{D} = \{1, \dots, d\}$ and we consider the stochastic process on \mathbb{D}^2 that has the same unnormalized density as the Potts model except that for points on the boundary (having one or the other coordinate equal to either 1 or d) nodes have fewer than four neighbors (those on sides have three neighbors, those at corners have two neighbors). We adjust the neighbor set accordingly

$$E = \{(s, t) \in \mathbb{D}^2 \times \mathbb{D}^2 : s_1 + 1 = t_1 \text{ and } s_2 = t_2 \text{ or } s_1 = t_1 \text{ and } s_2 + 1 = t_2\}. \quad (2)$$

Define

$$t_c(x) = \sum_{t \in \mathbb{D}^2} I(x_t = c)$$

and

$$t_*(x) = \sum_{(s,t) \in E} I(x_s = x_t)$$

Then the Potts model on the lattice \mathbb{D}^2 with free boundary conditions is the exponential family having canonical statistics t_c , $c \in C$ and t_* , unnormalized density function defined by

$$\log h(x) = \beta t_*(x) + \sum_{c \in C} \alpha_c t_c(x)$$

(now we are thinking of α as a vector rather than a function). As we said in the preceding section, this model is overparameterized. We can drop one term from the sum by setting one component of α to zero.

5.2 Toroidal Boundary Conditions

Having a boundary where we treat variables differently is inelegant. The elegant solution is to wrap. Let \mathbb{T} denote the same set as \mathbb{D} but with the neighbor relation defined differently. We glue together opposite edges to make a torus (the math term for donut). Formally,

$$E = \{ (s, t) \in \mathbb{T}^2 \times \mathbb{T}^2 : s_1 \oplus 1 = t_1 \text{ and } s_2 = t_2 \text{ or } s_1 = t_1 \text{ and } s_2 \oplus 1 = t_2 \}. \quad (3)$$

where \oplus denotes addition modulo d , which is the same as ordinary addition except that $d + 1 = 1$. Then everything else is as defined in the preceding section.

5.3 Conditioning on the Boundary

An alternative way to deal with the boundary is to fix variables on the boundary. Because of the spatial Markov property, a border of width one serves to separate a square region from the rest of the lattice. Let \mathbb{D} be as in the section on free boundary conditions. Now we add $\mathbb{G} = \{0, \dots, d+1\}$. So \mathbb{D}^2 is a subset of \mathbb{G}^2 and $\mathbb{G}^2 \setminus \mathbb{D}^2$ is the boundary of \mathbb{D}^2 (the set of all neighbors of points of \mathbb{D}^2 that are not themselves in \mathbb{D}^2).

We consider the conditional distribution of the variables at nodes in \mathbb{D}^2 conditional on the variables at nodes in $\mathbb{G}^2 \setminus \mathbb{D}^2$. Now we need the neighbor set

$$E = \{ (s, t) \in \mathbb{G}^2 \times \mathbb{G}^2 : s_1 + 1 = t_1 \text{ and } s_2 = t_2 \text{ or } s_1 = t_1 \text{ and } s_2 + 1 = t_2 \}. \quad (4)$$

Then again everything is the same as before. The distribution, of course, now depends not only on α and β but also on the values of the conditioning variables, the x_t for $t \in \mathbb{G}^2 \setminus \mathbb{D}^2$.

6 Infinite-Dimensional Distributions

We attempt to recover the infinite-dimensional distribution of the Potts model by taking limits of finite-dimensional distributions that we know how to simulate. Whether we choose, free, toroidal, or conditional boundary conditions, we do not have exactly the same situation as in the infinite-dimensional model. But we can hope that the influence of the boundary goes to zero as we get farther and farther away from the boundary (this is not quite true, more on this later).

So consider a fixed square region with node set A that is a subset of a much bigger square region with node set B_n whose boundary is equally

far away from A in all four coordinate directions, and we consider that as $n \rightarrow \infty$ we have a sequence of sets B_n whose union is \mathbb{Z}^2 . As before, let X_{B_n} denote the stochastic process which is the Potts model on B_n with the kind of boundary conditions we have adopted (which in the case of conditioning on the boundary includes a specification of the values on the boundary). Now let $X_{n,A}$ denote the marginal of X_{B_n} that is the marginal distribution of X_A for the model X_{B_n} .

We study the limiting distribution of $X_{n,A}$ as $n \rightarrow \infty$ for all subregions A . Denote the random vector having the limiting distribution as $X_{\infty,A}$, if the limiting distribution exists.

It follows from the fact that the marginals make sense for the finite-dimensional distributions that they will have the consistency property required to define an infinite dimensional distribution, that is, if $A \subset B$, then the distribution of $X_{\infty,A}$ will be the corresponding marginal of $X_{\infty,B}$. So the only nontrivial question is whether these limits exist.

It turns out that for sufficiently small values of β the limit exists and is unique, no matter what boundary conditions are used. But for sufficiently large values of β , the limit depends on the boundary conditions.

When $\alpha = 0$, the behavior changes at the *critical value*

$$\beta_{\text{crit}} = \log\left(1 + \sqrt{|C|}\right)$$

(Potts, 1952).

For $\beta < \beta_{\text{crit}}$ there is a unique limit distribution that does not depend on the boundary conditions. By symmetry, the marginal distribution of one X_t (when $\alpha = 0$) is the symmetric multinomial distribution (all colors equally likely) when free or toroidal boundary conditions are used, hence the limit distribution must have the same property. This is also true of the unique limit distribution.

For $\beta > \beta_{\text{crit}}$ there is a unique limit when free or toroidal boundary conditions are used. The marginal distribution of one X_t (still when $\alpha = 0$) must still be the symmetric multinomial distribution. But when one conditions on the boundary, the limit may depend on the values on the boundary. In particular, if one conditions on the entire boundary having the same color, then the limit depends on the color. The marginal distribution of one X_t is no longer symmetric but favors the color on the boundary. For $\beta \gg \beta_{\text{crit}}$ almost all nodes have this color; there is only a speckle of nodes having other colors.

For $\beta = \beta_{\text{crit}}$ the behavior is fractal, there are blobs of mostly one color of all sizes (in the infinite lattice) but no one color prevails. We cannot

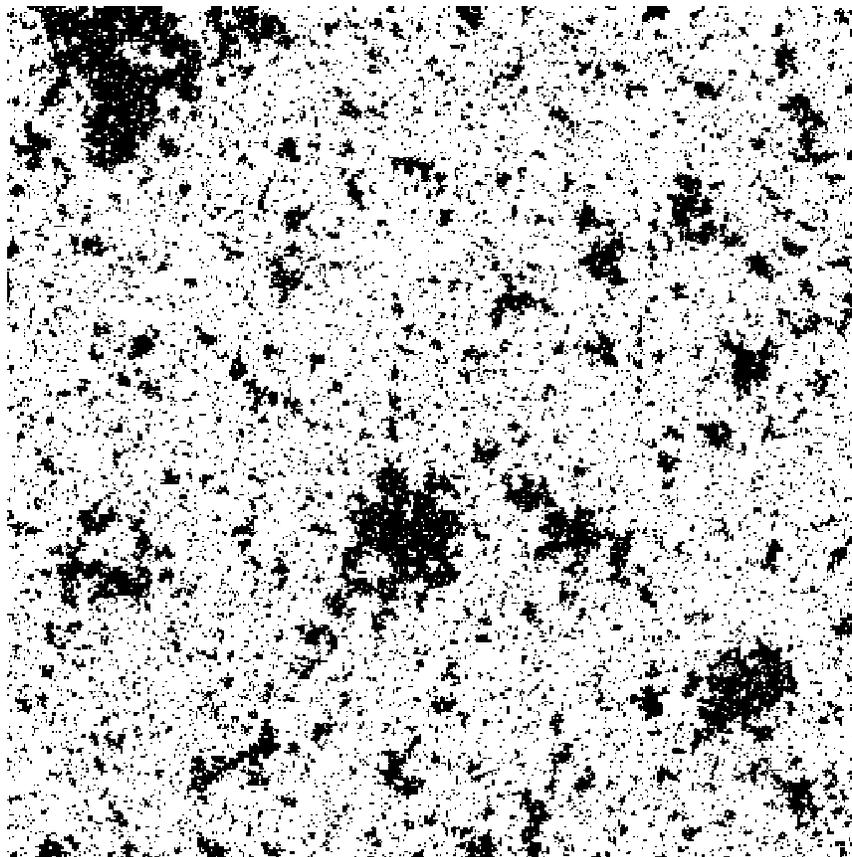


Figure 1: Ising model, 512×512 lattice, toroidal boundary conditions, at the critical value ($\alpha = 0$, $\beta = \beta_{\text{crit}}$).

simulate an infinite lattice, but we can simulate a large lattice and see what that looks like. Figure 1 shows one realization of an Ising model (two colors) at the critical value. This figure and all figures in this handout were made using the R function `potts` in the contributed package `potts` (Geyer and Johnson, 2014), which is available from CRAN.

As the dependence is lessened (β is lowered) the blobs get smaller and break up. The closer β gets to zero, the smaller the blobs get until at $\beta = 0$, the X_t are independent random variables, so the picture is almost uniformly gray (although made up of black and white pixels). Between the critical value of β and zero, the dependence (blobbiness) decreases smoothly

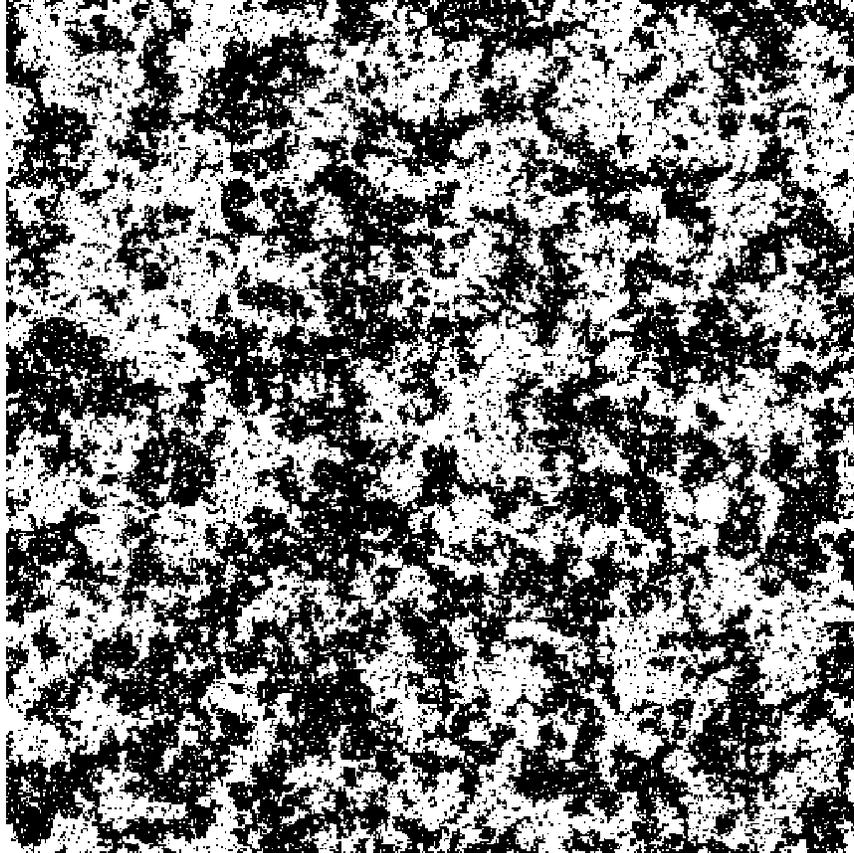


Figure 2: Ising model, 512×512 lattice, toroidal boundary conditions, somewhat below the critical value ($\alpha = 0$, $\beta = 0.95 \cdot \beta_{\text{crit}}$).

as a function of β . Figure 2 shows one realization of an Ising model (two colors) just a little below the critical value. Although the spatial dependence (blobbiness) is still clear, there are no longer blobs that stretch across the whole lattice.

As the dependence is increased from the critical value to infinity there is one infinite blob of one color with smaller blobs of other colors inside it. The closer β gets to infinity, the smaller all the blobs except for the infinite blob get. At $\beta = \infty$, there is only one color in the whole image. If we have used the symmetric limit, letting finite lattices with free or toroidal boundary conditions go to infinity, then all the colors are equally likely. If we have

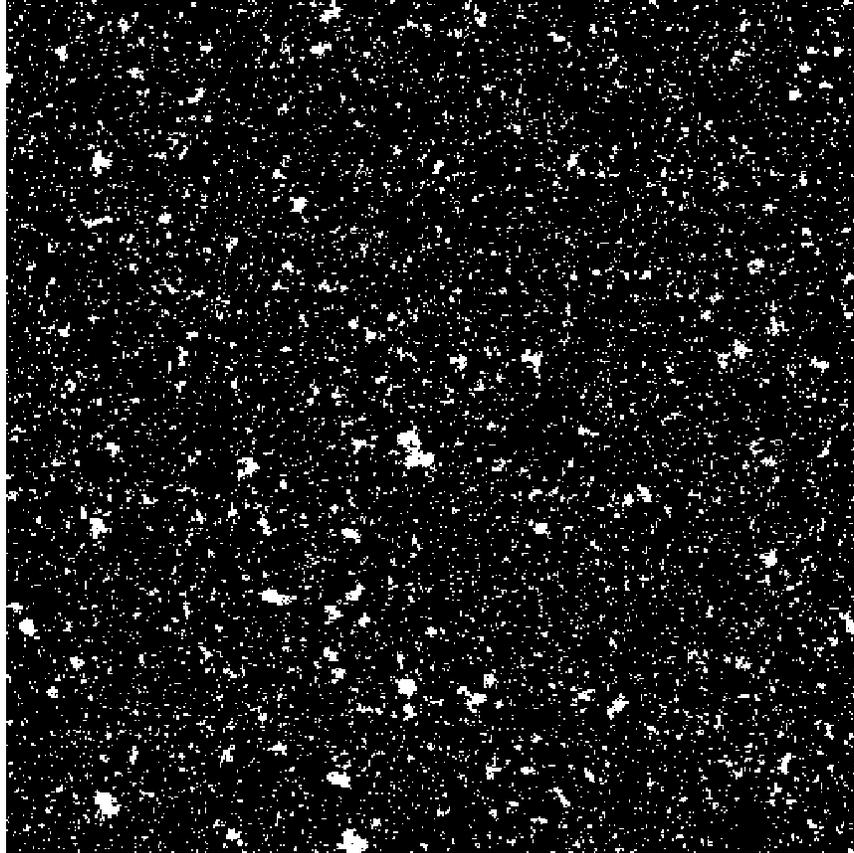


Figure 3: Ising model, 512×512 lattice, toroidal boundary conditions, somewhat above the critical value ($\alpha = 0$, $\beta = 1.05 \cdot \beta_{\text{crit}}$).

the limit of finite-dimensional distributions that condition on a boundary having just one color on the boundary, then the limit (still at $\beta = \infty$) has all pixels that color (the boundary color). Figure 3 shows one realization of an Ising model (two colors) just a little above the critical value. Now there is one clearly largest blob (although not infinite, of course, because the lattice is finite, and must be finite in order to be simulated).

This change of behavior at the critical value is an example of a *phase transition*. The most familiar phase transitions are solid to liquid to gas where there is a qualitative change in the behavior of the material at a certain temperature (melting point and boiling point). Magnets also undergo

phase transitions. At the Curie temperature a permanent magnet loses its magnetism. The Ising and Potts models model this sort of behavior.

7 Simulation

7.1 Naive Metropolis and Gibbs

It is easy to do Gibbs or Metropolis for a Potts model using the spatial Markov property.

Suppose we update one variable at a time (one X_t). Let x be the current state, and y be the proposal, then $x_s = y_s$ for all s except for $s = t$, if X_t is the variable being updated. As always, we can do either a random scan or a fixed scan (either works). Let \mathcal{Y} denote the possible values of the proposal, there are $|C|$ such possible values, differing only in the values of y_t , which takes as values each of the $|C|$ colors.

For the Gibbs update, the proposal distribution is the conditional distribution of X_t given the rest of the X_s . By the spatial Markov property, this only depends on the at most four X_s that are neighbors of X_t . If h is the unnormalized density of the process, then this conditional distribution is given by

$$\begin{aligned} & h(y) / \sum_{z \in \mathcal{Y}} h(z) \\ &= \exp \left(\alpha(y_t) + \beta \sum_{\substack{s \in \mathcal{D}^2 \\ s \sim t}} I(y_t = x_s) \right) / \sum_{z \in \mathcal{Y}} \exp \left(\alpha(z_t) + \beta \sum_{\substack{s \in \mathcal{D}^2 \\ s \sim t}} I(z_t = x_s) \right) \end{aligned}$$

The sums in the exponentials are only over the neighbors of node t (at most four of them) and the other sum (over \mathcal{Y}) has only as many terms as colors, and we want only a few of them (two for the Ising model).

There are many possible variable-at-a-time Metropolis proposals. The simplest uses the uniform distribution on all the colors. Another uses the uniform distribution on all the colors except the current color (the value of x_t). In any case the odds ratio is

$$\frac{h(y)}{h(x)} = \exp \left(\alpha(y_t) - \alpha(x_t) + \beta \sum_{\substack{s \in \mathcal{D}^2 \\ s \sim t}} [I(y_t = x_s) - I(x_t = x_s)] \right)$$

is simple to calculate because it only involves at most four neighboring nodes. The Metropolis update then does Metropolis rejection.

One-variable-at-a time Gibbs and Metropolis are simple, but neither works well for large lattices except at very small values of β (very near zero, very weak dependence). For very large values of β and very large lattices it would take these samplers billions of years to get from a state that is predominately one color to a state that is predominantly another color.

From 1953 when the Metropolis algorithm was invented until 1987 when the Swendsen-Wang algorithm was invented, the Ising model was considered the archetypical hard problem for MCMC. The Swendsen-Wang algorithm made it easy.

7.2 Swendsen-Wang

The Swendsen-Wang algorithm (Swendsen and Wang, 1987) works for any Potts model with positive dependence ($\beta \geq 0$). It works in a way that seems at first counterintuitive, even bizarre. It complicates the problem by introducing new variables. The new variables are Bernoulli random variables Y_{st} for $(s, t) \in E$ (the set of neighbor pairs). These Bernoulli random variables Y_{st} are often called *bonds*. To distinguish then, the old variables X_t are called *spins*.

Having increased the number of variables and hence the state space of the Markov chain, we now need to specify the desired equilibrium distribution on this new state space. We keep the same distribution (given by the Potts model) for the old variables (the X_t). Of course, this is now the marginal (for the vector X). We now specify a conditional for Y given X (where Y is the vector of all the bond variables), and that completes the specification of the joint distribution of X and Y .

The Y_{st} are conditionally independent given X and

$$P(y_{st} = 1|x) = \begin{cases} \gamma_{st}, & x_s = x_t \\ 0, & \text{otherwise} \end{cases}$$

where the γ_{st} are constants to be named later (we choose them after we see what makes the Swendsen-Wang algorithm simple).

The Swendsen-Wang algorithm is a block Gibbs algorithm. It samples from the conditional distribution of bonds given spins (Y given X), then from the conditional distribution of spins given bonds (X given Y). We have just seen one of these conditionals. Clearly it is simple to sample from because of the conditional independence of the Y_{st} .

The conditional distribution of X given Y is a bit more complicated. First note that any two nodes connected by a bond must be the same color ($Y_{st} = 1$ implies $X_s = X_t$). This is clearly also the case for any two nodes connected by a chain of bonds. Define a new neighbor relation \sim_Y as follows

$$s \sim_Y t \quad \text{if and only if} \quad s \sim t \text{ and } Y_{st} = 1$$

This defines a symmetric relation on the lattice ($s \sim_Y t$ if and only if $t \sim_Y s$). Now consider the reflexive transitive closure of this relation. Call that \approx_Y . Reflexive closure means $t \sim_Y t$ for all t , and transitive closure means the minimal relation that contains \sim_Y and satisfies the transitivity property

$$r \approx_Y s \text{ and } s \approx_Y t \quad \text{implies} \quad r \approx_Y t$$

That is $r \approx_Y t$ holds if and only if for some k there are nodes s_1, \dots, s_k such that

$$r \sim_Y s_1 \sim_Y s_2 \sim_Y \dots \sim_Y s_{k-1} \sim_Y s_k \sim_Y t$$

It is not easy to figure out what the relation \approx_Y is (more on that presently). But it is an equivalence relation (reflexive, symmetric, and transitive). Hence it defines equivalence classes

$$[t] = \{ s \in \mathbb{D}^2 : s \approx_Y t \}$$

which partition the lattice. We will call these equivalence classes *patches* for short. The algorithm for figuring these out is called the algorithm for maximal connected components of an undirected graph or just the algorithm for equivalence classes. The R function `weak` in the contributed package `pooh` (Geyer, 2013), which is available on CRAN, does this computation. The R function `potts` in the contributed package `potts` (Geyer and Johnson, 2014), which is available on CRAN, does simulates Potts models using the Swendsen-Wang algorithm and uses the same algorithm as the `weak` function in the `pooh` package to calculate the equivalence classes of \approx_Y .

We already know that nodes in a patch $[t]$ must have the same color ($X_s = X_t$ if $s \in [t]$). Thus we consider the probability distribution (given Y) of the colors of the patches.

The unnormalized joint density of spins and bonds is

$$h(x, y) = \prod_{t \in \mathbb{D}^2} e^{\alpha(x_t)} \prod_{(s,t) \in E} \left[e^{\beta} \gamma_{st}^{y_{st}} (1 - \gamma_{st})^{1-y_{st}} \right]^{I(x_s=x_t)} (1 - y_{st})^{1-I(x_s=x_t)} \quad (5)$$

In the last term we are using the convention $0^0 = 1^0 = 1^1 = 1$, so the last term is equal to 1 unless $x_s \neq s_t$ and $y_{st} = 1$, in which case it is $0^1 = 0$, which makes this impossible, as our specification of the conditional distribution of Y given X required.

Since there is no difference between an unnormalized joint density and an unnormalized conditional density (they differ only in their normalizing constants), this means (5) is also the unnormalized conditional of x given y considered as a function of x for fixed y .

We already know that x has probability zero given y unless it is constant on patches. Call x that have this property *patch respecting*.

For a patch respecting x we have $x_s \neq x_t$ implies x_s and x_t are in different patches and hence $y_{st} = 0$. Hence the last term in (5) is always equal to one (is never 0^1) for all patch respecting x . Since the other terms in (5) are never zero if we choose the γ_{ij} to be neither zero or one, we have proved that the conditional probability of x given y is nonzero if and only if x is patch respecting subject to this restriction on the choice of the γ_{st} .

Hence we now restrict attention to patch respecting x . Then we can write

$$\begin{aligned} h(x, y) = \exp & \left(\sum_{t \in \mathbb{D}^2} \alpha(x_t) \right. \\ & + \sum_{\substack{(s,t) \in E \\ s \approx_Y t}} \left[\beta + y_{st} \log(\gamma_{st}) + (1 - y_{st}) \log(1 - \gamma_{st}) \right] \\ & \left. + \sum_{\substack{(s,t) \in E \\ s \not\approx_Y t}} \left[\beta + \log(1 - \gamma_{st}) \right] I(x_s = x_t) \right) \end{aligned}$$

The simplification of the \approx_Y terms comes from $s \approx_Y t$ implies $x_s = x_t$, and the simplification of the $\not\approx_Y$ terms comes from $s \not\approx_Y t$ implies $y_{st} = 0$.

We now choose γ_{ij} to make the $\not\approx_Y$ terms go away

$$1 - \gamma_{st} = e^{-\beta}, \quad (s, t) \in E. \quad (6)$$

Then we note that the \approx_Y terms do not contain x and hence can be considered part of the normalizing constant for an unnormalized conditional density of x given y (that is, they can be dropped). Hence we can write (with this choice of γ_{st})

$$h(x | y) = \exp \left(\sum_{t \in \mathcal{D}^2} \alpha(x_t) \right)$$

and if we define the set of all patches

$$\mathcal{P} = \{ [t] : t \in \mathbb{D}^2 \}$$

this conditional density simplifies to

$$h(x | y) = \prod_{A \in \mathcal{P}} \exp\left(|A| \cdot \alpha(x_A)\right)$$

where we write $\alpha(x_A)$ for the value of the function α evaluated at any x_t for any $t \in A$ (since they all have the same color c , they all have the same value $\alpha(c)$). We are still insisting that x be patch respecting. This conditional distribution is remarkably simple. Note that because of the product over patches, the patch colors are conditionally independent given the bonds. It gets even simpler in the special case of interest to the physicists where $\alpha_c = 0$ for all c . Then all patch colors are equally likely.

We summarize.

- [Update y given x .] The bonds are conditionally independent given the pixels.
 - If $(s, t) \in E$ and $x_s = x_t$, then $Y_{st} = 0$ with probability $e^{-\beta}$.
 - If $(s, t) \in E$ and $x_s \neq x_t$, then $Y_{st} = 0$ with probability one.
- [Calculate patches.] Run an algorithm to determine the equivalence classes \mathcal{P} (which depend on y).
- [Update x given y .] The patch colors are conditionally independent given the patches. The probability of color c for patch A is proportional to

$$\exp\left(|A| \cdot \alpha(x_A)\right)$$

References

- Geyer, C. J. (2013). R package pooh, version 0.2-1. Partial orders and relations. <http://cran.r-project.org/package=pooh>.
- Geyer, C. J. and Johnson, L. T. (2014). R package potts, version 0.5-2. Markov chain Monte Carlo for Potts models. <http://cran.r-project.org/package=potts>.

- Hoel, P. G., Port, S. C. and Stone, C. J. (1972). *Introduction to Stochastic Processes*. Boston: Houghton Mifflin. Republished, Waveland Press, Prospect Heights, Illinois, 1986.
- Kindermann, R. and Snell, J. L. (1980). *Markov Random Fields and their Applications*. Providence: American Mathematical Society.
- Potts, R. B. (1952). Some generalized order-disorder transformations. *Mathematical Proceedings of the Cambridge Philosophical Society*, **48**, 106–109.
- Swendsen, R. H. and Wang, J. S. (1987). Nonuniversal critical dynamics in Monte Carlo simulations. *Physical Review Letters*, **58**, 86–88.