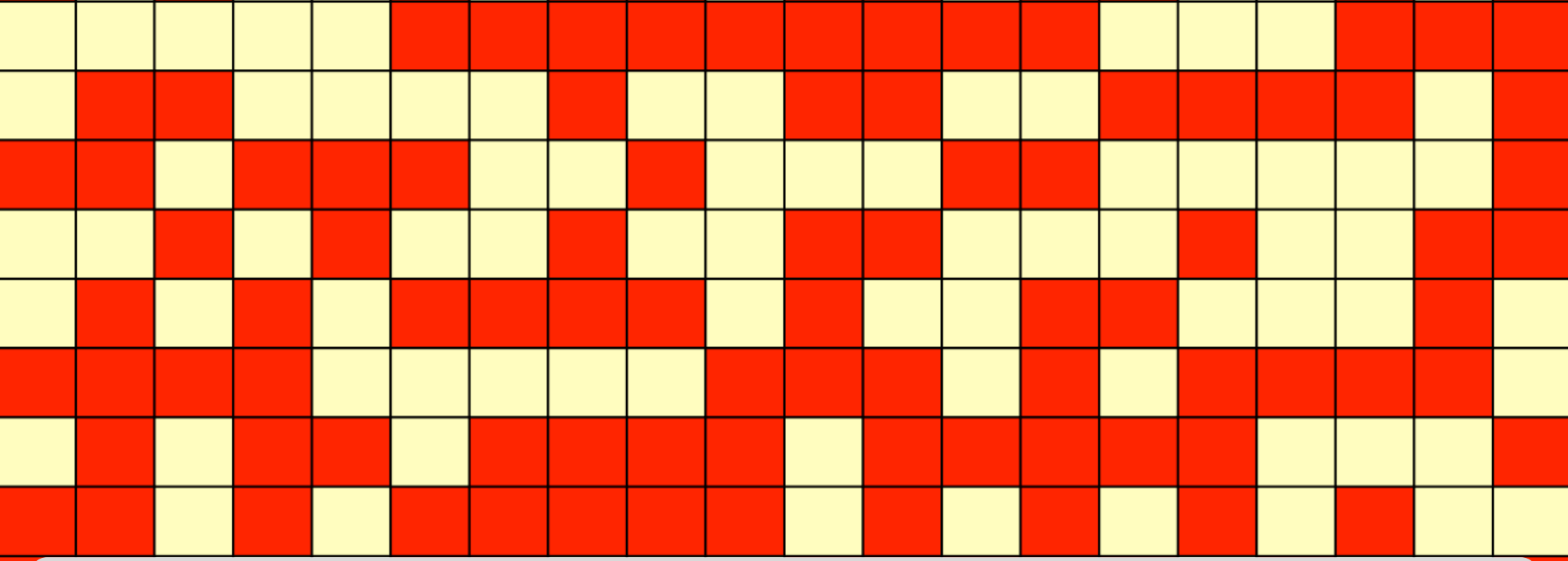


Stationary states of opinion diffusion

Project for the exam: AMS (DSE)

Paola Serra and Marzio De Corato

January 11, 2022

Theoretical Framework

Statistical Mechanics

"Ludwig Boltzmann, who spent much of his life studying statistical mechanics, died in 1906, by his own hand. Paul Ehrenfest, carrying on the work, died similarly in 1933. Now it is our turn to study statistical mechanics." States of Matter (1975), by David L. Goodstein

Concepts of statistical mechanics: ensemble [6]

- **Statistical ensemble**: a large number of virtual copies of a system ; each of them is a possible state of the real system (epistemic probability) . It is the formalization of a repeated experiment proposed by Gibbs (empirical probability)
- **Microcanical ensamble:** $p = 1/W$ W is the number of microstates
- **Canonical ensamble:** $p = \frac{1}{Z} \exp\left(-\frac{E}{kT}\right)$ where $Z = \sum_i \exp\left(\frac{-E_i}{k_b T}\right)$

$$\langle A(x) \rangle = \frac{1}{Z} \int dx_s A(x_s) exp\left[-\frac{H^{(S)}(x_s)}{k_b T}\right] \tag{1}$$

$$Z = \int dx_s exp\left[-\frac{H^{S}(x_s)}{k_b T}\right] \tag{2}$$

$$\left\langle H(x)^2 \right\rangle - \langle H(x) \rangle^2 = k_b T^2 C \tag{3}$$

Ising Model

Ising model [4, 3]

- An array of atoms that can take states ± 1. The energy of the system is given by $E(\mathbf{x}, J, H) = -\left[\frac{1}{2}\sum_{m,n} J_{mn} x_m x_n + \sum_n H x_n\right]$ where J is the coupling constant between two neighbour sites, and H is an external field.
- The probability of the system to be in the state x is given by $p(x|\beta, J, H) = \frac{1}{Z(\beta,J,H)} \exp[-\beta E(x, J, H)]$ (canonical ensemble) where $\beta = 1/k_b T \quad Z(\beta, J, H) = \sum_x exp\left[-\beta E(x, J, H)\right]$
- It is useful to characterize the order level of a lattice (macroscopic) with the (spatial) correlation functions (whose input are microscopic quantities). In particular, for the Ising model, these are given by the following expression (with $H = 0$)

$$g(m) = \frac{\langle \sigma_i \sigma_{i+m} \rangle - \langle \sigma_i \rangle \langle \sigma_{i+m} \rangle}{1 - \langle \sigma_i \rangle \langle \sigma_{i+m} \rangle} = \langle \sigma_i \sigma_{i+m} \rangle$$

Numerical simulations

"Never make a calculation until you know the answer. Make an estimate before every calculation, try a simple physical argument (symmetry! invariance! conservation!) before every derivation, guess the answer to every paradox and puzzle. Courage: No one else needs to know what the guess is. Therefore make it quickly, by instinct. A right guess reinforces this instinct. A wrong guess brings the refreshment of surprise. In either case life as a spacetime expert, however long, is more fun!" John Archibald Wheeler

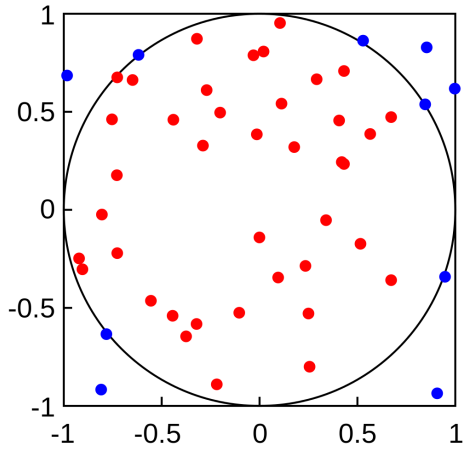
Monte Carlo method [2, 4, 5]

Basic Monte Carlo

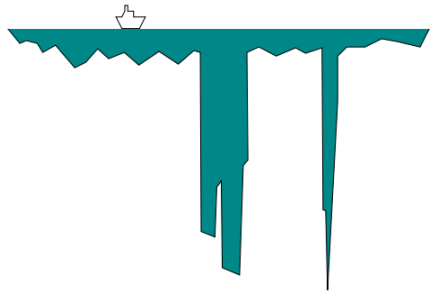
- $\mathbb{E}[f(X)] = \int f(x)p(x)dx \approx (1/S)\sum_{s=1}^{S} f(x_S)$
- Generate S samples from the distribution that we would evaluate
- Approximate the distribution using the empirical one
- Basic example the value of π

$$I = \int_{-r}^{r}\int_{-r}^{r} \mathbb{I}(x^2 + y^2 \leq r^2)dx\ dy$$

$$\approx 4r^2 \frac{1}{S}\sum_{s=1}^{S} f(x_s, y_s) \qquad (4)$$



But in this case ?



Images taken from [1, 4]

Monte Carlo method [2]

- Lets consider the generic integral $I = \int_a^b dx\ f(x)$
- This can be recast in the following form $I = \int_0^1 dx\ w(x)\frac{f(x)}{w(x)}$ where $w(x)$ is the probability distribution of the sampling points
- If w(x) is the derivative of u(x) (non-decreasing,non negative) we have $I = \int_0^1 du\ \frac{f[x(u)]}{w[x(u)]}$
- If one considers L random values of u uniformely distributed in the interval [0,1] we have $I \approx \frac{1}{L}\sum_{i=1}^{L}\frac{f[x(u)]}{w[x(x)]}$
- The choice of w is crucial since $\sigma = \frac{1}{L}\left[\left\langle\left(\frac{f}{w}\right)^2\right\rangle - \left\langle\frac{f}{w}\right\rangle^2\right]$
- Brute Force: w constant...not a good idea

Monte Carlo method [6]

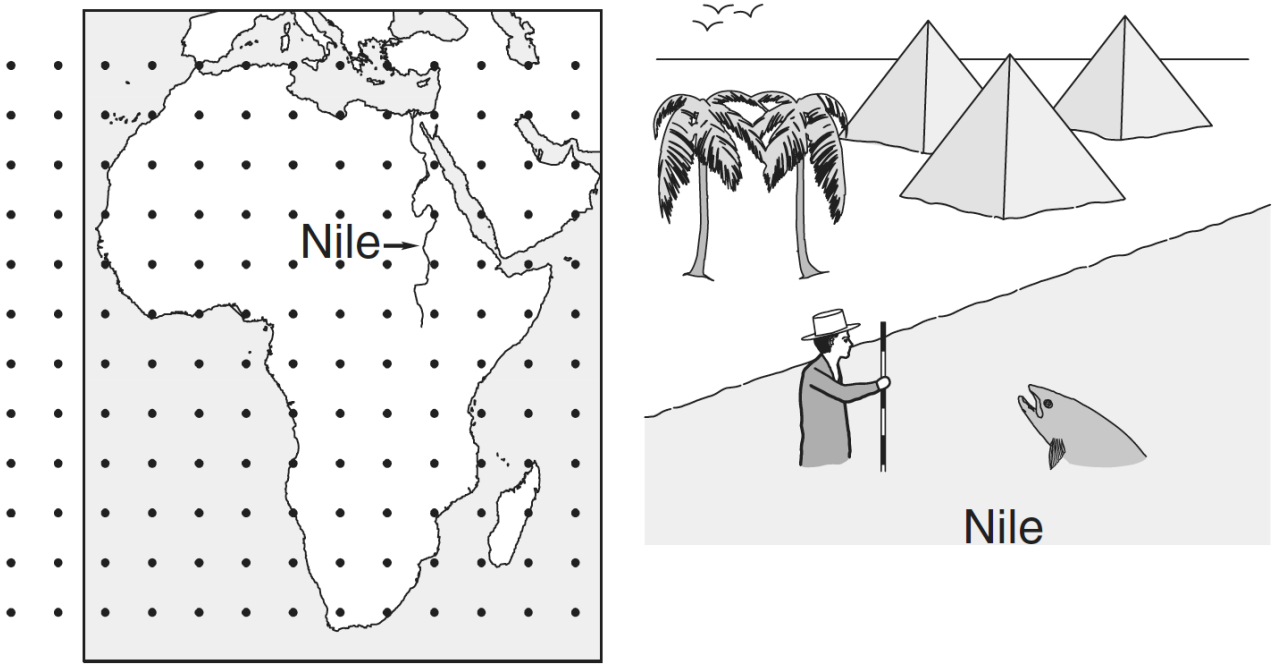


Image taken from [2]

Monte Carlo method: Metropolis idea [2]

$$\langle A \rangle = \frac{\int d\mathbf{r}^N \exp\left[-\beta E(\mathbf{r}^N)\right] A(r^N)}{\int d\mathbf{r}^N \exp\left[-\beta E(\mathbf{r}^N)\right]} \tag{5}$$

- We have a ratio between two integrals, therefore what we need to sample is the ratio and not the integrals alone
- PROBLEM: we do not know the form of the denominator (if we know it we do not need the Monte Carlo method)
- The probability density is $N(r^N) = \exp\left[-\beta E(\mathbf{r}^N)\right]/Z$
- Metropolis idea: randomly generate points with this last probability distribution. In this case we have $\langle A \rangle \approx 1/L \sum_{i=1}^{L} n_i A(\mathbf{r}_i^N)$

Monte Carlo method: Metropolis idea [2]

- How the points are generated ? With a Boltzmann weighted Markov chain
- $\pi(old \rightarrow new) = \alpha(old \rightarrow new) \times acc(old \rightarrow new)$ where π is the transition probability element from the old state to the new state,α is the matrix element of Markov Chain and acc is the acceptance ratio.
- Detailed balance condition at the equilibrium
- If the energy of the new state is lower with respect to the old one the move is always accepted. Otherwise a random number is generated from the uniform distribution between the interval $[0, 1]$, since $acc(old \rightarrow new) = \exp\left[-\beta\left(E(new) - E(old)\right)\right] < 1$. The move is accepted if the random number is lower than $acc(old \rightarrow new)$
- Note that since a ratio is involved the Z function is no more necessary
- $\pi(old \rightarrow new)$ should be ergodic

Goals and methods

Goals and methods

- Reproduce the main result for a 2D anti ferromagnetic lattice ($J = -1$) with no external magnetic field ($H = 0$) with the montecarlo-metropolis
- Once checked that the script provide the correct results apply it to a lattice ($J = +1$). In this case the spins represent an opinion and the sites people. The goal is to find the stationary states (at $T = 0$ and $T \neq 0$)
- Introduce in the lattice some blocks that never change their status. These islands represent groups that never change mind and only diffuse their ideas. (at $T = 0$ and $T \neq 0$)

Simulation features

- 10x10 lattice
- Periodic boundary conditions $\rightarrow$ the topology of a torus (genus equal to 1)
- 6000 steps

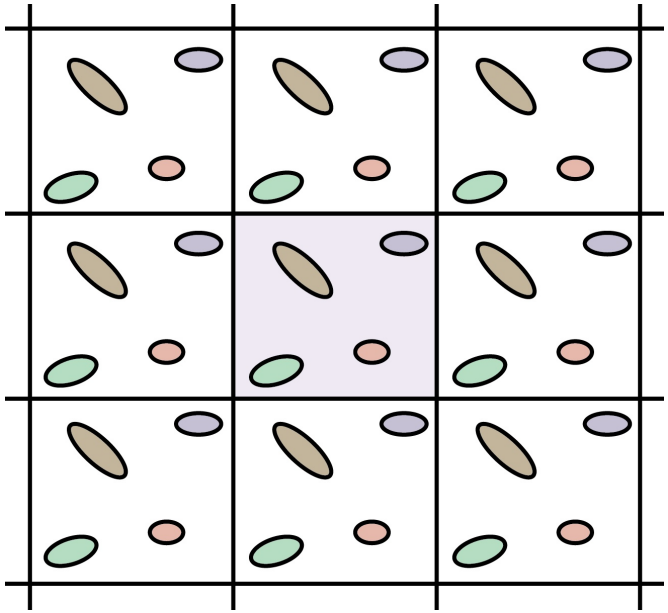


Image taken from [1]

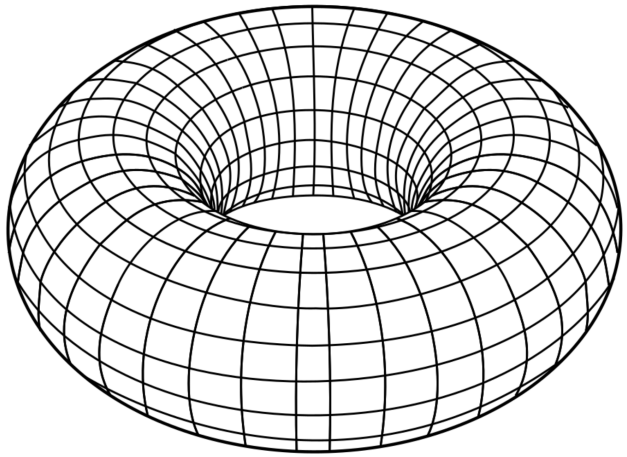
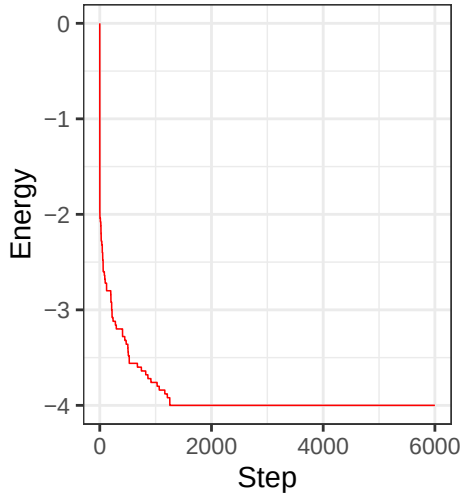
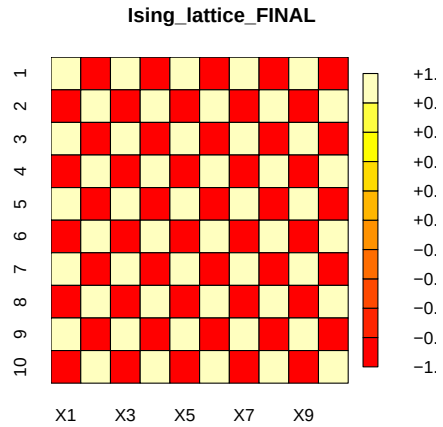


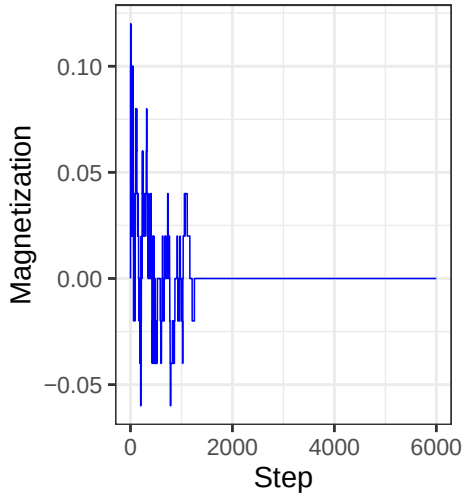
Image taken from [1]

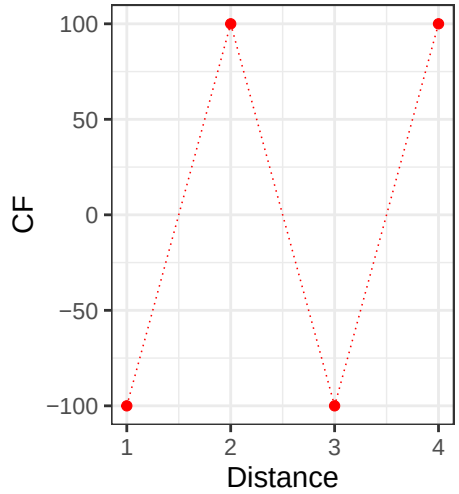
Results

Antiferromagnetic J=-1, T=0

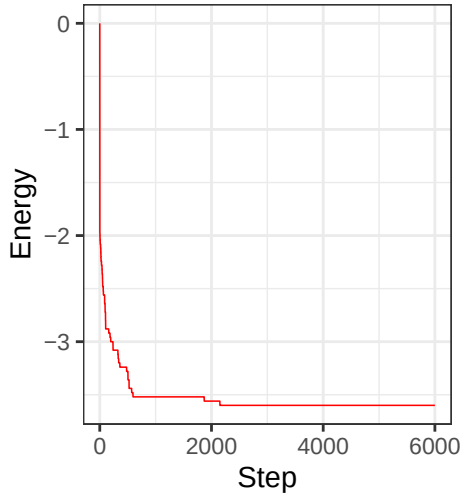


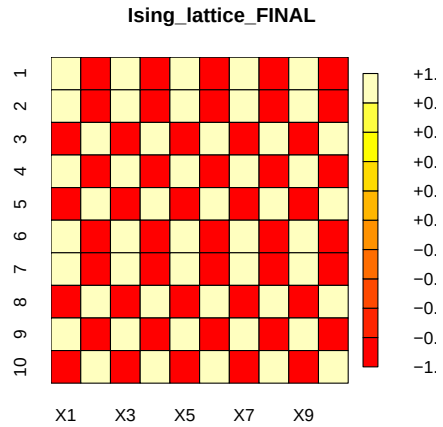


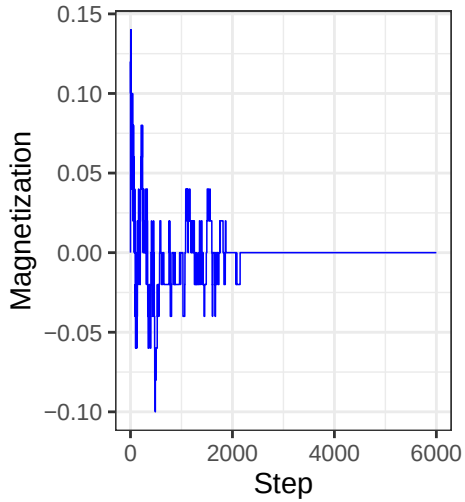


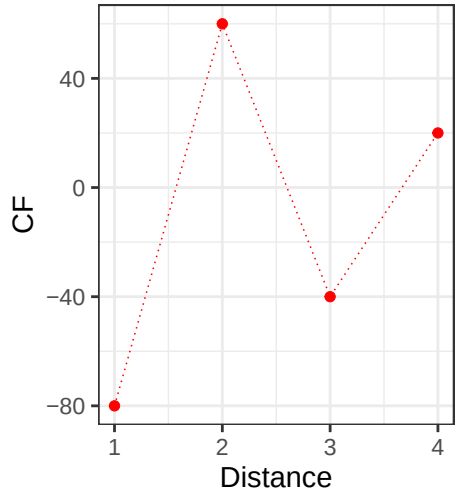


Antiferromagnetic J=-1, T=1.5

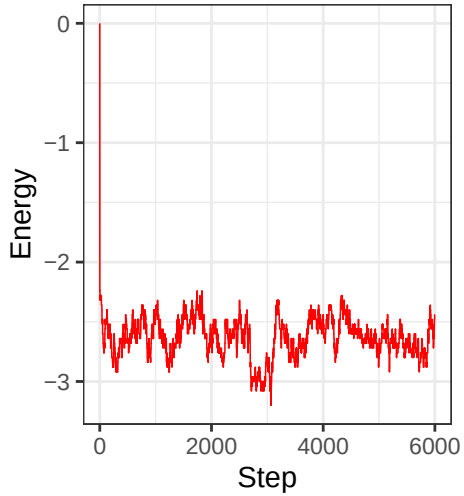


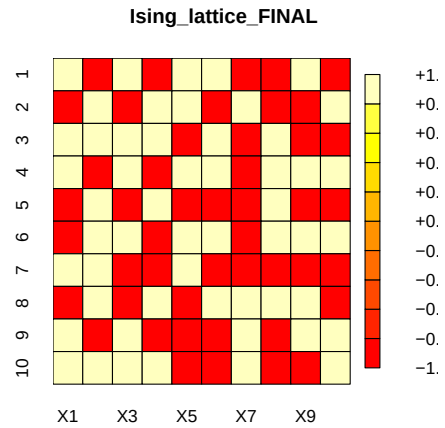


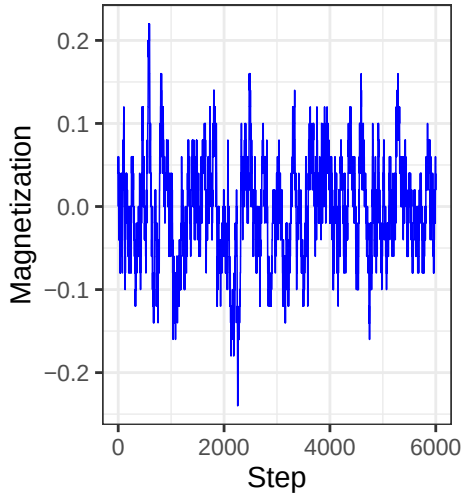


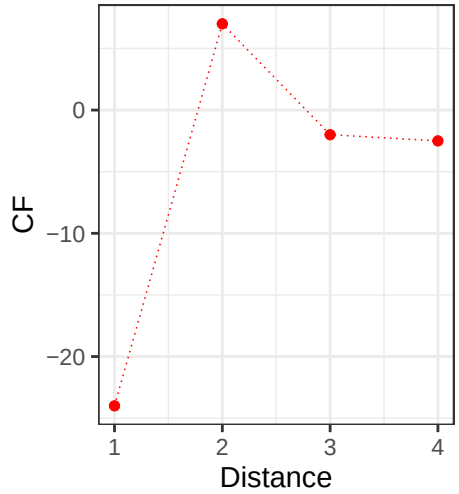


Antiferromagnetic J=-1, T=4

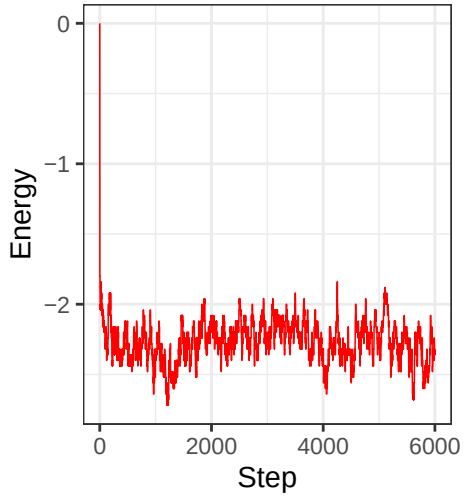


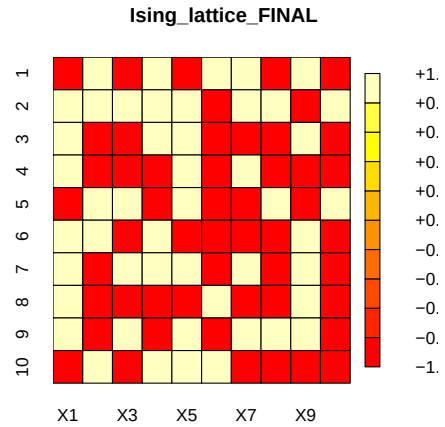


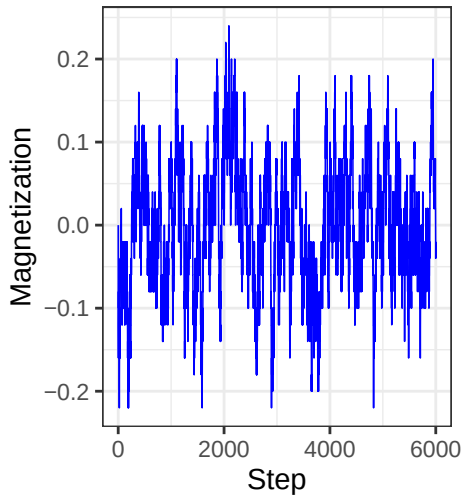


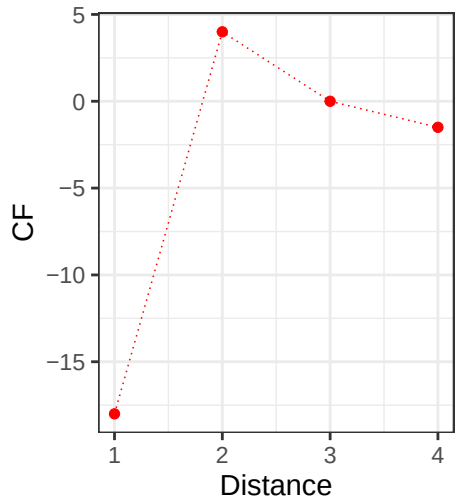


Antiferromagnetic J=-1, T=8









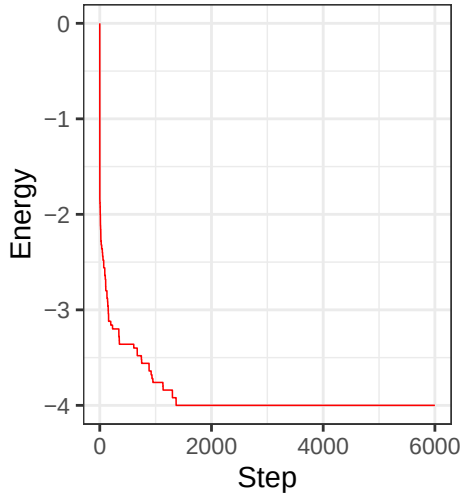
Antiferromagnetic - conclusions

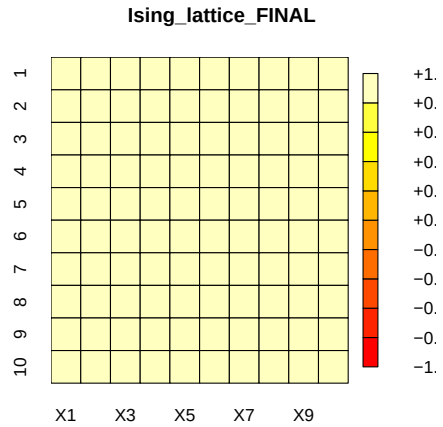
- As long as the temperature is increased the Monte-Carlo algorithm finds difficult to retrieve a stationary state. This is correct since with a higher temperature more state becomes accessible and the algorithm can escape from the trap states
- The correlation function has its absolute value diminished as the temperature is increased; moreover the correlation function tends to be much more similar to a flat curve (near 0). This is correct since the temperature increase makes accessible states in which the values are not alternated (as for T=0), therefore the coherence of the overall systems is destroyed as the temperature increase.
- This is similar to what happens to a block of a crystal when it melts: the ordered structure by which is formed is replaced by a less ordered one (liquid))

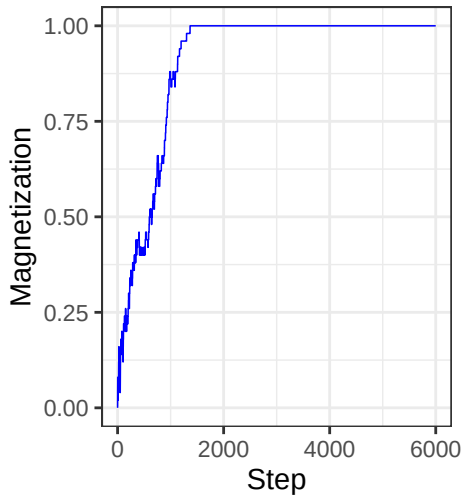
Ferromagnetic - introduction

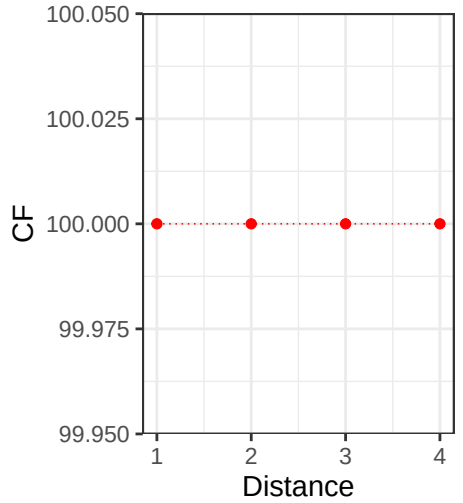
- Lets consider a lattice which models an array of people with a binary opinion ($+1 or -1$)
- This situation is identical to a ferromagnetic case of the Ising model
- For a sociological context the ferromagnetic behaviour can model the fact that people tends to follow the opinion of majority (mirror neurons) (but not always, see later ...)

Ferromagnetic J=+1, T=0

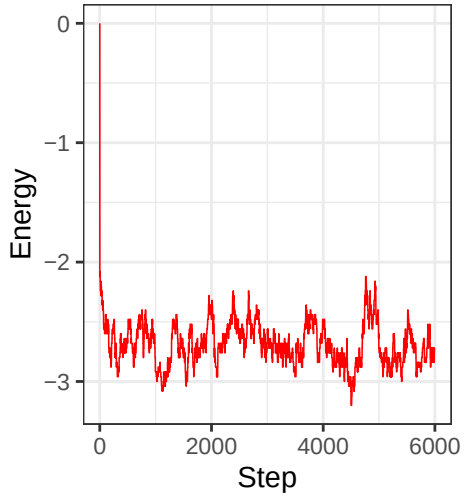


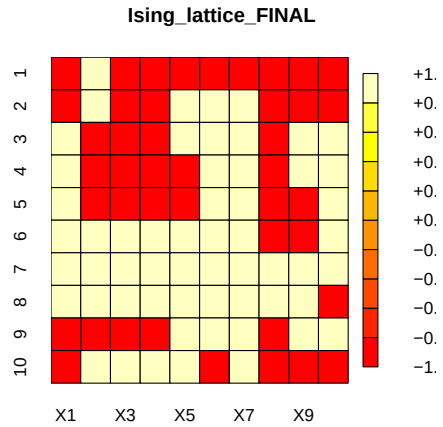


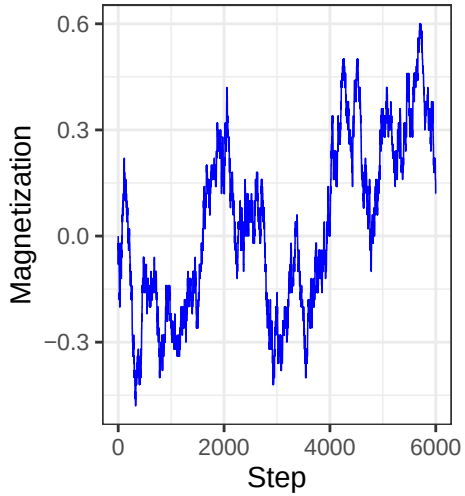


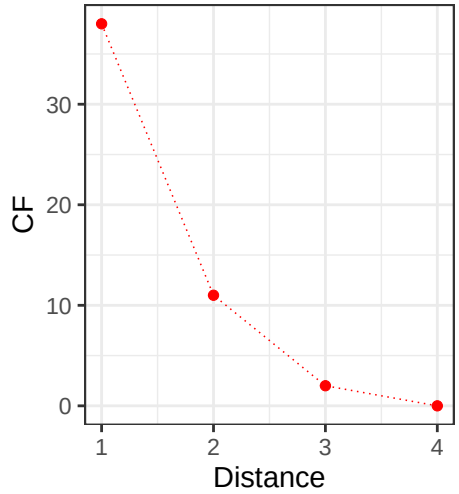


Ferromagnetic J=+1, T=8

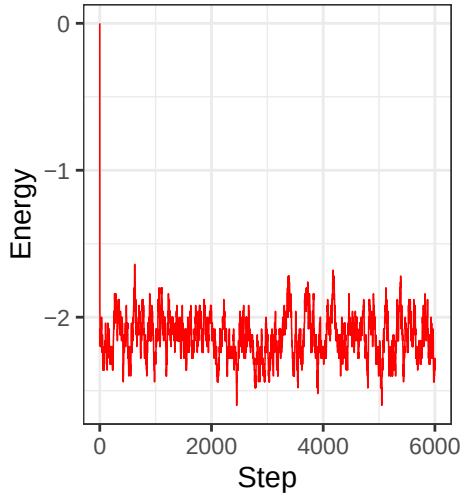


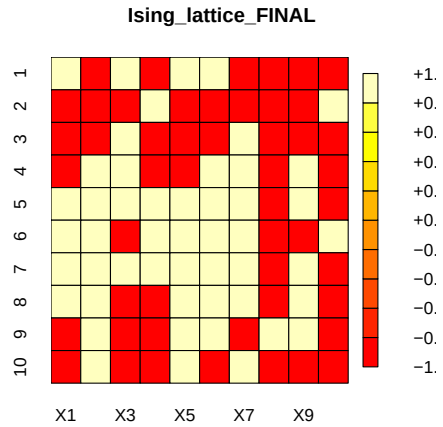


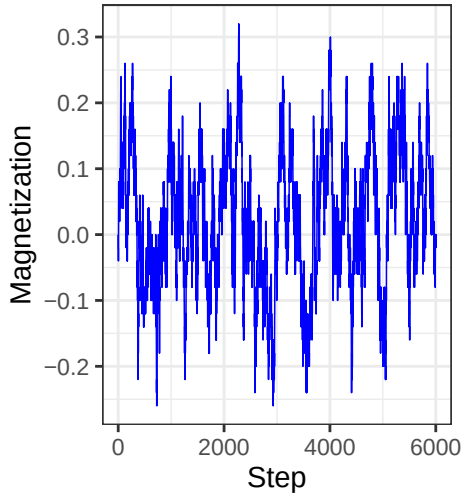


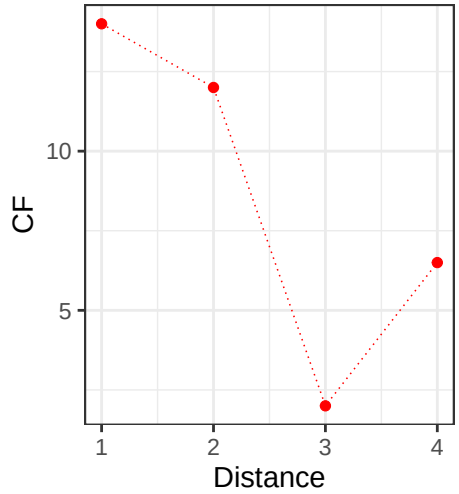


Ferromagnetic J=+1, T=16





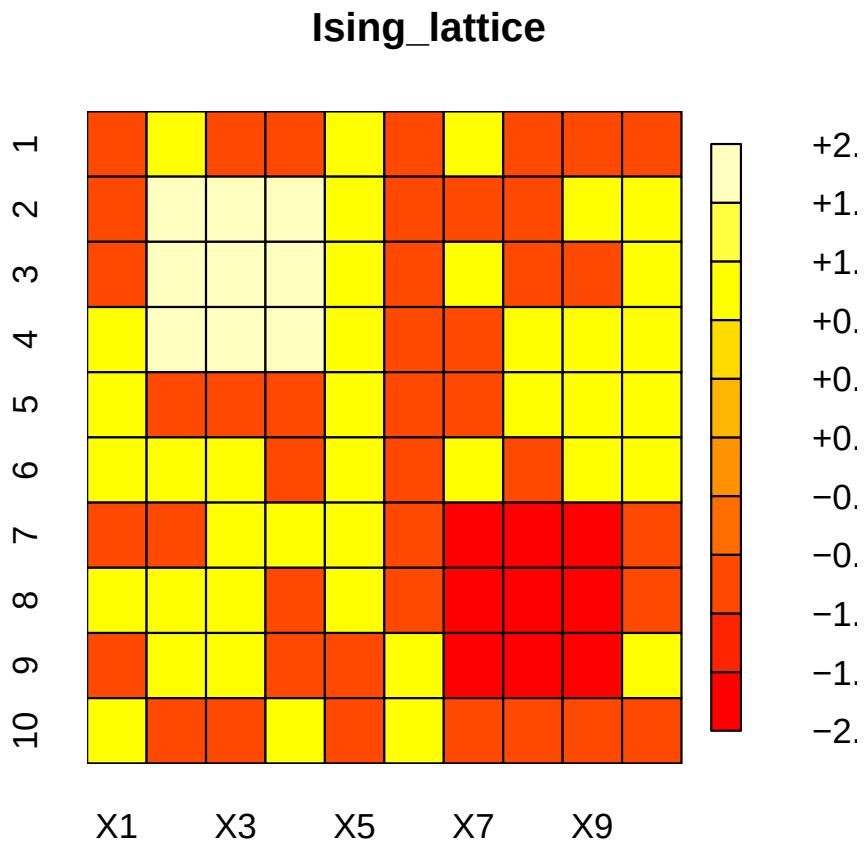




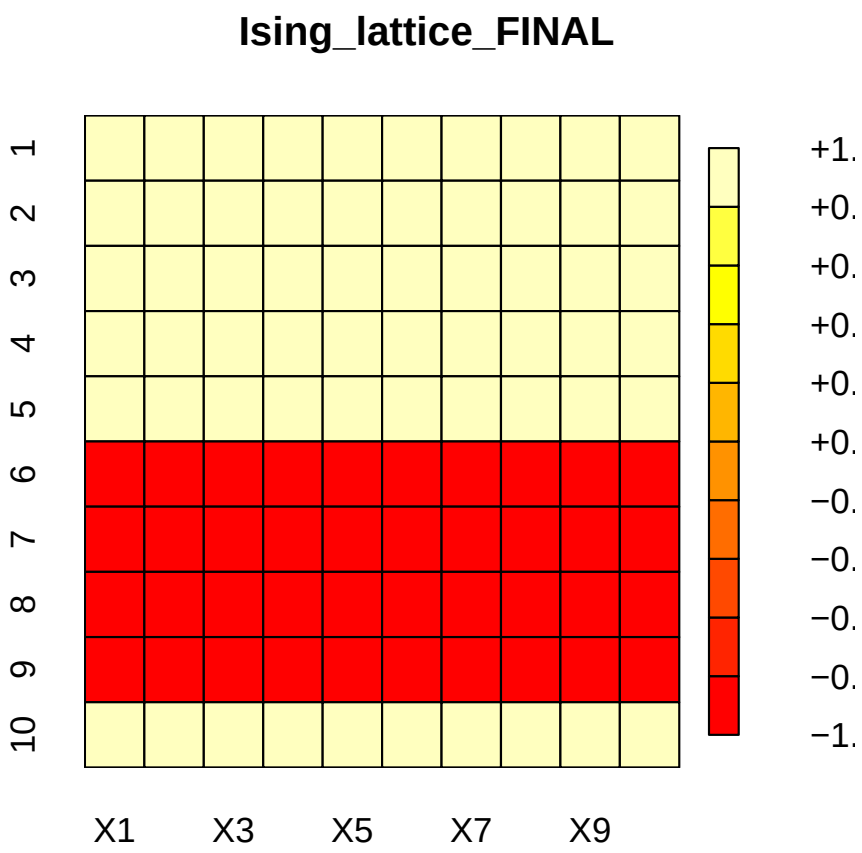
Ferromagnetic - remarks

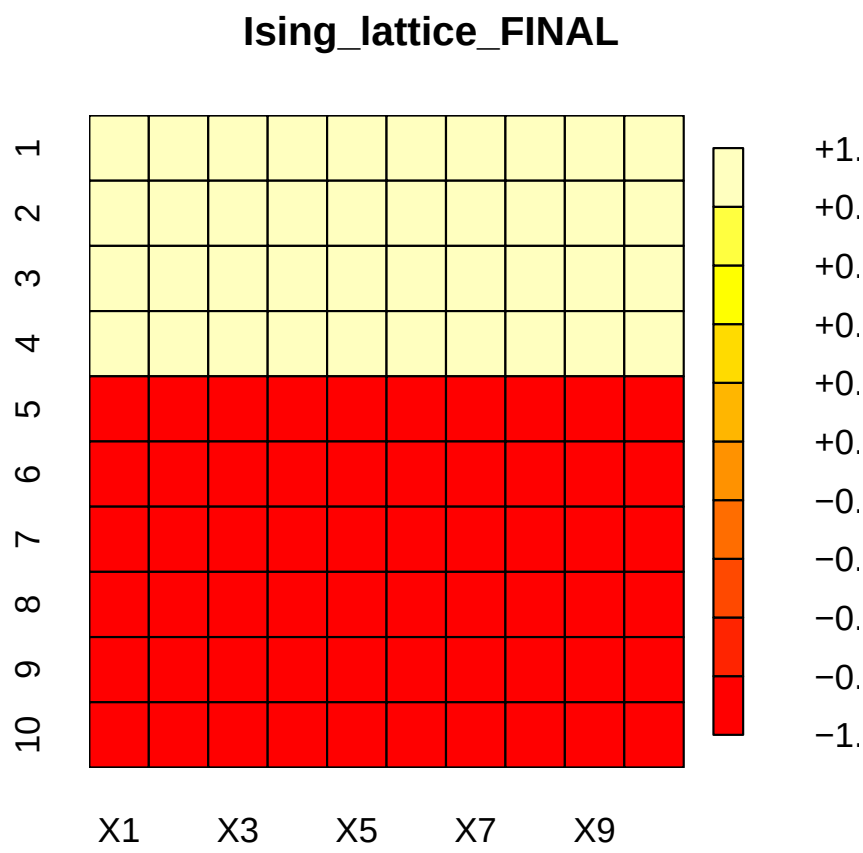
- In this case at temperature T=0 the whole lattice goes in one of the two states: all people at last will have the same opinion.
- As expected the temperature breaks also in this case the overall coherence: this may model the fact that some individuals are not always forced to follow the mass opinion
- In practice we know that opinions diffuses from some people that act as diffuser; furthermore these subject are very rare to change their ideas
- We modelled this behaviour by considering some clusters that has a fixed status

Ferromagnetic with two centers T=0

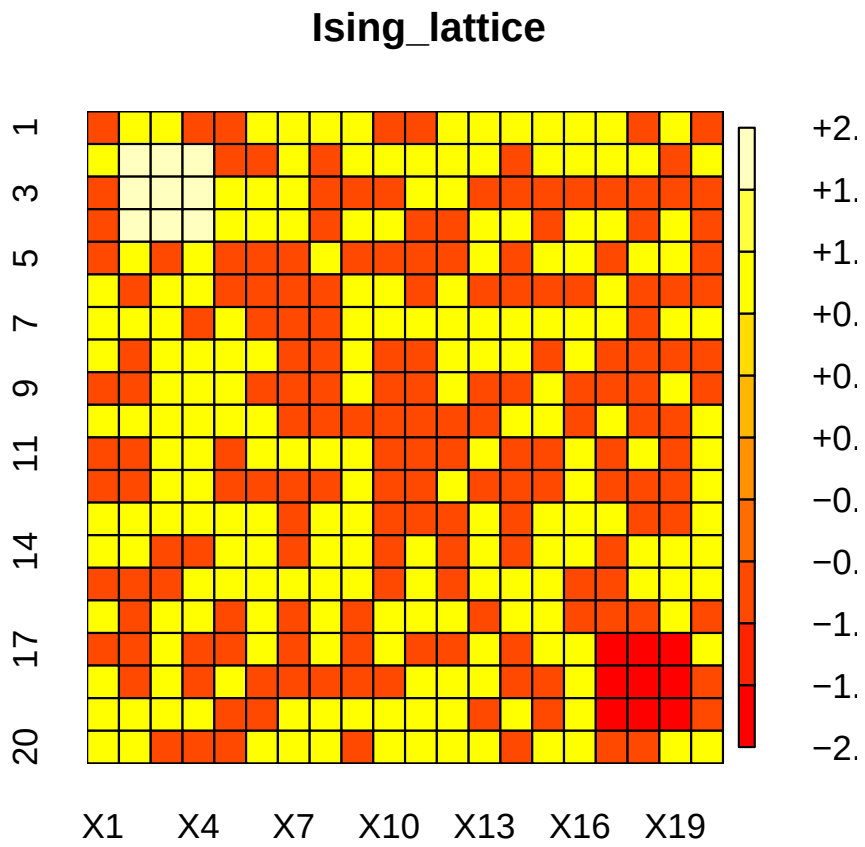


Ferromagnetic with two centers T=0

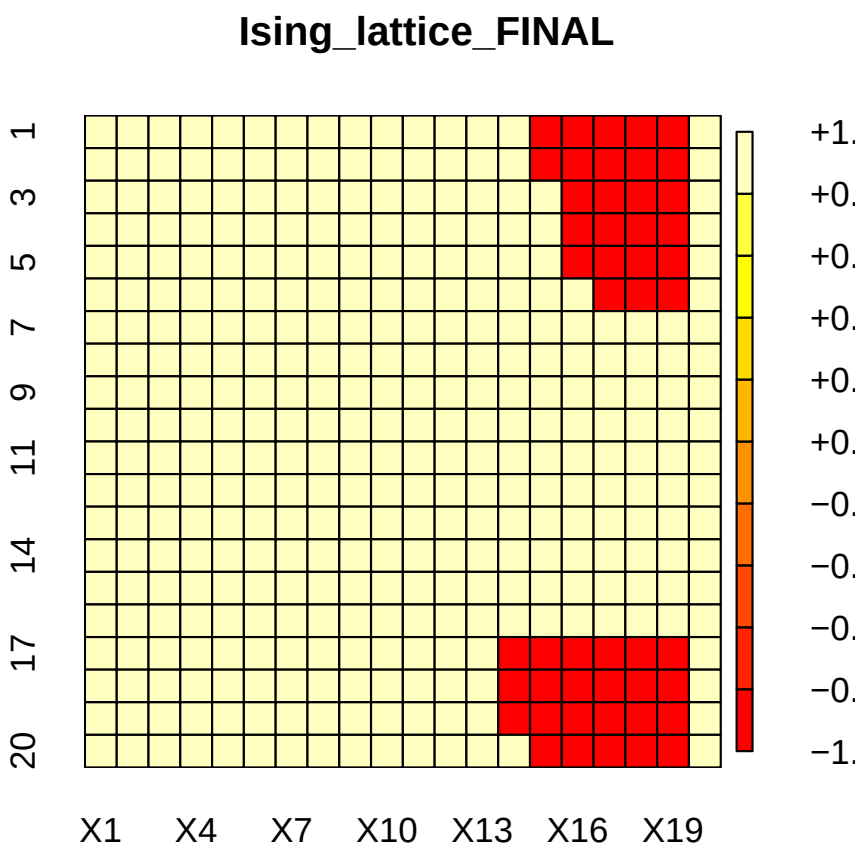


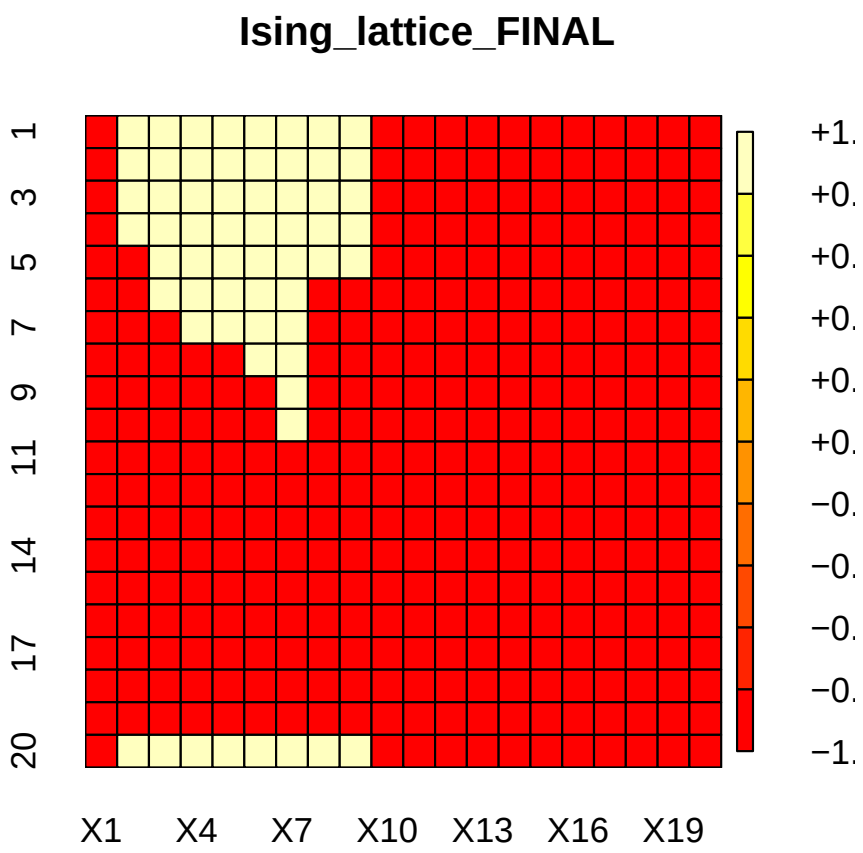


Ferromagnetic with two centers T=0, 20×20

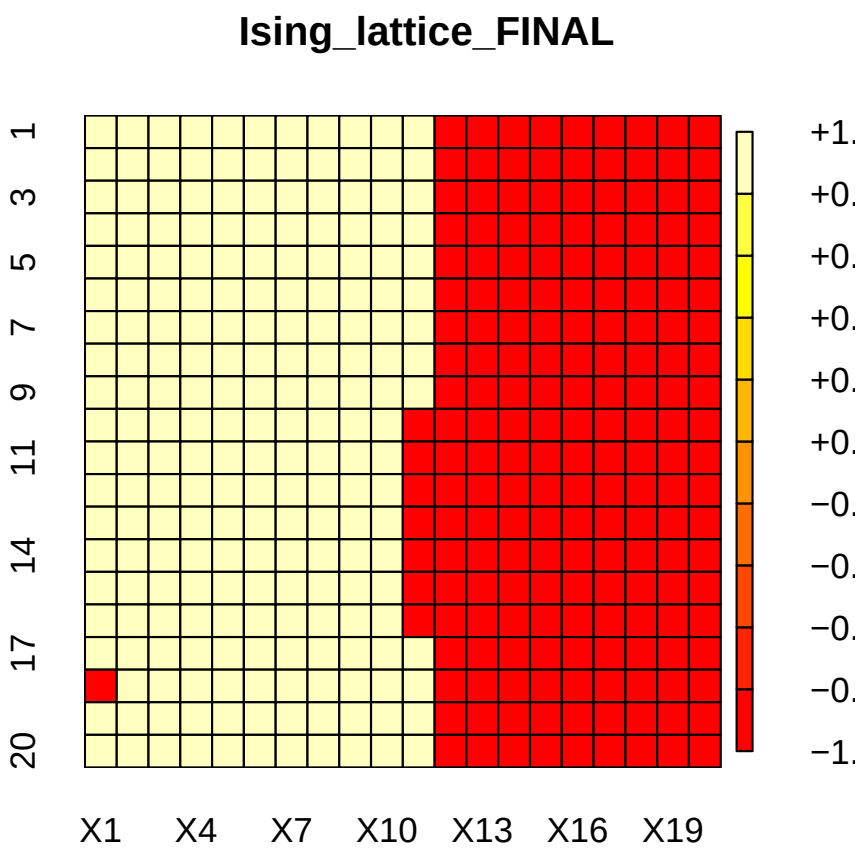


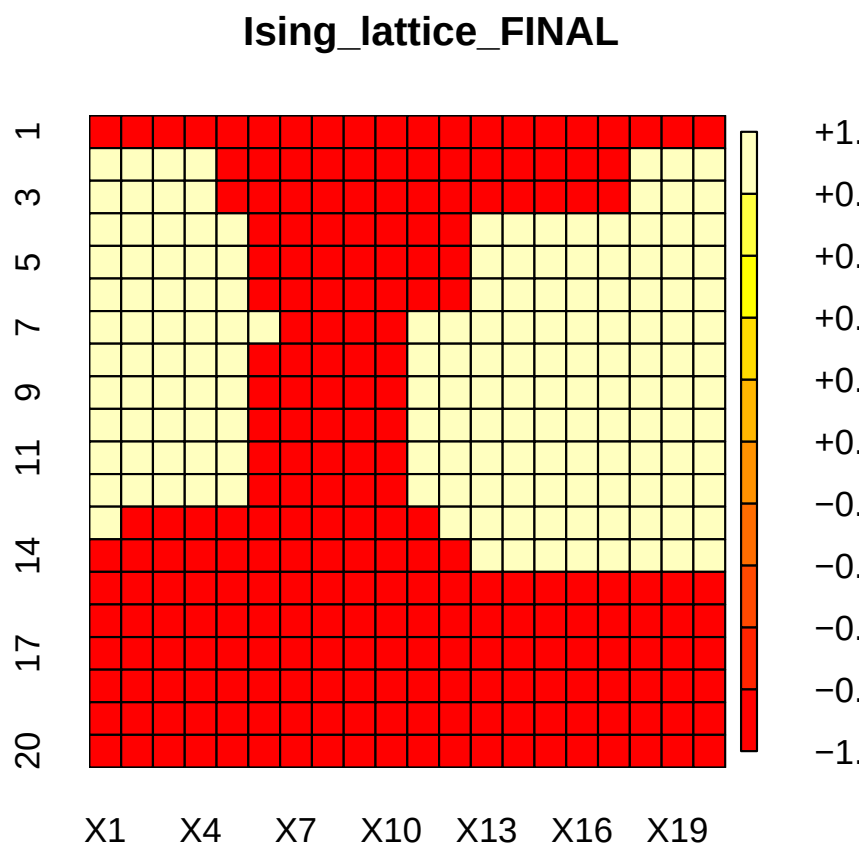
Ferromagnetic with two centers T=0, 20×20



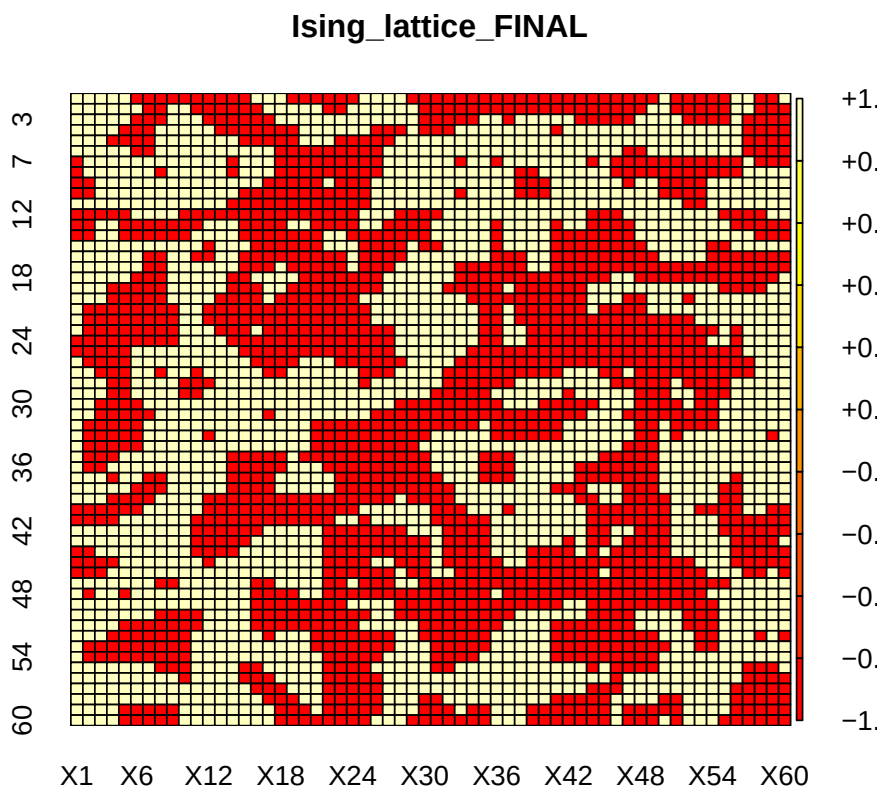


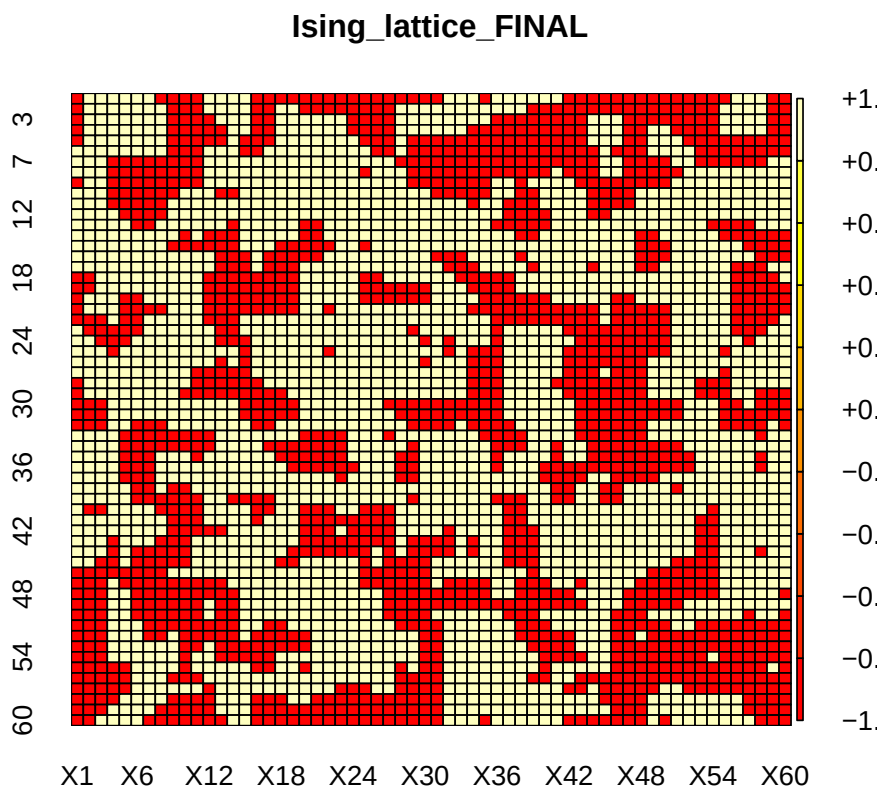
Ferromagnetic with two centers T=0.25, 20×20



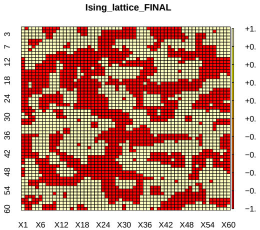


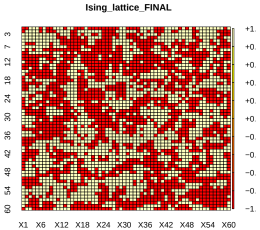
Ferromagnetic with two centers T=0, 60×60

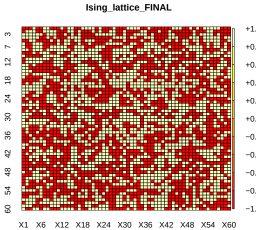


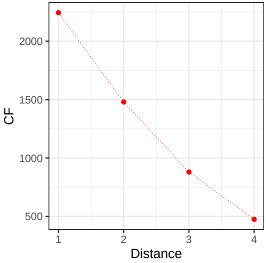


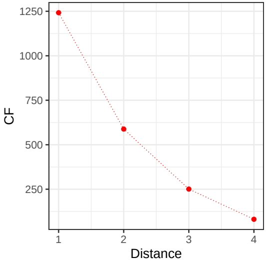
Ferromagnetic with two centers $T > 0$, 60×60

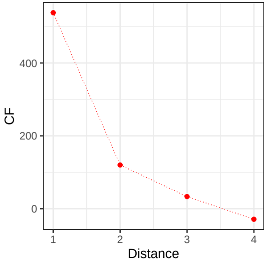




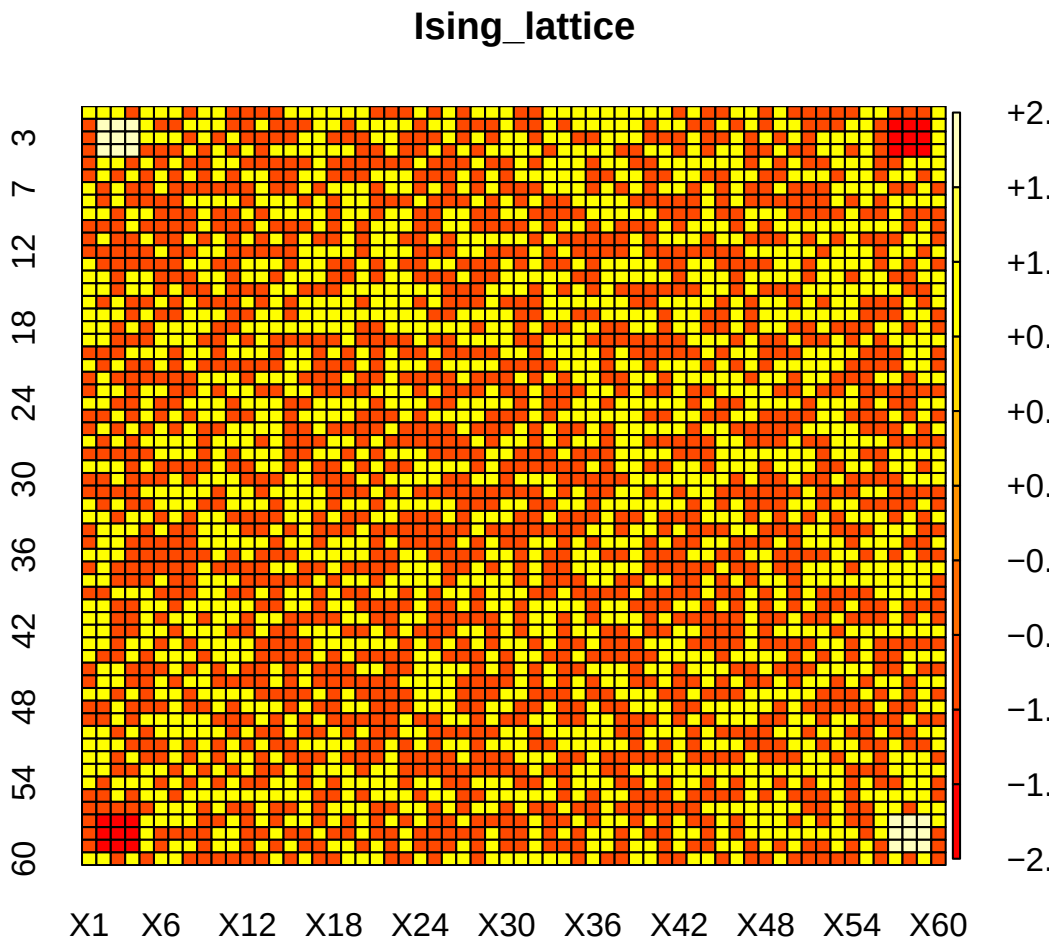




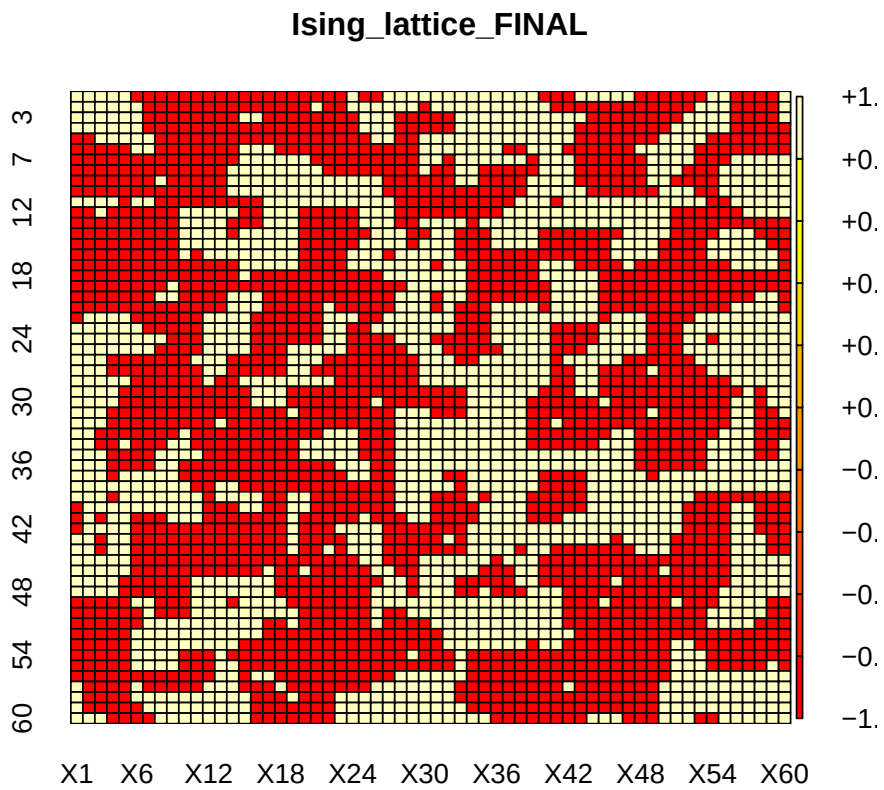


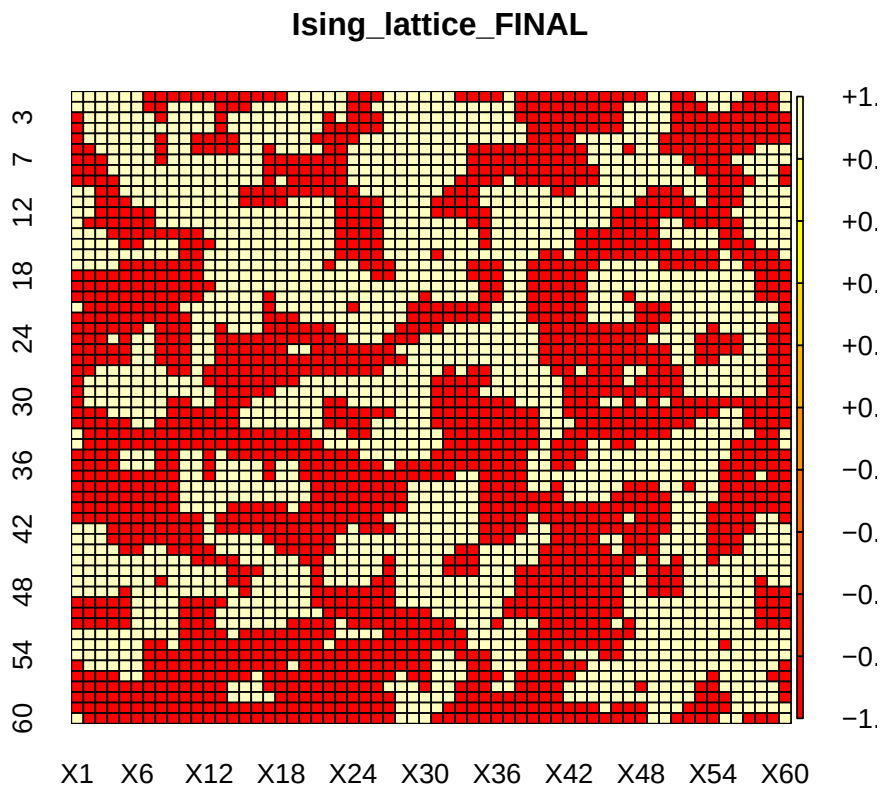


Ferromagnetic with four centers T=0, 60×60



Ferromagnetic with four centers T=0, 60×60





Ferromagnetic - conclusion

- The diffusion centrers avoid that the lattice flip in to an homogeneous state (all $+1$ or -1)
- With the diffusion center, al long as large lattice are considered (> 10) complex border/patterns comes out.
- In large lattices, the complexity of these patterns (which can be quantified with the correlation functions) vanishes as the temperature is increased
- **OUTLOOK**: perform different simulations for large lattices with 0 or low temperature in order to better exploits the ground/trap states. However for this purpose is better to use or C++ or Fortran since they have better performances in terms of computational cost

Bibliography I

[1] https://commons.wikimedia.org/wiki.

[2] Daan Frenkel and Berend Smit. *Understanding molecular simulation: from algorithms to applications*. Vol. 1. Elsevier, 2001.

[3] *Jeffrey C. Grossman and Elif Ertekin, NSE C242 Phys C203, Spring 2008, U.C. Berkeley, howpublished = https://nanohub.org/resources/4579/download/05-01-08-forweb.pdf,*

[4] David JC MacKay and David JC Mac Kay. *Information theory, inference and learning algorithms*. Cambridge university press, 2003.

[5] Kevin P Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.

[6] Luca Peliti. *Statistical mechanics in a nutshell*. Princeton University Press, 2011.

Connections with log-linear models and multivariate Gaussian [5]

- The joint probability function $p(\mathbf{y})$ can be represented with an undirected graph (Markov Random Field) on the basis of the Hammersley-Clifford Theorem (unlike the Bayesian network for the MRF no topological ordering is possible)
- **HK theorem** A positive distribution $p(y)$ satisfies the CI properties of an undirected graph G iif p can be represeted as a product of factors, one per maximal clique $p(y|\theta) = (1/Z)\prod_{c\in C}\phi_c(y_c|\theta_c)$ where $Z = \sum_{\mathbf{y}}\prod\phi_{c\in C}(\mathbf{y}_c|\theta_c)$
- **Gibbs distribution** $p(y|\theta) = (1/Z(\theta))\exp\left(-\sum_c E(y_c|\theta_c)\right)$; therefore $\phi_c(y_c|\theta_c) = \exp(-E(y_c|\theta_c))$
- For discrete variables one can use the maximum entropy or a log-linear model $\log p(y|\theta) = \sum\phi_c(y_c)^T\theta_c - \log Z(\theta)$

Connections with log-linear models and multivariate Gaussian [5]

- In this context the Ising model can be obtained with the following clique potential

$$\psi_{st}(y_s, y_t) = \begin{pmatrix} e^{w_{st}} & e^{-w_{st}} \\ e^{-w_{st}} & e^{w_{st}} \end{pmatrix} \tag{6}$$

- The analogy between the Ising and Gaussian graphical model can be captured by considering

$$\log p(\mathbf{y}) = -\sum_{s \sim t} y_s w_{st} y_t = (-1/2)\mathbf{y}^T \mathbf{W}\mathbf{t} + \sum_s b_s y_s = (-1/2)\mathbf{y}^T \mathbf{W}\mathbf{y} + \mathbf{b}^T \mathbf{y}$$

- Setting $\boldsymbol{\Sigma}^{-1} = \mathbf{W}$, $\boldsymbol{\mu} = \boldsymbol{\Sigma}\mathbf{b}$,and $c = (1/2)\boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ we can get the form of a multivariate gaussian (unnormalized)

$$p(y) \sim \exp\left((-1/2)(\mathbf{y} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu}) + c\right)$$

Numerical simulation [6]

- **Molecular dynamics:** the equation of motion are solved numerically
  PROS: information of both the dynamical and static properties of the system are explored
- **Monte Carlo:** a fictitious evolution process of the system is solved in order to get the equilibrium distribution
  - Also the systems whose dynamics is not defined can be explored
  - A fictitious dynamics can be considered in order to reach the equilibrium faster

Concepts of statistical mechanics: entropy and temperature [2]

- Consider a system with total energy E that is composed by two subsystems (with energy E_1 and E_2 that can exchange only energy - canonical ensemble)
- There are many way in which the energy can distribute in the two systems with the constrain $E = E_1 + E_2$
- In particular given the energy E_1 the total number of degenerate states is $\Omega_1(E_1) \times \Omega_1(E_2)$
- We would have a measure of the degeneracy of the system that is additive, therefore

$$\ln \Omega(E_1, E - E_1) = \ln \Omega_1(E_1) + \ln \Omega_2(E - E_1)$$

- Every energy state of the total system is equal likely. Therefore the most probable state is the one that maximizes $\ln \Omega(E_1, E - E_1)$

$$\left(\frac{\partial \ln \Omega(E_1, E - E_1)}{\partial E_1} \right)_{N,V,E} = 0$$

Concepts of statistical mechanics: entropy and temperature [2]

- Given that the system 1 is in the state E_1 the second system have energy $E - E_i$. Therefore the degeneracy of the second system is equal to $\Omega(E - E_i)$
- Thus the probability of 1 to be in the state i will be

$$P_i = \frac{\Omega(E - E_i)}{\sum_j \Omega(E - E_j)}$$

- Expanding around $E_i = 0$

$$\ln \Omega_B(E - E_i) = \ln \Omega_B(E) - E_i \frac{\partial \log \Omega_B(E)}{\partial E} + O(1/E)$$

$$\ln \Omega_B(E - E_i) = \ln \Omega_B(E) - \frac{E_i}{k_b T} + O(1/E)$$

$$P_i = \frac{\exp(-E_i/k_b T)}{\sum_j \exp(-E_j/k_b T)}$$

Monte Carlo method: Metropolis idea [2]

- How the points are generated ? With a Boltzmann weighted Markov chain
- $\pi(old \to new) = \alpha(old \to new) \times acc(old \to new)$ where π is the transition probability element from the old state to the new state,α is the matrix element of Markov Chain and acc is the acceptance ratio.
- Detailed balance condition at the equilibrium
  $N(old)\pi(old \to new) = N(new)\pi(new \to old)$
- With a symmetric Markov transition matrix we have
  $N(old) \times acc(old \to new) = N(new) \times acc(new \to old)$
- Therefore we have

$$\frac{acc(old \to new)}{acc(new \to old)} = \frac{N(n)}{N(o)} = \exp\left[-\beta\left(E(new) - E(old)\right)\right] \qquad (7)$$

- THE Z TERM IS NO MORE PRESENT ! We have only the difference between the two energies !!!

Monte Carlo method: Metropolis idea [2]

$$acc(old \to new) = \begin{cases} N(new)/N(old) & N(new) < N(old) \\ 1 & N(new) \geq N(old) \end{cases}$$

Therefore the overall transition probabilities are given by

$$\pi(old \to new) = \begin{cases} \alpha(old \to new) & N(new) \geq N(old) \\ \alpha(old \to new)\,[N(new)/N(old)] & N(new) < N(old) \end{cases}$$

$$\pi(old \to new) = 1 - \sum_{new \neq old} \pi(old \to new)$$

- In practice for each move a random number is generated from the uniform distribution between the interval $[0, 1]$, since $acc(old \to new) = \exp\left[-\beta\left(E(new) - E(old)\right)\right] < 1$. The move is accepted if the random number is lower than $acc(old \to new)$
- $\pi(old \to new)$ should be ergodic