

# Ising

April, 2020

Jose M. Camacho

## INTRODUCCIÓN A LOS FENÓMENOS CRÍTICOS

---

A escalas macroscópicas los materiales en equilibrio termodinámico presentan unas propiedades uniformes que determinan su estado. A las regiones dentro del espacio determinado por tales propiedades físicas se les llama **fase**. Los sistemas pueden cambiar abruptamente su comportamiento macroscópico incluso cuando los factores externos que determinan su estado (como la temperatura o la presión) cambian de forma suave y progresiva. A esos cambios en el comportamiento macroscópico del sistema se les llama **cambios de fase**. Las propiedades que cambian durante las transiciones de fase pueden hacerlo de manera continua (**transición de fase continua o de segundo orden**) o de manera discontinua o abrupta (**transiciones de fase discontinua**). El ejemplo más común de cambios de fase son las transiciones del agua que observamos a diario: hielo↔agua líquida↔vapor de agua (otros como la sublimación o la sublimación inversa también son posibles, si bien son más difíciles de observar en la cotidianidad diaria). A los puntos en los que ocurren los cambios de fase se les llama **puntos críticos**[1].

Cuando el sistema está lo suficientemente cerca del punto crítico la mayoría de las cantidades de interés exhiben un comportamiento de **ley de potencias**. Los exponentes que rigen esas leyes de potencias reciben el nombre **exponentes críticos** [1, 2]. Estos exponentes no dependen de los detalles de la física del sistema, sino de algunas de sus características generales como la dimensión o la distancia de correlación, de hecho los exponentes críticos de sistemas muy diferentes son iguales. Este fenómeno es llamado **universalidad** [2]. Podemos asignar cada sistema a una **clase de universalidad**, en el sentido de que dos sistemas pertenecen misma clase si tienen la misma dimensión  $d$  y parámetros de orden de la misma dimensional [2].

A continuación vamos a estudiar el modelo de Ising, el cuál es uno de los modelos más estudiados en la teoría de las transiciones de fase. Este modelo fue inicialmente propuesto para la descripción del ferromagnetismo, sin embargo, puede interpretarse en una gran cantidad de contextos, lo que lo convierte en un modelo con capacidad descriptiva para un amplio espectro de sistemas físicos. Por ejemplo, redes neuronales, procesos económicos, plegamiento de proteínas, o comportamiento social. Finalmente, el interés en el modelo de Ising también radica en que es considerado el modelo prototipo con una transición de fase continua no trivial [3].

## El Modelo de Ising

El hamiltoniano que rige la dinámica del modelo de ising viene dado por:

$$H = - \sum_{\langle i,j \rangle} J_{ij} s_i s_j - \sum_i h s_i$$

En el contexto de un material ferromagnético las magnitudes que en el Hamiltoniano aparecen se interpretan como sigue.  $s_i = \pm 1$  representa el spin de la partícula  $i$ ,  $J_{ij}$  es la constante de acoplamiento entre los spines  $s_i$  y  $s_j$ , esto es, determina, como interactúan los spines vecinos. En principio todos interactuarán de igual manera dos a dos, luego:  $J_{ij} = J = cte$ , además, para el caso del material ferromagnético,  $J > 0$ .  $\langle i, j \rangle$  indica que la suma solo corre sobre los vecinos próximos de la red (normalmente el modelo de Ising suele definirse sobre una red regular.), lo cual supone que los spines solo interactúan en corto alcance. Finalmente  $h$  es un campo magnético externo.

$$P(\sigma) = \frac{1}{Z} e^{-\beta H(\sigma)}$$

La interpretación de esta función es la siguiente:  $P(\sigma)$  proporciona la probabilidad de encontrar al sistema en un estado en equilibrio con configuración  $\sigma$  dada. La constante de normalización  $Z$  es llamada **función de partición**:

$$Z = \int d\sigma e^{-\beta H(\sigma)}$$

Y contiene las propiedades estadísticas del sistema en equilibrio termodinámico.

## Aproximaciones de Campo Medio

El campo medio es una familia de aproximaciones que consisten en aproximar todas las interacciones experimentadas por cada uno de los elementos del sistema por una interacción media o efectiva[4]. La potencia del campo medio reside en su sencillez, permitiéndonos resolución exacta de muchos problemas, y proporcionando resultados cualitativamente correctos en la mayoría de los casos.

Estas aproximaciones son útiles cuando las fluctuaciones no son importantes. Asimismo, la dimensionalidad del sistema juega un papel importante acerca de la cuestión de si el campo medio funcionará o no. Para dimensiones altas ocurrirá que cada spin tendrá un gran número de vecinos, por lo que las interacciones entre las partículas se promediarán, siendo correcto reemplazar esas interacciones por una sola interacción efectiva. Tendremos así un campo de interacción que puede ser escogido como la media de las interacciones de una partícula con sus vecinas. A continuación se presenta la aproximación de campo medio de Bragg-Williams para el modelo de ferromagnético de Ising.

Bien, en un sistema magnético como el del modelo de Ising podemos esperar que el momento magnético total sea proporcional al número total de spines abajo y arriba del sistema:

$$N = N_+ + N_-$$

$$m = \frac{1}{N} \sum_{i=1}^N s_i = \frac{N_+ - N_-}{N} = \frac{2 \cdot N_+ - N}{N}$$

Por definición la entropía de Boltzmann es igual al logaritmo del número de estados posibles:

$$S = \ln \binom{N}{N_+} = \ln \binom{N}{N(m+1) \cdot 0.5}$$

Después de simplificar usando las propiedades de los logaritmos y usando la aproximación de Stirling:

$$S = \ln(2) - \frac{1}{2} \cdot (1+m) \cdot \ln(1+m) - \frac{1}{2} \cdot (1-m) \cdot \ln(1-m)$$

En ausencia de un campo magnético recordemos que el hamiltoniano de Ising viene dado por:

$$H = -J \sum_{\langle i,j \rangle} s_i s_j$$

El campo local actuando sobre cada spin es:  $H_i = J \sum_{\langle j \rangle} s_j$ , donde  $\langle j \rangle$  indica suma sobre los vecinos más próximos de  $i$ . Ahora bien, usando la aproximación de campo medio procedemos a cambiar el valor del campo de los spines vecinos de  $i$  por el campo medio de todos los spines:  $m = \frac{1}{N} \sum_k s_k$ . De esta forma el campo sufrido por el spin  $i$  es:

$$H_i = Jmz$$

donde  $z$  es el número de coordinación de la red, esto es, el número de vecinos de cada spin en la red. Procedemos a calcular ahora la energía media de la red:

$$E = \langle H \rangle = -J \sum_{\langle i,j \rangle} \langle s_i s_j \rangle$$

La aproximación de campo medio supone la decorrelación estadística de los spines, por lo que el correlador a dos cuerpos será cero:

$$C(s_i, s_j) = \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle = 0 \Leftrightarrow \langle s_i s_j \rangle = \langle s_i \rangle \langle s_j \rangle$$

por lo que:

$$E = \langle H \rangle = -J \sum_{\langle i,j \rangle} \langle s_i s_j \rangle = -J \sum_{\langle i,j \rangle} \langle s_i \rangle \langle s_j \rangle = -\frac{1}{2} J m^2 z N$$

donde el factor  $\frac{1}{2}$  se añade para no contar dos veces los spines vecinos.  
por tanto la energía por spin vendrá dada por:

$$\epsilon = \frac{\langle H \rangle}{N} = \frac{E}{N} = -\frac{1}{2} J m^2 z$$

### Fluctuaciones Canónicas de la Energía

Energía media:

$$E = \langle H \rangle = \frac{1}{Z} \sum_{\sigma} H(\sigma) e^{-\beta H(\sigma)} = \frac{\sum_{\sigma} H(\sigma) e^{-\beta H(\sigma)}}{\sum_{\sigma} e^{-\beta H(\sigma)}} = \frac{-1}{Z} \cdot \frac{\partial Z}{\partial \beta} = \frac{\partial \ln(Z)}{\partial \beta}$$

El calor específico del sistema puede derivarse a partir de la energía:

$$\begin{aligned} \frac{\partial E}{\partial \beta} &= -\frac{\partial^2 \ln(Z)}{\partial \beta^2} = -k_B T^2 C_v = -\frac{\partial}{\partial \beta} \left( \frac{\sum_{\sigma} H(\sigma) e^{-\beta H(\sigma)}}{\sum_{\sigma} e^{-\beta H(\sigma)}} \right) = -\frac{\partial}{\partial \beta} \left( \frac{f}{g} \right) = \\ &= \frac{f'g - g'f}{g^2} = -\frac{-\sum_{\sigma} H^2(\sigma) e^{-\beta H(\sigma)} Z + \sum_{\sigma} H(\sigma) e^{-\beta H(\sigma)} Z^2}{Z^2} = \\ &= \frac{\sum_{\sigma} H^2(\sigma) e^{-\beta H(\sigma)}}{Z} - \frac{\sum_{\sigma} H(\sigma) e^{-\beta H(\sigma)}}{Z^2} = \\ &= \langle H^2 \rangle - \langle H \rangle^2 = k_B T^2 C_v \end{aligned}$$

Finalmente:

$$C_v = \frac{\sigma^2[H]}{k_B T^2} = \frac{1}{k_B T^2} (\langle H^2 \rangle - \langle H \rangle^2) \quad \text{-Ecuación de Einstein-}$$

Si queremos el calor específico por spin solo hay que dividir por  $N$ :

$$c_v = \frac{\sigma^2[H]}{k_B T^2 N} = \frac{1}{k_B T^2 N} (\langle H^2 \rangle - \langle H \rangle^2)$$

### Susceptibilidad Magnética

La susceptibilidad magnética es una medida de como cambia la magnetización de un sistema tras producirse un cambio en el campo magnético externo.

$$\chi_T = \frac{N}{T k_B} (\langle m^2 \rangle - \langle m \rangle^2)$$

## Cumulante de Binder

En el campo de las transiciones de fase el cumulante de cuarto orden del parámetro de orden (en el caso del modelo de Ising, el momento magnético  $m$ ) juega un papel importante, puesto que permite identificar calcular el exponente crítico de la distancia de correlación y de ahí identificar la clase de universalidad de la transición[5] . Este cumulante recibe el nombre de cumulante de Binder y viene definido por:

$$U_4 = 1 - \frac{\langle m^4 \rangle}{3\langle m^2 \rangle^2}$$

## Clase de Universalidad del modelo de Ising

Algunos de los sistemas físicos más importantes que pertenecen a la clase de universalidad del modelo de Ising: (I) Transiciones líquido-vapor, (II) Mezclas binarias, (III) Sistemas Coulombianos, (IV) Sistemas Micelares, (V) Sistemas Magnéticos uniaxiales[3].

Los valores de los exponentes críticos que caracterizan a estos sistemas vienen dados por:

Parameter	d = 2	d = 3	d = 4
$\alpha$	0	0.11008	0
$\beta$	$\frac{1}{8}$	0.326419	$\frac{1}{2}$
$\gamma$	$\frac{7}{4}$	1.237075	1
$\delta$	15	4.78984	3
$\eta$	$\frac{1}{4}$	0.036298	0

Table 1: Clase de Universalidad del Modelo de Ising

## MÉTODOS MONTE CARLO

Los Métodos Monte Carlo son aquellos métodos numéricos que emplean el uso de números aleatorios para la resolución de un determinado problema. Su principal ventaja es su flexibilidad para resolver diferentes problemas [3]. La expansión práctica de estos métodos ha sido posible gracias a la invención de los ordenadores, si bien ya existían algunas aplicaciones antes de la era de los ordenadores [6].

Historicamente el campo de las transiciones de fase ha hecho un amplio uso de los métodos Monte Carlo, de hecho, erróneamente a veces se piensa que es su único campo de aplicación [7]. El uso que le vamos a dar a los métodos MC en este trabajo es el de samplear una distribución de probabilidad arbitraria  $f_X(x)$ , esto es, para generar muestras de números aleatorios que sigan esa distribución de probabilidad.

De los cursos de estadística sabemos que un método general para la producción de números aleatorios no correlacionados que sigan cualquier distribución de probabilidad es el de la transformada inversa. Este método emplea la generación de números aleatorios uniformemente distribuidos en el intervalo  $(0, 1)$ :  $u \sim U(0, 1)$  para producir una muestra que siga una distribución de probabilidad arbitraria  $f_X(x)$ . Para empezar sea  $X$  una variable aleatoria con función de probabilidad acumulada  $F_X(x)$ , que admite inversa:  $u = F_X(x) \in [0, 1]$  entonces la variable transformada  $u$  sigue una uniforme  $u \sim [0, 1]$ . Este resultado se usa para generar una muestra aleatoria siguiendo cualquier distribución de probabilidad:  $u = F_X(x) \Leftrightarrow x = F_X^{-1}(u)$ .

El problema del método anterior, aunque sencillo y general, es que requiere de una buena implementación de  $F_X^{-1}(x)$ [7], lo cual, en general, es difícil de conocer. Se hacen necesarios otros métodos para el muestreo de distribuciones de probabilidad arbitrarias, tan necesario para la simulación de ciertos sistemas de naturaleza estocástica, como los modelados con el modelo de Ising. Veamos un método sencillo para la generación de números aleatorios siguiendo una distribución arbitraria  $f_X(x)$ : el método del rechazo.

## Método del Rechazo

Este método permite generar muestras de variable aleatoria arbitraria  $\mathbf{X} \sim f_{\mathbf{X}}(\mathbf{x}), \forall \mathbf{x} \in \mathbb{R}^p$  a partir de una variable aleatoria propuesta  $Y \sim g_Y(\mathbf{x}), \forall \mathbf{x} \in \mathbb{R}^p$  (fácil de generar). La idea es generar un valor de  $Y$  que se aceptará con probabilidad  $h(x) = \frac{f_{\mathbf{X}}(\mathbf{x})}{R \cdot g_Y(\mathbf{x})}$ , si el valor no es aceptado el proceso se repite hasta que se genere un valor de  $Y$  que sea aceptado. Así se consigue una variable aleatoria resultante que siga  $\sim f_{\mathbf{X}}(\mathbf{x})$ . Cualitativamente podemos razonar que aquellos valores que proporcionen los valores más grandes  $f_{\mathbf{X}}(\mathbf{x})$  se aceptarán con mayor probabilidad que los valores que hacen  $f_{\mathbf{X}}(\mathbf{x})$  más pequeño [7]. El valor  $R$  es una constante con valor  $\geq 1$  que cumple  $f_{\mathbf{X}}(\mathbf{x}) \leq R \cdot g_Y(\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R}^p$  [8].

1. Genera  $\mathbf{Y} \sim g_Y(\mathbf{x})$
2. Genera  $u \sim (0, R \cdot g_Y(x))$
3. Comprueba si  $U \leq f_{\mathbf{X}}(\mathbf{x})$ , en ese caso  $\mathbf{X} = \mathbf{Y}$ , en caso contrario vuelve al paso 1.

El problema evidente de este método es que si la probabilidad de aceptación es muy baja el tiempo de resolución se hará excesivamente largo, haciéndolo inviable en algunos casos. Una mejora a este método es el **método del rechazo con repetición**, el cual consiste en repetir el último valor aceptado cada vez con nuevo valor es rechazado. Sin embargo, el problema que implica la repetición es que si la probabilidad de aceptación es muy baja entonces tendremos muestras con largas cadenas de números repetidos.

## Métodos Monte Carlo Dinámicos

Los métodos MC dinámicos son métodos MC que permiten la simulación de sistemas con un conjunto de tasas de transición constantes y conocidas (también reciben el nombre de Métodos Monte Carlo Cinéticos, depende del área de aplicación) [9].

En este caso vamos a usar utilizar los MC dinámicos para hacer un *sampling* una distribución de probabilidad usando reglas sencillas, incluso cuando no sabemos que forma tiene la distribución [7]. La idea principal consiste en utilizar la información previa del sistema para las probabilidades de proposición y aceptación de los valores [7]. Sin embargo, como veremos, el problema que conllevan es que producen valores correlacionados para las variables aleatorias, lo cual puede conllevar una importante fuente de error [7].

El ejemplo más sencillo de un método MC dinámico para el muestro de una distribución de probabilidad  $f_X(x)$  es el método del rechazo con repetición antes mencionado. La pregunta evidente de este método es: Si se repiten los valores  $x_{k-1}$  y  $x_k$  ¿Podemos decir que los valores  $x_k$  siguen la distribución  $f_{X_k}(x_k)$ ? [7]

## Rechazo con repetición

El valor  $x_k$  se puede aceptar o rechazar. Podemos, entonces, ver el proceso de aceptación como una variable Bernoulli  $B$  con interpretación:  $b = 1$  aceptación,  $b = 0$  rechazo. De modo que podemos escribir la función de distribución de probabilidad como:

$$\begin{aligned} f_{X_k}(x_k) &= f_{X_k}(x|b=1)p(b=1) + f_{X_k}(x_k|b=0)p(b=0) \\ f_{X_k}(x_k) &= f_{X_k}(x|b=1)p(b=1) + f_{X_k}(x_k|b=0)(1-p(b=1)) \end{aligned}$$

Usando ahora el teorema de Bayes para la probabilidad conjunta de una variable discreta  $C$  y otra continua  $A$ :

$$\begin{aligned} q_{A,C} &= q(c|A=a)q_A(a) \\ q_{C,A} &= q(a|C=c)q_C(c) \end{aligned}$$

De donde podemos escribir:

$$q_C(c|A = a) = \frac{q_C(c)q_A(a|C = c)}{q_A(a)}$$

Donde  $A \rightarrow X_k$ ,  $C \rightarrow B$ ,  $q_C(c) \rightarrow p(b)$ ,  $q_A(a) \rightarrow g_{X_k}(x_k)$ ,  $q_C(c|A = a) \rightarrow \text{Prob}(b|x_k)$ ,  $q_A(a|C = c) \rightarrow f_{X_k}(x_k|b = 1)$ :

$$\text{Prob}(b|x_k) = \frac{p(b)f_{X_k}(x_k|b)}{g_X(x)}$$

Sustituimos en la expresión de  $f_{X_k}(x_k)$ :

$$f_{X_k}(x_k) = \text{Prob}(b = 1|x_k)g_{X_k}(x_k) + f_{X_k}(x_k|b = 0)(1 - p(b = 1))$$

Donde  $g_{X_k}(x_k)$  es la probabilidad de proposición de un valor  $x_k$  de la variable aleatoria  $X_k$ , y  $\text{Prob}(b = 1|x_k)$  es la probabilidad de aceptación de ese valor, la cual empezaremos a llamar  $h_{X_k}(x_k) \equiv \text{Prob}(b = 1|x_k)$ . Ahora bien, si el valor  $x_k$  es rechazado nos quedaremos con el valor anterior,  $x_{k-1}$ ,  $f_{X_k}(x_k|b = 0) = f_{X_{k-1}}(x_{k-1})$ . En cuando a  $p(b = 1)$  puede calcularse como la media de la probabilidad de aceptación (esto lo sabemos porque  $B$  es una variable de Bernoulli):

$$p(b = 1) = \int_{-\infty}^{\infty} dx_k h_{X_k}(x_k) g_{X_k}(x_k)$$

Por tanto nos queda:

$$f_{X_k}(x_k) = h_{X_k}(x_k)g_{X_k}(x_k) + f_{X_{k-1}}(x_{k-1}) \left[ 1 - \int_{-\infty}^{\infty} dx_k h_{X_k}(x_k)g_{X_k}(x_k) \right]$$

Llamo  $\epsilon = \int_{-\infty}^{\infty} dx_k h_{X_k}(x_k)g_{X_k}(x_k)$ :

$$f_{X_k}(x_k) = h_{X_k}(x_k)g_{X_k}(x_k) + f_{X_{k-1}}(x_{k-1}) [1 - \epsilon]$$

Esta ecuación es una expresión recurrente en  $k$  que puede resolverse facilmente en términos de la distribución inicial  $f_{X_0}(x_0)$ :

$$f_{X_k}(x_k) = (1 - \epsilon)^k \left[ f_{X_0}(x_0) - \frac{h_{X_k}(x_k)g_{X_k}(x_k)}{\epsilon} \right] + \frac{h_{X_k}(x_k)g_{X_k}(x_k)}{\epsilon}$$

Finalmente podemos escribir esta relación de recurrencia en términos de la verdadera distribución  $f_X(x)$  que queremos samplear. Para ello usamos la definición de probabilidad condicionada:

$$f_X(x|b) = \frac{f_{X,B}(x, b = 1)}{\text{Prob}(B = 1)} = \frac{f_{X,B}(x, b = 1)}{\int_{-\infty}^{\infty} f_{X,B}(x, b = 1)dx}$$

$$h(x) = \text{Prob}(B = 1|x) = \frac{f_{X,B}(x|b = 1)}{f_{X,B}(x, 0) + f_{X,B}(x, 1)} = \frac{f_{X,B}(x, b = 1)}{g(x)}$$

de donde obtenemos:  $f_X(x) \equiv f_{X,B}(x, 1) = \frac{h_X(x)g_X(x)}{\int_{-\infty}^{\infty} h_X(x)g_X(x)dx}$

$$f_{X_k}(x_k) = (1 - \epsilon)^k [f_{X_0} - f_X(x)] + f_X(x)$$

Podemos decir en tonces que, en general, los valores  $x_k$  no están distribuidos según  $f_X(x)$ . Para conseguir que  $X_k \sim f_X(x)$  tenemos dos opciones que anulan el término  $(1 - \epsilon)^k [f_{X_0} - f_X(x)]$ . La primera es que los valores iniciales ya estén distribuidos según  $f_X(x)$ , esto es,  $f_X(x) = f_{X_0}(x_0)$ . La segunda opción es que el término  $(1 - \epsilon)^k$  sea nulo. Puesto que  $0 < \epsilon \leq 1$  ocurre que  $(1 - \epsilon)^k$  decrece monotonamente con  $k$ , de manera que esperamos que  $\lim_{k \rightarrow \infty} f_{X_k}(x_k) = f_X(x)$ . Así que conviene esperar un cierto número de pasos del algoritmo antes de empezar a almacenar los valores de la variable aleatoria  $X$ . Cuanto esperar dependerá del valor de  $\epsilon$  de cada problema. A esta espera se le conoce como **termalización**.

## Métodos MC dinámicos: planteamiento general

Entendido el método del rechazo con repetición, podemos decir que los métodos MC dinámicos son una generalización de ese método cuando el nuevo valor depende de los anteriores valores aceptados, pero sin repetir el valor. Veamos, queremos generar un valor  $x_n$  de una variable aleatoria  $X_n$  en la  $n$ -ésima iteración y ésta depende de los números previos generados en iteraciones pasadas:

$$f_{X_n}(x_n) = f_{X_n}(x_n | x_0, \dots, x_{n-1}) \quad (1)$$

La forma más sencilla de implementar la dependencia con la historia del sistema es suponer que el nuevo paso solo depende del paso anterior  $f_{X_n}(x_n | x_0, \dots, x_{n-1}) = f_{X_n}(x_n | x_{n-1})$ . Este tipo de procesos que solo utilizan la información del estado inmediatamente anterior del sistema recibe el nombre de **procesos de Markov**. En general podemos decir que proceso de Markov es un método para la generación de nuevas configuraciones de un sistema a partir de la configuración actual del sistema. La ventaja de estos métodos es precisamente que no requieren del conocimiento de toda la historia previa del sistema, sino solo de la configuración inmediatamente anterior [2]. Otra simplificación que se toma es que el proceso de Markov sea homogéneo, esto es, no depende de  $n$ :  $f_{X_n}(x_n | x_{n-1}) = f_X(x_n | x_{n-1})$ . Por simplicidad llamemos:  $x_n \equiv x$ ,  $x_{n-1} \equiv y$ . Por tanto, ahora las probabilidades de aceptación  $h$  y de proposición  $g$  dependen del valor anterior  $y$ :

$$h = h(x|y) \quad g = g(x|y)$$

Asimismo, la probabilidad de aceptación de un valor  $x$  dependerá del valor del paso anterior  $y$ :

$$P(b = 1|y) = \int h(z|y)g(z|y)dz \equiv \epsilon(y)$$

Para la proposición de las funciones  $h$  y  $g$  la única condición que necesitamos es que la distribución estacionaria de la cadena de markov  $f_{X_n}(x_n)$  sea la distribución de probabilidad que queremos samplear:  $\lim_{n \rightarrow \infty} f_{X_n}(x_n) = f_X(x)$ , para lo cual debe cumplirse la condición de **balance detallado**:

$$h(x|y)g(y|x)f_X(x) = h(x|y)g(x|y)f_X(x)$$

Para ver que esta condición es necesaria y suficiente tenemos que usar la siguiente expresión de la probabilidad condicionada:

$$f_{X_{n+1}}(x) = \int_{-\infty}^{\infty} f(x|y)f_{X_n}(y)$$

En caso de que  $f_{X_n} = f_{X_{n+1}} \equiv f_X^{st}$ :

$$f_X^{st}(x) = \int_{-\infty}^{\infty} f(x|y)f_X^{st}(y)$$

usando que:  $\int_{-\infty}^{\infty} f(y|x)dy = 1$ :

$$\int_{-\infty}^{\infty} f(y|x)f_X^{st}(x) = \int_{-\infty}^{\infty} f(x|y)f_X^{st}(y) \Leftrightarrow f(x|y)f_X^{st}(x) = f(y|x)f_X^{st}(y)$$

La última expresión es la condición de balance detallado.

Esta es la ecuación que deben satisfacer las probabilidades de proposición  $g(x|y)$  y de aceptación  $h(x|y)$  para que la distribución estacionaria de la cadena de markov  $f_{X_n}(x_n)$  sea la distribución de probabilidad que queremos samplear. Los diferentes métodos dinámicos dan soluciones diferentes para esta ecuación. A continuación veremos una de esas soluciones, la dada por el **Algoritmo de Metrópolis**, uno de los MC más usados en la simulación de procesos de Markov [2].

## Algoritmo de Metrópolis

En este método se fija la distribución de probabilidad de proposición  $g(x|y)$  y se resuelve para la probabilidad de aceptación  $h(x|y)$ . En otras palabras, dada  $g(x|y)$  debemos encontrar una función  $h(x|y)$  que cumpla que:

$$0 \leq h(x|y) \leq 1$$

$$h(y|x)g(y|x)f_X(x) = h(x|y)g(x|y)f_X(y)$$

Para encontrar las soluciones de esta última ecuación introducimos las siguientes funciones:

$$q(x|y) = \frac{g(y|x)f_X(x)}{g(x|y)f_X(y)} \quad \text{cumpliendo} \quad q(y|x) = \frac{1}{q(x|y)}$$

$$h(x|y) = \sqrt{q(x|y)}w(x, y)$$

La función  $w$  debe cumplir  $w(x, y) = w(y, x)$ , puesto que:

$$h(y|x)g(y|x)f_X(x) = h(x|y)g(x|y)f_X(y) \Leftrightarrow q(x|y) = \frac{h(x|y)}{h(y|x)} \Leftrightarrow$$

$$\Leftrightarrow q(x|y) = \frac{h(x|y)}{h(y|x)} = \frac{\sqrt{q(x|y)}w(x, y)}{\sqrt{q(y|x)}w(y, x)}$$

Usando  $q(y|x) = \frac{1}{q(x|y)}$ :

$$q(x|y) = \frac{1}{q(y|x)} \cdot \frac{w(x, y)}{w(y, x)}$$

de donde se ve que es necesario que  $w(x, y) = w(y, x)$ . Si establecemos que la dependencia de  $w$  con  $x$  e  $y$  a través de  $q$ :

$$\text{Si } w(x, y) = w(q), \text{ entonces } w(x, y) = w(y, x) \Leftrightarrow w(q) = w\left(\frac{1}{q}\right)$$

Ahora bien, en términos de  $w$  la condición  $0 \leq h(x|y) \leq 1$  es equivalente a:

$$0 \leq \sqrt{q}w(q) \leq 1 \Leftrightarrow w(q) \leq \frac{1}{\sqrt{q}}$$

Estamos buscando entonces una función  $w$  que cumpla:

$$w(q) = w\left(\frac{1}{q}\right) \quad w(q) \leq \frac{1}{\sqrt{q}}$$

La solución del algoritmo de metrópolis es  $w(q) = \min\left\{\sqrt{q}, \frac{1}{\sqrt{q}}\right\}$ , de manera que si  $q \leq 1$  entonces  $w(q) = \sqrt{q}$ , y si  $q > 1$  entonces  $w(q) = \frac{1}{\sqrt{q}}$ . Esta elección de  $w$  conlleva la siguiente forma para la probabilidad de aceptación:

$$h(x|y) = \sqrt{q(x|y)}w(x, y) \Leftrightarrow h(x|y) = \min\{1, q(x|y)\}$$

Para el modelo de Ising diremos, en términos físicos, que, para cada cambio en la configuración del sistema desde un estado  $\alpha$  a un estado  $\alpha'$ , se calcula el cambio de energía producido  $\Delta E_{\alpha \rightarrow \alpha'} = E_{\alpha'} - E_{\alpha}$  [2]. Si  $\Delta E_{\alpha \rightarrow \alpha'} < 0$  el cambio en la configuración del sistema se acepta, pero si  $\Delta E_{\alpha \rightarrow \alpha'} > 0$ , entonces ese cambio se acepta con probabilidad  $e^{-\beta \Delta E_{\alpha \rightarrow \alpha'}}$ .

1. Genera una nueva configuración (cambia un spin  $i$  aleatoriamente)
2. Calcula el cambio en la energía del sistema  $\Delta E_{\alpha \rightarrow \alpha'}$
3. Calcula  $P(E_{\alpha'} \rightarrow E_{\alpha}) = \min\{1, e^{-\beta(E_{\alpha'} - E_{\alpha})}\}$

4. Genera  $u \sim [0, 1]$

5. Acepta el cambio en la configuración si  $u \leq P$

Los cambios en la energía solo puede tomar 5 valores en el caso de una red regular bidimensional con 4 vecinos más próximos:

$$\Delta E_{\alpha \rightarrow \alpha'} = -2 \cdot J \cdot s_{\alpha} \sum_{\mu=1}^z s_{\alpha_{\mu}}$$

Donde  $s_{\alpha} \sum_{\mu=1}^z s_{\alpha_{\mu}} = -4, -2, 0, +2, +4$

## RESULTADOS

Para la realización de las simulaciones se han llevado a cabo con :

- Pasos Monte Carlo para la termalización:  $N \cdot 3000$ , donde  $N = L \cdot L$  es el número de spines de la red.
- Pasos Monte carlo para :  $10^6$ .
- Pasos Montecarlo dejados entre cada medición para evitar correlaciones: 10.
- Tamaños de red simulados:  $L = 20, 60, 100, 140$ .
- Rango de temperaturas recurrido para cada tamaño:  $T_{min} = 1.5, T_{max} = 3.5$  con pasos de  $\Delta T = 0.042$  (en unidades naturales para las cuales  $J = k_B = 1$ )

### Magnetización

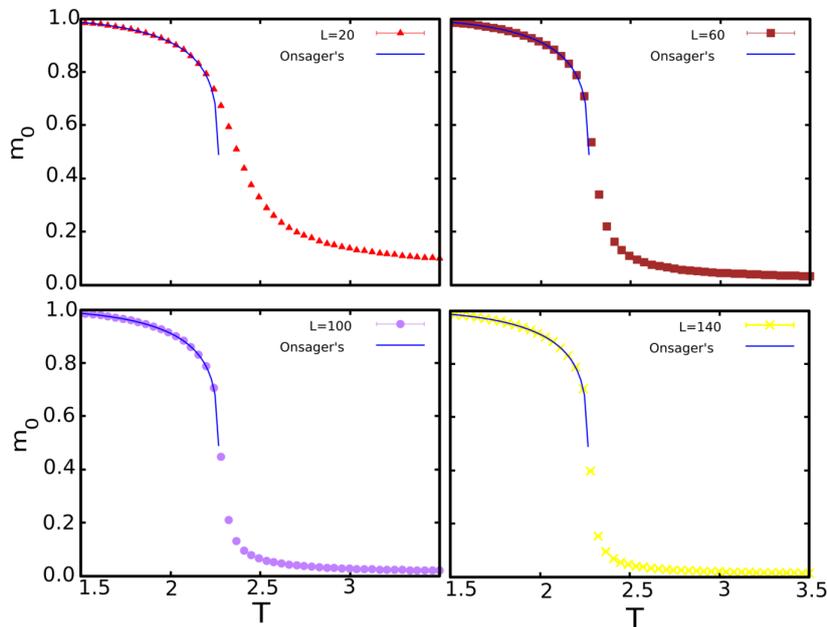


Figure 1: **Valor absoluto de la magnetización para diferentes tamaños de la red.** En el caso del modelo de Ising las propiedades microscópicas del sistema pueden resumirse en la posición de los spines de cada partícula (“up” or “down”). Mientras que las propiedades microscópicas emergentes son la magnetización  $m$ , energía  $E$ , el calor específico  $C_V$  o la susceptibilidad magnética  $\chi_T$ . De entre ellas el parámetro de orden,(parámetro cuyas propiedades macroscópicas cambian en la transición de fase) es  $m$ , la cual puede cambiar espontáneamente de 0 a  $-1$  o  $+1$ , es decir, todos los spines “down” o “up” respectivamente. En esta simulación estamos tomando el valor absoluto de  $m$  para que el sistema siempre colapse al mismo estado y sea más fácil el posterior tratamiento de los resultados. El cambio en la magnetización de 0 a  $+1$  ocurre en  $T = T_c$ . Sin embargo, puesto que estamos simulando un sistema finito, la transición de fase no ocurre en la temperatura crítica. De hecho, para tamaños pequeños, como  $L = 20$ , el cambio en la magnetization ocurre de forma más suave y no puede estimarse exactamente para que valor de  $T$  se produce la transición. Conforme aumenta el tamaño de la red más acerca ocurre la transición de  $T_c$  y más se acercan los resultados a la solución teórica de Onsager.

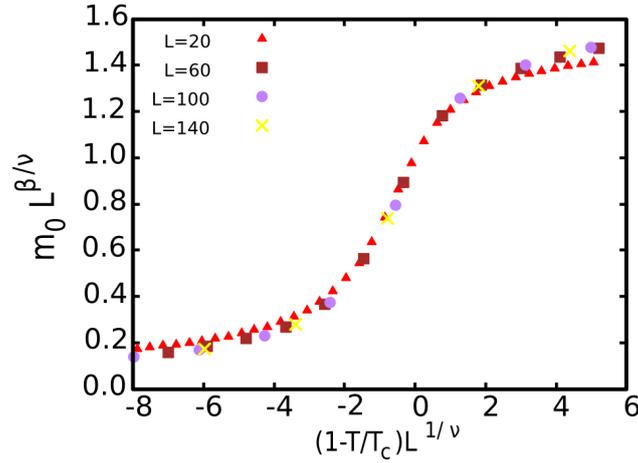


Figure 2: **Colapso de las curvas del valor absoluto de la magnetización.** para un sistema infinito la distancia de correlación,  $\xi(T)$ , (distancia a la cual spines en diferentes posiciones están correlacionados) se comporta como una ley de potencias  $\xi(T) \sim \left|1 - \frac{T}{T_c}\right|^{-\nu}$ . En el caso de un sistema finito, evidentemente,  $\xi \sim L$ , puesto que la distancia de correlación no puede exceder la dimensión del mismo. Si queremos obtener como se comportan a magnetización para tamaños finitos del sistema podemos tratar  $m$  como una función homogénea de  $\xi$  y  $L$ , de manera que se recupere la relación  $m \sim \left|1 - \frac{T}{T_c}\right|^\beta$  de un sistema infinito:  $m(\xi, L) = \xi^x \tilde{m}\left(\frac{\xi}{L}\right)$ . La relación de escala queda como:  $m = L^{-\frac{\beta}{\nu}} \tilde{m}\left[\left(1 - \frac{T}{T_c}\right) L^{\frac{1}{\nu}}\right]$ .

## Energía por Spin

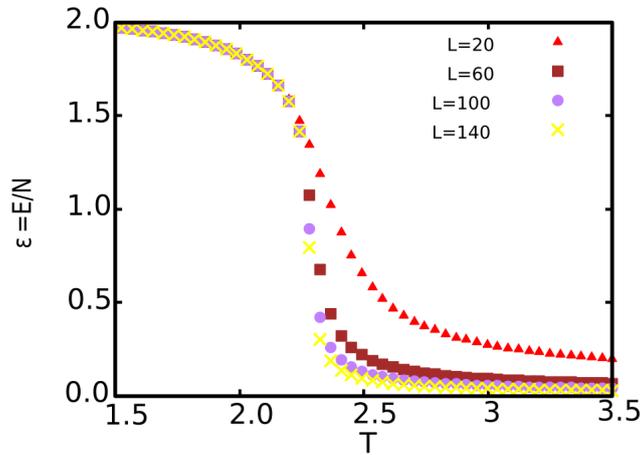


Figure 3: **Energía media por spin para diferentes valores del tamaño de la red.**  $\epsilon = \frac{E}{N} = -\frac{1}{2}Jm^2z$ , donde  $J = 1$ ,  $z = 4$ . El comportamiento de la energía por spin es similar al de la magnetización puesto que es la única magnitud de la que depende que no es constante. La energía por spin es cero cuando la magnetización es cero puesto que hay, en promedio, el mismo número de spines “up” que spines “down”. Cuando el sistema se acerca a la temperatura crítica se forman clusters de spines apuntando a una misma dirección, la magnetización promedio deja de ser cero, así como la energía media por spin, hasta que todos los spines del sistema apuntan a la misma dirección en la temperatura crítica. Los efectos de tamaño finito dan lugar a que el cambio de  $\epsilon$  por spin ocurre de forma más suave.

## Calor Específico

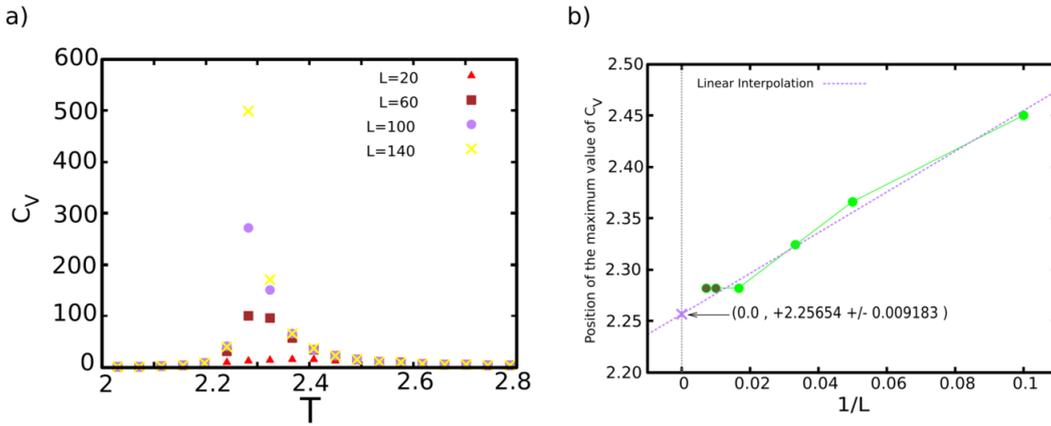


Figure 4: **a) Calor específico del sistema para diferentes tamaños de la red.** El calor específico diverge en  $T = T_c$ . Puesto que un ordenador no puede arrojar un valor infinito lo que obtenemos conforme nos acercamos a la temperatura crítica son valores de  $C_V$  mucho mayores que los obtenidos lejos de  $T_c$ . Este comportamiento divergente es más pronunciado cuanto mayor sea  $L$ . También es importante notar que, puesto que estamos simulando sistemas finitos, el valor máximo de  $C_V$  no se da exactamente en  $T_c$ . Este hecho puede usarse para estimar el valor de  $T_c$ , para ello basta dibujar el valor de la posición de máximo de  $C_V$  frente a  $\frac{1}{L}$  y extrapolar para  $L \rightarrow \infty$ . **b) Estimación de la temperatura crítica mediante la posición del máximo de  $C_V$ .** Puesto que la cantidad de temperaturas que podemos tomar es limitada el valor obtenido para las posiciones de los máximos vendrá afectada por la discretización que hayamos tomado (en este caso  $\Delta T = 0.042$  en unidades naturales de  $J$ ). Los puntos en negro se han omitido de la interpolación porque ya se ha llegado a la máxima aproximación posible de  $T_c$  para la discretización tomada para la temperatura (en caso contrario tendríamos tres puntos con la misma ordenada).

## Susceptibilidad Magnética

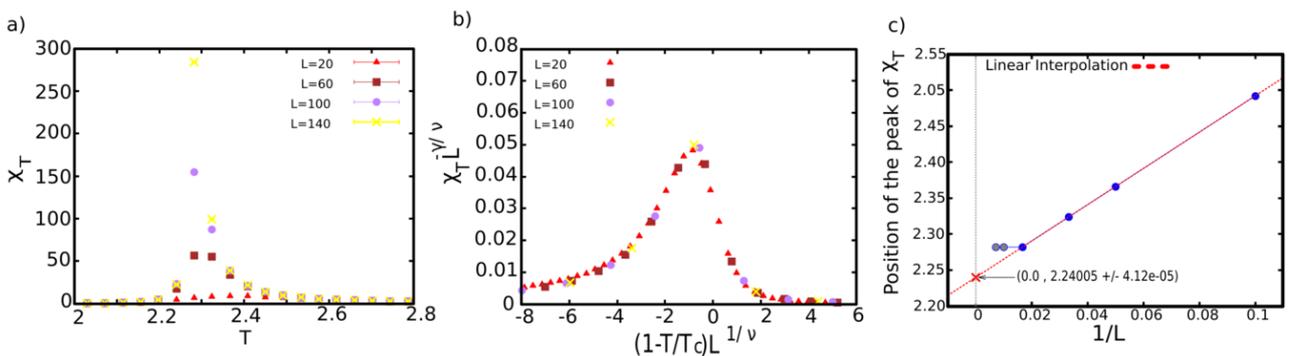


Figure 5: **a) Susceptibilidad magnética para diferentes valores del tamaño de la red.** Al igual que ocurría con el calor específico, la susceptibilidad magnética  $\chi_T$  diverge en  $T = T_c$ . Asimismo, como hemos visto para  $C_V$ , el valor máximo de  $\chi_T$  no se da exactamente en  $T_c$ , ya que estamos simulando sistemas finitos. Este hecho puede usarse para estimar el valor de  $T_c$ , para ello basta dibujar el valor de la posición de máximo de  $\chi_T$  frente a  $\frac{1}{L}$  y extrapolar para  $L \rightarrow \infty$ . **b) Colapso de las curvas de magnetización para diferentes tamaños de la red.** Para la susceptibilidad magnética la relación de escala viene dada por  $\chi_T(T, L) = L^{\frac{1}{\nu}} \tilde{\chi} \left[ \left(1 - \frac{T}{T_c}\right) L^{\frac{1}{\nu}} \right]$ . **c) Estimación de la temperatura crítica mediante la posición del máximo de  $\chi_T$**

## Cumulante de Binder

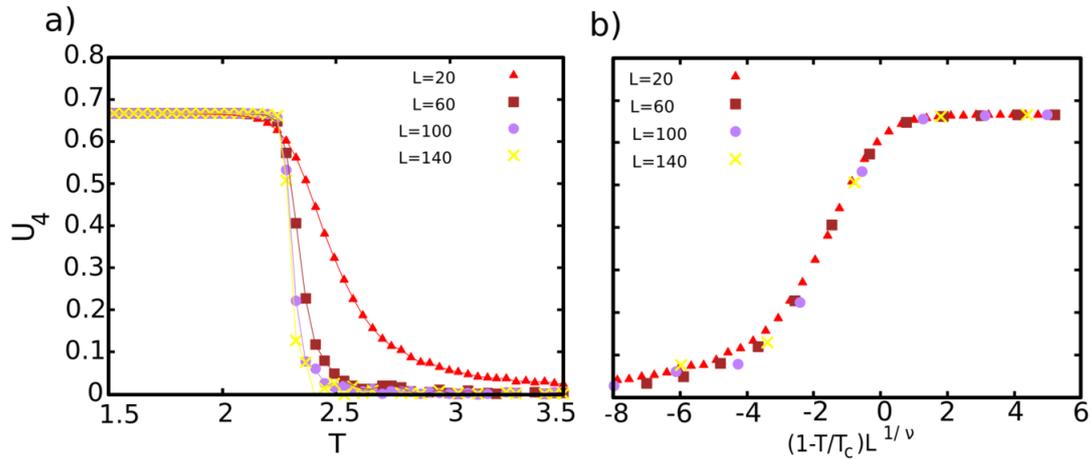


Figure 6: a) **Cumulante de Binder para diferentes tamaños de la red.** Como el cumulado de binder es constante para  $T = T_c$ ,  $U_4(T_c, L) = U_4(0, L) = cte \forall L$ , podemos determinar  $T_c$  como el punto de corte entre las gráficas para cada  $L$ . De nuevo, a la hora de determinar dicho punto de corte nos vemos limitados por la cantidad de valores de temperatura tomados durante la simulación. Una primera aproximación al valor de  $T_c$  sería tomar el valor de  $T'$  de nuestros puntos que da una diferencia  $U_4(L_1, T') - U_4(L_2, T')$  más cercana a cero (evidentemente ambas gráficas deben tener los mismos valores de temperatura). Sin embargo, si queremos conseguir una precisión por encima de los puntos tomados podemos usar métodos de interpolación.

### Estimación de $T_c$ a partir del cumulado de Binder

El punto de corte entre las curvas pueden estimarse “a ojo” gráficamente, sin embargo, como esto es poco matemático tres métodos numéricos distintos se han empleado:

- **Nearest data-point method:** Dados dos conjuntos de datos, este método encuentra el punto (de nuestros datos) más cercano al punto de intersección. Para ello, el método toma cada par de puntos  $(x, y_1)$  y  $(x, y_2)$  y calcula  $|y_2 - y_1|$ . Finalmente nos quedamos con el valor de  $x$  que haya proporcionado la diferencia  $|y_2 - y_1|$  más pequeña. La ventaja de este método es que es muy fácil de implementar, pero requiere que ambas gráficas tengan los mismos valores en el eje  $x$  para sus puntos.
- **Nearest data-point method, with linear interpolation:** Tras obtener los puntos (de nuestros datos) más cercanos al punto de corte, podemos utilizar una interpolación lineal para obtener una aproximación al punto de corte.
- **Fsolve Method:** Se interpolan las curvas usando splines, y después se resuelve la ecuación dada por la diferencia de los splines igualado a cero:  $\text{Spline1} - \text{Spline2} = 0$ . Uno de los problemas de este método es que falla si se pretende estimar varios puntos de corte. Además, otro problema de este método, es que requiere de una semilla inicial  $x_0$  relativamente cercana al valor real. Para mis datos he usado  $x_0 = 2.30$ . En consecuencia, este método puede usarse después de alguno de los dos anteriores.

Los dos últimos métodos tienen la ventaja de que nos permiten una precisión mayor que la propia discretización de las variables que estamos usando.

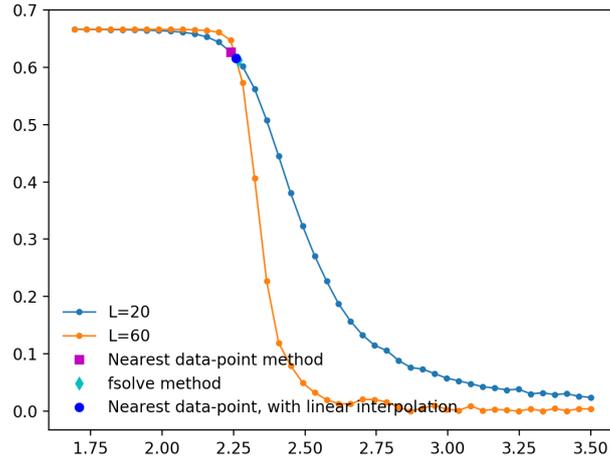


Figure 7: Ejemplo de aplicación de los métodos: Nearest data-point method with and without linear interpolation, and the fsolve method. El fsolve method requiere de una semilla inicial relativamente cercana al punto de corte real, por ejemplo, para el conjunto de datos de este trabajo solo da una estimación correcta de  $T_c$  cuando la semilla  $x_0 \in [2.24, 2.36]$ .

Sizes	NDPM	NDPM-l.i.	Fsolve
$L = 20, 60$	2.240	2.258	2.266
$L = 20, 100$	2.240	2.252	2.253
$L = 20, 140$	2.240	2.248	2.260
$L = 60, 100$	2.240	2.248	2.260
$L = 60, 140$	2.240	2.247	2.261
$L = 100, 140$	2.240	2.246	2.261

Table 2: Resultados de cada método aplicados para cada par de tamaños

## REFERENCES

- John Cardy. *Scaling and Renormalization in Statistical Physics*. Cambridge University Press, 1996.
- A.J. Fisher J.j. Binney, N.J. Dowrick. *The theory of critical phenomena: an introduction to the renormalization group*. Clarendon press - oxford, 1993.
- Andrea Pelissetto and Ettore Vicari. Critical phenomena and renormalization-group theory. *Physics Reports*, 368(6):549–727, oct 2002.
- T. C. Lubensky P. M. Chaikin. *Principles of Condensed Matter Physics*. Cambridge University Press, 1995.
- W Selke and L N Shchur. Critical Binder cumulant in two-dimensional anisotropic Ising models. *Journal of Physics A: Mathematical and General*, 38(44):L739–L744, oct 2005.
- Galin L. Hones Xiao-Li Meng Steve Brooks, Andrew Gelman. *Handook of Markov Chain Monte Carlo*. CRC Press, 2011.
- Pere Colete Raúl Toral. *Stochastic Numerical Methods: An Introduction for Students and Scientists*. Wiley- VCH, 2014.
- Bernard D. Flury. Acceptance–Rejection Sampling Made Easy. *SIAM Review*, 32(3):474–476, sep 1990.

Subhadip Raychaudhuri. Kinetic Monte Carlo Simulation in Biophysics and Systems Biology. In *Theory and Applications of Monte Carlo Simulations*. InTech, mar 2013.