

Física computacional - Modelo de Ising

Laboratorio 4

El mono

Resumen

Simulamos el modelo de Ising y calculamos algunas variables termodinámicas usando el método de Metropolis. Estudiamos la dependencia de estas variables con respecto a la temperatura del sistema y, en particular, observamos que existe una temperatura crítica T_c en la cual la magnetización presenta una bifurcación. Esto es, una transición de fase.

1. Introducción

En este trabajo simulamos el modelo de Ising y calculamos algunas variables termodinámicas usando Markov Chain Monte Carlo (MCMC) con una cadena generada mediante el método Metropolis [2]. Concretamente, simulamos un toro discreto de n^2 partículas, cada una de las cuales puede tener spin *up* o *down*; y calculamos, en función de la temperatura T^* (reducida) la magnetización M , la energía E , la susceptibilidad χ , el calor específico C y la temperatura crítica T_c . Estos cálculos son, en general, de la forma

$$Q = \mathbb{E}(f_Q(X)), \quad (1)$$

donde $X : \Omega \rightarrow S$ es una variable aleatoria que cae en el espacio S de estados del sistema con distribución de Boltzmann, y $f_Q : S \rightarrow \mathbb{R}$ es alguna función por medio de la cual se calcula la cantidad Q .

Dado que el espacio de estados S es de cardinal 2^{n^2} (pues cada una de las n^2 partículas puede tener spin *up* o *down*), es demasiado costoso calcular Q por definición. Por esta razón usualmente se opta por aproximar Q mediante integración Monte Carlo (MC). Sin embargo, no es trivial implementar este método de integración sobre S (pensándolo como un espacio de probabilidad con la distribución de Boltzmann). Una cantidad considerable de teoría de probabilidades es necesaria para entender (al menos, superficialmente) por qué funciona esta técnica y cuáles son sus limitaciones. En la siguiente sección presentamos un poco de esta teoría (para más detalle, ver [1, 4]).

2. Preliminares

En esta sección explicamos brevemente el método de Metrópolis y su utilidad en el contexto de la integración Monte Carlo. Esto es porque creemos que conocer ciertos detalles de la teoría de probabilidades puede ayudar a entender algunos de los fenómenos que se observan más adelante en las simulaciones numéricas.

2.1. Monte Carlo Markov Chain

Recordemos que la integración MC para calcular (1) requiere una forma de muestrear X , esto es, una forma de generar una sucesión de estados que sean realizaciones de variables X_n independientes e idénticamente distribuidas con la distribución de X . Hay varias formas de hacer esto, una de ellas consiste en generar una cadena de Markov tal que sus variables se aproximen a una muestra de X . De aquí el nombre: *Markov Chain Monte Carlo*.

Supongamos que existe una cadena de Markov $Y = (Y_n)$ sobre S con una única distribución invariante π igual a la distribución deseada (en este caso, la distribución de Boltzmann). Entonces, independientemente de la distribución inicial (generalmente Y_0 es uniforme o atómica), Y_n converge debilmente¹ a X . Así, a partir de cierto $N \in \mathbb{N}$, la cadena $Z = (Y_{N+n})$ se distribuye aproximadamente π . El fragmento restante de la cadena: Y_0, Y_1, \dots, Y_{N-1} se llama el *burn-in* y se descarta.

Ahora, si bien Z es un proceso de Markov donde cada Z_n se distribuye aproximadamente π , los Z_n podrían no ser independientes (recordemos que estamos en busca de una muestra de X , para lo cual la independencia es un factor crucial). En estas circunstancias es necesario extender la teoría de integración MC al caso MCMC con muestreo no independiente y calcular nuevamente la varianza de la integral. En este trabajo no discutimos estos cálculos, simplemente mencionamos que la varianza de la integral MC es ahora inversamente proporcional al número de muestras (como antes) y proporcional a τ_a , una variable llamada tiempo de autocorrelación que representa, intuitivamente, cuantas muestras Z_n debemos descartar para emular un muestreo independiente. En otras palabras, τ_a se define de forma que el proceso

$$W = (Z_0, Z_{\tau_a}, Z_{2\tau_a}, \dots)$$

no presente evidencia de dependencia.

Sobre el final del trabajo estimaremos algunos tiempos de autocorrelación mediante la ecuación (ver [1, 4])

$$\tau_a = \frac{1}{2} + \sum_{k \in \mathbb{N}} A(k), \quad (2)$$

donde

$$A(k) = \frac{\text{Cov}(f_Q(Z_0), f_Q(Z_k))}{\text{Var}(f_Q(Z_0))} = \frac{\mathbb{E}(f_Q(Z_0)f_Q(Z_k)) - \mathbb{E}^2(f_Q(Z_n))}{\mathbb{E}((f_Q(Z_0))^2) - \mathbb{E}^2(f_Q(Z_n))} \quad (3)$$

es la función de autocorrelación del proceso $f_Q(Z)$ y depende de la variable termodinámica Q que estemos estimando.

2.2. Método de Metropolis

Una forma ampliamente utilizada para construir una cadena de Markov con distribución invariante dada (en nuestro caso, la distribución de Boltzmann) es el método de Metropolis-Hastings. Este método define una cadena (Y_n) sobre el espacio de estados S a partir de una *propuesta* de transición, esto es, una familia de distribuciones de probabilidad $g(x' | x)$, que especifica la probabilidad de proponer un nuevo estado $x' \in S$

¹Esto es, $\mathbb{E}(h(Y_n)) \rightarrow \mathbb{E}(h(X))$ para toda h continua y acotada.

dado el estado actual $x \in S$.² A partir del estado actual $Y_n = x$, se genera un candidato $x' \sim g(\cdot | x)$, el cual es aceptado con probabilidad

$$\mathbb{P}(\text{aceptar } x' | x) = \min \left\{ 1, \frac{\pi(x') g(x | x')}{\pi(x) g(x' | x)} \right\}, \quad (4)$$

y rechazado (manteniendo el estado actual $Y_{n+1} = x$) con la probabilidad complementaria. Este procedimiento garantiza que la cadena resultante sea reversible con respecto a π y, bajo ciertas condiciones, ergódica. En particular, el método propuesto originalmente por Metropolis consideraba únicamente familias de distribuciones simétricas. Esto es, $g(x | x') = g(x' | x)$, de forma que la ecuación (4) se simplifica.

Será importante notar el siguiente comportamiento en el método de Metropolis. Hemos dicho que a partir de un estado $Y_n = x$ se samplea (con distribución $g(\cdot | x)$) un nuevo punto, al que más tarde decidiremos si movernos o no (con probabilidad dada por (4)). Intuitivamente, la sucesión Y_n navega por S pegando saltos de longitud dada por g . ¿Qué ocurre si la distribución invariante π es, por poner un ejemplo, bimodal con dos regiones de alta probabilidad separadas por un valle de baja probabilidad? Si g tiene poca varianza e $Y_n = x$ se encuentra en una de las regiones de alta probabilidad, es poco probable que *salte* hacia la otra región. Si bien está garantizado (por la convergencia débil) que eventualmente Y_n saltará hacia la otra región, en un experimento numérico, esto puede demorar mucho. Es posible que al hacer MCMC con este tipo de distribuciones, la realización de Y_n que usemos no represente π (como queremos) sino la restricción de π a una de estas regiones.

En resumen, el método de Metropolis es una forma de generar una cadena de Markov con la cual se aproxima una muestra de una distribución π . Con esta muestra se calcula la integral MC. Entre las limitaciones más importantes de esta técnica tenemos que: debemos descartar el *burn in*, cuyo tamaño es desconocido y difícil de acotar eficientemente, no tenemos independencia, aunque podemos simularla conociendo τ_a (también, desconocido y difícil de acotar eficientemente), y ciertas distribuciones pueden ser difíciles de representar, si no se modifica la distribución de propuestas g . En particular, la distribución de Boltzmann en el espacio de estados del modelo de Ising tiene forma bimodal si T es cercana a la temperatura crítica T_c , por lo que se necesitan muestreos extremadamente largos para obtener resultados confiables.

3. Resultados

En la Figura 1 presentamos, en función del número N de pasos MC, los resultados de la integral MC de la magnetización M y energía E de un modelo de Ising de tamaño 40×40 con condiciones de borde periódicas. Por simplicidad, llamaremos $\langle M \rangle$ y $\langle E \rangle$ a los resultados de estas integrales. Hacemos gráficos diferentes para diferentes temperaturas, y comparamos los casos donde el primer estado tiene todos los espines *up* (llamado *cold*

²Formalmente, se especifica una familia de medidas de probabilidad indexadas por S , pero por consistencia con la bibliografía, denotamos esta familia por medio de sus funciones densidad condicional $g(\cdot | x)$ a pesar de que no necesariamente son absolutamente continuas con respecto a la medida de Lebesgue.

start, porque su baja energía) y donde el primer estado tiene espines aleatorios (llamado *hot start*).³

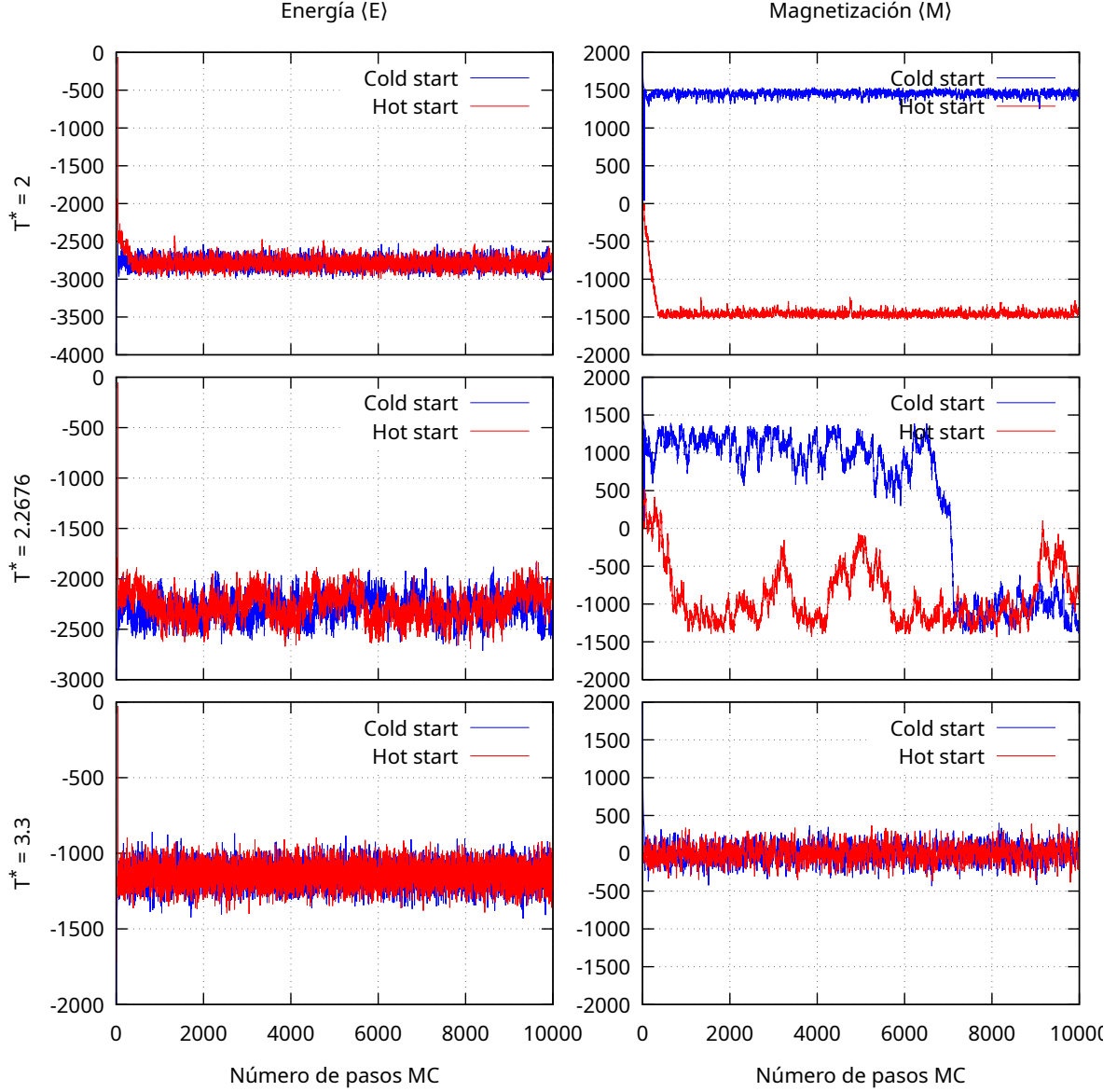


Figura 1: Valores de $\langle M \rangle$ y $\langle E \rangle$ en función del número de pasos MC.

En la Figura 2 presentamos,⁴ como función de la temperatura reducida T^* , el valor absoluto de la magnetización por partícula, la energía por partícula, la susceptibilidad y el calor específico. Estimamos todos estos valores por integración MC con $N = 10000$ pasos y un *burn in* de 1000 pasos. Hacemos esto para modelos de Ising de tamaños $n = 10$, $n = 20$ y $n = 40$, donde n es el número de partículas por lado. En la Figura 3 presentamos, para discutir posteriormente, un acercamiento de los gráficos de la Figura 2.

³Inciso a.

⁴Inciso b.

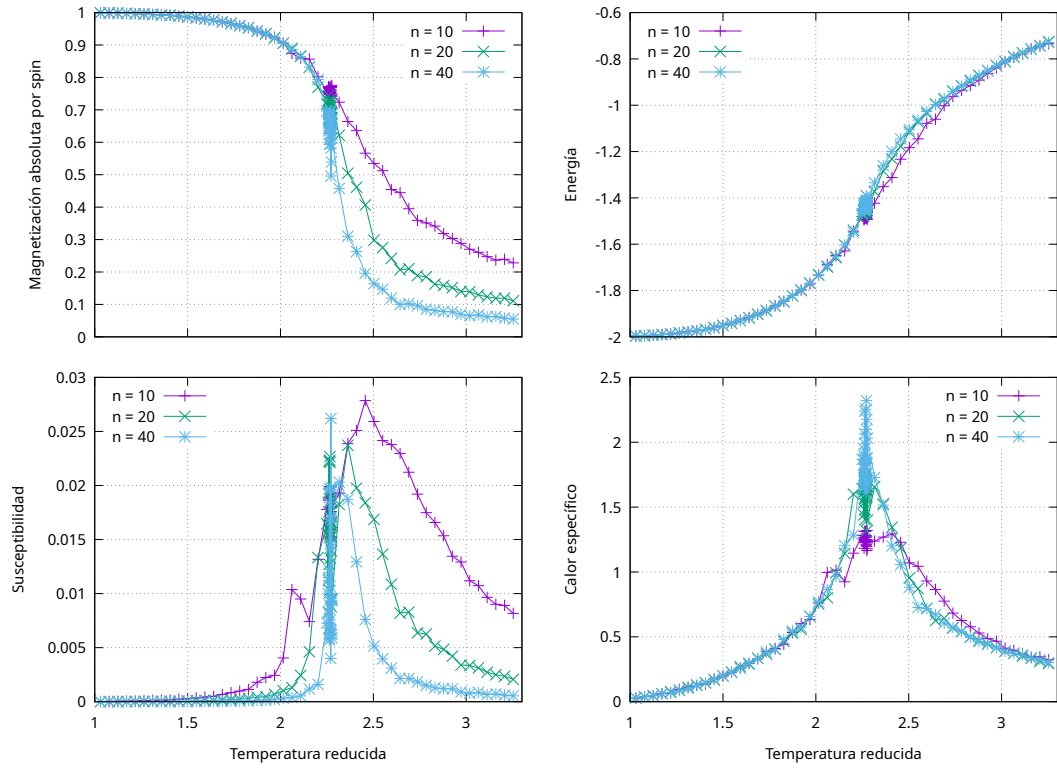


Figura 2: Algunas variables termodinámicas en función de la temperatura reducida T^* .

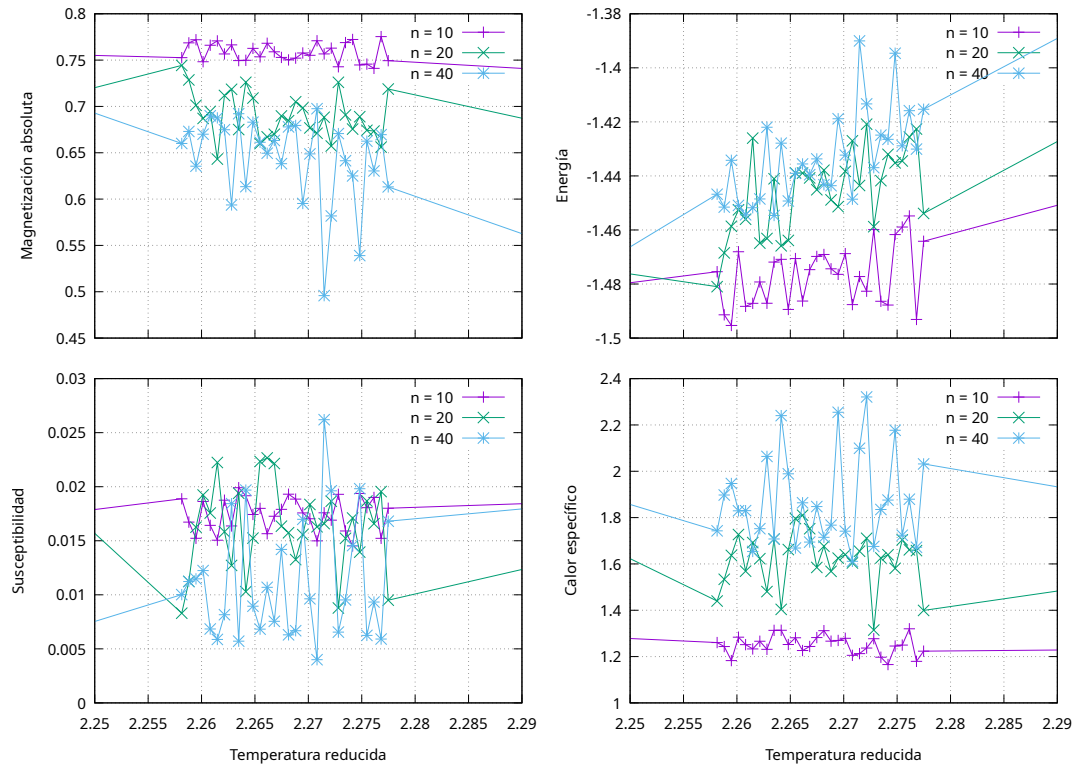


Figura 3: Acercamiento de la Figura 2.

En la Figura 4 presentamos,⁵ para dos temperaturas reucidas diferentes (una ligeramente por encima y otra ligeramente por debajo de la temperatura crítica $T_c = 2.2676$), los histogramas de la magnetización en el modelo de Ising de tamaño 40×40 . Para esto, como antes, excluimos un *burn in* de 1000 pasos MC y armamos el histograma con los datos de magnetización de $N = 10000$ estados del modelo.

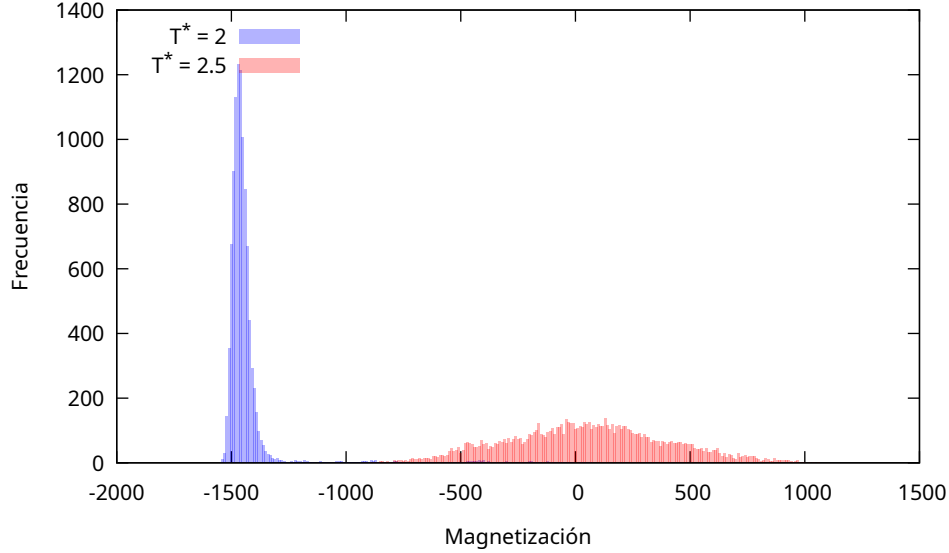


Figura 4: Histogramas de $\langle M \rangle$ a diferentes temperaturas.

En la Figura 5 presentamos,⁶ para distintos tamaños de modelo, los *cumulantes de Binder*. Calculamos estos valores a partir de datos de magnetización obtenidos por integración MC con *burn in* de 1000 pasos y tamaño de muestra de $N = 100000$ pasos MC. Notese el incremento en el número de pasos.

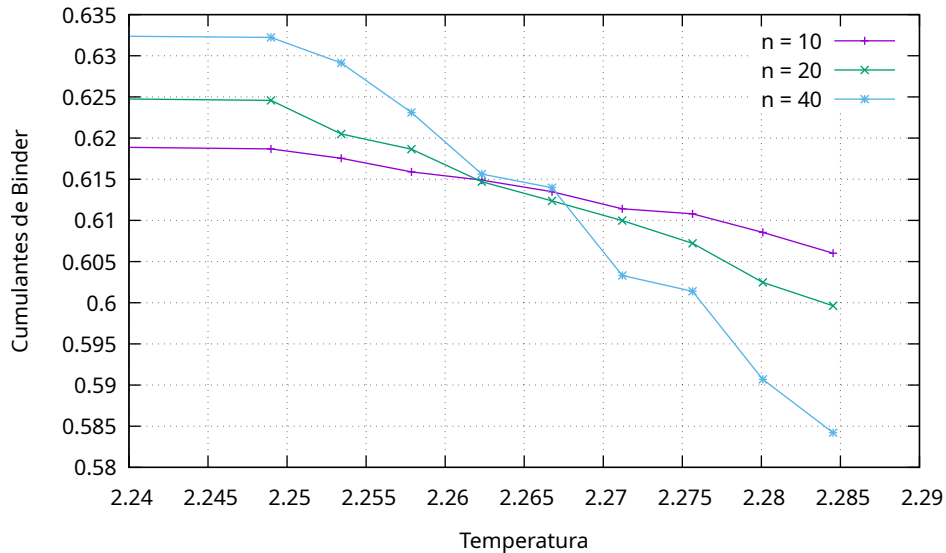


Figura 5: Cumulantes de Binder para modelos de diferentes tamaños.

⁵Inciso c.

⁶Indiso d.

En la Figura 6 mostramos los valores de $\langle |M| \rangle$ y χ obtenidos anteriormente por integración MC junto con las funciones $(T_c - T)^\beta$ y $|T - T_c|^{-\gamma}$. Se usan los exponentes críticos $\beta = \frac{1}{8}$ y $\gamma = \frac{7}{4}$ encontrados por Onsager [CITA A ONSAGER].

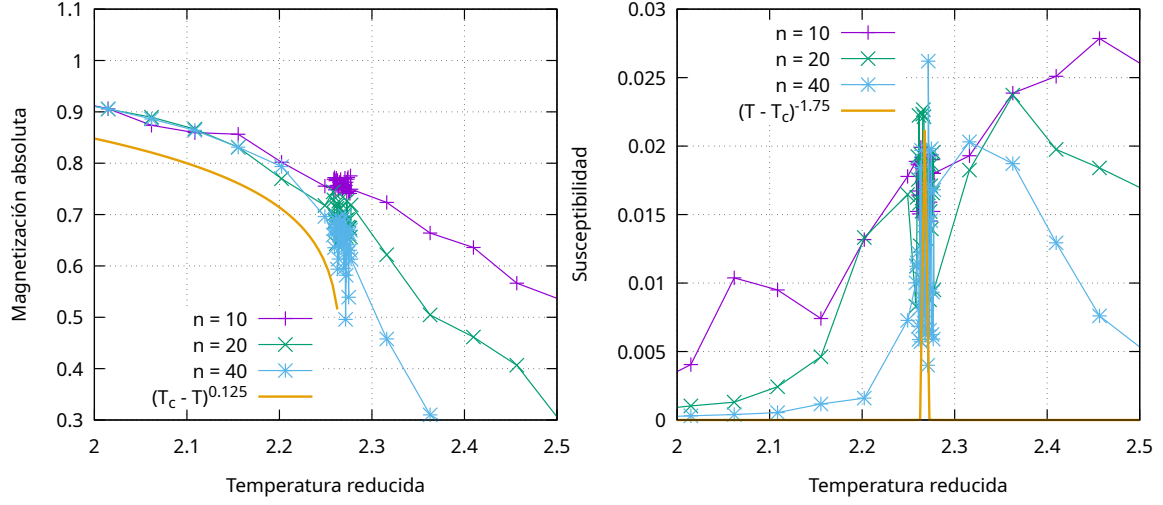


Figura 6: Exponentes críticos.

Finalmente, en la Figura 7 presentamos las funciones de autocorrelación de la magnetización $|M|$ y energía E calculadas a partir de los estados dados por el método de Metropolis. En función de estos datos, en la Tabla 1 presentamos cotas para los tiempos de autocorrelación de magnetización τ_M y energía τ_E .

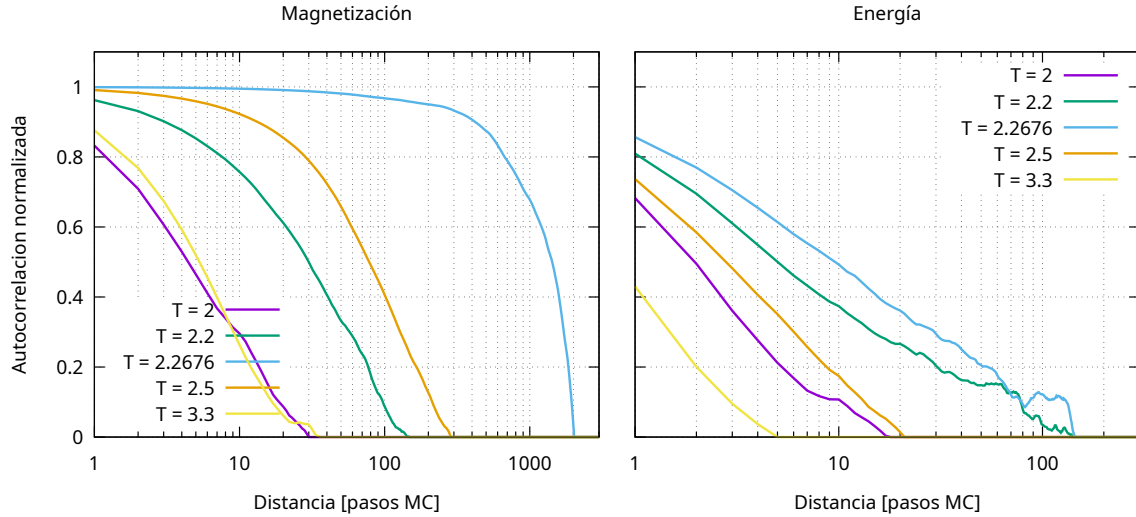


Figura 7: Funciones de autocorrelación de $|M|$ y E .

T	τ_M	τ_E
2	7.4	3.5
2.2	40.9	20.3
2.2676	1179	29.7
2.5	96	5
3.3	7.5	1.3

Tabla 1: Tiempos de autocorrelación τ_M y τ_E en función de T .

4. Discusión

Esta discusión podría ser extensa, así que partamosla en subsecciones.

4.1. Muestreo de S con el método de Metropolis

Aca ponemos comentarios sobre el inciso a, que justamente nos dice que tan estable es el muestreo. Y de donde estimamos que un burn in de 1000 puntos es razonable para la mayoría de las temperaturas (ya cerca de T_c no hay nada que hacer...).

4.2. Variables termodinámicas

Pasamos ahora al inciso 2, lo importante es notar que todo parece en orden salvo cerca de las temperaturas críticas, esto va a ser un factor común en todos los resultados (salvo los relacionados con los cumulantes de Binder, que son un poco mas estables). Pusimos un zoom in de estos graficos justamente para que se aprecie.

Analizamos tambien los histogramas aca, que estan indicando que la magnetización a temperaturas por encima de la crítica es desorganizada (lo cual es esperable) y temperaturas por debajo de la crítica son estables y se agrupan.

4.3. Estimación de T_c

Para esto existen los cumulantes de Binder, no se suficiente física así que no voy a emitir comentario al respecto. Aca es donde digo que tiene pinta que algo pasa entre 2.2625 y 2.2675. Sin embargo mi compu tiro corridas de $N = 100000$ y no dio para mas, así que hasta ahí llega nuestra precision.

4.4. Autocorrelación y estimación de errores

Como mencionamos en la sección 2, la varianza de la integración monte carlo en casos de muestreo no independiente es mayor que la de muestreo independiente. Se estima mayor en un factor τ_a , que estimamos en la Table 1, así que a pesar de que usamos $N = 10000$, con lo cual deberíamos esperar resultados con desviación de 0.01 unidades, la verdadera desviación debe multiplicarse por $\sqrt{\tau_a}$.

Por poner un ejemplo, en la Figura 4 vemos que la magnetización con $N = 10000$ tiene media 0 y hemos calculado que tiene varianza $1.19e5$. O sea una desviación de 345 mas o menos. La desviación si todo fuera independiente es 0.01 multiplicado por los valores de la variable que se estan tomando, que ronda los 1600 diría yo, entonces esperamos una

desviación de 16 unidades... pero vemos una de 345! Ahora, la tabla nos dice que el τ_M para $T = 2.5$ es de mas o menos 96! que tiene raíz mas o menos 10, lo cual nos lleva de 16 unidades de desviación a 160 mas o menos, mucho mas cerca de las 345 observadas experimentalmente. Este es un ejemplo de mierda, por supuesto, ojala pudiéramos hacer otro. De hecho, podemos.

Mentira, no podemos poner ningun ejemplo porque precisamos conocer justamente la desviación de la propia función de energía o magnetización en el espacio de estados con la verdadera distribución de Boltzman! Si pudieramos estimar eso, entonces tendríamos como para laburar un buen ejemplo de la desviación esperada... pero bueno, no quiero hacerlo. No estaría mal ponerlo como si pudieramos intentarlo en el futuro.

4.5. Cosas que me quedan por agregar.

A medida que aumenta n , sabemos que la magnetización aleatoria se pelea con la temperatura. Si la temperatura T^* es suficientemente grande, la probabilidad de aceptar un cambio crece, lo que conlleva mayor variabilidad de la magnetización en cada paso MC. Entonces, para temperaturas bajas, esperamos ver, en los *cold starts* que la magnetización general se mantiene, pues es poco propenso que el sistema acepte los cambios propuestos en un paso MC. Y en el *hot start*, esperamos ver uqe el sistema se decida rapidamente por una magnetización mas o menos uniforme y permanezca ahi. Esto es lo que se ve en los gráficos. Por mera casualidad, el hot start decidio estabilizarse con espin negativo...

Con respecto valores grandes de T^* esperamos que la magnetización sea completamente aleatoria a lo largo de los pasos MC, pues se aceptan casi todos los cambios, sin importar si empezamos frios o calientes.

Lo interesante pasa para temperaturas intermedias. En $T^* = 2.2676$ observamos que la energía no se estabiliza, sinó que oscila (de forma no periódica) en un rango de enrgías. Al mismo tiempo, la magnetización tampoco se estabiliza, sinó que oscila entre períodos de magnetización casi uniformemente positiva y casi uniformemente negativa. Notemos ademas, que parecen coincidir los valores de n para los cuales $\langle M \rangle$ oscila y $\langle E \rangle$ toma valores grandes. Esto nos indica que la distribución de Boltzmann para este valor de T^* es (hablando mal y pronto) bimodal, y la cadena de Markov del método de Metropolis debe pasar un tiempo en la región dominada por una de las modas y otro tiempo en la región dominada por la otra moda. A medida que la cadena transiciona entre ambas regiones, pasa por estados de energía mas alta, por esto las transiciones en $\langle M \rangle$ se corresponden con incrementos en $\langle E \rangle$.

Uno podría pensar que en este caso la integral MC no converge rápidamente, por lo que habría utilizar algún método alternativo, una cadena de Markov distinta quizas, o cualquier otra alternativa. Sin embargo, en la opinión del autor, el hecho de que $\langle M \rangle$ no converge es evidencia de que estamos haciendo la pregunta incorrecta, no de que somos incapaces de responderla. Me explico: si sabes que tu distribución es bimodal, qué sentido tiene preguntarte por la media? No es información representativa de la distribución. La integral MC calcula correctamente la media (si es cierto que tarda bastante mas en converger), pero preguntarse por la media en una bimodal es inútil. De hecho, en el caso $T^* = 0$ la distribución de Boltzmann también es bimodal pero la integral de MC se estabiliza en una de las modas: ESTE ES EL VERDADERO ERROR! La verdadera integral de M debería icluír información de ambas modas y equilibrarse en $\langle M \rangle = 0$. La falencia aquí es de la cadena de Markov del método de Metropolis, que no recorre correctamente el espacio de estados.

En principio, esto significa que no podemos confiar en los resultados calculados para temperaturas bajas! Sin embargo, dado que el sistema es simétrico con respecto al cambio de espines, aún podemos obtener resultados que ilustren el comportamiento del sistema si usamos el valor absoluto de la magnetización en lugar de la magnetización con signo. Repito: no cambiamos de M a $|M|$ porque las cosas *se terminan promediando en 0* sino justamente porque no lo hacen. El problema es mas serio que simplemente no obtener información! ESTO TAMBIEN PODRÍA FRASEARSE COMO QUE ESTAMOS CAMBIANDO LA PREGUNTA.

Bueno, con respecto a los demas graficos no hay mucho que decir porque yo mismo no se mucho sober lo que esta pasando, puedo interpretar las probabilidades y demas cosas, pero no tanto la física. O decir a que corresponde macroscópicamente.... así que estas partes del informe van a quedar cortiñas... jaja willie cortiñas.

5. Trabajo actual

Todavía nos queda estimar bien T_c , aclara que no estamos tomando los burn-ins, y una vez que tengamos T_c estimada, tenemos que aproximar los exponentes críticos usando regresión exponencial y presentar gráficos de eso, diría yo... aunque podemos cargarlo en los mismos gráficos que ya venimos teniendo, porque estan todos exponenciales...

Al discutir el error va a haber qu ecitar tanto el libro que me paso Guido que hay que rastrear como el paper de Jankel que nos paso la vero.

Disminuír cantidad de puntos y hacer simulaciones mas largas. La variabilidad parece aumentar con $n = 40$, eso podría ser porque a tamaños mas grandes, cuesta mas saltar entre las 2 modas de una distribución bimodal. Escribir eso... no se si tendremos tiempo para hacer una corrida mas larga para los experimentos. Sería mejor poner un poco de esfuerzo para estudiar lo que es el tema siguiente.

6. Trabajo futuro

[3] fue de mucha ayuda...

Para este trabajo no fueron tantas las herramientas computacionales las que aprendimos sino mas bien las físicas. Estudiamos un poco de mecanica estadística para saber que estaba pasando, y estudiamos un montón de estadística para entender mas los procesos aleatorios y sus caracterizaciones, como se prueban independencias y demas cosas. Es un trabajo agotador pero nuevamente, me encanta.

Otra vez podríamos haber usado CUDA pero no lo hicimos (aunque ya las cosas estaban tardando un pelin mas...).

ESTA ES LA ÚLTIMA SECCIÓN DE TRABAJO FUTURO, LA DEJO PARA TENER UNA REFERENCIA PARA ACTUALIZAR MAS TARDE.

Finalmente! Las imágenes tienen letra grande. Este es el más notorio de nuestros avances, pero hubo otros: adoptamos una estructura de carpetas *estandar* (en cierto sentido) para el código correspondiente a los ejercicios, automatizamos la compilación, testeo y generación de imágenes por medio de archivos `makefile`, agregamos control de versiones por medio de Git y almacenamos el código en un repositorio online con GitHub. También estudiamos lo básico de CUDA, un lenguaje similar a C++ desarrollado por NVidia que

permite programar para los procesadores de las GPUs, pero como los ejercicio de este laboratorio no eran demasiado costosos computacionalmente, no fue necesario implementar los generadores en este lenguaje. Esperamos, sin embargo, que cobre relevancia en laboratorios futuros.

Este ha sido, hasta el momento, el laboratorio más divertido. Entre las cosas que quedan por explorar, las más importantes son: el manual del compilador (opciones de optimización, tratar advertencias como errores, usar varios compiladores para detectar diferentes errores, etc), el lenguaje de scripting Bash, y temas básicos de mecánica estadística, para poder hacer los siguientes laboratorios.

Referencias

- [1] W. Janke. Statistical analysis of simulations: Data correlations and error estimation. 08 2009.
- [2] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.
- [3] G. Phillies. *Elementary Lectures in Statistical Mechanics*. Graduate Texts in Contemporary Physics. Springer New York, 2000.
- [4] W. von der Linden and E. Schachinger. Chapter 7: Markov chain monte carlo. https://itp.tugraz.at/LV/wvl/Comp_Simulationen/, 2007. Part of the lecture notes *Computer Simulationen*, Graz University of Technology.