

ALGORITMOS DE TRASMISIÓN DE MENSAJES EN EL FERROMAGNÉTICO DE ISING BIDIMENSIONAL

MESSAGE-PASSING ALGORITHMS FOR THE BIDIMENSIONAL ISING FERROMAGNET

E. DOMÍNGUEZ[†], A. LAGE Y R. MULET

Cátedra de Sistemas Complejos Henri Poincaré y Departamento de Física Teórica, Facultad de Física, Universidad de La Habana, CP 10400 Habana, Cuba. eduardo@sica.uh.cu†

† autor para la correspondencia

La técnica de los algoritmos de transmisión de mensajes permite aproximar la energía libre -y el resto de las magnitudes termodinámicas-, de sistemas discretos de espines. El ferromagnético de Ising bidimensional (FIB) es un sujeto de pruebas ideal para ilustrar el funcionamiento de los algoritmos *Belief Propagation* (BP) y *Generalized Belief Propagation* (GBP) así como la relación de estos con un método de la física estadística llamado *Cluster Variation Method* (CVM). La relativa sencillez del FIB facilita comparar los resultados obtenidos en diferentes aproximaciones de los algoritmos con un cálculo variacional exacto de la energía libre correspondiente (Bethe, Kikuchi, campo medio, etc).

Message-passing algorithms are exploited to find approximated solutions for discrete spin systems. Many thermodynamic quantities can be obtained, often accurately, by this method. The Ising bidimensional ferromagnet (FIB) is a good trial system used here to test different approximations. We compare algorithmic results with analytic calculations to illustrate the matching between the algorithms named *Generalized Belief Propagation* (GBP) and a family of free energy variational approximations called *Cluster Variation Method* (CVM).

PACS: Ising model magnetic ordering, 75.10.Hk; Statistical mechanics of phase transitions in model systems, 64.60.De; Information theory entropy in, 89.70.Cf

INTRODUCCIÓN

Un resultado fundamental de la física estadística de los sistemas en equilibrio a temperatura $T = \beta^{-1}$, es que la distribución de probabilidad de los microestados $\vec{\sigma} \in S$ es la conocida distribución de Boltzmann

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

donde $H(\vec{\sigma})$ es la energía del estado -que se supone conocida y fácil de calcular-. La constante de normalización Z es la llamada función de partición y se calcula según

$$Z = \sum_{\vec{\sigma} \in S} e^{-\beta H(\vec{\sigma})} \quad (2)$$

Los diferentes momentos de la distribución (1) permiten determinar los valores medios de las variables del sistema, las susceptibilidades, etc. Además, por medio de la función de partición se hallan los diferentes potenciales termodinámicos, y así toda la fenomenología del problema. Formalmente entonces puede decirse que en la ecuación (2) está contenida toda la información relevante del sistema en equilibrio.

Sin embargo, el cálculo exacto de Z , o de las distribuciones marginales de probabilidad, es posible solamente en un

número muy reducido de casos ya que, en general, la suma por todos los estados del sistema no se puede hallar analíticamente. Para la inmensa mayoría de los sistemas el tamaño del espacio de configuraciones, $|S|$, crece exponencialmente con el número de variables (partículas, entes) que lo componen, por lo cual se hace imposible realizar un cálculo exhaustivo de Z .

Como resultado de una fructífera interrelación entre la física estadística, la optimización computacional y la teoría de la información, han surgido métodos y enfoques alternativos. Se ha desarrollado una amplia gama de algoritmos de inferencia que permiten estimar las distribuciones de probabilidad de las diferentes variables de un sistema [1, 2, 3]. Una clase muy importante la constituyen los algoritmos de transmisión de mensajes.

Nuestro propósito fundamental es ilustrar el uso de estos algoritmos de transmisión de mensajes así como su íntima conexión con la física estadística. En particular nos interesa el algoritmo *Generalized Belief Propagation* (GBP) y su relación con un método de la física estadística conocido como *Cluster Variation Method* (CVM) [4, 5]. Aplicamos los algoritmos mencionados a un modelo muy

estudiado pero interesante como sujeto de pruebas: el sistema ferromagnético de Ising bidimensional (FIB). Este es uno de los pocos sistemas discretos no triviales que posee solución exacta conocida [6, 7]. Además, su relativa sencillez permite comparar los resultados de los métodos algorítmicos con los obtenidos mediante el CVM y con valores exactos de las diferentes magnitudes estimadas.

Para precisar, consideraremos una red cuadrada bidimensional de N espines $\sigma_i = \pm 1$ (ver figura (1)), que interactúan con sus primeros vecinos según el hamiltoniano

$$E(\vec{\sigma}) = -\sum_{\langle ij \rangle} J \sigma_i \sigma_j, \text{ con } J > 0 \quad (3)$$

donde J es la constante que caracteriza la intensidad de las interacciones entre espines y la suma se realiza por todos los pares de primeros vecinos. Se consideran condiciones de frontera periódicas. Es conocido que FIB exhibe una transición de fase de segundo orden de un estado paramagnético a ferromagnético a una temperatura $T \approx 2.2692 J$ para campo externo nulo [6, 7]. Mediante una aproximación de campo medio puede explicarse cualitativamente la presencia de esta transición, pero el valor de la temperatura crítica obtenido ($T_c = 4.0 J$) está lejos del resultado exacto. Esto es consecuencia de que dicha aproximación es particularmente mala en sistemas de dimensión finita y donde el orden local y la geometría son relevantes. En este trabajo se aplican las aproximaciones de Bethe y CVM, que sí tienen en cuenta la estructura del sistema, para obtener mejores resultados.

Para organizar la exposición se ha dividido el resto del trabajo en cinco partes o secciones. La primera sección presenta informalmente las ideas relativas al CVM, poniendo énfasis en su implementación en el FIB. Se propone un método variacional para aproximar la energía libre total, y así evitar el cálculo exhaustivo de Z . Seguidamente, se describe el algoritmo GBP y se plasman las ecuaciones que lo caracterizan. Este algoritmo es uno de los medios para enfrentar el cálculo propuesto en la sección anterior. En la sección siguiente se desarrolla, de manera exacta, el cálculo variacional propuesto. Se explica el papel de la simetría en la simplificación del proceso. Los principales resultados obtenidos por sendos métodos -algorítmico y variacional- se exponen en la última sección. Las gráficas que se muestran subrayan la correspondencia entre ambos. De forma breve se mencionan propiedades de la dinámica del algoritmo GBP. Por último se esbozan, a modo de resumen, las conclusiones.

DESCRIPCIÓN DEL MÉTODO CVM

La energía libre variacional de un sistema se define como un funcional de la distribución de probabilidades de los estados

$$F[p(\vec{\sigma})] = U[p(\vec{\sigma})] - TS[p(\vec{\sigma})]$$

De acuerdo con los postulados de la física estadística se deduce que la distribución $p(\vec{\sigma})$ real es aquella que minimice $F[p]$. La minimización de esta energía libre es un problema computacional al menos tan difícil como el cálculo de la función de partición. La idea del CVM, propuesto por Kikuchi, es minimizar una magnitud F_{Kik} que aproxime a F pero que haga los cálculos más asequibles. Para definir F_{Kik} se divide el sistema en regiones y se suma la contribución a la energía libre de cada una. El aporte de cada región depende de la distribución local de probabilidades de las variables de esta. Como consecuencia, se transforma la minimización por el espacio de las distribuciones globales de probabilidad a la minimización en el espacio de las distribuciones locales.

Con más precisión, en el CVM se divide la red en un conjunto de regiones maximales que constituyan un cubrimiento, es decir, que no quede ningún enlace ni espín fuera de región. Cada una de las regiones maximales debe contener algunos espines y todos los enlaces entre estos, por ejemplo, en la red cuadrada una región puede ser un cuadrado de 2×2 espines, incluyendo los enlaces que los "cierran". A este tipo de regiones de 2×2 le llamaremos plaquetas, ver figura (2). Un conjunto maximal sería el conjunto de todas las plaquetas posibles de la red.

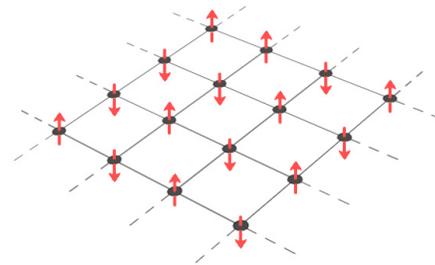


Figura 1. Red bidimensional de espines binarios.

Es evidente que las regiones tienen que solaparse entre sí, puesto que no debe quedar ningún enlace fuera del cubrimiento. Con los elementos compartidos entre dos regiones primarias podemos formar otra región; por ejemplo, dos plaquetas contiguas se solapan en dos espines y un enlace común; a esta región la llamaremos enlace. Al conjunto inicial de regiones le añadimos el conjunto formado por todas las intersecciones posibles, esto es, en nuestro ejemplo, adicionar al conjunto de todas las plaquetas, el conjunto de todos los enlaces. El proceso se repite añadiendo las intersecciones de las intersecciones y así sucesivamente hasta que ya no sea posible continuar. Mediante esta sencilla regla se genera el conjunto de regiones utilizado en el CVM. En el caso de que se comience con regiones maximales del tipo plaqueta, se tendrán al final regiones plaqueta, enlace y espín (como intersección de los enlaces). La aproximación obtenida si las regiones maximales son del tipo enlace se llama aproximación de Bethe.

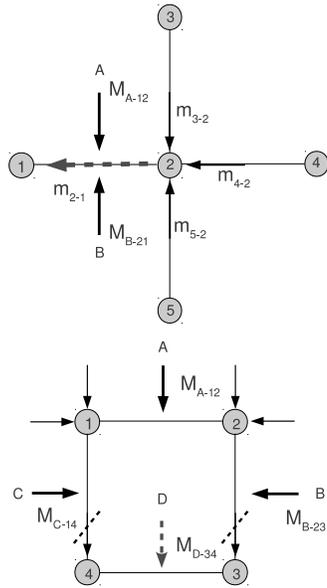


Figura 2. Representación esquemática del cálculo de los mensajes. El mensaje actualizado aparece en líneas discontinuas.

A cada región definida se le asocia una energía media U_R y una entropía S_R de la manera usual mediante las ecuaciones (5) y (6). La energía E_R está determinada por las interacciones que queden dentro de la región, por ejemplo a una plaqueta le corresponde la energía de cuatro enlaces; también debe considerarse la interacción de los espines con un posible campo externo h . La interacción con este campo se tendría en cuenta mediante un término aditivo de la forma $E_i = -h\sigma_i$ para cada espín. En este trabajo consideramos siempre el campo externo nulo. No obstante, hay que señalar que el cálculo propuesto para el CVM es posible también en el caso de un campo externo homogéneo. El algoritmo GBP (que veremos más adelante) es capaz de lidiar incluso con un campo externo desordenado.

$$U_R[p_R] = \sum_{\vec{\sigma}_R} p_R(\vec{\sigma}_R) E_R(\vec{\sigma}_R) \quad (5)$$

$$S_R[p_R] = -\sum_{\vec{\sigma}_R} p_R(\vec{\sigma}_R) \ln p_R(\vec{\sigma}_R) \quad (6)$$

$$\vec{\sigma}_R = (\sigma_i, \sigma_j, \dots, \sigma_l) \text{ con } i, j, \dots, l \in R.$$

En las ecuaciones anteriores p_R es la distribución de probabilidades de las variables que componen la región. Para la energía libre queda, naturalmente

$$F_R[p_R] = U_R[p_R] - T S_R[p_R] \quad (7)$$

Por último, la energía libre total del sistema en la aproximación de Kikuchi será la suma de la energía libre de cada región, multiplicada por un coeficiente c_R que tenga en cuenta el solapamiento entre regiones.

$$F_{Kik} = \sum_{R \in \mathcal{R}} c_R F_R \quad (8)$$

De esta manera se pretende aproximar la energía libre real por los aportes a ella de cada región. Este procedimiento es exacto para la energía total; siempre se puede escribir como la suma de la energía de cada región. Para la entropía, sin embargo esta última afirmación no es cierta en general. En muy pocos casos la entropía total es la suma de las entropías locales [3].

El método CVM puede parecer similar a una aproximación de campo medio, sin embargo existe una diferencia notable. La aproximación de campo medio propone una forma manejable, generalmente factorizable, de la distribución de probabilidades global para el cálculo de la energía libre. Puede demostrarse que la energía libre de campo medio obtenida es una cota superior del valor exacto. Por otro lado, el CVM no utiliza ni permite construir una distribución de probabilidades global a partir de las distribuciones locales. Por esta causa, el valor de F_{Kik} no es una cota superior del valor real de F .

Resumiendo, lo que interesa en el CVM es determinar la distribución de probabilidades de cada región de forma que se extreme la F_{Kik} . A primera vista este podría parecer un problema igual de formidable que el original pero, afortunadamente, aquí es donde entran los algoritmos de transmisión de mensajes a jugar su papel.

ALGORITMO GBP EN LA APROXIMACIÓN DE KIKUCHI

En la sección anterior vimos cómo se plantea el cálculo aproximado de la energía libre en términos de las distribuciones locales de probabilidad que extremen cierto funcional. En general, la minimización directa es un problema complicado, dada la no convexidad del funcional de interés (8).

En el ámbito de las ciencias de la computación se desarrollaron los algoritmos de transmisión de mensajes como herramienta que permite realizar cálculos de inferencia de las distribuciones de probabilidad locales de las variables en un grafo o red. Estas distribuciones son las que se necesitan para aproximar la energía libre de cada región. A continuación se describe las magnitudes fundamentales y las ecuaciones que rigen a estos algoritmos.

En cada región del sistema -definida previamente mediante el CVM- se definen dos conjuntos de magnitudes fundamentales: *beliefs* (opiniones) y mensajes. El *belief* de una región, $b_R(s_R)$, representa una especie de aproximación a la distribución real de probabilidades de cada estado de los espines contenidos en la región. Los mensajes son pedazos de información que las regiones intercambian en aras de “coordinar” sus opiniones. El algoritmo consiste en la propagación de los mensajes por la red, mediante ciertas reglas, hasta que se alcance un “consenso” (estabilidad). Este es el motivo del nombre *Generalized Belief Propagation*. Los algoritmos de transmisión de mensajes se han aplicado con buenos resultados en diferentes problemas, ofreciendo una solución aproximada en un tiempo corto en comparación con otros procedimientos[3].

La razón de este éxito no se pudo aclarar completamente hasta que se descubrió la equivalencia de este método con la minimización de la energía libre de las regiones. Se puede demostrar que cuando el algoritmo converge, lo hace a un mínimo local de la energía libre correspondiente [5, 8]. La función de los mensajes es equivalente a la de ciertos multiplicadores de Lagrange que surgen en el proceso de minimización. Las ecuaciones de iteración constituyen un método de punto fijo para calcular estos multiplicadores. Por otro lado, es importante destacar que la calidad de la solución depende de la estructura del sistema; por ejemplo, es conocido que para una red que no presente ciclos, la inferencia es exacta. Para sistemas con ciclos de larga longitud los resultados son igualmente satisfactorios. La aplicación del GBP en una red cuadrada bidimensional ha de ser cautelosa debido a la presencia de ciclos cortos. Los resultados obtenidos deben considerarse siempre una aproximación. Se considera una buena práctica el incluir en las regiones maximales al menos los ciclos más cortos del sistema.

En la aproximación de Kikuchi se definen tres tipos de regiones: plaqueta, enlace y espín. Los mensajes que se intercambian son de dos tipos, de plaqueta a enlace $m_i(\sigma_j)$ y de enlace a espín $M_{ij}(\sigma_p, \sigma_j)$. Cada una de las regiones construye su *belief* en función de los mensajes que le llegan de las regiones adyacentes.

El valor del *belief* del espín i será proporcional al producto de los mensajes que “entran” a este desde los otros espines

$$b_i(\sigma_i) \propto \prod_{j \in N(i)} m_{j \rightarrow i}(\sigma_i) \quad (9)$$

aquí $N(i)$ representa el conjunto de todos los vecinos del espín i . La fórmula anterior nos dice que un espín determina la probabilidad de estar en un estado u otro en dependencia de la información que recibe de sus primeros vecinos en la red.

El *belief* para un enlace se calcula según

$$\begin{aligned} b_{ij}(\sigma_i, \sigma_j) &\propto e^{-\beta J \sigma_i \sigma_j} \prod_{n \in N(i) \setminus j} m_{n \rightarrow i}(\sigma_i) n \\ &\times \prod_{n \in N(j) \setminus i} m_{n \rightarrow j}(\sigma_j) n \\ &\times \prod_{M_k \in P(\langle i, j \rangle)} M_k(\sigma_i, \sigma_j) \end{aligned} \quad (10)$$

donde $P(\langle i, j \rangle)$ es el conjunto de los mensajes que las plaquetas vecinas le envían al enlace $\langle i, j \rangle$.

De la misma manera se escribe una ecuación para la plaqueta donde intervienen los mensajes de plaquetas vecinas hacia los enlaces internos M_p , así como mensajes de los enlaces vecinos a los espines internos $m_{n \rightarrow p}$

$$\begin{aligned} b_{ijkl}(\sigma_i, \sigma_j, \sigma_k, \sigma_l) &\propto e^{-\beta H(\sigma_i, \sigma_j, \sigma_k, \sigma_l)} \prod_{M_p \in U} M_p(\mathbf{s}_p) \\ &\times \prod_{a \in \{ijkl\}} \prod_{n \in N(a)} m_{n \rightarrow a}(\sigma_a). \end{aligned} \quad (11)$$

En la ecuación (11) el conjunto U está formado por todos los mensajes provenientes de las plaquetas vecinas

El hamiltoniano “local” no es otra cosa que

$$H(\sigma_i, \sigma_j, \sigma_k, \sigma_l) = -J(\sigma_i \sigma_j + \sigma_j \sigma_k + \sigma_k \sigma_l + \sigma_l \sigma_i).$$

Las ecuaciones usadas para calcular los *belief* tienen una estructura similar en todas las regiones. Siempre se trata del producto de un factor de Boltzmann y un conjunto de mensajes externos. De esta forma se logra un balance entre las interacciones internas y el exterior de cada región.

Sobre el conjunto de magnitudes definidas anteriormente se impone una serie de restricciones que garantiza su consistencia. Primeramente se exigen condiciones de normalización para cada región

$$\sum_{\mathbf{s}_R} b(\mathbf{s}_R) = 1.$$

Además, es necesario que el *belief* de cada región sea congruente con el de las subregiones que contiene; por ejemplo, al sumar por las variables σ_k, σ_l , debe cumplirse que

$$\sum_{\sigma_k, \sigma_l} b(\sigma_i, \sigma_j, \sigma_k, \sigma_l) = b(\sigma_i, \sigma_j). \quad (12)$$

Este tipo de condición se llama de “marginalización”.

Cada una de las ecuaciones genera relaciones auto-consistentes que ligan a todos los mensajes entre si. Esas son las ecuaciones que se resuelven iterativamente pasando mensajes de región en región.

Como hay dos tipos de mensajes en la aproximación de Kikuchi, se tienen dos tipos de ecuaciones diferentes; a cada uno de los mensajes de la red le corresponde una diferente. La forma de la ecuación (13) aparece representada esquemáticamente en la parte superior de la figura 2,

$$m_{2 \rightarrow 1}(\sigma_1) \propto \sum_{\sigma_2} e^{\beta J \sigma_1 \sigma_2} M_{A \rightarrow 12} M_{B \rightarrow 21} m_{3 \rightarrow 2} m_{4 \rightarrow 2} m_{5 \rightarrow 2} \quad (13)$$

$$\begin{aligned} M_{D \rightarrow 34} &\propto \sum_{1,2} f(\sigma_1, \sigma_2, \sigma_3, \sigma_4) M_{A \rightarrow 12} M_{B \rightarrow 23} M_{C \rightarrow 41} \\ &\times \frac{m_{A \rightarrow 1} m_{A \rightarrow 2} m_{B \rightarrow 2} m_{C \rightarrow 1}}{m_{1 \rightarrow 4} m_{2 \rightarrow 3}}, \end{aligned} \quad (14)$$

donde

$$f(\sigma_1, \sigma_2, \sigma_3, \sigma_4) = e^{\beta J(\sigma_4 \sigma_1 + \sigma_1 \sigma_2 + \sigma_2 \sigma_3)}.$$

Para implementar en la práctica el algoritmo se definen las estructuras de datos convenientes y se programan las ecuaciones de iteración (13) y (14) en todas las regiones de un sistema $L \times L$. Los mensajes se inicializan aleatoriamente y se actualizan al azar hasta que la variación relativa tras cada iteración sea inferior a 10^{-5} .

Hemos descrito ya la esencia del método CVM y del algoritmo GBP que corresponde a la misma selección de regiones. En esta sección se realiza el cálculo variacional exacto de la energía libre en la aproximación de Kikuchi para el FIB. Se parametrizan las distribuciones de probabilidad de cada región y se extrema la energía libre en el espacio de los parámetros.

Explotando las propiedades de simetría del sistema podemos parametrizar la energía libre mediante un conjunto reducido de magnitudes y luego proceder a su minimización. Para una plaqueta arbitraria proponemos una función de la forma

Explotando las propiedades de simetría del sistema podemos parametrizar la energía libre mediante un conjunto reducido de magnitudes y luego proceder a su minimización. Para una plaqueta arbitraria proponemos una función de la forma

$$b_p(\sigma_1, \sigma_2, \sigma_3, \sigma_4) = \frac{1}{16} \left(1 + m \sum_{i=1}^4 \sigma_i + r \sum_{\langle ij \rangle} \sigma_i \sigma_j + k(\sigma_1 \sigma_3 + \sigma_2 \sigma_4) + s \sum_{\langle ijk \rangle} \sigma_i \sigma_j \sigma_k + t \sigma_1 \sigma_2 \sigma_3 \sigma_4 \right),$$

de la cual se deduce, tras marginalizar por las variables adecuadas, que

$$b_{ij}(\sigma_i, \sigma_j) = \frac{1 + m(\sigma_i + \sigma_j) + r\sigma_i\sigma_j}{4}$$

$$b_i(\sigma_i) = \frac{1 + m\sigma_i}{2}$$

Mediante un cálculo muy sencillo puede demostrarse que m es igual a $\langle s \rangle$, es decir, la magnetización por espín. También se ve que $r = \langle s_i s_j \rangle$, es la correlación entre espines vecinos. La forma propuesta para las distribuciones de probabilidad es general dada la simetría del sistema. Un elemento clave es que estamos suponiendo que las distribuciones son las mismas para todas las regiones del mismo tipo y que el sistema es homogéneo e isotrópico. Esta simetría es la llave para la simplificación del cálculo, que de otra forma habría de considerar una distribución diferente para cada región del sistema, como pasa en los sistemas desordenados.

Si ahora escribimos la energía, entropía y energía libre intensiva del sistema en función de estas distribuciones de probabilidad, obtendremos una parametrización de la energía libre mediante las magnitudes m, r, s, k, t . Utilizando las ecuaciones (5)-(8) se obtiene la F_{Kik} parametrizada.

Es importante notar que la ecuación (8) da la energía libre extensiva del sistema; basta dividir esta magnitud (y todas las extensivas) por N , el número de espines de la red, para obtenerlas en su forma intensiva.

Los coeficientes R que aparecen en (8) para cada región son[5]

- Plaqueta: $c_p = 1$, no es subregión de ninguna región definida.
- Enlace: $c_E = 1 - 2 = -1$, cada enlace pertenece a dos plaquetas.
- Espín: $c_s = 1 + (-4) + 4 = 1$, cada espín es subregión de cuatro enlaces y cuatro plaquetas.

La energía media de un enlace se calcula según su definición

$$U_E = - \sum_{\sigma_i, \sigma_j} J \sigma_i \sigma_j b_{ij}(\sigma_i, \sigma_j), \quad (13)$$

$$U_E = -Jr.$$

Si se considera que cada plaqueta contiene 4 enlaces y se tienen en cuenta las condiciones de marginalización, es elemental obtener que

$$U_P = \sum_{ijkl} -J(\sigma_i \sigma_j + \sigma_j \sigma_k + \sigma_k \sigma_l + \sigma_l \sigma_i) b_{ijkl}(\sigma_i, \sigma_j, \sigma_k, \sigma_l), \quad (16)$$

$$U_P = 4U_E,$$

$$U_P = -4Jr.$$

De esta manera, queda para la energía intensiva total (haciendo $J=1$)

$$u_{total} = c_P U_P + 2c_E U_E, \quad (17)$$

$$u_{total} = -2r,$$

y para la entropía, muy similar

$$s_{total} = c_P S_P + 2c_E S_E + c_S S_S$$

$$S_R = - \sum_{\vec{\sigma}_R} b(\vec{\sigma}_R) \ln b(\vec{\sigma}_R) \quad \text{con } R = P, E, S \quad (18)$$

$$s_{total} = S_P - 2S_E + S_S.$$

Finalmente, la energía libre intensiva queda

$$\frac{F_{Kik}}{N} = f_{Kik} = u_{total} - T \sigma_{total}. \quad (19)$$

La minimización de $f_{Kik}(m, r, s, k, t)$ es matemáticamente sencilla, basta resolver el sistema de ecuaciones

$$\nabla_{mrskt} f_{Kik} = \vec{0} \quad (20)$$

en algún sistema de cálculo simbólico-numérico como *Mathematica*. En estas ecuaciones la temperatura entra como parámetro por lo que el resultado depende de los valores de T que se tomen. Por ejemplo, pueden tomarse equiespaciados en el intervalo $[0,4]$, que es donde se espera que ocurran fenómenos interesantes. Para cada valor de temperatura se obtienen los valores de m^*, r^*, d^* , etc., que satisfagan la condición de mínimo (20). De esta forma podemos encontrar una dependencia de estos

parámetros con la temperatura: $m^*(T)$, $r^*(T)$, $d^*(T)$, ..., $f_{Kik}(T)$.

RESULTADOS. ALGORITMO GBP. COMPARACIÓN

Se implementó el algoritmo BP y GBP en sistemas de 8×8 , 16×16 , 32×32 y 64×64 ; cada tamaño se corrió con 10^3 condiciones iniciales diferentes para promediar los resultados. Para cada valor de temperatura utilizado, se inicializa aleatoriamente el conjunto de mensajes y se iteran las ecuaciones (13) y (14) hasta alcanzar la convergencia. En la gráfica (3) aparece el promedio del logaritmo del tiempo de convergencia para cada tamaño. Es de interés señalar que en la vecindad de la temperatura donde ocurre una transición de fase, el tiempo que demora alcanzar la convergencia tiene un máximo, en correspondencia con el hecho de que la susceptibilidad y la longitud de correlación divergen a esta temperatura. Esto hace que las pequeñas fluctuaciones se propaguen fácilmente y que al algoritmo le cueste encontrar la solución correcta.

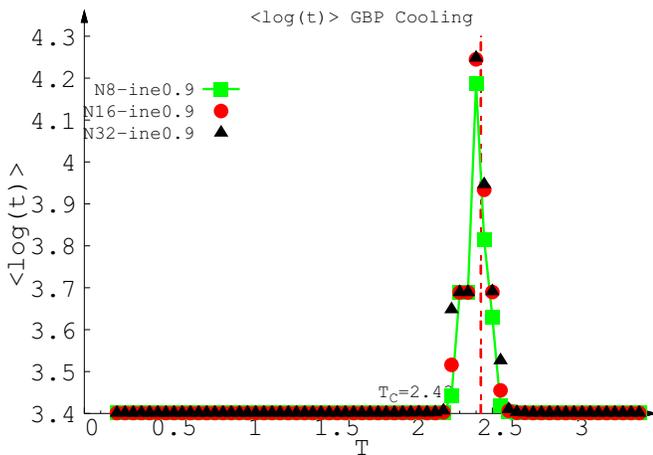


Figura 3: Logaritmo del tiempo de convergencia. Obsérvese que existe una divergencia para $T \approx 2.4J$, temperatura crítica en la aproximación de Kikuchi.

Los resultados de la magnetización, energía intensiva, entropía, etc., muestran independencia con el tamaño de la red utilizada. Esto es muy lógico, en fin de cuentas, las ecuaciones de iteración son las mismas en todos los sistemas y se obtienen siempre las mismas distribuciones de probabilidad. Las Figuras 4, 5 y 6 muestran en puntos redondos los valores obtenidos para diferentes magnitudes en dependencia de la temperatura.

En las mismas Figuras 4, 5 y 6 aparecen, en líneas continuas, las correspondientes magnitudes obtenidas por el procedimiento variacional. La coincidencia de ambos resultados es perfecta y no casual. Se ha demostrado que las distribuciones inferidas algorítmicamente se corresponden con aquellas que minimizan la energía libre en la aproximación correspondiente [5]. En la mayoría de los casos, por ejemplo en los sistemas desordenados, el cálculo variacional es imposible de realizar de la manera en que se hizo aquí. El principal motivo es que no se puede utilizar la simetría del sistema para lograr una simplificación. Los algoritmos del tipo de GBP son, en este caso, una

vía eficiente para encontrar el mínimo de F_{Kik} .

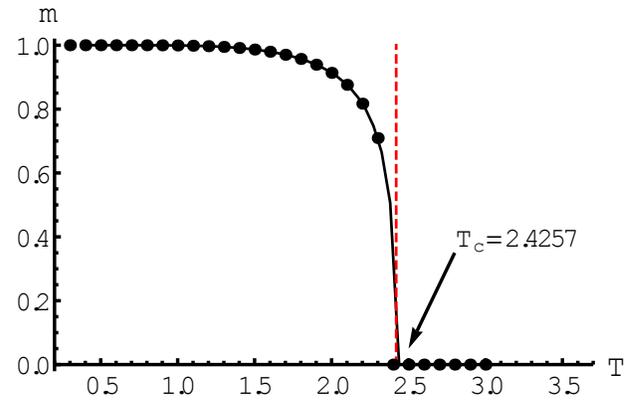


Figura 4: Magnetización vs. Temperatura. La T_C de transición predicha es $T_C \approx 2.426J$.

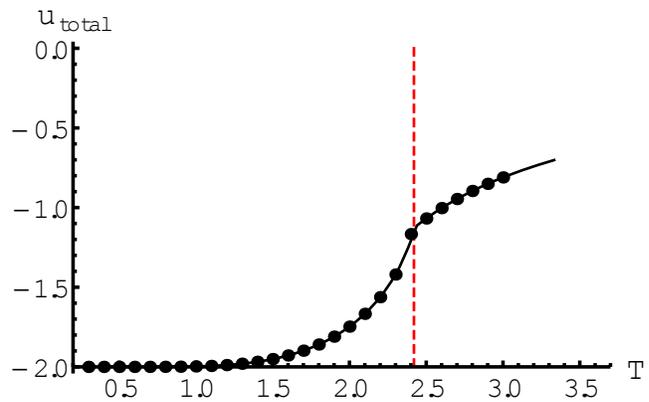


Figura 5: Energía intensiva vs. Temperatura. Como es natural, para $T \rightarrow 0$, $u_{total} \rightarrow -2$.

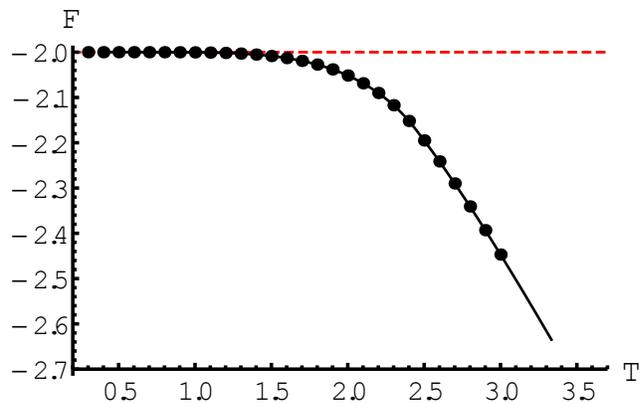


Figura 6: Energía libre vs. Temperatura. A baja temperatura tiende al valor de la energía mientras que para $T \rightarrow \infty$ tenemos que $F \rightarrow T \ln 2$.

Como se puede observar, por ambas vías se predice la existencia de una transición para-ferromagnética alrededor de una temperatura $T_C \approx 2.426J$. Este valor está mucho más cerca del correcto ($T_{Ons} \approx 2.269J$) que el obtenido mediante la aproximación de campo medio $T_{CM} \approx 4.0J$. Cuando se realizan los cálculos en la aproximación de Bethe (regiones enlace solamente), se obtienen cualitativamente los mismos

resultados, el valor de la temperatura crítica es un poco peor, $T_{Bethe} \approx 2.88J$. Si ordenamos los resultados quedaría $T_{CM} > T_{Bethe} > T_{Kik} > T_{Oms}$. Es posible seguir en esta serie de aproximaciones de modo tal que se obtengan resultados cada vez más cercanos a los exactos. Dentro del CVM podrían utilizarse como regiones primarias grupos de 3x3, 4x4, etc, espines.

En la Figura 5 se nota el cambio en la pendiente de la energía en la vecindad de C, como es característico de este tipo de transición.

CONCLUSIONES

Al utilizar la simetría del ferromagnético de Ising conseguimos realizar un cálculo exacto de su energía libre en la aproximación de CVM en condiciones de campo externo nulo. De esta manera obtuvimos un valor de la temperatura de transición que mejora la aproximación de campo medio y se acerca al exacto. Es importante señalar que el método variacional descrito es aplicable incluso con campo externo aplicado, donde no existe solución exacta conocida. Corroboramos también que existe una correspondencia exacta entre los resultados de los algoritmos de transmisión de mensajes del tipo GBP y el CVM con la correcta selección de regiones. Para el caso mucho más interesante de un sistema desordenado con interacciones $J_{ij} = \pm 1$ también se ha empleado GBP, donde permite describir una rica fenomenología [9].

Hay que mencionar que existen algoritmos de transmisión de mensajes, basados en otras ideas, que realizan la inferencia de manera exacta, por ejemplo, HAK, *Double-Loop*[1]. Sin em-

bargo estos algoritmos en general convergen más lentamente que GBP [3]. Por tanto, cuál usar para un problema específico está en dependencia del objetivo que se persiga: precisión o rapidez. También es posible utilizar algoritmos estocásticos como el método de Monte-Carlo para estimar las distribuciones marginales de probabilidad. Estos métodos estocásticos simulan la dinámica real del sistema en el espacio de configuraciones y requieren de un tiempo largo para samplearlo efectivamente. Los algoritmos vistos aquí son un artilugio matemático para encontrar rápida y aproximadamente los mínimos de la energía libre; sus resultados son válidos sólo cuando convergen.

-
- [1] T. Heskes, K. Albers and H. Kappen, Proceedings of the Nineteenth Conference Annual on Uncertainty in Artificial Intelligence (UAI-03), 313–320 (Morgan Kaufmann, San Francisco, CA, 2003).
 - [2] A. L. Yuille, *Neural Computation* **14**, 1691 (2002).
 - [3] A. Pelizzola, *J Phys A* **38**, R309 (2005).
 - [4] R. Kikuchi, *Phys Rev* **81**, 988 (1951).
 - [5] J. Yedidia, W. T. Freeman and Y. Weiss, *IT-IEEE* **51**, 2282 (2005).
 - [6] L. Onsager, *Phys Rev* **65**, 117 (1944).
 - [7] K. Huang, *Statistical Mechanics* (John Wiley and Sons, 1963).
 - [8] J. S. Yedidia, W. T. Freeman and Y. Weiss, *Intelligence* **8**, 236 (2002).
 - [9] A. Lage, R. Mulet, E. Domínguez, T. Rizzo and F. Ricci-Tersenghi, In preparation.