Содержание

Б. В. КРЫЖАНОВСКИЙ, Л. Б. ЛИТИНСКИЙ
Научно-исследовательский институт
системных исследований РАН, Москва
E-mail: litin@iont.ru

МЕТОД n–ОКРЕСТНОСТЕЙ И ВЫЧИСЛЕНИЕ СВОБОДНОЙ ЭНЕРГИИ

Аннотация

Методы статистической физики находят применение не только в физике, но и в задачах комбинаторной оптимизации [1], в теории нейронных сетей [2] и машинного обучения [3], при компьютерной обработке изображений [4]. Одна из центральных проблем здесь — вычисление свободной энергии, поскольку через нее выражаются такие макро-характеристики системы как намагниченность, теплоемкость, восприимчивость и т. д. Нами разрабатывается метод n-окрестностей для вычисления свободной энергии. Наш подход не зависит от специфики матрицы связи и позволяет включить в рассмотрение внешнее магнитное поле. Применение метода к модели Изинга на D-мерной кубической решетке дало простое аналитическое выражение для критических значений обратной температуры, хорошо описывающее компьютерный эксперимент для целого ряда размерностей: $D = 3, 4, 5, 6$ и 7. В докладе дается краткое введение в статфизический подход и излагается метод n-окрестностей.

B. V. KRYZHANOVSKY, L. B. LITINSKII
Institute of Optical Neural Technologies, RAS, Moscow
E-mail: litin@iont.ru

n-vicinity method and calculation of free energy

Abstract

Methods of statistical physics are used not in physics only but also in the combinatorial optimization [1], the theory of neural networks [2], machine learning [3] and computer processing of images [4]. In all these problems calculation of free energy is one of the key problems because it is necessary to calculate many important macroscopic characteristics of a system such as magnetization, heat capacity, susceptibility and so on. To calculate free energy we developed an n-vicinity method. Our approach does not depend on specifics of the connection matrix and allows us to introduce an external magnetic field. When applied to the Ising model on a D-dimensional cubic lattices, it gave a simple analytic solution for critical values of the inverse temperature that describes well computer simulations for a number of dimensions $D = 3, 4, 5, 6$ and 7. Here we explain briefly the statistical physics approach and discuss the n-vicinity method.

 УДК 001(06)+004.032.26 (06) Нейронные сети

Введение

Будем изучать систему, состоящую из большого числа N взаимодействующих спинов s_i:

$$s_i = \{\pm 1\}\,,\ i = 1, 2, \ldots, N,\ N \gg 1.$$

Взаимодействия между спинами задаются симметричной матрицей связи $\mathbf{T} = (T_{ij})$, самовоздействие отсутствует: $T_{ii} = 0$. Состояние системы как целого описывается конфигурационным вектором $\mathbf{s} = (s_1, \ldots, s_N)$. Если H — однородное внешнее поле, то энергия состояния в расчете на один спин равна:

$$E(\mathbf{s}, H) = -\frac{1}{2N}\sum_{i,j=1}^{N} T_{ij} s_i s_j - \frac{H}{N}\sum_{i=1}^{N} s_i. \tag{1}$$

Эффективный подход к описанию подобных систем большого размера доставляет статистическая физика. Центральная роль при этом отводится вычислению статистической суммы (*статсуммы*):

$$Z_N(\beta, H) = \sum_{all\ \mathbf{s}} \exp\left[-\beta E(\mathbf{s}, H)\right],$$

где $\beta = 1/T$ — обратная температура, а суммирование ведется по всем 2^N состояниям $\mathbf{s}$. Статистическая сумма есть функция обратной температуры и внешнего магнитного поля. Поскольку значение $Z_N(\beta, H)$ даже при $\beta = 0$ равно 2^N и при больших значениях $N \gg 1$ есть очень большое число, а интересуются термодинамическим пределом

$$Z(\beta, H) = \lim_{N\to\infty} Z_N(\beta, H),$$

удобнее работать со *свободной энергией*:

$$F(\beta, H) = -\frac{1}{\beta}\cdot\frac{\ln Z(\beta, H)}{N}. \tag{2}$$

Вслед за большинством исследователей мы будем использовать также *безразмерную* свободную энергию

$$f(\beta, H) = \beta\cdot F(\beta, H). \tag{3}$$

При $N \to \infty$ и $F(\beta, H)$, и $f(\beta, H)$ принимают конечные значения.

Важность свободной энергии определяется тем, что практически любая макрохарактеристика спиновой системы является производной от нее [5]. Например, среднее значение энергии при данной температуре (эту характеристику называют *внутренней энергией* U) есть

$$U(\beta, H) = \frac{\partial}{\partial \beta} f(\beta, H), \tag{4}$$

теплоемкость вычисляется как

$$C(\beta, H) = -\beta^2 \frac{\partial^2 f(\beta, H)}{\partial \beta^2},$$

намагниченность состояния есть

$$M(\beta, H) = -\frac{\partial F(\beta, H)}{\partial H}, \tag{5}$$

а магнитная восприимчивость равна

$$\chi(\beta, H) = -\frac{\partial^2 F(\beta, H)}{\partial H^2}.$$

Таким образом, чтобы изучать зависимость макрохарактеристик системы от внешних параметров задачи, необходимо уметь вычислять свободную энергию.

Важное место в статистической физике принадлежит изучению *фазовых переходов*. Поясним, что под этим понимается. Из самых общих соображений свободная энергия является непрерывной функцией своих аргументов. Но и не более того: производные свободной энергии могут иметь те или иные особенности — скачки и/или бесконечные разрывы. Пусть, например, первая производная свободной энергии по β при некотором значении β_c имеет скачок (для этого достаточно, чтобы на графике функции $F(\beta, H)$ при $\beta = \beta_c$ имелся излом). Тогда внутренняя энергия $U(\beta, H)$ в точке β_c меняется скачком — см. (4). Иными словами, для β меньших чем β_c, бесконечно малое приращение β приводило к бесконечно малому изменению внутренней энергии $U(\beta)$. А в точке β_c бесконечно малое приращение аргумента приводит к конечному скачку $U(\beta)$. Ясно, при $\beta = \beta_c$ у спиновой системы происходит какая-та внутренняя перестройка — либо меняется тип решетки, в узлах которой сидят спины, либо меняется характер распределения спинов и тому подобное. В точке β_c происходит глобальная перестройка спиновой системы. Говорят, что при переходе через *критическое значение* β_c система переходит из одного фазового состояния в другое.

Отыскание всех возможных фазовых состояний, их классификация и описание условий существования составляет предмет изучения спиновой системы. На формальном языке речь идет об исследовании особенностей свободной энергии. Если сингулярность имеется у первой производной $F(\beta, H)$ (как в рассмотренном случае) — говорят о фазовом переходе *первого рода*; если особенность имеет вторая производная $F(\beta, H)$ фазовый переход называют переходом *второго рода*. Фазовые переходы более высокого рода пока не наблюдались.

Вычисление свободной энергии, описание ее зависимости от внешних параметров β, H и др. — трудная задача, которую удается решить точно только для нескольких типов матрицы $\mathbf{T}$ $[1, 2, 5, 6]$. Как правило, при этом использовалась специфика матрицы связи. Мы развиваем подход к вычислению свободной энергии, не зависящий от конкретного вида матрицы связи [7–14]. Еще одно достоинство нашего подхода — возможность эффективного включения в рассмотрение магнитного поля H, что обычно составляет серьезную проблему. Кратко существо нашего подхода состоит в следующем.

Зафиксируем начальную конфигурацию (начальное состояние) $\mathbf{s}_0 \in \mathbf{R}^\mathbf{N}$, и определим ее n-окрестность Ω_n как множество состояний, отличающихся от $\mathbf{s}_0$ противоположными значениями n координат, где $n \in [0, N]$. Распределение энергий состояний, принадлежащих n-окрестности как правило неизвестно, однако можно получить точные выражения для среднего значения E_n и дисперсии σ_n^2 этого распределения $[8, 9]$. Эти два момента распределения выражаются через характеристики матрицы связи и начальной конфигурации $\mathbf{s}_0$. Далее, эмпирический факт состоит в том, что распределение энергий состояний из Ω_n является одномодовым, и хорошо аппроксимируется гауссовой плотностью с параметрами E_n и σ_n^2 (детальное обсуждение можно найти в [9]). Тогда в выражении для статистической суммы

$$Z_N = \sum\nolimits_{n=0}^{N} \sum\nolimits_{\mathbf{s}\in\Omega_n} \exp\left(-\beta E(\mathbf{s})\right)$$

суммирование по состояниям из Ω_n можно заменить интегрированием по гауссовой мере, а суммирование по n-окрестностям — интегрированием по непрерывной переменной $x = n/N$: $x \in [0, 1]$. После не очень сложных преобразований статистическая сумма принимает вид двойного интеграла от экспоненты, в показателе которой стоит большой множитель N:

$Z_N \approx \int dx \int dE \exp\left[-N \cdot \varphi(x, E)\right]$. Такие интегралы вычисляются стандартным методом перевала. Функция в экспоненте $\varphi(x, E)$ как от параметров зависит от обратной температуры β, внешнего магнитного поля H (если имеется) и характеристик матрицы связи. В общем виде выражение для $\varphi(x, E)$ получено в [8, 9].

Чтобы определить границы применимости нашего метода, мы исследовали с его помощью модель Изинга на D-мерной кубической решетке. За почти 100 лет изучения модели Изинга здесь получено много точных результатов, на которых можно тестировать новые методы и подходы. Основные полученные нами результаты таковы.

В предположении, что взаимодействие между ближайшими соседями вдоль различных направлений решетки может быть различным, получено уравнение состояния, которое естественным образом обобщает известное уравнение Брэгга-Вильямса [5].

Для решеток больших размерностей $D > 3$ нашим методом удается воспроизвести все известные для модели Изинга результаты: наличие фазового перехода второго рода и классические значения критических показателей [5]. Для критического значения обратной температуры получена неизвестная ранее простая аналитическая формула

$$\beta_c = \left(1 - \sqrt{1 - 2/D}\right)/2,$$

хорошо описывающая результаты компьютерного эксперимента для всех размерностей $3 \leqslant D \leqslant 7$, для которых проводилось компьютерное моделирование [15–18]. Согласие между нашей формулой и компьютерным экспериментом улучшается с ростом D.

Для двумерной модели Изинга ($D = 2$) наш подход дает приемлемое критическое значение обратной температуры, но предсказывает скачок намагниченности в критической точке, что является признаком фазового перехода первого рода. Это противоречит известному точному решению Онсагера, которое предсказывает здесь фазовый переход второго рода [5]. Для одномерной модели Изинга ($D = 1$) наш подход в нынешнем виде вообще неприменим.

Основные соотношения и D-мерная модель Изинга

Поначалу все рассмотрения проведем для матрицы связи $\mathbf{T} = (T_{ij})_1^N$ общего вида, а затем перейдем к матрице, отвечающей D-мерной модели

Изинга. Модель Изинга рассматривается на D-мерной кубической решетке с взаимодействием между ближайшими спинами и периодическими граничными условиями [5].

Метод n-окрестностей основан на том, что распределение энергий состояний из n-окрестности начальной конфигурации аппроксимируется гауссовой плотностью. Детальное обсуждение этого вопроса можно найти в [9]. Там же в самом общем виде приведены точные выражения для среднего значения энергии E_n и дисперсии σ_n^2.

В качестве начальной конфигурации возьмем $\mathbf{s}_0 = (1, 1, \ldots, 1) \in \mathbf{R}^N$. Ее n-окрестность Ω_n состоит из конфигураций, у которых в точности n спинов равны -1: $\Omega_n = \left\{\mathbf{s} : \sum_{i=1}^{N} s_i = N - 2n\right\}$, $n = 0, 1, \ldots, N$.

Через E_0 обозначим энергию начальной конфигурации $\mathbf{s}_0$ (без учета внешнего поля H):

$$E_0 = E(\mathbf{s}_0) = -\frac{1}{2N} \sum_{i \neq j}^{N} T_{ij}.$$

Если H — величина однородного внешнего поля, то энергия (1) состояния $\mathbf{s} \in \Omega_n$ равна

$$E(\mathbf{s}, H) = E(\mathbf{s}) - H \left(1 - \frac{2n}{N}\right).$$

В [8, 13, 14] показано, что асимптотическое выражение для статсуммы приводится к виду:

$$Z_N \sim \int_0^1 dx \int_{E_x(\min)}^{E_x(\max)} \exp\left[-N \cdot \varphi(x, E)\right] dE, \tag{6}$$

где $N \gg 1$, а функция в показателе экспоненты равна:

$$\varphi(x, E) = S(x) + \beta\left[E - H \cdot (1 - 2x)\right] + \frac{1}{2N} \left(\frac{E - E_x}{\sigma_x}\right)^2.$$

Здесь $x = n/N$ — непрерывный аналог дискретной переменной n, $S(x) = x \ln x + (1 - x)\ln(1 - x)$ — функция Шеннона, а $E_x = \lim\limits_{N \to \infty} E_n$ и $\sigma_x^2 = \lim\limits_{N \to \infty} \sigma_n^2$ суть асимптотические выражения для среднего E_n и дисперсии σ_n^2 (определены ниже).

Интеграл в выражении для Z_N вычисляется методом перевала, для чего необходимо минимизировать функцию $\varphi(x, E)$ по переменным x и E для каждой пары значений внешних параметров β и H. После несложных преобразований задача сводится к минимизации функции одной переменной

$$f(x) = S(x) + \beta E_x - \frac{\beta^2 N \sigma_x^2}{2} - \beta H(1 - 2x), \tag{7}$$

при условии $\beta N \sigma_x^2 + E_0 - E_x \leqslant 0$.

Введем новые обозначения, дающие более удобную нормировку:

$$b = \beta \frac{\sum_{ij} T_{ij}^2}{\sum_{ij} T_{ij}}, \quad q = \frac{\left(\sum_{ij} T_{ij}\right)^2}{N \sum_{ij} T_{ij}^2}, \quad h = \frac{2\sum_{ij} T_{ij}}{\sum_{ij} T_{ij}^2} H, \tag{8}$$

и перейдем к D-мерной модели Изинга, для чего воспользуемся асимптотическими выражениями для среднего значения энергии и дисперсии [9, 10]:

$$E_x = E_0(1 - 2x)^2, \quad \sigma_x^2 = \frac{8\sum_{ij} T_{ij}^2}{N^2} x^2 (1 - x)^2.$$

В окончательном виде задача принимает следующий вид: для каждого значения «безразмерной» обратной температуры b необходимо отыскать на интервале$(0, x_b)$ глобальный минимум функции

$$f(x) = L(x) - \frac{qb}{2}\left[(1 - 2x)^2 + 8b(x(1 - x))^2\right] - \frac{bh}{2}(1 - 2x), \tag{9}$$

где x_b =1/2, когда $b \leqslant 1$, и $x_b = \left(1 - \sqrt{1 - 1/b}\right)/2$, когда $b \geqslant 1$.

Свободная энергия $f(b, h)$ (3) равна:

$$f(b, h) = \min_{x \in (0, x_b)} f(x, b, h). \tag{10}$$

Сделаем несколько замечаний. Для D-мерной модели Изинга начальная конфигурация $\mathrm{s}_0 = (1, 1, \ldots, 1)$ является глобальным минимумом по энергии (*основным состоянием*). Такой выбор начальной конфигурации позволяет эффективно указать границы интегрирования по энергии $E_x(\min)$ и $E_x(\max)$ в выражении (6) для статсуммы Z_N.

Когда взаимодействия между ближайшими спинами одинаковы: $T_{ij} = (1 - \delta_{ij})J$, выражения (8) заметно упрощаются:

$$\sum\nolimits_{ij} T_{ij} = DNJ, \;\; \sum\nolimits_{ij} T_{ij}^2 = DNJ^2/2, \;\; b = \beta J, \;\; q = 2D, \;\; h = 2H/J.$$

Большую роль в дальнейшем играет параметр q, равный здесь числу ближайших соседей спина. Если взаимодействие вдоль различных направлений D-мерной решетки различно, параметр q не обязан принимать целочисленные значения. Например, для $2D$ модели Изинга с различными константами взаимодействия по вертикали (J) и горизонтали (K) имеем:

$$q = 2\left(1 + \frac{2}{K/J + J/K}\right).$$

Такое q может равняться любому числу из интервала $2 < q < 4$. Аналогично, для $3D$ модели Изинга с различными константами взаимодействия (J, K, L) получаем:

$$q = 2\frac{(J+K+L)^2}{J^2+K^2+L^2} \in (2, 6).$$

В общем случае *действительное* число q есть эффективное координационное число, характеризующее взаимодействие спина с его ближайшим окружением.

Наконец, положим $H = 0$ и будем решать задачу в отсутствие внешнего поля.

Уравнение состояния и его решение

Уравнение состояния

Для каждого значения b необходимо отыскать на интервале $[0, x_b]$ глобальный минимум функции

$$f(x) = L(x) - \frac{qb}{2}\left[(1-2x)^2 + 8b(x(1-x))^2\right],$$

для чего необходимо найти решения уравнения

$$\frac{\partial f}{\partial x} = \ln\frac{x}{1-x} + 2qb(1-2x)\left[1 - 4bx(1-x)\right] = 0. \tag{11}$$

Очевидно, что $x_0 = 1/2$ всегда является решением уравнения (11). Назовем решение x_0 *тривиальным*. Чтобы найти остальные решения, избавимся от тривиального решения, преобразовав уравнение (11) к виду:

$$\frac{\ln\frac{1-x}{x}}{2(1-2x)} = qb\left[1 - 4bx(1-x)\right]. \tag{12}$$

Уравнение (12) называется уравнением состояния. Обозначим через $l(x)$ его левую часть, а через $r(x,b)$ — правую:

$$l(x) = \frac{\ln\frac{1-x}{x}}{2(1-2x)}, \quad r(x,b) = qb\left[1 - 4bx(1-x)\right].$$

Функции $l(x)$и $r(x,b)$ ведут себя симметрично относительно точки $x_0 = 1/2$; будем изучать их поведение на интервале [0, 1/2]. Функция $l(x)$ монотонно убывает от $+\infty$ на левом краю интервала до значения, равного 1 на правом краю: $l(0.5) = 1$. Ее первая и вторая производные на правом краю равны $l'(x_0) = 0$ и $l''(x_0) = 8/3$ соответственно. Квадратный трехчлен $r(x,b)$ монотонно убывает от значения $r(0) = qb$ на левом краю до $r(x_0) = qb(1-b)$ на правом — см. нижние панели на рис.1. По мере возрастания параметра b у функции $f(x)$ будут возникать точки перегиба и/или локальные экстремумы — см. графики на верхних панелях рис. 1. Эти трансформации удобно изучать, анализируя взаимное расположение кривых $l(x)$ и $r(x,b)$.

Пока значение b невелико ($b \sim 0$), кривые $l(x)$ и $r(x,b)$ не имеют общих точек — см. нижнюю панель рис. 1a. В этом случае единственным решением уравнения состояния остается тривиальное решение $x_0 = 1/2$, а свободная энергия (10) имеет вид:

$$f(b) = f(x_0, b) = -\ln 2 - \frac{qb^2}{4}, \quad b \ll 1.$$

С ростом параметра b кривая $r(x,b)$ как целое поднимается вверх, приближаясь к кривой $l(x)$. При некотором значении b кривые $l(x)$ и $r(x,b)$ коснутся друг друга. Касание кривых может произойти либо во внутренней точке интервала $(0, x_0)$ (нижняя панель рис. 1b), либо на правом конце интервала, в точке $x_0 = 0.5$ (см. нижнюю панель рис. 1c).

Касание кривых $l(x)$ и $r(x,b)$ во внутренней точке интервала означает, что у функции $f(x)$ образовалась точка перегиба, в которой производная функции $f(x)$ равна нулю — см. верхнюю панель рис. 1b. Такая точка не является минимумом функции $f(x)$, поэтому данное решение уравнения (12) интереса не представляет. Однако стоит слегка увеличить значение параметра b, и вместо точки перегиба у функции $f(x)$ появятся два экстремума: локальный минимум x_1 и локальный максимум x_2 — см. верхние панели рис. 1b и рис. 1d. Новый локальный минимум является нетривиальным решением уравнения (11), и его необходимо сравнивать по глубине с минимумом в точке $x_0 = 1/2$.

В работах [13, 14] исчерпывающим образом проанализировано то, как решение уравнения состояния зависит от значения параметра q. Оказывается, имеются два критических значения $q_1 = 4\ln 2 \approx 2.77$ и $q_2 = 16/3$, и свойства решения уравнения (12) кардинально зависят от того, какому из трех интервалов принадлежит q: $q < q_1$, $q_1 \leqslant q < q_2$ или $q_2 \leqslant q$. Опуская детали, опишем полученные результаты.

Решение уравнения состояния для $q_2 \leqslant q$

В этом случае кривые $l(x)$ и $r(x,b)$ соприкасаются только на правом конце интервала, в точке $x_0 = 1/2$. Касание происходит, когда параметр b станет равен критическому значению

$$b_c = \frac{1-\sqrt{1-4/q}}{2}. \qquad (13)$$

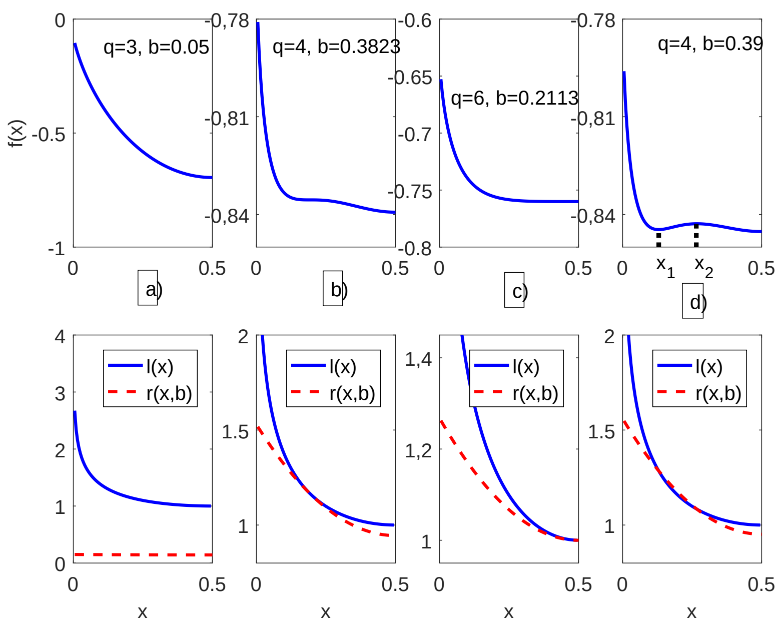

Рис. 1. Для различных значений q и b графики функции $f(x)$ (верхние панели) и кривые $l(x)$ и $r(x,b)$ (нижние панели)

В этот момент минимум функции $f(x)$ покинет точку $x_0 = 1/2$ и переместится в бесконечно близкую к ней точку $x_\varepsilon(b) = x_0 - \varepsilon(b)$. При этом произойдет фазовый переход второго рода. Дальнейшее увеличение обратной температуры будет сопровождаться стремлением точки минимума x_ε к левому концу интервала, в направлении к 0. При $b \to \infty$ точка глобального минимума $x_\varepsilon(b)$ асимптотически стремится к 0, никогда его не достигая. Так и должно происходить: спиновая система асимптотически стремится к основному состоянию, когда температура $T = 1/b$ стремится к 0.

Для модели Изинга больших размерностей $D = 3, 4, 5, 6, 7$ точные критические значения обратной температуры неизвестны, их оценки получены интенсивным компьютерным моделированием [15–18]. Эти данные мы использовали на рис. 2: сплошной линией показаны результаты компьютерного моделирования. Подставив в (13) значения $q = 6, 8, 10, 12, 14$ получим оценки для тех же величин нашим методом — на рис.2 эти данные показаны штриховой линией.

Точечной линией показана оценка $b_c = 1/2D$, получающаяся из уравнения Брэгга-Вильямса [5]. Видно, что наша формула описывает результаты компьютерного эксперимента много лучше. С ростом D согласие между теорией и компьютерным экспериментом улучшается. Напомним, что параметр q может принимать любые значения, не только целочисленные — см. замечание в конце предыдущего раздела.

Проводя стандартные вычисления, для $q \geqslant \frac{16}{3}$ можно получить значения критических показателей [5]. Для таких q критические показатели принимают у нас классические значения, характерные для модели среднего поля: $\alpha = 0, \beta = 1/2, \gamma = \gamma' = 1, \delta = 3$. Для $D > 3$ это совпадает с тем, что ранее было получено другими авторами. Для $D = 3$ общепринятыми являются неклассические значения критических показателей, приближенные значения которых получены методом ренорм-группы.

Решение уравнения состояния для $q_1 \leqslant q < q_2$

В этом случае касание кривых $l(x)$ и $r(x, b)$ происходит при некотором значении $b = b_t$ во внутренней точке $x_t \in (0, 1/2)$. При увеличении параметра b за значение b_t у функции $f(x)$ появляется локальный минимум $x_1(b)$, глубина которого растет с увеличением b, а сам минимум сдвигается в направлении начала координат (см. выше). При достижении некоторого критического значения $b = b_j$ глубина нового минимума станет равной глубине «тривиального» минимума в точке $x_0 = 1/2$. В этот момент произойдет перескок глобального минимума из x_0 в точку x_j, отвечающую критическому значению b_j ("j" от jump, т. е. прыжок). Данный фазовый переход сопровождается скачкообразным изменением намагниченности $m = 1 - 2x$, и является фазовым переходом первого рода — см. (5).

В [13, 14] получено трансцендентное уравнение, численное решение которого дает критическое значение b_t. Для $q = 4$ получаем, что $b_j \approx 0.3912$. В то же время, для планарной модели Изинга ($D = 2 \Rightarrow q = 4$) семьдесят лет назад Онсагером получено точное решение, которое дает фазовый переход второго рода и несколько большее критическое значение обратной температуры: $b_c \approx 0.4407$. Таким образом, для $q_1 \leqslant q < q_2$ к результатам, полученным методом n-окрестностей надо относиться с осторожностью.

С одной стороны, количественные результаты находятся на приемле-

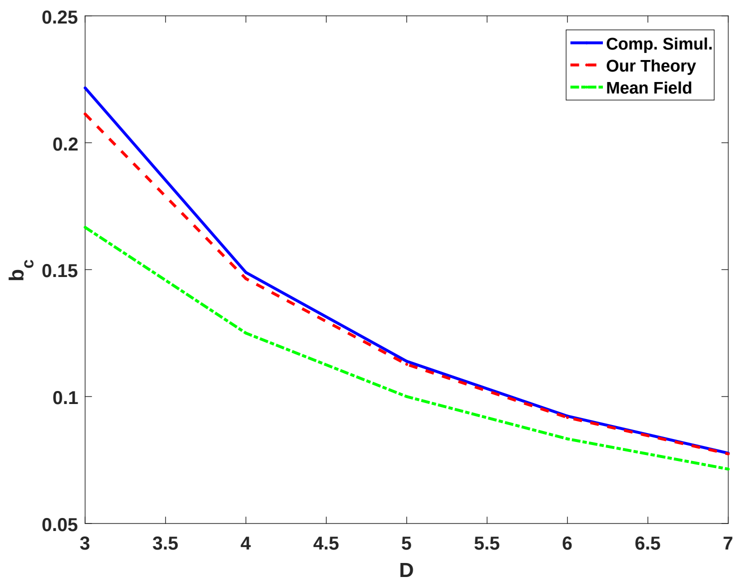

Рис. 2. Зависимость критической температуры b_c от размерности решетки D: компьютерное моделирование (сплошная линия), формула (9) (штриховая), приближение среднего поля (точечная)

мом уровне точности — различие между оценкой b_j и точным значением b_c составляет около 11%. На рис. 3 для планарной модели Изинга показана зависимость критического значения β_c от отношения констант взаимодействия по вертикали $J = 1$ и горизонтали (K меняется от 0.1 до 1). Мы видим, что наши результаты неплохо коррелируют с точным решением.

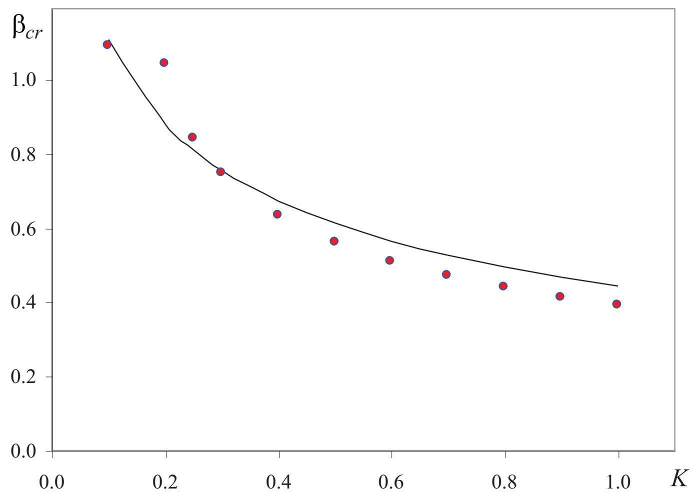

Рис. 3. Для $2D$ модели Изинга зависимость критической температуры от константы взаимодействия K при $J = 1$: сплошная кривая — теория Онсагера, кружки — наша теория

На рис. 4 приведены графики свободной энергии $f(\beta)$, построенные для различных значений K. Сплошной линией показана $f_{Ons}(\beta)$, вычисленная по теории Онсагера, а кружочками — то, что получается в нашем подходе: $f_{n-vic}(\beta)$. Видно, что отличие наших результатов от точной теории становится заметным только в районе критического значения обратной температуры, а относительная ошибка находится на уровне 1%. С другой стороны, смущает скачок намагниченности в момент фазового перехода. Этот скачок не предсказывается никакими теориями. Применение метода n-окрестностей для таких значений q требует дополнительных исследований.

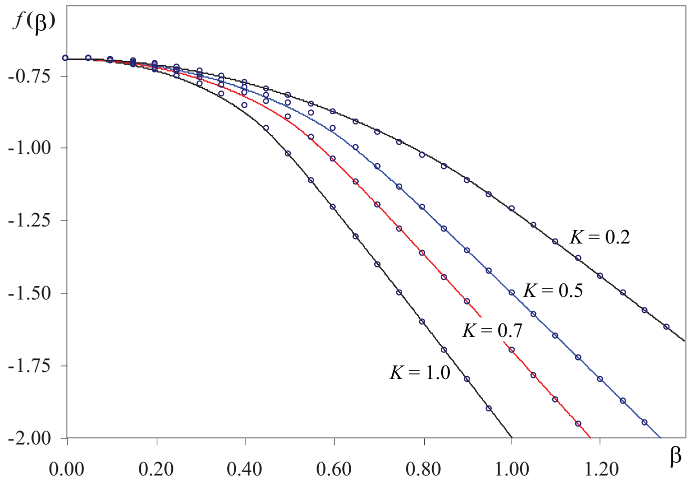

Рис. 4. Для $2D$ модели Изинга графики свободной энергии $f(\beta)$ при $K = 0.2$, 0.5, 0.7, 1.0: сплошные кривые — теория Онсагера, кружки — наша теория

Решение уравнения состояния для $q < q_1$

Для таких q метод n-окрестностей неприменим вовсе. Это следует из результатов работы [19], посвященной эффективному вычислению *полной спектральной плотности* модели Изинга[1]. Доказательство данного утверждения отложим до следующей публикации [20]. Здесь лишь констатируем, что неприменимость метода n-окрестностей в области $q < q_1$ является его ограничением.

Обсуждение и выводы

Основное приближение нашего метода состоит в том, что переходя от суммирования по n-окрестности Ω_n к интегрированию по Ω_x неизвестное распределение энергий $P_n(E)$ мы заменяем гауссовой плотностью со средним E_x и дисперсией σ_x^2. Центральная часть распределения $P_n(E)$ апрокси-

[1]Полная спектральная плотность — это функция распределения энергий всех 2^N состояний спиновой системы.

мируется при этом очень хорошо, но хвосты $P_n(E)$ всегда не гауссовы: для $D = 2$ и $D = 3$ они «толще» гауссовой плотности, для $D = 1$ напротив, тоньше. Эти различия на хвостах и являются источником ошибки.

В работах [13, 14] показано, что с ростом координационного числа q величина ошибки должна уменьшаться[2]. В то же время, аппроксимация центральной части распределения $P_n(E)$ гауссовым колоколом в целом передает какие-то существенные характеристики истинного распределения и дает разумные результаты по крайней мере, для значений q выше границы отсечения: $q > 4 \ln 2$.

Если говорить о физических приложениях метода n-окрестностей, то учет взаимодействия спинов со следующими соседями — или даже с более дальними соседями, — может привести к увеличению эффективного координационного числа q. Можно ожидать, что это только улучшит согласие результатов, полученных методом n-окрестностей с точным решением (либо компьютерным экспериментом).

Одной из областей применения нашего метода могут оказаться современные методики анализа сцен и компьютерной обработки изображений [4,21–23]. Здесь происходит многократное перевычисление так называемой *нормировочной константы*, которая и есть значение статистической суммы при фиксированных значениях внешних параметров. Затем значения внешних параметров меняются. Здесь востребованным оказывается быстрое — пусть даже и приближенное — вычисление статистической суммы, а критическая температура и характер фазового перехода никого не интересуют.

Безусловно, во всех приложениях перспективным является включение внешнего магнитного поля. Наши выражения (7), (9) позволяют рассчитывать на то, что это удастся сделать.

Работа выполнялась при финансовой поддержке Российского Фонда Фундаментальных Исследований (гранты 15–07–04861 и 16–01–00626).

Литература

1. *Martin O. C., Monasson R., Zecchina R.* Statistical mechanics methods and phase transitions in optimization problems // *Theoretical Computer Science*.– 2001. – **265**. – pp. 3–67.

[2]Это мы и наблюдаем, проверяя наш метод на модели Изинга.

2. *Amit D., Gutfreund H., Sompolinsky H.* Statistical mechanics of neural networks near saturation // *Annals of Physics*. – 1987.– **173**. – pp. 30–67.
3. *Hinton G. E., Osindero S., The Y.* A fast learning algorithm for deep belief nets // *Neural Computation*. – 2006. – **18**. – pp. 1527–1554.
4. *Wainwright M. J., Jaakkola T., Willsky A. S.* A new class of upper bounds on the log partition function // *IEEE Trans. on Information Theory*. – 2005. – **51**. – pp. 2313–2335.
5. *Бэкстер Р.* Точно решаемые модели статистической физики. – М.: Мир, 1985.
6. *Доценко В. С.* Универсальная случайность // *Успехи физ. наук*. – 2011. – v. 181 (3). – pp. 269–292.
7. *Kryzhanovsky B., Litinskii L.* Approximate method of free energy calculation for spin system with arbitrary connection matrix // *Journal of Physics: Conference Series*. – 2014. – v. 574. – pp. 012–017.
8. *Крыжановский Б. В., Литинский Л. Б.* Обобщенное уравнение Брэгга-Вильямса для систем с произвольным дальнодействием // *ДАН*. – 2014. – v. 459. – pp. 1–5.
9. *Kryzhanovsky B., Litinskii L.* Generalized approach to description of energy distribution of spin system // *Optical Memory & Neural Networks (Information Optics)*. – 2015. – v. 24. – pp. 165-185.
10. *Крыжановский Б. В., Литинский Л. Б.* Общий метод вычисления статистической суммы // *Сб. трудов XVII Всероссийской научно-технической конференции «Нейроинформатика-2015»*. – 2015. – **2**. – pp. 81–92.
11. *Крыжановский Б. В., Литинский Л. Б.* Обобщенный подход к описанию распределения энергий спиновой системы // *Труды НИИСИ*. – 2015. – v. 5(1). – pp. 5–21.
12. *Крыжановский Б. В., Литинский Л. Б.* Приближенное вычисление статистической суммы для модели Изинга на гиперкубе // *Труды НИИСИ*. – 2016. – v. 6(2). – pp. 5–10.
13. *Kryzhanovsky B., Litinskii L.* Applicability of n-vicinity method for calculation of free energy of Ising model // *Physica A: Statistical mechanics and its applications*. – 2017. – v. 468. – pp. 493–507.
14. *Крыжановский Б. В., Литинский Л. Б.* Метод n-окрестностей для вычисления статистической суммы в модели Изинга: границы применимости // *Труды НИИСИ*. – 2017. – v. 7(1). – pp. 4–17.
15. *Blote H. W. J, Shchur L. N, Talapov A. L.* The cluster processor: new results // *Int. J. Mod. Phys. C*. – 1999. – v. 10. – pp. 1137–1138.
16. *Aggkvist R. H., Rosengren A., Lundow P. H., Markstöm K., Andren D., Kundrotas P.* On the Ising model for the simple cubic lattice // *Advances in Physics*. – 2007. – v. 56(5). – pp. 653–755.

17. *Lundow P.H., Markstrom K.* The critical behaviour of the Ising model on the 4-dimensional lattice // *Phys. Rev. E*. – 2009. – v. 80. – pp. 031–104.

18. *Lundow P.H., Markstrom K.* The discontinuity of the specific heatfor the 5D Ising model // *Nuclear Physics B*. – 2015. – v. 895. – pp. 305–318.

19. *Крыжановский Б.В.* Вычисление спектральной плотности спиновой системы для точно решаемых моделей (в печати).

20. *Крыжановский Б.В., Литинский Л.Б.* Метод n-окрестностей и аппроксимация спектральной плотности (готовится к печати).

21. *Wang C., Komodakis N., Paragios N.* Markov random field modeling, inference & learning in computer vision & image understanding: A survey // *Preprint to Elsevier*. – 2013.

22. *Liu J., De Lamare R.* Low-latency reweighted belief propagation decoding for LDPC codes // *IEEE Communications Letters*. – 2012.

23. *Novikov A., Rodomanov A., Osokin A., Vetrov D.* Putting MRFs on a tensor train // In: *Proc. of the 31st Intern. Conf. on Machine Learning, Beijing, China. Journal of Machine Learning Research: W&CP*. – 2014. – v. 32.

Борис Владимирович КРЫЖАНОВСКИЙ, руководитель Центра оптико-нейронных технологий НИИСИ РАН (Москва), член-корреспондент РАН, доктор физико-математических наук. Президент Российской ассоциации нейроинформатики. Область научных интересов: нейронные сети, квантовая механика, нелинейная оптика. Имеет более 80 научных публикаций.

Леонид Борисович ЛИТИНСКИЙ, старший научный сотрудник Центра оптико-нейронных технологий НИИСИ РАН, кандидат физико-математических наук. Область научных интересов: ассоциативные нейронные сети, распознавание образов, многоэкстремальные задачи. Автор более 40 научных публикаций.