

Отчёт по проекту Решеточные модели макромолекул

Москаленко Р. Б.

24.05.2021

1 Введение

Рассмотрим конформацию (несамопересекающуюся последовательность узлов) на двумерной решётке. Такие конформации можно рассматривать как термодинамическую систему, основанную на модели Изинга, для которых существуют две фазы: плотная (глобулярная) и развёрнутая. Эти фазы соответствуют низким и высоким температурам системы.

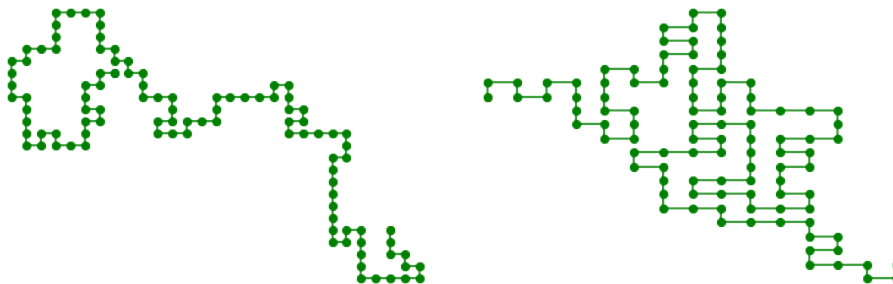


Рис. 1. Пример неплотной и плотной конформации

Если посмотреть на изображения конформаций каждого вида, хорошо видно, что плотные конформации по структуре близки с двумерным решёткам, где у каждого узла имеется множество соседей, и развёрнутые конформации наоборот близки к одномерным структурам, где узлы у которых больше 2 соседей встречаются редко. Соответственно можно предположить, что плотные конформации будут иметь свойства схожие с двумерными решётками, а развёрнутые с одномерными. В двумерных решётках наблюдается магнитный фазовый переход, в то время как в одномерных решётках переход не происходит. Цель данного исследования определить наличие магнитного перехода в плотных конформациях.

1.1 Модель

В данной модели мы рассматриваем ансамбли конформаций: множества конформаций одинаковой длины L , полученные при одинаковых температурах. Мы получаем конформации используя алгоритм SAW. На каждой из конформаций строится модель Изинга [2]. В каждой вершине размещается спин, который может принимать одно из двух значений: $+1, -1$. Гамильтониан данной системы имеет вид

$$H = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i$$

где i, j индексы соседних узлов, J - коэффициент взаимодействия h - воздействие внешнего поля.

Статистическая сумма

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

где $\{\sigma\}$ - множество всех возможных наборов значений спинов.

Намагниченность и энергия каждого состояния считаются по следующим формулам

$$E = -J \sum_{i,j} \sigma_i \sigma_j, M = \sum_i \sigma_i$$

Средняя намагниченность системы

$$\langle M \rangle = \frac{1}{Z} \sum_{\{\sigma\}} M e^{-H(\sigma)\beta}$$

1.2 Метод Монте-Карло

Для расчёта модели Изинга используется метод Монте-Карло. Были реализованы версии с односпиновым и кластерным апдейтом [3]. Код представлен в репозитории github [1]. В итоге для измерений используется кластерная версия. Благодаря отказоустойчивости она работает значительно быстрее, и быстрее сходится, особенно при низких температурах.

Алгоритм с кластерным апдейтом работает следующим образом. На каждой итерации мы выбираем случайный спин и начиная с него начинаем строить кластер из одинаково направленных спинов, добавляя новые спины в кластер с определённой вероятностью. затем мы меняем значения спинов в кластере на противоположные.

Чтобы вычислить намагниченность, мы сначала случайным образом инициализируем спины, затем делаем некоторое число шагов для отжига модели. Далее на каждом шаге мы измеряем намагниченность, и после выполнения определённого числа шагов, усредняем полученные значения. Так как средняя намагниченность равна 0, имеет смысл рассматривать модуль квадрат намагниченности.

$$\langle M^2 \rangle = \frac{1}{n} \sum_{\{\sigma\}} \left(\sum_i \sigma_i \right)^2, \langle |M| \rangle = \frac{1}{n} \sum_{\{\sigma\}} \left| \sum_i \sigma_i \right|$$

2 Алгоритмы

Реализованы два алгоритма обновления спинов. Односпиновый и кластерный апдейт. Оба алгоритма работают на произвольном графе, используя таблицу соседей. Алгоритмы реализованы как отдельные библиотеки для Python, и написаны с использованием технологии Cython для ускорения работы. Кластерный апдейт является более эффективным по времени работы и количеству шагов, которые необходимо выполнить для хорошей сходимости модели.

2.1 Проверка алгоритмов

Чтобы убедиться что алгоритмы работают правильно мы проверили, что оба алгоритма дают одинаковые результаты на одних и тех же конфигурациях, так же сравнил их с точными решениями для одномерной модели Изинга.

Результаты замеров кластерным и односпиновым апдейтом совпадают в пределах погрешности.

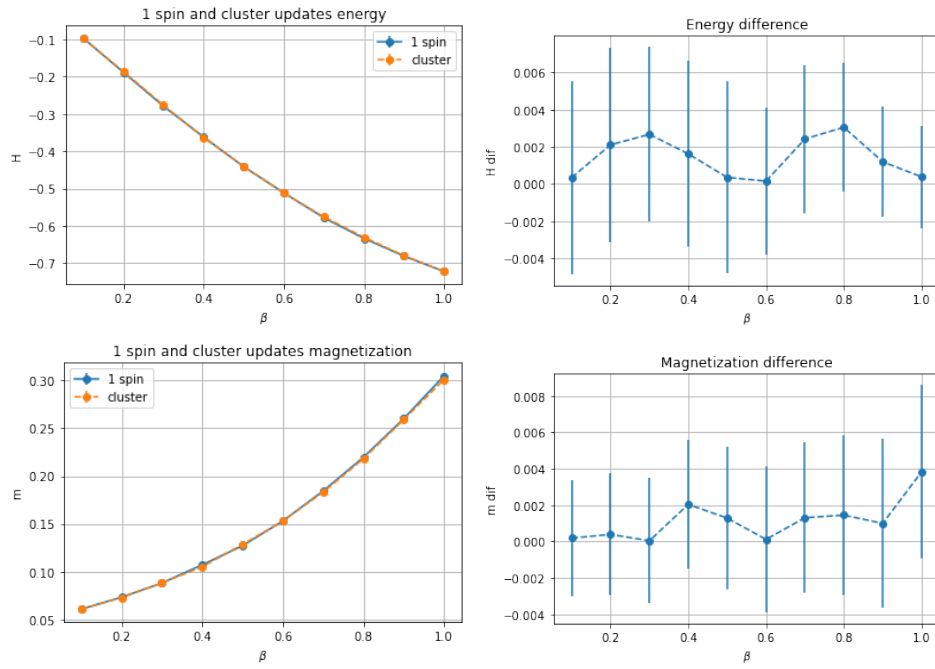


Рис. 2. кластерный и односпиновый апдейт

Для сравнения с точными значениями для одномерной модели Изинга, мы используем замкнутый квадратный контур. Данная конформация по свойствам полностью совпадает с одномерной моделью Изинга с открытыми граничными условиями.



(а) Конформация эмитирующая одномерную модель

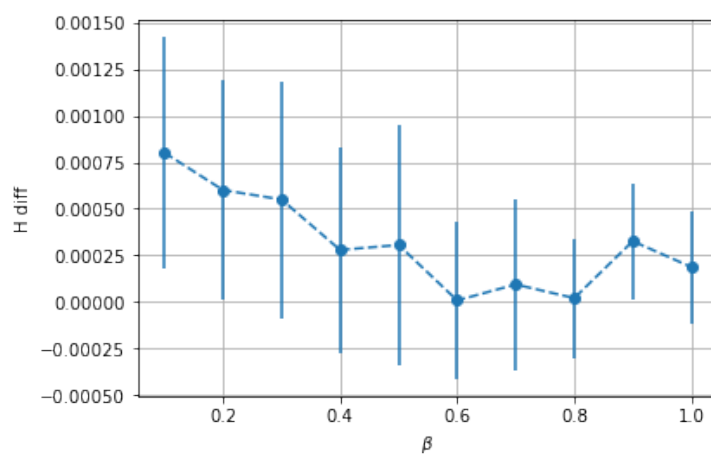
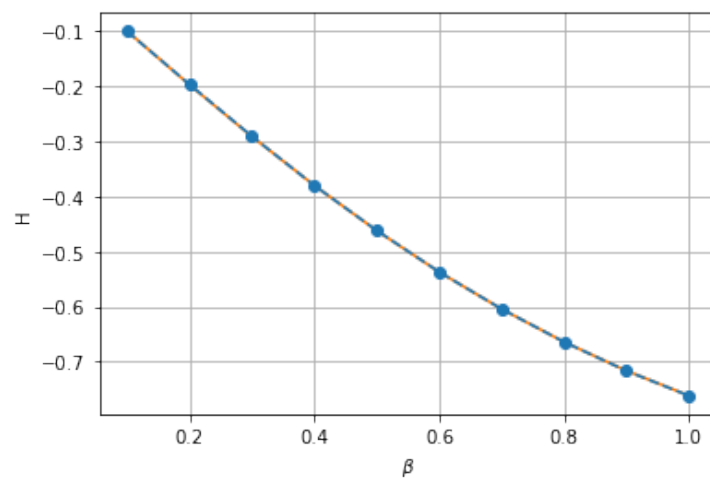


Рис. 3. Сравнение с точным решением одномерной модели

Так же был написан код, точно вычисляющий энергию системы путём полного перебора всех её состояний. Сравнение на маленьких конформациях (длина 10) даёт одинаковые результаты.

Примеры с использованием кластерного апдейта добавлены в библиотеку `mc_lib`.

3 Замеры

3.1 $U = 1$

Для вычислений были сгенерированы по 1000 реплик длины 250, 500, 1000, 2000. При моделировании методом Монте-Карло делалось 10000–30000 шагов на отжиг, и 50000–100000 шагов для замеров. Оказалось что достаточно большая часть этих конформаций неплотные, то есть их свойства ближе к свойствам одномерной решётки, чем двумерной. При попытке посчитать среднее значения кумулянта Биндера неплотные конформации Сильно влияли на значение кумулянта, увеличивая погрешность от реплики к реплике.

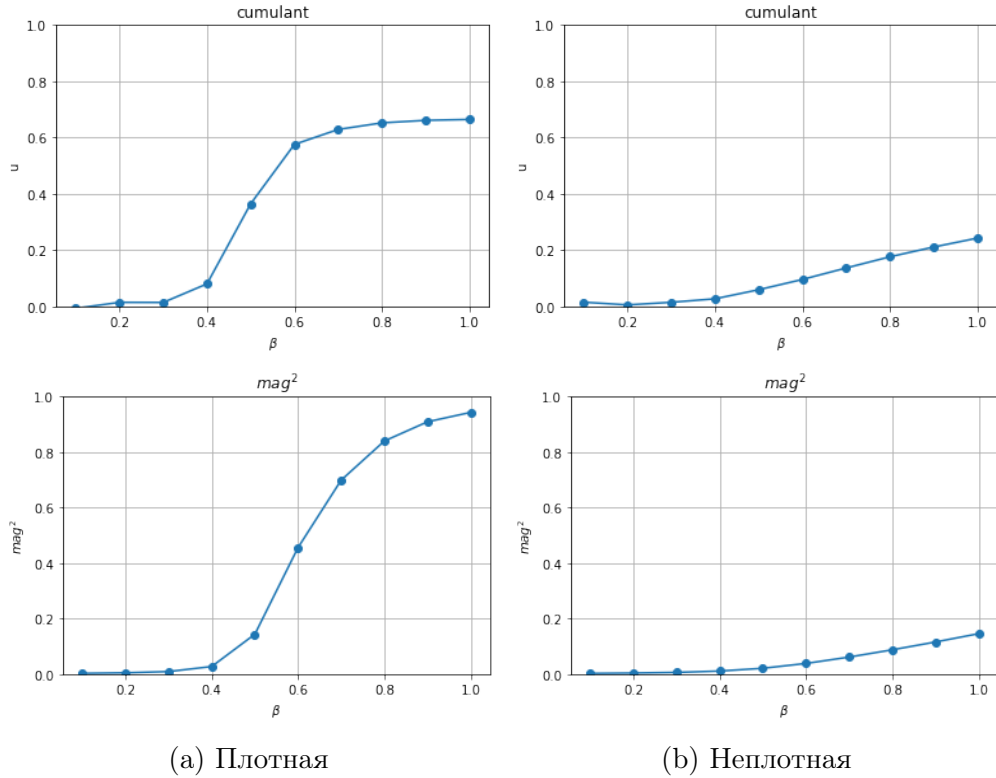


Рис. 4. Пример кумулянта и намагниченности плотной и неплотной конформаций

3.1.1 Разделение конформаций

Для отделения плотных конформаций от остальных было предложено вычислять их радиус инерции. $R = \sqrt{\frac{1}{n} \sum_{i=1}^n r_i^2}$, где r_i это расстояние от

узла конформации до её центра масс. Однако при рассмотрении большого количества конформаций оказалось, что маленький радиус инерции не гарантирует хорошую намагниченность конформации. Это хорошо видно при рассмотрении намагниченности конформаций при низких температурах $\beta = 1$

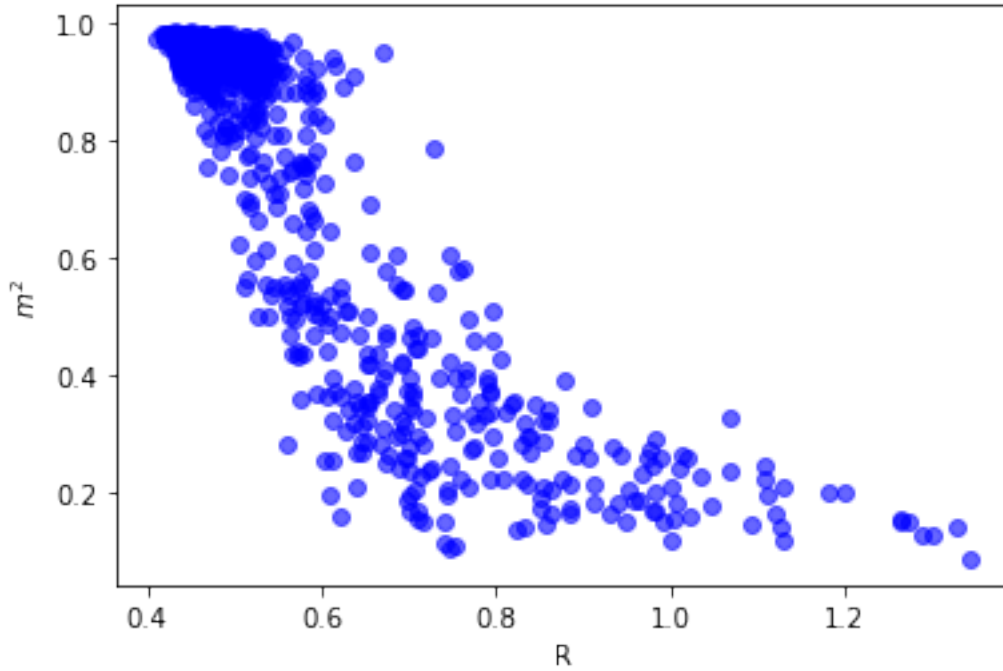


Рис. 5. Корреляция намагниченности конформаций при $\beta = 1$ и радиуса инерции для конформаций длины $L = 250$

На рис.5, при $R \approx 0.6$ m^2 принимают любые значения от 0.2 до 1.0. Значит, при разделении конформации только по радиусу инерции, мы либо будем отбрасывать намагничивающиеся конформации, либо оставлять не намагничивающиеся

Кластеризованные конформации На искусственном примере рис.6 показана одна из причин, по которой плотная конформация может плохо намагничиваться. Тут имеется несколько крупных двумерных кластеров, соединённых одномерной цепочкой. И не смотря на то, что сами по себе эти кластеры намагничиваются, направление спинов в них слабо связано, из-за чего спины в разных кластерах с большой вероятностью будут направлены в противоположные стороны. Далее соединяющие це-

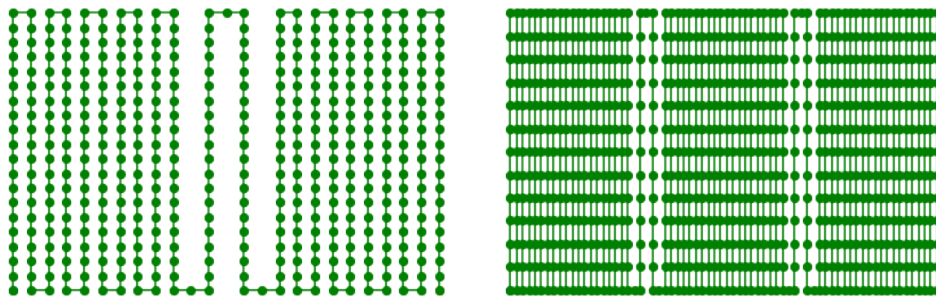
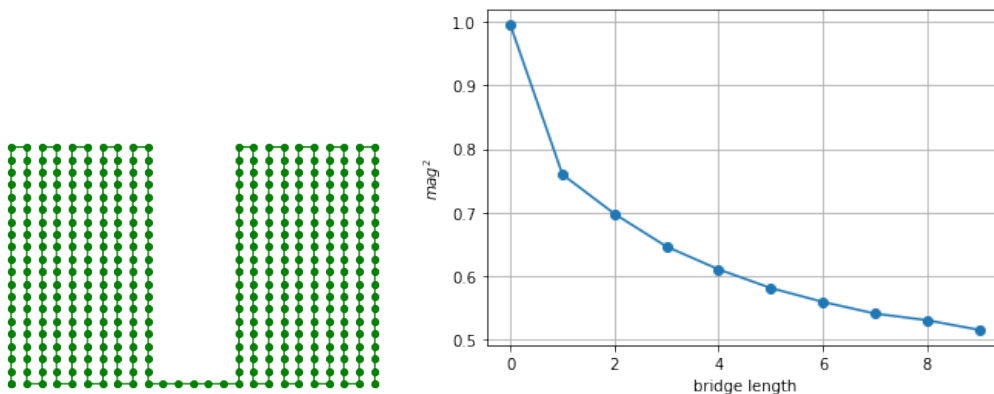


Рис. 6. Пример плотных немагнитных конформаций с двумя и тремя кластерами

почки будут называться мостами, длина моста - количество вершин, входящих в него.

На рис.6 приведён пример с очень длинным мостом, чтобы показать, что длина моста почти не влияет на радиус инерции конформации. Однако даже в конформациях с короткими мостами можно наблюдать большую разницу в намагниченности по сравнению с конформациями без мостов (рис.7).



(a) Пример с мостом длины 5 (b) График зависимости намагниченности от длины моста

Рис. 7. Значения намагниченности для конформации с двумя кластерами при различных длинах моста. Длина 0 означает, что рассматривалась конформация без моста, состоящая из одного кластера.

Было сделано предположение, что можно определять намагничивающиеся конформации используя кластеры и мосты. Следующей задачей стало проанализировать конформации на количество и размеры класте-

ров, а так же мостов. Однако пока мы не дали чёткого определения моста и кластера. Поэтому были рассмотрены несколько вариантов.

Первым вариантом было искать классические мосты – спины, при удалении увеличивается число компонент связности графа. Однако такой способ не дал желаемого эффекта, так как кластеры могут быть соединены более чем одним мостом. И например на конформации из рис. 8 данный способ не выделяет ни одного моста, хотя там очевидно есть структуры, отделённые друг от друга одномерными цепочками.

Следующий алгоритм выделял как мосты все цепочки спинов у которых 1 или 2 соседа, однако при таком подходе мы получаем мосты, которые соединяют один и тот же кластер. Такие мосты не разделяют кластеры и не оказывают на конформацию эффект описанный выше. Так же этим способом мы выделяем множество вершин на краях конформации как мосты, например вершины в углах прямоугольника будут считаться мостами, что очевидно неправильно.

Итоговая версия алгоритма выделяет как мосты все спины, которые имеют 1 или 2 соседа, и затем добавляет мосты, которые соединяют один и тот же кластер, к этому же кластеру. Таким образом мы оставляем только мосты, которые разделяют конформацию на отдельные плотные части, которые мы и называем кластерами. Данный алгоритм описан ниже.

Алгоритм разбиения на мосты и кластеры

1. Отметить все спины с 1 или 2 соседями как мосты.
2. Создаём массив, где отмечаем посещённые спины. Создаём массив где для каждого спина будем писать номер его кластера. И переменную отвечающую за текущую длину моста l . Изначально все спины не посещены, $l = 0$.
3. Начинаем идти по конформации от первой вершины.
 - (a) Если спин отмечен как мост, то увеличиваем l на 1
 - (b) Если спин не отмечен как мост, и не посещён. Увеличиваем счётчик кластеров на 1 и запускаем DFS(Алгоритм DFS описан ниже). Если $l > 0$ увеличиваем счётчик мостов на 1, длина нового моста $= l$. Обнуляем l
 - (c) Если спин не мост, уже посещён, последний встреченный спин, не являющийся мостом, принадлежит тому же кластеру и текущая длина моста $l > 0$. Значит этот мост соединяет один

и тот же кластер. Поэтому добавляем предыдущие l спинов к этому кластеру, обнуляем l .

(d) Если спин не мост, посещён, но номер кластера отличается от последнего встреченного кластера. Если $l > 0$ увеличиваем счётчик мостов на 1, длина нового моста = l . Обнуляем l

4. Проверяем первый и последний мост, если они соединяют один и тот же кластер, или один из их концов не соединён ни с каким кластером, добавляем их в кластер, с которым они соединены.

Алгоритм DFS

1. Заходим в вершину.
2. Отмечаем вершину как посещённую.
3. Отмечаем номер её кластера.
4. Увеличиваем счётчик размера текущего кластера на 1.
5. Заходим во все соседние не посещённые вершины не мосты.

В данном алгоритме мы пользуемся тем, что мосты обязательно образуются из подряд идущих вершин конформации. Поэтому чтобы определить соединяет ли мост один и тот же кластер, нам достаточно, идя по конформации, запоминать последний встреченный кластер и сравнивать его с новым.



Рис. 8. Пример реальных конформаций с маленьким радиусом инерции и маленькой намагниченностью, с отмеченными мостами

3.2 Статистика по кластерам и мостам

Ниже представлены гистограммы с размерами и числом кластеров и мостов в конформациях. Посчитано на 10000 конформациях с длинами 250, 500, 10000, конформации получены при $\frac{U}{T} = 1$. Размеры кластеров и длины мостов нормированы на длины конформаций.

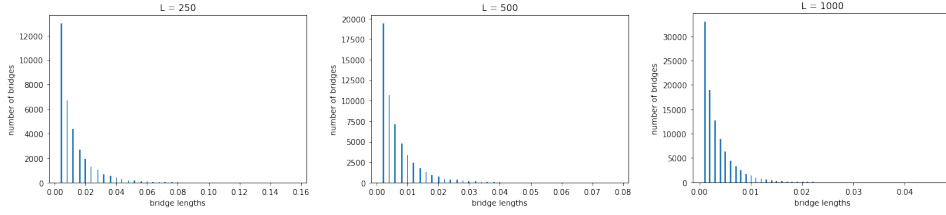


Рис. 9. Распределение длин мостов.

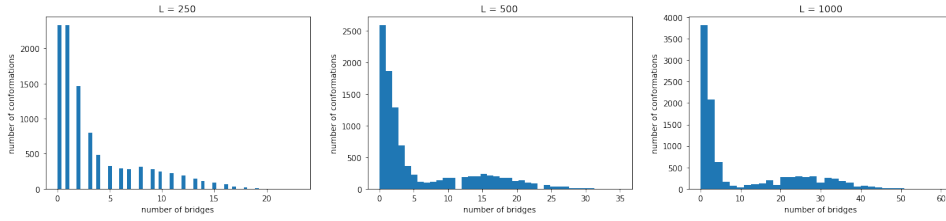


Рис. 10. Распределение числа мостов в конформации.

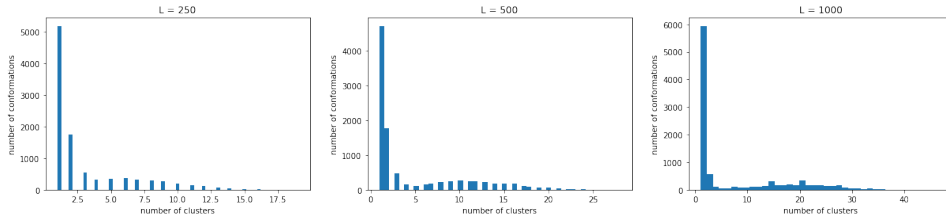


Рис. 11. Распределение числа кластеров в конформации.

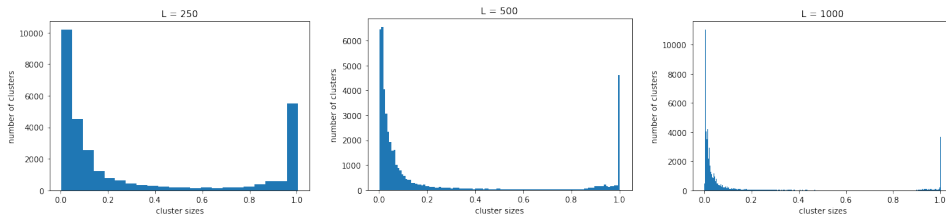


Рис. 12. Распределение размера кластеров.

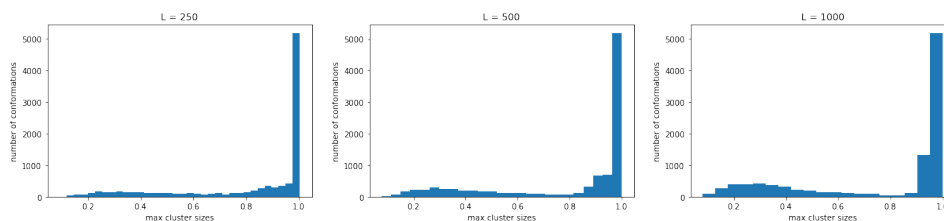


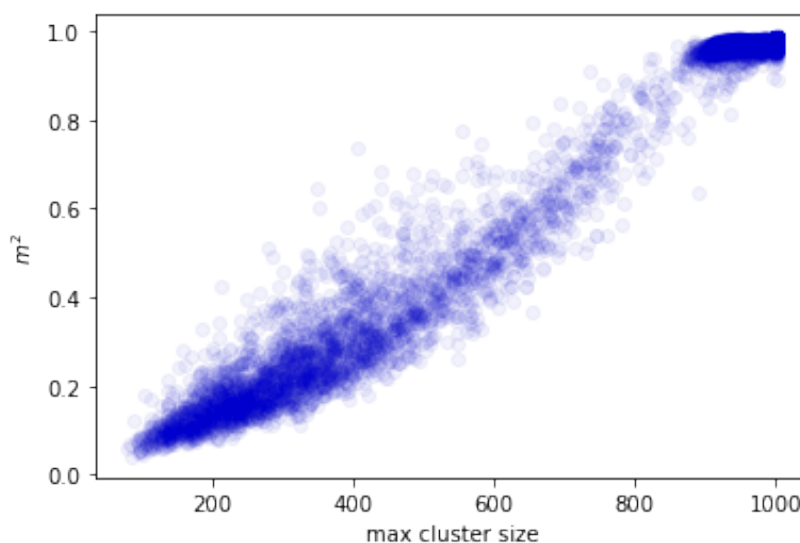
Рис. 13. Распределение размера наибольшего кластера в конформации.

3.3 Результаты разбиения на кластеры

Результаты анализа связи между намагниченностью и количеством и размерами кластеров и мостов подтверждают сказанное выше. У конформаций с большим числом кластеров обычно намагниченность ниже чем у конформаций с одним большим кластером.

Я рассмотрел несколько параметров: количество мостов, количество кластеров, суммарная длина мостов, размер наибольшего кластера. Наилучшим способом разделения конформаций на магнитные и немагнитные сейчас выглядит именно разделение по размеру наибольшего кластера. Как видно на рис.14 при разбиении по данному параметру разброс намагниченности значительно ниже, чем при разбиении по радиусу инерции. Данный параметр можно легко масштабировать для разных длин конформаций.

Рис. 14. График размера наибольшего кластера и квадрата намагниченности для 10000 конформаций длины 1000



Сравнение разделения по кластерам и по радиусу Чтобы оценить и сравнить эффективность разбиения конформаций при помощи размера наибольшего кластера и радиуса инерции воспользуемся следующим способом.

1. Зададим значение намагниченности μ , начиная с которого будем считать конформации намагниченными.
2. Из всех сгенерированных конформаций возьмём n конформаций с наименьшими радиусами инерции, и n конформаций с наибольшими размерами кластеров.
3. Среди выбранных конформаций посчитаем $k_{\mu,n}$ количество конформаций, намагниченность которых $< \mu$. Чем ниже это значение, тем лучше соответствующий способ разделения.
4. Повторяем предыдущие пункты для разных значений μ и n .

Сравнивая полученные значения $k_{\mu,n}$ можем определить, какой из способов эффективнее. На рис.15 видно как примерно ведут себя данные значения: до определённого n они равны 0, затем, дойдя до границы между магнитными и немагнитными конформациями, оно начинает расти, после чего рост становится линейным, так как все оставшиеся конформации не являются магнитными. Лучше разницу между двумя способами видно на рис.??, где, при всех значениях μ и n , разница $k_{\mu,n}$ остаётся отрицательной. То есть разделение по радиусу инерции всегда оставляет больше немагнитных конформаций, чем разделение по размеру кластеров.

3.4 Кумулянт и точка перехода

Кумулянт Биндера для одной реплики при заданной температуре вычисляется по формуле $U = 1 - \frac{\langle m^4 \rangle}{3\langle m^2 \rangle^2}$. Дальше значения усредняются между репликами при каждой температуре $\langle U \rangle = \frac{1}{n} \sum_{i=1}^n U_i$. Погрешность кумулянта от реплики к реплике вычисляется как среднеквадратичное отклонение по формуле $\sqrt{\frac{1}{n} \sum_{i=1}^n (\langle U \rangle - U_i)^2}$

Как видно на рис.17 вычисление кумулянта на всех сгенерированных конформациях даёт слишком большие погрешности от конформации к конформации, из-за этого становится невозможно определить точку перехода.

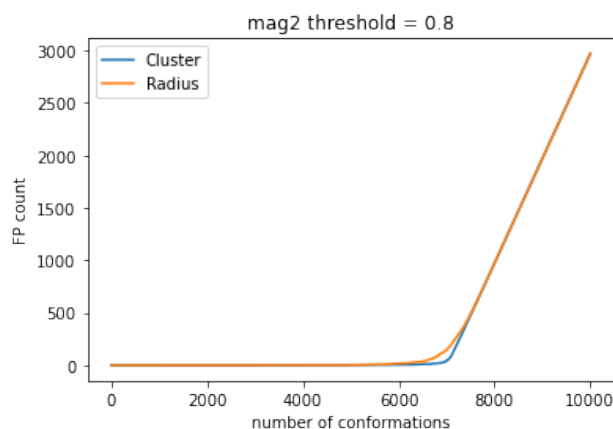


Рис. 15. График $k_{\mu,n}$ для разделения по кластерам и по радиусу, при $\mu = 0.8$.

Описанный выше способ разделения конформаций на магнитные и немагнитные должен позволить уменьшить погрешность при вычислении кумулянта. Чтобы подобрать значение параметра (размер наибольшего кластера), при котором будет происходить разделение, мы стали перебирать значения, и следить за поведением точки пересечения. Предположительно при увеличении параметра точка пересечения должна двигаться в сторону нуля, и начиная с определённого значения она должна остановиться.

Точка пересечения в данном случае вычисляется путём генерации отрезков по точкам замеров: рядом с предполагаемой точкой пересечения берём несколько замеров, в них генерируем точки из нормального распределения, распределение построено по значению и погрешности кумулянта.

Эксперименты с разделением конформаций по радиусу инерции показали, что таким образом можно выделить наборы конформаций разных длин так, что для них будет возможно найти точку перехода.

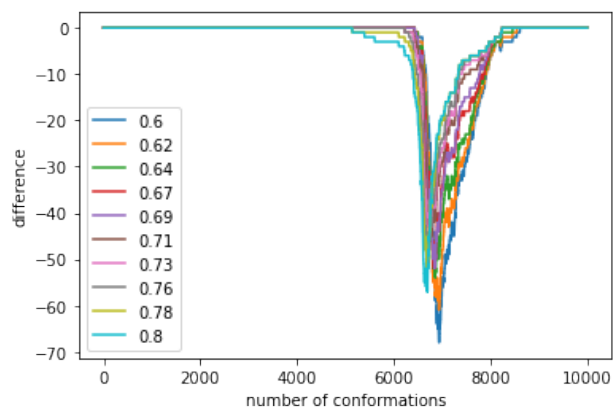


Рис. 16. График разности: $k_{\mu,n}$ при кластерном разделении и $k_{\mu,n}$ при разделении по радиусам. На 10000 конформаций длины 1000. Каждая линия соответствует одному значению μ .

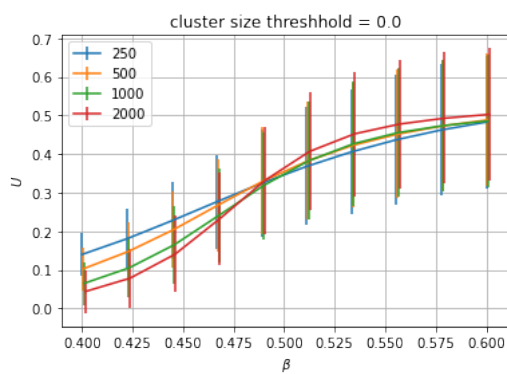


Рис. 17. кумулянты построенные на всех полученных конформациях

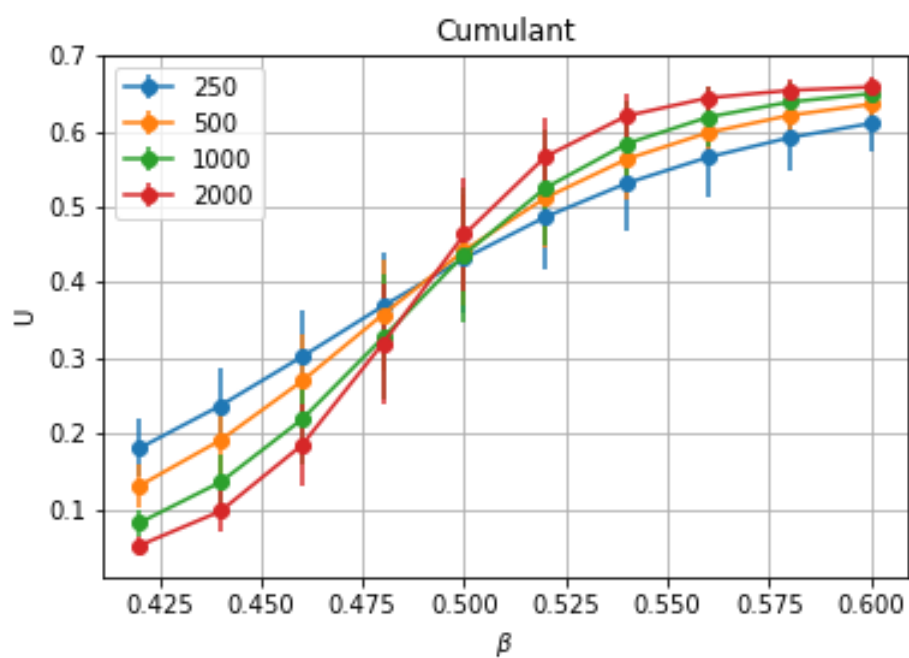


Рис. 18. Значения кумулянтов после отбора конформаций по радиусам инерции

Список литературы

- [1] Ссылка на репозиторий:<https://github.com/MoskalenkoRomanBorisovich/Ising-on-random-conformation>
- [2] Ernst Ising, Contribution to the Theory of Ferromagnetism, 1925
- [3] U. Wolff, Collective Monte Carlo Updating for Spin Systems. Physical Review Letters. 62 (4): 361–364, 1989
- [4] K. Binder, Finite size scaling analysis of ising model block distribution functions. Zeitschrift für Physik B: Condensed Matter. 43 (2): 119–140, 1981