

MD-моделирование воды

Поляченко Юрий

726

20 мая 2020 г.

1 Метод

Прежде чем описывать результаты расчетов, скажем о методе моделирования. Моделирование проводится в `lammps` при $T = 25\text{ }^\circ\text{C}$, $P = 1\text{ atm}$.

1.1 Единицы измерения

Используются `real` единицы. Чтобы сравнивать с результатами эксперимента, известными с СИ, надо перевести все в СИ. Были заданы следующие (нетривиальные) коэффициенты перевода

- $1\text{ atm} = 101325\text{ (Pa)}$ (в отличие от $1\text{ bar} = 10^5\text{ Pa}$)
- $1\text{ Kcal/mole} = 4184/N_a\text{ (J)}$
- $N_a = 6.02 \cdot 10^{23}\text{ (1/mole)}$

1.2 Расчет погрешностей

1.2.1 Погрешности средних

Дампы можно писать с любой частотой, но это не значит, что все эти «измерения» будут независимы и могут считаться для увеличения N в известной оценке $\sigma_x/\langle x \rangle \sim 1/\sqrt{N}$.

За число независимых измерений, содержащихся на траектории, предложено считать $N = t_{tot}/t_{mem}$, где t_{mem} - время потери корреляции между точным решением задачи коши и нашим численным интегрированием. Оно может быть оценено. Мы делали это для Л-Дж системы, и там получалось $t_m^d \sim 6 - 8\text{ ps}$. Здесь для верности я считал $t_m^d = 20\text{ ps}$. Тогда погрешность измеряемой величины $f(x_i, v_i, t_k)$ будет (при условии установившегося равновесия)

$$\sigma_f = \frac{\sqrt{\langle(\Delta f)^2\rangle}}{\sqrt{N_{uncor}}} = \frac{\sqrt{\langle(f_k - \langle f \rangle)^2\rangle}}{\sqrt{t_{tot}/t_m^d}} \quad (1)$$

1.2.2 Погрешности ошибок и флуктуаций

Хорошо бы уметь оценивать погрешность стандартного отклонения величины, т.к. `std` входит в теплємкость.

Оценка погрешности ошибки величины явно сложнее чем оценка погрешности самой величины. То, что здесь написано, придумано мной и ни где не доказано (по крайней мере я о такой не слышал).

Можно построить гистограмму попадания измеряемой величины $f(x)$ в интервалы $[y_i; y_i + dy]$. По нашему предположению оно будет гауссовым. Мы можем пофитить полученную гистограмму гауссовым распределением (μ, σ) и пара (μ_0, σ_0) , дающая наименьшее отклонение от гистограммы, будет теми числами, которые мы и указываем обычно в ответе $\langle f \rangle = \mu_0 \pm \sigma_0/\sqrt{N}$. На самом деле это уже не очень определенный алгоритм, т.к. не совсем ясно, что брать за «отклонение» от экспериментальной гистограммы. Обычно мы берем сумму квадратов отклонений, деленных на ошибки, но это не просто потому, что так удобно считать, а потому, что мы предполагаем, что истинное значение каждого конкретного измерения тоже распределено по гауссу, и мы на самом деле максимизируем вероятность того, что при истинных параметрах (μ_r, σ_r) экспериментальные точки лягут так как они легли на эксперименте. Это дает произведение экспонент, у каждой из которых в показателе стоит квадрат отклонения, поэтому максимизация вероятности реализации имеющегося события равносильно минимизации суммы квадратов отклонений.

Тут тоже надо бы пытаться максимизировать вероятность того, что при истинных параметрах измеряемой величины (μ_r, σ_r) реализуется та гистограмма, которая у нас получилась. Найдем вероятность того, что при гауссовом (μ, σ) распределении величины x реализуется наш случай. Для конкретного интервала $[x_i; x_i + dx]$ вероятность того что измеряемое значение попало в него именно k_i раз будет

$$P_i = p^{k_i}(x_i)(1 - p(x_i))^{(N-k_i)} C_{k_i}^N, \quad (2)$$

где $p(x) = f(x) \cdot dx = \exp(-(x - \mu)^2/2\sigma^2)/\sqrt{2\pi}\sigma \cdot dx$

Вероятность того, что в другом интервале произошло ровно k_j попаданий при том, что в i -ом все еще k_i , будет $P(k_j|k_i) = (P_j \cdot P_i)/P_i = P_j$ (по $P(A|B) = P(A \cap B)/P(B)$). Тогда вероятность реализации всей гистограммы одновременно

$$P(\mu, \sigma) = \prod_{i=1}^M P_i = \prod_{i=1}^M C_{k_i}^N \left(\frac{e^{-(\Delta x_i)^2}}{\sqrt{2\pi}\sigma} \right)^{k_i} \left(1 - \frac{e^{-(\Delta x_i)^2}}{\sqrt{2\pi}\sigma} \right)^{N-k_i}, \quad \Delta x_i = \frac{x_i - \mu}{\sigma\sqrt{2}}$$

Я не проверял, но было бы логично, чтобы $P(\mu, \sigma)$ имела бы максимум в $P(\langle \{x\} \rangle, std[\{x\}])$. Аппроксимировав $P(\mu, \sigma)$ до 2 порядка гессианом в пространстве (μ, σ) вблизи максимума, по идее можно оценить погрешность $\Delta(std[\{x\}])$.

Все описанное звучит логично для меня, но оно относительно сложно в исполнении, да и гарантий того, что ответ будет верный, нет. Момент

$P(k_j|k_i) = P_j$ особенно сомнителен. К тому же в работе есть гораздо более серьезные проблемы, чем отсутствие обоснования погрешностей. Поэтому эта процедура не реализовывалась.

1.3 Получение искомых величин из расчета

Плотность рассчитывается ламмписом автоматически. Для теплоемкостей известна их связь с тепловыми флуктуациями в системе:

$$C_V = \frac{\langle (\Delta_{nvt} E_{tot})^2 \rangle}{k_B T^2} \quad (3)$$

$$C_P = \frac{\langle (\Delta_{npt} (E_{tot} + PV))^2 \rangle}{k_B T^2} \quad (4)$$

В обоих случаях (nvt и npt) используется термостат Нозе-Гувера. Силы термостатов выбраны типичными согласно документации ламмписа. $\tau_P = 1000dt$, $\tau_T = 100dt$. Их вариации проводились – при слишком сильном термо- и баро-статировании система ведет себя неадекватно. При выходе же на стабильный режим получаемые ответы слабо зависят от силы ...статов. Используется значение по умолчанию $dt = 1$ ps. Проверено, что больше ($dt \geq 2$ ps) брать плохо – получаемые значения меняются, а меньше бессмысленно ($C_X(dt = 1ps) = C_X(dt = 0.5ps)$).

1.4 Методы работы с кулоном и потенциалы

Используется 3 ламмпис-потенциала для воды

- lj/cut/coul/long + ppm 1.0e-5
- lj/cut/coul/dsf, $\sim 1/R_{dump}$, 12.0
- lj/cut/tip4p/long ... 0.15 12.0 + ppm 1.0e-5

Параметры молекулы SPC/E: $q_O = -0.8476$ e, $q_H = 0.4238$ e, $d_{OH} = 1$ Å, $\angle(HOH) = 109.4^\circ$. Для удержания заданных связей использовался rattle с точностью 10^{-5} и максимальным количеством последовательных итераций для удовлетворения связи = 100. Параметры рекомендованы документацией. Точность – относительная ошибка в силах. Во всех случаях параметры Л-Дж берутся по документации ламмписа $\varepsilon = 0.1550$ Kcal/mole, $\sigma = 3.1536$ Å. В начале было проверено, что отличие ответов при этих параметрах от ответов в SPC/E модели незначительны (хотя

не очевидно, с чем сравнивать). Л-Дж взаимодействие вводится только между атомами кислорода.

Перейдем к результатам. Анализ верности ответа будет проводиться сравнением ответов друг с другом, т.к. сойтись с экспериментом не получилось.

2 Результаты RPRM

2.1 Допустимый размер системы

Как видно, делать линейный размер системы больше 5 межатомных расстояний не сильно необходимо (опять же, с чем мы сравниваем и какая точность нужна - не ясно. Было бы сравнение с экспериментом – тогда да, а так не ясно). Брать меньше 5 не хорошо, т.к. область обрезки Л-Дж начинает приближаться и к периодической границе, что явно плохо.

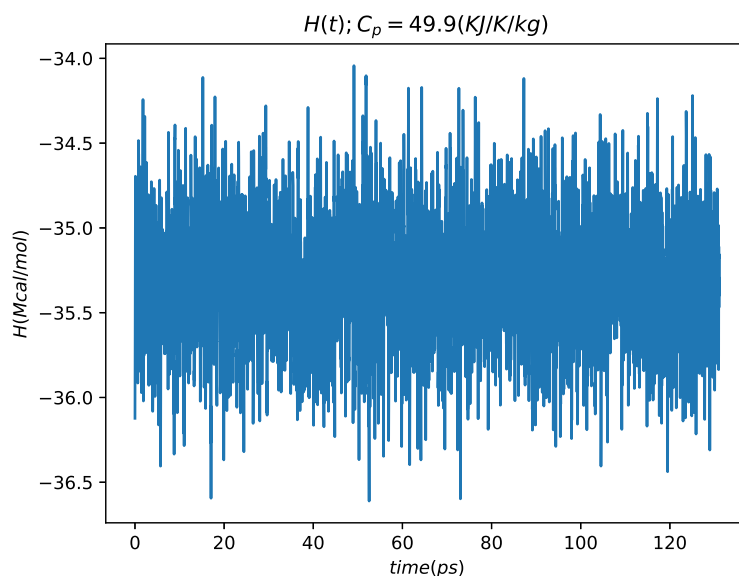


Рис. 1: rprm, $N = 15^3$

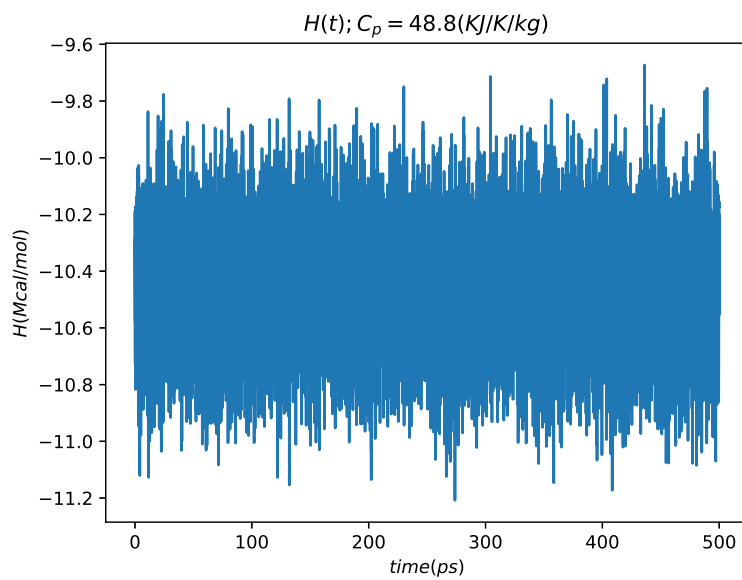


Рис. 2: pppm, $N = 10^3$

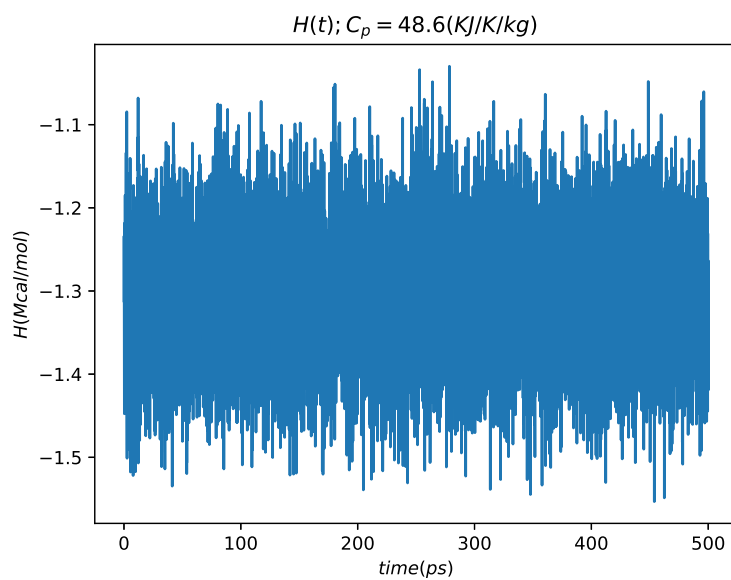


Рис. 3: pppm, $N = 5^3$

2.2 Другие равновесные характеристики

Были измерены так же плотность, давление, полная энергия и температура системы. C_V в заголовке $E_{tot}(t)$ посчитано в предположении pvt, поэтому не является релевантным, т.к. по факту расчет был prt.

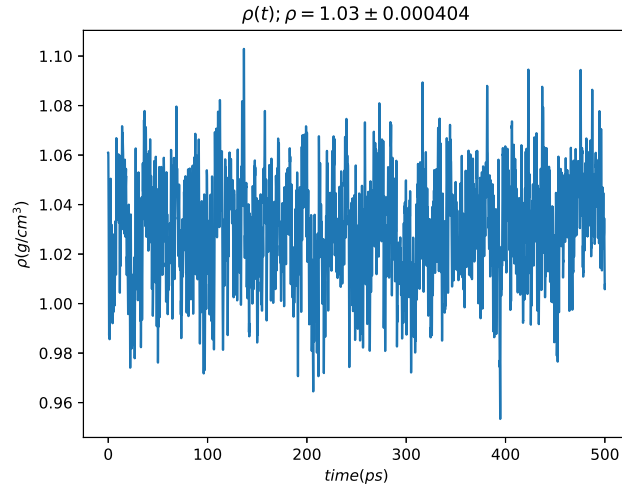


Рис. 4: pprn, $N = 5^3$

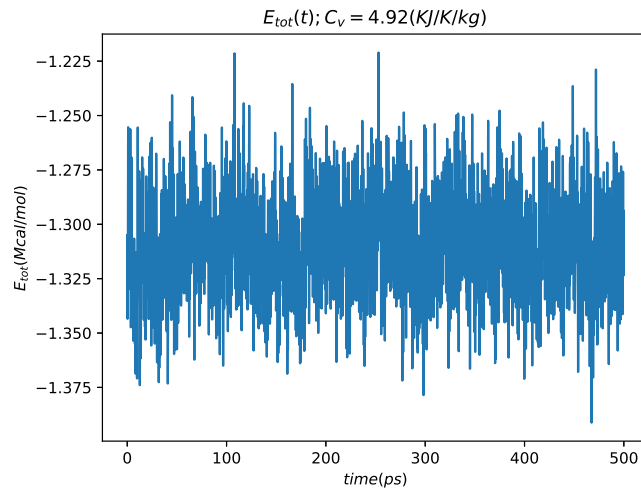


Рис. 5: pprn, $N = 5^3$

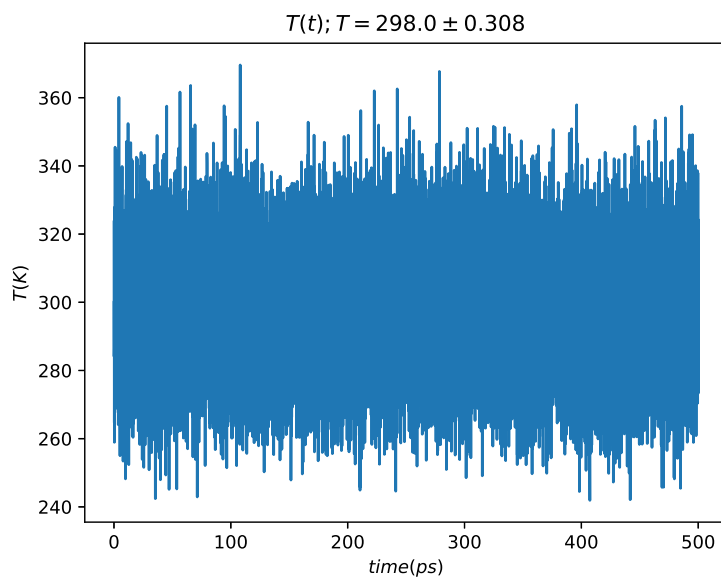


Рис. 6: pprn, $N = 5^3$

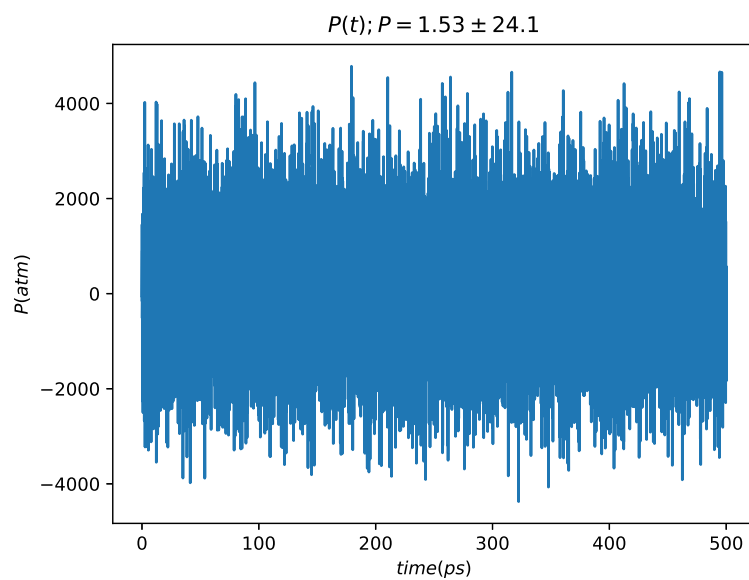


Рис. 7: pprn, $N = 5^3$

Можно заметить, что значение C_V даже посчитанное в `npt` вместо положенного `nvt` режиме гораздо меньше отличается от экспериментального, чем значения «правильно» посчитанного C_P . Среднее значение T так же прекрасно совпадает с тем, на что был настроен термостат. В отличие от значения среднего P , которое отличается от целевого на величину порядка себя. Из-за таких проблем с давлением я думаю, что возможно косяк в баростате, но я его пока не нашел.

3 Результаты DSF

Тут тоже 5 частиц в длину ячейки оказалось достаточно для не сильного относительного отклонения теплоемкости от значений для больших N .

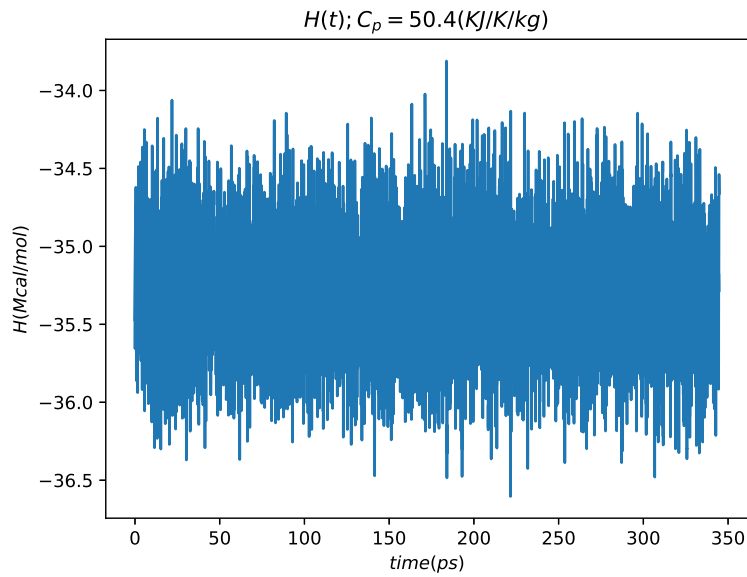


Рис. 8: dsf, $N = 15^3$

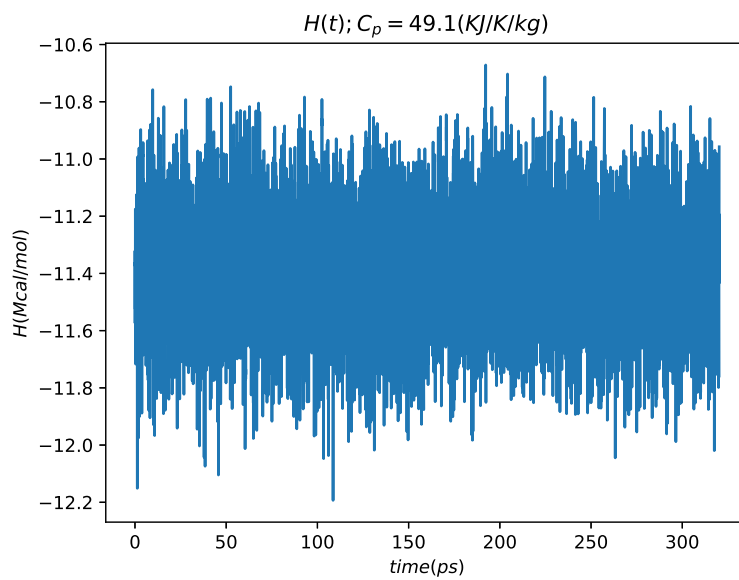


Рис. 9: dsf, $N = 10^3$

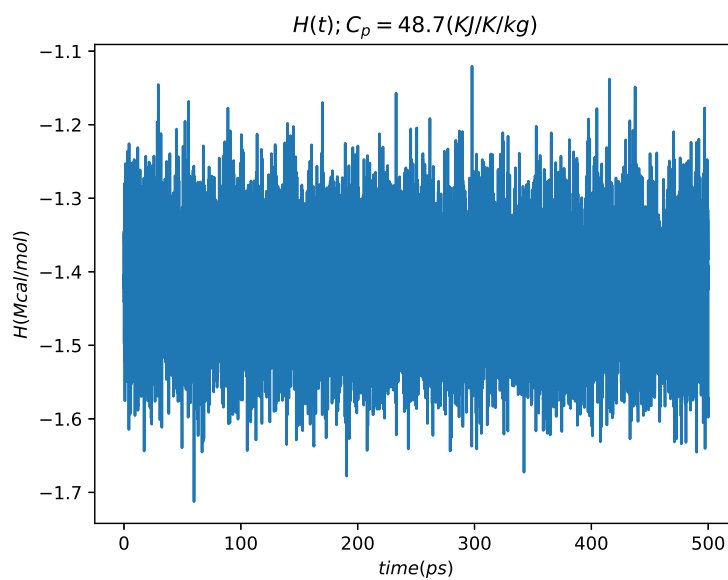


Рис. 10: dsf, $N = 5^3$

3.1 Другие равновесные характеристики

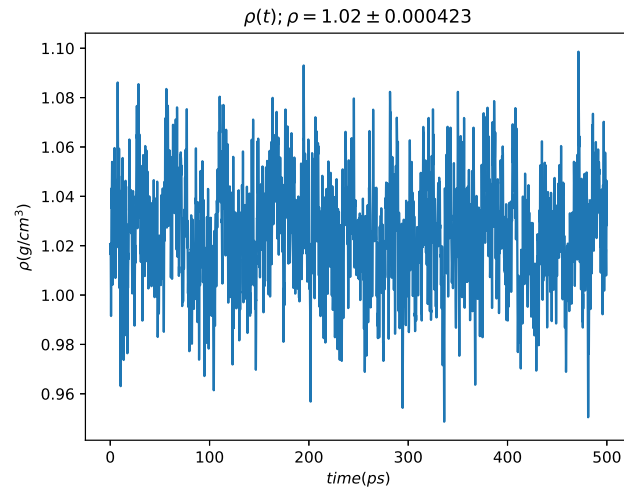


Рис. 11: dsf, $N = 5^3$

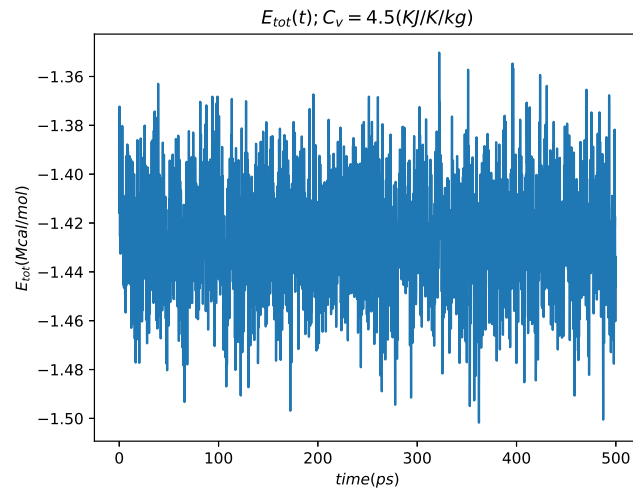


Рис. 12: dsf, $N = 5^3$

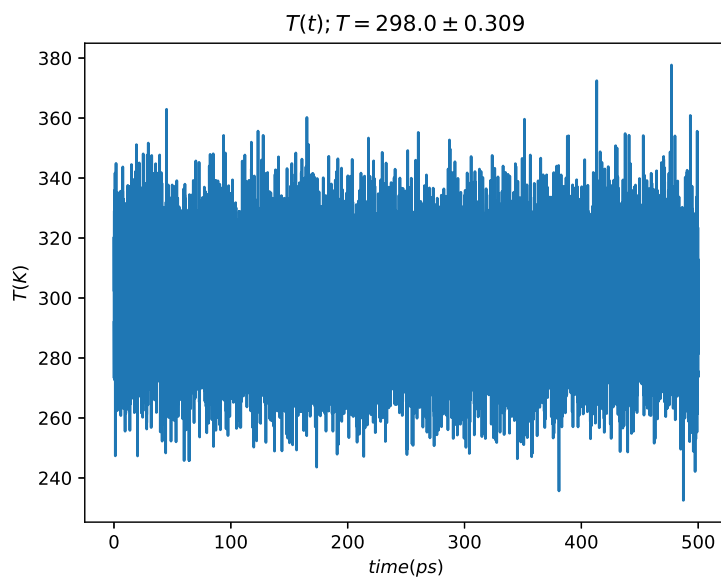


Рис. 13: dsf, $N = 5^3$

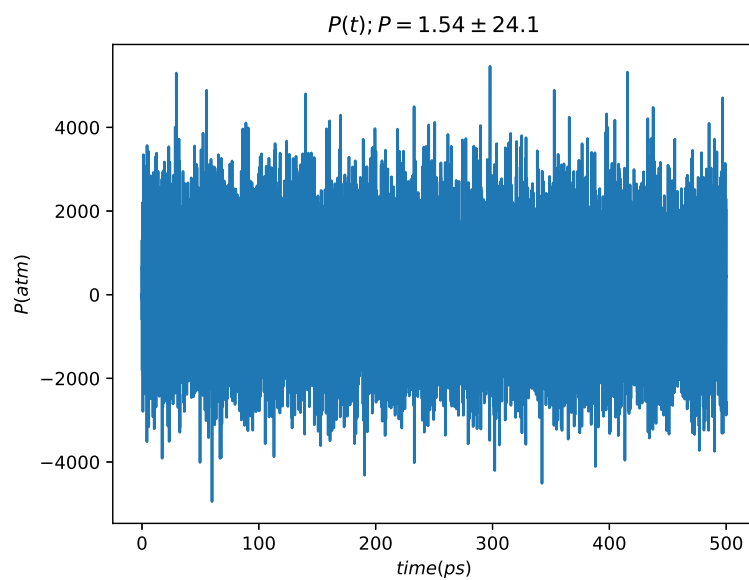


Рис. 14: dsf, $N = 5^3$

Можно заметить, что со значением кулоновского дамппнга 0.05 для dsf значения теплоемкости уже близко к значению для rrrm: 48.6 vs 48.7. Можно проварьировать этот параметр, попытавшись добиться еще лучшего согласия. Для плотности результат при любых значениях, близких к значениям из документации, не меняется, так что эффективно сравниваем только C_p . Как видно по рисункам ниже, значение $\alpha = 0.05$ изначально было хорошим попаданием, так что можно оставить его.

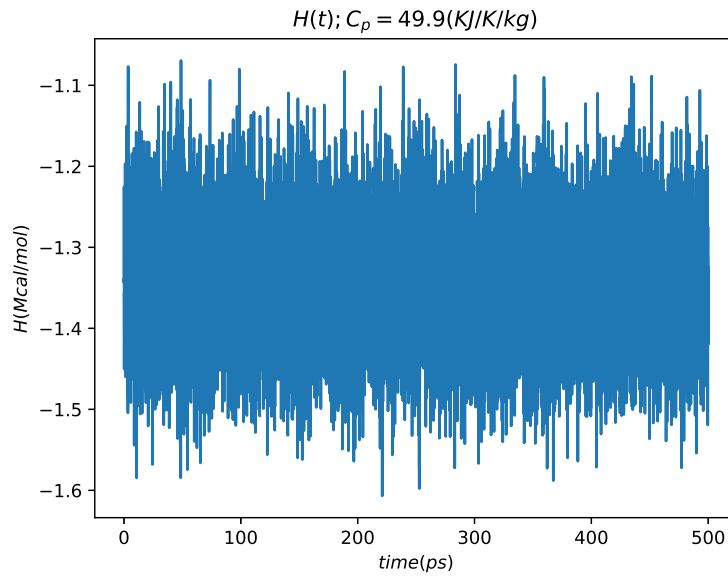


Рис. 15: dsf, $\alpha = 0.1$

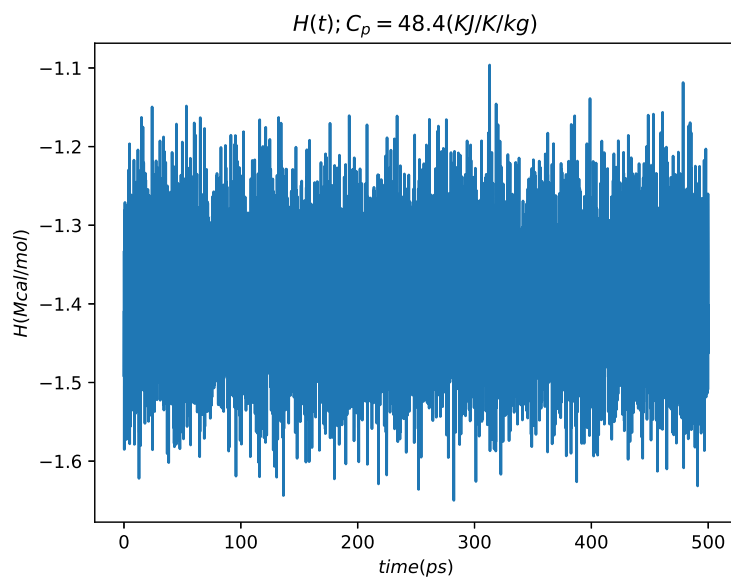


Рис. 16: dsf, $\alpha = 0.025$

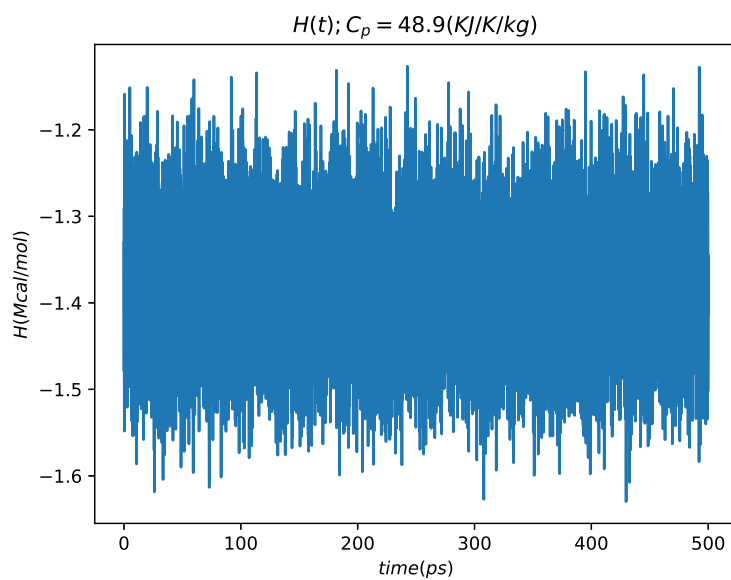


Рис. 17: dsf, $\alpha = 0.055$

4 Результаты tip4p

Значения теплоемкостей расходятся с экспериментальными, поэтому для подолнительной проверки был проведен расчет с потенциалом tip4p/rrrm. Тут можно отметить, что для воспроизведения результатов необходимо было заменить параметры молекулы на те, что рекомендует ламмпс. Приведем здесь это сравнение. Для H_2O из документации tip4p воспроизводит результаты, получаемые на SPC/E H_2O .

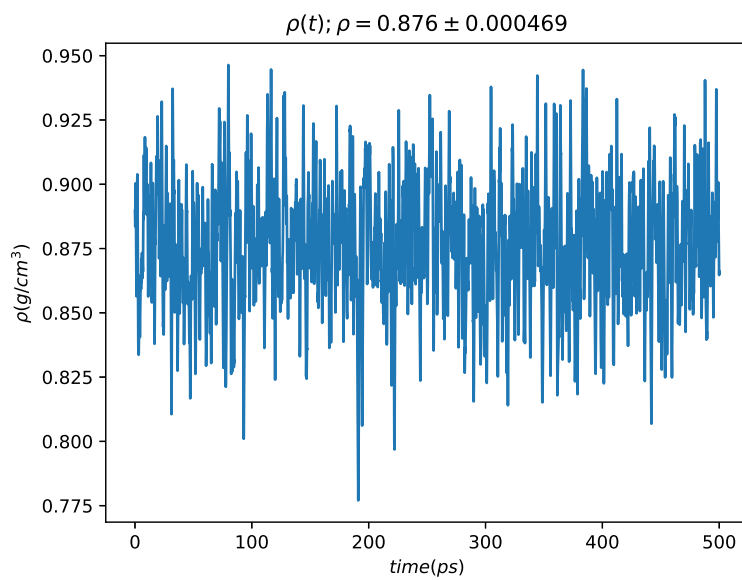


Рис. 18: tip4p/pppm, SPC/E H_2O

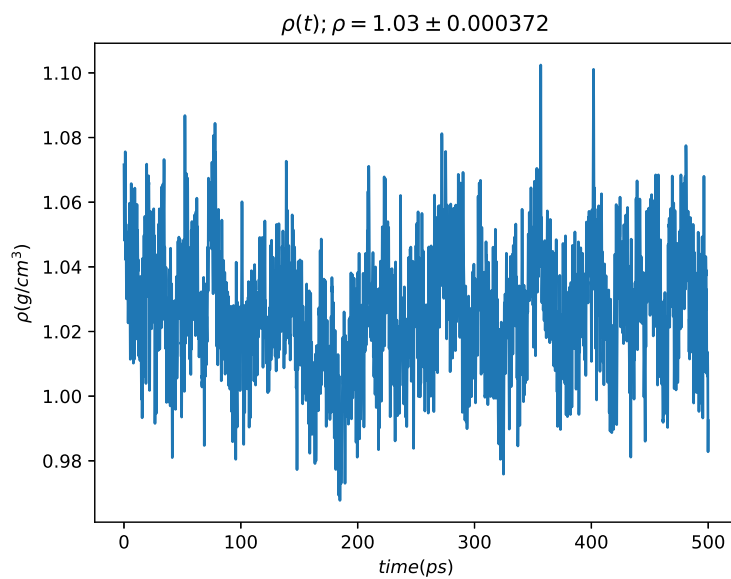


Рис. 19: tip4p/pppm, lammps docs H_2O

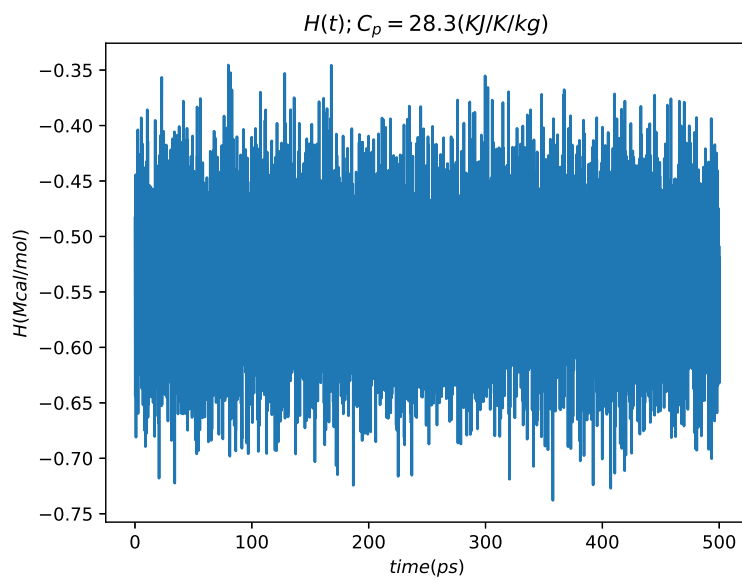


Рис. 20: tip4p/pppm, SPC/E H_2O

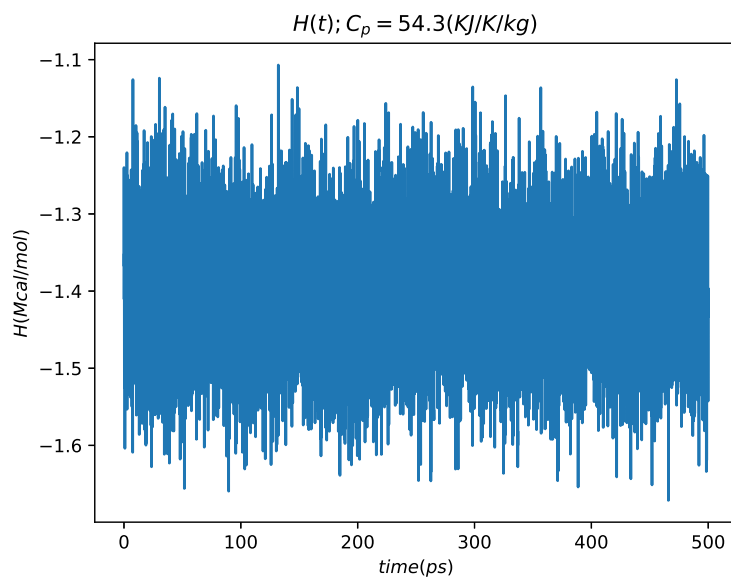


Рис. 21: tip4p/pppm, lammps docs H_2O